How Is the Liberalization of Food Markets Progressing? Market Integration and Transaction Costs in Subsistence Economies

Wouter Zant

We propose a modification of Baulch's parity bounds model to measure the market integration of food markets in developing countries. Instead of extrapolating a single observation of transaction costs, we estimate transaction costs. Predicted transaction costs compare well with survey data of traders. Probabilities of market regimes, computed on the basis of predicted transaction costs, fluctuate significantly and do not support fixed regime probabilities over time. The probability of market integration with trade decreases consistently during food shortages, increasing either the probability of no trade or loss-making trade or the probability of profitable but unexploited trade opportunities. The data further support a negative trend in market integration with trade. JEL codes: F14, Q13, Q17

In this paper, we investigate the measurement of market integration in domestic staple food markets in Malawi. Market integration is widely recognized as conducive to economic growth and poverty alleviation. A high degree of market integration implies smooth trade flows from surplus to deficit areas, improved transmission of price signals, less price volatility, production decisions that are made according to comparative advantage, gains from trade, and, hence, greater welfare. The increased integration of food markets in developing countries is considered to be of vital importance for agricultural transformation and economic growth (Fafchamps 1992). Market integration is also crucial from the perspective of food security. Many sub-Saharan countries face occasional food shortages as a result of crop failures, in turn caused by drought or other

Wouter Zant is researcher at VU University Amsterdam; email address: wouter.zant@vu.nl.

Acknowledgements: I would like to thank Jos van Ommeren for critical comments on previous versions of this paper. This paper also benefited from comments by Hans Binswanger, Stephan von Cramon-Taubadel, participants of the Annual Meeting of the Association of African Agricultural Economists in Cape Town, South Africa, September 2010, participants of the annual Conference of the Centre for the Study of African Economies in Oxford, UK, March 2011, colleagues at VU-University Amsterdam, Elisabeth Sadoulet, and three anonymous referees of this journal. The usual disclaimer applies. A supplemental appendix to this paper is available at http://wber.oxfordjournals.org/.

THE WORLD BANK ECONOMIC REVIEW, VOL. 27, NO. 1, pp. 28–54 doi:10.1093/wber/lhs017
Advance Access Publication July 9, 2012

climatic hazards. A well-integrated market for staple food potentially offers a mechanism for reducing the adverse impacts of these shocks by quickly moving food from surplus to deficit areas. Conversely, poorly integrated markets, such as those due to an inadequately functioning trading infrastructure, can hinder a smooth solution to or even deteriorate food shortages. The potential benefits of market integration have long been recognized. Under the auspices of the World Bank and the IMF, developing countries have adopted policies toward market liberalization, including the liberalization of domestic staple food markets.

How is this liberalization of food markets progressing? In this context, the question arises how to measure market integration. Popular approaches currently used to measure market integration focus on the co-movement of prices and make use of time-series techniques (cointegration, error correction and Granger causality). However, these approaches suffer from serious limitations. Attempts to link the measurement of market integration with spatial price equilibrium overcome these limitations (Baulch 1997a; Barrett 2001; Barrett and Li 2002; Park and others 2002; Negassa and Myers 2007; Moser and others 2009). With this alternative approach, known as the parity bounds model (PBM), probabilities for different regimes are estimated using a switching regression technique, and these probabilities characterize the market. A distinctive feature of Baulch's PBM is the incorporation of transaction costs. Transaction costs are not observed but are constructed by the extrapolation of a single observation of these costs.

In this paper, we propose a modification of the PBM. Instead of extrapolating fixed transaction costs, we estimate these costs. Furthermore, rather than estimating fixed regime probabilities for (a few) selected trade pairs using a switching regression technique, we propose a simple but economically justifiable technique of calculating regime probabilities that vary by location and time. This method allows us to track developments in market integration for any location in a network. Finally, and in contrast to previous work, we analyze and explain regime probabilities over the years. The empirical work is based on the Malawi maize market.

The paper is organized as follows. In section I, we discuss the empirical literature on market integration. In section II, we present a description of the Malawi economy, its maize market and maize trade. In section III, we propose a modification of the PBM. In section IV, we discuss estimation results of transaction cost equations, present calculated regime probabilities and explain variations in these regime probabilities. In section V, we present our conclusions.

Market integration in the literature

In this section, we briefly discuss and assess the two major lines of research on market integration, notably research using the co-movement in spatial prices and research using spatial price equilibrium as its starting point. The section aims to highlight where this paper fits in the literature.

Using Co-Movement in Prices to Assess Market Integration

A popular way to investigate market integration empirically is to use a time-series estimation of an equation that explains prices of one location using current and lagged prices of another location, employing the following formula: $p_t^j = \beta_0 + \sum_\tau \beta_{1\tau} p_{t-\tau}^k + \varepsilon_t$, where p_t^j (p_t^k) is the price in location j (k) at time t, and ε is a random error. Market integration is determined by (the sum of) the coefficient(s) of prices, which should be equal to one for complete market integration (Law of One Price), or $\sum_\tau \beta_{1\tau} = 1$. Alternatively, market integration is investigated by establishing that prices are integrated ($\varepsilon_t \sim I(0)$) using appropriate cointegration tests. Extensions allow for lagged prices in both markets, distinguish between short-run and long-run adjustments using an error correction framework, make use of Granger causality, and allow asymmetric adjustments toward equilibrium. Models allowing asymmetric adjustments are known as threshold co-integration models. Empirical and methodological examples of time-series approaches to market integration are surveyed in Fackler and Goodwin (2001), and recent examples include Lutz and others (2007) and Abdulai (2007).

There are several limitations to using the co-movement of prices for the study of market integration. Time-series techniques require the direction of trade to be fixed over time and cannot address trade reversals or discontinuities in trade. This issue is not trivial. Variations in production in rain-fed agriculture in developing countries are common and often large, and they may easily transform surplus areas into deficit areas from one crop year to another and back again with the accompanying trade reversals. Time-series techniques are also problematic if data are not complete, and incomplete data in developing countries' agriculture are the rule rather than the exception. The time-series approach also does not consider transaction costs. Because transaction costs often constitute a dominant component of market prices, especially in the case of staple foods with low production costs, omitting these costs will lead to flawed inferences. Transaction costs also tend to fluctuate independently from producer prices, both over time and across locations. Monte Carlo simulations show that conventional tests of market integration based on the co-movement of prices are flawed (Baulch 1997b; McNew and Fackler 1997). The disregard of transaction costs in time-series approaches and the need to adequately consider these costs has been highlighted in reviews (e.g., Fackler and Goodwin 2001).

Using Spatial Price Equilibrium to Assess Market Integration

One of the major limitations of approaches using the co-movement of prices is the lack of a theoretical foundation. A small body of empirical literature rejects the time-series approach on the abovementioned grounds and uses the Enke-Samuelson-Takayama-Judge spatial price equilibrium conditions as its starting point.[1] A seminal contribution is Baulch (1997a), which builds on earlier work from Sexton and others (1991) and Spiller and Wood (1986). Following the competitive spatial price equilibrium conditions, Baulch (1997a) identifies three regimes in his parity bounds model (PBM): in regime I, at the parity bounds, the price difference between two locations is exactly equal to the transaction cost ($p_t^j - p_t^k = tc_t^{kj}$); in regime II, inside the parity bounds, the price difference between two locations is lower than the transaction cost ($p_t^j - p_t^k < tc_t^{kj}$) and in regime III, outside the parity bounds, the price difference exceeds the transaction cost ($p_t^j - p_t^k > tc_t^{kj}$), where tc^{kj} represents the transaction cost of trading goods from k to j. Regimes I and II are consistent with competitive spatial price equilibrium and market integration. In the absence of adequate transaction cost data, Baulch (1997a) uses the observed transaction costs of a single cross-sectional observation and extrapolates them to all periods and all locations. Regimes are characterized by price differences less the extrapolated transaction costs. A switching regression technique is used to estimate the probabilities of the three regimes under the assumption of normally distributed errors within the parity bounds and half normally distributed errors above and below the parity bounds. The PBM is a flexible tool to diagnose market integration that overcomes the limitations of price co-movement techniques.[2] Extensions have focused on relaxing the assumption on the distribution of errors (Barrett and Li, 2002), complementing the analysis with trade flow data (Barrett 2001; Barrett and Li 2002), and shifts in probabilities (Park and others 2002; Negassa and Myers 2007). In this paper, we propose a modification of Baulch's original contribution (Baulch 1997a) and apply the proposed approach to the Malawi maize market.

Facts and Figures on Malawi

Prior to explaining and applying the proposed modifications to the PBM, we present a description of the Malawi economy, its maize market, and its domestic maize trade. The description provides background information that supports several assumptions underlying the empirical application.

