

Zant, Wouter

Working Paper
Bridges

Tinbergen Institute Discussion Paper, No. 17-117/V

Provided in Cooperation with:

Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Zant, Wouter (2017) : Bridges, Tinbergen Institute Discussion Paper, No. 17-117/V, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/177685>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

TI 2017-117/V
Tinbergen Institute Discussion Paper



Bridges

Wouter (W.) Zant¹

1: VU Amsterdam, The Netherlands; Tinbergen Institute, The Netherlands

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

Contact: discussionpapers@tinbergen.nl

More TI discussion papers can be downloaded at the [Tinbergen Site](#)

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 598 4580

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900

Bridges

Wouter Zant*

Abstract

We estimate to what extent bridges in Mozambique lead to transport cost reductions and attribute these reductions to key determinants, in particular road distance, road quality and crossing borders. For identification we exploit the introduction of a road bridge over the Zambezi river, in August 2009 between Caia and Chimuara, with the simultaneously completed rehabilitation of the railway bridge between Vila de Sena and Mutarara, which jointly create variation in optimal trading itineraries between markets. Estimations, based on monthly maize prices in 24 markets, for up to five years before and after the (re)introduction of the bridges, are conditioned on spatial price equilibrium with market integration. Bridges explain a reduction of transport costs, averaged over itineraries, from 17% to 42% (ATT). Results are shown to be robust for the non-random bridge placement and various other threats. Reduction in transport costs for particular itineraries is occasionally even larger, and is, generally, mainly due to road distance and modestly to road quality.

JEL code: D23, D61, O13, O18, Q13, R41

Key words: agricultural markets, transport costs, bridges, Mozambique, sub-Saharan Africa

* Wouter Zant is associate professor at the VU University Amsterdam and research fellow of the Tinbergen Institute, the Netherlands; mailing address: Vrije Universiteit, De Boelelaan 1105, room 10A-79, 1081 HV Amsterdam, The Netherlands; e-mail address: wouter.zant@vu.nl; tel: +31 20 598 9592.

Introduction

Trade costs are important for developing countries, both for rural development and economic growth. Most sub-Saharan African countries are plagued with high to extremely high trade costs, which has major implications for the operation of these economies. Since the seminal theoretical work of Key, de Janvry and Sadoulet (2000) and de Janvry and Sadoulet (2006) there is a consistent basis for explaining the rationality of subsistence farming and the key role of transport costs in this outcome, where transport costs affect both inputs and output. Their framework goes a long way in explaining low input levels, low productivity and low technological progress in agriculture. Various subsequent contributions have supplied empirical evidence supporting the idea that transport costs constitute a major driver in subsistence farming (see for example Omamo, 1998). Conversely, low trade costs or major reductions in trade costs lead to improvements in the operation of markets. Market efficiency will increase as agricultural produce moves more easily from low-price mostly rural surplus areas to high-price mostly urban deficit areas. A wide range of studies have empirically investigated to what extent infrastructure – be it road infrastructure (roads, bridges, etc.), rail infrastructure (railroads) or communication infrastructure (ICT, mobile phones) – have improved the efficiency of markets and raised the welfare of households. A major question in these studies is who benefits from these infrastructure improvements? Are the benefits flowing to consumers, to traders or to producers? And are these infrastructure improvements helpful in raising welfare of households and alleviation of poverty? Results so far are mixed, suggesting in some cases that both producers and consumers realize welfare gains (see for example Jensen, 2007), while in others welfare gains for farmers are not evident and traders are more likely to benefit (see for example Fafchamps and Minten, 2012; Zant, 2017). Apart from the question who gets the benefits it is clear that improved operation of markets, notably of food markets, is helpful in increasing food security.

In this study we investigate to what extent bridges lead to lower trade costs, using the spatial variation of maize prices over major maize markets in Mozambique. For identification we exploit the introduction of a road bridge over the Zambezi river, in August 2009 between Caia and Chimuara, jointly with the simultaneous completion of the rehabilitation of the railway bridge between Vila de Sena and Mutarara. Correspondence in timing allows to mark a period in which absence of the Zambezi bridges overlaps, creating the required variation in optimal trading itineraries between markets. Monthly maize price data used for this study extend to around five years before and after the introduction of the new bridge and the (re)introduction of the rehabilitated bridge. Estimations are restricted to observations that represent spatial price equilibrium with integrated markets.

The topic of the current study is related to several lines of research in the empirical literature, on the impact of transport infrastructure (Donaldson, 2010; Banerjee, Duflo and Qian, 2012; Casaburi et al. 2013; Brooks and Donovan, 2017; Volpe Martincus, Carballo and Cusolito, 2017), on the operation of markets (Tostão and Brorsen, 2005; Cirera and Arndt, 2008; Zant, 2017), on prices and transport cost (Minten and Kyle, 1999), and on transport costs and behaviour of households (Jacoby, 2000; Renkow, Hallstrom and Karanja, 2005; Jacoby and Minten, 2009). Our study connects the literature on estimation of impacts of infrastructure with the empirical studies on spatial price equilibrium: The key innovation of this study is the use economic determinants of transport costs to find the parity bounds of spatial price equilibrium regimes, which is made feasible by the identification strategy. The outcomes allow us to measure the benefits of bridges in terms of reduced transport costs, averaged over itineraries and for specific itineraries. Additionally we can measure the extent to which markets are integrated. In contrast with much of the spatial economics empirical work the underlying framework of our estimations is structural rather than reduced form.

In the remainder of this paper we discuss, in Section 1, the maize market in Mozambique, the Mozambique road network and transport services, and the role of the Zambezi river and Zambezi bridges in the Mozambique economy. In Section 2 we explain the theory underlying the empirical estimation, we supply details on data and data sources, we develop the empirical strategy and discuss the identification of causal impacts. In Section 3 we present and discuss estimation outcomes and robustness checks. In Section 4 we discuss potential threats and alternative explanation to our estimation outcomes. We conclude with quantifying benefits of the Zambezi bridges for a few specific itineraries, and a summary of findings in Section 5.

1. Maize Markets, Transport Services. Prices and the Zambezi

Maize Markets

Maize is the most important staple food of Mozambique: it is widely produced, marketed, exported and consumed. In all provinces two third of all rural households produce maize. Despite widespread subsistence farming – only around 30% of production is traded on the market – maize is three times more marketed than cassava. Also, maize has a budget share of similar size as all other staple foods¹ together (Tschirley et al., 2006). The calorie share of maize in the average Mozambique diet ranges from 25% to 39%, corresponding with a per capita (annual) consumption of 60 to 85 kg, although, particularly in the south, and in the Maputo region, the maize share is lower due to substitution with rice (Tschirley et al., 2006). Per capita dietary needs also form an indication of the share of production available for trade between regions: Niassa in the north and Manica and Tete in the center are therefore clearly in a position to supply maize to other regions (see Figure 2).

¹ Staples in Mozambique are maize, rice, cassava, wheat, sorghum, millet, sweet potatoes beans and groundnuts.

This partly reflects the geography of maize production: production of maize is concentrated in the central and northern part of Mozambique. The Northern provinces Niassa, Cabo Delgado, and Nampula have better rainfall distribution (see Figure 2) and better soil fertility, while the Southern region has unfavorable weather conditions (see Figure 2) and suffers from occasional pests (Abdula, 2005). Most agricultural production in Mozambique is rain-fed. Drought and also flooding cause occasional drops in production. In the 1999-2000 crop season, maize production declined 18 percent, primarily due to floods that devastated large areas of the Center and South of the country (Abdula, 2005). Major production, assembly and wholesale markets in the central region are Manica, Chimoio and Gorongosa, and in the north Alto Molocue, Montepuez, Mocuba and Ribaue (Figure 1). The major retail terminal markets, nearly all located on the coastline, are, from south to north, Maputo (including Matola), Xai-xai, Maxixe, Massinga, Beira, Quelimane, Nacala and Pemba (see Figure 1). The relatively large concentration of population in the Maputo province (see Figure 2) clearly makes this area a major destination of maize trade.

How is the value chain from maize producers to consumers organized, what is size and structure of the various stages in this value chain, and which agents in particular are the key drivers of spatial arbitrage and integration between markets? From producer to consumer, we find itinerant traders, large scale assemblers, wholesale traders, millers and retailers active in the maize value chain. Close to the end of the chain, near consumers, are retailers and millers: these are more involved in earning returns by value addition rather than earning returns on trade and transport. Wholesale traders take an intermediate position: they buy from assemblers and supply to mills of various sizes. This activity may entail gains from price differences between markets, but is likely to have a large component of value added through collecting, sorting, quality control and distribution. The key agents in Mozambique that drive market integration are traders – mostly informal itinerant traders but also large scale

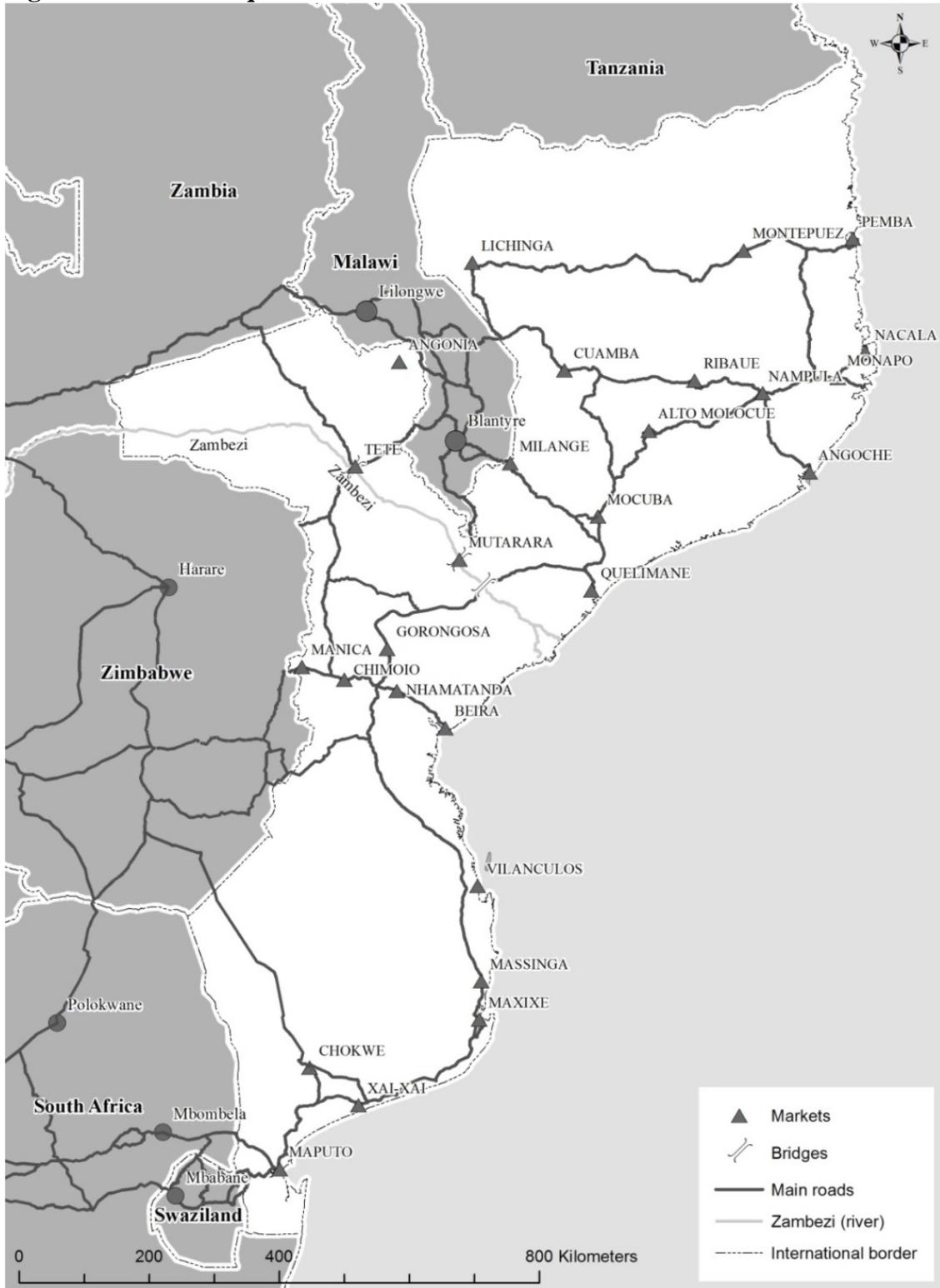
assemblers – and transporters (Zovala, 2014; De Vletter and Polana, 2001). Farmers sell most of their surplus maize to informal itinerant small-scale traders right after harvest. In many markets in Mozambique, both north, central and south, most of the maize traded in assembly and retail markets is supplied by informal traders throughout the year. Informal traders also carry out most of the marketing functions between the rural producers and the urban consumers: they supply their own working capital, hire storage facilities in source / assembly markets and arrange truck transport once a sufficient quantity / number of bags with maize is collected. The following quote characterizes the informal itinerant trader: *“Itinerant traders, commonly known as “Mamas”, are normally comprised of women coming from the southern region of the country, mainly from Maputo city, to buy and bulk maize in surplus maize villages of Central and Northern Mozambique with the objective of reselling it in the maize deficit markets of southern Mozambique. Itinerant traders set up buying points in surplus areas. After acquiring large volumes of maize, they rent trucks to transport their maize to southern markets”* (Zavale, 2014). Competition for informal traders comes from formal large warehouse trader companies, who source maize from the same locations and often operate their own (fleet of) trucks (de Vletter and Polana, 2001).

