Market Integration with Transaction Costs in Developing Country Staple Food Markets: the Case of the Malawi Maize market

By

Zant, Wouter

Contributed Paper presented at the Joint 3rd African Association of Agricultural Economists (AAAE) and 48th Agricultural Economists Association of South Africa (AEASA) Conference, Cape Town, South Africa, September 19-23, 2010.

DRAFT VERSION – INCOMPLETE - NOT TO BE QUOTED - COMMENTS WELCOME

initial version: 15/11/2009, this version: 1/3/2010

Market Integration with Transaction Costs in Developing Country Staple Food Markets: the Case of the Malawi Maize market

Wouter Zant[1]

Abstract

We investigate measurement of market integration of staple food markets in developing countries. The analysis takes the Parity Bound Model as starting point and modifies this model by parameterizing and estimating transaction costs. The specification of transaction costs takes account of transport costs, fixed source costs, fixed destination costs, ad valorem taxes & levies and seasonality an is implemented on the basis of a specific sub-sample of price differentials. Price differentials combined with predicted transaction costs enable the measurement of market integration for each location and each period. The proposed method is applied to the Malawi maize market with monthly data from June 1999 to October 2009 for 26 districts. This period covers two major food shortages (2001-2003 and 2005-2006) as well as the impact of the global rise in food prices in 2008. Predicted transaction costs are shown to be in the same order of magnitude as survey observations of transaction costs of domestic maize trade in Malawi. The evidence indicates that markets are particularly well integrated during food shortages and substantially less in periods without food shortage. However, this result is largely due to high transaction costs during food shortages that make trade economically unattractive. Market integration develops markedly different between regions and districts: various districts in the south show signs of barriers to trade while districts in the north suffer from frequent high transaction costs. Low levels of market integration observed outside periods of food shortages, as well as high transaction costs during food shortages both suggest that trading infrastructure is not sufficiently developed.

JEL code: F14, Q13

Key words: food markets, transaction costs, trade, market integration, Parity Bound Model, Malawi, Africa

---

[1] Wouter Zant is associated with the VU University Amsterdam; mailing address: Wouter Zant, Vrije Universiteit, De Boelelaan 1105, room 5A-36, 1081 HV Amsterdam, The Netherlands; e-mail address: wzant@feweb.vu.nl. My work on this subject started with an EU funded evaluation project in 2006 in which I participated (Review of Stakeholders' Responses to the 2005-2006 Food Crisis in Malawi, May/June 2006). I would like to thank my colleague consultants in this project and all stakeholders in Malawi for their contribution to my understanding of this subject. Errors and omissions in this study are the sole responsibility of the author.

Contents

1. Introduction

In this paper we study measurement of market integration in domestic staple food markets in developing countries. Market integration is widely recognized as conducive to economic growth and poverty alleviation. A higher degree of market integration entails less restrictions to trade, smooth trade flows from surplus areas to deficit areas, better transmission of price signals, less price volatility, production decisions according to comparative advantage, realization of gains from trade and, in summary, higher welfare. Higher degrees of market integration are also likely to stimulate a quicker response to policies and to induce more adequate reactions to shocks. Many developing countries have embarked on a process of market liberalization of domestic markets including domestic staple food markets, partly motivated and guided by the World Bank and IMF. A sound method designed for developing countries to assess accurately to what extent these efforts have changed the degree of market integration would be informative and useful. Additionally the assessment of the degree of market development since the dismantling of state trade organizations and the impact of major supply and demand shocks, over the years and between regions, could also benefit from such a method. Popular methods currently practiced to measure market integration focus on the co-movement of prices and make use of time series techniques (Granger causality, error correction and co-integration). Using time series techniques for this purpose suffers from serious drawbacks: these techniques cannot disaggregate market integration in time and space, fail to deal with discontinuities in trade and trade-reversals, do not offer a measure of integration in degrees - integration is either accepted or rejected - and are not compatible with any theoretical insight of spatial price equilibrium. The Parity Bound Model (Baulch, 1997) overcomes these drawbacks. The key distinctive feature of the Parity Bound Model is the incorporation of transaction costs - constructed on the basis of a single observation of these costs - in the assessment of market integration. Along the lines of the competitive spatial equilibrium, price differentials relative to constructed transaction costs are used to assess market integration. Probabilities of price differentials to be at par with, below or above transactions costs are estimated with a switching regression technique. In this paper a modification of the Parity Bound Model is proposed: instead of extrapolating transaction costs we propose to estimate transactions costs. Out-of-sample predictions of transaction costs for the entire market combined with price differentials are used to assess market integration.

The paper is organized as follows. In Section 2 we review approaches and methods to assess market integration documented in the literature, including a discussion of their merits and drawbacks. Next we explain and evaluate the Parity Bound Model (PBM) and subsequent extensions and applications of this model. The concept of market integration is clearly defined in this model. An modification of this model is proposed for staple food markets in developing countries. In Section 3 we apply the proposed modification to the Malawi maize market. For this purpose we first present an empirical description of Malawi, the Malawi economy and the Malawi maize market from 1999 to 2009, highlighting supply and demand of maize, maize prices, domestic trade infrastructure and domestic trade practices. Next, we elaborate the empirical specification of transaction costs and elaborate on the selection of observations used for estimation and present the market integration indicator. In Section 4 we present and discuss estimation results of transaction cost equations. We predict transaction costs for all periods and all trade pairs and verify predicted transactions costs with survey data on transaction costs. Next we calculate to what extent price differentials are equal, above or below estimated transactions costs, which establishes the key indicator to assess market integration. We look specifically at trade from rural to urban (deficit) areas, from rural surplus to rural deficit areas and during major food shortages.

2. Market integration with transaction costs

Using co-movement in prices to assess market integration

A popular way to investigate market integration is to use time series estimation of an equation that explains prices of one location with prices of other locations and possibly lagged[2] prices of other locations. In most instances these time series analyses of market integration are purely based on price data and make no use of data on transaction costs. Market integration is determined by (the sum of) the coefficient of prices which should be equal to one for complete market integration (Law of One Price), or by establishing that prices are integrated and hence that the error term is stationary using appropriate tests (co-integration). Formally estimation of a variant of the equation below is used to assess market integration:

$$p_t^i = \beta_0 + \sum_k \beta_{1k} p_{t-k}^j + \varepsilon_t \tag{1}$$

$$\text{where } p_t^i = \text{price in } i \text{ at time } t$$

For market integration (Law of One Price) it is required that

$$\sum_k \beta_{1k} = 1 \tag{2}$$

Alternatively one may postulate market integration if prices of two markets are integrated for the same order and for this purpose one needs to test the stationarity of the error term with appropriate tests, or:

$$\varepsilon_t \sim I(0) \tag{3}$$

More sophisticated – but essentially similar – approaches are applied, with more flexible specifications and using more elaborated techniques. After conversion of equation (1) to an error-correction specification – where short run price fluctuations in differences depend on lagged deviations from equilibrium, there own lagged price differences, and lagged price differences of other locations - it is possible to distinguish short run and long run. An error correction specification may be specified as follows:

$$\Delta p_t^i = \beta_0 + \beta_1 p_{t-1}^i + \beta_2 p_{t-1}^j + \sum_k \beta_{3k} \Delta p_{t-k}^i + \sum_m \beta_{4m} \Delta p_{t-m}^j + \varepsilon_t \tag{4}$$

Similar tests on the error term confirm or reject the co-movement of spatial prices, and hence the integration of spatial markets.

Again alternatively, one may use an error correction framework to investigate Granger causality between (pairs of) spatial prices. In the above equation a rejection of the hypothesis that $\beta_4=0$

---

[2] If the realization of arbitrage transactions requires more time than the frequency of the price data used in estimation, lagging of explanatory variables is needed.

indicates that prices in market j Granger cause prices in market i. The literature abounds with empirical and methodological examples of these time series approaches. A highly arbitrary selection of these types of studies is: Ravallion (1986), Baffes (1991), Dercon (1995), Lutz et al. (2007). Further and additional empirical examples can be found in nearly any study on market integration.

There are several important drawbacks in assessing market integration by investigation of the co-movement of prices. Time series techniques require the direction of trade to be fixed and cannot deal with trade reversals or discontinuities in trade. This is not a trivial issue: variations in production in rain fed agriculture in developing countries are common and often large. Transformations of surplus areas into deficit areas from one crop year to another and back again, including the accompanying reversals of the direction of trade are not rare occasions but occur with some frequency. Time series techniques also run into trouble if data are not complete and incomplete data are common in agriculture in developing countries (as is confirmed in the empirical part of this study!). Absence of data is obviously a problem for any analysis but incomplete data do pose more problems for time series techniques than available alternatives (see e.g. the remainder of this study). In most studies only prices are used in the analysis and transaction costs are not properly accounted for, simply because such data are not available. Relating market integration on the co-movement of prices alone and ignoring transaction costs will lead to flawed inferences since transaction costs constitute a substantial component of prices and fluctuate independently over time and between locations. Monte Carlo simulations show that conventional tests of market integration based on the co-movement of prices are flawed (Baulch (1997b)).

The Parity Bound Model: theory, application, improved versions and proposed extension

An additional major drawback of the approaches that use time series techniques to establish the co-movement of prices in order to identify market integration, is its lack of theoretical base. There is a small literature that discards the time series approach on the above mentioned grounds and takes an alternative route to find proper indicators of market integration. This literature uses competitive spatial equilibrium as its starting point. The Enke-Samuelson-Takayama-Judge spatial and temporal equilibrium is characterized by the so-called spatial equilibrium conditions (also complementarity conditions, or complementary slack condition). These conditions are:

$$p_t^j - p_t^k - tc_t^{jk} \leq 0 \quad \perp \quad m_t^{jk} \geq 0 \tag{5}$$

$where\ p_t^j\ is\ the\ price\ in\ the\ importing\ market\ j\ at\ time\ t$

$p_t^k\ is\ the\ price\ in\ the\ exporting\ market\ k\ at\ time\ t$

$tc_t^{jk}\ is\ the\ transction\ \text{cost}\ of\ trade\ from\ k\ to\ j\ at\ time\ t$

$m_t^{jk}\ is\ the\ trade\ flow\ from\ k\ to\ j\ at\ time\ t$

For competitive spatial and temporal equilibrium it is required that one of the conditions in the above equation is satisfied with equality. The basic insight from these conditions is that in competitive equilibrium the price differential cannot exceed transaction costs. If trade is taking place ($m_t^{jk}>0$) than the price differential must equal transaction costs ($p_t^j-p_t^k=tc_t^{jk}$). If there is no trade ($m_t^{jk}=0$) competitive spatial and temporal equilibrium requires that the price differential to be below transaction costs ($p_t^j-p_t^k<tc_t^{jk}$). Equilibrium is achieved as - once the price differential is

sufficient to cover transaction costs and offering incentives for trade flows - the gap between prices in import and export markets plus transaction costs narrows - either by a decrease in import prices, an increase in export prices, an increase in transaction costs or a combination of these - until the price differential is exactly equal to transaction costs and all trade opportunities are exhausted. A price differential smaller than transaction costs (p_t^j-p_t^k <tc_t^{jk}) is also consistent with competitive equilibrium but there will be no positive trade flows (m_t^{jk} =0). In this case the price difference between the importing market j and the exporting market k is not sufficient to offset the costs of moving goods from k to j. Consequently there is no incentive to trade. Formally it is not possible to identify if this "no-trade" situation is due to too low prices at import markets, too high prices in export markets or too high transaction costs. However, once transaction costs are known we are able to make statements about the relative contribution of transaction costs to changes in market integration. This will be further elaborated in the following sections.

Conversely we may characterize a violation of the spatial arbitrage conditions or, equivalently, a condition that characterizes markets which are not integrated as follows:

$$p_t^j - p_t^k - tc_t^{jk} > 0 \quad and \quad m_t^{jk} \geq 0 \qquad (6)$$

The price differential is sufficiently high to cover transaction costs, but for some reason or another no or only limited trade takes place. At least not to the extent that the price in location j is equalized with the price in location k plus transaction costs. Again it is not possible to determine the cause of the lack of market integration. It is suggested that this could be caused by barriers and impediments to trade, government controls on production, oligopolistic pricing, transportation bottlenecks, etc. A higher incidence of the last condition (equation (3)) indicates a low degree of market integration.

A seminal contribution to the empirical line of research using the competitive spatial equilibrium conditions as starting point is Baulch (1997a), which was preceded by Sexton et al. and Spiller and Huang (1986), and has seen a number of extensions and improvements (Barrett (2002), Barrett and Li (2002), Fackler and Tastan (2008)). Following the competitive spatial equilibrium conditions Baulch (1997a) identifies three regimes in his so-called Parity Bound Model: in regime I, between the Parity Bounds, the price difference between two locations is exactly equal to transaction costs (p_t^j-p_t^k -tc_t^{jk}=0); in regime II, inside the Parity Bound, the price difference between two locations is lower than transaction costs (p_t^j-p_t^k -tc_t^{jk}<0) and in regime III, outside the Parity Bound, the price differential exceeds transaction cost (p_t^j-p_t^k -tc_t^{jk}>0). Regime I and II are consistent with competitive spatial market equilibrium and market integration. In the empirical implementation Baulch proceeds with the attribution of trade pairs to regimes. The problematic part in the empirical application is the lack of information on transaction costs. Data on transaction costs are usually not available. Therefore, Baulch (1997) proposes to use observed transactions costs on one single moment and extrapolates these to all periods and all locations. The difference between price differentials and extrapolated transaction costs characterizes the regimes. A switching regression equation is used to estimate the probabilities to be in either of the regimes, or to estimate the upper and lower parity bounds. Estimation of probabilities in a switching regression framework is based on techniques developed for estimation of stochastic productivity frontier estimation. Within the Parity Bounds the error is normally distributed, while it is assumed to be half-normally distributed above and below the parity bounds.

The Parity Bound Model (PBM) is a flexible tool to assess market integration as it allows trade reversals and discontinuity in trade flows, and it takes account of transaction costs. It allows transaction costs to fluctuate over time and between locations. The PBM enables geographical and temporal desaggregation, thereby allowing comparisons between areas within a country and between specified periods, and measures market integration in continuous degrees, rather than by accepting or rejecting market integration. Finally, unlike Granger causality tests and co-integration tests the PBM explicitly takes an economic model of spatial price determination as its starting point: the conditions for market integration are consistent with the Enke-Samuelson-Takayama and Judge spatial price equilibrium. Popular conventional methods using co-movement of prices to measure market integration fail on these accounts and therefore the PBM model is believed to be superior.

