TI 2015-055/V
Tinbergen Institute Discussion Paper

Trains, Trade and Transaction Costs: How does Domestic Trade by Rail affect Market Prices of Malawi Agricultural Commodities?

Wouter Zant

Faculty of Economics and Business Administration, VU University Amsterdam, and Tinbergen Institute, the Netherlands.

Trains, Trade and Transaction Costs:

How does Domestic Trade by Rail affect Market Prices

of Malawi Agricultural Commodities?

Wouter Zant*

Abstract

We measure the impact of low cost transport by rail in Malawi on the dispersion of agricultural commodities prices across markets, by exploiting the quasi experimental design of the nearly total collapse of domestic transport by rail in January 2003, due to the destruction of a railway bridge at Rivirivi, Balaka. Estimations are based on monthly market prices of four agricultural commodities (maize, groundnuts, rice and beans), in 27 local markets, for the period 1998-2006. Market-pairs connected by rail when the railway line was operational, are intervention observations. Railway transport services explain a 14% to 17% reduction in price dispersion across markets. Geographical reach of trade varies by crop, most likely related to storability and geographical spread of production. Perishability appears to increase impact reflecting the lack of intertemporal arbitrage. Overall, impacts are remarkably similar in size across commodities.

JEL code: D23, F14, N77, Q13, O18, O55

Key words: trade, crop prices, transaction costs, rail infrastructure, Malawi, sub Saharan Africa

* Wouter Zant is associated with VU University and fellow of the Tinbergen Institute, both Amsterdam, The Netherlands. Email: wouter.zant@vu.nl. I am grateful for comments from conference / seminar participants at Milan (ICAE2015), Oxford (CSAE2016) and VU University, and from three anonymous referees of this journal. I am also grateful to Hans Quené for assistance in constructing the data, and to Jasper Dekkers for constructing a maps.

1. Introduction

High transport costs make trade between markets unprofitable and force farmers into subsistence farming. Conversely, low transport costs increase trade, lead to lower prices and lower price dispersion, and offer farmers incentives to commercialize. In the longer run low transport costs also increase supply response, improve allocative efficiency, accelerate technology adoption and innovation, and enhance economic growth. Sub-Saharan countries, particularly landlocked ones, face high transport costs and suffer from high and volatile food prices and poorly functioning markets. Railway transport is a low cost alternative to road transport, and a feasible and useful complement to other modes of transport. Rail transport contributes to lower food prices through its impact on the operation of markets, increases welfare of households and improves food security. Rail transport may also lead to a reduction of overall transport prices through rail-road competition.

In this study we investigate the hypothesis that, due to its low cost, railway transport services increase domestic trade in agricultural commodities. With increased trade flows of agricultural produce from surplus to deficit areas, prices will increase in surplus markets and decrease in deficit markets. As a result dispersion of agricultural commodity prices across markets reduces, and, consequently, the availability of low cost rail transport services is associated with lower price dispersion across markets. The key empirical challenge in measuring impacts of infrastructure is to find an identification strategy that allows to separate the impact of railway services from other factors. For the purpose of this research we exploit a natural experiment, notably the disruption of a railway bridge in the heart of the Malawi railway network, caused by a tropical storm in January 2003. The bridge collapse also caused a nearly total collapse of domestic freight by train. In view of these developments we assume that markets located along the rail line were connected with each other through rail transport services until January 2003 and lost this connection from January 2003 onwards. Natural experiments in transport infrastructure are a rare event (cf. Jacoby and Minten, 2008) and experimental designs in infrastructure are usually not feasible[1]. Consequently, and apart from historical studies that exploit the rollout of the railway network as identification strategy (Donaldson, 2012; Burgess and Donaldson, 2012; Jedwab et al., 2014; Jedwab and Moradi, 2015) impact studies on railway services are not common.

For the empirical estimations we make use of monthly market prices for a few selected crops (maize, rice, groundnuts and beans) in selected markets. These selected crops are grown in all Malawi districts and the selected markets are evenly spread throughout Malawi. A number of these markets are located close to the rail line while others are remote from the rail line. The sample period stretches a number of years before and after the date of the bridge collapse. In the estimations we explain (absolute) price dispersion of a number of food crops across markets. Price dispersion is assumed to be determined by a series of market specific factors (seasonality, trend, and fixed effects) and market-pair specific factors (e.g. transport infrastructure). The estimation results support a 14% to 17% reduction in price dispersion across markets as a result of rail transport services. Geographical reach of trade varies by crop, most likely related to perishability, storability and geographical spread of production. Perishability appears to increase impact reflecting the lack of intertemporal arbitrage.

The empirical literature on impacts of infrastructure is large. A substantial part is macro in nature, often takes a historical perspective, uses rollout of infrastructure as identification and derives its theoretical underpinning from either new economic geography (including Davis and Weinstein) or trade theory (Ricardian comparative advantage, Heckscher-Ohlin, Eaton-Kortum).

[1] Casaburi et al., 2013 is an exception.

The impact of transport infrastructure research in the tradition of new economic geography and Davis & Weinstein centers around the question if the equilibrium distribution of economic activity across space is determined by locational fundamentals (geographical endowments) or economies of scale and scope due to concentration and clustering of activities and (past)investments (Jedwab and Moradi, 2015; Jedwab et al., 2014; Redding and Sturm, 2008; Bleakley and Lin, 2012; Ahlfeldt et al., 2014; Storeygard, 2016). Jedwab and Moradi (2015) and Jedwab et al. (2014) exploit railroad construction respectively in Ghana and Kenya over the last century to show path dependence of economic activity and local increasing returns. Both Redding and Sturm (2008) and Ahlfeldt et al. (2014) exploit the fall of the Berlin wall and the division and reunification of Germany as a natural experiment, to explain the decline in population of east-west border cities with changes in market access (Redding and Sturm, 2008) and to explain differences in productivity and amenities across city blocks in the city of Berlin (Ahlfeldt et al., 2014). Bleakley and Lin (2012) find support for persistence of activities around natural obstacles to water navigation; Storeygard (2014) employs night time lights satellite data to measure city level economic activity combined with roads data, to study the role of transport costs in determining income of sub-Saharan cities and find a large transport cost elasticity of city income for cities away from a port. Trade theory based work looks at the impact of infrastructure on trade volumes, goods and labour markets, trade costs and responsiveness to shocks (e.g. Michaels, 2008; Feyrer, 2009; Donaldson, 2010; Burgess and Donaldson, 2012; Atkin and Donaldson, 2015; Allen and Atkin, 2016). Donaldson (2010) exploits, in his extensive study on colonial Indian railways (1861-1930), the roll-out of the railroad network and finds, on the basis of a general equilibrium trade model that railroads reduce trade costs, reduce price dispersion between regions, increase trade volumes and welfare and decrease income volatility. Burgess and

Donaldson (2012), using the same identification strategy, find that railroads in India made prices and income less responsive to shocks. Michaels (2008) uses the construction of the US Interstate Highway System (1956-1975) to show an increase in trucking activity and retail sales and, in line with Heckscher-Ohlin, increases the wage bill of high skilled workers relative to low skilled workers. Feyrer (2009) exploits the closing of the Suez canal and the related variation of trading distances to find more accurate estimates of trade costs using a gravity framework.

A different strand of the literature is directed to the impact of infrastructure in contemporaneous developing countries and makes use of data with a shorter time dimension and a smaller space dimension, often micro survey data (e.g. Yamauchi et al., 2011; Casaburi et al., 2013; Asher and Novosad, 2016). Using (quality) improvements in road quality jointly with Indonesian household panel and village census data (1995-2007), Yamauchi et al. (2011) claim that the increase on household income growth and non-agricultural labour supply due to improved connectivity supports complementarity between education and quality of local roads. Casaburi et al. (2013) employs a genuine experimental design in rural road rehabilitation in Sierra Leone to estimate the impact on transport costs and market prices and test alternative models of price formation. They find price reductions in rice and cassava, larger for cassava and for locations remote from urban centres, and smaller in markets located in production areas. Part of the research on developing countries particularly considers the impact of transport and trading infrastructure on transaction costs, and, related to this, trade margins & costs, (dispersion of) commodity prices, household income, welfare and supply response (Minten and Kyle, 1999; Jacoby, 2000; Jacoby and Minten, 2008; Goyal, 2010; Zant, 2014). A key motivation for this research is the impact of high transaction costs on low input, low productivity and low growth agriculture, and thereby on welfare of rural households. Studying food prices in Zaire, both across regions and between products, Minten and Kyle (1999) show that transportation costs explain differences in food prices between producer regions and urban Kinshasa, and that road quality is the key determinant of transportation costs. Inspired by the difficulty of experimental designs in infrastructure Jacoby (2000) and Jacoby and Minten (2008) take the reverse route and use the economic impact of transport costs on various economic variables and economic behaviour (wages, value of agricultural land, household income, migration), in the absence of road infrastructure, in order to measure the potential gains of putting road infrastructure in place. Empirical application to Nepal data support a substantial benefit to the poor which is, however, not large enough to reduce income equality (Jacoby, 2000). Madagascar household data suggest large gains in income from improved infrastructure for remote households but gains are small relative to the improved non-farm earning opportunities in town (Jacoby and Minten, 2008). Focusing on marketing infrastructure in the soy market in Madhya Pradesh, Goyal (2010) finds increased soybean prices, decreased price dispersion and increased area under soy cultivation, due to the introduction of a direct marketing channel by a major private company. Zant (2014) also finds a positive supply response, both in the intensive and extensive margin, of tobacco growers in Malawi due to improved market access, realised with the introduction of a new auction floor.

Although different from transport costs, the work on the impact of search costs on (dispersion of) market prices due to the rollout of mobile phone networks is also related to this study (e.g. Aker, 2010; Jensen, 2007; Muto and Yamano, 2009; Fafchamps and Aker, 2014). After the introduction of mobile phones, price dispersion on fish markets in Kerala, India has reduced, while fishermen’s profits as well as consumer welfare increased (Jensen, 2007). Easy and timely access to information also prevents waste, inefficiency and spoilage of perishable crops (see also Muto and Yamano, 2009). Aker (2010) and Aker and Fafchamps (2012) estimate a 10% to

16% reduction in price dispersion across grain markets in Niger (2001, 2006, retail and wholesale) due to mobile phones. Better storability of grain and less perishability is claimed to account for lower reduction in price dispersion compared to Jensen (2007). Muto and Yamano (2009) find that mobile phone coverage in Uganda (2003,2005) has induced market participation of banana farmers in remote areas, but find no impacts on maize marketing. Similarly Fafchamps and Minten (2012) find no effects of SMS based services in Maharashtra, India on prices received by farmers, value added, crop losses, crop choices and cultivation practices. Asymmetric information between traders and farmers, and comparative advantage in transport causes benefits to accrue to traders and not to producers.

In this study we provide empirical evidence of the impact of rail transport services on dispersion of market prices of agricultural commodities. The estimation strategy is similar to the one used in Aker (2010) and the topic is closely related to literature that explains the impact of trade costs on prices (e.g. Minten and Kyle, 1999). The remainder of the paper is organized as follows. In Section 2 we briefly characterize the Malawi economy, present details on Malawi rail infrastructure and rail freight and document the collapse of domestic rail freight since January 2003. In Section 3 we explain the methodology underlying the empirical estimation and the identification strategy. In Section 4 we present and discuss estimations and in Section 5 we give a summary and conclusion.

2. The Malawi economy, transport by rail and domestic trade in agricultural commodities

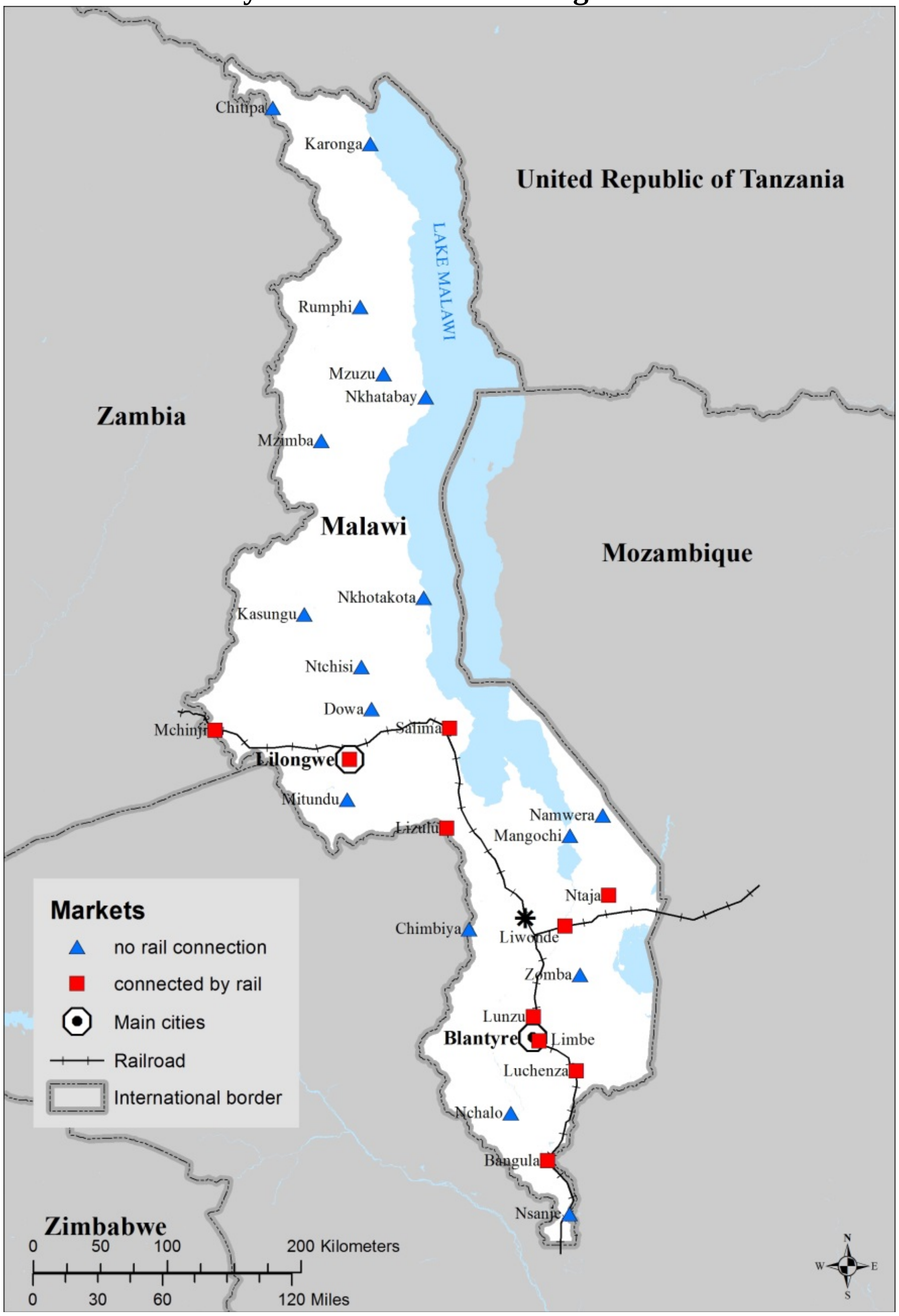
The Malawi economy

Malawi is a relatively small landlocked country in the south of Africa, measuring around 800km from north to south and around 150 km from east to west, in area size around half the UK, bordering in the northwest with Zambia, in the northeast with Tanzania and in the south with Mozambique. A large lake, Lake Malawi, part of the Great Rift Valley, stretches from north to south, along the east border of the country (see Figure 1). During the study period (1998-2006) the population increased from close to 10 million to 13 million and is mostly rural: only a small fraction (11% to 15%) lives in the cities Lilongwe, Blantyre, Mzuzu and Zomba. More than 80% of the Malawi population depends for food and income on subsistence farming. The incidence of poverty is high: more than 50% of the population in Malawi is poor (Integrated Household Survey 2005 (IHS-2), National Statistical Office), poverty is extremely high in remote rural districts (e.g. Chitipa: 67.2%; Nsanje: 76.0% and Chikwawa: 65.8%), and in the southern region at least 10%-points higher relative to other regions. The key food crop is maize, followed with a distance by cassava, rice, groundnuts and beans. Malawi suffers from occasional food shortages due to drought and poor harvests (see Zant, 2012, 2013). Tobacco is by far the most important cash crop. Just like the other major cash crops, sugar and tea, tobacco cultivation dates back to the colonial period. Tobacco, however, has become nearly completely smallholder based in the course of the 1990s (see Zant, 2014), while tea and sugar production is (still) mainly on account of estates. Nearly every city, town or larger village has one or more markets for agricultural food crops on a regular basis, often daily or weekly. Both local farmers and traders operate on these markets. The major export crops are marketed differently: through auctions in the case of tobacco and tea, and through a large processing company (Llovo) in the case of sugar. Tobacco and sugar exports are also transported by rail.