Maize prices

Maize market prices by region, shown in Figure 3 for the period from 2005 to 2014, have a number of characteristics that have implications for the empirical work in this study. The parallel movement in prices of different regions clearly suggests that prices are correlated. However, prices in the south are overall higher than in the central and northern region, reflecting the difference between surplus and deficit areas. Prices also appear to fluctuate systematically over the season. Figure 4 summarizes the pattern of seasonality for the years 2005 to 2014. Prices consistently peak between November and March, to drop subsequently when the new harvest enters the market. During the lean season prices rise to levels more than

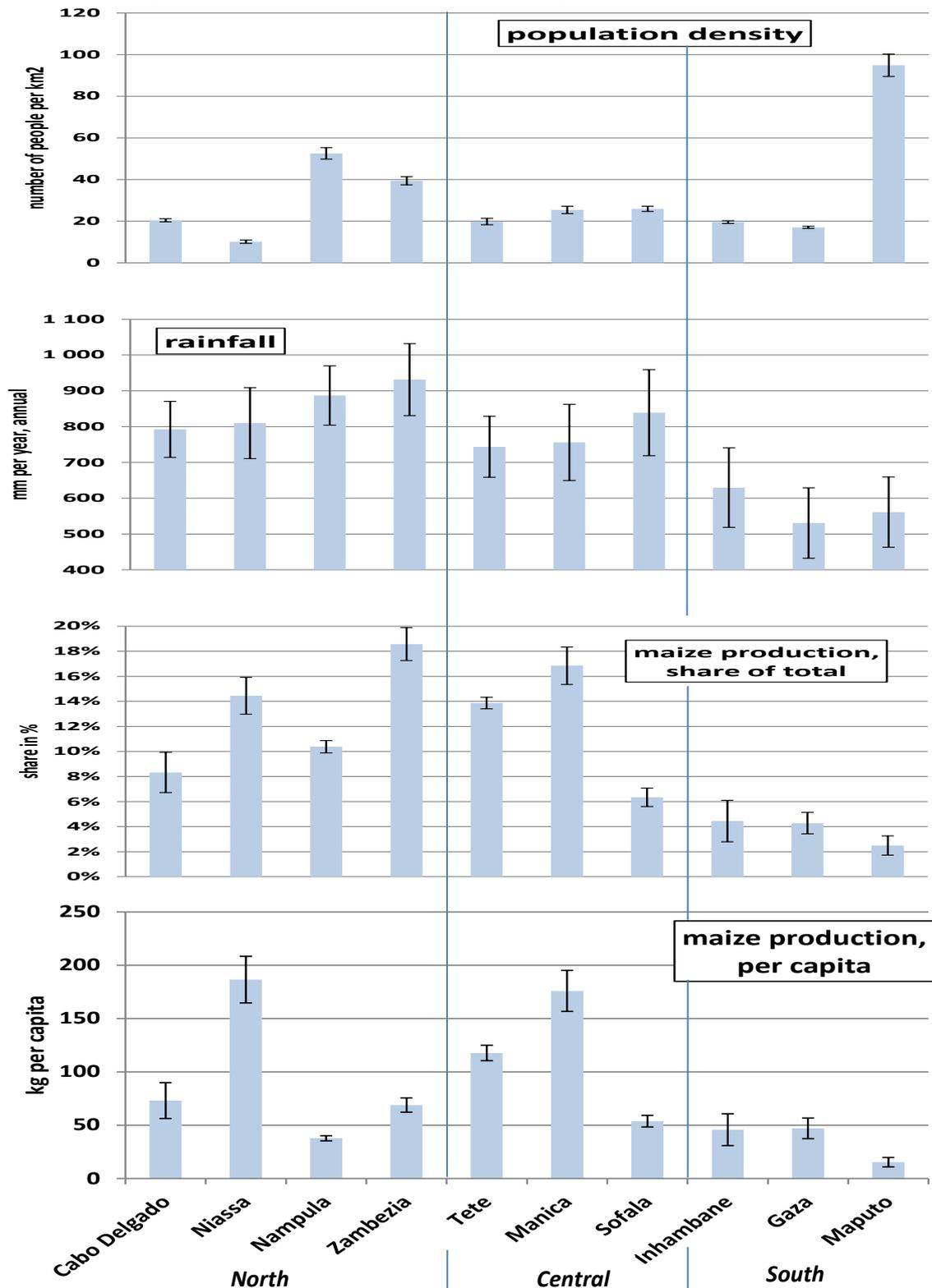
twice as high as directly after harvesting. Seasonality also shows up in spatial price differences (see Figure 5): these peak from April to June.

Figure 1 Mozambique: markets and roads



Source: VU-SPINlab

Figure 2 Population density, rainfall and maize production by province



Source: (author's calculations based on data from) Instituto Nacional de Estatística Moçambique, FEWSNET and Ministry of Agriculture, Early Warning Unit (Aviso Previo); The figure is based on aggregate (average) annual province data. Error bars show 95% confidence intervals. See the maps in the appendix for the location of provinces.

Note: population data are from 2005 to 2014, rainfall from 1996 to 2012 and production from 2005 to 2012.

Figure 3 Mozambique maize prices, by region, Jan 2005-Dec 2014 (meticals per kg)

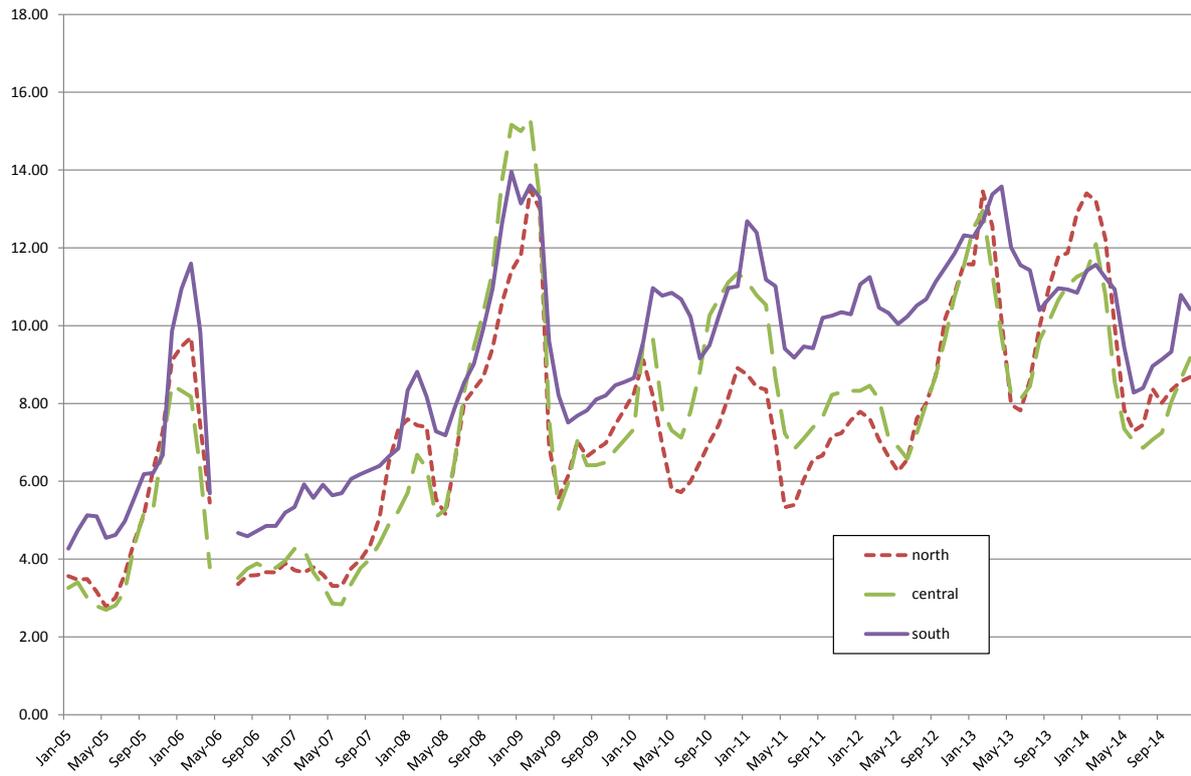
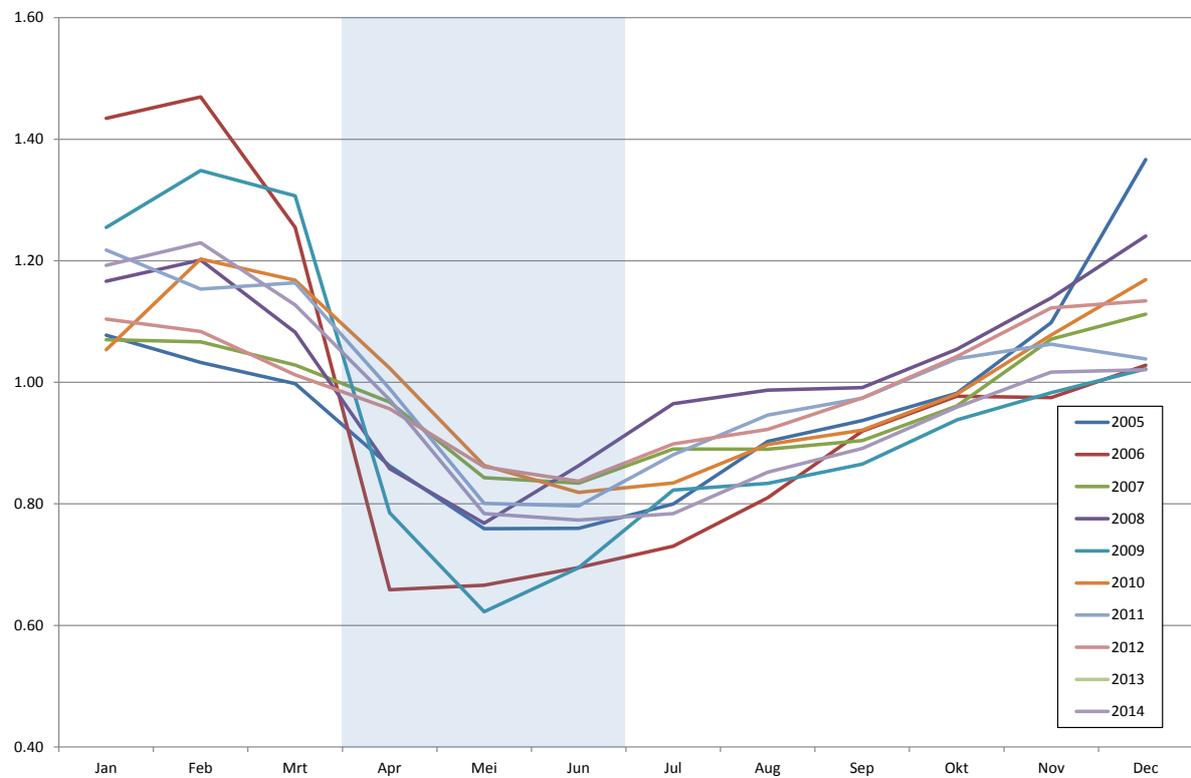
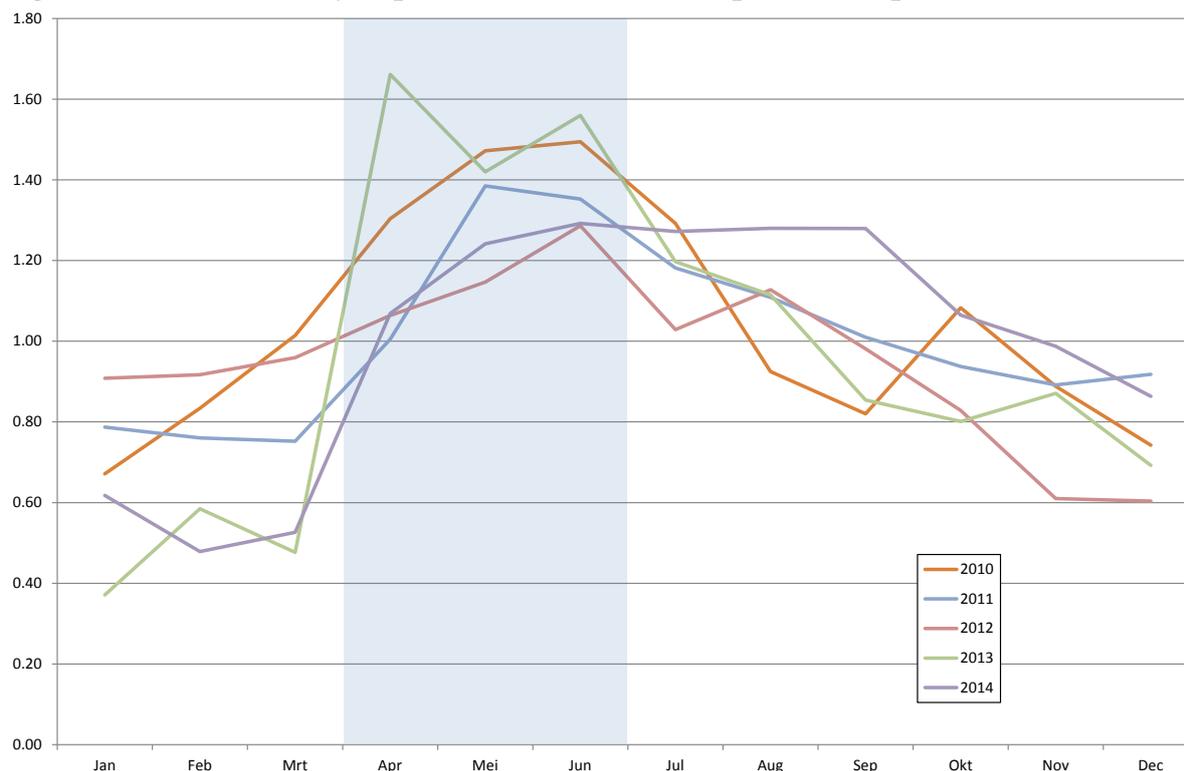


Figure 4 Seasonality in maize prices (index)



Source: authors calculations based on SIMA data; the shaded area indicates the months when trade is most likely.

Figure 5 Seasonality in price difference with Maputo maize prices (index)



Source: authors calculations based on SIMA data; the shaded area indicates the months when trade is most likely.

Road network and Transport Services

Maize is mainly transported by trucks. The road infrastructure is, however, not well developed, and especially during the rainy season many roads cannot be used. The Zambezi river creates a natural barrier to domestic trade: major domestic trade flows of maize are from the central area to the south, since Southern Mozambique and most notably the Maputo-Matola area is the major maize deficit area, while the coastal cities north of the Zambezi are supplied by the more inland production centers in the north. Angonia in the northeast, also a major production, assembly and wholesale market, supplies Tete, while maize surpluses in both Angonia and Lichinga region are exported to Malawi (USGS / FEWS NET) where prices often are higher (see Appendix, Figure A3). Maize available for sale in wholesale markets in Maputo (Xiquelene and others) is, amongst other locations, sourced from the central region, Nhamatanda, Chimoio or Manica, around 1100 km by road (Abdula, 2005;

SIMA data from 1999-2001), or even as far as Tete, around 1500km by road from Maputo (Tostão and Brorsen, 2005; SIMA trade flow data from 1998-2001). Transport cost in SIMA data pertain to itineraries as far as the distance between Lichinga and Maputo (by road around 2300km!). Southern Mozambique, and the Maputo-Matola area in particular, also rely on South Africa as a major supplier of maize (see Haggblade et al., 2008), where prices often are lower (see Appendix, Figure A2).