The Malawi Economy and the Malawi Maize Market

Malawi is a landlocked country in sub-Saharan Africa that measures more than 800 km from north to south and 100 to 170 km from east to west. It is

1. In particular, $p_t^j - p_t^k - tc_t^{kj} \leq 0 \perp m_t^{kj} \geq 0$, where $p_t^{j(k)} =$ the price in the importing (exporting) market at time t, $tc_t^{kj} =$ per unit transaction cost from k to j at time t and $m_t^{kj} =$ trade flow from k to j at time t. The equation states that equilibrium requires that at least one of the conditions in the equation is satisfied with equality.

2. Another promising alternative direction in market integration research, in contrast to the PBM but also using the spatial price equilibrium as its starting point, is proposed by Fackler and Tastan (2008). They measure the degree to which an excess demand shock in one location is transmitted to another location.

bordered to the north and northeast by Tanzania, to the northwest by Zambia, and to the south, the southeast, and the southwest by Mozambique. Lake Malawi, a large Great Rift Valley lake called Lake Nyasa outside of Malawi, forms part of the country's eastern border and stretches from the northern tip to the lower third of the country. Malawi's population increased from 10.5 million in 2000 to 13.9 million in 2010. The central and, particularly, the southern regions are relatively densely populated. Lilongwe, in the central region, and Blantyre, in the south, are the two largest cities. Each city had approximately 0.5 million people in 2000, increasing to slightly below 0.9 million in 2010. The majority of the population lives in rural areas (approximately 87 percent, from 2000 to 2010) and earns income from agriculture (77 percent of the population above 15 years of age, 84 percent in rural areas). Per capita GDP in 2009 (2008), expressed in purchasing power parity US$, is between 840 and 900 US$, making Malawi among the poorest countries in the world.[3] The poverty rate in Malawi is high: 52.4 percent of the population is poor, and 22 percent of the population is very poor (IHS 2005). There is considerable geographical spread in poverty. Reported poverty rates vary from 70 percent in remote rural districts to 25 percent in urban centers.

In Malawi, maize is dominant in both production and consumption. It is the primary staple food of households. Between 52 and 65 percent of the total per capita calorie intake is from maize (MVAC 2003). Because of its higher population density, the south is the largest market for maize. Simultaneously, nearly all households—an estimated 97 percent—grow maize (IHS 2005). The production of maize is undertaken by households on subsistence grounds. A share of 81 percent of the population in rural areas is classified as subsistence farmers (IHS 2005). The availability of maize in the market is determined by the size of production and the degree of subsistence farming. Estimates of the share of domestic maize production sold on the market are in the range of 5 to 25 percent (IHS 2005; Jayne and others 2008). Variation in rainfall and occasional droughts cause large fluctuations in the production of maize, and countrywide shortages of maize occurred in the 2001/2002, 2002/2003 and 2005/2006 marketing seasons. These shortages led to inflows of emergency food aid from international institutions and donors. In times of food shortages, domestic maize prices peak to import parity levels or even exceed these levels. Observed peaks in maize prices confirm this assessment. However, large maize price increases in 2008 were caused by other factors. Similar to the prices of most agricultural products, Malawi maize prices follow a distinct seasonal pattern, with highs at the end of the marketing season just before harvesting (January to March) and lows directly after harvesting (May to July). Increases from the minimum to the subsequent peak in market prices vary from 30 to 40

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

percent, which is very large in light of the importance of maize as a staple food for Malawi households.

The privatization of agricultural trade was accomplished in the 1980s and 1990s (Christiansen and Stackhouse 1989; Smith 1995). The Agricultural Development and Marketing Corporation (ADMARC), which previously employed a countrywide storage, marketing, and distribution network to buy and sell maize with the aim to guarantee food security, particularly in remote areas, has become redundant in the maize market in recent decades. Since 1994, ADMARC maize purchases have been significantly below 10 percent of total production and were negligible from 2000 to 2005. Also, in terms of marketed output, which is much smaller than production but more relevant in the case of subsistence farming, the role of ADMARC is negligible. Survey investigations for the year 2008 indicate that on average, only 8 percent of the maize sold by farmers was purchased by ADMARC (Jayne and others 2008). Private traders were the main buyers of maize from small holders, accounting for about 75 percent of all maize sold. Outside of Lilongwe and Blantyre, ADMARC purchased marginal quantities in the rural areas surveyed.

Domestic Trade and Transport in Malawi

The transport of goods is undertaken by ship, by train, and by truck. Transport by ship between various ports of Lake Malawi is growing but accounts for a marginal share of all maize traded (less than 1 percent). The Malawi rail network is a rail line with a total length within Malawi of 797 km that runs from Zambia in the west through Lilongwe to the south, where it splits into a line further south to Blantyre and Beira in Mozambique and a line to the east, to Nacala in Mozambique. This rail network has not been fully operational in recent decades because of civil war in Mozambique, among other factors.[4] The bulk of domestically traded goods are transported by road. Transport by trucks is reported to account for 70 percent of the domestic freight of all goods (Lall and others 2009).

The maize trade in Malawi is undertaken by large farmers, small, medium, and large traders, wholesalers, maize processing firms, and ADMARC. Most district-to-district trade of maize is from farmers to small and medium traders and occasionally to larger traders and wholesalers. Approximately 75 percent of all traders buy directly from farmers and sell as retailers (Fafchamps and others 2005). Survey evidence indicates that less than 1 percent of all traders are involved in wholesaling as a stand-alone business. The dispersion of the size distribution of trader businesses and the prevalence of many small-scale businesses suggest constant returns to scale in trade (Fafchamps and others 2005). Evidence further suggests that the number of small traders operating in

4. Transport by ship and by train are currently of negligible importance, but their potential is significant because unit costs of transport by ship and by train are only a fraction of the unit cost of road transport.

rural areas is increasing (Jayne and others 2008). Trading channels vary by location, but without exception, only marginal quantities are sold to ADMARC. The bulk of the maize trade is in the hands of the private sector. There is a distinct pattern of trade over the season that is partly influenced by the timing of the harvest, liquidity constraints of farmers, and ADMARC's participation in the market (Jayne and others 2008).

By far the largest component of the cost of domestic maize trade is transport costs, with an estimated share of 48 to 57 percent of total transaction costs (Fafchamps and others 2005). Transport costs in Malawi are high because the secondary road infrastructure (the so-called feeder roads) is not well developed (Lall and others 2009). The quality of the major trunk road network (international routes) is good and is not a constraint, but for rural areas, the conditions are fundamentally different. Average unit transport costs per ton per kilometer from rural to urban areas are approximately 20 times higher than from urban to urban areas, which is caused by differences in the average distance (85 km in rural transport, more than 2000 km in international transport), the average truckload (2.5 tons in rural transport, more than 20 tons in international transport), the demand for transport services and the availability of backhaul cargo (Lall and others 2009). The type of trucks used on the international routes is also not appropriate for transport in rural areas (Lall and others 2009). Transport costs are also high because fuel for transport is expensive. The import and distribution of fuel for transport are managed by a parastatal of the Government of Malawi, which also sets pan-territorial pump prices. Domestic pump prices are high relative to import unit values and thus relative to world market prices, and they have increased stepwise over the years. The gap between import unit values and fuel prices at the pump is 80 to 220 percent. Annual increases in pump prices average approximately 26 percent and reached a peak between June 2008 and January 2009. High transport costs clearly restrict the maize trade. Survey data indicate that the average distance between the purchase and sale location of maize transactions is approximately 55 km, with a maximum of 200 km (Fafchamps and others 2005). Not surprisingly, the problem of high transport costs is particularly acute in remote and isolated areas.

All traders, regardless of size of operations, face high lending interest rates. High interest rates make the funding of commercial trading activities through bank credit virtually impossible. Trading transactions are therefore primarily self-funded. Survey information on traders indicates that traders supply, on average, 96 percent of the working capital of their trading business themselves and that approximately 88 percent of all traders are 100 percent funded with their own capital (Fafchamps and others 2005). Because the scale of trading operations is closely related to the availability of working capital, traders put considerable effort in preventing working capital from being tied up in investments for long periods. In practice, this situation leads to the inefficient and costly back-to-back trade of small quantities. Therefore, and out of economic necessity, the turnover time of working capital is short. The average (median) number of days between purchase and sale is nearly eight days (three days), with approximately 45 percent of transactions completed within two days and less than 10 percent completed in more than 14 days (Fafchamps and others 2005).[5]

Methodology

In this section, we explain the proposed adjustments to the PBM. A key element of this modification, the estimation of unit transaction costs, is extensively discussed. Next, we show how we measure market integration. Finally, we specify how we intend to explain fluctuations in market integration.