Extensions and improvements of the PBM model have been inspired by complementing price and transaction cost information with trade flow information. Barrett and Li (2002) argue that both data are needed to investigate the efficiency of trading behavior. With trade flow information at hand one may distinguish six regimes: the regimes identified in Baulch (1997a), but either with positive or without trade flows. Regimes are combined and labeled perfect integration (p_t^j-p_t^k -tc_t^{jk}=0 and m_t^{jk}>0), segmented equilibrium (p_t^j-p_t^k -tc_t^{jk}<0 and m_t^{jk}=0), imperfect integration (p_t^j-p_t^k -tc_t^{jk}> <0 and m_t^{jk}>0) and segmented disequilibrium (p_t^j-p_t^k -tc_t^{jk}>0 and m_t^{jk}=0): the first two regimes are consistent with competitive spatial equilibrium, while the first and the third regime require positive trade flows and thus describe integrated markets. Probabilities for regimes are estimated using a switching regression specification, along the same lines as Baulch (1997a). The methodology is demonstrated with international soybean meal trade between Canada, Japan, Taiwan and the United States. Imperfect competition is only identified to be substantial and significant in trade with Japan and attributed to non-tariff barriers. Next, the incidence of segmented disequilibrium is, as expected in well developed international markets, negligible. Probabilities for perfect integration and segmented equilibrium are, both, often large and significant.

A more fundamental approach is proposed by Fackler and Tastan (2008). They describe three measures of market integration that are directly linked to a model of spatial price determination and propose an econometric approach to estimate a spatial price model that gives estimates of these measures of market integration. Complementary slack conditions are derived from net excess demand function, transaction costs and an equilibrium condition, for each location in a network. Measures of market integration are defined in terms of the price transmission ratios, which measures the degree to which an excess demand shock in one location is transmitted to another location. The following measures are proposed: the expected value of price transmission ratios, the probability of price transmission ratios >0 and the share of isolated submarkets in the market as a whole. Obtaining an explicit expression of the price distribution in terms of latent variables is essentially impossible with many locations. Therefore indirect inference techniques (or Simulated Methods of Moments) are used for estimation. Estimation is implemented with simulated data and with data of soybean trade between the US, Brazil and Rotterdam. The application with empirical data support (nearly) complete integration between US and Rotterdam, and between Brazil and Rotterdam, which trade links are assumed to be active. The approach is claimed to offer integration measures with a direct economic interpretation (assuming that they can be identified) and avoids the problem of market integration due to indirect trading (two locations may never trade with each other but yet be integrated through a third location[3]). Three limitations of this approach are mentioned: the assumption of normality of structural variables, the assumption of linear excess demand and linear transport supply functions and the spatial price model ignores storage and transport timing issues.

Despite the intelligent, sophisticated and appealing approaches described in Barrett et al. (2002) and Fackler and Tastan (2008), we do not develop our work in these directions[4]. Barrett and Li (2002) complement price data with (comprehensive) data on trade flows and on transaction costs and this establishes a key element of their analysis. Domestic trade flow and transaction cost data are usually not available for developing countries domestic markets. In fact, even in developed countries comprehensive domestic trade flow data and transaction data are difficult to get and often of poor quality[5]. Moreover, this lack of data is not likely to change in the near future. Our contention is that, if ever at any moment in the future these data will become available - as is the case in a few highly developed international markets – it is likely that the problem of market integration in developing countries has become redundant. Hence, we believe that approaches and methods proposed in these studies are not helpful for the analysis of market integration in most developing countries. The approach of Fackler and Tastan (2008) appears fruitful and interesting since only prices are needed to estimate measures of integration. Also the distinct economic interpretation of their market integration measures is elegant. Nevertheless, a number of issues stand in the way of practical applications. Identification of the market integration measures is unclear. The computational burden is large and the technique is advanced and not easy accessible. Relaxing the restrictive assumptions will be methodologically demanding and, combined with a full and empirical modeling of a spatial network, increase the computational burden further. Tractable theoretical solutions are not compatible with accurate specifications of real world excess demand and transaction cost. The approach estimates fixed and time invariant market integration measures rather then describing market integration period-to-period for each location. In short, it is not likely that this will become a transparent and easy access approach to evaluate market integration, useful for assessing developing country domestic markets.

Instead of extending PBM model along the lines of Barrett and Li (2002) and Fackler and Tastan (2008) we propose a modification of the approach applied in Baulch (1997). Using observed transactions costs at one moment in time as a basis for extrapolation to all periods and all locations is the backbone to the approach implemented in Baulch (1997a). Data on transaction costs are not easily available for all periods and trading pairs in a spatial network, and this applies especially in an environment with poor recording of statistical data as is the case in most developing countries. Under these conditions extrapolation appears to be a sensible strategy. However, it does entail a large degree of inflexibility and error: extrapolation can only account for limited and rather restricted time and trade pairs specific variations in transaction costs. Therefore we propose to model transaction costs and use the predictions of this model as estimates of transaction costs. Empirical modeling of transactions cost equation allows for trade pair specific and time specific variation, based on a parametrization that can be justified economically. Next, we also avoid the distributional assumptions required to estimate the

---

[3] This issue is solved in our empirical application since we aim at covering all locations in a network.

[4] In the approach proposed in the remainder this paper – and unlike Baulch (1997a) – there are no distributional assumptions underlying the identification of regimes. Hence, we do take account of one of the points of criticism to the approach of Baulch (1997a), which is addressed in Barrett and Li (2002).

[5] Barrett and Li (2002) need around two pages to explain difficulties, arbitrary decisions in construction and selection and remaining potential errors in their price, transaction cost and trade flow data from an internationally traded commodity (soybean meal) and for a selection of highly developed countries (Canada, Japan, Korea and US).

probabilities of the different regimes within a switching regression framework: we simply construct the distribution of price differentials minus transactions costs and use calculated values to assess market integration over time and between locations. Rather than estimating fixed and time invariant regime probabilities - as is standard in Baulch (1997a), Barrett and Li (2002) and Tostão and Brorsen (2005) - we maintain full flexibility, let the data speak and do not enforce a fixed value of the market integration. Variation over time and between locations is shown to be substantial and informative. Finally, it is attempted to fully exploit the possibilities to disclose the impact of transaction costs on market integration.

In the next section we discuss the specification and estimation of such a transaction cost equation. We present an empirical application of the proposed approach with data of the Malawi maize market and, hence, we begin with a descriptive background of the Malawi maize market.

3. Facts and figures on Malawi

The Malawi economy and the Malawi maize market

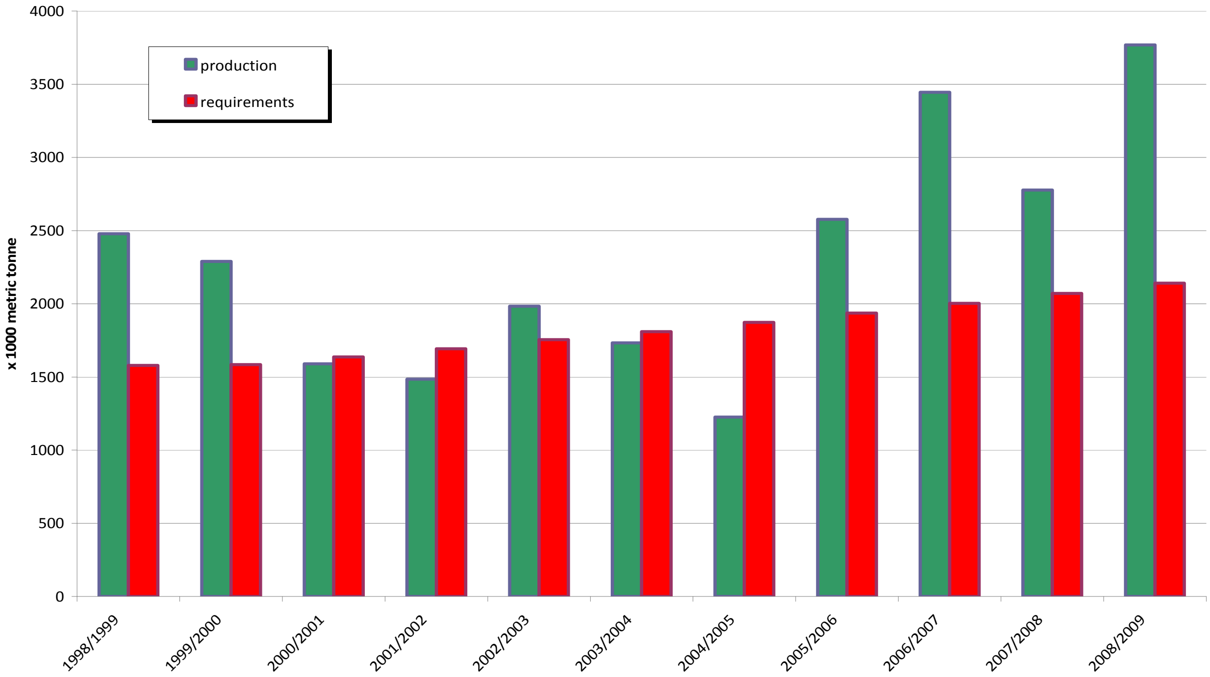
In this section we present an empirical description of Malawi, the Malawi economy and the Malawi maize market from 1999 to 2009, highlighting supply and demand of maize, maize prices, domestic trade infrastructure and domestic trade practices. Malawi is a landlocked country in sub-Sahara Africa, measuring more than 800 kilometers from north to south, and 100-170 km from east to west, bordering in the north and northeast by Tanzania, on the upper west side by Zambia and enclosed by Mozambique on both the lower west side, the lower east side and the south, has a total land area of 94,276 km2, around 2.5 times the size of The Netherlands and of similar size as Portugal. A large lake – Lake Malawi - borders the country on the eastern side and stretches from the upper north to around 2/3 down south, culminating in the Shire river, which exits the country in the outermost southern point of the country to eventually feed the Zambezi river (see Annex for a Map of Malawi). The population has increased from 10.5 million in 2000 to 14.5 million in 2010, of which around 12% lives in the Northern region, around 42% in the Central Region and around 46% in the Southern region. The Central and specifically the Southern region are relatively densely populated. There are two major cities (Lilongwe and Blantyre, with both around 0.5 million people in 2000 increasing to slightly less than 0.9 million in 2010). The larger part of the population lives in rural areas (around 87% from 2000-2010). On average 77% of population above 15 years of age – and 84% in rural areas - earns an income in agriculture, and this peaks to above 90% is a few rural districts. The composition of GDP is also, but much less, biased towards agriculture: agriculture accounts for 40% of GDP, services for around 40% of GDP (retail and distribution 15%, government services 6%, transport 6% and financial services 8%) and manufacturing around 10%. The government budget varies from 26 to 35% of GDP of which 8 to 14%-points is accounted for by grants. In years without a food shortage economic growth is moderate (3 to 6%). The major export products are tobacco and tea (54 to 98% of total merchandise exports) and the major import products are fuel and fertilizer (15-30% of total merchandise imports). The balance of trade has shown an increasing deficit in the period from 2000 to 2009 (in terms of GDP increasing from 8% in 2002 to 24% in 2004), partly due to deteriorating terms of trade. Per capita GDP, expressed in purchasing power parity US$, in 2009 (2008) is between 840 and 900 US$, making Malawi one of the poorest countries in the world, with a ranking in the bottom 14 of all countries[6]. The poverty rate is high and

---

[6] Malawi per capita GDP in purchasing power parity US$ according to IMF, 2009: 881 US$ (ranking: 170 in a total of 181 countries), according to the World Bank, 2008: 837 US$ (rank: 156 in a total of 166 countries) and according to CIA, 2009: 900 US$ (rank: 180 in a total of 193 countries).

varies from around 25% in the urban areas (eg. Lilongwe: 24), to around 70% in remote rural areas (e.g. in the north Chitipa (67.2%) and in the south Nsanje (76.0%) and Chikwawa (65.8%), see appendix for regions and districts). For the country as a whole 52.4% of the population in Malawi is poor, and 22% of the population is ultra poor (source: Integrated Household Survey 2005 (IHS-2)). The urban-rural and between region & district differences also extend to other topics. Life expectancy at birth ranges for the whole country from 43 to 49 from 2000 to 2010, but these figures vary from 60 to 65 for urban districts and from 42-47 in rural districts (NSO, Statitical Yearbook 2008). Above 15-years-of-age literacy rates are substantially higher in rural areas of the Northern region (77%) than in rural areas of the Central and Southern region (58% and 54%; source: NSO, 2007). The urban areas (Lilongwe, Blantyre, Zomba and Mzuzu) have the highest literacy rates (87%). Household size ranges from 4.1 in the Southern region, 4.5 in the Central region to 5.1 in the Northern region (Source: Benson et al. 2002).

Figure 1 Annual maize production and maize requirements in Malawi*



* Maize production by crop season and maize requirements by marketing season
Source: (own calculations based on data from) FEWS NET; Ministry of Agriculture and Food Security, Malawi

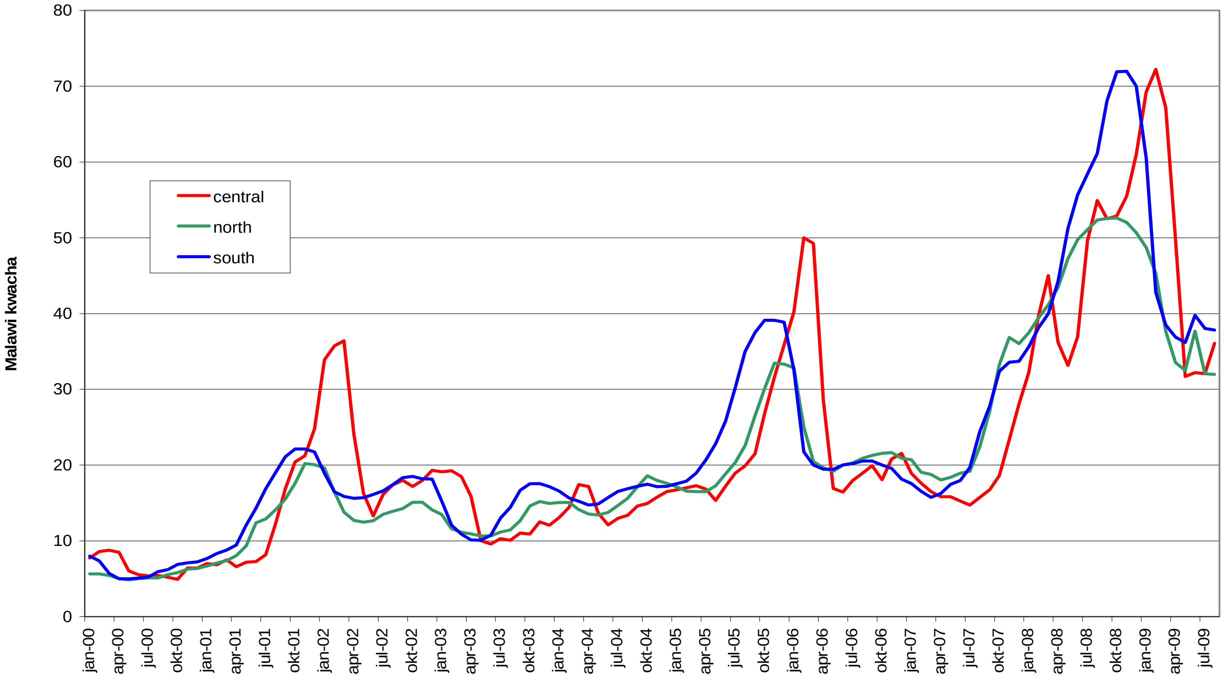
Maize in Malawi dominates both in agricultural production and in consumption of households. Maize is the major and dominant staple food in Malawi: between 52% and 65% of the total per capita calories intake is due to maize (FAO, 1990, 2002). Also, nearly all households – an estimated 97% - grow maize (IHS (2005)). Production of maize in Malawi is primarily undertaken by households on subsistence grounds. According to the same household survey data from 2005 81% of the population in rural areas is classified as subsistence farmers. It is claimed that, as a result, only a small part of maize production is sold on the market[7]. The largest market for white maize is the southern part of the country which has a high population density.