The Zambezi river and the bridges at Caia and Mutarara

The Zambezi is a large river, that originates in the outer north of Zambia, near the Congo border, runs through Angola, re-enters Zambia, than forms the border between Zambia and Namibia (partly), and Zambia and Zimbabwe, enters Mozambique in the east at the Cahora Bassa reservoir, and runs subsequently southwest to empty in the Strait of Mozambique (Indian Ocean). In the world the Zambezi ranks 31 in terms of length, 16 in terms of drainage area and 37 in terms of average discharge. Relative to the river Rhine in Europe, the Zambezi is more than twice as large in terms of length (2693km/1233km) and average discharge (4880m³/s/2330 m³/s) and 6.7 times as large in terms of the drainage area (1330000km²/198735 km²) (source: Wikipedia). The Zambezi is the fourth largest river in Africa, on all accounts. The within Mozambique length of the Zambezi is, however, only around 800km, nearly a third of its total length. Also the width of the river is substantial: in broad valleys the river spreads out over a wide area and is 5km to 8km wide and, not unimportant for transport by river, shallow in many places.

How does the Zambezi river affect the economy of Mozambique? A key contribution of the Zambezi river is power generation at the Cahora Bassa hydropower plant (completed in 1973), that supplies power to both Mozambique and South Africa. However, in terms of transport the river offers much less. Although the river is navigable, commercial long-distance transport by river is not well-developed, mainly due to its unreliability. As indicated

above, since crossing the river is a major barrier, the Zambezi river does have a negative impact on north-south trade and transport by road, within Mozambique.

This brings us to the, currently, four bridges that cross the Zambezi. The bridge at Tete is a suspension bridge that is in operation since 1973. It is integrated in the highway network and forms a major gateway to Zambia, Malawi and the northern part of Mozambique on the northern side, and to Zimbabwe, South-Africa and the southern part of Mozambique, on the southern side. Since the bridge at Tete for a long time has been the only way to cross the Zambezi, the bridge has become a major road transport bottleneck often plagued by congestion. In the course of the past decade an additional bridge in Tete has been constructed, which was opened for traffic in 2014.

On August 1, 2009 a new road bridge over the Zambezi River was opened², between Caia and Chimuara, linking Sofala and Zambezia provinces in the centre of the country. The bridge is part of the main north-south highway. Tete, where the other road bridges over the Zambezi are located, is around 300km kilometers to the northwest. The construction of the bridge began in March 2006. Already in 1979/80 work began on the access roads but war and conflict caused major delays in completing the work. The new bridge at Caia has replaced a ferry service between Caia and Chimuara that operated until the completion of the bridge. The ferry was widely perceived as inefficient due to long waiting times and extensive queues of trucks, causing high risks of spoiling perishable crops³. The toll for the new bridge is the

² Mozambique has a tradition to tag (presidents) names to bridges. The bridge between Caia and Chimuara is named after president Armando Emilio Guebuza. The original bridge at Tete is known as the Samora Machel bridge and the bridge between Vila de Sena and Mutorara as the Dona Ana bridge. In this study we identify bridges by the nearby town or village.

³ Trucks transport often encountered delays for days or weeks for the ferry trip between Caia and Chimuara. Engine breakdowns and sensitivity to tides made the ferry connection unreliable. Tostão and Brorsen (2005) report: "...in early 2001 the ferry was shut down for nearly two months because the Zambezi river was flooding, and in July 2001 the ferry service was interrupted again because there was not enough water in the river".

same as motorists had to pay for using the ferry – 800 meticaais (around US\$30 dollars) for trucks and 80 meticaais for light vehicles⁴.

Only around 60 km upstream, there is another bridge, spanning the lower Zambezi River between the towns of Vila de Sena and Mutarara. This bridge was originally built by the Portuguese in 1934 during the Portuguese rule of Mozambique and constructed as a railway bridge to link Malawi and the Moatize coal fields to the port of Beira. At the time, this 3.7 kilometers long Vila de Sena – Mutarara bridge was the longest railway bridge in Africa. Although not located on a primary highway, it provided an alternative route to pass the Zambezi river, next to the bridge at Tete and the former ferry between Caia and Chimuarara. Hence, the Vila de Sena – Mutarara bridge used to be the last downstream bridge before the completion of the bridge at Caia in 2009. The bridge was rendered unusable in the 1980s, during the Mozambican Civil War. After the ending of the civil war in 1992, USAID assisted with the repairs and it was converted to a single-lane bridge for vehicle traffic. However, more recently in October 2006 the bridge was completely closed to vehicular traffic for rehabilitation and (re-)conversion into a rail bridge, and was reopened as a rail bridge in August 2009.

2. Theory, Data and Empirical Strategy

Underlying theory

We exploit the Law of One Price which states that differences in prices of the same commodity at different locations should be equal to transaction costs⁵. Under spatial price equilibrium (see for example Fackler and Goodwin, 2001, and Anderson and Wincoop, 2004) spatial price dispersion, i.e. the price difference across markets, is defined as the sum of

⁴ These are 2009 tariffs.

⁵ Throughout this paper we make the following approximation: transaction costs \approx trade costs \approx transport costs, and use these concepts interchangeable.

transaction costs. Recently developed general equilibrium trade models (Eaton and Kortum, 2002) also predict that for goods that are traded between regions, price differences between regions can be used to measure trade costs. Equality of price difference between locations and transaction costs is established by arbitrage between locations (the so-called no-arbitrage condition). Empirical studies on the maize market in Mozambique confirm spatial efficiency (see Tostão and Brorsen, 2005; Cirera and Arndt, 2008).

In the spatial price literature (see for example Fackler and Goodwin, 2001) it is standard to distinguish three regimes: I transaction costs are larger than the price difference between markets ($tc_{jk} > p_j - p_k$), there is no trade since this is not profitable and markets are autarkic and not integrated. Under regime II transaction costs are close or equal to the price difference between markets ($tc_{jk} \leq p_j - p_k$), there are positive trade flows and markets are integrated. Both regime I and II are labelled as spatial price equilibria. Only under regime II the no-arbitrage condition applies. Finally, under regime III transaction costs are well below the price difference between markets ($tc_{jk} \ll p_j - p_k$), but there is no trade. This regime is characterized by profitable but unexploited trade opportunities. Markets in regime III are not integrated and also spatially inefficient. Regime I points at transaction costs as the major cause of the lack of integration between markets, while in regime III the causes of the lack of integration are not related to transaction costs. Various suggestions are made in the literature for the regime III spatial inefficiency: lack of information, barriers to entry, market power or trade quota's.

The Parity Bound model attempts to measure the size of the respective regimes by estimating the probability of each regime under the assumption of a (truncated) normal distribution of the expression $(p_j - p_k) - \widehat{tc}_{jk}$ where \widehat{tc}_{jk} represents a prediction of transaction costs, often a simple extrapolation based on a small number of observations (see

Baulch 1997a, b; Tostão and Brorsen, 2005; Cirera and Arndt, 2008). The question arises why maximization of a likelihood function that is based on the assumption of a (truncated) normal distribution, yields economically meaningful estimates of the different regimes. Hence, a major drawback of this empirical strategy is the lack of economics that drives the distinction between regimes. In this study we overcome this drawback by exploiting determinants of transaction costs (road distance, road quality, crossing borders, fuel prices and wages of truck drivers) to find the size of the distinctive regimes or, equivalently, the Parity Bounds. It should be noted that it is merely by virtue of the introduction of the bridges and the resulting variation in road distance, road quality and crossing borders, that we are able to implement this strategy.

Data and data sources

Market prices for maize are sourced from Sistema de Informação de Mercados Agrícolas de Moçambique (SIMA; www.masa.gov.mz/sima), from the weekly publication *Quente-Quente*⁶, in particular the weekly quotations of white maize retail prices⁷, and are originally recorded for 27 markets⁸. We use monthly prices for the period from January 2005 to December 2014. This period covers a timespan of up to five years without and with bridges⁹. The price data are collected by interviewing randomly selected retailers in each market. Unfortunately, there are missings in the monthly price data. A few markets of the original data (Angoche, Monapo and Vilanculos) are dropped altogether since these markets have no data

⁶ SIMA information is also available in newspapers.

⁷ Quadro 3, preço e mudança percentual a nível de mercado retalhista (MT/kg), grão de milho branco (Table 3, Prices and percentage price changes in retail markets (meticals per kg), white maize grain)

⁸ Alto Molocue, Angoche, Angonia, Beira, Chimoio, Chokwe, Cuamba, Gorongosa, Lichinga, Manica, Maputo, Massinga, Maxixe, Milange, Mocuba, Monapo, Montepuez, Mutarara, Nacala, Nampula, Nhamatanda, Pemba, Quelimane, Ribaue, Tete, Vilanculos en Xai-Xai. Figure 1 shows the locations of these markets in Mozambique.

⁹ It is not exactly clear to what extent the Vila de Sena - Mutarara bridge was effectively used as a road bridge, before it was closed for traffic to be rehabilitated as a railway bridge in October 2006 (see descriptive section). This bridge is, however, not integrated in the major trunk roads of Mozambique. Whatever is the case, the “closing for rehabilitation date” suggests a clear period of being closed to traffic of 34 months before the re-opening in August 2009. In the estimation we have experimented with several time spans before and after, that take these dates into account.

in the period under consideration. After dropping these markets, we have around 88% of the potential number of monthly price observations, in total slightly more than 2500 ($2533/(24 \times 12 \times 10)$). The town of Quelimane has less than 40% of the total observations, mainly in the months after the bridge introduction. Finally, maize prices data of a few months are totally missing (May and June 2006)¹⁰. The construction of spatial price differences – the price dispersion measure that is used in the estimations – blows up the number of available observations to large numbers (around 25,000)¹¹.

Data on transports costs are from the same source (SIMA). Collection of these data is organized similarly to the collection of price data, by asking quotations from randomly selected wholesale traders in major source and destination markets. Transport costs are specified by product¹² and by the weight of the bags transported. Transport costs are recorded for the period from August 2001 to December 2010, with nearly three quarter of the observations before 2005. After 2010 the publication of these series stops. Moreover, the transport costs series are also only available for a limited number of itineraries¹³. In all we have only around 800 observations of transport cost data, that cover the period of study (January 2005 -December 2014) only to a limited extent.

Road distance and traveling time are both taken from Google Maps, at the time of writing of this study (2017). Since our study period is from January 2005 to December 2014, this may entail measurement error: we do not incorporate road construction, road maintenance and rehabilitation that have taken place in the road infrastructure (other than changes do to the introduction and rehabilitation of the bridges), or alternatively, we have assumed that these changes are small and negligible relative to the change realized by the

¹⁰ SIMA staff reports that missings are recorded for markets and periods if no maize transactions have taken place.

¹¹ Price data for n markets in a specific month, yields n^2 market pair data, of which $(n^2-n)/2$ are economically relevant; hence, with 24 markets a month without missings yields 276 market pair data $((24^2-24)/2)$.

¹² Differences in volumes for different crops explains separately recorded per kg transport costs.

¹³ See the Appendix Table A2 and A3 for information on the availability of transport costs by itinerary and year.

bridges. Note that some developments in road infrastructure are controlled for through the inclusion of trends.

A few other variables are used, in particular in the estimations with covariates and PSM estimations: population data, both by province and by city/town, are from three censuses (1997, 2007, 2016) from the Instituto Nacional de Estatística Moçambique. Population data for intermediate months and years are constructed by interpolation. Fuel prices are from the International Financial Statistics from the IMF. Maize production by province is from Trabalho de Inquérito Agrícola / Anuário de Estatística Agrária, Ministry of Agriculture. Maize production data are incomplete: missing years are constructed (see Appendix, Table A1). We have used quotations of the nearest SAFEX white maize futures contract as representative for maize prices in South Africa. These series are taken from the SAFEX web site and converted to Mozambique meticaís with the help of IMF/IFS monthly exchange rates (period average, domestic currency per US\$). Finally, we use market prices of maize grain from Lilongwe, Malawi, taken from FAO.

Empirical strategy

We have assumed that (nominal) transport costs are mainly determined by road distance and road quality, while fuel prices and wages contribute to nominal increases. Moreover, various of the itineraries between markets, especially itineraries from the northeast to the center and south and back, run through Malawi. Since crossing borders is likely to increase costs, it appears useful to control for the impact of crossing the Mozambique–Malawi borders on transport costs¹⁴. In summary, in order to measure the impact of the Cena and Mutarara

¹⁴ These border costs are pass-through border costs (meaning passing the border two times rather than one time), and, consequently, border crossing does not entail international trade related costs like exchange rate costs and non-tariff barriers. The border costs that are identified in this study are thereby different from those documented in the literature (see e.g. Morshed, 2003, 2007; Gopinath et al., 2011 and Azam, 2007). Including exchange rate variability or proxies for non-tariff barriers or non-traded services in order to control for border effects in the estimations is, hence, without meaning.

bridges on the determinants of transport costs, we start with the following difference-in-difference specification:

$$(1) \quad y_{jkt} = \beta_0 + \beta_1 \text{roaddistance}_{jkt} + \beta_2 \text{roadquality}_{jkt} + \beta_3 \text{border}_{jkt} + \\ + \beta_4 \text{fuel price}_t + \beta_5 \text{consumer price}_t + \\ \mathbf{X}_{jkt} \gamma + \eta_{jk} + \theta_t + \varepsilon_{jkt}$$

In this equation $\text{roaddistance}_{jkt}$ is the shortest road distance between markets. We expect that an increase in road distance increases transport costs, and hence $\beta_1 > 0$. The variable roadquality is specified as the number of kilometres realised per hour of travel time: a high value indicates a high road quality, a low value a low road quality. We expect that increased road quality decreases transport costs, and hence $\beta_2 < 0$. Costs of crossing the Malawi-Mozambique border enter as lump sum costs that can be parametrized by including a binary variable. The variable border is a binary variable that takes the value 1 if the shortest itinerary crosses the Malawi-Mozambique border, and zero elsewhere. We expect that crossing the Malawi-Mozambique border increases transport costs, and hence $\beta_3 > 0$. These three variables – road distance, road quality and border crossings – are the real determinants of transaction costs (rather than the nominal determinants). The remaining two – fuel prices and consumer prices – are the nominal determinants. The variable fuelprice is an index of diesel or gasoline pump prices and these are expected to be positively correlated with transaction costs, and hence $\beta_4 > 0$, and the same applies to the variable consumerprice (hence $\beta_5 > 0$), which is the national consumer price index and is included to approximate, amongst other things wage cost of truck drivers etc. The vector \mathbf{X}_{jt} represents variables that influence price dispersion, such as rainfall in source areas, population densities, foreign maize prices, distance to seaports, access to imports, domestic transport by sea, etc. Parameters η_{jk} and θ_t are market pair and time fixed effects, and ε_{jkt} is an error term with zero mean and constant

variance. The parameters of interest are β_1 , β_2 and β_3 which measure the contribution of road distance, road quality and border crossings to transport costs.