Proposed Adjustments to the Parity Bounds Model

To investigate market integration, we propose a modification of Baulch's original contribution (Baulch 1997a). In particular, we intend to model and estimate transaction costs. Transaction costs are postulated to be determined by transport costs, source costs, destination costs, trade pair costs, trends, and seasonality. We use predictions of transaction costs to calculate regime probabilities. To our knowledge, there is no behavioral regularity or theory that justifies the estimation of fixed regime probabilities. Therefore, our next contribution is to allow for the possibility that regime probabilities fluctuate over time and to conjecture that these fluctuations are informative. For this purpose, we exploit all available information in a network of trade, calculate the indicator of market integration $(p_t^j - p_t^k - tc_t^{kj})$ for each trade pair in each location and period, and track fluctuations in resulting regime probabilities over time for each location. The final contribution of this study is to proceed beyond merely measuring market integration and to explain developments in regime probabilities.[6] In particular, we investigate how regime probabilities correlate with food shortages and trend developments after controlling for district fixed effects.

Empirical Specification of Transaction Costs Equations

What are the determinants of transaction costs? The largest component of transaction costs involves transport costs. We use transport fuel prices as an approximation of unit transport costs. Pump prices of transport fuel in Malawi are pan-territorial: they vary over time, not between locations. Another aspect

5. These numbers support the appropriateness of investigating spatial arbitrage behavior with monthly data.

6. Conventional techniques to measure market integration, including both the PBM and co-movement techniques, are purely diagnostic tools. They have little to offer in explaining fluctuations in market integration. Nevertheless, speculations on potential determinants of market integration abound in these studies.

of unit transport costs is related to the distance between locations. Distance may account for costs, such as information costs, search costs, economies of scale and scope, road-block costs, and bribes. Therefore, the distance between locations is used to approximate these transport costs. Partial derivatives with respect to transport fuel price and distance are expected to be positive because costs in competitive markets will be passed on.

Transaction costs are partly related to the location where the merchandise is sourced. These costs are associated with the collection of information, the collection, bagging, loading, and storage of staple food and interest on working capital. Likewise, transaction costs are partly related to the location where the merchandise is sold, the destination market. Examples of these costs are information costs, costs of unloading, taxes, and levies on transactions, market fees, and quality and weight verification costs. In the estimation of unit transaction costs, we control for source costs and destination costs by including a set of seller and buyer dummy variables. Seller and buyer dummies are binary variables with a value of one for a specific buyer or seller and zero elsewhere. In the empirical estimations, we make use of district data: the district maize balance establishes whether a district is a buyer of seller district (see below).

To capture developments in road infrastructure and road quality, developments in transportation technology, changes in the structure of the trading sector, and developments in telecommunication infrastructure, a general trend variable is included. The trend variable takes a value of one at the start of the sample period and increases by one each period. Partial derivatives with respect to trends are expected to be negative, reflecting cost reductions caused by structural developments, such as the roll-out of the mobile phone infrastructure, which are shown to reduce the information and search costs of trade (e.g., Aker 2008; Jensen 2007; Muto and Yamano 2009).

There is a distinct seasonality in maize prices. Because the domestic maize trade is determined by maize prices in different locations, it is likely that seasonality in maize prices also translates into seasonality in transaction costs. Therefore, a set of monthly dummy variables—binary variables with a value of one for a specific month and zero elsewhere—is included to account for seasonality in transaction costs.

To further generalize the specification, we allow for seller-specific and buyer-specific seasonal patterns by interacting monthly dummies with seller and buyer dummies. The estimated equation is summarized as follows:

$$\begin{aligned}\ln(tc_t^{kj}) = \beta_0 + \beta_1 \ln(dis^{kj}) + \beta_2 \ln(p_{f,t}) + \beta_3 trend_t \\ + \mu_k + \nu_j + \delta_m + \lambda_{km} + \tau_{jm} + \varepsilon_t^{kj}\end{aligned}, \qquad (1)$$

where tc^{kj} is the transaction cost of trading goods from k to j; dis^{kj} is the distance between location j and k; p_f is the transport fuel price; trend is a trend variable; μ_k, ν_j, and δ_m are seller, buyer, and monthly fixed effects; λ_{km}, and

τ_{jm} are seller-monthly and buyer-monthly fixed effects; and ε^{kj} is an error term with a zero mean and a constant variance ($\varepsilon^{kj} \sim (0,\sigma^2)$).[7]

Selecting Price Differences for Estimation of Unit Transaction Costs

Observations of unit transaction costs are inferred from observed price differences or, more precisely, a subsample of these observed price differences. We assume competitive markets of maize and domestic trade services and profit-maximizing traders.[8] Under competitive conditions with nonzero trade flows between buyer and seller locations, a positive price difference is exactly equal to unit transaction costs:

$$tc_t^{kj} \equiv (p_t^j - p_t^k)_{|competitive\,equilibrium\,with\,positive\,tradeflows}. \tag{2}$$

Therefore, for each trade pair in each period, we must find exogenous conditions that are likely to correspond with competitive equilibrium with positive trade flows. First, because only positive transaction costs allow an economic rationalization for trade, we only consider positive price differences between locations. Next, we assume that trade flows are most likely to occur from surplus to deficit locations or from potential export locations to potential import locations. Under subsistence farming, the possibility of export and the need to import are determined by the previous season's production of maize available at the start of the marketing year relative to the expected requirements of maize for consumption in the course of the marketing year. Production above (below) the level of requirements characterizes a potential exporter (importer) or a surplus (deficit) location. The second condition for selecting observations of price differences that are likely to be equal to transaction costs is that we formally restrict the price differences to those trade pairs that connect surplus with deficit locations where potential surplus and deficit locations, respectively, satisfy the following conditions:

$$Q_{0t}^i > (1+\varphi)\sum_m E\left(fr_{mt}^i\right) \tag{3a}$$

$$Q_{0t}^i < \sum_m E(fr_{mt}^i), \tag{3b}$$

7. Additional interactions are possible but are set aside because they lead to complex interpretations of coefficients. Trade-pair specific trends are also not used in the estimations because of multicollinearity with distance.

8. Both maize markets and domestic trade services markets involve a homogeneous good, have a large number of buyers and sellers, have negligible entry or exit barriers, have no increasing returns to scale, and have no (or negligible) government involvement. The description in the previous section supports these assumptions.

where Q_0 is the available annual production of maize from the previous crop year at the start of the marketing year; fr_{mt} is the maize consumption requirement in month m for the current marketing year; φ is the excess over the maize requirements above which exports are triggered; i is the location; t is the year; and E is the expectation operator.

The available production of the previous season is known, and the expected maize consumption requirements are a transformation of population size, daily calorie intake, and dietary preferences (see appendix for details). We assume that there is no carry-over of maize between crop seasons, either because home production is fully exhausted or because surplus production is sold to satisfy the cash needs of households (see, e.g., reports of the Malawi Vulnerability Assessment Committee). Additionally, despite abundantly available storage capacity (Kutengule and others 2006), liquidity constraints, credit constraints, high interest rates, and a shortage of working capital make arbitrage between crop seasons economically unattractive and thus uncommon (see also the previous section). We assume that production must exceed total requirements with a number of months of food requirements (φ) in order to trigger exports, and we explore the sensitivity to different values of φ.

A countrywide food shortage is likely to disturb normal trading patterns. Food shortages lead to large price increases, which induce speculative and rent-seeking behavior on the supply side. In addition, food shortages give rise to inflows of food aid. Such inflows affect price formation in unpredictable ways. Both factors increase uncertainty in market prices, may lead to erratic price movements and create distress among producers, traders, and consumers. Conversely, regular patterns of domestic arbitrage trade prevail in periods when the production of staple food is sufficient to meet the requirements of the entire population. For these reasons, we condition the sample of price differences for the estimation of transaction costs on a countrywide maize surplus. Formally this condition can be written as follows:

$$\sum_i Q_{0t}^i > \sum_i \sum_m E\left(fr_{mt}^i\right). \tag{4}$$

Both the theoretical and empirical literature emphasize transaction costs, mainly transport costs, as the key determinant of the decision of households to sell produced staple food on the market (Key and others 2000; Fafchamps and Vargas Hill 2005). Assuming that the production level of maize is sufficient to feed the entire population, and assuming that the cultivation of maize occurs throughout the country (see the previous section), economizing on transaction costs is achieved by trading with the closest possible locations. Survey evidence also suggests that domestic short distance trade is very costly (Lall and others

2009) and that agricultural merchandise is only traded, on average, over moderate distances (Fafchamps and others 2005). This implies the following:

$$m_t^{kj} = 0 \text{ if } dis^{kj} > dis^{max}, \tag{5}$$

where m^{kj} is the trade flow from k to j, dis^{kj} is the distance between j and k and dis^{max} is the maximum distance of trade transactions.