Railway infrastructure, operations and transport costs

The Malawi rail network consists of a single rail line with a total within Malawi length of 797 km, running from Zambia in the west (where it runs 25 km into Zambian territory to Chipata), to the east via Lilongwe and Salima, and next to the south where – in the district of Balaka – the line splits into a line further south to Blantyre and Beira in Mozambique, and a line to the east, to Nacala in Mozambique (see Figure 1).

Figure 1 Malawi railway network and selected agricultural markets



Source: VU SPINlab; Note to Figure: the asterisk on the map indicates the railway bridge at Rivirivi, Balaka

The entire network is single track throughout, unsignalled and train operations are managed by radio and by electro-mechanical token system. Historically, the line is operated by a government owned railway company, Malawi Railways. However, on December 1, 1999 a 20-year concession for the operation of the network and supply railway transport services has been given to Central East African Railways (CEAR). This concession, an integrated part of a larger concession (with the US based Railroad Development Corporation, as main concessionary), known as the Nacala corridor, further consists of the port of Nacala in Mozambique, a railway line that runs from Nacala to the Malawi border, and a 26km railway line from Mchinji, Malawi to Chipata in Zambia. The Mozambique components are owned by the parastatal Mozambique railway company CFM. The joint venture that operates this concession is known as Corredor de Desenvolvimento do Norte (CDN)[2]. Investment in railway lines in the region is driven, in the first place by the exploration of quarrying and mining companies[3].

The rail network has not been fully operational in the past, amongst other things because of civil war in Mozambique, destruction by floods and poor maintenance. Lall et al. (2009) characterize the role of railroads in Malawi historically and against the background of the civil war in Mozambique (1977-1992) as follows: "*…rail has historically been the main mode for international freight transport, connecting Malawi with its southern neighbors of Mozambique, Zimbabwe and South Africa. However, the civil war in Mozambique from the mid-seventies cut off the two main rail arteries – the Nacala and Beira-Sena lines. With the Nacala line being mined and the destruction of the main bridge across the Zambezi River on the Beira-Sena lines, the importance of rail in Malawi's international freight movements has declined*".

---

[2] See B.J.Knapp and H.Posner III, 'A luta continua!', Railway Gazette International, June 2004, 160.6, 363. Since privatization in 1999 and up to the time of writing several other private sector parties have participated in the CDN concession.

[3] A search on SSA railways in the archives of news sources (All Africa (www.allafrica.com), The International Railway Journal, Railway Gazette International) generates nearly exclusively articles in some way related mining.

However, the (potential) importance of the railway for domestic trade in the Malawi economy is acknowledged in policy documents: *"Rail transportation is also an important mode of transport for rural farmers who usually use the train to move their farm produce to main markets in the cities or trading centers. Such commodities include tomatoes, pigeon peas and other vegetables. In 2006, CEAR (Central East African Railways) recorded approximately 480,000 passengers moved largely from smallholder farmers. Since 2000, CEAR moved over 250,000 tons of local products to main markets locally but has been experiencing reduced usage by the locals to transport their commodities using rail transportation."* (taken from Millennium Challenge Corporation, 2011).

Since unit transport costs by train is a fraction of the cost of road transport, transport of goods by rail remains an attractive alternative. World Bank (2006) compares local road transport costs in 2003 with average per ton-kilometer tariffs for a number of SSA corridors / rail operators. Unit road transport costs are calculated to be a factor 1.4 to 3.1 higher[4]. Donaldson (2010) claims that road transport in India is 4.5 times more costly relative to rail transport. Various studies further stress the role of rail transport in keeping road tariffs in check. This is particularly relevant to Malawi which has notoriously high transport costs, due to both poor feeder roads as well as limited competition (see Lall et al. 2009). Moreover, in many SSA countries road infrastructure is supplied at less than full recovery cost, creating a road-rail competition imbalance (World Bank, 2006; Teravaninthorn and Raballand, 2008).

Transport by rail and the January 2003 collapse of domestic freight by rail

What can we learn from data on transport by rail? Figure 2 shows statistics of transport by rail (freight (tonnes, ton/km) and passengers) in the period 1997-2007. Despite the huge fluctuations in freight, the figure reveals a clear structural break in trade volume, starting in January 2003,

[4] Average tariffs are (resp. corridor, rail operator, road tariff and rail tariff in US$c): Senegal-Mali, Transrail, 7.9, 5.3; Cote d'Ivoire-Burkina/Mali, Sitarail, 7.9, 5.5; Cameroon-Chad, Camrail, 11.2, 6.3; Mozambique, CCFB/CFM, 10.0, 5.5 and Tanzania-Great Lakes, TRC, 13.5, 4.3.

when a nearly total collapse of domestic trade occurred. The collapse is further corroborated by the development of domestic freight (Figure 2b) and, somewhat less pronounced, by the development of passenger trade (Figure 2c).

Export and import freight by rail develops differently. Malawi exports tobacco, sugar and beans, and imports fuel and fertilizer, all items of key importance to the Malawi economy. Also, in times of food shortages, food aid is most efficiently and most cost effectively transported by rail (for example in 2002 and 2003, see the archive of All Africa (allafrica.com)). CEAR annual freight data (tonnes or ton/kilometer moved) show large increases in exports in 1999 and 2000, possibly associated with the privatization of railway services and enhanced by the 1998 devaluation of the Malawi kwacha. The data suggest that CEAR prioritizes the more profitable imports and exports rather than local freight and passenger services[5]. Export and import freight were also less affected by the 2003 bridge collapse, although tobacco exports have come to complete standstill. The average share of local freight (vis-à-vis international trade) has dropped from around 50% before 2003, to around 10% from 2003 onwards.

The collapse in rail transport expressed in the figures was for a large part on account of the disruption of a bridge at RiviRivi in Balaka district, at a central location in the rail network (see Figure 1). The disruption of the bridge was caused by Delfina, a tropical storm[6]. Delfina started at the northwest coast of Madagascar on December 30, 2002, intensified while moving westward before hitting northeast Mozambique on December 31, and weakened while moving inland by January 1, 2003, into Mozambique and Malawi. By January 9 Delfina had died out.

[5] The bias towards international freight is apparent from the operational strategy of CEAR reported in various newspapers (see e.g. All Africa.com). Alternatively, domestic passenger services were suspended for periods and for segments in the network, partly due to disagreements with the Malawi government about the extent of subsidy for these services (see The Chronicle, February 7, 2006).

[6] The drop coincides with the decline in local freight due to the closure of a quarry at Changalume (86km north of Blantyre) in 2002, which supplies clinker to a Blantyre cement plant. Unfortunately, we are unable to construct local freight data excluding freight of clinker. The closure is unlikely to affect market prices of agricultural commodities.

Figure 2a All trade by rail, monthlies

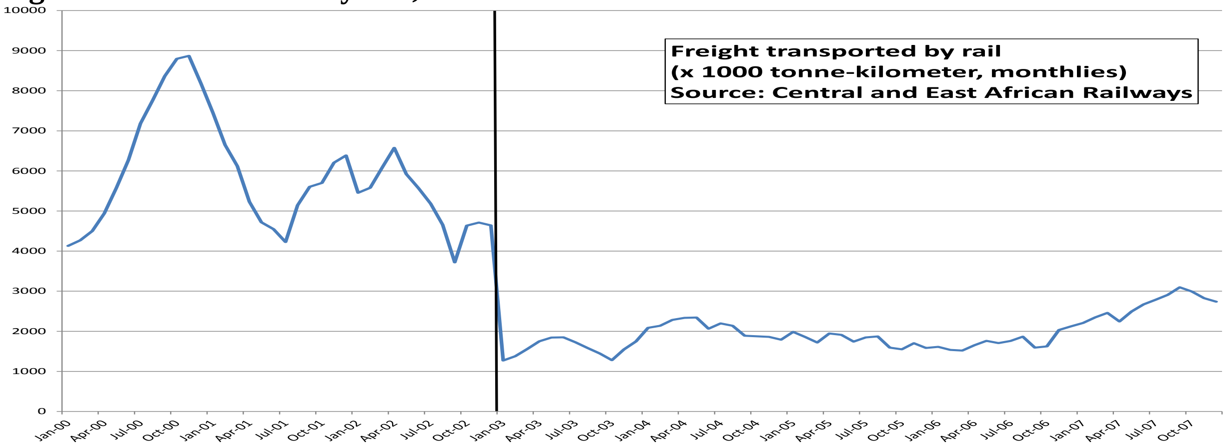
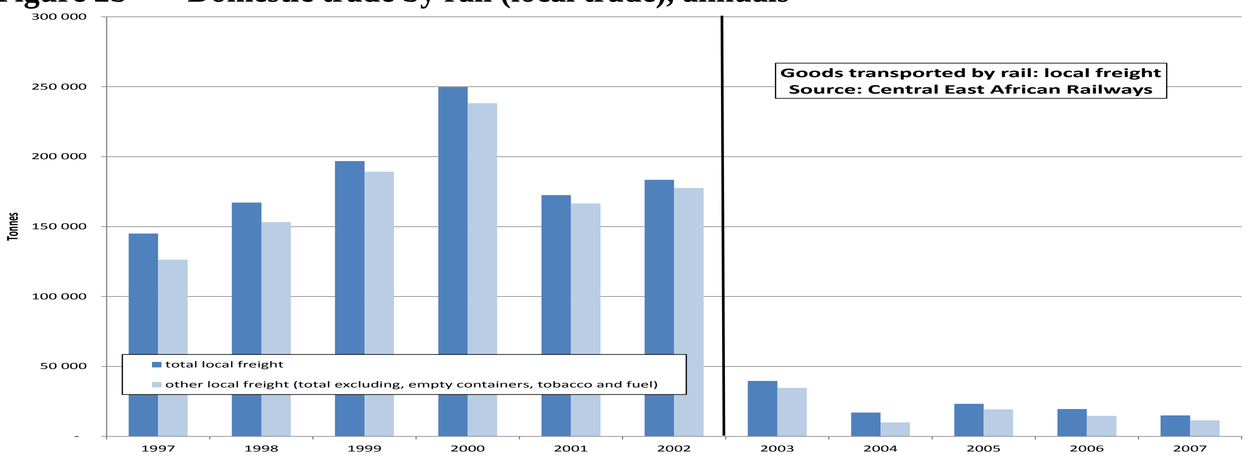
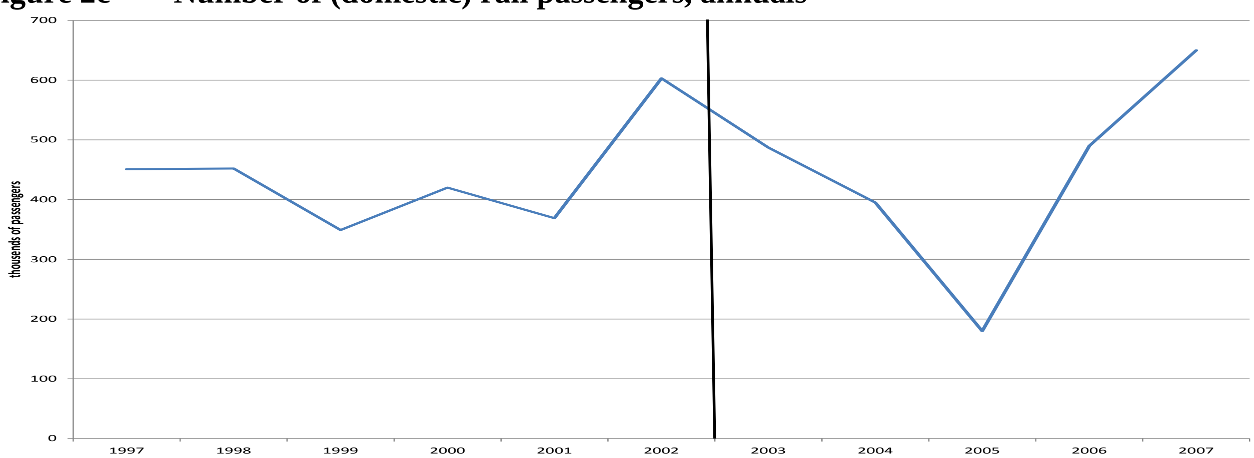


Figure 2b Domestic trade by rail (local trade), annuals*



* For the year 2004 we only have data on aggregate tonnage: composition is computed by interpolation

Figure 2c Number of (domestic) rail passengers, annuals



The major damage of the storm was done in Mozambique. In Malawi, the storm's remnants caused flooding in several districts, although not widespread. Delfina damaged roads, and, most importantly for our study, destroyed a railway bridge at RiviRivi, in Balaka district. Also a train derailed due to effects from the storm, which cut rail travel between northern Mozambique and neighboring Malawi. The storm further destroyed about 3,600 houses, and around 30,000 people were forced to leave their homes. The floods affected 57,000 properties, damaging 23,500 ha of agricultural land. Delfina killed eight people in Malawi, prompting President Muluzi to declare the country as a disaster area on January 11 (source: Wikipedia)[7]. Whether CEAR or the Malawi government should pay to rebuild the bridge, caused substantial delay in making the railway line operational again. In May 2005, close to two and a half years later, and with support from USAlD and the UK DfID, the Rivirivi railway bridge was reconstructed and rail transport operations were resumed[8].

The dramatic sequence of events has created an interesting opportunity to measure the impact of rail transport services on markets. Ideally one would need records of bilateral trade flows by rail, including prices by market and traded quantities of agricultural products by source and destination, both before and after the disaster had taken place, in order to investigate impact rigorously. Unfortunately, such data on trade by rail are not available. As a matter of fact, we also do not know to what extent trade in agricultural commodities by rail takes place in the form of passenger trade – farmers and traders travelling by train to nearby town and city markets on

[7] Despite the casualties and damage caused by the storm, we were unable – using annual aggregate data sourced from the Ministry of Agriculture and Food Security – to find an adverse impact on agricultural production of the affected districts.

[8] Personal communication of the author with CEAR staff. In January 2005 The Railway Gazette International (161.1,12-14) reported: '*work to repair damage to the Rivirivi Bridge caused by Cyclone Delfina in January 2003 is nearing completion, and CEAR hopes to restore train services shortly*'.

order to sell their produce – or in the form of formal freight[9]. What is available is data on market prices of agricultural commodities for a large number of markets in Malawi. Hence, in this study we aim to measure the impact of railway services on market prices of agricultural products.

Malawi markets for agricultural commodities: production, prices, domestic trade and demand

In the empirical part we look at prices of a few specific crops, notably maize, rice, groundnuts and beans. These food crops are all important crops in the Malawi context, where maize takes an outstanding position accounting for 50 to 60% of the diet of most people in Malawi and is produced by nearly all farm households. Maize, rice and beans are primarily consumed domestically, although in case of bumper crops small quantities are exported. Groundnuts are partly exported, but, again, the bulk of production is consumed domestically. There is a remarkable even spread of production of these crops over Malawi: with the exception of rice, they are cultivated in all Malawi districts (see Appendix 7). Groundnut cultivation is most concentrated in the central region and least in the southern region, and cultivation of pulses is most concentrated in the southern region and least in the north and central region.

Whatever the cause of the even distribution of production – high transport and trading costs, inefficient subsistence agriculture, little comparative advantage of regions – it may also, and simultaneously, limit the scope for domestic trade. Supply from local production will often be a cheaper and preferred alternative, rather than "imports" from neighboring regions and districts. In the end the (relative) balance between supply from local production and local demand over the season determines the scope for trade.

[9] Fafchamps et al, 2005 reports that in their 2000 Malawi trader survey none of the traders made use of transport by train. It should be noted, however, that the impact that we are after, also concerns farmers (rather than traders) who take up trading activities, or, more precisely, farmers who take the trouble of selling their output in a nearby town or city to fetch a higher price rather than selling at the farm gate or the local market. Evidence on Ugandan coffee farmers suggests that such activities are not uncommon (see Fafchamps and Vargas-Hill, 2005). A nearby railway station may make this proposition even more profitable for a farmer.

Seasonality in production and constant demand easily translates into seasonality in prices. Prices tend to be low in the months after harvesting (April-June), and subsequently increase continuously to reach (very) high levels just before the next harvest is available. Differences between highs and lows often are larger than 100%. Seasonality in prices is more pronounced in urban areas, due to higher income, larger population and lower local supply. Seasonality in prices of maize, rice, groundnuts and beans is confirmed by the data (see Appendix 2) and its importance to food security highlighted in other research (see e.g. Kaminski et al., 2014). Large and predictable seasonality in agricultural prices is likely to be partly explained by high search and transport costs. Price seasonality in Malawi is on average largest in maize, smallest in rice, and groundnuts and beans are in between[10]. Maize is also an exception in terms of value: rice, beans, and groundnut are high value crops with prices of on average 4 to 5 times the price of maize (but with distinct variations by crop, by year and by market, see Appendix 6). Trade in high value crops is likely to be more attractive for farmers due to relatively lower trade costs.