Further empirical issues: measuring optimal road distance

The key variable that governs impact is the minimum road distance between markets. We have assumed that prior to August 2009 – the month when the bridges became operational – any north-south road transport crossing the Zambezi (and vice versa) was directed through Tete and crosses the Zambezi via the bridge in Tete. This rule unambiguously determines the optimal transport routes for north-south trade in maize in Mozambique.

In practise this implies, for example, that maize transport from Alto Molocue, north of the Zambezi, to Beira, south of the Zambezi (see Figure 1) involved a 1268km (18 hours, 39 minutes) journey before, and a 749km (11 hours, 5 minutes) journey after the introduction of the bridge, a decrease of more than 40% in both road distance and travel time. Likewise, maize transport from Chimoio (south of the Zambezi) to Quelimane (north of the Zambezi) involved a 1005km journey before and a 564km journey after the introduction of the bridge, a decrease of more than 44% in road distance and 42% in travel time.

By implication all freight from north to south (or vice versa) that is transported either through the Vila de Sena-Mutarara bridge – before March 2006, when it operated as a road bridge – or through the ferry that operated between Caia and Chimuará, is assumed to be of negligible size and have a negligible impact on maize markets, maize market prices and geographical price dispersion. Moreover, we ignore road construction, and improvements in the road network by rehabilitation and maintenance of roads, in the course of our sample period¹⁵. This approximation is partly justified by the relatively short period of study. Also, to a certain degree these changes are accounted for by including trends. However, more

¹⁵ Road rehabilitation and maintenance is likely to affect road quality (=kilometers travelled per hour travel time). The employed identification strategy jointly creates independent variation in road distance and road quality. This makes it possible to quantify cost reductions due to road rehabilitation and maintenance, relative to cost reductions due to changes in road distance.

generally, we assume these changes to be of negligible size relative to the introduction of the bridges.

Further empirical issues: trends and seasonality in maize prices

Agricultural crop prices are notorious for seasonality, and maize prices in Mozambique are no exception: these prices start to rise around September and reach a maximum from January to March (see Figure 4). Prices in the lean season have values that are twice as high compared to the post-harvesting months. This corresponds with observed seasonality in staple food prices in other sub-Saharan countries (see Kaminski et al., 2016; Zant, 2017). Moreover, price developments over time reflect the rain-fed character of agriculture with occasional peaks in prices due to drought. If seasonality is important for prices development, it is likely to affect price dispersion as well. This is further supported by the seasonality in prices relative to Maputo prices (see Figure 5): this figure clearly suggest that the largest price differences are realized from April to July. To accommodate for the strong seasonality in maize prices, we have included source and destination specific monthly dummies in the estimations.

The base specification includes indices of national fuel prices and national consumer prices. These variables capture some trend developments but are incapable to account for trend developments that differ by source market or by destination market. Such trends may reflect developments in road infrastructure, transportation technology and networks that are different between markets and regions. Therefore we have included source and destination specific trends to control for geographically diverging price and quality developments. The empirical estimations indicate that the contribution of seasonality and trends to the variation in price dispersion is large (see below).

Further empirical issues: distribution of information

The no-arbitrage condition is more plausible if information dissemination, particularly information on prices, is not costly. In the case of the Mozambique maize market, SIMA maize price data, used for the empirical estimations in this study (see below for further details), are widely disseminated throughout the country, on a frequent basis. Dissemination of this price information is organized through weekly published bulletins and through newspapers (which was in place well before the sample period of this study). In this context we further note that mobile phones have considerably improved market transparency and availability of price information: to a much larger extent than traditional sources of information, mobile phones make information feasible that is completely customized to individual needs. The roll-out of the mobile phone network in Mozambique has taken place from 1997 to 2009, and all major markets identified in the current study had access to mobile phone services, already in 2006 (see Zant, 2017).

3. Empirical Estimation, Covariates and Robustness Checks

Estimation of a basic specification

For the empirical estimations we are keen to filter out observations that fall under regime III and I, and maintain observations under regime II (see *Underlying theory*¹⁶). However, we do not observe maize trade flows – the distinctive feature of regime II – and therefore cannot verify if trade actually takes place. Instead we have assumed, as a start, that a positive price difference between two markets pins down source and destination market, since profitable trade requires destination markets need to have higher prices. This decision rule is, obviously,

¹⁶ Regime I: transaction costs are larger than the price difference between markets ($tc_{jk} > p_j - p_k$), there is no trade (since this is not profitable) and markets are autarkic and not integrated. Under regime II transaction costs are close or equal to the price difference between markets ($tc_{jk} \leq p_j - p_k$), there are positive trade flows and markets are integrated. Finally, under regime III transaction costs are well the price difference between markets ($tc_{jk} \ll p_j - p_k$), there is no trade, and there is a situation of profitable but unexploited trade opportunities.

too broad and thereby not accurate: the price difference may be too small to justify profitable trade or even when the price difference is sufficient for profitable trade there could be other barriers to traded. Hence, we need to restrict the sample of price differences to regime II observations.

Restricting the sample of price differences to regime II observations is implemented by constructing trade costs for each market pair and each month – as the first step – and subsequently selecting those observations of price differences for which the price difference is approximately equal to the constructed trade costs ($p_j - p_k \approx \widehat{tc}_{jk}$), which is the second step. In the first step we construct predictions of transport costs for all market pairs and months on the basis of observed transport costs. The available data on transport costs are limited: we have around 800 monthly observations, unevenly spread over the period 2001-2010 (see data section for details), a number that is relatively small compared to the number of observations on price differences (around 25,000). Moreover and unfortunately, the transport cost data only marginally overlap with the period of study (January 2005-December 2014). Nevertheless, on the basis of this limited set of transport cost data we estimate essentially the same relationship as set out in equation (1). There are, however, a few differences: since we have less observations, and thereby less power, we have to economize on explanatory variables. Consequently, we include simple seasonality variables and a simple trend variable (rather than including seasonality and trend variables by source and destination) and we omit month-year fixed effects (since these are not helpful in predictions), assuming that part of this variation will be captured by fuel prices and the national consumer price index¹⁷. Estimation results are documented in the Appendix, Table A4. The out-of-sample predictions of this estimation generate transport cost estimates (\widehat{tc}_{jk}) for close to 200

¹⁷ Estimations of price differences confirm that a considerable part of the month-year fixed effects is captured by national fuel prices and the national consumer price index.

market pairs. Predictions for the remaining market pairs are obtained by multiplying road distance with the average predicted per km transport cost by source and month.

With estimates of transport costs for all market pairs and months, we calculate $(p_j - p_k) - \widehat{tc}_{jk}$. Following the prescriptions of spatial price theory the value of this metric should be zero if markets are integrated. Hence, the second step is to evaluate this metric $(p_j - p_k) - \widehat{tc}_{jk}$ in order to establish which observations satisfy this condition, or in other words, to determine the parity bounds. We implement this by conditioning the estimations to those with statistically significant coefficients for each of these cost determinants, proposed in equation (1). We estimate equation (1), set out in the previous section without covariates (X_{jkt}), restricting the sample to observations around $(p_j - p_k) - \widehat{tc}_{jk} = 0$, using a grid of varying values above and below 0.

Table 1 Impact of distance, road quality and borders: seasonality and trends

Dependent variable: $\ln(p_i - p_k)$ if $p_i > p_k$					
Variables	sample period: 34 months before, 34 months after ^a				
ln(road distance)	0.982*** (0.197)	1.007*** (0.173)	0.919*** (0.216)	0.884*** (0.198)	
ln(road quality)	-3.554*** (0.402)	-4.207*** (0.387)	-2.085*** (0.510)	-2.189*** (0.509)	
border	-0.176*** (0.047)	-0.185*** (0.045)	-0.157*** (0.051)	-0.116** (0.049)	
ln(fuel price)	0.700*** (0.193)	0.462** (0.213)	0.828*** (0.199)	0.604*** (0.223)	
ln(consumer prices)	0.920*** (0.105)	1.019*** (0.112)	2.758*** (0.449)	2.209*** (0.501)	
seasonality at source	no	yes	no	yes	
seasonality at destination	no	yes	no	yes	
trend at source	no	no	yes	yes	
trend at destination	no	no	yes	yes	
adj R ²	0.3197	0.3685	0.3322	0.3827	
observations	11418	11418	11418	11418	
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara.

The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

All equations are estimated with OLS and include market pair fixed effects. In all estimations observations with extreme price differences (around 3% of all observations) and extreme road distances (>2500km) are omitted. Robust standard errors, clustered by market pair, are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Estimation output of this basic specification is reported in Table 1. Table 1 shows several interesting outcomes. In the first place the estimations confirm the robustness of the

coefficients of road distance and road quality: these coefficients are statistically significant and have the expected signs irrespective of the inclusion of seasonality and trends. Secondly, the impact of border crossing on transport costs is negative and statistically significant suggesting that travelling through Malawi does not add additional transport cost but rather reduces these costs. Thirdly, also coefficients of fuel price and consumer prices are statistically significant and have the expected sign. Fourthly, the estimations show that both seasonality and trends contribute substantially to the explanation of the variation in price dispersion, improving adjusted R^2 with more than 5%-points. Including trends clearly affects coefficients of road quality (and also consumer price index). Finally, we find that the contribution of our impact variables in the explanation of the total variation in maize price dispersion is small relative to fixed effects, seasonality and trends¹⁸.

For the estimations in Table 1 we have started our sample in October 2006 (the month in which the Vila de Sena - Mutarara bridge was closed for traffic to be rehabilitated) and chosen an equal number of months before and after the introduction of the new bridge and the completion of the rehabilitation of the old bridge. Hence, we aimed at a “clean” strategy by including a period “without bridge facilities” jointly with a period “with bridge facilities” of an equal number of months¹⁹. However, forward (and possibly also backward) extension of the sample period is arbitrary: to accommodate for these arbitrary decisions we have re-run estimations with varying sample periods. A selection of estimations is reported in Table 2.

¹⁸ This can be verified by estimating the specification in column 4 of Table 1 without the impact variables.

¹⁹ The number of months before and after is equal but, due to differences in missings, the number of observations is not necessarily equal.

Table 2 Impact of distance, road quality and borders: varying months before and after

Dependent variable: $\ln(p_i - p_k)$ if $p_i > p_k$				
Variables	65 months before, 65 months after ^a	65 months before, 34 months after ^a	34 months before, 65 months after ^a	34 months before, 34 months after ^a
ln(road distance)	0.843*** (0.175)	0.918*** (0.179)	0.882*** (0.181)	0.884*** (0.198)
ln(road quality)	-1.498*** (0.465)	-1.298*** (0.466)	-2.894*** (0.427)	-2.189*** (0.509)
border (binary)	-0.215*** (0.045)	-0.104** (0.046)	-0.233*** (0.045)	-0.116** (0.049)
ln(fuel price)	0.244** (0.123)	0.253* (0.132)	0.693*** (0.198)	0.604*** (0.223)
ln(consumer prices)	1.896*** (0.268)	2.067*** (0.452)	2.442*** (0.314)	2.209*** (0.501)
adj R ²	0.3449	0.3648	0.3560	0.3827
observations	18614	14060	15972	11418
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0
regime I	13.5%	12.7%	13.8%	12.9%
regime II	73.0%	74.9%	71.8%	73.6%
regime III	13.5%	12.4%	14.4%	13.5%

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara.

The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

All equations are estimated with OLS and include market pair fixed effects, and seasonality and trend, by source and destination. In all estimations observations with extreme price differences (around 3% of all observations) and extreme road distances (>2500km) are omitted. Robust standard errors, clustered by market pair, are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The estimation results reported in Table 2 confirm estimations in previous table: coefficients of distance, road quality, fuel price and consumer price are all statistically significant and have the expected sign. Coefficients of road distance and consumer price are relatively stable, while coefficients of both road quality and fuel price appear to increase over time. The impact of border crossing on transport costs is again statistically significant and negative, in all estimations.

Estimations with covariates

Following the specification formulated in equation (1), we now include X variables into the basic specification and investigate if the previous estimation results are robust for the inclusion of covariates. We employ covariates reflecting developments on the demand side, on the supply side, associated with trade and associated with imports or exports of food. Developments on the demand side are approximated with population. Data on population by market are constructed as interpolations of census data of population by market, which are available for a few designated years. We expect a larger population in destination market, ceteris paribus, to increase prices in destination markets and thereby to increase price

differences between source and destination markets. For population in source markets we expect the reverse.

Data on developments on the supply side are less easy and straightforward: we employ rainfall by district of source markets, in the previous production season, and per capita production by province of source markets, of the previous production season. Both variables have their drawbacks: rainfall by district may be a reasonable predictor of rain-fed maize production but is less suitable to indicate maize available for trade to other markets and regions. Per capita production more adequately reflects the availability for trade, but maize production data are only available by province and for a limited set of years²⁰. We expect higher per capita maize production in source markets, *ceteris paribus*, to decrease prices in source markets and thereby, *ceteris paribus*, to increase price differences between source and destination markets.