In summary, we obtain observations of unit transaction costs—or a sample of price differences, which are likely to be equal to unit transaction costs—by conditioning price differences as follows:

$$tc_t^{kj} \equiv p_t^j - p_t^k$$

conditional on

$$\begin{gathered}
\text{(i)} \quad p_t^j > p_t^k \\
\text{(ii)} \quad Q_{0t}^j < \sum_m E\left(fr_{mt}^j\right) \; \textit{and} \; Q_{0t}^k > (1+\varphi)\sum_m E\left(fr_{mt}^k\right) \\
\text{(iii)} \quad \sum_i Q_{0t}^i > \sum_i \sum_m E\left(fr_{mt}^i\right) \\
\text{(iv)} \quad dis^{kj} \leq dis^{max}.
\end{gathered} \tag{6}$$

Measurement of Market Integration

Once we have estimates of transaction costs, the measurement of market integration and other regimes is straightforward. The indicator of market integration is the price difference minus the predicted transaction cost $p_t^j - p_t^k - \widehat{tc_t^{kj}}$. We calculate the market integration indicator $p_t^j - p_t^k - \widehat{tc_t^{kj}}$ for each trade pair, in each location and in each period; hence, for all j ≠ k, for all t, and for both buyers and sellers (districts). For a typical seller or seller district, spatial price differences and, thus, market integration indicators will be negative. In a network with n locations, this results in n-1 indicator values per location and per period. We claim that movements in the distribution of the market indicator values for a specific location are informative about the changes in the operation of the market of that location. In line with Baulch, we identify three regimes for each location j, empirically calculated as follows:

regime I, integrated with trade flows, efficient market:

$$\kappa < p_t^j - p_t^k - \widehat{tc_t^{kj}} < \kappa \tag{7}$$

regime II, integrated, no trade or loss making trade:

$$p_t^j - p_t^k - \widehat{tc_t^{kj}} \leq -\kappa \tag{8}$$

regime III, not integrated, unexploited arbitrage opportunities:

$$p_t^j - p_t^k - \widehat{tc_t^{kj}} \geq \kappa \tag{9}$$

and regime probabilities for location j at time t are defined as follows:

$$Prob(s)_t^j = N(s)_t^j \ / \sum_s N(s)_t^j \tag{10}$$

where s is regime I, II, or III, and $N(s)_t^j$ is the number of cases of regime (s) in location j at time t.

To attribute observations to regimes, we need to specify a value of κ. What is a value of κ that can be justified? Only one of the components of the market indicator $(p_t^j - p_t^k - \widehat{tc_t^{kj}})$, notably unit transaction costs, is estimated; the other two, the prices in both locations, are observed. Therefore, the standard deviation of the market indicator is equivalent to the standard deviation of estimated unit transaction costs. Consequently, to find the values of the market indicator at the parity bounds and thereby attribute observations to all three regimes, it appears rational to use a cut-off value of κ equal or proportional to the standard deviation of estimated transaction costs.

Once we have a numerical value for κ, we can calculate regime probabilities for each location and each period by dividing the number of observations of each regime by the total number of trade pairs in each location, following equation (10). A regime probability of 100 percent signifies that the location (district) is fully characterized by this regime. The incidence of a regime is proportional to its probability. By definition, the three regime probabilities total 100 percent and are, hence, not independent.

Explaining Market Integration

The developed strategy for constructing regimes has little to offer in explaining regime fluctuations and in measuring the importance of different determinants. At this stage, we can only describe regularities in observed patterns and speculate about their causes. Nevertheless, we are eager to learn how the regime probabilities of deficit districts behave, especially during periods of food shortages. How can we design a strategy to identify causes and explain fluctuations in regime probabilities? Although a full-fledged identification strategy is beyond the current work, we can derive insights from estimating simple correlations. On the basis of inspection of the figures (see below), we conjecture that regime probabilities are correlated with food shortages. Therefore, we propose to estimate regime probabilities of deficit districts on a constant, a food shortage dummy, a trend, buyer fixed effects, and buyer-food shortage fixed effects. In formula form, this yields[9]:

$$Prob(s)_t^j = \gamma_0 + \gamma_1 \, fs_t \; + \gamma_2 \; trend_t \; + \nu_j + \gamma_j \, x \, fs_t + \; \varepsilon_t^j, \tag{11}$$

where fs_t is a binary dummy for a food shortage in period t ($fs = 1$ under a food shortage and $fs = 0$ elsewhere), ν_j are buyer fixed effects and $\gamma_j \, x \, fs_t$, are buyer food shortage fixed effects.[10] Food shortages are defined as periods in which the aggregate production of staple food is insufficient to meet the requirements of the entire population ($\sum_i Q_{0t}^i < \sum_i \sum_m E(fr_{mt}^i)$; see also equation (4)).

Estimations of transaction costs and market integration

In this section, we present and discuss the estimation results of transaction cost equations. Subsequently, we assess these estimation results on the basis of survey data of traders. We continue with presenting selected regime probabilities computed on the basis of predicted transactions costs. The section is completed with estimations that attempt to explain fluctuations in regime probabilities.

Estimation of Transaction Cost Equations

To apply the proposed approach, we use Malawi district-level monthly maize price data complemented with other variables over the period from June 1999 to October 2009 (see appendix for details). The district is the unit of observation: buyer (seller) or deficit (surplus) locations refer to deficit districts (surplus districts), where the difference between district production and district requirements determines whether a district is a seller district (production − requirements >0) or a buyer district (production − requirements <0). For the estimation of transaction cost equations, we use the specification proposed in equation (1) conditioned on a selection of observations specified in equation (6). For comparisons over time, unit transaction costs and fuel prices are

9. The inclusion of surplus districts into the estimation is a potential strategy to identify the impact of food shortages on regime probabilities of deficit districts. Examination of this strategy is left for future research.

10. Strictly speaking, the term $\gamma_j \, x \, fs_t$ includes both $\gamma_1 fs_t$ and ν_j, and the latter two terms need not be presented separately. However, we have adopted the notation convention to present both terms (see also, for example, equation (1)). In equation (11), separate notation is particularly relevant because we are interested in identifying the common correlation of regimes with food shortages (γ_1).

TABLE 1. Estimating Transaction Cost Equations: Sample Restricted by Distance

Dependent variable: natural logarithm of per kilogram transaction cost or of price difference ($\ln tc_t^{kj}$) or $\ln(p_t^j - p_t^k)$)

Assumptions		[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]
Excess supply	φ	2/12	2/12	4/12	4/12	2/12	2/12	4/12	4/12
minimum calorie pppd		2200	2300	2200	2300	2200	2300	2200	2300
maximum distance		200	200	200	200	300	300	300	300
ln(distance)	β_{dis}	0.451	0.489	0.384	0.472	0.480	0.449	0.410	0.443
		(4.3)	(5.2)	(3.3)	(4.1)	(8.2)	(8.3)	(6.1)	(5.9)
ln(fuel price)	β_{fp}	1.332	1.113	0.979	0.957	1.284	1.102	1.083	1.164
		(3.4)	(3.1)	(2.4)	(2.5)	(4.0)	(3.6)	(3.1)	(3.5)
Trend	β_t	.0057	−0.0062	−0.0042	−0.0054	−0.0050	−0.0052	−0.0040	−0.0050
		(1.8)	(2.2)	(1.3)	(1.8)	(1.9)	(2.1)	(1.5)	(1.9)
Dbuy × dmonth	ν_j, δ_m, τ_{jm},	yes	yes	yes	yes	yes	yes	yes	yes
Dsel × dmonth	μ_k, λ_{km}	yes	yes	yes	yes	yes	yes	yes	yes
Adjusted R^2		0.3335	0.3652	0.3510	0.3618	0.3886	0.4044	0.3986	0.3868
Number of observations		766	802	639	649	1090	1131	895	867

The table reports transaction cost estimations on the basis of district maize prices from June 1999 to October 2009. Estimation is by OLS (district fixed-effect estimations are available from the author on request). The transaction cost and fuel price are deflated with the Malawi consumer price index. The sample is restricted to observations below a maximum distance. Absolute t-statistics are given in parentheses (.) below the coefficient. We do not report the coefficients and t-values of the constant term and dummy variables. Columns reflect different assumptions with respect to the minimum per person per day calorie intake (either 2200 or 2300 kcal), excess supply over requirements before districts start to export (either two- or four-month district requirements; see also φ in equation (2a)), and maximum distance, the distance beyond which trade is assumed to be unlikely (either 200 or 300 km).

deflated with the Malawi consumer price index. Additionally, the unit transaction cost, distance, and fuel price are transformed into natural logarithms. Table 1 reports estimated transaction cost equations. Absolute t-statistics are given in parentheses below the coefficient. We have omitted the coefficients and statistics of constant terms and dummy variables. Columns in the table represent estimation results with different values for three parameters: minimum per person per day caloric intake (2200 or 2300 kcal), excess production over requirements above which district exports are triggered (two or four months of maize requirements, respectively; φ in equation (3a)), and the distance above which domestic trade is unlikely to occur (200 or 300 km). The transaction cost equations are estimated by OLS. Trade pair fixed-effect estimations (not reported) generate a similar output for transport fuel price and trend and thus confirm the OLS results. In the trade pair fixed-effect estimation, the distance variable drops out because of the within-transformation. Because we are keen to quantify the coefficient of distance, because this variable reflects a variety of influences on transaction costs, we maintain the presentation of OLS results.