The predominantly small scale domestic trading business is undertaken by farmers, small, medium and large traders, wholesalers, maize processing firms and ADMARC. 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 et al. 2005). Most "district to district" trade of maize is from farmers to small and medium traders, and occasionally to larger traders and wholesalers. Around 75% of all traders buy directly from farmers and sell as a retailer (Fafchamps et al. 2005). Trading channels vary by location, but the bulk of maize trade is in the hands of the private sector. Survey data indicate that average distance between purchase location

[10] We calculate seasonality in prices by dividing monthly prices with average price. The average price is a centered 12 months average with a moving window. The resulting number that measures the extent to which monthly prices diverge from the season average, is dimensionless and automatically takes account of price changes over time.

and sale location of maize transactions is around 55km with a maximum of 200km (Fafchamps et al. 2005).

Although transport of agricultural produce by rail is a cheap and thereby attractive alternative, the dominant mode of transport of trade in agricultural products in Malawi is transport by road[11]. For a variety of reasons, however, domestic transport costs of transport by truck are very high: the main causes include poor (secondary) roads, high petrol prices, relatively inefficient small loads / no scale economies, no backloads, and undercapitalized and inefficient back-to-back funding (see Lall et al., 2009; Zant, 2013). The trunk road network (see Appendix 1 for a map of the road network), connecting cities and district towns, functions reasonably well, but the lack of good secondary roads leaves many locations underserved. Cheap transport services in Malawi potentially creates large welfare gains and enhances the scope for economic growth. If fully operational, the (extended) railway system in Malawi is an attractive and cheap alternative and an interesting complement to the currently dominant mode of transport.

3. Measuring the impact of railway services on dispersion of market prices

Data

For the empirical work we use monthly retail market prices of agricultural commodities taken from the Agro-Economic Survey, of the Ministry of Agriculture and Food Security. We have these data for a long period (from 1991/92 to 2008/09), for a large number of markets (around 70) and for a large number of agricultural commodities and livestock products (around 20). However, for the purpose of this study we use a limited subset: we use price data of only four crops (maize, rice, groundnuts and beans prices), for 27 markets and for the period from January

[11] The dominant mode of transport in domestic trade varies by country. In India for example the dominant mode of transport is rail (Donaldson 2010) and in Zaire road and river (Minten en Kyle, 1999). Most often, various modes complement each other.

1997 to December 2007. We have selected crops that are widely produced and consumed, markets that have the most complete price data[12] and a sample period that covers a distinct number of years before and after the bridge collapse. Even the selected price data are not complete, and, more troublesome, especially in the period of key interest to our research (around January 2003), a substantial drop in the completeness of the data occurs, most likely due to the food crises at the time (see Appendix 5)[13]. Therefore we carefully verified the number of available observations in the estimations, especially in the period from January 2003 onwards.

All distances used in the empirical part are distances as the crow flies, calculated using standard *Great Circle Distances*, and based on latitude-longitude coordinates of locations (markets, railway stations). We are aware that distance measured as the crow flies differs from road distance and that road distance is the relevant concept for transport costs. However, since we do not exactly know the (changes in) road distance at the time, we rather avoid the likely but uncertain error, and prefer the clear and transparent approximation of distance.

In the estimations we include the following three covariates ($X_{jk,t}$): rainfall, population density and per capita (gross) income. Rainfall is an annual index of crop season rainfall in mm. normalised with the long run average crop season rainfall in mm. Rainfall is recorded in around 30 weather stations and attributed to markets on the basis of proximity. We expect that above average rainfall increases crop production and the availability of agricultural commodities after harvest, and increased supply will reduce prices. Hence, we expect current crop season rainfall to have a negative impact on next year price. In terms of price dispersion we expect that a large difference in rainfall between locations increases price dispersion. Rainfall data are from

[12] In order to maintain the informational content of prices we have refrained from imputing values for missings.
[13] Malawi and Malawi agriculture suffered from floods and inundation in 2001 and from droughts in 2002. Under these circumstances one may expect that markets do not operate.

Department of Climate Change and Meteorological Services, Ministry of Natural Resources, Energy and Environment in Blantyre.

Population density is the number of people per square kilometer, by Extension Planning Area (EPA) or district / Rural Development Project (RDP). Population size varies between districts, but moves only gradually over time. Higher population densities are associated with more trade, and more efficient trade (see Fafchamps et al. 2005). Hence, we expect price dispersion between locations to decrease the larger the population density in both locations. Population data are from the National Statistical Office in Zomba and district area is taken from www.geohive.com.

Per capita income is a constructed annual variable by district: in order to calculate gross income from agriculture we multiply agricultural production with average retail market prices, both by crop season and district and summed over crops and livestock products[14]. All prices are deflated with the consumer price index for rural areas (source: National Statistical Office, Zomba, Malawi). Than we exploit data on rural and urban population by district: we first calculate the per capita agricultural income by using rural population data by district. Next, we assume that the highest per capita agricultural income in the region is related to per capita income of the urban population: we impute n times the region highest per capita income to the urban population, to construct (average) per capita income by year and by district, where n reflects the productivity differences between urban and rural workers[15]. There is a multitude of explicit and implicit assumptions in this per capita income calculation with many arbitrary

[14] We distinguish the crops maize (local, composite and hybrid), rice, millet, sorghum, cassava, sweet potatoes, groundnuts, pulses, cotton, tobacco, tea and sugar, and the livestock products steak, pork, mutton and goat meat.

[15] Urban population only refers to a fraction of the population of the districts of Lilongwe, Blantyre, Zomba and Mzimba. We use n = 1.5; However, a range of values varying between 1 and 3 did not fundamentally change the estimation results. We cannot calibrate the value of n with GDP data because of the subsistence character of the Malawi economy: home consumed production is included in our per capita income concept but it does not show up in per capita GDP.

elements, which many researchers will label as “heroic”. Nevertheless the constructed data should give a sensible order of magnitude for per capita income. Per capita income reflects demand and we expect that a higher per capita income increases demand and pushes up prices of agricultural commodities. Large differences in income between locations will, ceteris paribus, increase price dispersion. All covariates are expressed in terms of the natural logarithm of the absolute value of the difference of between markets ($\ln|x_k - x_j|$).

Theoretical considerations

Costs of railway transport are relatively low compared to transport costs by road which is the standard mode of transport in Malawi. Farmers and traders in Malawi based in areas near to a railway station potentially benefit from these cheap transport services. The lower transport costs enhances trade of agricultural commodities markets along the railway line, increasing flows of goods from surplus to deficit areas and, thereby raising low prices in surplus areas and reducing high prices in deficit areas. Hence, availability of railway transport services should reduce price dispersion across markets along the railway line. The key mechanism that drives this process is standard profit maximising producer behaviour with transaction costs.

Empirical specification

For the estimation of impacts we apply a panel fixed effect strategy and use the following regression model:

$$Y_{jk,t} = \beta_0 + \sum_i \beta_{1i} Y_{jk,t-i} + \beta_2 connected\,by\,rail_{jk,t} + \sum_n \beta_{4n} X_{n,jk,t} + \sum_l \beta_{5l}(time\,.\,\omega_l) + \sum_{sl} \beta_{6sl}(season_s\,.\,\omega_l) + \omega_l + \varphi_t + \eta_{jk} + \zeta_{jk,t}$$

where $Y_{jk,t}$ is the dispersion of prices across markets *j* and *k*, at time *t*, ‘*connected by rail*’ is a variable with a value (ln(*distance*)) if both markets are less than 20 km away from a railway station and 0 otherwise, $X_{n,jk,t}$ is a vector of n market-pair variables at time t, affecting the dispersion of prices across markets j and k, *time* is a time trend, *season* is seasonal dummy, ω_l, φ_t and η_{jk} are market, time and market-pair fixed effects and $\zeta_{jk,t}$ is a cluster robust error term.

There are several ways to measure dispersion of prices across markets ($Y_{jk,t}$), like, for example, the coefficient of variation or the maximum minus the minimum. We follow Aker (2010) and use (the natural logarithm of) the absolute price difference across markets ($\ln|p_j-p_k|$). In the estimated specification we have included the lagged dependent variable as explanatory variable in order to filter out lagged responses. The vector of $X_{jk,t}$ variables are assumed to be determinants of price dispersion between markets j and k, associated with either transaction costs (like transport costs, gasoline prices and economies of scale), or (relative) local supply and demand balances. In the data description the covariates used in estimations are discussed in detail. In the estimated equations we interacted a time trend and seasonality with markets[16]. The data description in the previous section supports a crop and market specific seasonal price pattern. Monthly fixed effects are included to control for country wide variations in agricultural production between years (caused by bumper crops and droughts).

Identification strategy

The exogenous collapse of railway transport services due to the disruption of the railway bridge at Rivirivi, documented in the previous section, creates a quasi-experimental design that offers an opportunity to identify the impact of railway services on the dispersion of agricultural commodity prices. The regression equation represents a panel fixed effect model with markets connected by rail when the railway was operational, as intervention observations. The coefficient

[16] We discarded the option to interact a time trend and seasonality with market-pairs since this nearly exhausts the degrees of freedom in estimation. Moreover we do have empirical support for trends and seasonality in prices (which explains part of the trend and seasonality is price dispersion), but trends and seasonality in prices dispersion are less evident from the data.

of interest in the regression equation is β_2: this coefficient reflects the impact of the availability of railway transport services on the dispersion of agricultural commodity prices. We expect the dispersion of these prices to be lower in the locations that have access to railway transport services. Hence, we expect β_2 to be negative.

Intervention locations

Interventions in this study are the market-pairs that are connected with each other by rail, when the railway is operational. Moreover, we have assumed that after the bridge collapse all segments of the network were affected, and, hence, all market pairs formerly connected were no more connected after the bridge collapse. This assumption is supported by communication with CEAR staff and freight data on domestic trade (see para on rail operations). We have defined intervention markets as markets that are less than 20km away from the nearest railway station[17]. Of course 20km is an arbitrary cut-off: we have verified the robustness of the estimation results by taking different cut-off distances (see Robustness of estimation results).

4. Estimations and discussion

Other empirical issues

In running the estimations we have assumed that the impact of railway services is geographically restricted. Transport costs increase more or less proportionally with transport distance (and are, in this respect, different from search costs). As a result the impact of availability of transport services is spatially restricted: transport costs are high and even become prohibitive for markets that are a large distance away. As transports costs translate into higher prices of traded goods, there is a clear trade-off between "import" and local supply: local production or the use of close

[17] Practically this implies we have 10 markets connected by rail (and thereby 45 intervention market-pairs). These markets are: Bangula, Lilongwe, Limbe, Liwonde, Lizulu, Luchenza, Lunzu, Mchinji, Ntaja, and Salima (see also Figure 1; see Appendix 8 for a list with markets, distance to station and latitude-longitude coordinates).

local substitutes could be cheaper alternatives than "import" from far away locations, especially if there is no advantage from specialization and if the agricultural product may be produced anywhere. Therefore we assume that (potential) domestic trade only takes place between markets that are located a limited distance away from each other.

How far this distance is needs to be investigated empirically. Aggregate data on freight by rail indicate that the average distance of freight by rail in case of local freight is 80-115 km and in case of export or import freight, 180-220km (calculations based on aggregate CEAR data). For domestic passengers the average distance travelled varies from 40 to 80km. Survey data on domestic trade and domestic traders, generally using (pick-up) trucks as mode of transport, indicate that average distance between location of purchase and sale location of maize transactions is around 55km with a maximum of 200km (Fafchamps et al., 2005).

The distance over which crops are traded is likely to be influenced by perishability and storability of crops, whether the crop is a high value or low value crop, the geographical spread of production and consumption, and how large expected gains from trade are. Since high value crops have relatively smaller transport costs, these crops are likely to be traded over longer distances. Next, trade in perishable crops is, by nature, spatially restricted: these crops simply degenerate if transported over long distances and thereby become unsaleable (we ignore the possibility of cooled transport which is typically not feasible for a SSA farmer-trader.) Conversely, storable crops are more suitable to be traded over longer distances. Trade over longer distances may arise if production area is more dispersed and/or more remote from consumption locations. Finally, the larger the gains the larger the distance over which crops are traded. These expected gains from trade depend on the difference in prices in each location, which in turn depends on differences in the supply and demand balance across locations[18]. A shortfall of local production relative to local demand will potentially give rise to high prices and is typical for urban areas. Seasonality in production will also lead to seasonality in prices, which will be more pronounced if demand is higher. Hence, differences in seasonality of prices across locations will affect expected gains from trade.

Moreover, we have also assumed that the period without railway services between the market-pairs connected by rail is restricted. With the disruption of the railway bridge at Rivirivi, Balaka a period started without railway services for markets along the railway line. The start of this period is 100% accurate. However, it is not clear when this period ended. From personal communication with CEAR staff we know that the railway bridge at Rivirivi was repaired and rail transport operations were resumed in May 2005. However, in the course of time the lack of railways services will lead to adjustments like the use of alternative modes of transport, increased local cultivation of food and shifts in consumption to local substitutes. The speed of adjustment will depend on the availability of cheap alternative modes of transport, supply response of local production and the resilience of demand to shift to other food. These adjustments are likely to have taken place since domestic trade did not recover after rail transport operations were resumed (see Figure 1, Figure 2a and 2b). The CEAR operational strategy favoring international freight may also play a role.

In the estimations we have assumed on the basis of data on passenger and freight transport by rail (see Section 3), that the period in which the railway is effectively not operational, is at least 2 years and at most 3 years. The minimum of 2 years is further motivated by the limited availability of price data around 2003 (see Appendix 2). The maximum of 3 years

[18] Availability of information on prices in several markets is obviously a key determinant of trade flows. Several studies highlight the importance of search costs and the availability of price information (see literature review).

is motivated by the fading out of the impact (see estimation section) and is determined empirically, using a grid procedure.

Selecting maximum trading distance and sample period

The estimation results of a basic specification are summarized in Table 1. The basic specification is uniform across crops with respect to maximum trading distance (100km), sample period (48 months before and 36 months after collapse) and cut-off distance to rail (20km). These assumptions will be relaxed in following estimations. All estimations include season-market dummies, a time trend for each market, month dummies and market-pair dummies.

Table 1 Impact (ATE) of rail transport services on price dispersion: uniform specification

dependent variable: $\ln(\text{abs}(p_{jt}-p_{kt}))$	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.107**	-0.074	-0.019	-0.031
	(0.0494)	(0.0596)	(0.0770)	(0.0305)
lagged dependent variable (t-1)	0.187***	0.203***	0.236***	0.236***
	(0.0297)	(0.0344)	(0.0344)	(0.0459)
lagged dependent variable (t-2)		0.095***	0.0664*	0.116***
		(0.0243)	(0.0399)	(0.0366)
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
Covariates	no	no	no	no
R^2	0.3479	0.4948	0.5232	0.5071
max trading distance (km)	100	100	100	100
sample period	1/99-12/05	1/99-12/05	1/99-12/05	1/99-12/05
months before 1/03 and after 12/02	48; 36	48; 36	48; 36	48; 36
number of observations	2146	1806	1435	1853

Note to table: The dependent variable: $\ln(\text{abs}(p_{jt}-p_{kt}))$ is the natural logarithm of the price difference between locations j and k, in month t. *Connected by rail* has the value ln(distance) if both markets are less than 20km away from a railway station, while the railway was operational, and zero elsewhere (in case of groundnuts less than 10km). Prices are deflated with the rural consumer price index (source: National Statistical Office, Zomba, Malawi). Robust standard errors in parentheses below the coefficient are clustered by market-pairs. Estimation results with two-way clustered standard errors (by markets of each market-pair) are shown in Appendix 3. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Higher order lags of the dependent variable are included if statistically significant at acceptable levels of accuracy. This strategy made us include a number of lagged dependent variables which varied by crop. In all estimations the coefficients of lagged dependent variables are positive, decreasing in size with the order of the lag, and statistically significant.

Estimation results reported in Table 1 are fine for maize but not convincing for the other crops. We need to relax to assumption of a uniform maximum trading distance and sample period / period before and after bridge collapse, to allow for heterogeneity between crops. We determine appropriate values of the maximum trading distance and the relevant period before and after the date of the collapse empirically, using a simple grid procedure: we estimate with a maximum trading distance varying from 70km to 300km (with a 10km step), and with 24 to 48 months before, and 24 to 36 months after January 2003. We use this procedure in order to find, simultaneously, the appropriate maximum trading distance and relevant period, but also to assess the robustness of the estimations. Selected output of this exercise is reported in Appendix 3.