---

[7] Availability of maize to sell on the market obviously depends on the size of production. Maize harvests tend to

The main maize crop in Malawi is planted in September and October and harvested from March to June. Malawi has two distinct seasons, a wet, warm season, running from November to April and a dryer, cooler season, from May to October. Annual rainfall varies from 700 to 1600mm, with a median of around 1100mm, depending on location and year. The size of harvest is critically affected by the incidence of droughts, which causes enormous fluctuations in production of maize (see Figure 1). On a national level a good harvest can be as high as three times the size of a poor harvest. However, on district level this figure is reaches peaks of 6 times to 8 times, suggesting scope for district to district trade. The figure also shows country wide shortages of maize in crop season 2000/2001, 2001/2002 (marketing season 2001/2002, 2002/2003) and crop season 2004/2005 (marketing season 2005/2006). In both instances this has given rise to large inflows of emergency food aid from international institutions and donors.

**Figure 2	Maize market prices in Malawi**



Source: FEWS NET; Ministry of Agriculture and Food Security, Malawi
* The figure shows median prices by region of a total number of 72 maize markets

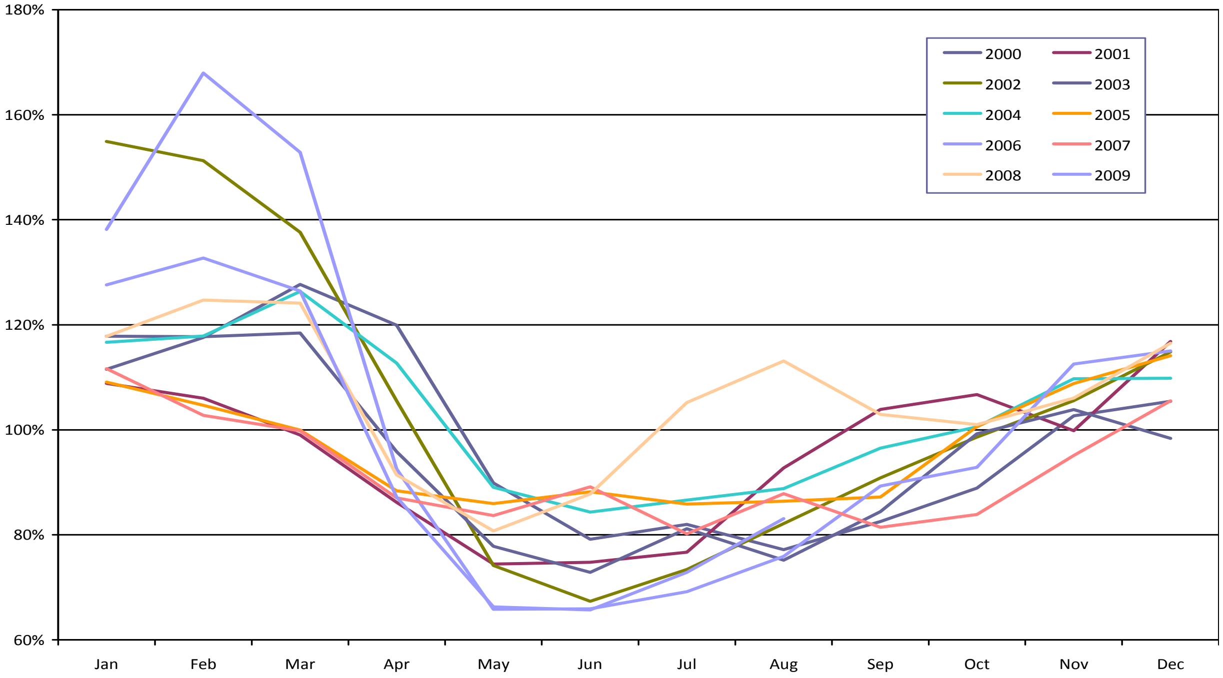
Periods with food shortages are clearly identified as high price periods where domestic maize prices peak to or even exceed import parity levels (see Figure 2). However, this does not apply to the large price increases in 2008 which were accompanied by bumper harvests substantially in excess domestic requirements, for a number of crop years in a row (see Figure 1): 2008 domestic maize prizes were clearly affected by the 2008 price increase in global food prices. The data

---

show dramatic variations, mainly due to weather conditions. Reported estimates of the share of production that is marketed are in the range of 5-10% of domestic production. However, this figure is not entirely convincing, since it is based on the IHS-2 of 2005, in which crop season maize production was at an historically unprecedented low level (see also Figure 1). This illustrates an important drawback of using survey data: answers to questions are a reflection of circumstances at the time of the survey and cannot always be extrapolated beyond this period.

further suggest that price increases during shortages are consistently lower in the north and that price increases in the south precede price increases in other parts of the country (see Figure 2). Like all agricultural products around the world, there is a distinct seasonal pattern in maize Malawi prices. This pattern shows highs at the end of the marketing season, just before harvesting, during the months of January, February and March and lows after harvesting during the months May, June and July (see Figure 3). Even in marketing seasons with moderate circumstances (no food shortages, no global rise of food prices) the increase from low to subsequent high averages from 30% to 40%[8]. In general, maize price movements in Malawi are large in view of the importance of maize as a staple food for households in Malawi and in view of the claim on household budgets of expenditure on maize.

Figure 3 Seasonality in Malawi maize prices



Source: own calculations based on data from FEWS NET; Ministry of Agriculture and Food Security, Malawi

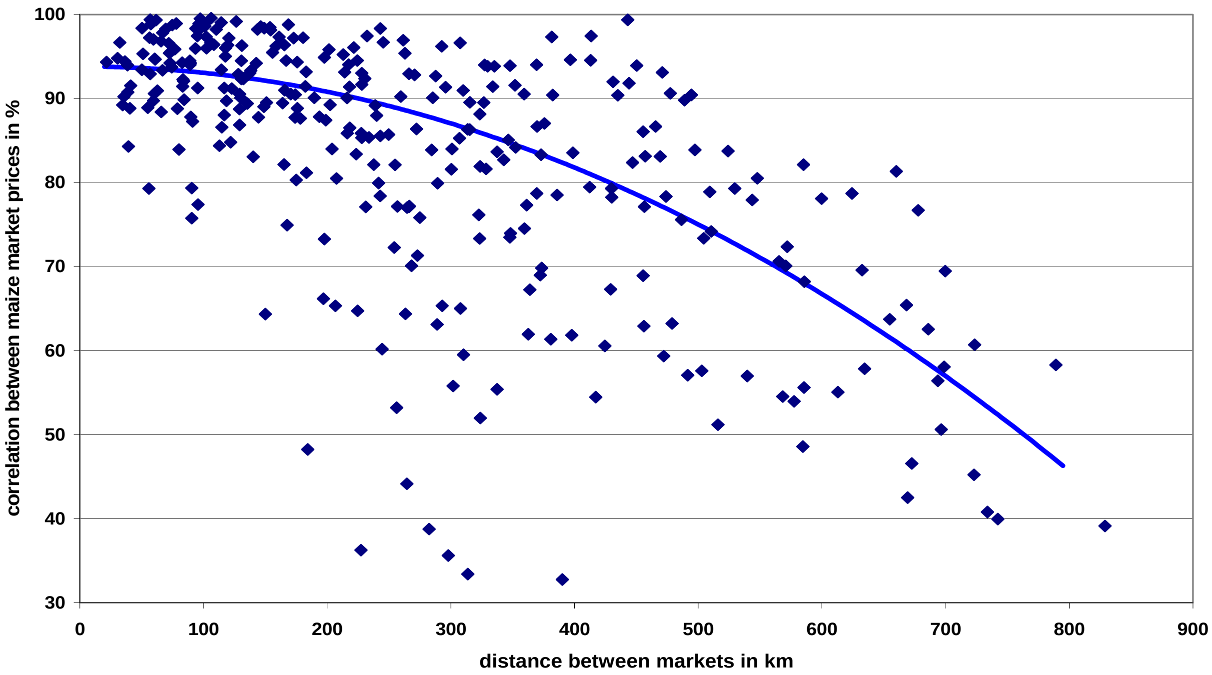
The average (median) distance between district capitals is 264 km (224 km), and has a maximum of 829 km and a minimum of 22 km. The distribution of district capitals over the country – the density of the spatial network - is not even. All districts in the north are typically remote and more dispersed or even isolated (see Table A3 in the appendix for a characterization of the spatial density of the network). Spatial prices – prices in different markets within Malawi – differ from each other – amongst other things - due to transaction costs. This is to some extent confirmed by the higher correlation of prices in different markets the closer these markets are located to each other. Beyond 250 km the correlation decreases rapidly (see Figure 4).

---

[8] Figure 3 shows country averages of maize market prices. Aggregation over districts dampens fluctuations at district level, or, conversely, fluctuations on district level are even higher than what is apparent from Figure 3. This is confirmed by our data.

Production technology and production practices – and consequently cost of producing maize - are similar throughout all districts of Malawi[9].

Figure 4 Distance between districts and correlation of maize district prices



Source: own calculations based on data from FEWS NET; Ministry of Agriculture and Food Security, Malawi; The Travel Distance Calculator, http://www.mapcrow.info

Domestic trade and transport in Malawi

Transport of goods is undertaken by lake, by rail and by road. Cargo lake services are available between several ports on the lake: Malawi Lake Services offers regular cargo services between the ports of Chipoka, Nkhata Bay, Chilumba and Monkey Bay, with four vessels with capacity of 300, 210, 600 and 720 tonnes. Although transported freight by lake is growing, freight of maize still accounts for a marginal share of total maize traded (less than 1%). Malawi also has a limited rail network: the rail line – with a total within Malawi length of 797 km - runs from the Zambian border at Mchinji to the east via Lilongwe and Salima, to turn south to Chipoka, Balaka, Nkaya where it splits into a line to the south, to Blantyre and Mkanga and enters into Mozambique to end in Beira (the Beira corridor), and a line to the east via Liwonde and Mayuchi, to Nacala in Mozambique (the Nacala corridor). The Nacala corridor is operational since 2005 and used for freight transport of maize, containers, fuels, fertilizer, cement, tobacco and sugar. The Beira

---

[9] Competition in the maize market drives down the market price to the level of production costs. Consequently under conditions of ample supply (without district to district trade), maize market prices will reflect unit cost of production (and the transaction cost component in maize market prices will be negligible). The hypothesis that unit maize production costs are equal between districts may be tested by calculating minimum market price of maize over the season, for each district and for each time period and by implementing a t-test on each district pair. Test results (not reported) are mixed and sensitive to the calculation of minimum prices and the use of alternative district specific consumer price indices for deflation. Equality is, however, supported in a substantial number of instances.

corridor has been out of service (since 1983) due to civil war in Mozambique and flooding in the south of Malawi, but is rehabilitated in 2007-2010 with the help of World Bank and other donors (the Sena rail rehabilitation). It is not entirely clear if (and which part of) the Beira corridor in Malawi has been operational and during which period. It appears, however, that the possibilities to benefit from transport opportunities by lake or rail – which are estimated to be around 10% of the cost of road transport – are limited and not fully exploited[10]. Both rail and lake transport are not reliable, undercapitalized, inefficient and cover a limited area of the country. Volume of transport by rail (in terms of ton-kilometers) is much larger for rail but it is not clear from the data to what extent this freight concerns maize.

The bulk of domestically traded goods is transported by road, and most "district to district" trade is done by small traders and farmers with larger holdings. Small scale traders and farmer-traders operate in villages, transport maize over limited distances, and use small sized pick-up trucks (maximum 1000 kg). Traded quantities per transaction are vary from 1 to 10 tons. These traders are large in number (many thousands) and have an elementary level of business organization, management and communication. Other major features of this trade business are lack of trust, personal travel for inspection of goods, payment of cash upon delivery, breach of contract, theft, and absence of brand names, trademarks or certified quality. Small scale traders are also active in cross border trade, have transactions with ADMARC unit markets and are most directly affected by ADMARC price policies and export bans. Their efficiency is constrained by lack of finance and lack of trust, resulting in much time being spent on travel in connection with collection and inspection of produce. Intermediate brokers are seldom used by informal traders. Transport of the produce represents by far the largest component of variable cost. Few informal traders make use of weighing equipment: instead volume is the preferred measure. This has the advantage that a high moisture content, which is often non-desirable, does not add to the sales price. A group of medium scale traders operate in urban centers, district towns and local trading centers, source maize from local traders and large smallholders and procure maize from the rural district markets which are in operation one, two or more days in the week. At the upper end there are a few large scale commercial traders (APEX traders, wholesale traders), that operate in urban centers, handle supplies of maize above 1000 tons and are mostly involved in trade of a range of commodities apart from maize. A number of larger traders are involved in processing (maize milling, corn-soy blending, etc.) and are active in importing maize from abroad through tenders from NFRA, donors and NGOs and related large scale operations. These large traders lease or own warehouses, operate a fleet of trucks and are involved with long distance deals. They further maintain a domestic commercial network of traders / suppliers in order to source maize but do seldom have a network of rural buying & sales depots of any significance. However, survey evidence indicates that less than 1% of all traders do pure wholesaling as a stand-alone business. Around 75% of all traders buy directly from farmers and sell as a retailer (Fafchamps et al. (2005)).

All traders - regardless of size of operations - face high commercial lending interest rates. Combined with the prevailing market uncertainty due to government interventions on the maize market, this makes funding of commercial trading activities through credit close to impossible. Consequently, trading operations are primarily self-funded: Survey information on traders in Malawi indicate that on average traders supply 96% of the working capital of the trading business. Around 88% of all traders are 100% funded with own capital (Fafchamps et al. (2005)).

---

[10] Controlling for lower transport costs for pairs of locations at lake Malawi, or along the railway line did not affect the estimations.