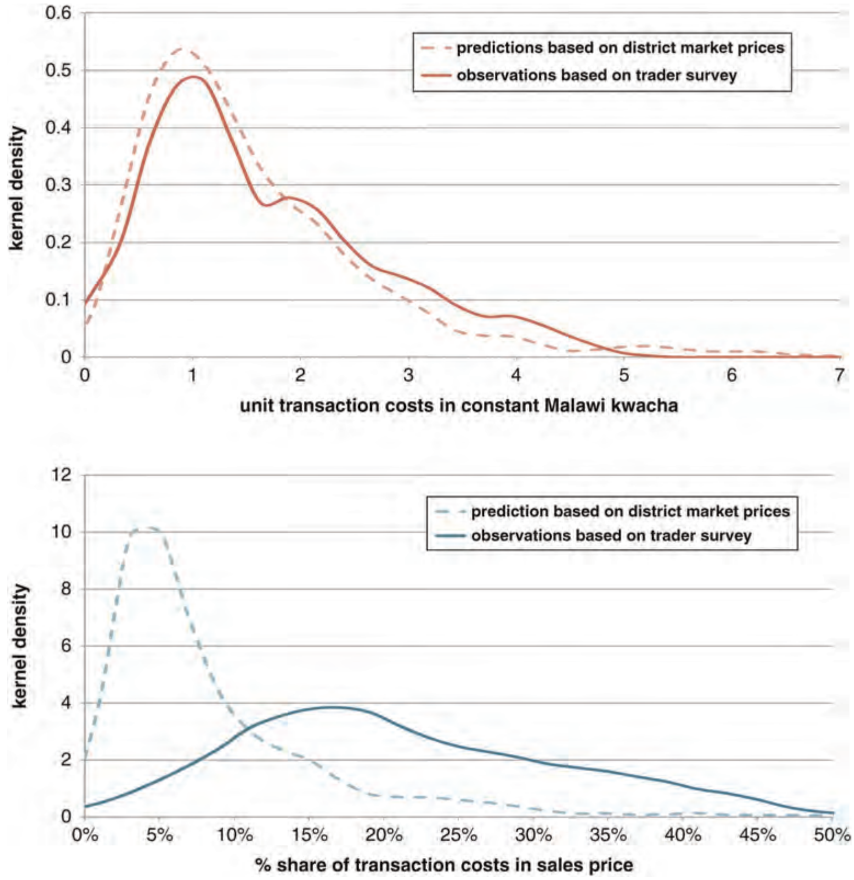
The estimation results are well behaved. The reported estimation output of all equations indicates that distance, fuel price, and the common trend are statistically significant (or close to significant, in the case of the trend variable) and have the expected sign. The elasticity of transaction costs with respect to fuel price approaches unity in most equations, with values between 0.96 and 1.33. This outcome is reasonable because fuel prices reflect the variable transport cost, which must be passed on in a competitive market and because transport costs constitute a large component of transaction costs. The elasticity of transaction costs with respect to distance varies from 0.38 to 0.49. A significant negative common trend is observed in most estimates, reflecting a variety of developments in structure and technology. A large part of the negative trend development is likely to be better captured with trade-pair specific trend variables, but for presentation, we prefer the common trend. The overall goodness of fit, between 33 and 40 percent with 639 to 1131 observations, is reasonable. A substantial portion of the variation is explained by dummy variables. The comparison of estimation of a spline by distance with an unrestricted estimation (see appendix) supports the restriction of the observations to those below a maximum distance.

Further Assessment of Estimation Results

A different way of assessing the estimation results of transaction cost equations is to compare predicted values of transaction costs with estimates of transaction costs obtained from surveys. For this purpose, we use a survey of traders in Malawi covering the period from August 1999 to February 2000 that is documented in Fafchamps and others (2005).[11] This survey is a representative

11. The data of this survey were kindly made available by Marcel Fafchamps.

FIGURE 1. Comparison of Predicted and Observed Transaction Costs*



Source: Author's calculations.
* The figure shows kernel density estimations of predicted and observed unit transaction costs in Malawi kwacha per kilogram of maize (constant 2000 prices), in levels (upper panel) and expressed as a percentage of sales prices of the supplying district (lower panel). Predictions are based on district market prices, and observations are obtained from a trader survey (Fafchamps and others, 2005).

sample of 738 Malawi traders. However, the survey contains only 275 observations of maize transactions by maize traders[12]. In figure 1, we show the density functions of unit transaction costs and transaction costs relative to sales prices, both from this survey and from predictions on the basis of the estimated transaction cost equations. In table 2, we summarize the average and standard deviation of transaction costs by rural region and urban center, in level and share, from

12. We have extracted data on traders who reported having traded maize during the last 12 months. Within this group, we used the data on maize transactions documented under "variable marketing costs of a completed purchase and sale transaction."

TABLE 2. Comparison of Predicted and Observed Transaction Costs*

Rural region, urban areas	Level			Share		
	Predicted	Observed	Mean test	Predicted	Observed	Mean test
Rural north	1.547 (49; 1.064)	2.298 (19; 1.322)	2.4	0.221 (49; 0.204)	0.213 (17; 0.111)	0.2
Rural central	1.829 (518; 1.170)	1.794 (52; 0.990)	0.2	0.098 (518; 0.092)	0.208 (49; 0.078)	8.1
Rural south	1.279 (439; 1.036)	1.462 (116; 1.008)	1.7	0.059 (439; 0.053)	0.205 (112; 0.116)	19.7
Urban areas	1.610 (84; 1.311)	1.694 (55; 0.993)	0.4	0.079 (84; 0.052)	0.198 (53; 0.091)	9.7

* The table reports the average predicted and observed transaction costs in Malawi kwacha per kilogram of maize (constant 2000 prices) in levels and shares (expressed as a percentage of sales prices, all summarized by supplying district). The predicted levels (the first column) are within sample predicted per unit transaction costs from estimated transaction cost equations, and the observed levels (the second column) are averages based on survey data on transaction costs from Fafchamps and others (2005). Likewise, shares are based on predicted transaction costs and on trader survey data. Numbers of observations and standard deviation are given in parentheses (.) below the mean. The mean test reports the (absolute) t-statistic of the difference in average values.

this survey and from predictions on the basis of the estimated transaction cost equations. The table also reports test results on the equality of unit transaction costs and shares.

