TI 2017-021/V
Tinbergen Institute Discussion Paper

Impact of Mobile Phones on Staple Food Markets in Mozambique: Improved Arbitrage or Increased Rent Extraction?

REVISION: January 2018

Wouter Zant[1]

[1] Vrije Universiteit Amsterdam; Tinbergen Institute, The Netherlands

Impact of Mobile Phones on Staple Food Markets in Mozambique: Improved Arbitrage or Increased Rent Extraction?

Wouter Zant*

Abstract

We investigate to what extent the roll-out of the mobile phone network in Mozambique reduced search costs, and thereby lowered transport costs and improved the efficiency of agricultural markets. Estimations are based on data of both maize market prices and transport costs of maize grain. Evidence suggests improved arbitrage jointly with increased rent extraction by traders: the rollout explains a 4.5-7 percent reduction in maize price dispersion, and a slightly larger reduction in per ton km transport costs. Benefits of increased market efficiency are not biased towards either producer or consumer markets. Results are robust for non-random rollout of the mobile phone network and several other threats.

JEL code: O13, O33, Q11, Q13

Key words: search costs, transport costs, mobile phones, agricultural markets, maize prices, Mozambique, sub-Sahara Africa

***** Wouter Zant is associate professor at the Vrije Universiteit and research fellow of the Tinbergen Institute, both Amsterdam, the Netherlands; mailing address: Wouter Zant, Vrije Universiteit, De Boelelaan 1105, room 10A-79, 1081 HV Amsterdam, The Netherlands; email: wouter.zant@vu.nl; tel: +31 20 598 9592. I am grateful to Hans Quené for assistance with data compilation, and to Jasper Dekkers for constructing maps. I also thank Taryn Dinkelman and conference and seminar participants at UCLA in Berkeley (ABCA2017), Lissabon (EEA-ESEM2017) and at the VU Amsterdam for useful comments. All remaining errors are my own.

Introduction

This paper investigates the hypothesis that a decrease in the costs of information, due to the introduction of mobile phones, improves spatial arbitrage and the efficient operation of markets, and reduces transport costs. In particular we estimate the impact of mobile phones in Mozambique on dispersion of maize prices and on grain transport costs. The roll-out of mobile phone infrastructure in Mozambique started in 1997 in the Maputo area and around ten years later all major cities and towns had access to the mobile phone network. Traditionally, information on maize prices across markets in poor countries is collected by traders travelling to markets, through word-of-mouth and through personal and professional networks. In the Mozambique context fairly reliable information on agricultural prices and markets is supplied on a frequent weekly basis by Sistema de Informação de Mercados Agrícolas (SIMA). The newly available mobile phone technology, however, allows traders to assess maize prices in many distant markets instantaneously, efficiently, at low costs and customized to personal needs. Improved information as a result of lower search costs leads to a reduction of transport costs and to a reduction of price dispersion across markets.

For the identification of the impact of mobile phones on dispersion of agricultural prices, we use the rollout of the mobile phone infrastructure. A standard difference-in-difference approach (diff-in-diff) with fixed effects is applied to estimate impacts. Since the roll-out is unlikely to be random, we need to control for possible selection bias. The impact estimations are, therefore, complemented with propensity score matching. For the impact estimations we use weekly recorded data of retail market prices of white maize grain and grain transport costs, respectively for the period 1997-2007, and 2001-2010 (source: SIMA). For a variety of transformations, specifications and checks, these core data are complemented with data on distance between markets, population, rainfall, fuel prices, poverty and consumer prices. A few variables, like mobile phone network density and road quality are constructed on the basis of these variables. With the exception of mobile phone rollout and rainfall data, all data are obtained from public domain sources.

We find that the introduction of mobile phones in Mozambique has reduced maize price dispersion by 4.5-7 percent and per ton kilometer transport costs slightly more. Assuming that spatial price differences are equivalent to transactions costs, and approximately the same as the sum of transport costs and rents of traders, this outcome suggests improved arbitrage jointly with increased rent extraction by traders. The claim of higher rents is also confirmed by explaining price dispersion data with transport costs. Increased efficiency of maize markets is further shown to be evenly spread between producer and consumer markets.

The research in this paper is related with various other contributions on the impact of mobile phones on agriculture in developing countries (see e.g. Jensen, 2007; Muto and Yamano, 2009; Aker, 2010; Fafchamps and Minten, 2012; Fafchamps and Aker, 2014; Aker and Ksoll, 2016). Similar to previous work we show that spatial price differences decreases with the introduction of mobile phones. More importantly and unlike previous work, the current study investigates, explains and quantifies the different impact of mobile phones on price dispersion and transport costs. Then we disentangle price impacts by source and destination markets to further quantify benefits of mobile phones for different groups.

The rest of this study proceeds as follows. Section 1 discusses the empirical literature on search costs and on the role of mobile phone technology. Section 2 presents the background on maize marketing, maize trade and maize prices in Mozambique, and discusses the introduction of mobile phones in Mozambique. Section 3 sets out the conceptual framework and the model, discusses data and data sources, and elaborates methodology and empirical strategy. Section 4 presents the impact estimations and presents robustness checks.

Section 5 highlights potential threats and alternative explanations. Section 6 measures benefits to consumers, traders and producers. Section 7 presents the summary and conclusion.

1. Search Costs, Mobile Phones, Transport Costs and Staple Food Prices

There is a growing body of empirical work on the impact of mobile phones and related information technology, on trade and agriculture in developing countries (Jensen, 2007; Muto and Yamano, 2009; Aker, 2010; Fafchamps; Fafchamps and Minten, 2012; Aker and Fafchamps, 2014; Tadesse and Bahiigwa, 2015; Aker and Ksoll, 2016). This empirical work is based both on experimental (RCTs) and non-experimental data, in the latter case exploiting the roll out of mobile phone infrastructure, often jointly with estimation techniques designed for non-experimental data. The overall conclusion, thus far, is that the introduction of mobile phones has caused a decrease in price dispersion (and hence an increase in efficiency of markets), most likely due to increased trader activities. However, there is much less consensus if farmers are benefiting from access to mobile phones or if behaviour of farmers is affected.

Jensen (2007) makes use of micro level survey data to show that price dispersion on fish markets in Kerala, India has dramatically reduced after the introduction of mobile phones, increasing fishermen's profits and also consumer welfare. Easy and timely access to information is also shown to prevent waste, inefficiency and spoilage of production of perishable crops (Overa, 2006; Jensen, 2007; Muto and Yamano, 2009). Muto and Yamano (2009) investigate marketing costs of maize and bananas during the introduction of mobile phones in Uganda, using household data for 2003 and 2005, and show increased market participation of farmers in remote areas, but no other impacts on maize marketing. Asymmetric information between traders and farmers is suggested to block potential benefits for farmers. Aker (2010) finds that price dispersion across Niger millet markets experienced a 10-16% reduction after the introduction of mobile phones, due to traders' activities. The reduction in price dispersion is shown to be stronger for market pairs that are farther apart and if roads have lower quality. Reduction in price dispersion is also shown to be larger once a critical mass of market pairs has mobile phone coverage. The lower reduction in price dispersion compared to Jensen (2007) is attributed to better storability of grain and less perishability relative to fish. Fafchamps and Minten (2012) estimate the benefits for farmers of SMS based agricultural information in Maharashtra, India, using a randomized controlled trial. The information includes prices, weather forecasts, crop advice and new items. They find no effect of this service on the prices received by farmers, value added, crop losses, crop choices and cultivation practices. These results are in line with the limited commercial take-up of the information service, but difficult to reconcile with previous investigations on the impact of information (as documented above). A comparative advantage in transport is suggested as an explanation why benefits accrue in the first place to traders and not to producers. Aker and Fafchamps (2014) find that the introduction of mobile phones in Niger reduced dispersion of producer prices for a semi-perishable crop (cow peas), but does not affect price dispersion of storable crops (millet and sorghum). Also levels of producer prices are not affected, while variation in prices over the years is reduced.