We have further included foreign prices, since prices in neighbouring countries, but especially in South Africa and Malawi. Note that South Africa and Malawi play a different role: South Africa is a major source of maize grain imports, while Malawi is a major destination of maize grain exports (see Zavale, 2014)). In the case of South African prices we have used quotations of the nearest SAFEX white maize contract²¹. Specifically we have calculated monthly averages of daily quotations and converted these to Mozambique meticaïs. The influence of foreign prices is ambiguous: it depends on whether maize is imported from foreign markets and foreign prices are lower (South-Africa), or maize is exported to foreign markets and foreign prices are higher (Malawi). Under the assumption of

²⁰ For the period of our study (2004-2014) annual maize production data are available for the years 2002-2003, 2005-2008 and 2012. The remaining years are predictions on the basis of a simple model with province and year fixed effects, province trends and rainfall shocks (see Appendix Table A1).

²¹ See Appendix, Figure A4 for the development of domestic maize prices vis-à-vis SAFEX white maize.

negligible import tariffs²², South African maize prices are often well below Maputo maize prices (see Appendix, Figure A2). In the southern Mozambique terminal markets these foreign prices are likely to exert a downward pressure on domestic maize prices. If this is the dominant situation we expect a negative impact SAFEX prices on price differences between markets. For Malawi maize prices the reverse applies (see Appendix, Figure A3): these prices are expected to exert an upward pressure on price differences between Mozambique markets.

Intensive trade between terminal markets is not very likely. However, there is an exception: most terminal markets are located along the coast (Pemba, Nacala, Quelimane, Beira, Massinga, Maxixe, Xai-Xai and Maputo) and this offers a low-cost alternative to transport freight by sea. Especially markets that are well connected with the inland surplus markets may have opportunities for profitable trading transactions with deficit terminal market on the coast in the south, Maputo in the first place. Likewise one could argue that distance of inland markets to the nearest seaport is important. This distance approximates the potential impact of access to other markets, either domestic or foreign. There is, however, a problem: we do not have an adequate identification strategy for impacts on transport costs through these channels and as a result all potential impacts are captured by the market pair fixed effects.

²² Maize imports are subject to a 2.5% import tariff and a 17% Value Added Tax, which is not levied on domestic production (see Zavale, 2014). These import duties, however, do not offset the price difference with Maputo (but this is different for other markets).

Table 3 Impact of distance, road quality and borders: including covariates

Dependent variable: $\ln(p_i - p_k)$ if $p_i > p_k$				
Variables	34 months before, 34 months after ^a			
ln(road distance)	0.876 ^{***} (0.198)	0.884 ^{***} (0.198)	0.870 ^{***} (0.196)	0.863 ^{***} (0.196)
ln(road quality)	-2.185 ^{***} (0.507)	-2.206 ^{***} (0.509)	-2.154 ^{***} (0.500)	-2.161 ^{***} (0.498)
border (binary)	-0.119 ^{**} (0.049)	-0.116 ^{**} (0.049)	-0.123 ^{**} (0.049)	-0.126 ^{**} (0.049)
ln(fuel price)	0.638 ^{***} (0.220)	0.655 ^{***} (0.218)	1.040 ^{***} (0.252)	1.097 ^{***} (0.245)
ln(consumer prices)	2.204 ^{***} (0.501)	2.255 ^{***} (0.507)	2.951 ^{***} (0.544)	2.984 ^{***} (0.552)
ln(population at source)	5.228 [*] (3.154)			4.872 (3.179)
ln(population at dest.)	-3.739 (9.519)			-2.451 (9.337)
ln(pcmaiz production _{t-1}) ^b		0.189 (0.123)		0.163 (0.121)
drought _{t-1} ^b		-0.007 (0.056)		-0.012 (0.058)
ln(SAFEX white maize)			-0.459 ^{***} (0.092)	-0.446 ^{***} (0.093)
ln(Lilongwe maizeprice)			-0.052 (0.059)	-0.046 (0.060)
adj R ²	0.3831	0.3829	0.3908	0.3862
observations	11418	11418	11418	11418
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0
Variables	65 months before, 65 months after ^a	65 months before, 34 months after ^a	34 months before, 65 months after ^a	34 months before, 34 months after ^a
ln(road distance)	0.916 ^{***} (0.173)	0.894 ^{***} (0.178)	0.919 ^{***} (0.192)	0.863 ^{***} (0.196)
ln(road quality)	-1.359 ^{***} (0.444)	-1.318 ^{***} (0.455)	-2.324 ^{***} (0.472)	-2.161 ^{***} (0.498)
border (binary)	-0.164 ^{***} (0.044)	-0.124 ^{***} (0.046)	-0.183 ^{***} (0.046)	-0.126 ^{**} (0.049)
ln(fuel price)	0.654 ^{***} (0.187)	0.530 ^{***} (0.194)	1.243 ^{***} (0.231)	1.097 ^{***} (0.245)
ln(consumer prices)	2.163 ^{***} (0.404)	3.060 ^{***} (0.507)	2.607 ^{***} (0.461)	2.984 ^{***} (0.552)
ln(population at source)	0.963 (0.667)	0.855 (0.714)	4.561 (3.088)	4.872 (3.179)
ln(population at dest.)	0.308 (1.112)	-0.054 (1.231)	1.558 (8.778)	-2.451 (9.337)
ln(pcmaiz production _{t-1}) ^b	-0.069 (0.086)	-0.054 (0.089)	0.177 (0.114)	0.163 (0.121)
drought _{t-1} ^b	0.084 [*] (0.049)	0.036 (0.051)	0.019 (0.051)	-0.012 (0.058)
ln(SAFEX white maize)	-0.217 ^{***} (0.057)	-0.235 ^{***} (0.057)	-0.449 ^{***} (0.092)	-0.446 ^{***} (0.093)
ln(Lilongwe maizeprice)	0.076 (0.049)	0.106 ^{**} (0.051)	-0.068 (0.056)	-0.046 (0.060)
adj R ²	0.3623	0.3677	0.3788	0.3862
observations	15041	14060	12399	11418
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara. The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

^b By source market.

All equations are estimated with OLS and include market pair fixed effects, and seasonality and trend, by source and destination. In all estimations observations with extreme price differences (around 3%) and extreme road distances (>2500km) are omitted. Robust standard errors, clustered by market pair, are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Estimations including covariates, reported in Table 3, further confirm the robustness of distance, road quality, fuel prices and consumer prices: coefficients are all statistically significance and, in most cases, of similar size as estimations without covariates. Also as before, coefficients of border are negative and significant (not strongly though). With respect to the covariates we find that population in source markets and, most importantly, SAFEX

prices are significant and have the expected sign. Most other covariates are not significant (but often do have the expected sign).

4. Potential threats and alternative explanations

OLS with a binary impact variable

Since the placement of the bridge is non-random, and, hence, since the selection of market pairs that realize a reduction in road distance, an improvement in road quality and a change in border crossing is also non-random, OLS regression comparing this selection with other market pairs is unlikely to consistently estimate the causal impact of the bridge on transport costs. In the spatial economics literature the major strategy to address this is to develop plausible instruments that meet the requirements²³. In empirical studies this has resulted in the so-called planned route IV, the historical route IV and the inconsequential place approach (Redding and Turner, 2014). Since we do not have data on planned or historical routes and inconsequential places are not common in the Mozambique trunk road network, these IV strategies cannot be implemented. Instead, we have implemented propensity score matching (PSM). To this end we need to convert the impact variables into one binary impact variable. Prior to discussing the details of the PSM estimation, we present OLS results of a standard difference-in-difference specification that includes this binary impact variable, for reference:

$$(2) \quad y_{jkt} = \beta_0 + \beta_1 \text{bridge}_{jkt} + (\beta_4 \text{fuel price}_t + \beta_5 \text{consumer price}_t +) \mathbf{X}_{jkt} \gamma + \eta_{jk} + (\theta_t) + \varepsilon_{jt}$$

where bridge_{jkt} is equal to 1 in period t if the shortest route from j to k runs via Caia-Chimuara bridge when this bridge was operational, and zero otherwise. Impact is now

²³ Instruments have to satisfy the exclusion restriction, meaning that the excluded exogenous variables – the instrument – is correlated with the change in infrastructure, but only affects transport costs through this channel.

expected to have a negative coefficient reflecting the reduction in transport costs, averaged over itineraries. Since fuel price and consumer prices are national prices, these variables capture to a certain extent and are, thereby, interchangeable with month-year time fixed effects (θ_t). Estimation results are reported in Table 4.

Table 4 Impact of bridges: OLS with a binary impact variable

Dependent variable: $\ln(p_i - p_k)$ if $p_i > p_k$				
Variables	sample period: 34 months before, 34 months after ^a			
bridge (binary)	-0.104 ^{**} (0.042)	-0.148 ^{***} (0.044)	-0.102 ^{**} (0.042)	-0.124 ^{***} (0.042)
ln(fuel price)	-	0.416 [*] (0.219)	-	0.856 ^{***} (0.234)
ln(consumer prices)	-	1.802 ^{***} (0.493)	-	2.907 ^{***} (0.528)
month-year (binary)	yes	no	yes	no
ln(population at source)			3.908 (2.979)	0.749 (2973)
ln(population at dest.)			-13.45 (9.057)	-10.80 (9.097)
ln(pcmaize production _{t-1}) ^b			0.180 (0.140)	0.028 (0.127)
drought _{t-1} ^b			-0.118 (0.047)	-0.072 (0.046)
ln(SAFEX white maize)			-	-0.488 ^{***} (0.084)
ln(Lilongwe maize price)			-	0.037 (0.059)
adj R ²	0.5034	0.4891	0.5047	0.4957
observations	10032	10032	10032	10032
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0
sample variables	65 months before, 65 months after ^a	65 months before, 34 months after ^a	34 months before, 65 months after ^a	34 months before, 34 months after ^a
bridge (binary)	-0.121 ^{***} (0.034)	-0.132 ^{***} (0.039)	-0.105 ^{***} (0.037)	-0.124 ^{***} (0.042)
ln(fuel price)	0.423 ^{**} (0.183)	0.313 [*] (0.183)	1.008 ^{***} (0.229)	0.856 ^{***} (0.234)
ln(consumer prices)	2.112 ^{***} (0.411)	3.201 ^{***} (0.481)	2.327 ^{***} (0.459)	2.907 ^{***} (0.528)
ln(population at source)	-0.540 (0.594)	-0.607 (0.657)	-0.177 (2.775)	0.749 (2973)
ln(population at dest.)	-1.051 (1.282)	-1.606 (1.292)	-4.861 (8.346)	-10.80 (9.097)
ln(pcmaize production _{t-1}) ^b	-0.082 (0.084)	-0.102 (0.086)	0.118 (0.127)	0.028 (0.127)
drought _{t-1} ^b	0.083 ^{**} (0.041)	0.044 (0.040)	-0.017 (0.046)	-0.072 (0.046)
ln(SAFEX white maize)	-0.141 ^{**} (0.056)	-0.168 ^{***} (0.056)	-0.485 ^{***} (0.084)	-0.488 ^{***} (0.084)
ln(Lilongwe maize price)	0.186 ^{***} (0.049)	0.222 ^{***} (0.060)	-0.000 (0.056)	0.037 (0.059)
adj R ²	0.4790	0.4884	0.4841	0.4957
observations	13033	12124	10941	10032
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara.

The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

^b By source market.

All equations are estimated with OLS and include market pair fixed effects. In all estimations observations with extreme price differences (around 3% of all observations) and extreme road distances (>2500km) are omitted. Robust standard errors, clustered by market pair, are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

In the upper panel of Table 4 estimations results are shown for 34 months before and after the introduction of the bridges, with and without covariates, either with fuel prices and consumer prices, or with month-year dummies. Coefficients of the price variables are

statistically significant, have the expected sign, but vary considerably. SAFEX white maize spot prices have a statistically significant negative coefficient. With month-year dummies the fit of the estimation is slightly better. Impact of bridges on transport costs is also statistically significant and varies from a reduction of 10 to 15%. It should be noted that the bridge coefficient reflects the impact of the bridges on transport costs, averaged over itineraries: this should correspond with the average treatment effect (ATE). In the lower panel of Table 4 estimations results are shown for varying sample periods and including covariates in all estimations (but without month-year dummies). Results are similar: impact of bridges on transport costs is statistically significant and varies from a reduction of 11 to 13%. Coefficients of the price variables are consistently significant, have the expected sign, but, again, vary considerably. Next to SAFEX white maize spot prices, also Lilongwe maize and last season drought have significant coefficients with the expected sign.

Propensity Score Matching estimation

The matching strategy builds on the Conditional Independence Assumption (CIA) meaning that the outcome in both treatment and control group, is independent of treatment assignment given the propensity score. In order to implement Propensity Score Matching estimations, we first estimate the propensity score i.e. the probability of treatment, hence, in this analysis the probability to make use of the Zambezi bridge. For this purpose we employ a logit model²⁴. We make extensive use of the covariates included in the previous estimation for the estimation of the propensity score. In particular we assume that the propensity score is determined by a trend, whether source and destination market are on different side of the Zambezi²⁵, population size in source and destination markets, per capita production in source and destination markets and SAFEX white maize spot prices. We also include trend

²⁴ Probit or logit are likely to give similar outcomes. However, the logit distribution has more density mass in the bounds and this corresponds with our empirical setting (see also Caliendo et al., 2005)

²⁵ Since a number of market pairs of which source and destination market are on either side of the Zambezi, have an optimal itinerary that crosses the Zambezi at Tete, this variable is not the same as the *bridge* variable.

interacted with the “across the Zambezi” dummy. We claim that the selected variables meet the requirements since they simultaneously influence assignment into treatment or control group and outcome, while assignment into treatment does not affect these variables. Results of the propensity score estimation are reported in the Appendix (Table A5). Coefficients of the covariates in the propensity score estimation have expected signs: positive for being located on opposite sides of the Zambezi, positive for population, negative for per capita maize production and negative for SAFEX spot prices. The pseudo R2 indicates how well variables explain the probability to make use of the Zambezi bridge and is thereby a formal test of the model. These statistics are comfortably high.