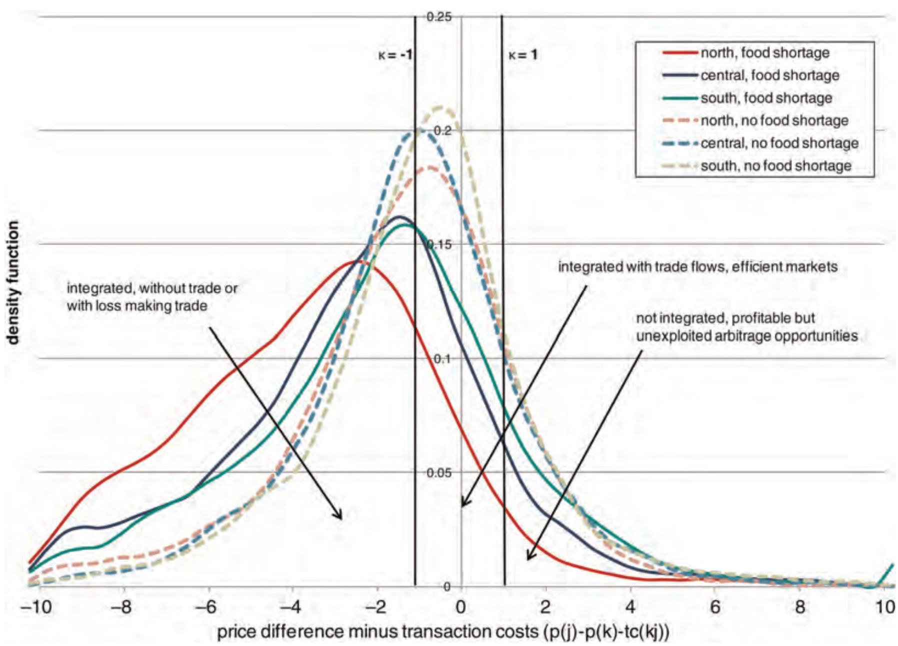
Predictions of per unit transaction costs by rural region or urban center vary from 1.3 to 1.8, with a standard deviation between 1 and 1.3. In this exercise, all prices are Malawi kwacha per kilogram in constant 2000 prices. The overall average of predicted unit transaction costs is 1.6 (median: 1.3). Survey observations have per unit transaction costs by rural region or urban center varying from 1.5 to 2.3. The overall average of unit transaction costs from the trader survey is 1.7 (median: 1.4). Predicted shares of transaction costs in terms of selling price by rural region or urban center vary from 6 to 22 percent, with an overall average of 9 percent (median 6 percent). Survey observations of transaction costs shares by rural region or urban center vary from 20 to 21 percent, with an overall average of 21 percent (median 20 percent). The equality test confirms equality for the unit transaction costs and rejects equality for shares (both with the exception of the rural north, possibly because of the number of observations).

We attribute the rejection of equality for shares to the difference in the type of data used for this exercise. Survey data on purchase and sale prices reported by traders are (and should be) well below market prices. The available evidence indicates that market prices, on average, are 9 to 38 percent above the sales prices reported by traders. In addition, the trader data refer to a single completed trade transaction. Such documented transactions are likely to bias transaction costs upward. The observed differences between purchase and sales prices in the survey data also contain a component of trader profit that is not well captured in the estimations. Conversely, monthly aggregate district market prices are averages computed over one month and a range of transactions among a range of traders and other market participants. The process of aggregation will dampen fluctuations between different locations; likewise, estimates of unit transaction costs and transaction costs relative to sales prices on the basis of aggregate price data will be lower. Despite the clear rejection of the equality of shares, we find that the test reported in table 2 indicates that levels of predicted unit transaction costs and survey observations of unit transaction costs are of the same order of magnitude. Therefore, we conclude that predicted unit transaction costs may be considered good approximations for the assessment of market integration.

Calculated Regime Probabilities

With predicted unit transaction costs, we can calculate the value of the market indicator, the price difference less predicted transaction costs, for each location (and thus for both directions of trade) and for each month. A summary presentation of the results, the density function of the market integration indicator

FIGURE 2. Market Regimes by Region: With and without Food Shortage*



Source: Author's calculations using κ = 1.

* The figure shows kernel density estimations of price differences minus predicted unit transaction costs ($p^j - p^k = tc^{kj}$) in buying districts, aggregated by region. Values below the upper κ value indicate market integration. The periods of food shortage are from April 2001 to March 2003 and from April 2005 to March 2004.

shown in figure 2, aggregates observations over months and buyer districts, partitioned into regions and periods with and without food shortages. The area under the density functions and below $-\kappa$, above $+\kappa$ and between $-\kappa$ and $+\kappa$ characterizes the incidence of the various regimes.

For all regions, both with and without food shortages, the density function is nicely shaped with a peak below zero. Most of the observations are below the $-\kappa$ value and hence are categorized within regime II (integrated, without trade or with loss making trade). The other two regimes (regime I, integrated with trade flows, efficient markets; and regime III, not integrated, profitable but unexploited arbitrage opportunities) have a lower incidence. The figure also indicates that during food shortages, the density functions flatten and shift to the left.

Of particular interest is how market integration has developed in typical deficit districts. Urban areas are notorious deficit districts (Blantyre and Lilongwe), but a number of rural districts also frequently face food shortages.[13] Regime probabilities for regimes I, II, and III over time and for the major urban districts, Lilongwe and Blantyre, are shown in figure 3. Periods of food shortages, indicated in the figures with vertical lines, are from April 2001 to March 2003 and from April 2005 to March 2006. The computation of regime probabilities in the figures is based on $\kappa = 1$.[14]

The large fluctuations in the regime probabilities over the years are immediately apparent. In view of these fluctuations it is difficult to maintain that fixed regime probabilities are sensible approximations for the Malawi maize market. We further observe that food shortages coincide systematically with a substantially lower incidence of regime I. Moreover, over the period under investigation—from April 1999 to October 2009—we cannot observe a clear upward trend of regime I. On the contrary, we observe a slight tendency to trend downward after 2007. Similar observations are made for rural deficit districts. The inspection of these figures (see appendix) indicates overall large fluctuations in regime probabilities, consistently lower regime I probabilities during food shortages and a tendency of regime I probabilities to trend downward, especially after 2007. The downward trend after 2007 may reflect the increased involvement of domestic policies (price controls and export restrictions) in the wake of worldwide increases in food prices.

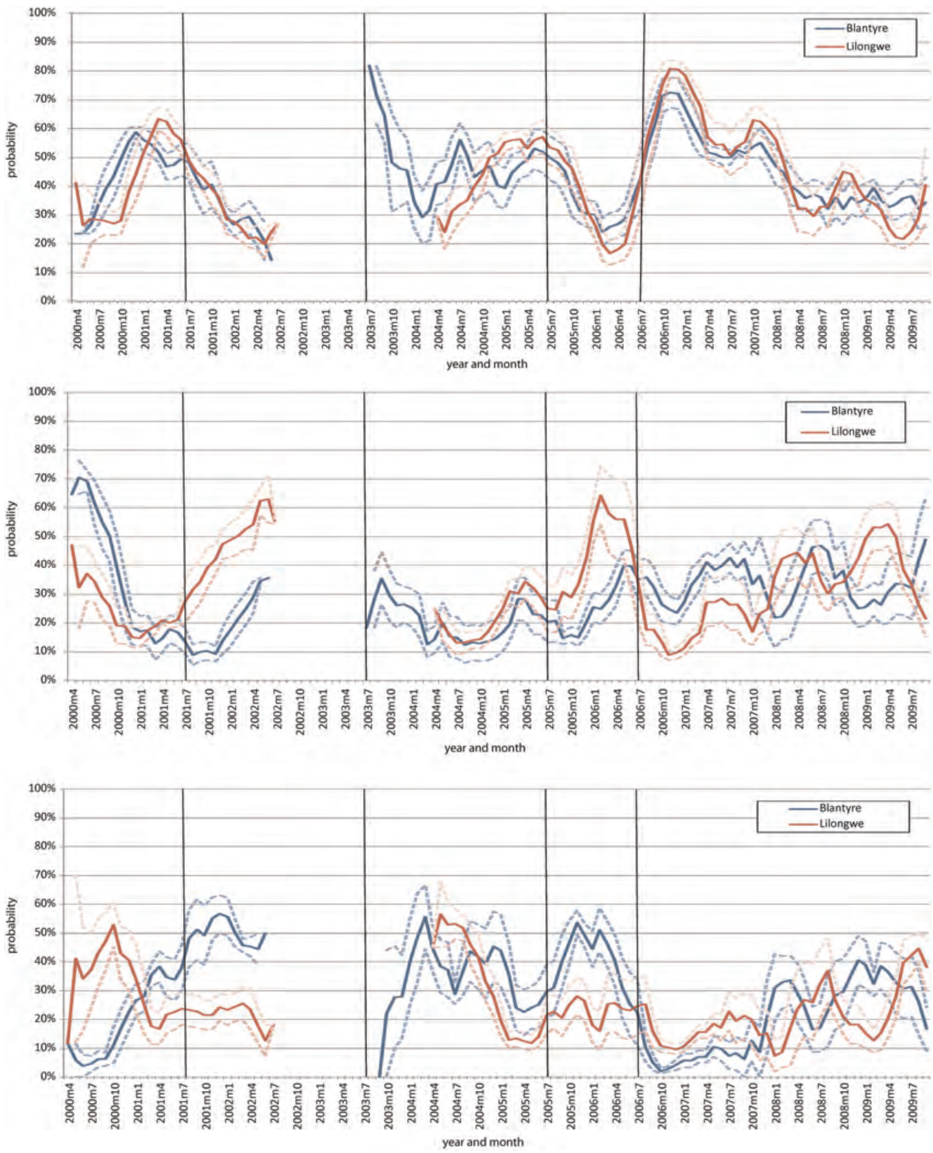
Food Shortages and Market Integration

We proceed with the explanation of variations in regime probabilities proposed in the previous section and estimate regime probability equations for urban areas and for rural deficit districts by region, following equation (11) (see table 3). To

13. Typical rural deficit districts are Nsanje, Chikwawa, Machinga, and Thyolo in the south; Dowa, Nkhotakota, and Ntchisi in the central region; and Karonga, Nkhata Bay. and Rumphi in the north.