The current study further verifies the impact of mobile technology on price dispersion for a different country (Mozambique), and for a crop that is key to food security. Moreover, with the help of observed transport costs, the impact on price dispersion is compared with the impact on transport costs. This allows to measure directly to what extent transporters benefit from mobile phones. With prior knowledge on typical source and destination markets, it is further investigated to what extent increased market efficiency leads to price changes at the production side or at the consumption side. This allows to quantify if consumers capture a disproportional part of the benefits of improved market efficiency. Finally, we touch upon the issue how maize growers are affected by the introduction of mobile phones.

2. Mozambique Maize Production and Marketing, and Mobile Phone Rollout

Maize production and marketing

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, maize is three times more marketed than cassava and maize has a budget share of similar size as all other staple foods[1] together (Tschirley et al., 2006)[2]. 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. However, 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: with per capita production well above 100kg, the provinces Niassa, Tete and Manica are clearly in the position to supply other parts of the country, or other countries (see Appendix, Figure A8 and Table A3).

Domestic production of maize is concentrated in the central and northern part of Mozambique (for a map of Mozambique, see Appendix, Figure A1). The Northern provinces Niassa, Cabo Delgado, and Nampula have better rainfall distribution and better soil fertility, while the Southern region has unfavourable weather conditions and suffers from occasional pests (Abdula, 2005; Appendix, Figure A8). Most agricultural production in Mozambique is rain-fed. Drought and also flooding cause drops in production. In the 1999-2000 crop season, maize production declined 18 percent, primarily due to floods that devastated large areas of the centre and south of the country (Abdula, 2005). Due to widespread subsistence farming only a limited share of production (around 30% of total production) is traded on the market.

[1] Staples in Mozambique are maize, rice, cassava, wheat, sorghum, millet, sweet potatoes beans and groundnuts.
[2] The current study mainly investigates the period from the end of the 1990s to 2007, and with a few extensions to 2010. This explains the relevance and justifies reference to slightly older policy reports and articles.

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 (Appendix, Figure A1, A8). The major terminal retail markets, nearly all on the seaside, are, from south to north, Maputo (including Matola), Xai-xai, Maxixe, Massinga, Beira, Quelimane, Nacala and Pemba (see Appendix, Figure A1).

Transport of maize in Mozambique is implemented mainly with trucks, and makes use of a modest road network. In 2008 Mozambique's total road network length is 32500km, of which about 22500km is classified network (primary and secondary networks each less than 5000km, and a tertiary network of around 12700 km), while the remaining part is unclassified network (around 6700km) and urban network (3300 km)[3]. Classified and total road network density (km road per 1000km2 land area) are 29 and 37 respectively, which is extremely low, even for low income countries. From the early 1990s the percentage of roads in good or fair condition has increased from 30% to 83%, above the average of other Sub-Saharan low-income countries. Rural accessibility is very low: only around 25% of rural Mozambicans live within 2 km of a classified network road, while 70% of the population is living in rural areas and 22% of Mozambique's GDP comes from the agriculture. Moreover, the poor condition of the rural network – 40% of the rural roads is in poor condition – stand in sharp contrast to the good condition of Mozambique's primary and secondary network. In summary, Mozambique's road infrastructure is not well developed, the trunk roads connecting cities and towns have improved over the past decades and are in good or fair condition, but secondary, tertiary and rural roads are in poor condition, and especially during the rainy season many of these roads cannot be used.

Trade in maize grain – the standard white maize grain quality – takes place throughout Mozambique. However, the Zambezi river (see Figure A1) creates a natural

[3] All numbers on road infrastructure sourced from Dominguez-Torres and Briceño-Garmendia, 2011.

barrier to domestic trade[4]: consequently, major domestic trade flows of maize are from the central area to the south while the northern cities at the seaside are supplied by the more inland production centres in the north. Southern Mozambique, and most notably the Maputo-Matola area, is a major maize deficit area. Maize available for sale in wholesale markets in Maputo (Xiquelene and others) is primarily sourced from Chimoio or Manica in the central region, around 1100 km by road (Abdula, 2005; SIMA data from 1999-2001), but also from markets further away[5]. Southern Mozambique, and the Maputo-Matola area in particular, also rely on South Africa as supplier of maize (see Haggblade et al., 2008; Zovala, 2017). Angonia, a major production area in the northeast, supplies Tete and also occasionally exports maize to Malawi. Exports to Malawi also take place from the Cuamba and Milange region (USGS / FEWS NET; Zovala, 2017). Transport cost data used in the current study (source: SIMA) which are only recorded for itineraries where trade of maize grain takes place, confirm these stylized facts (see also Appendix, Table A3).

The trading sector consists of itinerant traders, large scale assemblers, wholesale traders, millers and retailers. Retailers and millers are at the end of the value chain and are primarily 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 geographically dispersed markets, but is likely to have a large component of value added through collecting, sorting, quality control and distribution. The key agents in Mozambique

[4] Since 2009 – at the far end of the period of study – the Zambezi bridges between Chimuara and Caia, and Vila de Sena and Mutarara became operational. The Chimuara-Caia bridge was newly built and is part of the main north-south highway. The Vila de Sena-Mutarara bridge, around 60 kilometers upstream, originally a railway bridge, converted to a bridge for vehicles in the 1990s, is not connecting a primary highway and was closed for repair from 2006-2009, to be re-opened in 2009 after rehabilitation as a railway bridge. Hence, during the period of study, the Tete bridge was the only fully operational road bridge on a major highway connection. The north-south barrier due to transport costs is sufficient ground to investigate if maize markets north and south of Zambezi are practically separated (see also Zant, 2017).

[5] For example, from Tete, around 1500km by road from Maputo (Tostão and Brorsen, 2005 using SIMA trade flow data from 1998-2001). The largest distance for which trade costs are recorded in the SIMA transport cost data used in this research is from Lichinga to Maputo (by road around 2300km!).

that drive arbitrage between geographically dispersed markets, 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 directly after harvest (April-June). Likewise, in many markets in Mozambique, both north, central and south, most of the maize traded in assembly and retail markets is supplied by informal itinerant traders. Informal itinerant 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 transport once a sufficient quantity / number of bags with maize is collected. Itinerant traders make several trips per season (De Vletter and Polana, 2001). Barriers to enter the trading business appear to be low. We have no information on actual trade flows of white maize grain or number of trading agents actively involved maize trade in Mozambique.

Maize prices over time (see Appendix, Figure A4) reflect the rain-fed character of agriculture and occasional climatic hazards. Prices peaked in 2002 and 2006 due to droughts. Moreover, there is strong seasonality in maize prices: prices begin rising around September, to reach a maximum around March. The degree of seasonality (see Appendix, Figure A5 to A8) is substantial with prices in the lean season twice as high compared to the post-harvesting months and corresponds with observed seasonality in staple food prices in other sub-Saharan countries (see Kaminski et al., 2016). Seasonality in maize prices also appears to be stronger – with higher highs and lower lows – and with a diverging timing in rural areas compared to urban areas (see Appendix, Figure A8).