Since the scale of trading operations is closely related to the availability of working capital, a lot of effort is made by private sector agents to avoid that working capital is tied up in investments for long times. Hence, and out of economic necessity, the turnover time of working capital is short. Practically it implies that traders are mainly in involved in back to back to trade. The average (median) number of days between purchase and sale is nearly 8 days (3 days), with around 45% of the transactions completed in one or two days and less than 10% completed in more than 14 days (Fafchamps et al. (2005))[11]. The constraints on funding apply to all types of traders, although APEX traders may have some leverage at commercial banks, and on occasion pre-finance transactions for short periods. The shortage of capital is apparent throughout the whole trading chain, and in fact throughout the whole private sector. Additionally, all traders - regardless of size of operations - face high transport costs. Transport costs are high, especially because the secondary road infrastructure (feeder roads) is not well developed. As a result traders do not send vehicles to the rural areas, without a clear indication that their network counterpart has secured the requested quantities of maize. The problem of high transport costs becomes particularly acute in remote areas. Average distance between purchase location and sale location of maize transactions is just under 55km with a maximum distance between locations of purchase and sale of 200km (Fafchamps et al. (2005)).

Transaction costs: empirical specification and selecting a sample for estimation

Modeling transaction costs

Transaction costs for trade of maize between districts has various components. A major component is transport costs, which usually includes the costs of loading and unloading. Additionally, one can identify collection and bagging costs, storage costs, interest cost on working capital or opportunity cost of capital, information costs, taxes and levies on transactions, market fees, quality verification costs, etc.

In the empirical specification of transaction costs we have used several variables to approximate these cost components. A large component of transaction costs – in the range of 48 to 57% (see Fafchamps et al. (2005)) - concerns transport costs. Transport costs, in turn, are determined for a substantial part by distance traveled and transport fuel prices. Fuel for transport is fully imported into Malawi. In fact, fuel import is the single largest import product with a share of 11 to 17% of total merchandise imports (WDI, 2009). In order to account for transport costs we have included distance between districts and international diesel fuel prices, converted to Malawi kwacha as a determinant of transaction costs. It should be noted that other costs, like e.g. information costs, may also be related to distance and transport fuel prices.

Part of transaction costs concerns costs of labor, either hired or in the form of benefits from self employment and / or payment for services from labor. Unfortunately we do not have data on wages or unit prices of labor services in trade or the shadow unit price of benefits from self employment in trade. We have approximated the development of this cost component by using the (natural log of the) South African real wholesale price index as an explanatory variable. South Africa is the dominant economic power in the region and a large chunk of Malawi merchandise imports originate from South Africa (31 to 44%, 2000-2007). Hence, we assume that producer prices in South Africa – properly converted to control for South African rand – Malawi kwacha exchange rate movements - are a suitable approximation for the development of costs in the Malawi domestic trade industry. Moreover, it is also investigated if South African

---

[11] Since nearly all trade transactions are completed well within the period of a month, it appears appropriate to investigate arbitrage behavior with monthly (or even weekly) data.

rand – Malawi kwacha exchange rate independently explains transaction costs, reflecting the costs of (other) imported inputs.

A part of the costs of domestic trade is related to the location where the trader is based or the merchandise is sourced. Examples of these costs are information costs, collection and bagging costs, costs of loading, storage costs and interest cost on working capital or opportunity cost of capital. In the estimation of transaction costs we have controlled for source costs by including a set of seller dummies. Another part of the costs of domestic trade is related to the location where the merchandise is sold, the destination market. Examples of these costs are the costs of unloading, information costs, costs of unloading, taxes and levies on transactions, market fees, quality and weight verification costs, etc. In the estimation of transaction costs we have controlled for destination costs by including a set of buyer dummies. Since costs on seller and buyer side could very well be related to the price of maize we have also included either seller price or buyer price as an explanatory variable[12]. Including the buyer price will account for ad valorem taxes on sales of maize.

From the descriptive work on the Malawi maize market we know that there is distinct seasonality in maize prices (see Figure 3 in the previous section). Since arbitrage due to domestic trade flows are determined by prices in different locations it is likely that seasonality affects transaction costs as well. We control for seasonality by including a complete set of monthly dummies as explanatory variables in the transaction equation. We have also investigated the possibility of seller specific or buyer specific seasonal patterns by interacting seller or buyer dummies with monthly dummies.

Finally we have added a trend variable to account for technological developments and trend-wise developments in infrastructure (marketing infrastructure, transport infrastructure, etc). The trend variable takes a value of 1 at the start of the sample period and increase with one each month.

We cannot account for economies of scale since our transaction costs are by definition unit transaction costs and we do not have data on trade flows between districts. However, the widely dispersed size distribution of trader businesses in Malawi offers clear evidence of constant returns to scale in trade in Malawi (see Fafchamps et al. (2005)).

Formally, we estimate the following equation (all prices and unit values are converted to constant prices and all variables are transformed into natural logarithms):

---

[12] Since the dependent variable is the price differential – the price in the buyer market minus the price in the seller market, including both seller and buyer price will spoil the estimation.

$$\begin{aligned} tc_t^{ij} = {} & \beta_0 + \beta_1 dis^{ij} + \beta_2 pf_t + \beta_3 wpi_t + \\ & \beta_4 p_t^j + \Sigma\beta_5^j dsel^j + \beta_6 p_t^i + \Sigma\beta_7^i dbuy^i + \Sigma\beta_8^k dmonth^k + \beta_9 trend_t + \varepsilon_t^{ij} \end{aligned} \qquad (7)$$

where tc_{ijt} = *transaction cost of trading goods from i to j at time t*; dis_{ij} = *distance between district i and district j*; pf = *transport fuel price*; wpi = *wholesale price index*; p_j = *price level at seller side*; $dsel_j$ = *dummy for seller j*; p_i = *price level at buyer side*; $dbuy_i$ = *dummy for buyer i*; $dmonth_k$ = *dummy for month k*; $trend$ = *trend variable*; ε *is an error term with a zero mean and constant variance* $(\varepsilon \,\square\, (0,\sigma^2))$

$$\begin{aligned} and: & \ \partial tc/\partial dis > 0; \partial tc/\partial pf > 0; \partial tc/\partial wpi > 0; \partial tc/\partial p_j > 0; \partial tc/\partial sel_j > 0; \partial tc/\partial p_i > 0; \\ & \partial tc/\partial dbuy_i > 0; \partial tc/\partial trend < 0 \end{aligned}$$

In equation (7) ε_{ijt} is an error term with zero mean and constant variance ($\varepsilon_{ijt} \sim (0,\sigma^2)$. Partial derivatives with respect to distance, transport fuel price, whole sale prices, price level at seller / buyer side to be positive, and with respect to trend to be negative. The sign of the seller and buyer dummies and the sign of the monthly dummies are either positive or negative.

Selection of price differentials for estimation of transaction cost equation

Transaction costs need to be inferred from observed price differentials. However, transaction costs can only be observed from a sub-sample of these observed price differentials. If two locations are known to be connected by trade flows, the price differential between them will be equal to (unit) transaction costs. Therefore we proceed by identifying locations that can be expected with high probability to be connected by trade flows and we explain in the following lines how we identify locations.

Production of staple food in developing countries is dominated by subsistence farming: only a limited quantity of produced staple food is marketed. Given subsistence conditions, the decision to sell produced staple food to the market and to trade with other districts is related - in the first place - to the district staple food requirements relative to district staple food production[13]. It should be noted that both theoretical and empirical literature emphasize transaction cost as a major determinant to sell to the market (see Key, Sadoulet and de Janvry (2000); Fafchamps and Vargas Hill (2005) and also below). However, with limited or too little maize supplies for home consumption, risk averse households will be reluctant to sell their own maize, especially if future maize prices are uncertain. Excess availability of staple food is a necessary condition for trade – or "export" – to other districts. Conversely, districts that are eager to buy maize will, in general, be districts that face a shortage of maize. The per district balance between last season district production of maize available at the start of the new crop year and expected district requirements of maize for consumption in the course of the crop year establishes if a district is a surplus district (a positive balance) or a deficit district (a negative balance). Hence, potential seller districts satisfy the following condition:

---

[13] We assume that cultivation practises and production technology are similar between districts and hence unit production costs of the staple food are more or less similar between districts.

$$Q_{0,t}^{i} > \sum_{m} E(fr_{t+m}^{i}) \qquad (8a)$$

$$\begin{aligned} &\textit{where } Q_0 = \textit{available production of staple crop from previous season,} \\ &\quad fr = \textit{food requirements per period,} \\ &\quad i = \textit{district},\ t = \textit{time},\ \mathrm{m} = \textit{number of periods till next harvest} \\ &\quad E = \mathrm{exp}\,\textit{ectation operator} \end{aligned}$$

And potential buyer districts satisfy:

$$Q_{0,t}^{i} < \sum_{m} E(fr_{t+m}^{i}) \qquad (8b)$$

Crop estimates of maize production by district are made on a regular basis in various rounds in the course of the crop season and with increasing accuracy. Food requirements per district per period are calculated by multiplying population size with the daily calorie requirement, the number of days per month and calorie share of the staple food by district and divided by the calorie content per kg of the staple food (see the data account in the appendix for further details on the empirical choices made). We assume that there is no carry- over of stored maize between crop seasons, either because home district production is fully exhausted or surplus production is sold elsewhere to satisfy cash needs.

Next, we further assume that a country-wide food shortage is likely to confuse normal trading patterns, creates distress among producers, traders and consumers and disturbs prices in highly erratic ways. Conversely, we assume that regular patterns of domestic trade dominate in periods with sufficient supply of food to meet the requirements of the entire population. Periods with a country wide food shortage are identified by the condition:

$$\sum_{i} Q_{0,t}^{i} < \sum_{i}\sum_{m} E(fr_{t+m}^{i}) \qquad (9a)$$

And periods with country wide surplus are identified with the condition:

$$\sum_{i} Q_{0,t}^{i} > \sum_{i}\sum_{m} E(fr_{t+m}^{i}) \qquad (9b)$$

As mentioned above and on top of excess availability, the decision of households to sell produced staple food on the market trade depends on transaction costs (see Key, Sadoulet and de Janvry, 2000). Assuming that the production level of staple food is sufficient to feed the entire population and assuming that cultivation of the staple food takes place in every district (see section on the Malawi maize market) - economizing on transaction costs is achieved by obtaining staple food from nearby locations, or from locations as nearby as possible. In practice average distances of traded agricultural merchandise appear to be very limited (see Fafchamps, Gabre-Madhin and Minten, 2005).

$$m_t^{jk} = 0 \quad if \quad dis^{jk} > dis^{\max} \quad (and \quad m_t^{jk} \geq 0 \quad if \quad dis^{jk} \leq dis^{\max}) \qquad (10)$$

Hence, on the basis of these considerations, we postulate that a positive price differential between two districts ($p_i - p_j > 0$ where p = market price of the staple food per kg and $i \neq j$) exactly reflects transaction costs if district i is a deficit district and district j is a surplus district, if aggregate domestic production of food is sufficient to meet the expected requirements of the population and if the districts where domestic trade takes place are less than a certain maximum distance apart.

$$
\begin{aligned}
& tc_t^{jk} = p_t^j - p_t^k \\
& \quad \textit{subject to} \\
& \qquad (1) \quad p_t^j > p_t^k \\
& \qquad (2) \quad Q_{0,t}^j < \sum_m E(fr_{t+m}^j) \quad \textit{and} \quad Q_{0,t}^k > \sum_m E(fr_{t+m}^k) \\
& \qquad (3) \quad \sum_i Q_{0,t}^i > \sum_i \sum_m E(fr_{t+m}^i) \\
& \qquad (4) \quad dis^{jk} < dis^{\max}
\end{aligned}
\tag{11}
$$

In the review of the literature we referred to three regimes that characterize market integration: I. $p_{it} - p_{jt} - tc_{ijt} = 0$ (integrated with positive trade flows); II $p_{it} - p_{jt} - tc_{ijt} < 0$ (integrated, without trade flows) and III $p_{it} - p_{jt} - tc_{ijt} > 0$ (not integrated). In this section we propose a procedure to identify a subset of regime I observations in order to be able to estimate transactions costs. It is likely that substantially more observations than the ones selected for estimation of the transaction cost equation are part of regime I.

Measurement and identification of market integration

We measure market integration by evaluating the price differential minus the predicted transaction cost ($p_t^j - p_t^k - t\hat{c}_t^{jk}$). Values of this indicator that are equal to zero or negative are considered to reflect integrated markets. Hence, we have:

$$
\textit{markets are integrated if}: \qquad p_t^j - p_t^k - \hat{tc}_t^{jk} \leq 0 \tag{12}
$$

Likewise, the degree of market integration is simply the number of times the market integration indicator is equal to or below zero, or:

$$
\textit{degree of market integration}: \qquad \psi_t^{jk} = [n(p_t^j - p_t^k - \hat{tc}_t^{jk} \leq 0)]/n \tag{13}
$$

In case of the difference between the price differential and transaction costs equals zero theory tells us that market integration coincides with positive trade flows. Identification of the zero values of our indicator, within certain confidence intervals, is in principle possible if one assumes a particular distribution. Baulch (1997a) uses a normal and halfnormal distributions to estimate the bounds of the three regimes. However, elsewhere it is shown that the assumption of a normal distribution is restrictive and may not be supported by the data (see Barrett and Li (2002)). Therefore we propose to avoid this and keep the analysis distribution free by simply counting values that are equal to or below zero.

Finally, it should be noted that trade pairs with negative price differentials will always be identified as "integrated markets" and this applies to a substantial number of observations. Assuming that prices for all trade pairs are different ($p_{it} \neq p_{jt}$ for all $i \neq j$), 50% of all price differentials will be negative. With strictly positive (predicted) transaction costs the difference between negative price differentials and transactions costs will automatically also be negative and, hence, aggregated over all trade pairs a minimum of 50% of all observations will indicate market integration by definition. The spatial distribution of positive and negative price differentials is, however, informative. For the assessment of market integration one need to recall that it is also economically rational not to trade if the price differentials are negative. Negative price differentials are likely for typical surplus districts.