14. Estimations based on the alternative values of κ are available in the supplemental appendix at http://wber.oxfordjournals.org/. Alternative values (e.g., $\kappa = 0.5$ and $\kappa = 1.5$) change the probability levels, but the pattern over time remains the same. Not surprisingly, regime I probabilities increase (decrease) somewhat if κ increases (decreases).

FIGURE 3. From Rural Surplus Districts to Urban Areas*
Upper Panel: Integrated Markets and Efficient Trade ($-\kappa < p^j$-p^k-$tc^{kj} < +\kappa$)
Middle Panel: Integrated Markets, no Trade or Trade at a Loss (p^j-p^k-$tc^{kj} <$ -κ)
Lower Panel: Not Integrated, Unexploited Arbitrage Opportunities (p^j-p^k-$tc^{kj} > \kappa$)



Source: Author's calculations using $\kappa = 1$.
* For the purpose of calculating confidence intervals for regime probabilities, we use the per-location variation of regime probabilities around a six-month moving average.

TABLE 3. Explaining Developments in Regime Probabilities

Deficit districts	Regime		[1] REGIME I	[2] REGIME II	[3] REGIME III
Rural north	constant	γ_0	0.427 (12.9)	0.440 (9.8)	0.133 (3.2)
	fs	γ_1	−0.148 (3.6)	0.200 (2.8)	−0.070 (0.1)
	trend	γ_2	−0.00128 (3.5)	−0.00042 (0.8)	0.00171 (3.6)
	dbuy × fs	ν_j, $\gamma_j.fs$	yes	yes	yes
	R^2		0.1445	0.2164	0.1435
	chi^2 (parameters)		50.2 (6)	82.0 (6)	49.8 (6)
	# of observations		297	297	297
Rural central	constant	γ_0	0.472 (14.9)	0.256 (6.8)	0.272 (5.6)
	Fs	γ_1	−0.196 (4.7)	0.038 (0.8)	0.126 (1.9)
	trend	γ_2	−0.00084 (2.7)	0.00096 (2.6)	−0.00013 (0.3)
	dbuy × fs	ν_j, $\gamma_j.fs$	yes	yes	yes
	R^2		0.1202	0.0638	0.0622
	chi^2 (parameters)		53.4 (8)	26.6 (8)	25.9 (8)
	# of observations		391	391	391
Rural south	constant	γ_0	0.534 (15.5)	0.306 (6.7)	0.159 (2.4)
	fs	γ_1	−0.190 (3.0)	−0.008 (0.1)	0.197 (2.4)
	trend	γ_2	−0.00168 (5.7)	−0.00023 (0.6)	0.00192 (5.0)
	dbuy × fs	ν_j, $\gamma_j.fs$	yes	yes	yes
	R^2		0.1968	0.0806	0.1137
	chi^2 (parameters)		114 (12)	40.9 (12)	59.9 (12)
	# of observations		467	467	467
Urban	constant	γ_0	0.537 (16.3)	0.118 (3.1)	0.375 (9.2)
	Fs	γ_1	−0.238 (4.6)	0.277 (4.6)	0.120 (2.4)
	trend	γ_2	−0.00073 (2.1)	0.00261 (6.2)	−0.00188 (4.4)
	dbuy × fs	ν_j, $\gamma_j.fs$	yes	yes	yes
	R^2		0.2036	0.2001	0.1312
	chi^2 (parameters)		74.9 (6)	73.3 (6)	44.2 (6)
	# of observations		293	293	293

The table reports correlations of regime probabilities with a constant, a food shortage dummy, and a trend variable. REG I represents integrated markets and efficient trade, REG II represents integrated markets without trade or with trade at a loss, and REG III represents nonintegrated markets with unexploited and profitable arbitrage opportunities. Regime probabilities are calculated assuming $\kappa = 1$ (see equations (7) to (9)). Estimation is by Seemingly Unrelated Regression. Values of regime probabilities are never at the lower limit (0) or upper limit (1); therefore, there is no need to control for censoring of the dependent variable. Absolute z-statistics are given in parentheses (.) below the coefficient.

account for the mutual dependency of regime probabilities (the sum of the probabilities is one), we estimate using Seemingly Unrelated Regression.[15]

The results in table 3 indicate that the average probability for regime I (integrated markets with positive trade flows) is between 43 and 54 percent, between 12 and 44 percent for regime II (integrated, without trade or with loss-making trade), and between 13 and 38 percent for regime III (not

15. Additionally, we should consider censoring of the dependent variable. However, doing so is not necessary because probabilities never have the lower- or upper-bound values.

integrated, profitable but unexploited arbitrage opportunities). The average probability of regime II appears to be relatively high in the rural north, which may be a reflection of relatively high transaction costs. Both urban and rural central areas have relatively high average probabilities for regime III. In all regions, the sum of the regime II and regime III probabilities—thus, the probability of not being in regime I (integrated markets with positive trade flows)—is between 46 and 57 percent.[16]

Food shortages are consistently associated with lower regime I probabilities, with an average decrease due to food shortages of 15 to 24 percent points. The complement of this decrease in the probability of regime I is less consistent: in the northern region, the decrease in regime I probability coincides with an increase of the regime II probability, whereas in the rural central and rural southern regions, this decrease coincides with an increase in the regime III probability. In urban centers, food shortages decrease the regime I probability and increase the probabilities of regimes II and III. In addition, the common trend is consistently and significantly negative in the case of regime I probabilities. This common negative trend is the largest in the northern and southern regions. Conversely, a positive common trend in regime II dominates in rural central and urban districts and is insignificant elsewhere. A positive common trend in regime III is observed in the rural north and the rural south, and a negative common trend in regime III is observed in urban districts.

It is not easy to evaluate these results as reflections of the increasing integration of markets and the successful liberalization of the domestic trading sector. On the contrary, the level of fully integrated markets with positive trade flows (regime I) is moderate at best (approximately 50 percent), whereas food shortages are shown to have a substantial and significant negative impact, and trends are consistently negative and significant. The food shortage impact suggests that markets fail when we most need them to function well. Finally, trend developments in regime probabilities away from market integration with positive trade flows are also not very comforting.

Our reading of the empirical evidence presented in this paper is as follows. In periods of relative abundance, when district production is sufficient to meet district demand, there is limited need to trade to feed people. In these circumstances, trade is seldom profitable, with prices similar and close to production costs in most districts and with high costs of trade. Without food shortages and without the threat of a humanitarian disaster, districts tend to move into autarky, and subsistence households remain subsistence households. The limited trade that takes place under these conditions can easily be serviced by the generally poorly developed trading infrastructure. Alternatively, in periods of food shortage, districts are

16. It must be added that the absolute levels of average regime probabilities are dependent on the value of κ in the calculations and therefore are less informative than their relative level and their correlation with food shortages and trends. A sense of the magnitude of the variation is obtained by estimating with alternative values of κ (see the supplemental appendix at http://wber.oxfordjournals.org/).

forced to trade with each other in an environment that lacks an adequate trading infrastructure for larger volumes, to make large outlays on transport and to embark on uncommonly practiced and expensive district-to-district trade to remote rural areas, which is even more expensive because of the covariance of food shortages and congestion in trade. These circumstances induce (even) high(er) transaction costs, which lead to loss-making trade or block trade despite attractive trade opportunities.

Summary and Conclusion

We have investigated the measurement of market integration in staple food markets in Malawi. Consistent with competitive spatial market equilibrium and in the tradition of the PBM, we assess market integration by calculating the spatial price difference minus transaction costs. Transaction costs are predicted on the basis of estimations of transaction cost equations. Estimated transaction cost equations take account of transport costs, fixed source costs, fixed destination costs, seasonality, and technological and structural change. Empirical transaction cost estimates are well behaved in the key explanatory variables, and the predicted unit transaction costs are of the same order of magnitude as those reported in survey data. With the help of predicted transaction costs, we compute the distribution of market indicator values for each location and assess market integration. The presented evidence on market integration indicates that regime probabilities are not fixed. The probabilities of integrated markets consistently and substantially decrease during food shortages, increasing either the probability of no trade or loss-making trade or the probability of profitable but unexploited trade opportunities. The data further support a significant negative trend in market integration with trade. If we control for food shortages and trends, the average probabilities of market integration with trade are moderate.