Mobile phone rollout

Similar to most other sub-Saharan countries, where mobile phone technology was introduced at the end of the 1990s and early 2000s (ITU, 2016), mobile phone technology was introduced in Mozambique in 1997 in the Maputo area. In the following years the network expanded rapidly and around ten years later nearly all major cities and towns had access to the mobile phone network. During the first three years (1997-1999) mobile phone towers were installed exclusively in the Maputo area: in observing sound returns to investments, mobile phone companies concentrated on locations with high population density combined with low construction and maintenance costs for cell phone towers. Further (visual) inspection of the roll-out (see Appendix, Figure A2) suggests that new mobile phone towers have been installed nearly exclusively along the existing trunk roads, most likely also to reduce construction and maintenance costs. In later years the network was extended to more remote and less populated areas. However, rural areas in general, and the province of Niassa in the north in particular, remain typically underserved, both in terms of area and population. Other determinants of rollout, like distance to Maputo or other urban centers, proximity to the existing network, network density, urban status, per capita income, etc., are also likely to be important. However, since the determinants of rollout are key to understanding selection bias in rollout, we elaborate more formally on the observable determinants in the context of the propensity score estimation.

In the 2000s average mobile phone network density[6] in Mozambique as whole increased 5 to 6 fold[7]. The number of phone customers (mobile-cellular telephone subscriptions) in Mozambique increased from 51,065 in 2000 to 7,224,176 in 2010 (ITU, 2016), corresponding with an increase in the share of the population with access from 0.3% in 2000 to 30.1% in 2010. According to western standards still a modest share, but well above the stagnant land line coverage of less than 0.4% (fixed telephone subscriptions in 2010: 88,062). The success of the introduction of mobile phones in sub-Saharan African countries is due to the low prices of mobile phones, the low cost of mobile phone use, the

[6] Cell phone network density = the sum of surrounding locations with cell phone facilities divided by the distance to these locations, for each location with cell phone facilities.
[7] The 2009 network, the final year of our mobile phone network data, is shown in the Appendix, Figure A2.

widespread promotion of the pre-payment system which solved the cashing problem – a key problem with land lines – and the distribution of pre-paid cards for very small amounts. Despite the reasonably low costs of mobile phones and mobile phone use[8], it is likely that use and access to mobile phone services is biased against the poor.

3. Conceptual Framework, Model, Data and Empirical Strategy

Conceptual framework

The mechanism underlying the impact of mobile phone services on price dispersion and transport costs is associated with arbitrage activities of traders in agricultural commodities, and efficiency in transporting activities of drivers and transporters. Traders in agricultural commodities monitor prices of agricultural prices in various markets, searching for profitable arbitrage opportunities, and base their decisions on what and where to buy or sell, on these prices. Price information is typically distributed on a regular basis by public authorities, often a department of the Ministry of Agriculture[9]. Access to mobile phone technology enables these traders to obtain direct and more accurate information, at low cost, and customized to personal needs, from a network of geographically dispersed contacts. Moreover, mobile phone communication may also help to establish agreements on transactions, leading to selling and buying of predetermined quantities at predetermined prices. Thereby mobile phone technology potentially reduces costs associated with selling or buying under uncertainty and optimizes trade decisions.

Transporters earn an income from selling transport services. Transported merchandise could be anything, but in the current developing country context, concerns transport of agricultural commodities. Similar to the case of the commodity trader, transporters monitor

[8] At the time of writing (2016) the price of a simple mobile phone is around 400 Mt (around 5-6US$) and a local phone call around 6Mt per minute (less than US$ 0.10).

[9] In Mozambique SIMA is responsible for distributing price information (see also section on data sources).

potential flows of merchandise and related transport opportunities for several itineraries and base their decision on what to transport and to which market, on this information. But unlike the commodity trader there is no publicly accessible source of information (like SIMA) that records and disseminates information on potential freight. Consequently transporters need to rely on information obtained through their own network or through traders. Access to mobile phone technology clearly allows transporters to better identify transport opportunities, to better identify potential flows of merchandise in geographically dispersed markets, to make arrangements for return cargo more easily and to avoid possible asymmetric information issues with traders / arbitrageurs.

In practice trading and transporting activities are often combined. In Mozambique wholesalers earn an income both from geographical difference in prices of agricultural commodities, but also from undertaking transportation of merchandise between markets. Under sufficiently competitive conditions in transport services all cost reductions in transport will automatically and metical by metical translate into smaller price differences between markets. Conversely, a lack of competition will trigger traders and transporters to exercise their market power and capture rents.

Model

Following the large literature on the identification of trade costs through spatial price differences (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 transaction costs and a mark-up between these markets. In formula:

(1) $$p_j - p_i = tc_{ij} + \mu_{ij}$$

where p_i (p_j) is the market price in location i (j), tc_{ij} are transaction costs of trade from location i to location j, and μ_{ij} is the mark-up imposed by traders for trade from location i to location j.

Since transaction costs are primarily transport costs, transaction costs are approximated with transport costs in the empirical work. A key requirement for the formula to apply is that the empirical analysis should consider identical products, products that are homogeneous over time and across locations, and have negligible quality differences. White maize grain is produced, consumed and traded throughout Mozambique and, in fact, white maize grain is the dominant type of maize (see also previous section). Therefore, and without denying possible quality differences, we have assumed that the requirement of a homogenous product is satisfied.

Next, we identify source and destination markets, initially, by simply exploiting price differences: if the price in location j is higher than the price in location i, it is assumed that j is a destination market and i a source market. In formula:

(2) if $p_j > p_i$ then $p_j = p_{destination}$ and $p_i = p_{source}$

Since both prices and transport costs are observed, it is now possible to estimate, independently from each other, how shocks or exogenous changes like the introduction of mobile phone technology, affect price dispersion and transport costs. Deriving the impact on the mark-up of traders is than straightforward and directly follows from the previous equilibrium expression (equation (1)). Specifically:

(3) $\Delta\mu_{sd} = \Delta(p_d - p_s) - \Delta tc_{sd}$

Data sources

The data on the rollout of mobile phone infrastructure, sourced from the Ministry of Transport and Communication of Mozambique[10], contain 547 names of locations of mobile phone towers, their corresponding latitude and longitude coordinates and the first year of operation. The rollout data that we have stretch from 1997 to 2009. It is unlikely that further extension of the mobile phone network has stopped in 2009. However, with the limited

[10] Cell phone roll-out data were kindly made available by Jenny Aker.

number of markets in major towns and cities, identified in the empirical estimations, the roll-out in our data set has reached all markets already in 2006. The range of a mobile phone tower (or Base Transceiver Station) is, roughly, limited to 35km, but could vary with the height of antenna over surrounding terrain, the frequency of signal in use and various other parameters[11]. We employ a range of 35 km around the mobile phone tower (as the crow flies) to identify markets that have mobile phone facilities. Additionally, we require both source and destination markets to have mobile phone facilities, in order to identify market pairs between which mobile phone communication is feasible.