4. Estimations of transaction costs and market integration

Estimation of transaction cost equations

Transaction cost equations are estimated using the specification derived in the previous section (equation 7). Table 1 reports estimates of transaction cost equations. Estimation is by OLS. With the exception of trends and dummies, all variables are transformed into natural logarithms and all price variables are deflated with the Malawi consumer price index. Absolute *t*-statistics are given in parentheses (.) below the coefficient. We have implemented F tests to investigate equality of coefficients above and below the distance threshold and on equality of the coefficient of distance and fuel price or transport price index. For a part of the estimations the sample is restricted to surplus-deficit pairs ((1), (2), (5) and (6)). In the remaining equations we added nearly autarkic districts as deficit districts. Half of the cost equations ((5) to (8)) are estimated with inclusion of an (almost) complete set of monthly, seasonal and seller dummies of which we have omitted coefficients and statistics. We also do not report coefficient and t-values of the constant term. The restrictions for identification of a sub-sample of observations where equality of transactions costs and price differentials applies, are implemented as follows: a positive price differential (equation 11, first restriction), the deficit-surplus restriction (equation 11, second restriction) and the "no country-wide maize shortage" restriction (equation 11, third restriction) are directly imposed on the data set. The remaining sample is separated in observations below and above a threshold distance (related coefficients are labeled resp. 1 and 2). The threshold is determined with a grid procedure, from 100 to 250(km), with a step of 10(km), using the statistical significance of the key determinants as a selection criteria. Empirical evidence suggests that nearly all domestic trade is taking place within a radius of 250km (see evidence documented in Section 3, in particular the relation between distance and correlation of market prices and survey evidence on the distance between locations of purchase and sales of Malawi maize traders). Estimation results indicate that distance, fuel price, price level at buyer side and trend – within the maximum distance - are significant in all equations and have the right sign. Results on testing the equality of the coefficients are mixed: equality of the coefficients of fuel price and price level at buyer side, and SA wholesale price, for observations below and above threshold distance, is consistently rejected, in most cases at 5% level of confidence, in a few cases only at 10% level of confidence. The equality of the coefficient of distance is only rejected (and in some case weakly) in the specification without dummies.

Table 1 Estimating Transaction Cost Equations (data from June 1999 to October 2009)

Dependent variable: real price differential $\ln(p_{j,t}-p_{i,t})$

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
distance (1)	β_{dis1}	0.1858	0.1688	0.1148	0.0934	0.1559	0.1541	0.1725	0.1723
		(3.4)	(2.9)	(2.3)	(1.8)	(2.0)	(1.9)	(2.4)	(2.4)
fuel price (1)	β_{fp1}	0.2054		0.2662		0.3702		0.3664	
		(2.5)		(3.6)		(2.5)		(2.8)	
SA wholesale prices (1)	β_{wpi1}		0.2053		0.3063		1.0776		1.1408
			(1.3)		(2.0)		(3.0)		(3.6)
maize price at buyer side (1)	β_{mp1}	1.2330	1.2008	1.3029	1.2701	2.051	2.087	2.072	2.0970
		(13.1)	(12.4)	(14.9)	(14.0)	(13.4)	(13.5)	(14.5)	(14.6)
trend (1)	β_{t1}	-0.0103	-0.0089	-0.0103	-0.0086	-0.0410	-0.0634	-0.0314	-0.0593
		(6.2)	(5.7)	(6.6)	(5.8)	(2.9)	(3.0)	(2.3)	(3.0)
distance (2)	β_{dis2}	0.3053	0.3304	0.3318	0.3685	-0.0003	0.0163	0.0444	0.0678
		(4.5)	(4.1)	(5.1)	(4.7)	(0.0)	(0.1)	(0.4)	(0.6)
fuel price (2)	β_{fp2}	-0.2328		-0.2387		-0.3215		-0.3102	
		(3.3)		(3.6)		(3.1)		(3.1)	
SA wholesale prices (1)	β_{wpi2}		-0.2405		-0.2180		-0.1300		-0.1159
			(1.7)		(1.6)		(0.5)		(0.5)
maize price at buyer side (2)	β_{mp2}	1.6062	1.6711	1.6163	1.6944	2.4285	2.4701	2.4538	2.4996
		(16.7)	(16.9)	(17.1)	(17.4)	(16.1)	(16.3)	(17.1)	(17.4)
trend (2)	β_{t2}	-0.0093	-0.0111	-0.0093	-0.0113	-0.0170	-0.0178	-0.0174	-0.0181
		(5.6)	(7.0)	(5.7)	(7.3)	(8.8)	(9.0)	(9.1)	(9.3)
Adjusted R2		0.2302	0.2276	0.2191	0.2159	0.3235	0.3215	0.3175	0.3164
Nobs		2299	2299	2840	2840	2299	2299	2840	2840
threshold distance		250	250	250	250	220	220	220	220
$F(\beta_{dis1} = \beta_{dis2})$		2.55	2.79	9.86	9.00	1.24	0.97	0.95	0.63
		(0.1102)	(0.0952)	(0.0017)	(0.0027)	(0.265)	(0.3260)	(0.3294)	(0.4262)
$F(\beta_{fp1} = \beta_{fp2})$		22.78		34.9		14.71		17.19	
		(0.0000)		(0.0000)		(0.0001)		(0.0000)	
$F(\beta_{tp1} = \beta_{tp2})$			15.49		23.29		7.28		9.87
			(0.0001)		(0.0000)		(0.0070)		(0.0017)
$F(\beta_{mp1} = \beta_{mp2})$		8.75	14.00	6.68	12.33	3.08	3.12	3.55	3.93
		(0.0031)	(0.0002)	(0.0098)	(0.0005)	(0.0794)	(0.0776)	(0.0595)	(0.0476)
$F(\beta_{t1} = \beta_{t2})$		0.20	1.18	0.21	1.93	2.70	4.45	0.97	4.19
		(0.6527)	(0.2780)	(0.6450)	(0.1644)	(0.1003)	(0.0350)	(0.3249)	(0.0409)
$F(\beta_{dis1} = \beta_{fp1})$		0.03		2.56		1.71		1.80	
		(0.8540)		(0.1099)		(0.1914)		(0.1796)	

The table reports estimates of transaction cost equations. Estimation is by OLS. All variables are in natural logarithms and all price variables are deflated with the Malawi consumer price index. The sample is separated in observations below and above a threshold distance (related coefficients are labeled resp. 1 and 2). The threshold is determined with a grid procedure, from 100 to 250(km), with a step of 10(km). Absolute *t*-statistics are given in parentheses (.) below the coefficient. Adjusted R2 = coefficient of determination adjusted for degrees of freedom and nobs = number of observations. F tests (with P values in brackets below the F statistic) on equality of coefficient above and below the distance threshold and on equality of the coefficient of distance and fuel price or transport price index are reported. For column (1), (2), (5) and (6) the sample is restricted to surplus-deficit pairs; in (3), (4), (7) and (8) we use a wider interpretation of deficit districts. Column (5) to (8) are estimated with inclusion of an (almost) complete set of monthly, seasonal and seller dummies of which we have omitted coefficients and statistics. We also do not report coefficient and t-values of the constant term.

Assessment of estimation results

In the previous section we discussed the statistical properties of the estimated equations, combined with the economic restriction that one may impose on coefficients. Another way to assess the estimation results of transaction cost equations is to compare predicted values of transaction costs with values of transaction costs that are obtained from surveys. For this purpose we make use of a survey of traders in Malawi from August 1999 – February 2000, and documented in Fafchamp et al. (2005)[14]. This survey is a representative sample of 738 Malawi traders. We have extracted data of traders who reported to have traded maize during the last 12 months, and within this group we used the data on maize transaction documented under "variable marketing costs of a completed purchase and sale transaction". This subset contains a number of 275 observations of maize transactions (before deleting outliers). We calculate transaction costs as sales price minus purchase price which follows the definition in the previous analysis and the estimations, but which is a slightly broader concept than transaction costs per se (as it includes trader profits). Average and standard deviation of transaction costs by supplying district, in level and in share, from this survey and from the predictions on the basis of the estimated transaction cost equations are summarized in Table 2.

Within-sample predictions show per unit transaction costs varying from 1.0 to 2.4 Kwacha per kg (with an average of 1.4 Kwacha per kg), and shares of transaction costs in terms of sales price varying from 10% to 28% (on average 16%). Since out of sample predictions include trade with remote areas it is no surprise to observe higher per unit transaction costs and shares of transaction costs of sales prices. In that case per unit transaction costs vary from 0.2 to 4.1 Kwacha per kg (on average 1.7 Kwacha per kg), and shares of transaction costs in terms of sales price vary from 2% to 44% (on average 15%). Survey observations have per unit transaction costs varying from 0.5 to 2.6 Kwacha per kg (on average 1.8 Kwacha per kg), and a share of transaction costs of sales price varying from 10% to 32% (on average 22%).

A closer look at Table 3 reveals that unit transaction costs and shares of transaction costs in sales prices are marginally higher in the survey data (around 22% in the survey data, around 15% in the predictions). This bias could be caused by construction of the survey data and the data underlying the predictions. The survey data are based on a single completed purchase and sale transaction per interviewed trader. The use of a single transaction, but also the choice of the transaction – which presumably entails a certain degree of selection bias – will give an upward bias to profitability of the transaction. Conversely, predictions are based on average monthly market prices. The process of averaging over a month will dampen the size of price differences. However, both data from the survey and prediction have large standard deviations and are not very accurate. Predicted transaction costs and predicted transaction cost shares are in the same order of magnitude as survey observations of transaction costs of domestic maize trade in Malawi. Statistically values of these variables from both sources will be the equivalent (though we did not test this formally). We conclude on the basis of the comparison reported in Table 3 that equality of the unit transaction costs and share of transaction costs, based on our predictions and based on survey data cannot be rejected.

---

[14] The data of this survey were kindly made available by Marcel Fafchamps.

Table 2 Comparison of predicted and observed transaction costs*

district (RPD)	**level**			**Share**		
	(1)	(2)	(3)	(1)	(2)	(3)
Dedza	2.4 (1.3)	3.3 (3.0)	1.3 (0.6)	0.236 (0.151)	0.310 (0.200)	0.190 (0.072)
Dowa		3.2 (3.3)	1.5 (0.7)		0.254 (0.181)	0.226 (0.084)
Kasungu		2.5 (2.3)			0.214 (0.144)	
Lilongwe		0.3 (0.3)	1.8 (1.1)		0.020 (0.015)	0.206 (0.101)
Mchinji	1.6 (0.6)	2.9 (3.0)		0.175 (0.082)	0.238 (0.165)	
Nkhota Kota		1.0 (1.1)			0.078 (0.056)	
Ntcheu	1.5 (0.8)	2.5 (2.8)	2.3 (1.1)	0.166 (0.099)	0.234 (0.186)	0.234 (0.106)
Ntchisi	1.4 (0.5)	2.3 (2.4)	2.2 (1.6)	0.129 (0.041)	0.178 (0.123)	0.202 (0.117)
Salima	1.6 (1.0)	2.1 (2.2)	2.6 (0.5)	0.173 (0.150)	0.181 (0.133)	0.297 (0.044)
Chitipa	1.4 (0.7)	4.1 (4.4)	1.2 (0.5)	0.286 (0.192)	0.439 (0.366)	0.181 (0.044)
Karonga		0.3 (0.3)	2.5 (1.5)		0.029 (0.027)	0.226 (0.163)
Mzimba		2.2 (2.3)	1.6 (1.0)		0.181 (0.130)	0.210 (0.109)
Nkhata Bay		1.2 (1.1)	3.3 (0.9)		0.099 (0.080)	0.323 (0.106)
Rumphi		3.4 (3.7)			0.305 (0.270)	
Balaka	1.0 (0.5)	1.3 (1.2)	2.5 (1.1)	0.117 (0.041)	0.113 (0.093)	0.305 (0.120)
Blantyre		0.2 (0.2)	1.0 (0.6)		0.020 (0.015)	0.122 (0.063)
Chikwawa		0.2 (0.2)	1.2 (0.4)		0.018 (0.013)	0.158 (0.048)
Chiradzulu		0.2 (0.2)			0.019 (0.014)	
Machinga		1.5 (1.7)			0.122 (0.096)	
Mangochi	1.0 (0.6)	1.6 (1.8)	1.6 (0.9)	0.124 (0.054)	0.134 (0.106)	0.251 (0.096)
Mulanje		0.3 (0.2)	0.5 (0.5)		0.020 (0.018)	0.099 (0.085)
Mwanza	1.1 (0.5)	1.6 (1.5)	2.4 (1.5)	0.101 (0.061)	0.122 (0.098)	0.282 (0.197)
Nsanje		0.3 (0.2)	1.6 (1.0)		0.023 (0.019)	0.237 (0.123)
Phalombe	1.3 (0.6)	2.2 (2.1)	1.1 (0.5)	0.115 (0.040)	0.154 (0.119)	0.167 (0.064)
Thyolo		1.3 (1.5)	1.6 (0.9)		0.115 (0.115)	0.235 (0.115)
Zomba	1.4 (0.7)	2.4 (2.3)	2.2 (0.9)	0.187 (0.087)	0.209 (0.154)	0.245 (0.042)

* The table reports predicted transaction costs in Malawi kwacha per kg of maize (constant 2000 prices), levels and shares expressed as a percentage of sales prices and summarized by supplying district. Column (1) are within sample predicted values, column (2) are within and out-of-sample predictions and column 3 is based on survey data on transaction costs from Fafchamps et al. (2005). Predictions are based on estimations presented in Table 1.

Assessment of market integration of Malawi maize markets

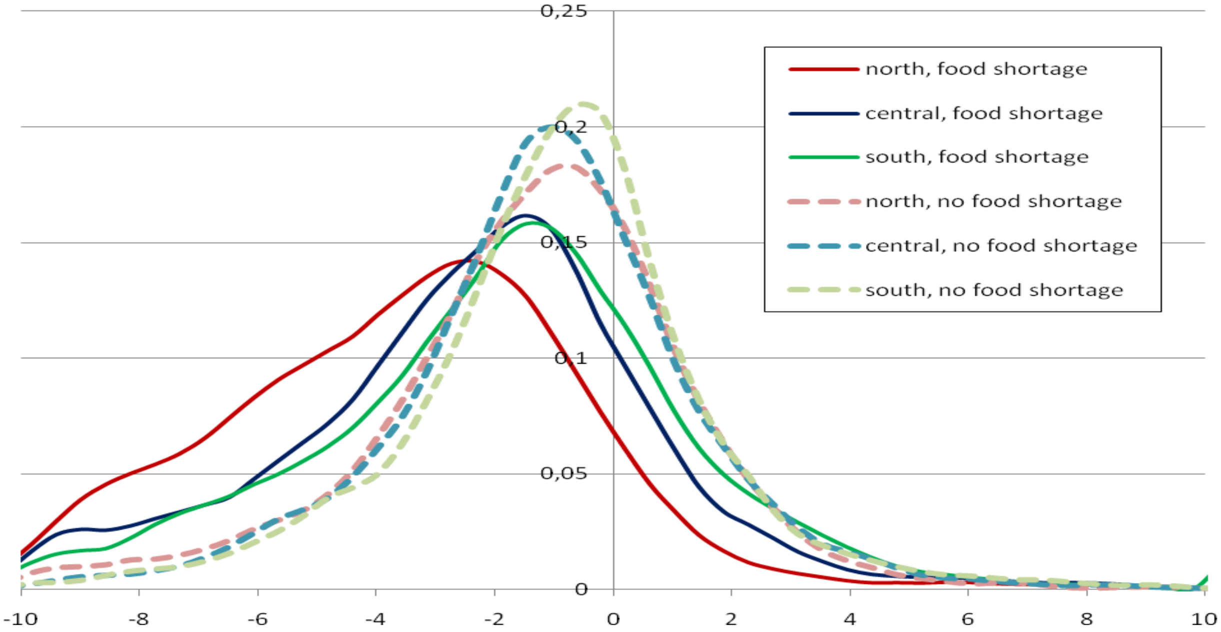
With estimates of transaction costs we are in the position to calculate the value of the price differential minus estimated transaction costs, for each trade pair, for both directions of trade and for each period. The share of values - per trade pair, district, region, period, etc - equal to and below zero characterizes the degree of market integration. Complete market integration is realized with values of 100%. The further below the 100% level the lower the degree of market integration. Any aggregate – either over time or over space – may be constructed in order to look at developments in market integration over time or between locations. The availability of a market integration indicator for each trade pair, for both directions and for each period allows to assess and compare market integration in many dimensions (between regions, urban versus rural, between periods of abundance and shortage, before and after implementation of economic policies, between seller and buyers, etc) and is, hence, a flexible tool to make assessments and comparisons of market integration.