On the basis of the tables and the underlying estimations reported in the appendix, we have the following observations. For all four commodities we observe sets of estimations with statistically significant ATEs with the required negative sign, around a specific combination of maximum trading distance and sample period. The regularity of these estimation outcomes – both across commodities and for each individual commodity, across combinations – offers comfort in and credibility of the estimations. For all commodities we see impacts disintegrate if the period without railway services is extended to the year 2006 (not shown, available from the author on request): apparently the period effectively without railway services – i.e. the period with effectively higher dispersion of agricultural market prices – is limited to the years 2003, 2004 and 2005, which is consistent with the fact that railway services resumed operations in May 2005. In fact, this motivated us to include estimations that exactly matched this period (January 2003 to April 2005), which further improved estimations in the case of beans, groundnuts and, especially, maize. Also, with a maximum trading distance of 70km or less, and with only two full years of observations before and after the bridge collapse (January 2003) the estimations tend to generate spurious outcomes. Next, the impact in the case of rice becomes significant with a maximum trading distance of 140-180 km. and for groundnuts with a maximum trading distance of 200-300 km., opposed to around 100 km. for maize and beans. The larger trading distance is possibly the result of the uneven spread of rice cultivation, compared to the other commodities. Also good storability (of rice and groundnuts) will make trade over longer distances easier. The impact coefficients of all commodities with the exception of beans are around -11%, while in the case of beans, the impact tends to be slightly stronger (see Appendix 2). We assume that this should be associated with the higher perishability of beans, and the related reduced scope for intertemporal arbitrage (relative to storable commodities). Other studies also confirm a higher impact in case of perishable crops (see e.g. Jensen 2007; Muto and Yamano, 2009; Aker and Fafchamps, 2014).

Table 2 shows a selection of the estimation results that allow for different maximum trading distances and sample periods across commodities and summarizes the key results of this study. The estimations confirm a statistically significant reduction in the dispersion of agricultural commodity prices across markets of around 9-12%. The size of the reduction is remarkably similar across commodities. The inclusion of the lagged dependent variable allows the distinction of short and long run impact, where the long run impact is calculated as $\beta_2/(1-\sum_i \beta_{1i})$, using the notation from the regression model. Long run impacts range from a reduction of 14% to a reduction of 17%. These results point at substantial welfare effects for consumers, from the enhanced efficiency of markets for agricultural commodities. The reduction in price dispersion is also likely to affect growth since lower food prices in subsistence economies constitute an important transmission mechanism to higher productivity (see De Janvry and Sadoulet, 2010).

Table 2 Impact (ATE) of rail transport services on price dispersion: selected output

dependent variable: ln(abs (p_{jt}-p_{kt}))	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.119***	-0.109***	-0.094**	-0.116***
	(0.0395)	(0.0416)	(0.0374)	(0.0346)
lagged dependent variable (t-1)	0.183***	0.182***	0.245***	0.193***
	(0.0329)	(0.0236)	(0.0210)	(0.0552)
lagged dependent variable (t-2)		0.096***	0.056***	0.095**
		(0.0214)	(0.0223)	(0.0441)
lagged dependent variable (t-3)			0.053**	
			(0.0262)	
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
Covariates	no	no	no	no
R^2	0.3898	0.4234	0.4202	0.4863
max trading distance (km)	110	160	250	110
sample period	1/99-4/05	1/99-4/05	1/99-4/05	1/00-4/05
months before 1/03 and after 12/02	48; 28	48; 28	48; 28	36; 28
number of observations	2062	2748	2930	1534
no. of intervention pairs	251	366	365	157
no. of control pairs connected by rail	161	190	210	107
no. of other controls (not connected)	1650	2192	2355	1270
long term impact	-0.146	-0.150	-0.146	-0.164

Note to table: See Table 1; The long term effect is calculated as $\beta_2/(1-\sum_i \beta_{1i})$ (see regression model above).

Robustness of estimation results

A number of robustness checks are implemented. We have repeated the estimations of Table 2 with inclusion of covariates, notably (relative) per capita gross income, (the sum of market-pairs) population density and (relative) rainfall. Estimated impacts come close to the ones reported in Table 2. Apparently the covariates are either independent of the intervention variable or well captured by the set of fixed effects applied in the basic estimations (or both).

Table 3 Impact (ATE) of rail transport services on price dispersion: including covariates

dependent variable: $\ln(\text{abs}\,(p_{jt}-p_{kt}))$	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.111^{**}	-0.124^{***}	-0.088^{**}	-0.097^{***}
	(0.0422)	(0.0408)	(0.0372)	(0.0342)
lagged dependent variable (t-1)	0.159^{***}	0.189^{***}	0.244^{***}	0.171^{***}
	(0.0304)	(0.0238)	(0.0219)	(0.0637)
lagged dependent variable (t-2)		0.090^{***}	0.062^{***}	0.090^{*}
		(0.0229)	(0.0231)	(0.0342)
lagged dependent variable (t-3)			0.053^{*}	
			(0.0273)	
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
Covariates	yes	yes	yes	yes
R^2	0.4098	0.4301	0.4246	0.5026
max trading distance (km)	110	160	250	110
sample period	1/99-4/05	1/99-4/05	1/99-4/05	1/00-4/05
months before 1/03 and after 12/02	48; 28	48; 28	48; 28	36; 28
number of observations	1834	2554	2801	1364
no. of intervention pairs	212	317	335	127
no. of control pairs connected by rail	139	172	196	84
no. of other controls (not connected)	1483	2065	2270	1153
long term impact	-0.132	-0.172	-0.137	-0.132

Note to table: See Table 1 and 2

The performance of covariates themselves is mixed. We expect positive signs for all three covariates: larger differences in either supply or demand conditions will induce higher price dispersions. The same applies to higher (sum of) population densities across markets. This result shows up in a few instances but not for all covariates and also often insignificant or weakly significant. We assume that this outcome is due to endogeneity of the covariates, measurements error in the construction of covariates and interaction of covariates with fixed effects. We will not further pursue this issue in this study.

Next, we have run the estimations with two subsets of controls as a robustness check: first we have left out those market-pairs of which at least one market is connected with the rail line (either operational or not) and used fully non connected market-pairs as controls; secondly we have used as controls only those market-pairs of which at least one is connected with the rail line. There are various justifications: one may argue that using remote market-pairs is merely measuring the difference between remote market-pairs and market-pairs along the rail line. Market-pairs near to the rail line but not connected are more likely to be similar to market-pairs that are connected to the rail line. Conversely one may argue that market-pairs of which one market is connected with the rail is, to some extent, benefitting from rail connection. Since this will blur the result, estimation with non-connected market pairs as controls should be preferred. The results (reported in the Appendix 4) are different but to a large extent confirm previous results: coefficients of the impact variable and lagged dependent variables have the right sign and statistical significance of these coefficients is acceptable to good in most estimations.

As a final robustness check, we have re-run the estimations with a smaller / larger number of intervention locations (see previous section). As a starting point we have used 20km as cut-off to define locations to be intervention locations. Alternatively we used as cut-off less than 10km, leading to 8 intervention locations and 28 intervention market-pairs, and less than 30km, leading to 13 intervention locations and 78 intervention market-pairs. The outcome of this exercise shows that the estimated impact remains more or less the same with a smaller number of intervention points, but deteriorates substantially with a larger number of intervention points.

Alternative explanations and potential threats

As the railway track in Malawi was not randomly constructed, outcomes may be the result of differences in markets already existing and not related to the railway line. For this reason – as is standard in similar type of exercises – we need to show that variables develop along a common trend and have similar means and distributions outside the intervention period. Table 4 shows tests on the common trend assumption, on the differences in unconditional means and distributions for relevant variables, outside the intervention period (January 2003 to April 2005).

The common trend tests suggest that we cannot reject the hypothesis of a common trend in all cases with the exception of groundnuts. In defense of the groundnut result we put forward that all estimations (not only groundnuts) control for market specific trends and hence also control for differences in trends.

Table 4a Common trends in price dispersion outside the intervention period*

	differences in trends	
	F-statistic (n,m)	p-value
market-pair data		
maize price: $\ln\lvert p_j-p_k\rvert$	F(2, 67) = 0.52	0.5985
rice price: $\ln\lvert p_j-p_k\rvert$	F(2, 103) = 0.16	0.8564
groundnuts price: $\ln\lvert p_j-p_k\rvert$	F(2,180) = 5.11	0.0069***
beans price: $\ln\lvert p_j-p_k\rvert$	F(2, 67) = 0.82	0.4435
maize price: $\lvert p_j-p_k\rvert$	F(2, 67) = 2.32	0.1063
rice price: $\lvert p_j-p_k\rvert$	F(2, 103) = 0.25	0.7764
groundnuts price: $\lvert p_j-p_k\rvert$	F(2,180) = 9.02	0.0002***
beans price: $\lvert p_j-p_k\rvert$	F(2, 67) = 1.09	0.3410

Table 4b Mean and distribution of (un)treated observables outside the intervention period*

	unconditional mean (SE)		difference in means	diff. in distributions	
	connected by rail	not connected by rail	F test (p-value)	D-statistic	p-value
market-pair data					
maize price: $\lvert p_j-p_k\rvert$	4.096 (0.252)	4.338 (0.228)	F(1,75): 0.51 (0.477)	0.0466	0.801
rice price: $\lvert p_j-p_k\rvert$	16.474 (1.642)	17.774 (0.964)	F(1,212): 0.47 (0.496)	0.0515	0.410
groundnuts price: $\lvert p_j-p_k\rvert$	41.697 (5.960)	45.797 (1.982)	F(1,103):0.43 (0.515)	0.1433	0.027**
beans price: $\lvert p_j-p_k\rvert$	34.665 (3.322)	36.706 (1.888)	F(1,75): 0.29 (0.595)	0.0475	0.177
distance: $dist_{jk}$	4.568 (0.101)	4.550 (0.049)	F(1,121): 0.03 (0.872)	0.0935	0.000***
rainfall: $\lvert rf_j-rf_k\rvert$	0.178 (0.017)	0.190 (0.010)	F(1,121): 0.38 (0.539)	0.1073	0.000***
population density:	10.706 (0.156)	10.176 (0.082)	F(1,21): 8.97 (0.003)***	0.3162	0.000***
pc gross income:	0.570 (0.086)	0.501 (0.037)	F(1,121): 0.55 (0.461)	0.1892	0.000***
market data					
maize price: p_j	3.180 (0.019)	3.152 (0.032)	F(1,26):0.55 (0.466)	0.1069	0.107
rice price: p_j	4.507 (0.036)	4.557 (0.028)	F(1,26): 1.19 (0.285)	0.1039	0.130
groundnuts price: p_j	4.865 (0.072)	4.844 (0.050)	F(1,26): 0.06 (0.807)	0.0829	0.460
beans price: p_j	4.795 (0.054)	4.751 (0.057)	F(1,26): 0.30 (0.586)	0.1445	0.011**
rainfall: rf_j	-.0353 (0.036)	0.004 (0.022)	F(1,26): 0.85 (0.365)	0.2303	0.000***
population density: pd_j	5.407 (0.206)	4.828 (0.117)	F(1,26):5.95 (0.022)	0.4826	0.000***
pc gross income: gi_j	6.913 (0.152)	6.817 (0.095)	F(1,26): 0.29 (0.596)	0.1913	0.000***

* Outside the intervention period is from January 2003 to April 2005. Market-pairs "connected by rail" are market-pairs where *both* markets are located within a distance of 20km of a railway station. Market-pairs "not connected by rail" are market-pairs of which at least one market is located more than 20 km located away from a railway station. The number of markets is 27 and hence the (potential) number of market-pairs 351, but practically less due to imposing a maximum trading distance. Robust standard errors are clustered by market-pairs or markets, according to the type of data tested. Prices are deflated with the rural consumer price index (source: National Statistical Office, Zomba, Malawi).

Tests on means and distribution of intervention and non-intervention market-pairs, outside the intervention period reveal a mixed picture: some means tests are indeterminate. Many variables, however, both at the level of markets and market-pairs, have different means and distributions. But how surprising is this? These markets are different locations with different population, different endowments and climate. In fact, we need these differences between locations to generate trade. In summary, the test outcomes reported in Table 4 do not invalidate the impact estimations, especially as long as we can adequately condition the variation in price dispersion on relevant covariates.

5. Summary and conclusion

In this study we have measured the impact of railway services on the dispersion of market prices of agricultural commodities in Malawi. For this purpose we have exploited the quasi experimental design of the nearly total collapse of domestic transport by rail in January 2003, due to the destruction of a railway bridge at Rivirivi, Balaka. Estimations are based on monthly market prices of four agricultural commodities (maize, groundnuts, rice and beans), in 27 local markets, for the period 1998-2006. The measured impact varies from a reduction in price dispersion of 9.5% to 12% in the short run, to 14% and 17% in the long run. Perishable and low value crops (respectively beans and maize) tend to be traded over smaller distances, and storable high value crops over larger distances (rice and groundnuts). There is some support for a relatively larger impact on perishable commodities (beans) reflecting the limited scope for intertemporal arbitrage. Results depend critically on the maximum distance between market-pairs, the period included before and after the collapse and which markets are assumed to be connected by rail. Estimations are robust for including covariates and various subsets of control groups.

References

Ahlfeldt, G.M., S.J.Redding, D.M.Sturm and N.Wolf, 2014, ‘The Economics of Density: Evidence from the Berlin Wall’, NBER working paper 20354.

Aker, J.C., 2010, ‘Information for Markets Near and Far: Mobile Phones and Agricultural Markets in Niger’, *American Economic Journal: Applied Economics,* 2 (July), 46-59.

Allen, T. and D. Atkin, 2015, ‘Volatility, Insurance, and the Gains from Trade’, working paper, NBER, NorthWestern, MIT.

Asher, S. and P. Novosad, 2016, ‘Market Access and Structural Transformation: Evidence from Rural Roads in India’, working paper.

Atkin, D. and D. Donaldson, 2012, ‘Who’s Getting Globalized? The Size and Implications of Intranational Trade Costs’ ,Working Paper, Stanford / MIT / NBER.

Bleakley, H. and J. Lin, 2012, ‘Portage and Path Dependence’, *Quarterly Journal of Economics*, 127, 587-644.

Burgess, R. and D. Donaldson, 2012, ‘Railroads and the Demise of Famine in Colonial India’, working paper, LSE / NBER.

Casaburi, L., Glennerster R. and T. Suri, 2013, ‘Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone’, mimeo.

Donaldson, D., 2010, ‘Railroads of the Raj: Estimating the Impact of Transportation Infrastructure’, Asia Research Centre, Working Paper 41 (also forthcoming in *American Economic Review*).

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 Travelling to the Market’, *American Journal of Agricultural Economics*, 87, 3, August, 717-734.

Fafchamps, M. and B. Minten, 2012, 'Impact of SMS-Based Agricultural Information on Indian Farmers', *World Bank Economic Review*, 27, February, 1-32.

Fafchamps, M. and J.C.Aker, 2014, 'Mobile Phone Coverage and Producer Markets: Evidence from West Africa', *World Bank Economic Review*, 29, 2, 262–292.

Feyrer, J., 2009, 'Distance, Trade, and Income – The 1967 to 1975 Closing of the Suez Canal as a Natural Experiment', NBER, working paper 15557.

Goyal, A., 2010, 'Information, Direct Access to Farmers and Rural Market Performance in Central India', *American Economic Journal: Applied Economics*, 2 (July), 22-45.

Jacoby, H.G., 2000, 'Access to Markets and the Benefits of Rural Roads', *The Economic Journal*, 110, 713-737.

Jacoby, H.G. and B. Minten, 2008, 'On measuring the Benefits of Lower Transport Costs', World Bank, Policy Research Working Paper 4484.

De Janvry, A. And E. Sadoulet, 2010, 'Agriculture for Development in Africa: Business-as Usual or New Departures', *Journal of African Economies*, 19, AERC, Sup. 2, ii7-ii39.

Jedwab, R. and A.Moradi, 2015, 'The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa', *Review of Economics and Statistics*, forthcoming.

Jedwab, R., E. Kerby and A.Moradi, 2014, 'History, Path Dependence and Development: Evidence from Colonial Railroads, Settlers and Cities in Kenya', mimeo.

Jensen, R., 2007, 'The Digital Provide: Information (Technology), Market Performance, and Welfare in the South Indian Fisheries Sector', *Quarterly Journal of Economics*, 72, 3, 879-924.

Kaminski, J., L. Christiaensen and C.L. Gilbert, C.L, 2014, 'The End of Seasonality? New Insights from Sub-Saharan Africa', *World Bank*, Policy Research Working Paper 6907.

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.

Michaels, G., 2008, 'The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System', *Review of Economics and Statistics*, 90, 4, 683-701

Millennium Challenge Corporation, 'Millennium Challenge Corporation: Malawi Compact Program Development (2011 – 2016), Project Concept Paper For The Transport Sector: Promoting Economic Growth And Poverty Reduction Through Addressing Transport Infrastructure Constraints In Malawi', Millennium Challenge Account - Malawi Country Office Secretariat (website www.mca-m.gov.mw, accessed September 2013).

Minten, B. and S. Kyle, 1999, 'The effect of Distance and Road Quality on Food Collection, Marketing Margins, and Traders' Wages: Evidence from the Former Zaire', *Journal of Development Economics*, 60, 467-495.

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, 1887-1896.

Redding, S.J. and D. M. Sturm, 2008, 'The Costs of Remoteness: Evidence from German Division and Reunification', *American Economic Review*, 98, 5, 1766-1797.