Maize prices are from Sistema de Informação de Mercados Agrícolas de Moçambique (SIMA; www.masa.gov.mz/sima), from their weekly publication Quente-Quente. SIMA, which started as a USAID / Michigan State University funded initiative, weekly distributes price bulletins, by email, covering amongst others farmer organizations and traders, by SIMA's provincial offices that further reproduce and disseminate the bulletins, through the Ministry of Commerce that uses the information in their own bulletins, and through broadcasts on the national radio and television news to whom SIMA contractually offers weekly input to market programs. Traders' interviews support the effectiveness of the SIMA price information[12]. From Quente-Quente we use in particular the weekly retail market quotations of white maize grain (Quadro 3, Preço e Mudança Percentual a Nível de Mercado Retalhista (MT/kg), grão de milho branco), recorded for 27 markets[13], from January 1999 to December 2007. White maize grain is the dominant quality of maize produced, traded and

---

[11] These other parameters include special equipment, the transmitter's rated power, uplink/downlink data rate of the subscriber's device, directional characteristics of the site antenna array, reflection and absorption of radio energy by buildings or vegetation, local geographical or regulatory factors and weather conditions.

[12] See "In Mozambique, Market Information publishes its 500th weekly bulletin, a Cause for Celebration", February 2005 posted on the internet (www.masa.gov.mz/sima/).

[13] 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. A map in the Appendix (Figure A1) shows the locations of these markets in Mozambique.

consumed throughout the country, the 27 markets form a set of markets that is representative for both rural producer and urban consumer areas, and the period covers the effective period of the roll-out of mobile phone infrastructure. The price data are collected by interviewing each Monday three randomly selected traders in each market and for each commodity.

Overall we have in total more than 6000 observations of prices, more than 50% of the potential number of weekly observations. Hence, and unfortunately, there are missing observations in the price data (see Appendix, Table A1, for an overview of the availability of data by year). However, missing observations are common in agricultural price data: they are highly correlated with the season and with occasional droughts (see Appendix, Table A1) and, consequently, reported by SIMA staff to be due to a lack of transactions. There could be a concern that the missing observations are correlated with mobile phone status: formal tests, however, show that missing observations in both prices, price differences and transport costs are not correlated with the mobile phone rollout (see Appendix, Table A2).

Data on transports costs are from the same source (SIMA). These data are only available for a limited number of itineraries. Collection of these data is organized similarly to the collection of price data, by asking quotations from randomly selected traders and wholesalers in major source and destination markets. Transport cost data are specified by itinerary, by product[14] and by the weight of the bags transported. To find possible measurement errors in the transport cost data, we have verified the nominal values of transport costs by regressing these values on a set of determinants (see Appendix, Table A4). Transport costs are recorded for the period August 2001 to December 2010, with nearly three quarter of the available observations before 2005 (see Appendix, Table A1). After 2010 the publication of these series stops. Again, similar to the case of prices, missing observations in transport costs data are not correlated with mobile phone rollout (see Appendix, Table A2).

[14] Differences in volumes for different crops explains separately recorded per kg transport costs.

Then, a number of miscellaneous variables from different sources are used. Distance, both road distance and Euclidian distance ("as the crow flies") in kilometres, and traveling time in hours, is obtained from GoogleMaps, accessed at the time of implementing this study (2016)[15]. Road distance is relevant for transport costs, while we use Euclidian distance to measure the coverage of mobile phone towers. Road quality is obtained by combining road distance and traveling time. Rainfall data by district, in units of 10 days (so-called decadal data), from 1995 to 2012, are from FEWSNET[16]. We use these data to capture shocks on the supply side due to flooding or drought. Data on population by city are from three censuses (August 1997, September 2007, July 2016), published by Instituto Nacional de Estatistica Moçambique. Monthly series are obtained by interpolation. Population is used to model (relative) demand. Jointly with road distance between cities and mobile phone access, we also use population data to construct network densities. Fuel prices (annuals, country aggregates), exchange rates and consumer price indices are from International Financial Statistics of the IMF, an used as covariates in the estimations. Jointly with road distance, road quality and consumer prices, we use fuel data also to verify and clean transport cost data (see Appendix, Table A4). Poverty head count data are based on household surveys and sourced from van de Boom (2010) and Alfani et al. (2012). Various of these variables are used to model the probability of access to mobile phones, the propensity score.

Methodology and empirical strategy

In order to compare market pairs with and without mobile phone coverage, we estimate – as a start – the following difference-in-difference specification with fixed effects,:

(4) $$y_{jk,t} \text{ (or } z_{jk,t}) = \beta_0 + \beta_1\, cell_{jk,t} + \boldsymbol{X}_{jk,t}\gamma + \eta_{jk} + \theta_t + \varepsilon_{jk,t}$$

[15] Hence, changes in road infrastructure during the period of study (1999-2007) are not properly accounted for.

[16] Rainfall data from FEWSNET were made kindly made available by Benedito Cunguara.

where $y_{jk,t}$ ($z_{jk,t}$) is the price difference (transport costs) between markets k and j, $cell_{jk,t}$ is a binary variable equal to 1 in period t if both markets k and j have mobile phone facilities, and zero otherwise. The vector $\boldsymbol{X}_{jk,t}$ represents variables that influence price dispersion and transport costs, such as drought and flooding in sources markets, fuel prices and differences in demand (due to population size and income). Parameters η_{jk} and θ_t represent market pair and time fixed effects, and $\varepsilon_{jk,t}$ is an error term with zero mean and constant variance. The parameter of interest is β_1 which measures the impact of mobile phones on either spatial price dispersion or on transport costs. Estimations based on a log transformation of the dependent variable, allowing proportional rather than fixed impacts, are included in the Appendix.

The positive difference in price between markets – assumed to be approximately equal the sum of transport costs and trade rents – is expressed per km to allow full comparison with transport costs by itinerary[17]. Hence, in summary, the measure for price dispersion per km is $y_{jk,t}= (p_{k,t} - p_{j,t})/d_{jk}$ where d_{jk}=road distance between source market j and destination market k. In the case of transport costs, and similar to price differences, we express (gross) transport costs between market j and k in period t, per ton and per km. Hence, $z_{jk,t}= tc_{ijk,t}/(d_{jk}.w_{i,t})$ where $tc_{jk,t}$ is gross transport costs for unit i, between market j and k, in period t and $w_{i,t}$ is unit weight, usually the weight of bags. All values (prices, costs) are deflated with the consumer price index to allow comparisons over time.

We can assess the impact separately for price dispersion and transport costs. The results may be challenged because of differences in estimation samples. Also, price dispersion is, of course, not independent of transport costs (see equation (1)). To assess rigorously how price dispersion and transport costs are related to each other we need to regress price dispersion both on per ton kilometer transport costs and the cell phone

[17] Since distance between markets is a market pair fixed effect this transformation is not needed for the estimations. However, it is useful for comparison and required if data are combined (see below and Section 6).

intervention variable, jointly with fixed effects and other controls. Implemting this estimation empirically avoids the problem of different estimation samples. In formula (and for convenience omitting trends and seasonality by source and destination) this is:

(5) $$y_{jk,t} = \beta_0 + \beta_1\, cell_{jk,t} + \beta_2\, z_{jk,t} + \eta_{jk} + \theta_t + \varepsilon_{jk,t}$$

A positive significant impact (β_1) jointly with a positive and significant coefficient of transport costs (β_2) would offer support that traders capture a larger part of the benefits of access to mobile phones.