In order to match treatment and control observations, we use Kernel Matching as a matching algorithm. This is motivated by the availability of a large number of control observations. Kernel Matching is a non-parametric estimator that uses a weighted average of all control group observations to construct the counterfactual outcome. Weights depend on the distance between each observation from the control group and the treatment observation for which the counterfactual is estimated. Higher weights are placed on observations close in terms of propensity score and vv . As more information is used compared to, for example, Nearest Neighbour matching, Kernel Matching results in a lower variance, and, thus, higher precision estimates. Kernel Matching is also very time consuming since for each treatment observation an appropriate set of weighted controls is constructed. The Kernel function is the Epanechnikov kernel. Following accepted practise we use a bandwidth of 0.06^{26} .

We have tested the robustness of the matching algorithm by also implementing Nearest Neighbour (NN) as a matching algorithm, using 5, 10 and 50 of the nearest controls, with replacement, combined with a caliper threshold, where the caliper takes values 0.2, 0.5

²⁶ This bandwidth value is the default value in the STATA routine `psmatch2` (E. Leuven and B. Sianesi, 2003, ‘PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing’.)

and 1. Replacement is justified because the distribution of the propensity score is very different in the treatment and control group. The diverging distributions are also apparent from the common support figures shown in the Appendix (Figure A4). Restricting matches to those within the caliper threshold – a maximum distance of the propensity score of treatments and matched control observations – decreases the possibility of bad matches and hence bias. A problem is, however, that the literature does not give a clue which values for the tolerance level are appropriate. Further, ordering is done randomly since estimations with NN matching are dependent on the ordering of the data. For both matching algorithms the number of lost treatment observations is very small, in all cases less than 3% (see Table 5 and Table A6). Estimations with NN matching generate similar results (see Appendix, Table A6) as with Kernel Matching, with the notable difference that the ATEs are slightly higher (in absolute terms) and move further away from the OLS estimates. Nevertheless, the relative similarity of the estimations with different types of matching offers confidence about the robustness of the matching procedure.

Table 5 Impact of bridges: Propensity Score Matching (PSM), Kernel Matching

Dependent variable: $p_i - p_k$ (if $p_i > p_k$)				
	65 months before, 65 months after ^a	65 months before, 34 months after ^a	34 months before, 65 months after ^a	34 months before, 34 months after ^a
est. technique	PSM / KM	PSM / KM	PSM / KM	PSM / KM
bridge (ATT)	-0.422*** (0.090)	-0.166** (0.082)	-0.423*** (0.091)	-0.193** (0.084)
bridge (ATU)	-0.112	-0.160	-0.160	-0.193
bridge (ATE)	-0.165	-0.161	-0.212	-0.204
treated, on support	2128	1471	2128	1461
treated, off support	2	349	2	359
untreated, on support	10240	9078	8718	7702
untreated, off support	3913	4259	2737	2937
observations	16283	15157	13585	12459
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara. The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009. Equations are estimated with propensity score matching (PSM), using kernel matching as matching algorithm. Estimates of the propensity scores are in the Appendix (Table A5). Matching algorithm: Kernel Matching, Epanechnikov kernel and bandwidth 0.06. In all estimations observations with extreme price differences (around 3% of all observations) and extreme road distances (>2500km) are omitted. Standard errors are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

The overlap and region of common support between treatment and comparison group is shown graphically in the Appendix (see Figure A4). The cut-off is the straightforward and standard “minima and maxima criterion”: treatment (control) observations whose propensity score is higher than the maximum or less than the minimum propensity score of the controls (treatments) are dropped. The distribution of the propensity score in treatment and control are clearly extremely different. However, visual inspection of the figures confirm that the range of values of the matched propensity score have both treatment and control observations with probabilities between 0 and 1. Hence, the overlap condition is satisfied. In order to assess the quality of the matching procedure we use the standardised bias, before and after matching, as suggested by Rosenbaum and Rubin (1985)²⁷. The results of this exercise, reported in the Appendix (Table A7) indicate that matching on the estimated propensity score balances the covariates in the matched samples reasonably well.

Next, we consider the economics of the PSM impact estimates. The PSM results indicate an average treatment (ATE) that is somewhat larger than the ones estimated with OLS, but not extremely out of line (see Table 5 vis-à-vis Table 4). Further, we observe that including post bridge placement data, increase the average treatment of the treated (ATT). In all, we conclude that the PSM results further confirm the impact of bridge on transport costs. In the following section we assess if the estimated impact of road distance and road quality are in line with the estimated average treatment of the treated (ATT).

The diff-in-diff approach requires that that pre-intervention outcomes of intervention and control groups have a common trend. Following standard practise (see Autor, 2003) we test the common trend assumption by replacing the binary impact dummy with a set of dummies which are interactions of market pair dummies of those market pairs that benefit

²⁷ $B = \frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{(V_1(X) + V_0(X))/2}}$ where \bar{X}_1 (\bar{X}_0) and $V_1(X)$ ($V_0(X)$) are, respectively, the average and variance of covariate X in the treatment (control) group. The standardised bias, B, is calculated before and after matching, for each covariate X.

from the bridge, both before and after the bridge introduction, with time-period dummies (either annuals, semi-annuals, quarterlies or monthlies). The set of coefficients in the subsequent estimation support the common trend assumption if coefficients are statistically insignificant before and statistically significant (and negative) after the introduction of the bridge. The outcome of this exercise is shown in Figure A5. On the basis of the Figures we cannot reject the hypothesis of a common trend in the pre-treatment period for treated and not treated observations.

5. Discussion of outcomes, summary and conclusion

The estimation outcome achieved so far offers insight into the impact of bridges, averaged over all itineraries, and how the reduction in transport costs can be attributed to the key determinants of transport costs. These averages are interesting and useful²⁸ but less informative about realised cost reductions of particular itineraries that benefit from the new bridge. However, with the estimated elasticities we are now in the position to also measure benefits for specific itineraries: for a selection of itineraries this is shown in Table 6. Total reduction in transport costs due to the bridges range from 18% to 50%. In most instances the cost reduction is for the larger part due to the shorter distance: with a few exceptions (in particular itineraries to Nacala) change in quality mostly contributes only modestly to transport cost reduction.

Comparing the constructed reduction in trade costs (Table 6) with the Averaged Treatment Effect on the Treated (ATT) estimated with propensity score matching (Table 5, Table A6), suggests that constructed impacts, averaged over itineraries, will come close to estimated impacts (ATT), provided that a sample is used that extends up to five year after the

²⁸ To measure the benefit of the bridge for the Mozambique maize market, we ideally need to weigh reductions in transport cost per itinerary with the size of freight on transported through these routes. But again, we do not have these trade flow data.

introduction of the new bridge. This makes us confident that the OLS estimates explaining trade costs with road distance and road quality are not off the mark.

Table 6 Reduction in trade costs by itinerary due to the introduction of bridges

from north to south		distance	cost	quality	cost			
From	to	% Δ	% Δ	% Δ	% Δ	total	distance	quality
AltoMolocue	Maputo	-16.2%	-15.4%	1.4%	-2.3%	-17.7%	87.0%	13.0%
AltoMolocue	Beira	-40.9%	-39.3%	-0.6%	1.0%	-38.3%		
Mocuba	Maputo	-17.7%	-16.9%	2.4%	-3.9%	-20.8%	81.4%	18.6%
Mocuba	Beira	-48.1%	-46.3%	-4.4%	7.8%	-38.5%		
Ribaue	Maputo	-15.4%	-14.7%	2.5%	-3.9%	-18.6%	78.8%	21.2%
Ribaue	Beira	-37.4%	-35.9%	-0.1%	0.2%	-35.7%		
Nampula	Maputo	-14.8%	-14.1%	3.2%	-5.1%	-19.2%	73.6%	26.4%
Nampula	Beira	-35.2%	-33.8%	1.0%	-1.6%	-35.4%	95.5%	4.5%
from south to north								
Chimoio	Nampula	-18.5%	-17.6%	3.4%	-5.4%	-23.0%	76.6%	23.4%
Chimoio	Nacala	-17.1%	-16.3%	12.8%	-18.1%	-34.4%	47.3%	52.7%
Gorongosa	Nampula	-29.7%	-28.4%	8.5%	-12.7%	-41.1%	69.1%	30.9%
Gorongosa	Nacala	-26.8%	-25.6%	17.5%	-23.5%	-49.1%	52.2%	47.8%
Manica	Nampula	-14.5%	-13.8%	3.2%	-5.1%	-19.0%	72.9%	27.1%
Manica	Nacala	-7.1%	-6.7%	12.5%	-17.8%	-24.5%	27.4%	72.6%

Source: authors' calculations

In this study we have investigated the impact of bridges in Mozambique on transport costs. For this purpose we exploited the introduction of the bridge between Caia and Chimuará and the rehabilitation of the railway bridge between Vila de Sena and Mutarara. These events generated the required variation in optimal trading distances between markets, needed to attribute impact to road distance, road quality and crossing borders. Crossing borders appeared to have a negative impact on trade costs: travelling through Malawi does not add additional transport cost but rather reduces these costs. The key finding is that, averaged over all itineraries, bridges have caused a 9% to 15% reduction in transport costs. For specific itineraries this reduction in costs is much larger and varies from 18% to 50%. The larger part

of the cost reduction is due to the reduction in road distance. Results are robust for inclusion of covariates and for controlling for non-random assignment of road placement.

References

- Abdula, D.C., 2005, 'Improving Maize Marketing And Trade Policies To Promote Household Food Security In Southern Mozambique', thesis submitted for the degree of Master of science, Department of Agricultural Economics, Michigan State University.
- Autor, 2003, 'Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing', *Journal of Labor Economics*, 2003, 21, 1.
- Azam, J.P., 2007, *Trade, Exchange Rate and Growth in Sub-Saharan Africa*, Cambridge Cambridge University Press, UK.
- Banerjee, A., E. Duflo and N.Qian, 2012, 'On the Road: Access to Transportation Infrastructure and Economic Growth in China', working paper.
- Baulch, B. 1997a, 'Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration', *American Journal of Agricultural Economics*, 79, 2, 477–487.
- Baulch, B. 1997b, 'Testing for Food Market Integration Revisited', *The Journal of Development Studies*, 33, 4, 512–34.
- Brooks, W. and K. Donovan, 2017, 'Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua', working paper.
- Casaburi, L., Glennerster R. and T. Suri, 2013, 'Rural Roads and Intermediated Trade: Regression Discontinuity Evidence from Sierra Leone', mimeo.
- Cirera, X. and C. Arndt, 2008, 'Measuring the impact of road rehabilitation on spatial market efficiency in maize markets in Mozambique', *Agricultural Economics*, 39, 17-28.
- De Janvry, A. and E. Sadoulet, 2006, 'Progress in the Modelling of Rural Households' Behaviour under Market Failures', Chapter 8 in De Janvry, A. and R.Kanbur, 2006.
- 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*).

- Eaton, J. and S. Kortum, 2002, 'Technology, Geography and Trade', *Econometrica*, 70, 5, 1741-1780.
- Fackler, P.L. and B.K. Goodwin, 2001, 'Spatial Price Analysis', *Handbook of Agricultural Economics*, 1, 971-1024.
- Fafchamps, M. and B. Minten, 2012, 'Impact of SMS-Based Agricultural Information on Indian Farmers', *World Bank Economic Review*, 27, February, 1-32.
- Gopinath, G., P.-O. Gourinchas, C-T. Hsieh, N. Li, 2011, 'International prices, costs and mark-up differences', *American Economic Review*, 101, 6, 2450–2486.
- 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, 2009, 'On measuring the Benefits of Lower Transport Costs', *Journal of Development Economics*, 89, 1, 28-38.
- Jensen, R., 2007, 'The Digital Divide: 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, 2016, 'Seasonality in Local Food Prices and Consumption: Evidence from Tanzania', *Oxford Economic Papers*, 68, 3, 736-757.
- Key, N., E.Sadoulet and A. de Janvry, 2000, 'Transactions Costs and Agricultural Household Supply Response', *American Journal of Agricultural Economics*, 82, May, 245-259.
- 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.
- Morshed, A.K.M.M., 2003, 'What can we learn from a large border effect in developing countries?', *Journal of Development Economics*, 72, 353-369.

- Morshed, A.K.M.M., 2007, 'Is there really a 'border effect'?', *Journal of International Money and Finance*, 26, 7, 1229–1238.
- Omamo, S.W., 1998, 'Transport Costs and Smallholder Cropping Choices: An Application to Siaya District, Kenya', *American Journal of Agricultural Economics*, 80, 116-123.
- Redding, S.J. and M.A Turner, 2014, 'Transportation Costs and the Spatial Organization of Economic Activity', Working Paper 20235, National Bureau of Economic Research, Cambridge, United States.
- Renkow, M., D.G. Hallstrom and D. D. Karanja, 2004, 'Rural infrastructure, transactions costs and market participation in Kenya', *Journal of Development Economics*, 73,1, 2004, 349-367.
- Rosenbaum, P.R. and D. B. Rubin, 1985, 'Constructing a Control Group Using Multivariate Matched Sampling Methods That Incorporate the Propensity Score', *The American Statistician*, 39, 1, 33-38
- Tostão, E. and B.W. Brorsen, 2005, 'Spatial price efficiency in Mozambique's post-reform maize markets', *Agricultural Economics*, 33, 205-214.
- Tschirley, D., D. Abdula and M.T.Weber, 2006, 'Toward Improved Maize Marketing and Trade Policies to Promote Household Food Security in Central and Southern Mozambique', Research Report 60E, Ministry of Agriculture, Mozambique.
- De Vletter, F and E. Polana, 2001, 'Female Itinerant Maize Traders in Southern Mozambique: A Study of a Higher-End Informal Sector Activity and Its Potential for Poverty Reduction', ILO-SAMAT discussion paper 17.
- Volpe Martincus, C., J. Carballo and A. Cusolito, 2017, 'Roads, exports and employment: Evidence from a developing country', *Journal of Development Economics*, 125, 21-39.

Zant, W., 2013, 'How Is the Liberalization of Food Markets Progressing? Market Integration and Transaction Costs in Subsistence Economies', *World Bank Economic Review*, 27, 1, 28-54. DOI: 10.1093/wber/lhs017.

Zant, W., 2017, 'Impact of Mobile Phones on Staple Food Markets in Mozambique: Improved Arbitrage or Increased Rent Extraction?', Tinbergen Discussion Paper 17-021/V, Tinbergen Institute, Amsterdam.