Yamauchi, F., M. Muto, S. Chowdurry, R.Dewina and S.Sumaryanto, 2011, 'Are Schooling and Roads Complementary? Evidence from Income Dynamics in Rural Indonesia', World Development, 39, 12, 2232-2244.

Teravaninthorn, S. & G. Raballand, 2008, Transport Prices and Costs in Africa: A Review of the

Main International Corridors, July, Africa Infrastructure Country Diagnostic, report no14.

World Bank, 2006, Sub-Saharan Africa Review of Selected Railway Concessions, Report No. 36491, Washington DC, World Bank.

Zant, W., 2012, ‘The economics of food aid under subsistence farming with an application to Malawi’, *Food Policy,* 37, 1, 124-141.

Zant, W., 2013, ‘How is the Liberalization of Food Markets Progressing? Market Integration and Transaction Costs in Subsistence Economies’, *World Bank Economic Review*, 2013, 27, 1, January, 28-54.

Zant, W., 2014, ‘How does Market Access affect Smallholder Behavior? The Case of Tobacco Marketing in Malawi’, *Tinbergen Institute*, Discussion Paper 12-088/V.

APPENDICES

Appendix 1 Malawi road network

Source: VU SPINlab.

Appendix 2 Samples varying over time and geographical space by commodity

MAIZE		domestic trade between markets within a maximum distance of:				
Period	←|→	80km	90km	100km	110km	120km
1/98-12/04	60,24	-0.069 (.057)	-0.056 (.055)	-0.102** (.050)	-0.096** (.041)	-0.057 (.044)
		1191; 0.4551	1449; 0.4322	1678; 0.3913	2024; 0.3768	2288; 0.3620
1/99-12/04	48,24	-0.107* (.063)	-0.107* (.061)	-0.137** (.055)	-0.125*** (.043)	-0.079 (.049)
		1067; 0.4849	1310; 0.4604	1519; 0.4233	1820; 0.4004	2054; 0.3831
1/00-12/04	36,24	-0.103 (.077)	-0.095 (.069)	-0.135** (.057)	-0.111** (.044)	-0.061 (.051)
		919; 0.5350	1144; 0.5038	1330; 0.4666	1584; 0.4411	1794; 0.4148
1/98-12/05	60,36	-0.053 (.047)	-0.042 (.041)	-0.076* (.045)	-.063* (.035)	-0.039 (.037)
		1614; 0.3835	2009; 0.3592	2306; 0.3324	2772; 0.3186	3123; 0.3089
1/99-12/05	48,36	-0.076 (.055)	-0.056 (.048)	-0.092* (.050)	-0.071* (.040)	-0.050 (.040)
		1490; 0.3924	1870; 0.3673	2147; 0.3453	2568; 0.3277	2889; 0.3163
1/00-12/05	36,36	-.078 (.058)	-0.043 (.051)	-0.093* (.051)	-0.058 (.042)	-0.037 (.040)
		1342; 0.4159	1704; 0.3887	1958; 0.3686	2332; 0.3487	2629; 0.3338
1/98-4/05	60,28	-0.084* (.050)	-0.086* (.049)	-0.113** (.048)	-0.095** (.037)	-0.063 (.041)
		1329; 0.4458	1630; 0.4253	1882; 0.3856	2266; 0.3716	2558; 0.3610
1/99-4/05	48,28	-0.126** (.051)	-0.130** (.050)	-0.147*** (.051)	-0.119*** (.039)	-0.083*(.044)
		1205; 0.4665	1491; 0.4453	1723; 0.4097	2062; 0.3898	2324; 0.3775
1/00-4/05	36,28	-0.124** (.050)	-0.106** (.047)	-0.139*** (.052)	-0.102** (.040)	-0.065* (.043)
		1057; 0.5009	1325; 0.4757	1534; 0.4410	1826; 0.4199	2064; 0.4031

RICE		domestic trade between markets within a maximum distance of:				
Period	←|→	140km	150km	160km	170km	180km
1/98-12/04	60,24	-0.072 (.047)	-0.100** (.049)	-0.115** (.049)	-0.076* (.046)	-0.077* (.042)
		2529; 0.4415	2678; 0.4290	2791; 0.4249	3086; 0.4114	3292; 0.3945
1/99-12/04	48,24	-0.108** (.045)	-0.133*** (.047)	-0.138*** (.046)	-0.088* (.045)	-0.085** (.042)
		2259; 0.4539	2393; 0.4411	2494; 0.4364	2752; 0.4250	2938; 0.4077
1/00-12/04	36,24	-0.069 (.052)	-0.091* (.054)	-0.091* (.053)	-0.063 (.047)	-0.056 (.043)
		1892; 0.4677	2007; 0.4537	2102; 0.4549	2317; 0.4424	2476; 0.4289
1/98-12/05	60,36	-0.058 (.040)	-0.078* (.041)	-0.088** (.041)	-0.062* (.035)	-0.064* (.032)
		3305; 0.4187	3506; 0.4106	3647; 0.4084	4036; 0.3986	4308; 0.3868
1/99-12/05	48,36	-0.073* (.038)	-0.092** (.039)	-0.094** (.038)	-0.063* (.035)	-0.060* (.032)
		3035; 0.4235	3221; 0.4161	3350; 0.4141	3702; 0.4066	3954; 0.3957
1/00-12/05	36,36	-0.040 (.042)	-0.058 (.042)	-0.058 (.041)	-0.035 (.036)	-0.031 (.033)
		2668; 0.4345	2835; 0.4273	2958; 0.4303	3267; 0.4221	3492; 0.4121
1/98-4/05	60,28	-0.062 (.043)	-0.086* (.044)	-0.100** (.044)	-0.072* (.039)	-0.075** (.037)
		2761; 0.4323	2924; 0.4203	3045; 0.4168	3372; 0.4058	3595; 0.3915
1/99-4/05	48,28	-0.081* (.042)	-0.102** (.042)	-0.109*** (.042)	-0.073* (.038)	-0.072** (.036)
		2491; 0.4387	2639; 0.4269	2748; 0.4234	3038; 0.4142	3241; 0.4005
1/00-4/05	36,28	-0.042 (.046)	-0.061 (.046)	-0.063 (.046)	-0.042 (.040)	-0.040 (.037)
		2124; 0.4483	2253; 0.4370	2356; 0.4390	2603; 0.4284	2779; 0.4174

GROUNDNUTS		domestic trade between markets within a maximum distance of:				
period	←|→	200km	225km	250km	275km	300km
1/98-12/04	60,24	-0.073 (.047)	-0.068 (.043)	-0.064 (.040)	-0.063* (.037)	-0.056* (.033)
		2329; 0.4386	2735; 0.4256	3035; 0.4077	3356; 0.3995	3679; 0.3902
1/99-12/04	48,24	-0.105** (.051)	-0.104** (.046)	-0.098** (.041)	-0.086** (.038)	-0.083** (.035)
		2007; 0.4645	2359; 0.4544	2622; 0.4385	2901; 0.4346	3174; 0.4269
1/00-12/04	36,24	-0.101* (.051)	-0.104** (.047)	-0.098** (.042)	-0.082** (.038)	-0.085** (.037)
		1582; 0.5116	1869; 0.4963	2089; 0.4854	2316; 0.4733	2540; 0.4633
1/98-12/05	60,36	-0.021 (.035)	-0.038 (.032)	-0.033 (.030)	-0.035 (.028)	-0.029 (.026)
		3280; 0.4039	3845; 0.3958	4291; 0.3830	4747; 0.3725	5204; 0.3688
1/99-12/05	48,36	-0.035 (.035)	-0.056* (.032)	-0.051* (0.029)	-0.044 (.027)	-0.042 (.026)
		2958; 0.4240	3469; 0.4185	3878; 0.4081	4292; 0.3981	4699; 0.3950
1/00-12/05	36,36	-0.014 (.038)	-0.039 (.036)	-0.034 (.033)	-0.027 (.031)	-0.039 (.030)
		2533; 0.4511	2979; 0.4415	3345; 0.4349	3707; 0.4190	4065; 0.4148
1/98-4/05	60,28	-0.071 (.044)	-0.065 (.039)	-0.059 (.037)	-0.062* (.034)	-0.055* (.030)
		2554; 0.4175	3003; 0.4080	3343; 0.3906	3700; 0.3785	4066; 0.3709
1/99-4/05	48,28	-0.104** (.047)	-0.102** (.042)	-0.094** (.037)	-0.080** (.034)	-.077** (.031)
		2232; 0.4418	2627; 0.4363	2930; 0.4202	3245; 0.4103	3561; 0.4040
1/00-4/05	36,28	-0.094** (.044)	-0.097** (.040)	- 0.091** (.037)	-0.075** (.034)	-0.083** (.033)
		1807; 0.4794	2137; 0.4691	2397; 0.4580	2660; 0.4392	2927; 0.4310

BEANS		domestic trade between markets within a maximum distance of:				
Period	←|→	80km	90km	100km	110km	120km
1/99-12/04	48,24	-0.074 (.077)	-0.060 (.066)	-0.078(.054)	-0.107***(.039)	-0.038 (.059)
		988; 0.5452	1190; 0.5512	1401; 0.5243	1690; 0.4953	1917; 0.47111
1/00-12/04	36,24	-0.140* (.071)	-0.129** (.060)	-0.098* (.050)	-0.122*** (.038)	-0.048 (.060)
		813; 0.5602	991; 0.5601	1168; 0.5267	1393; 0.4973	1586; 0.4797
1/01-12/04	24,24	-0.183*** (.059)	-0.179*** (.054)	-0.141*** (.042)	-0.166*** (.039)	-0.072 (.081)
		597; 0.6367	742; 0.6146	873; 0.5829	1032; 0.5427	1182; 0.5181
1/99-12/05	48,36	-0.007 (.052)	-0.006 (.044)	-0.036 (.030)	-0.081*** (.027)	-0.026 (.040)
		1259; 0.5104	1581; 0.5343	1833; 0.5103	2206; 0.4766	2483; 0.4553
1/00-12/05	36,36	-0.028 (.058)	-0.028 (.052)	-0.038 (.036)	-0.085** (.033)	-0.041 (.039)
		1084; 0.5173	1382; 0.5400	1600; 0.5125	1909; 0.4806	2152; 0.4652
1/01-12/05	24,36	-0.096 (.074)	-0.082 (.066)	-0.095** (.041)	-0.131*** (.045)	-0.054 (.052)
		868; 0.5640	1133; 0.5786	1305; 0.5534	1548; 0.5172	1737; 0.4742
1/99-4/05	48,28	-0.064 (.064)	-0.069 (.057)	-0.078* (.044)	-0.099*** (.033)	-0.034 (.050)
		1060; .5271	1298; 0.5402	1518; 0.5182	1831; 0.4830	2068; 0.4644
1/00-4/05	36,28	-0.121** (.055)	-0.122** (.055)	-0.094** (.041)	-0.116$^{* **}$ (.034)	-0.040 (.051)
		885; 0.5402	1099; 0.5479	1285; 0.5147	1534; 0.4833	1752; 0.4709
1/01-4/05	24,28	-0.196*** (.049)	-0.201*** (.053)	-0.162*** (.043)	-0.198*** (.040)	-0.115 (.073)
		669; 0.5987	850; 0.5925	990; 0.5601	1173; 0.5244	1333; 0.5060

Note to table: The table reports Population Average Treatment Effects (ATE) with data that are restricted to market-pairs within a range of specified maximum distance of each other and restricted to a varying pre- and post-intervention sample period. ←|→ is the number of months before and since January 2003; * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Next to the ATE coefficient in brackets are robust standard errors clustered by market-pairs, and below the coefficient the number of observations and R^2. Estimated specification is identical to the specification reported in Table 1.

Appendix 3 Cluster-robust standard errors: two-way clustering

Table A3 Impact (ATE) of rail transport services on price dispersion: selected output with two-way clustered standard errors

dependent variable: $\ln(\text{abs}(p_{jt}-p_{kt}))$	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.119***	-0.109***	-0.094***	-0.116***
	(0.0288)	(0.0352)	(0.0316)	(0.0154)
lagged dependent variable (t-1)	0.183***	0.182***	0.245***	0.193***
	(0.0203)	(0.0249)	(0.0134)	(0.0426)
lagged dependent variable (t-2)		0.096***	0.056***	0.095**
		(0.0238)	(0.0110)	(0.0149)
lagged dependent variable (t-3)			0.053*	
			(0.0282)	
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
covariates	no	No	no	no
centered R^2	0.3280	0.3926	0.3908	0.4623
max trading distance (km)	110	160	250	110
sample period	1/99-4/05	1/99-4/05	1/99-4/05	1/00-4/05
months before 1/03 and after 12/02	48; 28	48; 28	48; 28	36; 28
number of observations	2062	2748	2930	1534
no. of intervention pairs	251	366	365	157
no. of control pairs connected by rail	161	190	210	107
no. of other controls (not connected)	1650	2192	2355	1270
long term impact	-0.146	-0.150	-0.146	-0.164

Note to table: See Table 1 and 2. Robust standard errors are clustered by each market of market pairs.

Appendix 4 Testing robustness of the impact of rail transport services on price dispersion

Table A4-1 Using remote market-pairs as control group

dependent variable: $\ln(\text{abs}(p_{jt}-p_{kt}))$	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.177**	-0.172**	-0.151**	-0.189*
	(0.0817)	(.0769)	(0.0715)	(0.1110)
lagged dependent variable (t-1)	0.140***	0.182***	0.312***	0.225***
	(0.0498)	(0.0317)	(0.0312)	(0.0529)
lagged dependent variable (t-2)		0.105***	0.099***	
		(0.0317)	(0.0715)	
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
covariates	no	no	no	no
R^2	0.5038	0.4803	0.5194	0.6405
max trading distance (km)	110	180	200	110
sample period	1/99-4/05	1/99-4/05	1/99-4/05	1/00-4/05
months before 1/03 and after 12/02	48; 28	48; 28	48; 28	36; 28
number of observations	998	1635	1559	846
no. of intervention pairs	251	502	422	178
no. of control pairs connected by rail	161	235	219	135
no. of other controls (not connected)	586	898	918	533
long term impact	-0.206	-0.242	-0.257	-0.244

Table A4-2 Using nearby market-pairs as a control group

dependent variable: $\ln(\text{abs}(p_{jt}-p_{kt}))$	(1)	(2)	(3)	(4)
crop / commodity	maize	rice	groundnuts	beans
connected by rail	-0.130**	-0.174***	-0.109*	-0.203**
	(0.0572)	(0.0649)	(0.0581)	(0.0846)
lagged dependent variable (t-1)	0.170***	0.194***	0.265***	0.178**
	(0.0519)	(0.0278)	(0.0255)	(0.0743)
lagged dependent variable (t-2)		0.087***	0.056**	0.120**
		(0.0279)	(0.0276)	(0.0578)
season x market dummies	yes	yes	yes	yes
time trend x market dummies	yes	yes	yes	yes
market-pair dummies	yes	yes	yes	yes
month dummies	yes	yes	yes	yes
covariates	no	no	no	no
R^2	0.4675	0.4572	0.4111	0.4903
max trading distance (km)	90	140	225	110
sample period	1/99-4/05	1/99-4/05	1/99-4/05	1/00-4/05
months before 1/03 and after 12/02	48; 28	48; 28	48; 28	36; 28
number of observations	1036	1859	2225	1058
no. of intervention pairs	175	342	451	128
no. of control pairs connected by rail	110	178	241	97
no. of other controls (not connected)	751	1339	1533	833
long term impact	-0.157	-0.243	-0.160	-0.290