Of particular interest is, of course, how market integration has developed in typical deficit districts. From the maize production and population data we derived the surplus and deficit districts: urban areas are notorious deficit districts (Blantyre and Lilongwe), but there are also a number of rural areas that face continuous food shortages. Rural deficit districts are Nsanje, Chikwawa, Machinga and Thyolo in the Southern Region, Dowa, Nkhotakota and Ntchisi in the Central Region and Karonga, Nkhata Bay and Rumphi in the Northern Region.

We have presented a selection of results in several formats: a probability density function of the market integration indicator by geographic region (Figure 5) and by urban and rural areas (Figure 6), a table with average degrees of market integration by buyer district (Table 3), and the development of the degree of market integration over time for the major (deficit) cities Lilongwe and Blantyre (Figure 7) and for several deficit districts over time, ordered by region (8 a to c).

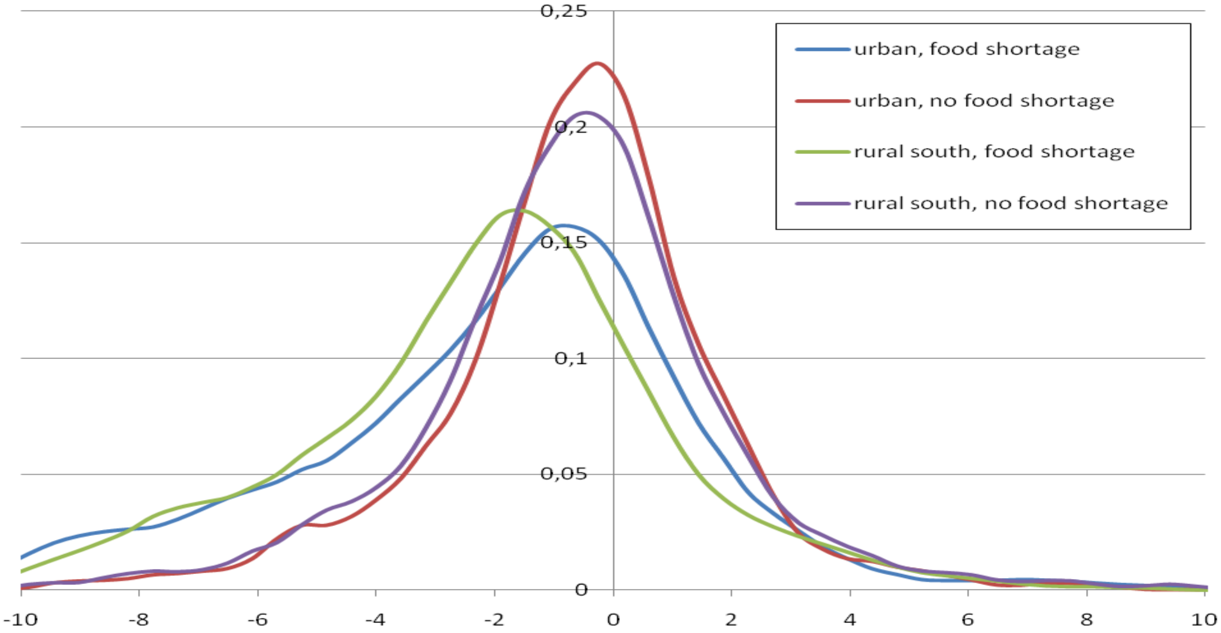
The density functions of the market integration indicator aggregate observations over time, partitioned in periods with and without food shortage (Figure 5 and 6). The area under the density function and below or equal to zero indicates market integration. For all regions, and both with and without food shortage the density function is nicely shaped with a peak somewhat below zero. Under food shortages the density functions flatten and shift to the left - for all regions, but most prominently the Northern region - indicating a larger proportion of observations below zero and, hence, a higher degree of market integration. This change is to a large extent due to increased transaction costs (see below), and it is therefore questionable if the improved market integration should be assessed positively. It may merely point at price difference that are not accompanied with trade flows. The comparison of urban deficit areas (Blantyre and Lilongwe) and rural deficit areas shows a similar pattern: the evidence supports a higher impact of higher transaction costs in rural areas (see Figure 6).

Figure 5 Market integration by region: with and without food shortage*



Source: own calculations
*The figure shows kernel density estimations of price differentials minus predicted unit transaction costs (p^j-p^k-tc^{jk}) by buying district and aggregated by region. Values below zero indicate market integration. The period of food shortage is from April 2001 to March 2003, and April 2005 to March 2004.

Figure 6 Market integration in urban and rural area: with and without food shortage



Source: own calculations
* See previous figure for explanations

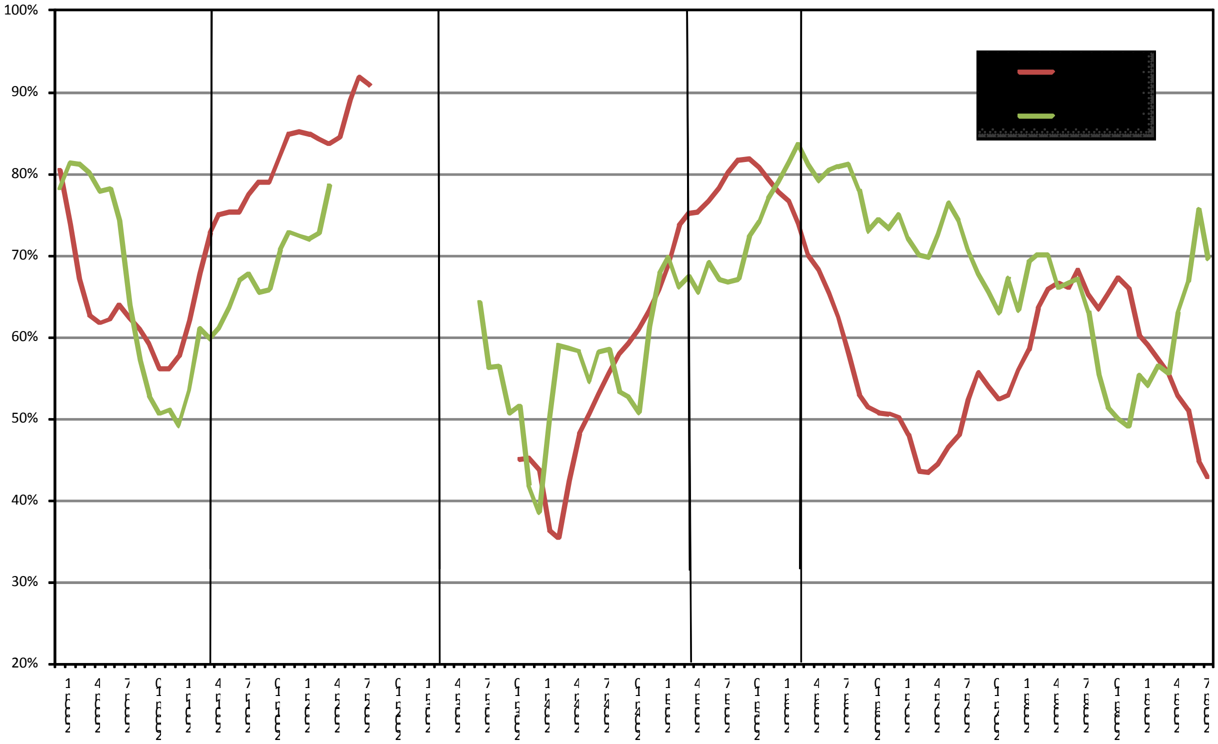
Table 3 Market integration by buyer district

district	no food shortage	food shortage	change in unit points		district	no food shortage	food shortage	change in unit points	
	mean	mean	total	due to tc		mean	mean	total	due to tc
Dedza	0.908 (0.29)	0.952 (0.21)	0.044	0.101	Rumphi	0.604 (0.49)	0.892 (0.31)	0.288	0.034
Dowa	0.686 (0.46)	0.781 (0.41)	0.096	0.136	Balaka	0.617 (0.49)	0.741 (0.44)	0.124	0.114
Kasungu	0.736 (0.44)	0.860 (0.35)	0.124	0.213	Blantyre	0.649 (0.48)	0.710 (0.45)	0.061	0.233
Lilongwe	0.578 (0.49)	0.807 (0.40)	0.229	0.075	Chikwawa	0.550 (0.50)	0.718 (0.45)	0.168	0.153
Mchinji	0.777 (0.42)	0.769 (0.42)	-0.008	0.201	Chiradzulu	0.792 (0.41)	0.794 (0.41)	0.001	0.229
Nkhotakota	0.575 (0.49)	0.748 (0.43)	0.173	0.182	Machinga	0.622 (0.49)	0.809 (0.39)	0.188	0.197
Ntcheu	0.847 (0.36)	0.884 (0.32)	0.038	0.115	Mangochi	0.723 (0.45)	0.746 (0.44)	0.022	0.203
Ntchisi	0.583 (0.49)	0.858 (0.35)	0.275	0.122	Mulanje	0.718 (0.45)	0.798 (0.40)	0.080	0.247
Salima	0.675 (0.47)	0.836 (0.37)	0.160	0.153	Mwanza	0.674 (0.47)	0.653 (0.48)	-0.021	0.182
Chitipa	0.875 (0.33)	0.980 (0.14)	0.105	-0.075	Nsanje	0.712 (0.45)	0.889 (0.31)	0.177	0.107
Karonga	0.664 (0.47)	0.954 (0.21)	0.291	-0.019	Phalombe	0.736 (0.44)	0.593 (0.49)	-0.143	0.304
Mzimba	0.774 (0.42)	0.848 (0.36)	0.074	0.129	Thyolo	0.643 (0.48)	0.807 (0.39)	0.164	0.132
Nkhata Bay	0.604 (0.49)	0.868 (0.34)	0.263	0.071	Zomba	0.765 (0.42)	0.892 (0.31)	0.127	0.113

* The table reports market integration as the average share of cases where the price differential minus unit transaction costs is equal or below zero $(n(p^j-p^i-tc^{ij})<=0)/n(total))$; food shortage periods are April 2001-March2003 and April 2005-March 2006; the change attributed to transaction costs is calculated as the difference between the change of the number of negative values of $(p^j-p^k-tc^{jk})$, from food shortage to no food shortage, and the change of the number of negative values of (p^j-p^k), also from food shortage to no food shortage. In formula: $[n((p^j-p^i-tc^{ij})<=0)_{\text{food shortage}} - n((p^j-p^i-tc^{ij})<=0)_{\text{no food shortage}}] - [n((p^j-p^i)<=0)_{\text{food shortage}} - n((p^j-p^i)<=0)_{\text{no food shortage}}]$. Typical deficit districts (according to production and population data) are shaded.

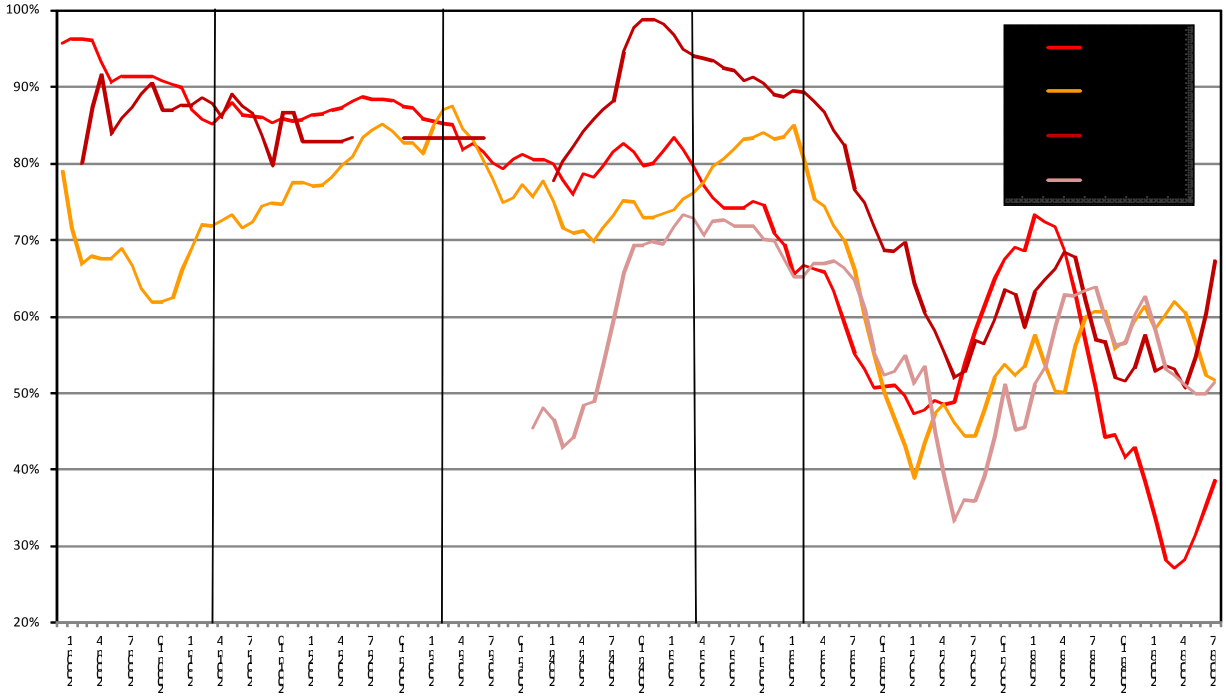
It is possible to quantify the exact area under the density function with values of the market integration indicator below zero. These numbers are presented in Table 3 sorted by buyer district, under food shortage and without food shortage. Typical deficit districts are shaded. The table supports the assertion that market integration is higher during food shortage periods: with average market integration during food shortages of around 82% and without food shortage of around 69%. For specific districts, and especially in the northern region this difference is more pronounced (up to 30% points for Karonga, Ntchisi Rumphi and Nkhata Bay), but also other districts show large differences (e.g. Lilongwe, Machinga, Nsanje). Surplus districts show relatively high market integration (a typical surplus district will have a negative price differential) and relatively small differences between food shortage and no food shortage periods (see e.g. Dedza, Ntcheu, Chitipa, Phalombe and Mwanza).