With regard to policy implications, the results suggest targeting both high transaction costs and spatial market inefficiencies. Transaction costs can be reduced by improving the physical infrastructure for trade, by lowering taxes and prices that directly affect trading costs, by removing regulations that create a burden on trade, and by establishing regulations supporting trade. Market inefficiencies can be addressed by improving the market orientation of players in the market.

APPENDIX

Data sources and data construction

We use monthly price data by district or Rural Development Project covering 26 districts/Rural Development Projects of Malawi[17] from June 1999 to

17. Malawi has a total of 28 districts. The tiny district Likoma, an island in Lake Malawi, is ignored and the district Neno is aggregated with to the district Mwanza.

October 2009. Data on monthly maize prices by district are publicly available through the Food Security Updates of the Famine Early Warning System Network (www.fews.net), which obtains these data from the Ministry of Agriculture and Food Security in Malawi. Malawi maize prices are reported for a gradually increasing number of individual markets (up to 68 different markets in 2009). We have selected the market for each district with the largest number of observations over the years, which were major district towns in nearly all cases. The price series by district are not complete. Of the potential $26 \times 116 = 3016$ monthly price observations (26 districts, 116 months), we have 2304 independent observations, approximately 77 percent of the total, with substantial variation in data availability across districts. We have refrained from filling white spots by interpolation and extrapolation because this may hinder the verification of the spatial arbitrage that governs domestic trade and may lead to poorly based inferences.[18] The available price data transform to a total of 46,386 observations of price differences. For the estimation of transaction costs, only a limited part of the data is relevant (see equation (6)). Recall, however, that the distribution of all price differences, both positive and negative, is used in the calculation of regime probabilities. Some sense of the spatial dimension of the data set may be useful: if we restrict the data to price differences between locations less than 150 km apart, a subsample of approximately 7000 observations remains.

Data on annual maize production by district or Rural Development Project are also made publicly available through the Food Security Updates of the Famine Early Warning System Network (www.fews.net), which obtains this information from the Ministry of Agriculture and Food Security in Malawi. The production of maize is estimated in several rounds in the course of the crop year, and the bulk of the data refer to final (mostly third-round and a few fourth-round) estimates.

The monthly maize requirements by district in kilograms are calculated by multiplying the monthly population by the daily calorie requirement, the number of days per month, and the maize calorie share by district and dividing by the calorie content per kilogram of maize. The annual population data by district are obtained from the National Statistical Office of Malawi (Malawi in Zomba; www.nso.malawi.net), which are converted to monthly data by interpolation. Assumptions regarding daily calorie requirements vary in empirical work. The Malawi Vulnerability Assessment Commission uses a minimum dietary requirement of 2100 kcal per day per head (Malawi Vulnerability Assessment Committee, May/June 2003, Malawi Baseline Livelihood Profiles, 2003). From 1999 to 2005, the observed daily per capita energy supply varies

18. Measurement error may be an additional issue. We do not know the exact quality of the price data, but it is likely—as in most price data—that there is a certain degree of measurement error in these series, which may affect the results. However, because we mainly use spatial price difference—price differences between different districts—a certain type of measurement error is eliminated.

from 2157 to 2217 kcal according to the food balances reported by the FAO (FAOSTAT). The Government of Malawi estimates that the actual daily caloric intake is 2366 kcal per person per day (Government of Malawi/World Bank 2006). We use different values of per capita per day caloric intake in the same range and report the applied values in the tables with estimations. The district maize calorie share in consumption is obtained from the Malawi Vulnerability Assessment Commission (MVAC 2003, see above) and ranges from 32 to 87 percent, with an average of 64 percent. The calorie content per kilogram of maize is equal to 3570. 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).

For comparisons over time, all prices and unit values are deflated with the consumer price index. Monthly consumer price index numbers are obtained from the National Statistical Office of Malawi, which differentiates between consumer price index numbers for urban and rural areas. Prices of fuel for transport are pan-territorial domestic prices of diesel at the pump, and data on these prices are taken from BP-Malawi. Distances in kilometers 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 used various data sources from the National Statistical Office of Malawi, particularly the Statistical Yearbook, Integrated Household Survey (2005), several Welfare Monitoring Surveys 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 taken from the World Development Indicators database (2009).

References

Abdulai, A. 2007. "Spatial Integration and Price Transmission in Agricultural Commodity Markets in Sub-Saharan Africa." *In Agricultural Commodity Markets and Trade, New Approaches to Analyzing Market Structure and Instability*, ed. A. Sarris and D. Hallam, Chapter 7, 163–186. FAO/Edward Elgar, Cheltenham, UK/Northampton, US.

Aker, J. C. 2010. "Information from Markets Near and Far: Mobile Phones and Agricultural Markets in Niger." *American Economic Journal: Applied Economics* 2: 46–59.

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): 292–307.

Baulch, B. 1997a. "Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration." *American Journal of Agricultural Economics* 79(2): 477–487.

——— 1997b. "Testing for Food Market Integration Revisited." *The Journal of Development Studies* 33(4): 512–34.

Christiansen, R. E., and L. A. Stackhouse. 1989. "The Privatization of Agricultural Trading in Malawi." *World Development* 17(5): 729–40.

Fackler, P. L., and B. K. Goodwin. 2001. "Spatial Price Analysis." *In Handbook of Agricultural Economics, Volume I*, ed. B. Gardner, and G. Rausser Chapter 17, 971–1024, Elsevier, Amsterdam, The Netherlands.

Fackler, P. L., and H. Tastan. 2008. "Estimating the Degree of Market Integration." *American Journal of Agricultural Economics* 90(1): 69–85.

Fafchamps, M. 1992. "Cash Crop Production, Food Price Volatility and Rural Market Integration in the Third World." *American Journal of Agricultural Economics* 74(1): 90–9.

Fafchamps, M., E. Gabre-Madhin, and B. Minten. 2005 "Increasing Returns and Market Efficiency in Agricultural Trade." *Journal of Development Economics* 78: 406–42.

Fafchamps, M., and R. Vargas Hill, 2005 "Selling at the Farmgate or Traveling to Market" *American Journal of Agricultural Economics* 87(3): 717–34.

Government of Malawi/World Bank 2006. "Poverty and Vulnerability Assessment: Investing In Our Future", World Bank Report No. 36546-MW.

Integrated Household Survey 2004–2005 (IHS) 2005. National Statistical Office, Zomba, MalawiJayne, T. S., J. Mangisoni, and N. Sitko. 2008. "Social Analysis of Malawi's Maize Market Reforms." Report prepared for the World Bank/Malawi, Michigan State University, East Lansing, Michigan.

Jensen, R. 2007. "The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector." *Quarterly Journal of Economics* 72(3): 879.

Kutengule, M., A. Nucifora, and H. Zaman. 2006. "Malawi Agricultural Development and Marketing Organization Reform." *In Poverty and Social Impact Analysis of Reforms, Lessons and Examples from Implementation*, ed. A. Coudouel, A.A. Dani, and S. Paternostro, Chapter 12, 415–452, The World Bank, Washington.

Lall, S. V., H. Wang, and T. Munthali. 2009. "Explaining High Transport Costs Within Malawi, Bad Roads or Lack of Trucking Competition?" The World Bank, Policy Research Working Paper 5133.

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): 102–33.

Malawi Vulnerability Assessment Committee (MVAC), 2003. Collaboration with the SADC FANR Vulnerability Assessment Committee, May/June, Malawi Baseline Livelihood Profiles, Lilongwe.

McNew, K., and P. L. Fackler. 1997. "Testing Market Equilibrium: Is Cointegration Informative." *Journal of Agriculture and Resource Economics* 22(2): 191–207.

Muto, M., and T. Yamano. 2009. "The Impact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda." *World Development* 37(12): 1887–96.

Moser, C., C. Barrett, and B. Minten. 2009. "Spatial Integration at Multiple Scales: Rice Markets in Madagascar." *Agricultural Economics* 40: 281–94.

Negassa, A., and R. J. Myers. 2007. "Estimating Policy Effects on Spatial Market Efficiency: An Extension to the Parity Bounds Model." *American Journal of Agricultural Economics* 89(2): 338–52.

Park, A., H. Jin, S. Rozelle, and J. Huang. 2002. "Market Emergence and Transition: Arbitrage, Transaction Costs, and Autarky in China's Grain Markets." *American Journal of Agricultural Economics* 84(1): 67–82.

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): 568–80.

Smith, L. D. 1995. "Malawi: reforming the state's role in agricultural marketing." *Food Policy* 20(6): 561–71.

Spiller, P., and R. Wood, 1988, "The Estimation of Transaction Costs in Arbitrage Models." *Journal of Econometrics* 35(2): 131–45.

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