For the estimation of equation (4) and (5) to generate valid estimates of the impact of mobile phone on price dispersion or transport costs, it is required that both observations of market pairs with and without access to cell phones are random samples. This is unlikely to be the case: the description of the rollout of mobile phone technology clearly reveals several drivers that guided investments in the expansion of the network. To address potential selection bias that arises because of this, we employ a standard technique for non-experimental data, notably propensity score matching. The first step in this technique is to model the probability (not) to have access to mobile phones, the propensity score: observable determinants of the rollout of the mobile phone network are exploited to establish a well performing probability model of access to mobile phones. In the second step a matching algorithm is employed to select observations for comparison, with a similar propensity score, both with and without access to mobile phones. Next, we assess the quality of the PSM estimation: we discuss if the determinants of the propensity score meet the requirements and how well the propensity score is explained, we consider if the matching algorithm is robust, we show if the common support condition is met, and we assess the quality of the matching outcome.

Empirical issues: source and destination markets, trends and seasonality and covariates

We measure price dispersion as the positive maize price difference between market j and k in period t, in formula $p_{k,t} - p_{j,t}$ for all $p_{k,t} > p_{j,t}$, for all markets for which j≠k. As indicated previously we assume that the higher price pertains to destination markets and the lower price to source markets. This mechanic and simple procedure will entail errors due to erratic and incidental fluctuations that bear no relationship with trade flows. To avoid this, we therefore use, additionally, prior information on which markets can be labelled as typical source and typical destination markets, in order to identify price differences that are likely to reflect regular trade flows. Our prior beliefs on the type of market are based on long run values of per capita production, the availability of data on growers' prices, source and destination markets in trade cost data, population size by market and location, notably inland or on the seaside (for details on the choices made, see Appendix, Table A3). Identified source and destination markets that follow our priors, closely correspond with the stylized facts from the Mozambique maize marketing reports (see e.g. Abdula, 2005; Tschirley et al., 2006, FEWSNET, 2010).

With a deflation procedure that is not likely to fully capture price and quality developments and in order to control for technology and network developments, we have included source and destination specific trends to the diff-in-diff specification. Likewise, with the strong seasonality in maize prices (see previous section and Appendix), we have also included source and destination specific monthly dummies. Estimations reported in the Appendix (see Appendix, Table A5 and A6) confirm that including source and destination specific trends and seasonality clearly improve the performance of the estimations. Consequently, trends and seasonality, by source and destination, are included in all estimations.

Covariates ($\boldsymbol{X}_{jk,t}$) are included to control for shocks in demand, supply and trade. In particular we have used the size of population by city (market) to account for differences in demand in source and destination markets. Next, we have used excess and shortage of rainfall as key determinants of supply shocks, in view of the predominantly rain-fed nature of agriculture. Drought is specified as the (log) of a threshold rainfall relative to actual rainfall, conditional on below threshold rainfall levels[18]. In this way the lower seasonal rainfall below the threshold level, the larger the influence of drought, while there is no influence if rainfall is above the threshold level. Flooding is specified as a dummy reflecting the occurrence of rainfall intensities of above 150 to 200mm per 10 days. The influence of rainfall shocks, both drought and flooding, is assumed to extend over the entire subsequent marketing season (from April to March). Finally, since fuel prices are a major contributor to transport costs, we have used (real) fuel prices interacted with source market dummies as covariate.

4. Empirical Estimation and Robustness Checks

Impacts on price dispersion

Table 1a and 1b report the estimation results of the empirical specification proposed in the methodology section (equation (4)). We have included market pair fixed effects, year-month fixed effects, and, following observed patterns in prices, a full set of seasonality and trend, by source and destination, in the estimations of both Table 1a and 1b. Controlling for seasonality and trend is shown to contribute substantially to variation of the dependent variable (see Appendix, Table A5 and A6)[19]. The first column in Table 1, based on the full sample , shows a statistically significant reduction in price dispersion between markets of 228 meticais per

---

[18] Threshold rainfall levels refer to a minimum level of seasonal rainfall required for agricultural crop output. Values of threshold rainfall levels vary from 600mm to 700mm of total rainfall over the rainy season.

[19] Estimations including market pair trends (see Appendix, Table A7) generate impacts of similar size and significance. Market pair trends cannot be included in the transport cost estimations because of lack of data.

1000kg per km. This implies a reduction in price dispersion of around 4.5%, evaluated at the average pre-mobile phone price dispersion.

Table 1a Impact of mobile phones on dispersion of maize prices: basic specification

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$			
	(1)	(2)	(3)
cell phone dummy	-228***	-280.0***	-302***
	(85)	(92.5)	(101)
adj R^2	0.498	0.502	0.491
no. of observations	39498	31265	29296

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): full sample [20]; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets below the coefficient are clustered by market-pairs. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Columns (2) and (3) in Table 1a respectively report impact estimations with data restricted to typical destination markets and typical source markets. Equality of spatial price difference and trade costs (including trade rents) requires positive trade flows. Such trade flows are more likely between typical source and typical destination markets. Therefore we exploit prior information on typical source and destination markets (see Appendix, Table A3) and estimate with restricted samples. With these restrictions we omit, for example, observations of market pairs that have cities like Maputo, Beira and Xai-Xai as source markets, or market pairs that have cities like Chimoio, Mocuba and Cuamba as destination markets. Both sub samples show larger reductions in price dispersion ranging from 280 to 302 meticais per 1000kg per km, a reduction of close to 7% relative to price dispersion before the introduction of mobile phones.

Impacts on transport costs

We proceed with estimating the impact of mobile phone introduction on transports costs. These estimations are reported in Table 1b. Using all available data and controlling for trends

[20] Around 3% of the observations at the right tail of the distribution of the dependent variable are dropped.

and seasonality, the impact is statistically significant at the 1% level (column (1)). Next, and analogous to the estimations on price dispersion, we have applied the same restrictions to the estimation sample. Estimations based on these restricted samples generate a very similar impact: the introduction of mobile phones leads to a statistically significant reduction in transport costs varying from 674 to 692 meticais.

Table 1b Impact of mobile phones on transport costs: basic specification

dependent variable: real transport costs of maize grain per ton-km ($tc_{ijk,t}/(d_{jk}.w_{i,t})$)			
	(1)	(2)	(3)
cell phone dummy	-674.1***	-692***	-674.1***
	(234.7)	(239.0)	(232.8)
R^2	0.8704	0.8663	0.867
no. of observations	1090	1035	1065

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): full sample [21]; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The size of the reduction is larger in the case of transport costs than in the case of price differences, where the difference between these impacts ranges from 674 to 692, equivalent to around 13% of average maize prices. Since the price difference between markets is approximately equal to transport costs and rent extraction (see previous section), the stronger reduction in transport costs needs corresponds with an increase in rent extraction. Hence, the estimated impacts suggest that, jointly with improved arbitrage between markets, traders also have realised increased their rent income.

Robustness checks: equality of pre-intervention trends of treated and non-treated

The diff-in-diff approach requires that that pre-intervention outcomes of intervention and control groups have a common trend. Since all market-pairs obtain access to mobile phone facilities in the course of the rollout, there is no strict distinction between intervention and

[21] Analogous to the estimations of price dispersion, around 3% of the observations at the right tail of the distribution of the dependent variable are dropped.

control groups. However, it is possible to test the common trend assumption in the pre-intervention period for market pairs that obtain access in year t, with market pairs that obtain access only in year t+k. For example, for market pairs obtaining access in 2003 we test if the trend for the years 1999 to 2002 is different from the trend (for the same period) of market pairs obtaining access only in 2004 or later. We have tested if the estimated coefficients of trends for treated and not (yet) treated differ, using a standard F-test.