Figure 7 Market integration: from rural surplus districts to urban deficit areas



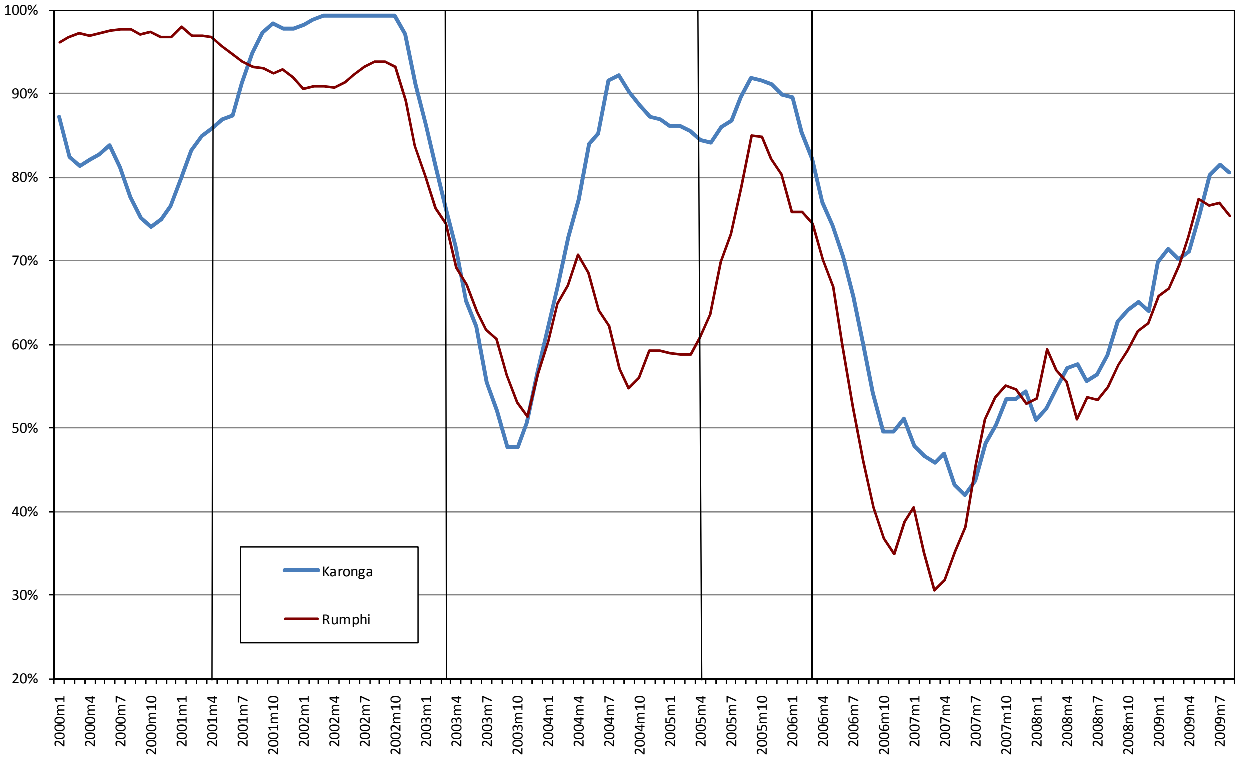
Source: own calculations

Figure 8a Market integration: from rural surplus districts to rural deficit districts (S)



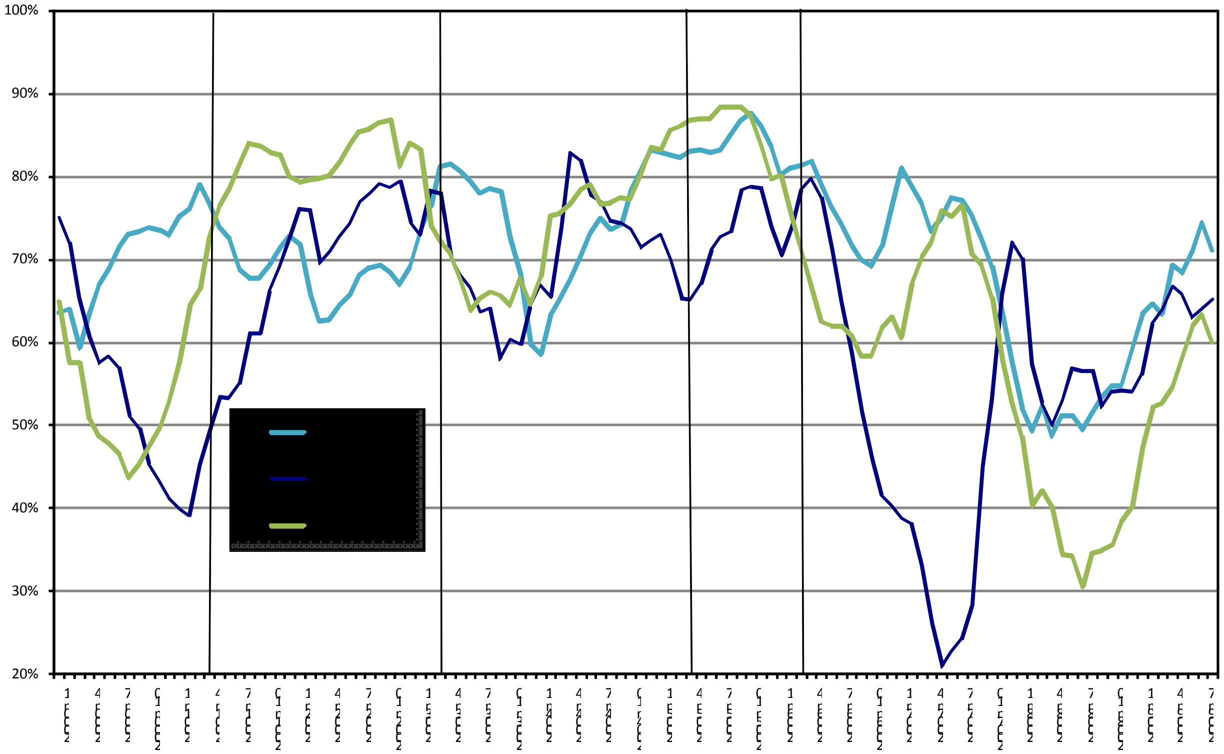
Source: own calculations

Figure 8b Market integration: from rural surplus districts to rural deficit districts (N)



Source: own calculations

Figure 8c Market integration: from rural surplus districts to rural deficit districts (C)



Source: own calculations

The contribution of transaction costs in the change in market integration from food shortage to no food shortage periods is calculated by comparing the observed change with the change that would arise when market integration is calculated on the basis of price differentials alone and, hence, by ignoring transaction costs. In formula we calculate : $[n((p^j-p^i-tc^{ij})<=0)_{\text{food shortage}} - n((p^j-p^i-tc^{ij})<=0)_{\text{no food shortage}}] - [n((p^j-p^i)<=0)_{\text{food shortage}} - n((p^j-p^i)<=0)_{\text{no food shortage}}]$. In nearly all districts a large part of the change in market integration in times of food shortage can be attributed to transaction costs, leading on average to a 14%-points increase in market integration.

The degree of market integration by month offers insight in developments over time. Calculated series by buyer district are highly volatile which is most likely due to the limited quality of the data (both measurement error and missing values). Recall also that for each month each district has - at most - 25 values of price differentials minus transaction costs ($p^j-p^i-tc^{ij}$): one should expect large fluctuations if the number of observations decreases. To avoid such wild fluctuations we have calculated 12-months centered moving averages. Results are presented for the major urban deficit areas Lilongwe and Blantyre (Figure 7) and selected rural deficit in the Southern region (Figure 8a), in the Northern region (Figure 8b) and in the Central Region (Figure 8c). Periods of food shortages are April 2001-March2003 and April 2005-March 2006: vertical lines in the figures indicate starting and ending dates. The figures confirm higher degrees of market integration during food shortages and overall higher degree of market integration in the Northern region, both indications of high transactions costs. Another common pattern is the overall lower degree of market integration in 2007, 2008 and 2009. With the available information we can only speculate about the reasons for this increase in market failure: large harvests during these periods possibly have made districts more self-sufficient and autarkic, allowing price patterns that justify trade, but without an infrastructure to support trade flows.

Table 4 Unexploited but profitable trade opportunities

District	(1) $n(p^j-p^i-tc^{ij}>0)$		(2) as (1), no food shortage		(3) as (2), surplus / deficit district	
	number	share	Number	Share	number	share
Nsanje	436	25.1%	125	27.1%	99	79.2%
Thyolo	666	31.1%	164	32.9%	14	8.5%
Machinga	713	32.6%	226	43.1%	51	22.6%
Blantyre	662	33.7%	164	32.0%	133	81.1%
Karonga	548	25.4%	170	33.1%	37	21.8%
Chiradzulu	294	20.7%	86	28.7%	40	46.5%

Source: own calculations

A final piece of evidence concerns violations of the competitive spatial arbitrage conditions, or trade pairs which are characterized with a positive difference between the price differential and transactions costs ($p^j-p^i-tc^{ij}>0$). If $p^j-p^i-tc^{ij}>0$ potentially profitable arbitrage opportunities are not exploited. However, there are little gains from trade to be made if there is no merchandise to trade. In marketing years with a country-wide food shortage there will be little maize to trade. Moreover prices in thin markets tend to be very volatile and less reliable to base arbitrage decisions on. Hence, to reveal real unexploited trade opportunities we better find to what extent a positive difference between price differentials and transaction costs ($p^j-p^i-tc^{ij}>0$) occurs outside periods of food shortages. Likewise, it is interesting to summarise to what extent a positive difference between price differentials and transaction costs ($p^j-p^i-tc^{ij}>0$) coincides with a surplus in district i and a deficit in district j[15]. Combinations of such circumstances ($p^j-p^i-tc^{ij}>0$, no country-wide food shortage and a surplus in district i and a deficit in district j) are likely to indicate barriers to trade. With the calculated degree of market integration we can detect the exact location of these barriers (by buyer, by seller or by trade pair). Our evidence reveals several buyer districts - in particular Nsanje, Blantyre and Chiradzulu. - with values of the market integration indicator falling into this category (see Table 4). Although we cannot identify causes of market failure, the evidence does suggest barriers to trade. These barriers to trade are most likely due to missing trade and marketing infrastructure, inadequate information and communication infrastructure, lack of credit for financing trade, poor roads and transport facilities and a shortage of trucks for transportation of agricultural merchandise.

The indicator of the degree of market integration suggests that markets appear to be well integrated during food shortages, with values often close to 100%, and much less during periods without a country-wide food shortage. We do not evaluate this as a positive sign since we believe that this is largely a reflection of high transaction costs during food shortages that make trade economically unattractive. Calculations of the contribution of transaction costs to these changes in the degree of market integration support this claim. Conversely we observe lower degrees of market integration if there is no food shortage. The significant drop in the degree of market integration in the years 2007, 2008 and 2009 – all years with substantial excess maize production – is also awkward: the impossibility to exploit trade opportunities suggests increased self-sufficiency of districts or barriers to trade, or both. Low levels of market integration observed outside periods of food shortages, as well as high transaction costs during food shortages both suggest that trading infrastructure is not sufficiently developed. Similar conclusions were reached in a study on market integration in Mozambique (see Tostão and Brorsen, 2005).

5. Summary and conclusion

We have investigated measurement of market integration in staple food markets in developing countries. Consistent with competitive spatial market equilibrium and in the tradition of the Parity Bound Model, we assess market integration by calculating the difference between price differentials and transaction costs. The required data on transaction costs are obtained from estimated transaction cost equations. Estimated transaction cost equation take account of transport costs, fixed source costs, fixed destination costs, ad valorem taxes & levies, seasonality and technological change. In the estimations on the basis of data from the Malawi maize market all key explanatory variables are statistically significant. Predicted unit transaction costs and the share of transaction cost are of the same order of magnitude as those observed in survey data. The proposed approach to assess market integration makes it possible to exactly locate market failures and to compare regions and periods. The degree of market integration in the Malawi maize market is shown to be higher during food shortages, especially in deficit districts and the evidence suggests that this is caused by higher transaction costs. The northern region also shows notoriously high levels of market integration which are likely to be due to high transaction costs. A remarkable drop in market integration is seen in many deficit markets in the years 2007, 2008 and 2009, years with relative abundance of maize. A few buyer districts are identified to have a large incidence of market failure suggesting barriers to trade. Overall we conclude that high transaction costs during

---

[15] It should be noted that we cannot verify to what extent excess supply by district is exhausted by sales to other districts in the course of the marketing year.

food shortages and low levels of market integration observed outside periods of food shortages both suggest that trading infrastructure is not sufficiently developed.

References

Baffes, J., 1991, 'Some Further Evidence on the Law of One Price', American Journal of Agricultural Economics, 73 (4), 1264-1273.

Barrett, C.B., 2001, 'Measuring Integration and Efficiency in International Agricultural Markets', Review of Agricultural Economics, 23 (1), 19-32.

Barrett, C.B. and J.R. Li, 2002, 'Distinguishing between Equilibrium and Integration in Spatial Price Analysis', American Journal of Agricultural Economics, 84 (2), May, 292-307.

Baulch, B., 1997a, 'Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration', American Journal of Agricultural Economics, 79 (2), May, 477-487.

Baulch, B., 1997b,'Testing for Food Market Integration Revisited', The Journal of Development Studies, 33, 4, April, 512-534.

Dercon, S., 1995, 'On Market Integration and Liberalization: Method and Application to Ethiopia', Journal of Development Studies, 32, 1, 112-143.

Fackler, P.L. and H. Tastan, 2008, 'Estimating the Degree of Market Integration', American Journal of Agricultural Economics, 90 (1), February, 69-85.

Fafchamps, M. E. Gabre-Madhin and B. Minten, 2005, 'Increasing Returns and Market Efficiency in Agricultural Trade', Journal of Development Economics, 78, 406-442.

Fafchamps, M. and R. Vargas Hill, 2005 'Selling at the Farmgate or Traveling to Market' American Journal of Agricultural Economics, 87, 3, 717-734.

Lutz, C., W. E. Kuiper and A. van Tilburg , 2007, 'Maize Market Liberalisation in Benin: A Case of Hysteresis', Journal of African Economies, 16, 1, 102-133.

Tostão, E. and B.W. Brorsen, 2005, 'Spatial price efficiency in Mozambiques's post-reform maize markets', Agricultural Economics, 33, 205-214.

Ravallion, M., 1986, ' Testing Market Integration', American Journal of Agricultural Economics, 68, February, 102-109.