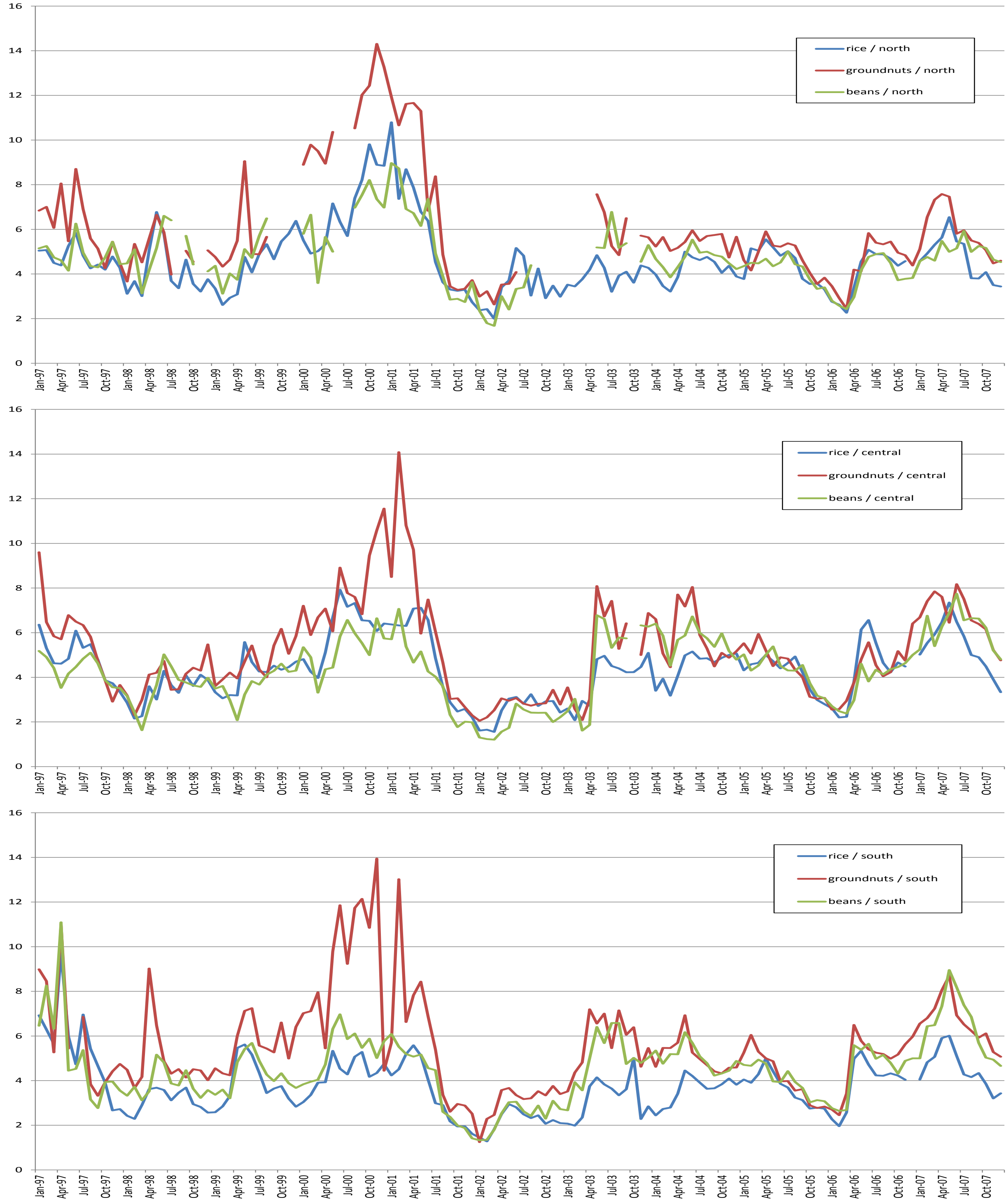
Note to table: see Table 1

Appendix 5 Availability of market price data of agricultural commodities (27 markets)

Figure A5 Number of market price observations per year (upper) and per month (lower)

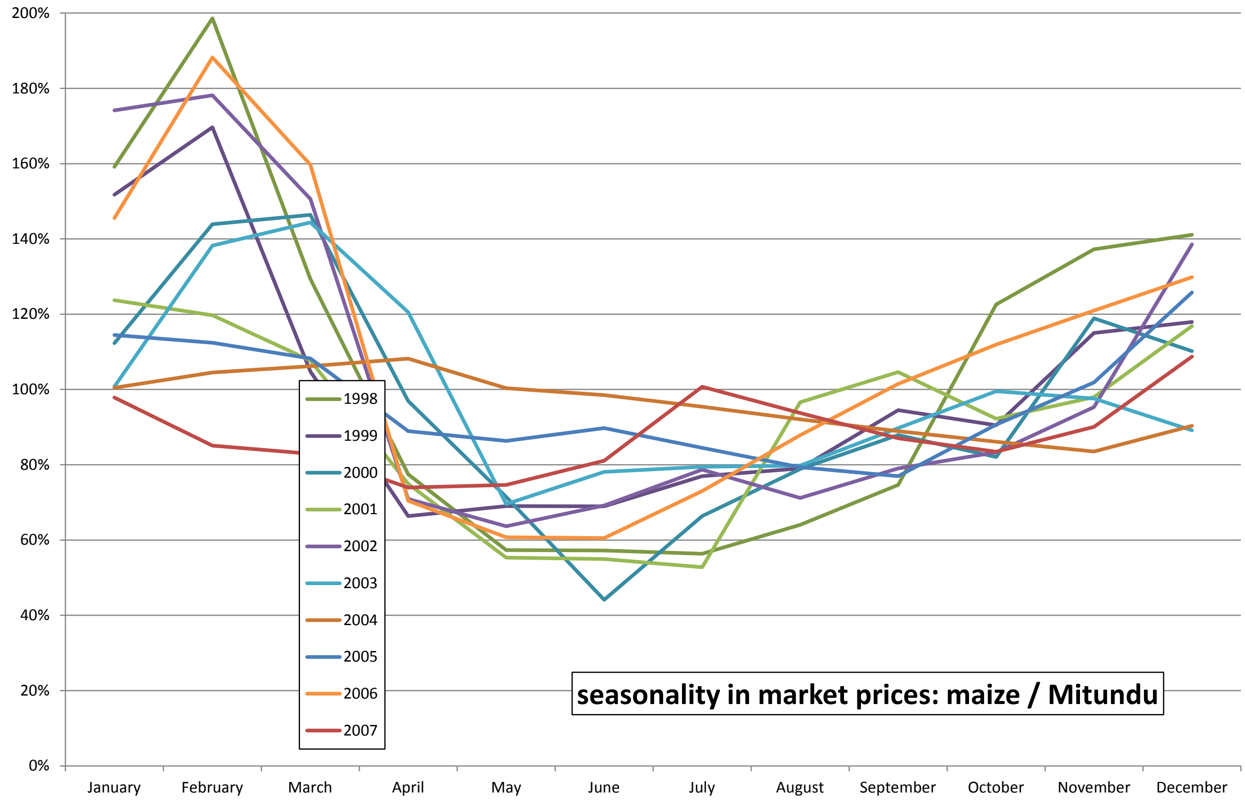
Appendix 6 Prices of agricultural commodities

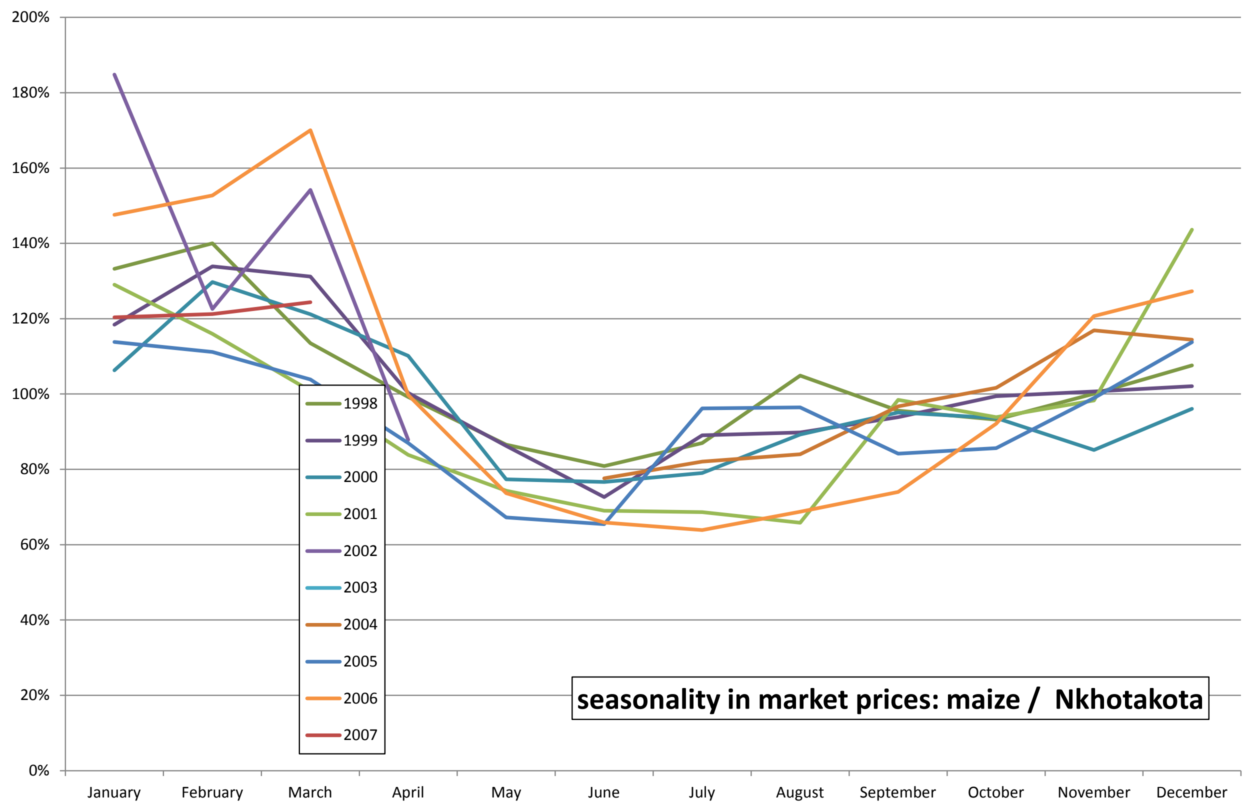
Figure A6-1 Rice, groundnuts and bean prices relative to maize prices

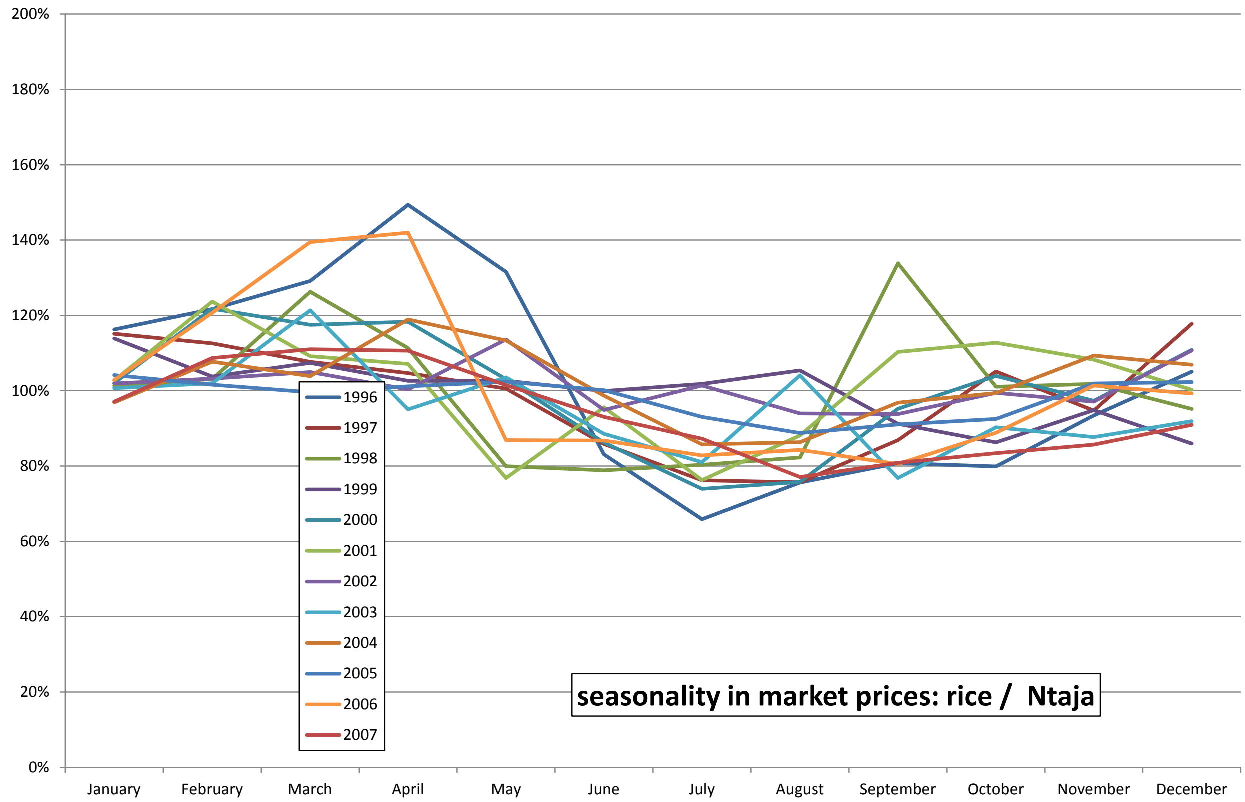


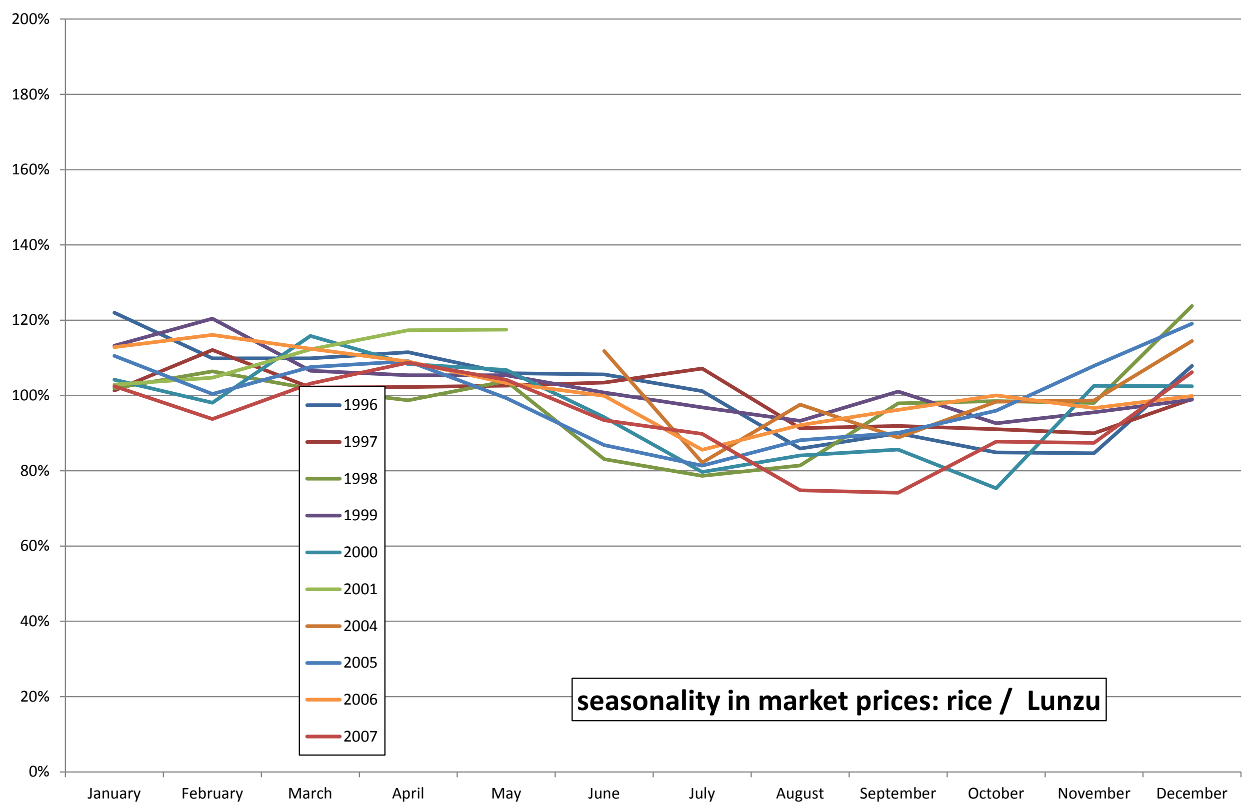
Note to figure: monthly price series of rice, groundnuts and beans are expressed relative to maize prices, for each month and each market and these relative prices are averaged over all markets.