Zavale, H., 2014. 'Analysis of price incentives and disincentives for maize in the Republic of Mozambique for the time period 2005-2013', MAFAP, FAO, Rome.

Appendix

Table A1 Maize production by province (x1000 tonnes)

year	CabDel	Gaza	Inhamb	Manica	Maputo	Nampul	Niassa	Sofala	Tete	Zambez
2002	85651	66921	18455	162822	21769	117435	175233	76091	205199	185198
2003	93071	56453	16691	172190	7622	89112	159660	104119	183433	298928
2004	<i>93743</i>	<i>89445</i>	<i>20724</i>	<i>184375</i>	<i>20376</i>	<i>109505</i>	<i>165788</i>	<i>87553</i>	<i>213827</i>	<i>230428</i>
2005	80363	40818	18013	162180	10400	102544	121748	52651	173989	178811
2006	104987	102091	32456	204026	29265	124000	222590	102489	260331	213241
2007	85655	60941	29049	211935	10891	93911	103820	96837	211826	229045
2008	76120	63815	36890	187079	26556	99623	170402	105093	238901	209090
2009	<i>78292</i>	<i>55351</i>	<i>33686</i>	<i>207662</i>	<i>21358</i>	<i>106359</i>	<i>146640</i>	<i>102677</i>	<i>226168</i>	<i>199351</i>
2010	<i>75522</i>	<i>54044</i>	<i>34541</i>	<i>212661</i>	<i>33227</i>	<i>105741</i>	<i>143084</i>	<i>106001</i>	<i>228720</i>	<i>193658</i>
2011	<i>72850</i>	<i>52768</i>	<i>24699</i>	<i>217780</i>	<i>36049</i>	<i>105126</i>	<i>139615</i>	<i>109434</i>	<i>231301</i>	<i>188127</i>
2012	68410	48675	20625	227748	31570	112494	143761	118346	226912	178848

Source: Trabalho de Inquérito Agrícola (TIA) / Anuario de Estatística Agrária, Ministry of Agriculture; numbers in italics are constructed using a simple specification with year and province fixed effects, province specific trends and seasonal rainfall. Province names are abbreviated in some instances.

Table A2 Availability of transport cost data: by itinerary

itinerary	n	in %	by source	itinerary	n	in %	by source
Alto Molocue – Nampula	62	7.5%		Montepuez – Nacala	25	3.0%	
Alto Molocue-Quelimane	47	5.7%		Montepuez – Pemba	24	2.9%	
Alto Molocue – Maputo	18	2.2%		all from Montepuez			6.1%
Alto Molocue – Maxixe	16	1.9%		Nampula – Maxixe	21	2.5%	
Alto Molocue – Beira	10	1.2%		Nampula – Maputo	18	2.2%	
total from Alto Molocue			19.9%	Nampula – Xaixai	15	1.8%	
Angonia - Tete	48	5.8%		Nampula – Beira	9	1.1%	
total from Angonia			7.6%	total from Nampula			10.0%
Chimoio – Tete	21	2.5%		Nhamatanda – Maputo	53	6.4%	
Chimoio – Xaixai	19	2.3%		Nhamatanda – Beira	16	1.9%	
Chimoio – Massinga	12	1.5%		Nhamatanda – Xaixai	15	1.8%	
Chimoio – Maputo	11	1.3%		total from Nhamatanda			12.1%
total from Chimoio			8.4%	Ribaue – Beira	12	1.5%	
Gorongosa – Beira	39	4.7%		total from Ribaue			2.5%
Gorongosa – Maputo	14	1.7%		Tete – Chimoio	21	2.5%	
total from Gorongosa			8.6%	Tete – Maputo	16	1.9%	
Lichinga – Maputo	12	1.5%		Tete – Maxixe	16	1.9%	
total from Lichinga			2.3%	Tete – Massinga	11	1.3%	
Manica – Beira	25	3.0%		total from Tete			8.5%
total from Manica			4.7%				
Mocuba – Quelimane	16	1.9%					
Mocuba – Nampula	13	1.6%					
total from Mocuba			3.9%				

Source: calculations based on SIMA data

Note: Shaded cells indicate itineraries that experienced a road distance change due to the Caia-Chimuara bridge.

Table A3 Availability of transport cost data: by year

year	n	in %	year	n	in %
2001	39	4.7%	2006	68	8.3%
2002	213	25.8%	2007	51	6.2%
2003	184	22.3%	2008	63	7.6%
2004	54	6.6%	2009	45	5.5%
2005	87	10.6%	2010	20	2.4%
all years	824				

Source: calculations based on SIMA data

Table A4 Explaining transport cost with transport cost data

Dependent variable: $\ln(tc_{ijk})$				
Variables	Sample period: Aug. 2001 – Dec. 2010 (unevenly distributed, see Table A4)			
$\ln(\text{road distance})$	0.608 (0.804)	3.495*** (1.043)	0.838*** (0.297)	0.548*** (0.154)
$\ln(\text{fuel price})$	0.164 (0.129)	0.112 (0.127)	0.119 (0.131)	0.226** (0.115)
$\ln(\text{consumer prices})$	1.437** (0.664)	1.307** (0.663)	1.591** (0.677)	0.550*** (0.210)
seasonality	yes	yes	yes	yes
trend at source	yes	no	yes	no
trend at destination	yes	yes	no	no
adj R ²	0.779	0.770	0.761	0.749
observations	824	824	824	824

All equations are estimated with OLS and include market pair fixed effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5 First stage logistic estimation of propensity score: probability of an optimal route crossing the new Zambezi bridge

Dependent variable: probability of using bridge (d_bridge (binary))				
Sample	65 months before, 65 months after ^b	65 months before, 34 months after ^{b,c}	34 months before, 65 months after ^{b,c}	34 months before, 34 months after ^b
estimation technique	logit			
variables				
year (or trend)	0.702*** (0.082)	0.567*** (0.083)	0.594*** (0.092)	0.393*** (0.099)
across Zambezi (binary)	1.597*** (0.502)	-0.178 (0.481)	1.192** (0.580)	-1.141* (0.615)
across Zambezi (trend)	0.069*** (0.007)	0.010*** (0.008)	0.075*** (0.008)	0.114*** (0.010)
$\ln(\text{population at source})$	0.372*** (0.038)	0.365*** (0.042)	0.384*** (0.038)	0.372*** (0.042)
$\ln(\text{population at dest.})$	0.293*** (0.033)	0.282*** (0.036)	0.292*** (0.033)	0.281*** (0.036)
$\ln(\text{pc production at source})$	-1.743*** (0.076)	-1.791*** (0.083)	-1.771*** (0.077)	-1.812*** (0.085)
$\ln(\text{pc production at dest.})$	-0.056 (0.058)	-0.077 (0.064)	-0.066 (0.058)	-0.088 (0.064)
$\ln(\text{SAFEXspot})$	-2.120*** (0.282)	-2.075*** (0.305)	-2.130*** (0.292)	-2.209*** (0.325)
pseudo R ²	0.6802	0.6974	0.6602	0.6772
observations	16283	15157	13585	12459

^b Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuara.

^c The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

Standard errors are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6 Impact of bridges: Propensity Score Matching (PSM), Nearest Neighbour

Dependent variable: $p_i - p_k$ (if $p_i > p_k$)				
est. technique	65 months before, 65 months after ^a	65 months before, 34 months after ^a	34 months before, 65 months after ^a	34 months before, 34 months after ^a
	PSM / NN	PSM / NN	PSM / NN	PSM / NN
bridge (ATT)	-0.384*** (0.110)	-0.242** (0.100)	-0.381*** (0.107)	-0.260** (0.102)
bridge (ATU)	-0.231	-0.267	-0.239	-0.110
bridge (ATE)	-0.258	-0.263	-0.267	-0.138
treated, on support	1982	1471	1980	1461
treated, off support	148	349	150	359
untreated, on support	9269	7402	7896	6283
untreated, off support	4884	5935	3559	4356
observations	16283	15157	13585	12459
parity bounds	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0	-2.0, +3.0

^a Number of months before and after August 1, 2009, the opening of the bridge between Caia and Chimuará.

The period without bridges is from October 2006 to July 2009, 34 months before August 1, 2009.

Equations are estimated with propensity score matching (PSM), using Nearest Neighbor (n=5), with replication, combined with Caliper threshold (0.5). Estimates of the propensity scores are in the Appendix (Table A5). In all estimations observations with extreme price differences (around 3% of all observations) and extreme road distances (>2500km) are omitted. Standard errors are in brackets next to the coefficient. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

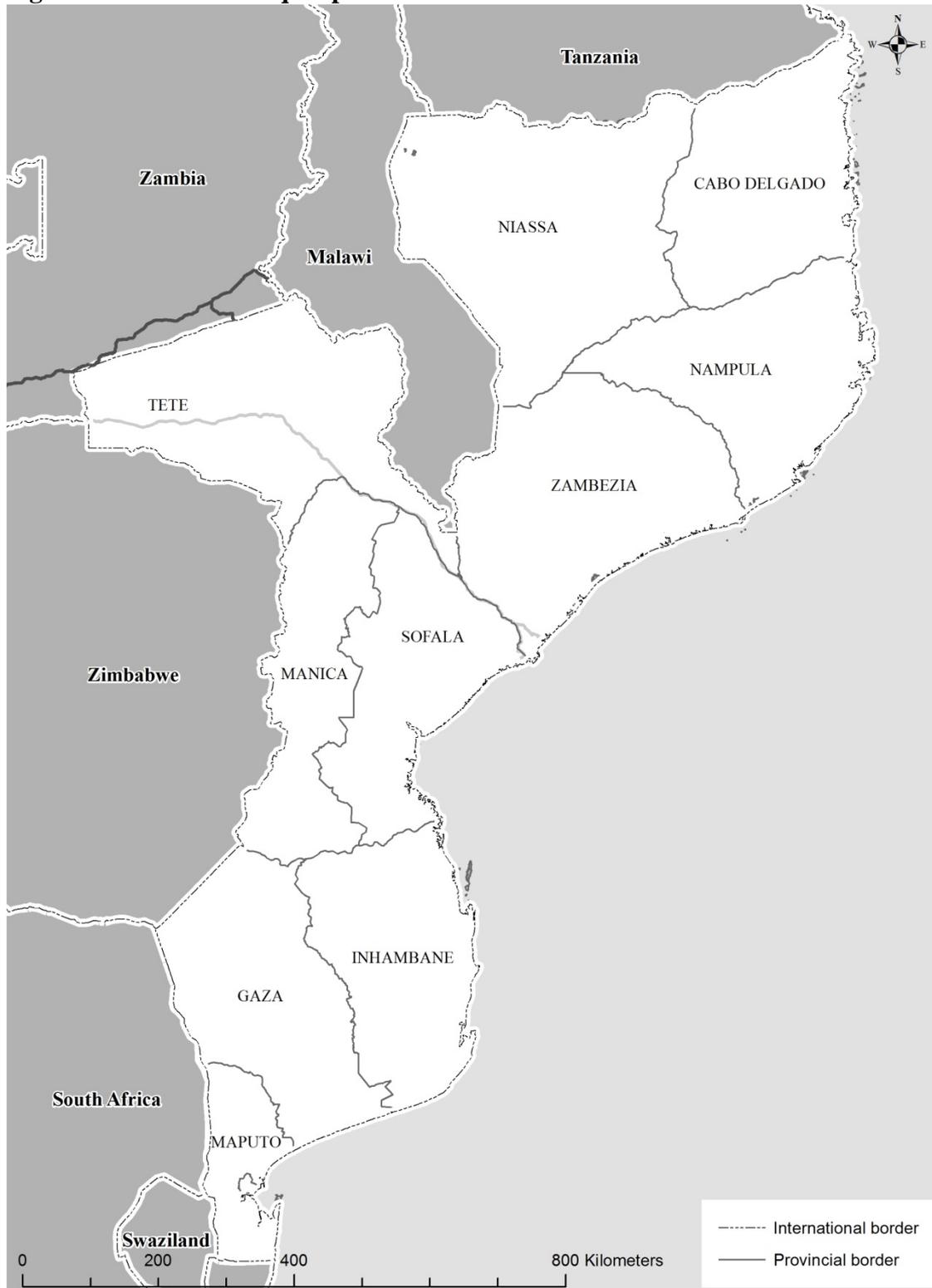
Table A7 Standardized Bias of Covariates, before and after Matching

before and after bridge introduction	65months before, 65 months after		65 months before, 34 months after		34 months before, 65 months after		34 months before, 34 months after	
	before	after	before	after	before	after	before	after
year (or trend)	1.425	1.103	1.523	1.084	1.180	0.937	1.279	0.960
across Zambezi (binary)	2.017	1.640	1.985	1.687	2.120	1.760	2.084	1.739
across Zambezi (trend)	3.322	2.873	3.414	2.753	3.043	2.660	3.105	2.545
ln(population at source)	0.102	0.039	0.082	-0.105	0.115	0.057	0.095	-0.094
ln (population at dest.)	0.329	0.240	0.333	0.073	0.312	0.225	0.317	0.0580
ln(pc production at source)	-0.570	-0.315	-0.626	-0.079	-0.584	-0.353	-0.642	-0.125
ln(pc production at dest.)	-0.162	-0.101	-0.124	0.030	-0.186	-0.135	-0.154	-0.002
ln(SAFEXspot)	0.846	0.730	0.840	0.634	0.589	0.528	0.568	0.425

Note to Table: $B = \frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{(V_1(X) + V_0(X))/2}}$ where \bar{X}_1 (\bar{X}_0) and $V_1(X)$ ($V_0(X)$) are, respectively, the average and variance