Table 2a Testing equality of pre-treatment trends in dispersion of maize prices

dependent variable: real positive maize price difference between markets, per km ($(p_{j,t} - p_{k,t})/d_{jk}$)

	(1)				(2)				(3)			
smpl[a]	treated	non treated	F() F p>F	# obs	treated	non treated	F() F p>F	# obs	treated	non treated	F() F p>F	# obs
9900	65.4	19.0	1;495	8997	52.4	54.4	1;258	5258	102	46.4	1;278	6299
	(44)	(33.4)	1.51		(44)	(21)	0.00		(35)	(15)	2.08	
			(0.22)				(0.97)				(0.15)	
9901	-20.2	-45.8	1;516	12761	-15.8	-48.5	1;265	7555	-22.0	-45.9	1;287	8682
	(12)	(15)	3.24*		(17)	(20)	2.77		(17)	(19)	1.70	
			(0.07)				(0.10)				(0.19)	
9902	-6.7	-7.9	1;538	15575	-21.1	-17.3	1;281	9118	-9.6	-18.3	1;297	10419
	(5.5)	(5.8)	0.06		(7.4)	(7.6)	0.28		(8.2)	(7.8)	1.85	
			(0.81)				(0.60)				(0.17)	
9903	-5.2	2.6	1;552	19492	-3.1	6.7	1;294	11975	-6.7	-1.7	1;298	12886
	(4.6)	(4.8)	1.97		(6.0)	(6.4)	2.06		(5.3)	(5.7)	0.61	
			(0.16)				(0.15)				(0.43)	
9904	2.1	11.9	1;559	24007	2.5	15.7	1;297	15463	0.6	7.1	1;298	15709
	(3.3)	(4.0)	4.07**		(3.6)	(4.8)	5.58**		(4.4)	(4.3)	1.23	
			(0.04)				(0.02)				(0.27)	
9905	4.0	10.8	1;592	28711	3.3	10.0	1;322	18912	-1.5	8.4	1;306	18295
	(3.7)	(4.0)	1.82		(7.6)	(4.4)	0.66		(3.9)	(4.2)	4.03**	
			(0.18)				(0.42)				(0.05)	

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. Equations are estimated using OLS. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Column (1): full sample ; column (2) is based on a subset of market pairs that excludes typical producer areas / assembly markets as destination market, while column (3) is based on a subset of pairs that excludes typical terminal markets as source markets (see also Table 1a). *Treated* (*non-treated*) is a trend variable for the pre-treatment period (up to t-1) for market pairs that obtained access to mobile phones in year t (*treated*) and for those that do not (yet) have access (*non-treated*) in year t. F is F-test of coef(*treated*)=coef(*non-treated*) with p values in brackets below the F statistic. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

a. Sample period 9900 is from Jan. 1999 to Dec. 2000, 9901 from Jan. 1999 to Dec. 2001, etc.

Results are reported in Table 2a and 2b. For three price dispersion test outcomes (estimation (1), 9904; (2), 9904 and (3), 9905) equality is rejected at the 5% level. In the course of the rollout, the balance of the observations is skewing towards the treated group, and this easily leads to statistically significant and different trends, given a small non-treated group. Hence, this outcome should not be a major concern. On the basis of the test results we cannot reject the hypothesis of a common trend in the pre-treatment period for treated and not treated, both in price dispersion and in transport costs.

Table 2b Testing equality of pre-treatment trends in transport costs

dependent variable: real transport costs of maize grain per ton-km

	(1)				(2)				(3)			
smpl[a]	treated	non treated	F() F p>F	#obs	treated	non treated	F()	#obs.	treated	non treated	F()	#obs
0102	89.6 (178)	67.2 (12)	1; 61 0.02 (0.90)	378	94.6 (158)	66.8 (19)	1; 51 0.03 (0.86)	322	168.1 (5.4)	169.1 (10.0)	1; 29 0.04 (0.85)	202
0103	-42.5 (9 2)	24.2 (19)	1; 83 0.49 (0.49)	638	-28.9 (93)	29.2 (27)	1; 70 0.35 (0.55)	548	-77.0 (96.4)	53.9 (95.2)	1; 43 0.56 (0.46)	386
0104	15.2 (17)	-43.0 (69)	1; 85 0.72 (0.40)	707	13.1 (22)	48.6 (95)	1; 72 0.14 (0.71)	617	22.6 (28)	59.3 (111)	1; 44 0.11 (0.74)	435

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight. Equations are estimated using OLS. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, while column (3) is based on a subset of pairs with typical production areas as source markets (see also Table 1b). *Treated* (*non-treated*) is a trend variable for the pre-treatment period (up to t-1) for market pairs that obtained access to mobile phones in year t (*treated*) and for those that do not (yet) have access (*non-treated*) in year t. F is F-test of coef(*treated*)=coef(*non-treated*) with p values in brackets below the F statistic. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

a. Sample period 0102 is from Aug. 2001 to Dec. 2002, 0103 from Aug 2001 to Dec 2003, etc.

Robustness checks: impacts with propensity score matching estimations

In order to address possible selection bias, we proceed with estimating impact using propensity score matching, a powerful and well established technique for non-experimental data to address these issues. We start with modelling the propensity score. The propensity score (the probability of treatment) is the probability to have access to mobile phone technology, in both markets of each market pair. Since, the treatment is a binary variable – 1 if both markets have access and zero elsewhere – we employ a logit model to estimate the propensity score. We model the propensity score by tapping from the description of the rollout, by inspecting the geographical pattern of the rollout over the years and by considering the likely drivers of investment in mobile phone infrastructure by mobile operating firms. These firms assess potential demand, driven by population and income, jointly with the costs of installing new mobile phone towers, i.e. the costs of expanding the mobile phone network, which are assumed to increase with the distance to their operational bases, located in big cities. The geographical pattern of the rollout over the years supports these determinants: highly populated cities and towns, with high income inhabitants, are first served, and remote high-cost locations, usually with a low per capita income and with a high incidence of poverty, follow, but with a quite substantial delay. The location of markets in the cell phone network both affects costs and potential demand, and is, hence, also likely to be an important determinant. Following these considerations we estimate the propensity score with population, poverty, distance to big cities and network density, averaged over source and destination market. These variables simultaneously influence assignment into treatment or control group and the outcome variable in both price dispersion and transport costs, and are themselves unaffected by assignment into treatment / control. We have further included a time trend. Results of the propensity score estimation are reported in the Appendix (Table A10). Coefficients of the covariates in the propensity score estimation have expected signs: positive for population and negative for poverty and distance to big cities. The pseudo R2 indicates how well variables explain the probability to have access to mobile phone technology in both markets of each market pair 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, for example, compared to Nearest Neighbour matching, Kernel Matching results in a lower variance, and, thus, higher precision estimates. Kernel Matching is also more 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[22].