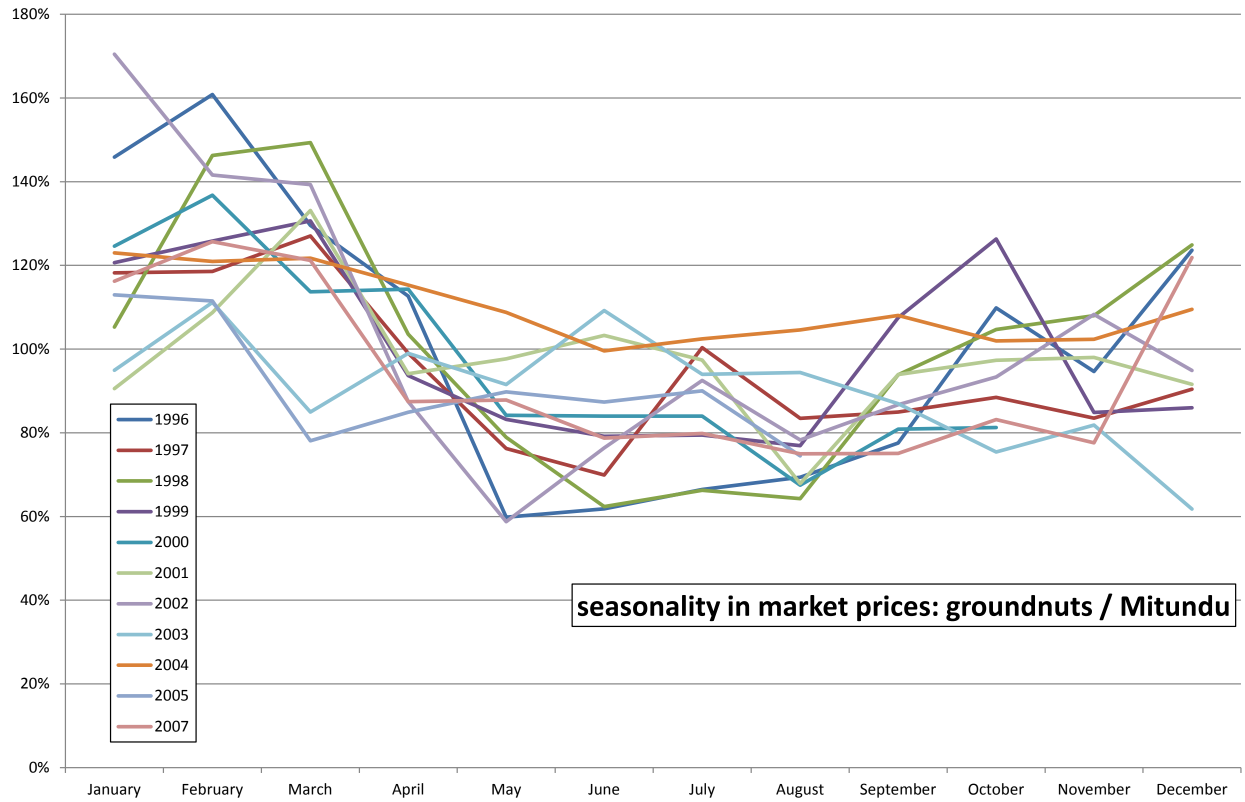
Figure A6-2 Seasonality of market prices of agricultural commodities

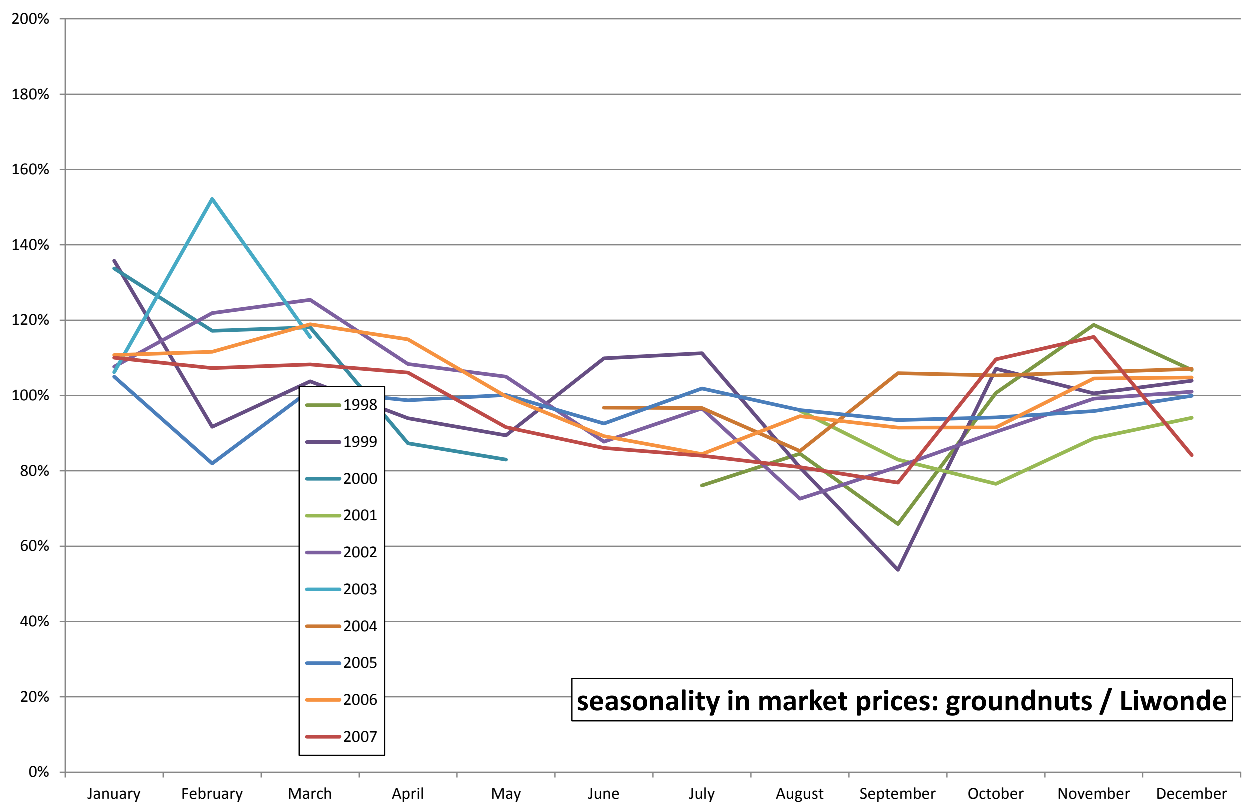


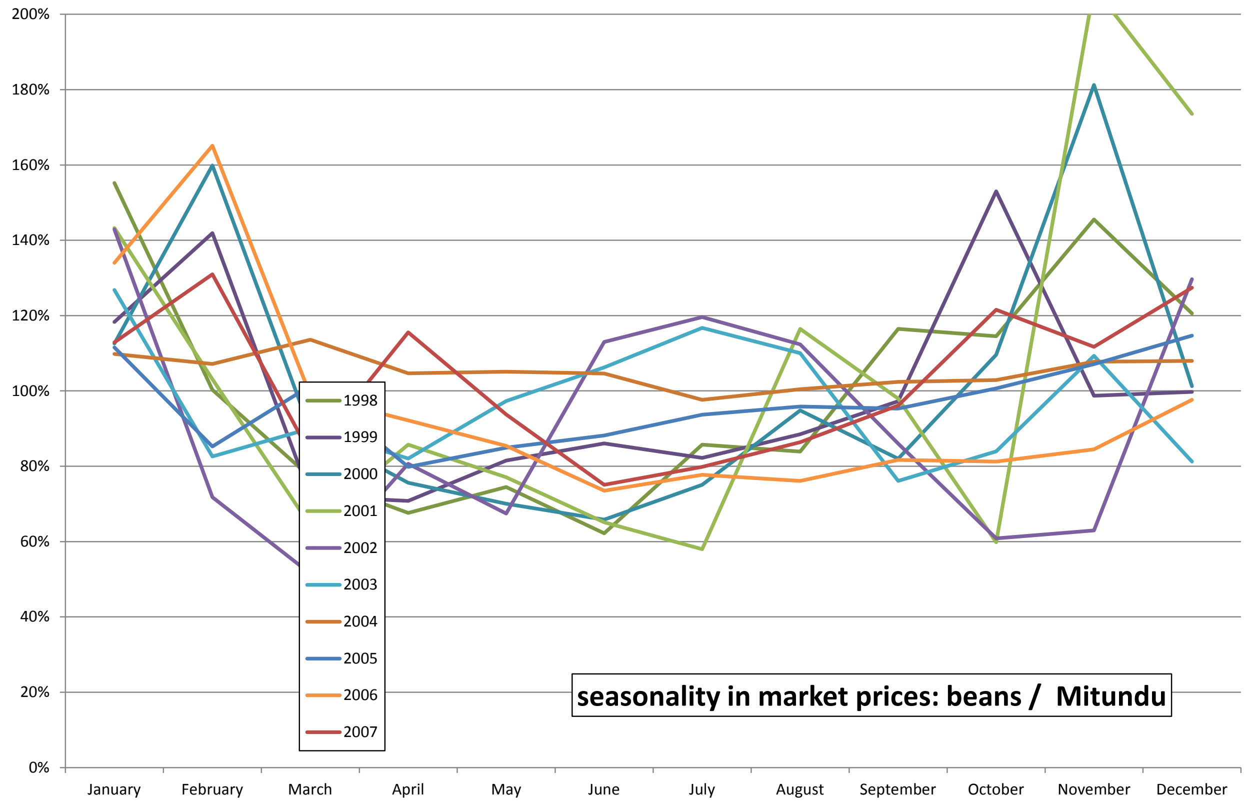


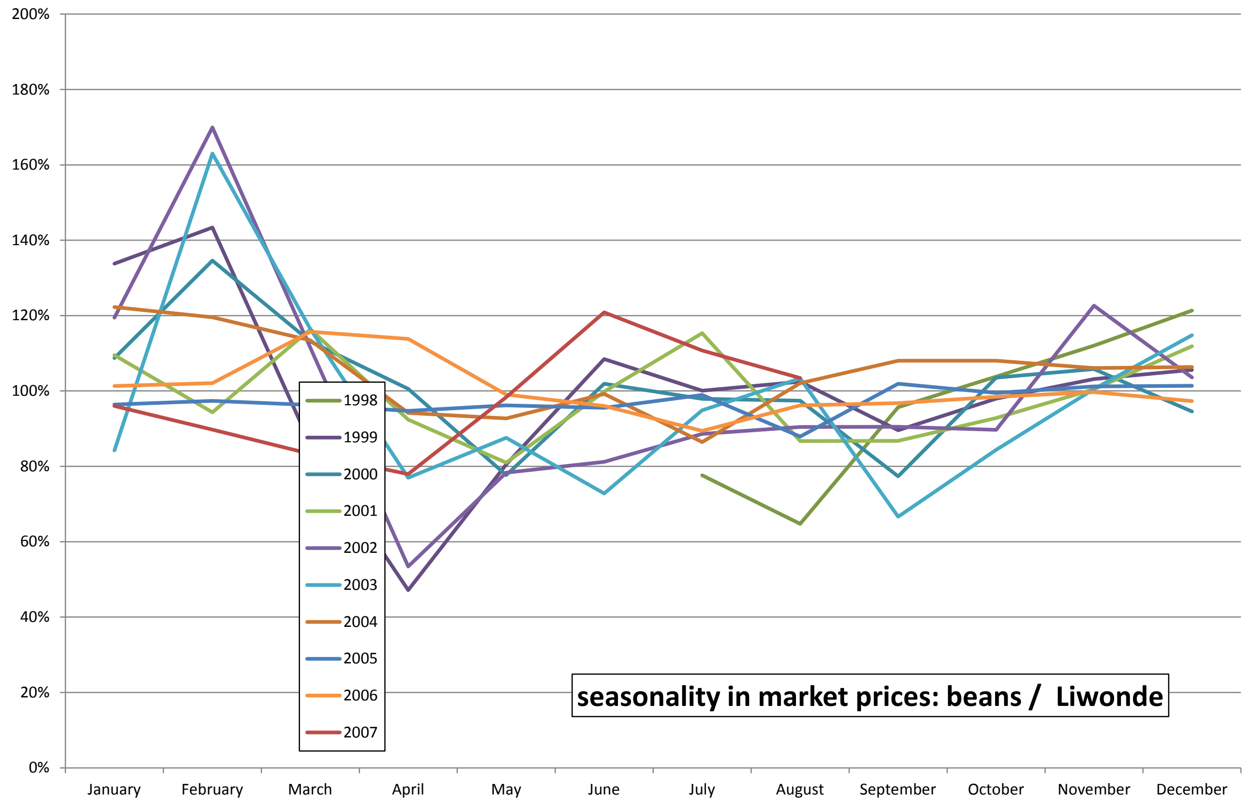

seasonality in market prices: rice / Ntaja
200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
January
February
March
April
May
June
July
August
September
October
November
December
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007


seasonality in market prices: rice / Lunzu
200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
January
February
March
April
May
June
July
August
September
October
November
December
1996
1997
1998
1999
2000
2001
2004
2005
2006
2007


180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2007
seasonality in market prices: groundnuts / Mitundu
January
February
March
April
May
June
July
August
September
October
November
December


200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
seasonality in market prices: groundnuts / Liwonde
January
February
March
April
May
June
July
August
September
October
November
December


200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
January
February
March
April
May
June
July
August
September
October
November
December
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
seasonality in market prices: beans / Mitundu


200%
180%
160%
140%
120%
100%
80%
60%
40%
20%
0%
January
February
March
April
May
June
July
August
September
October
November
December
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
seasonality in market prices: beans / Liwonde

Appendix 7 Geographical distribution of crop production

Figure A7-1 Per capita production by district (averages of annuals 1995/96-2007/08)

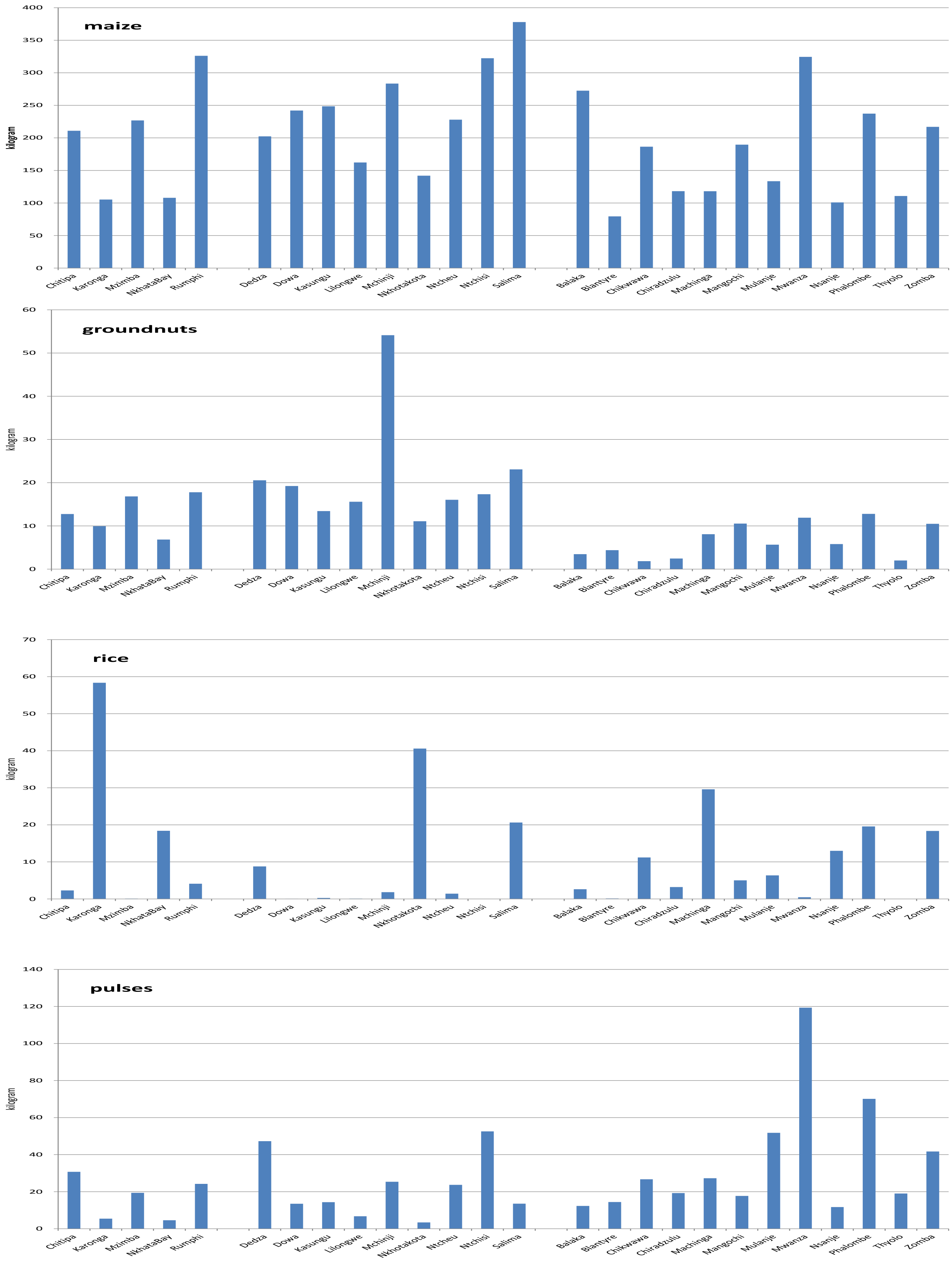
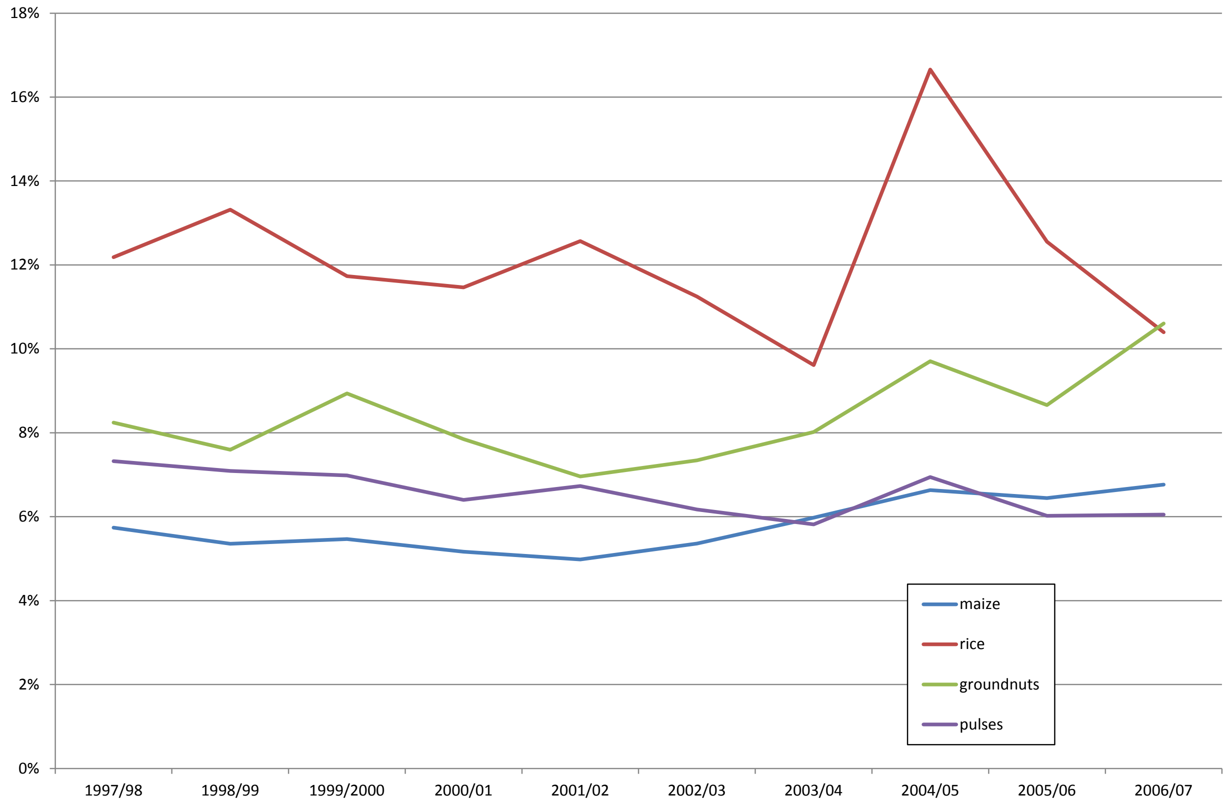