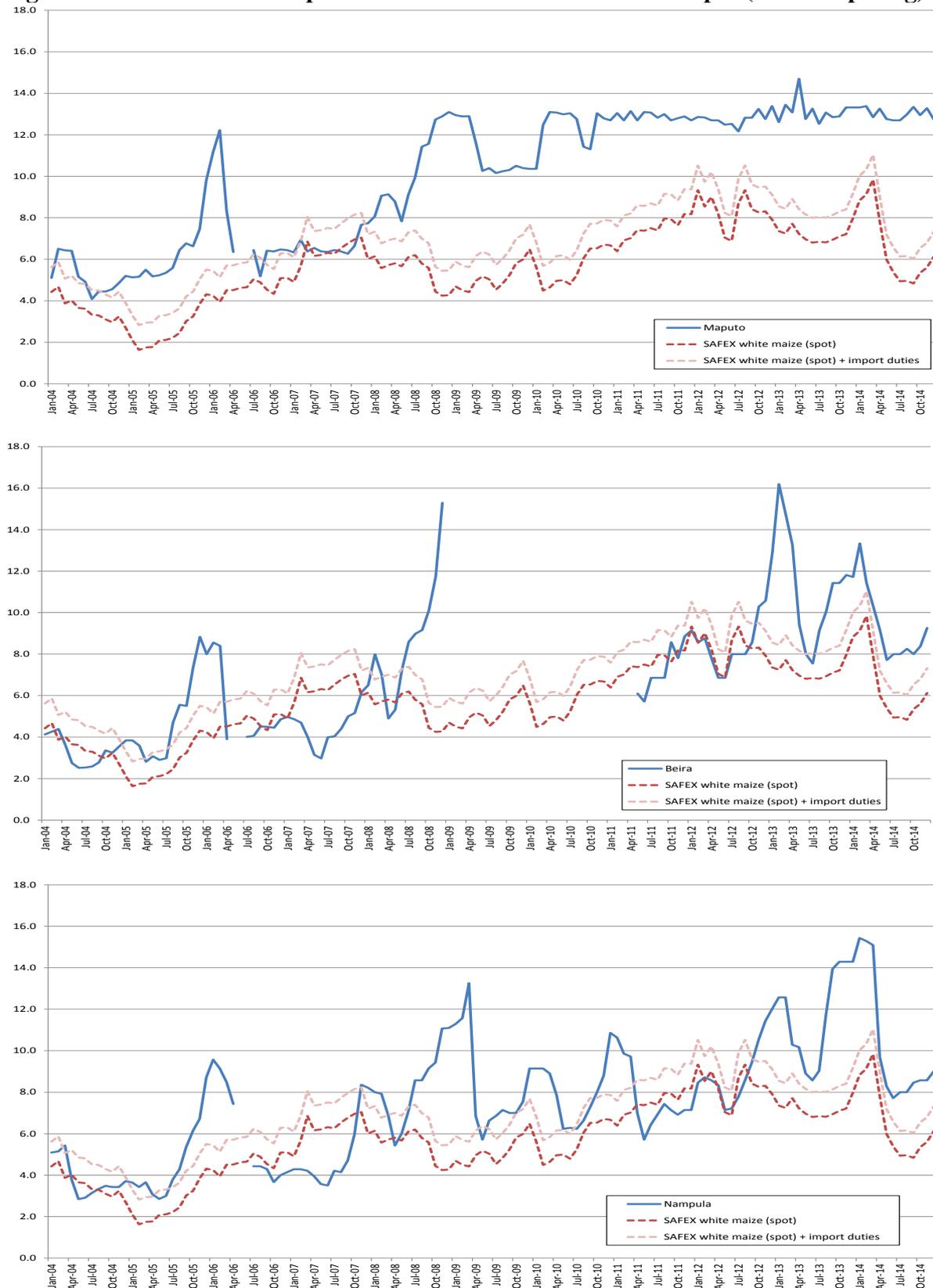
of covariate X in the treatment (control) group. The standardised bias, B, is calculated before and after matching, for each covariate X. The statistics in the table correspond with the Propensity Score Matching estimates with Kernel Matching reported in the main text.

Figure A1 Mozambique: provinces



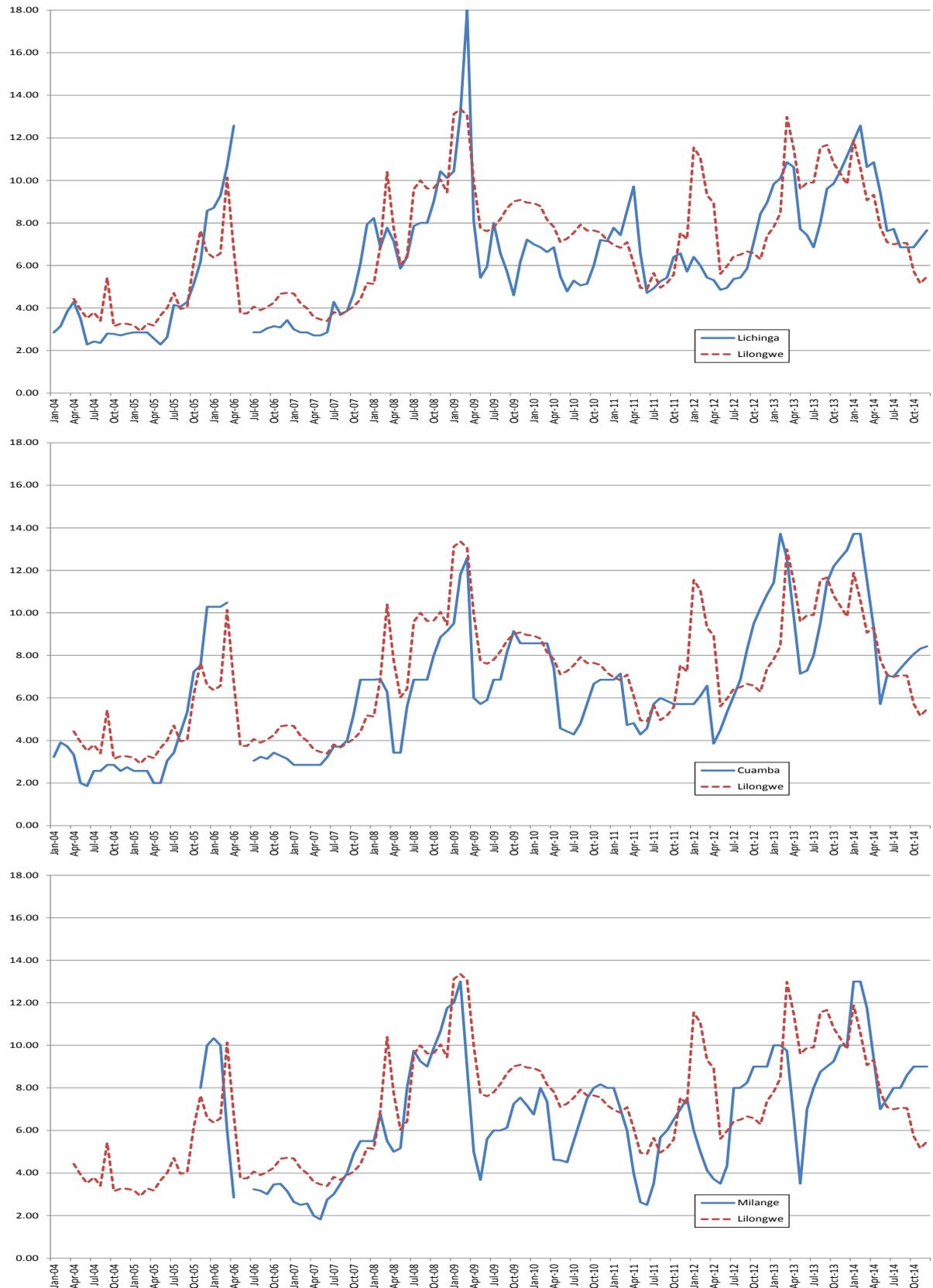
Source: VU SPINlab

Figure A2 Domestic maize prices vis-à-vis SAFEX white maize spot (meticais per kg)



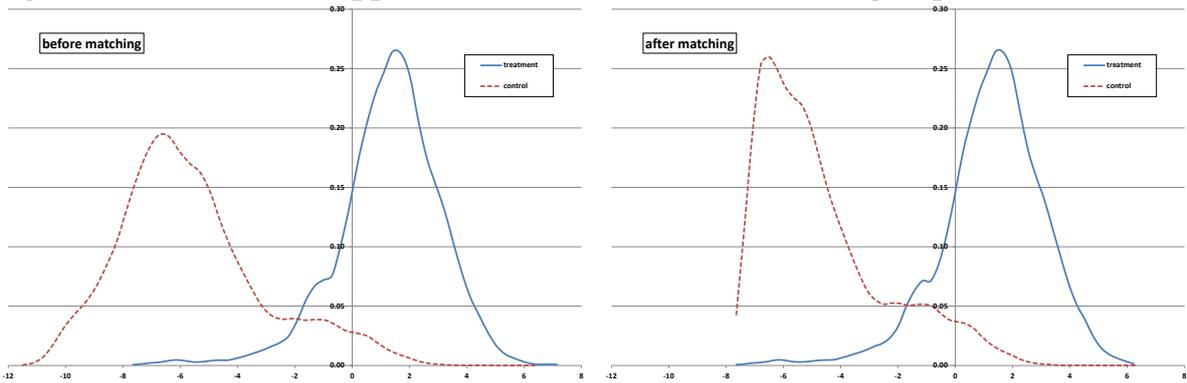
Source: SIMA, IFS (IMF), SAFEX (SAFEX white maize (spot) are monthly averages of daily quotations of the nearest contract), all prices in Mozambique meticais per kg. The light dotted line includes 2.5% import tax and 17% VAT (Tschirley et al., 2006; Zavale, 2014).

Figure A3 Domestic maize prices vis-à-vis Lilongwe maize prices (meticals per kg)

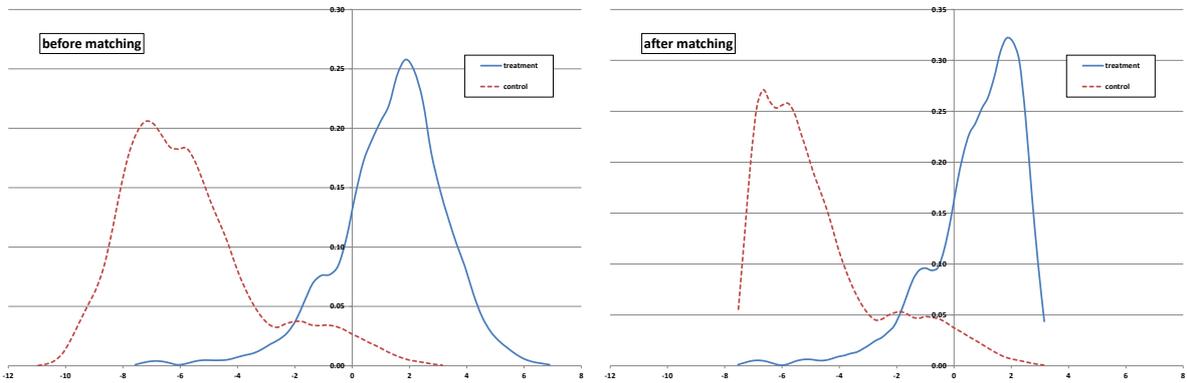


Source: SIMA, IFS (IMF) and FAO, all prices in Mozambique meticals per kg.

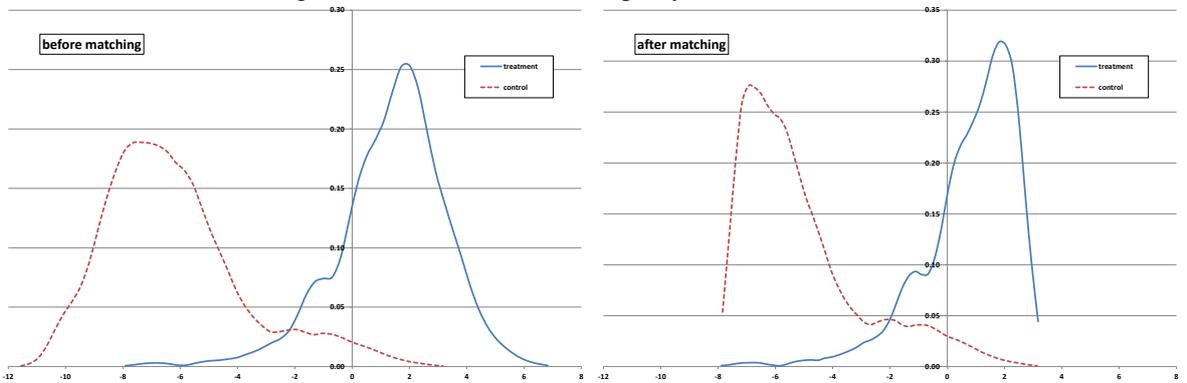
Figure A4 Common support between treatment and control group



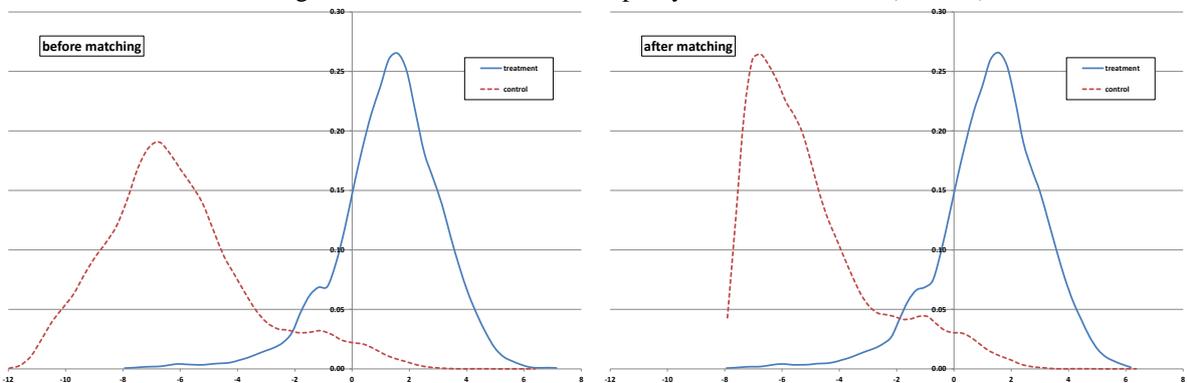
Note: PSM, Kernel Matching, months before/after: 34/65, parity bounds: -2.0/+3.0 (Table 5)



Note: PSM, Kernel Matching, months before/after: 34/34, parity bounds: -2.0/+3.0 (Table 5)



Note: PSM, Kernel Matching, months before/after: 65/34, parity bounds: -2.0/+3.0 (Table 5)



Note: PSM, Kernel Matching, months before/after: 65/65, parity bounds: -2.0/+3.0 (Table 5)

Figure A5 Testing for a common trend in the pre-treatment period