Sexton, R.J., C.L. Kling and H.F. Carman, 1991, 'Market Integration, Efficiency of Arbitrage, and Imperfect Competition: Methodology and Application to U.S.Celery', American Journal of Agricultural Economics, 73 (3), August, 568-580.

Data account and data construction

We use a comprehensive set of monthly data by district or Rural Development Project (RPD), covering all 26 districts / RPDs of Malawi (see Table A1 in this appendix for an overview of districts / RPDs, ADDs and regions), from June 1999, at the earliest, to December 2009 at the latest. Data on monthly maize prices by district are publicly available through the Food Security Updates of the Famine Early Warning System Network (FEWS NET; www.fews.net), which obtains these data from the Ministry of Agriculture and Food Security in Malawi (MoAFS; formerly the Ministry of Agriculture). FEWS NET has reported Malawi maize prices of an - over the years - increasing number of individual markets in Malawi (up to 68 different markets): we have selected the market for each district with the largest number of observations over the years, which were - in nearly all cases - maize prices in major district towns. The maize price data are from January 2000 to August 2009. The price series by district are not complete: Table A2 in this Appendix summarizes the characteristics of the price data, including the number of missing observations by district and for Malawi as a whole. Out of the potential of 26x116 = 3016 monthly price observations (26 districts, 116 months) we have 2304 independent observations, around 77%, with substantial variation of data availability between districts (see Table A2). We have refrained from filling white spots by interpolation since this may disturb the verification of the spatial arbitrage that governs domestic trade.

Annual maize production by district (RPD) are also made publicly available through the Food Security Updates of the FEWS NET (www.fews.net), that also obtains this information from MoAFS. Production of maize is estimated in several rounds in the course of the crop year. The bulk of the data on annual production by district refer to fourth and final round crop production estimates. For a few crop years – notably crop years 2006/2007, 2007/2008 and 2008/2009 – third round estimates on Agricultural Development Division (ADD) level are converted to RPD / district level on the basis of previous year shares. Annual population data by district are from the National Statistical Office of Malawi (NSO, Malawi in Zomba; www.nso.malawi.net). Annual population series are converted to monthly data by interpolation. Monthly maize requirements by district in kg are calculated by multiplying monthly population with the daily calorie requirement, the number of days per month and maize calorie share by district and divided by the calorie content per kg of maize. The maize calorie content is 3570 kcal per kg of maize. The maize calorie share in consumption by district is from the Malawi (MVAC, 2003). Several values of the daily calorie requirements can be used. The Malawi Vulnerability Assessment Commission (MVAC) uses a minimum dietary requirement of 2100 kcal per day per head (see MVAC, 2003). From 1999 to 2005 the actually observed per day per capita energy supply varies from 2157 to 2217 kcal according to the food balances of FAO (see FAOSTAT). Since we run these calculations to identify surplus and deficit districts and to calculate excess maize available for trade, we stayed on the safe side and we have used a per capita per day calorie intake of around 2300 in the calculations of district requirements. We calculate maize requirements per district over the marketing season and compare this to available district production at the start of the marketing season. If production exceeds /falls short of demand a district is labeled as a surplus / deficit district. The assessment of the balance between requirements and production by district is made at the start of the marketing season (when crop estimates are known)[16].

---

[16] We expect the price difference between deficit and surplus districts to be positive: this is true for around 65% of the observations (hence, for around 35% of the observations we have $p_{deficit}-p_{surplus}< 0$). We lack information to explain negative price differences between deficit and surplus districts.

Opportunities for district-to-district trade are assumed to arise, certainly, between surplus districts and deficit districts.

All prices and unit values are deflated with consumer price index to allow comparisons over time. Monthly consumer price index numbers are obtained from the National Statistical Office of Malawi (NSO, Malawi in Zomba; www.nso.malawi.net), which differentiates between consumer price index numbers for urban and rural areas, and from the International Financial Statistics of the International Monetary Fund. Monthly Malawi kwacha – US$ and South African rand - US$ exchange rates (both period averages), South African Wholesale prices (index) and Monthly US gasoline prices in US$ per gallon are from the International Financial Statistics of the International Monetary Fund. We assume that monthly US gasoline prices in US$ per gallon, converted to Malawi kwacha adequately reflect Malawi domestic transport fuel prices since these prices - on an annual basis - have a close-to-one correlation with annual Malawi import unit prices of diesel (taken from NSO Statistical Yearbook (see www.nso.malawi.net)). Distances in km between district towns are obtained from the Travel Distance Calculator (www.mapcrow.info). In the descriptive section on Malawi and the Malawi maize market we have used various data sources from the National Statistical Office (of Malawi), in particular the Statistical Yearbook, Integrated Household Survey (2005), Welfare Monitoring Survey 2007 and the publication "Malawi, An Atlas of Social Statistics" (Benson, T., J. Kaphuka, S. Kanyanda and R. Chinula, 2002, "Malawi, An Atlas of Social Statistics", NSO, Malawi / IFPRI, Washington). Data for fuel import and total merchandise imports are from World Development Indicators (2009).

The dimension of the data set of price differentials is large: with 26 districts, we can calculate (26 x 26 -26) = 650 price differentials for each period. Our price data start in June 1999 and extend to October 2009, in all 116 monthly periods and, hence, we have 116 x 650 = 75,400 observations of price differentials. As the available price series are not complete (and in order to avoid poorly based inferences due to measurement error[17] we have deliberately refrained from filling these missing observations) a total number of 46,386 observations of price differentials remain. Trade from district j to district i can only be economically rationalized if p_i-p_j > 0, and, conversely, no trade is also economically rational if p_i-p_j < 0. Assuming that prices in all districts are different ($p_i \neq p_j$ for all $i \neq j$), at most 50% of the observed price differentials indicate potential trade opportunities. Hence, only half of these price differentials – 23,184 observations in total - are relevant for the analysis of transaction costs. It should be noted, however, that the spatial distribution of positive and negative price differentials – hence, all price differentials - is informative and is used in the assessment of market integration. Table A2 in the appendix summarizes the price data used in the empirical work including the enumeration of missing observations. Of course, a considerable number of the price differentials are from pairs of districts that are remote from each other and hence are not likely to trade with each other. This, however, should result from the empirical investigations. To get some sense of spatial dimension of the data set: if we impose the restriction to price differentials between locations less than 150 km apart, a sub-sample of around 7000 observations remains.

---

[17] Additionally, we do not know the quality of the price data: it is likely – as in most price data - that there is a certain degree of measurement error in these series, which may affect results. However, since we mainly use price differentials – differences of prices in different districts – certain type of some measurement error is swept out..

Table A1 From RDP to ADD and region

RDP (district)	RPD acronym	ADD	REGION
Dowa	dow	Kasungu	Central
Kasungu	kas		
Mchinji	mch		
Ntchisi	nsi		
Dedza	ded	Lilongwe	
Lilongwe	lil		
Ntcheu	neu		
Nkhota Kota	nkk	Salima	
Salima	Sal		
Chitipa	ctp	Karonga	North
Karonga	kar		
Mzimba	mzi	Mzuzu	
Nkhata Bay	nkb		
Rumphi	rum		
Blantyre	bla	Blantyre	South
Chiradzulu	crz		
Mulanje	mul		
Mwanza	mwa		
Phalombe	pha		
Thyolo	thy		
Balaka	bal	Machinga	
Machinga	mac		
Mangochi	man		
Zomba	zom		
Chikwawa	cww	Shire Valley	
Nsanje	nsa		

* RPD = Rural Development Project, ADD = Agricultural Development Division

Table A2 **Descriptive statistics for monthly maize prices in Malawi by district, Jan2000-Aug2009 (nominal monthly maize in Malawi kwacha per kg)**

RPD / District	max	min	mean	median	std	obs	mis(#)	mis(%)
Dowa	66.3	4.1	24.0	19.0	16.1	98	18	15.5%
Kasungu	66.2	6.2	24.5	18.5	15.5	75	41	35.3%
Mchinji	64.0	5.4	22.7	18.4	14.4	111	5	4.3%
Ntchisi	72.9	6.6	24.8	18.9	16.9	104	12	10.3%
Dedza	78.3	9.0	25.7	17.4	16.4	63	53	45.7%
Lilongwe	72.0	5.8	25.6	20.7	16.4	91	25	21.6%
Ntcheu	76.9	3.3	21.2	16.2	16.1	115	1	0.9%
Nkhotakota	79.5	5.9	24.8	19.9	16.8	103	13	11.2%
Salima	81.0	2.6	23.7	18.7	18.0	98	18	15.5%
Chitipa	66.7	2.8	19.3	14.9	14.1	109	7	6.0%
Karonga	64.2	4.4	21.9	16.4	15.8	112	4	3.4%
Mzimba	62.3	6.4	22.0	17.0	14.0	114	2	1.7%
Nkhata Bay	56.5	5.1	26.2	21.7	15.0	81	35	30.2%
Rumphi	89.3	3.3	23.4	19.8	17.2	109	7	6.0%
Blantyre	79.9	4.4	25.1	19.8	17.9	95	21	18.1%
Chiradzulu	70.0	10.4	29.8	20.3	16.3	60	56	48.3%
Mulanje	78.2	9.4	34.9	31.0	18.7	46	70	60.3%
Mwanza	92.0	13.2	32.4	24.6	19.1	61	55	47.4%
Phalombe	85.1	11.0	36.2	32.7	19.3	45	71	61.2%
Thyolo	95.5	3.8	23.7	17.0	18.4	110	6	5.2%
Balaka	80.5	4.7	26.1	20.0	18.0	84	32	27.6%
Machinga	79.3	5.2	23.8	18.4	16.0	116	0	0.0%
Mangochi	81.8	4.7	23.5	18.3	17.3	103	13	11.2%
Zomba	78.8	4.3	28.0	20.7	19.2	58	58	50.0%
Chikwawa	79.9	12.7	31.5	23.2	16.5	62	54	46.6%
Nsanje	91.3	4.5	26.3	21.1	18.8	81	35	30.2%
MALAWI	95.5	2.6	24.9	19.3	17.0	2304	712	23.6%

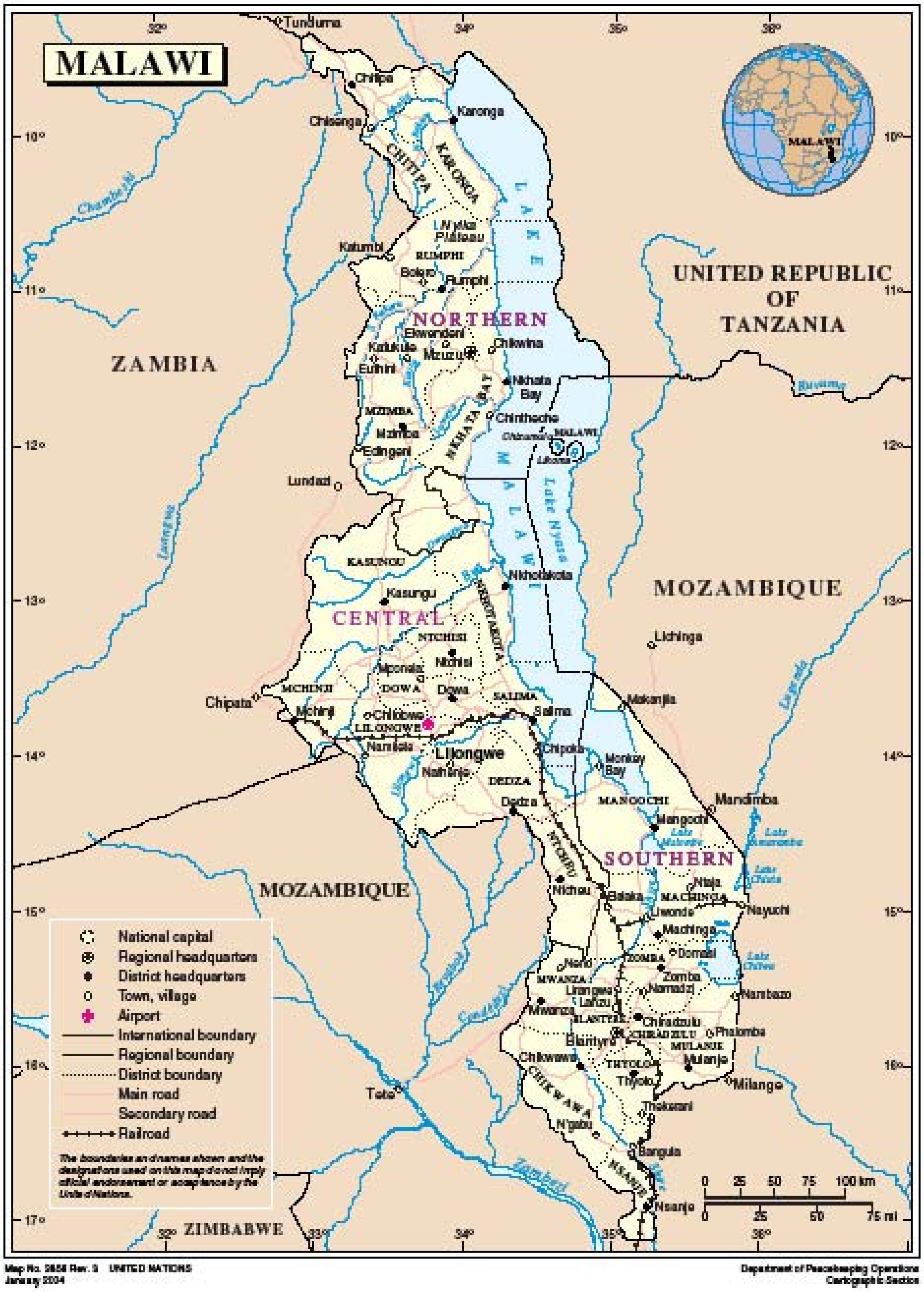
Source: calculations based on data from FEWS NET, MoAFS

Table A3 Spatial density of network

	number of district capitals within:		
	100 km	200 km	300 km
Dedza	5	16	20
Dowa	6	11	20
Kasungu	3	9	13
Lilongwe	5	11	21
Mchinji	1	7	13
Nkhotakota	4	9	15
Ntcheu	6	16	20
Ntchisi	5	11	18
Salima	5	11	21
Chitipa	1	2	4
Karonga	1	3	4
Mzimba	1	6	12
Nkhata Bay	2	6	11
Rumphi	1	4	8
Balaka	8	15	20
Blantyre	8	13	17
Chikwawa	7	13	16
Chiradzulu	9	13	17
Machinga	6	13	18
Mangochi	3	16	20
Mulanje	6	12	15
Mwanza	7	14	18
Nsanje	1	8	12
Phalombe	7	12	15
Thyolo	8	12	16
Zomba	10	13	18

Source: Travel Distance Calculator (www.mapcrow.info)

Figure A1 Map of Malawi



Source: United Nations, 2004