Figure A7-2 Concentration of crop production by district (Hirschman-Herfindhal index)

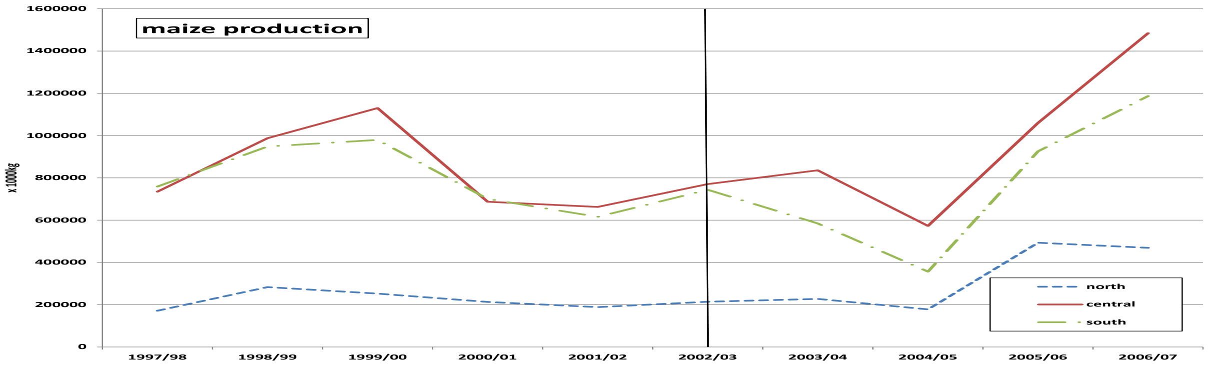


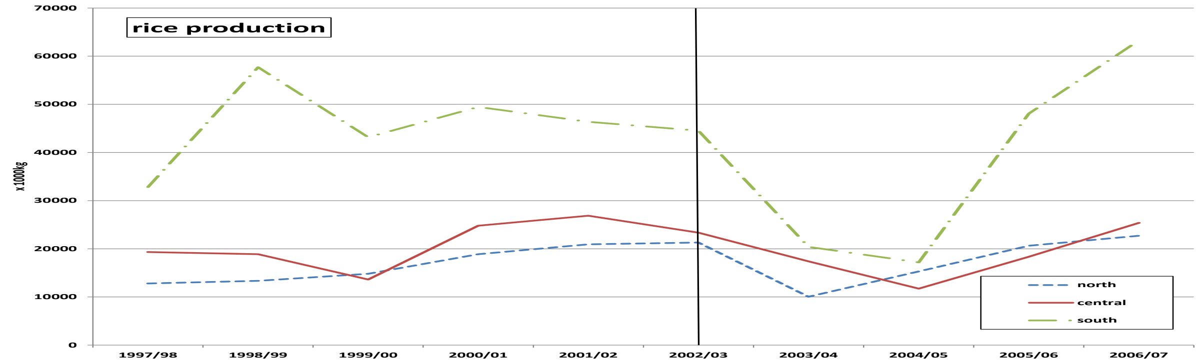
Appendix 8 Markets: location names, district, coordinates and distance to railway station

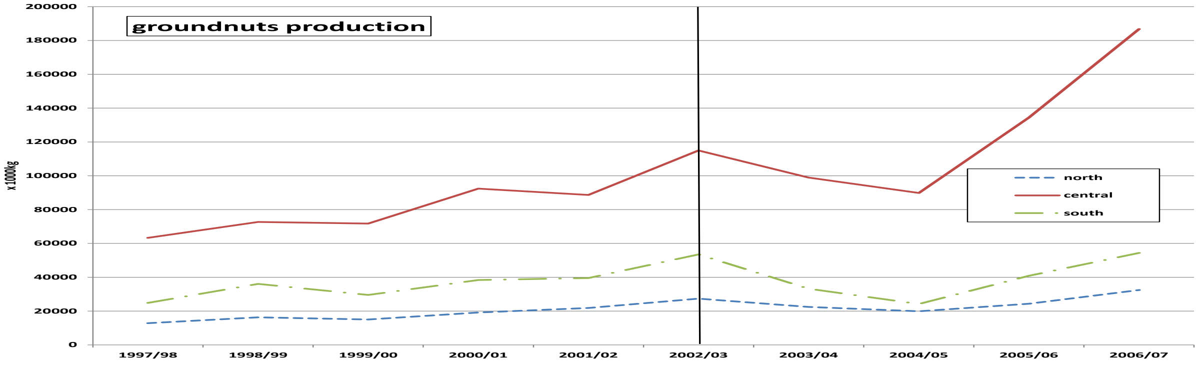
market	district/RDP*	coordinates		distance to nearest railway station
		latitude	longitude	
Chitipa	Chitipa	-9.69958	33.27001	459.6
Karonga	Karonga	-9.93926	33.92713	431.4
Mzimba	Mzimba	-11.8946	33.59682	225.2
Mzuzu	Mzimba	-11.4561	34.01450	263.0
Nkhatabay	Nkhata Bay	-11.6074	34.29762	242.6
Rumphi	Rumphi	-11.0155	33.85760	314.4
Chimbiya	Ntcheu	-15.0822	34.58972	39.5
Dowa	Dowa	-13.6532	33.93434	32.1
Kasungu	Kasungu	-13.0332	33.48348	104.4
Lilongwe	Lilongwe	-13.9810	33.78668	5.8
Lizulu	Ntcheu	-14.4377	34.42248	19.3
Mchinji	Mchinji	-13.7998	32.88052	1.6
Mitundu	Lilongwe	-14.2418	33.77091	24.4
Nkhotakota	Nkhotakota	-12.9254	34.28384	96.8
Ntchisi	Ntchisi	-13.3761	33.86522	63.8
Salima	Salima	-13.7796	34.45818	2.7
Bangula	Nsanje	-16.5817	35.11641	4.0
Limbe	Blantyre	-15.8082	35.05741	1.2
Liwonde	Machinga	-15.0662	35.23374	0.2
Luchenza	Thyolo	-16.0018	35.30928	0.7
Lunzu	Blantyre	-15.6515	35.02027	3.4
Mangochi	Mangochi	-14.4777	35.26370	57.5
Namwera	Mangochi	-14.3449	35.48377	72.2
Nchalo	Chikwawa	-16.2727	34.86774	40.5
Nsanje	Nsanje	-16.9213	35.26095	22.6
Ntaja	Machinga	-14.8667	35.52608	16.9
Zomba	Zomba	-15.3805	35.33286	27.0

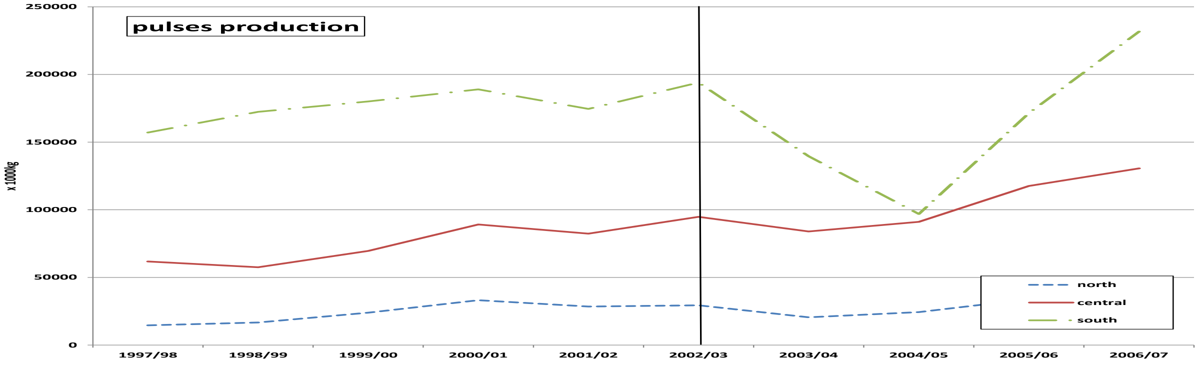
Note to table: RDP = Rural Development Project; Source: Euclidean distance calculated using lat-lon coordinates from www.geonames.org. Distance Nsanja-railway station = distance Nasanje to Tengani station, which is the nearest station.

Appendix 9 Price & production of maize, groundnuts, rice & pulses, 1997/98-2006/07

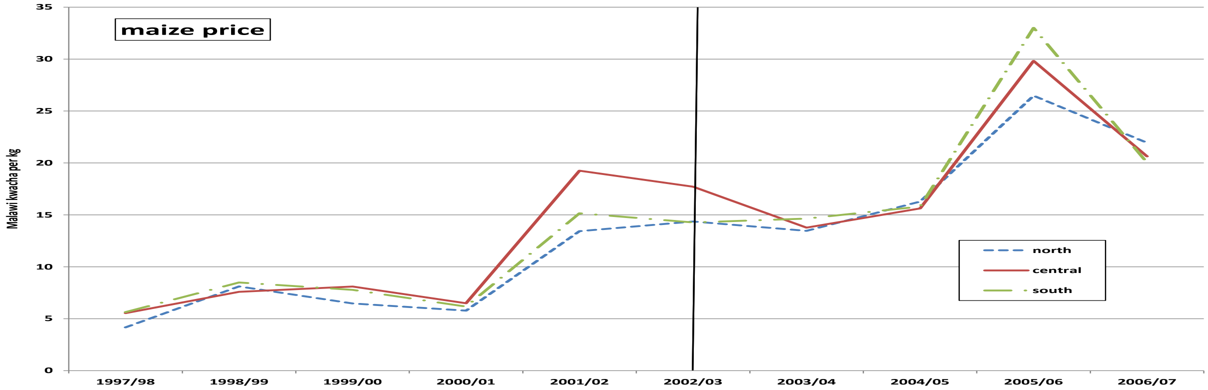


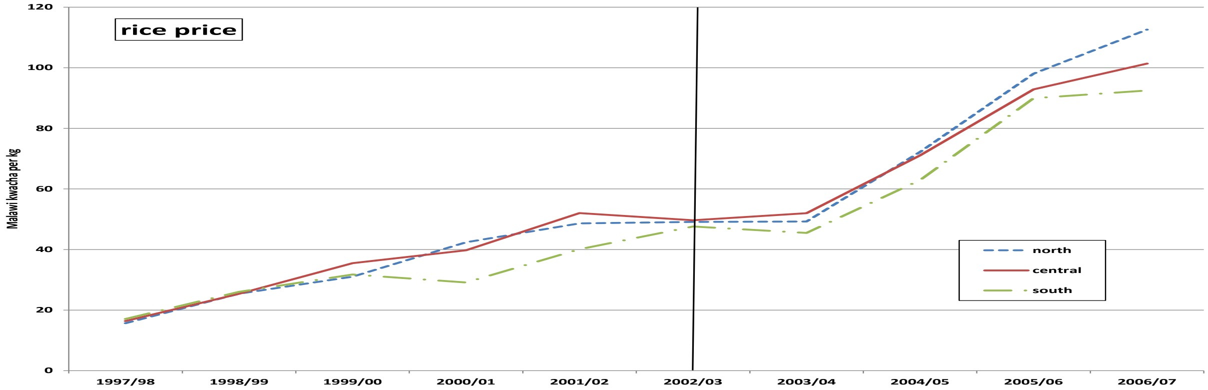


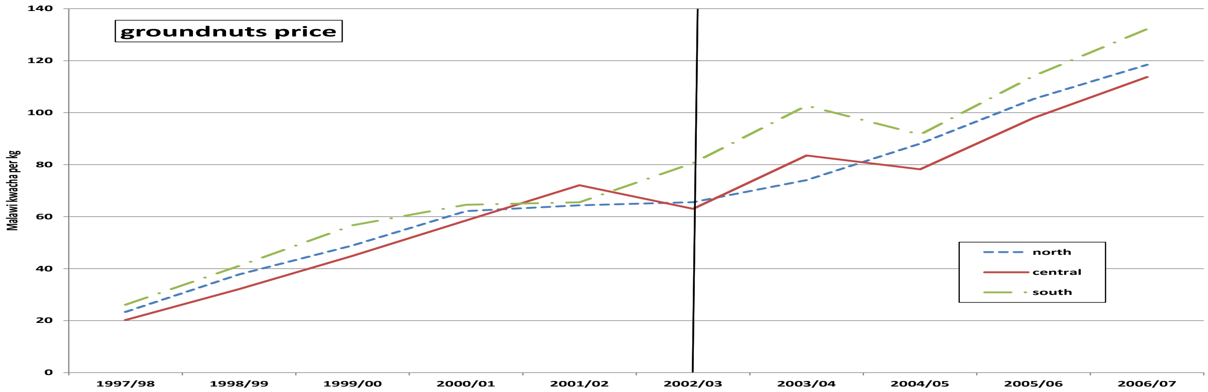


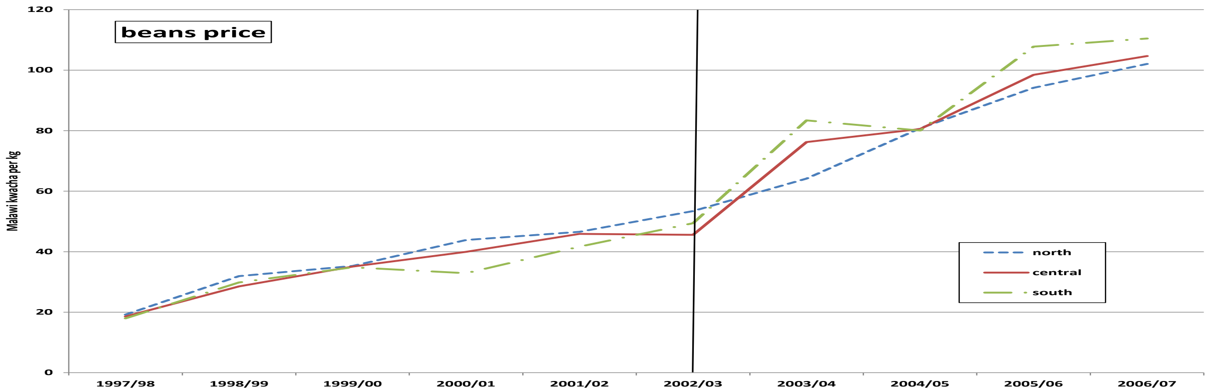


Appendix 9 Price & production of maize, groundnuts, rice & pulses, 1997/98-2006/07(cont.)









Appendix 10 Freight by rail