We have tested the robustness of the matching algorithm by also implementing Nearest Neighbour (NN) as a matching algorithm. For these estimations we employ 2 to 10 of the nearest controls, with replacement, combined with a caliper threshold, where the caliper takes values 0.005 and 0.01. Replacement is justified because the distribution of the propensity score is very different in the treatment and control group, which may lead to selection of distant counterfactuals. The diverging distributions are also apparent from the common support figures shown in the Appendix (Figure A9 and A10). 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 small, in all cases less than 5% (see Table 3A, 3B, and Appendix, Table A10, A11). Estimations with NN matching generate similar results (see Appendix, Table A10) as with Kernel Matching, though with a lower accuracy: the relative similarity of

[22] 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’.)

the estimations with different types of matching offers confidence about the robustness of the matching procedure.

Table 3a Impact of mobile phones on price dispersion: PSM, Kernel Matching

outcome variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$			
	(1)	(2)	(3)
ATT	-117.3** (57.5)	-165.4*** (60.2)	-137.1** (64.4)
ATU	732.1	363.4	558.7
ATE	1095.9	32.6	113.9
treated, on support	17742	13578	13106
treated, off support	7337	6446	5717
untreated, on support	10365	8126	7397
untreated, off support	4054	3115	3076
no. of observations	39498	31265	29296

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. Equations are estimated using Propensity Score Estimation with Kernel Matching (Epanechnikov kernel; bandwidth=0.06; see main text for details). Column (1): full sample ; (2) is based on a subset of market pairs that excludes typical producer areas / assembly markets as destination market, while column (3) is based on a subset of market pairs that excludes typical terminal markets as source markets. Standard errors are in brackets below the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3b Impact of mobile phones on transport costs: PSM, Kernel Matching

outcome variable: real transport costs of maize grain per ton-km $(tc_{ijk,t}/(d_{jk}.w_{i,t}))$			
	(1)	(2)	(3)
ATT	-426.5*** (143.0)	-552.8*** (143.3)	-465.8*** (143.3)
ATU	-329.5	-336.3	-402.1
ATE	-382.9	-452.8	-436.7
treated, on support	322	297	307
treated, off support	488	467	482
untreated, on support	263	255	259
untreated, off support	17	16	17
no. of observations	1090	1035	1065

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight. Equations are estimated using Propensity Score Estimation with Kernel Matching (Epanechnikov kernel; bandwidth=0.06; see main text for details). Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, while column (3) is based on a subset of pairs with typical production areas as source markets. Standard errors are in brackets below 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 A12 and A13). In the case of price dispersion, the propensity score distribution before matching is very different in the treatment and the control group: in the control group the propensity score is highly skewed towards the lower end and in the treatment group to the higher end, largely due to the increase in treatment over time. 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. Visual inspection of the figures confirm that in the range of treatment values of the matched propensity score all treatment observations have control observations with a similar propensity score, with positive probabilities, both for the case of price difference and transport costs. 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)[23]. The results of this exercise, reported in the Appendix (Table A12) indicate that matching on the estimated propensity score balances the covariates in the matched samples reasonably well.

The PSM impact estimations by and large confirm the results obtained through OLS / diff-in-diff. If anything, the results suggest an impact on price dispersion that is slightly lower than in the case of OLS / diff-in-diff estimates (see Table 1a), and this also applies to the PSM estimated impact on transport costs (see Table 1b). Estimated PSM impacts are also higher in transport costs relative to price dispersion: if this difference is real it can be explained as an increase in traders’ rents that is caused by the introduction of mobile phones.

Robustness checks: OLS with covariates

Variation in price dispersion and transport costs may also be explained by other observables. In order to take account of this we have re-estimated the OLS estimation of Table 1 and included covariates. To control for a variety of demand, supply and trade effects we include in the estimations population at source, the incidence of drought and flooding and fuel prices. Estimations, reported in Table 4a and 4b, further confirm previous results: impact are all

[23] $B = \frac{(\bar{X}_1 - \bar{X}_0)}{\sqrt{(V_1(X)+V_0(X))/2}}$ 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.

statistically significant and substantially higher in the case of transport costs. Overall, and both for price dispersion and transport costs, impacts are slightly stronger in covariates are included.

Table 4a Impact of mobile phones on dispersion of maize prices: OLS with covariates

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$			
	(1)	(2)	(3)
cell phone dummy	-354.0*** (83.4)	-415.5*** (88.2)	-454.8*** (98.2)
ln(population)	-11491.3 (8449.8)	18681.4** (9738.0)	-8913.9 (9161.8)
drought	-984.0*** (314.0)	-1043.4** (358.9)	-2189.9*** (638.7)
flooding	-156.5*** (57.8))	-184.6*** (64.3)	-92.9 (79.8)
ln(gasoline prices)[a]	yes	yes	yes
adj R^2	0.5033	0.5092	0.4975
no. of observations	39498	31265	29296

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets next to the coefficient are clustered by market-pairs. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

[a] Gasoline prices are interacted with source markets.

Table 4b Impact of mobile phones on transport costs: OLS with covariates

dependent variable: real transport costs of maize grain per ton-km $(tc_{ijk,t}/(d_{jk}.w_{i,t}))$			
	(1)	(2)	(3)
cell phone dummy	-703.5*** (253.0)	-760.9*** (247.3)	-703.5*** (250.9)
ln(population)	-3365.3 (5916.7)	-2460.4 (6486.5)	-3365.3 (5867.8)
drought	-4844.7 (4809.7)	-6752.9 (5195.7)	-4844.7 (4769.9)
flooding	59.7 (269.3)	45.3 (268.1)	59.7 (267.1)
ln(gasoline prices)[a]	yes	yes	yes
R^2	0.8750	0.8712	0.8717
no. of observations	1090	1035	1065

Note to table: See also note to table above. Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight.

5. Potential threats and alternative explanations

We discuss a number of concerns that may jeopardize the interpretation of the estimated impacts. The first concern relates to 'other factors': other factors may have taken place in the course of time that have triggered both the placement of mobile phone towers and supply and/or demand fluctuations in the maize markets. Since installing a mobile phone infrastructure does not take place overnight, requires an extensive preparation phase and a long run perspective on commercial viability, these investments are unlikely to be triggered by year to year fluctuation of any ‘other factor’. Hence, factors underlying fluctuations in supply and demand of the maize market can safely be assumed to be independent of long run decisions on mobile phone investments.

A second concern is about possible migration of traders in response to availability of mobile phone services. Traders may transfer their activities to markets and itineraries that have access to mobile phone services. Increased trader activity will reduce price differences between markets. Formally we cannot rule out this possibility: data on the number of traders active in different markets and on different itineraries are lacking. Also the size of trade flows between markets, and their fluctuations over the years, is unknown. Nevertheless, we consider migration of traders on a large scale unlikely given market uncertainties and potential costs. A third concern, related to the previous one, is increased entry (or exit) of traders in response to the availability of mobile phone services. Again, we cannot rule out this possibility. At the same time it appears logical to explain such a development as part and parcel of increased efficiency of markets. Where the estimates are assumed to pick up the short run response to increased information (changes in price dispersion and transport costs, with no change in the number of traders), changes in trade intensity (increase in trading capacity, number of traders, migration) may reflect long run response, that eventually will also impact on prices and costs.

A fourth concern is that the availability of mobile phone services (and increased trade) may trigger a supply response from maize growers, since maize growers possibly benefit from increased transparency of market prices of inputs and outputs and trade costs, leading to higher productivity and higher farm gate prices. Increased supply of maize grain would dampen price dispersion. Higher farm gate prices and increased supply from maize growers would be an attractive outcome since – in that case – part of the benefits of mobile phones are flowing to the rural poor. Estimating the impact of mobile phones on farm gate prices and on the gap between farm gate and market prices would offer a measure of the size of the benefits to producers that nicely compares with the measures of benefits to traders and consumers. This is, however, beyond the scope of the current work. Indirect evidence, based on highly aggregated (and also not undisputed[24]) data on maize area and maize production by province indicates a modest, but statistically significant supply response of 4% increase in area and a 11% increase in production, with the placement of every 10 mobile phone towers (see Appendix, Table A14)[25]. These calculations are, however, not accurate and need further work. With a high prevalence of subsistence farming and small shares of maize production sold on the market, price elasticities of supply are notoriously low in sub-Saharan Africa. Therefore, a large supply response from maize growers that precisely coincides with the availability of mobile phone services would be an unprecedented behavioural reaction from maize growers, rarely observed in sub-Saharan Africa.

A similar concern – the fifth concern – arises with respect to demand: in response to the availability of mobile phone services demand in general, and demand for maize in particular, may increase. With a large share of maize in the Mozambique consumption diet, one would expect an influence, if overall demand increases. Increased demand for maize in destination markets will have an increasing effect on price differences. This implies that estimated impact are biased downwards. Since reported impact estimates are statistically significant and of reasonable size, this concern is not a major worry. Formally, survey data on consumption are needed to further verify possible changes in demand due to the introduction of mobile phone services. It hardly needs mention that such an abundance of data is rare.

---

[24] Series collected with different methodologies (respectively Ministry of Agriculture, Aviso Previo, and Trabalho de Inquérito Agrícola (TIA) are reported to show large discrepancies (see Donovan, 2008).

[25] The elasticity of maize area and maize production is respectively around 0.0039 and 0.011 (see Appendix, Table A7): each additional mobile phone tower generates a maize area and production increase of, respectively, 0.0039% and 0.011%. Hence, 10 additional mobile phone towers, covering an average area of nearly 40,000 km2 ($10x[\pi 35^2]$), a quite substantial share of the area size of most provinces, have given rise to a 4% increase in maize area and a 11% increase in production (see Appendix for underlying estimations).

A final concern is about collusion: mobile phone services allow collusion between traders by directly facilitating communication and coordination. To a certain extent such a collusion appears to be supported by the impact estimates. Traders succeed in capturing a larger part of the benefits of this improvement in information. However, with a large number of traders, dispersed over a vast country, over a multitude of itineraries and involving millions of trade transactions, it is difficult to believe that this outcome is due to collusion: it is more likely due to asymmetric information.

6. Who benefits (more) from access to mobile phones?

A welfare analysis in order to assess which group – farmers, traders or consumers – benefits from access to mobile phones, would be most attractive but is clearly beyond the domain of the available data. There are, however, a few steps that can be made to further explore how benefits are distributed among these groups. Of particular interest is, in the first place, the size of the benefits accruing to traders (due to less transport costs) relative to the reduction in price dispersion, and, secondly, the extent to which the reduction in price dispersion can be attributed to producer locations or consumer locations.

Are traders' benefits from mobile phones larger than the efficiency increase of markets?

The larger reduction in transport cost due to mobile phones found in the separately estimated equations in the previous sections could be the result of differences in estimation samples. To investigate if traders' benefits from mobile phones are really larger than the increase in efficiency of markets, we can match the price difference data with the transport cost data, and estimate how price differences respond to transport costs jointly with the cell phone dummy. The matching of price and transport cost data singles out the possibility that different impacts are due to different samples and allows stronger conclusions. Hence, the empirical specification, repeating equation (5), to overcome this problem is:

$y_{jk,t} = \beta_0 + \beta_1\ cell_{jk,t} + \beta_2\ z_{jk,t} + \eta_{jk} + \theta_t + \varepsilon_{jk,t}$.

A positive significant impact (β_1) jointly with a positive and significant coefficient of transport costs (β_2), possibly restricted to 1, would lend support that traders capture a larger part of the benefits of access to mobile phones.

Table 5 Combining maize price dispersion with transport costs

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$

	OLS	OLS with β_2=1	PSM / KM
cell phone dummy	2226.8*** (601.8)	2448.8*** (678.9)	
per tkm transport cost	0.499*** (0.172)	-	
Trend	-25.150* (13.1)	-27.949*** (11.4)	
seasonality	yes	yes	
market pairs (dummies)	yes	yes	
year (dummies)	yes	yes	
R^2	0.7317	0.6481	
no. of observations	306	306	
ATT			2035.3*** (606.0)
ATU			1342.6
ATE			1847.7
treated, on support			105
treated, off support			22
untreated, on support			39
untreated, off support			4
no. of observations			170

Note to table: Maize price data are from January 1999 to December 2007 and transport cost data are from August 2001 to December 2010 and maize price data are from January 1999 to December 2007 (source: SIMA). Due to merging of data and conversion to monthlies only a limited number of observations remain: this prevents the inclusion of seasonality and trends by source and destination in the estimations. Nominal series are deflated with the consumer price index. In the case of OLS estimations, robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. The propensity score estimation, common support, and standardized bias of covariates of the PSM are not reported but available from the author on request.

Unfortunately, due to the matching of price dispersion and transaction cost data, estimation of this equation suffers from a low number of observations. To increase the number of observations we have converted the data from weeklies to monthlies, implicitly assuming that transport costs and prices have negligible within month variation. Even then we cannot include year-month dummies, jointly with seasonality and trends by source and destination, but have to apply a more simple specification with year dummies, market pair dummies, and seasonality and trend (general, not specified by source or destination). Since we are interested to estimate to what extent the excess reduction of transport costs over reduction of price dispersion is associated with the cell phone rollout, we have also estimated equation (5) under the restriction that $\beta_2 = 1$. Under the assumption of spatially competitive markets this estimation is equivalent to estimation the traders' rents directly (see also equation (3)). The propensity score matching also follows this specification and is further built up along the same lines as the PSM estimation of Table 3A and 3B[26].

The estimations, summarized in Table 5, all support a positive and statistically significant impact of cell phones on traders' rents, while per ton km transport costs are also positive and statistically significant in the unrestricted specification. The size of the increase in rents is larger than expected: we attribute this to sample differences. In summary: the evidence of estimated impacts, estimated independently for price dispersion and transport costs, clearly suggest a larger impact on transport cost. This result is further supported by the estimation results that join the price dispersion data with the transport cost data and thereby avoids the differences in impact that might arise out of different samples.

Is mobile phones induced reduction of price dispersion the result of an increase of prices in source markets or a decrease of prices in destination markets?

We proceed with exploiting the distinction in source and destination markets to find more evidence on which group benefits from access to mobile phone technology. One may consider the reduction of price dispersion the result of either an increase of prices in source markets or a decrease of prices in destination markets. If the price difference is equal to $p_{destination} - p_{source}$ (by assumption), a decrease in this expression comes about either through a (relative) decrease in $p_{destination}$ or a (relative) increase in p_{source}. If the entire decrease in price dispersion is due to a decrease of prices in destination markets, consumers of maize capture

[26] Details on the propensity score estimation, common support, and standardized bias of covariates are not reported but available from the author on request.

the benefits of access to mobile phones. Alternatively, if the entire decrease in price dispersion is due to an increase of prices in source markets, maize traders (or producers of maize[27]) capture the benefits of access to mobile phones. This is investigated by estimating essentially the same specification as in the case of price dispersion (see Table 1a), with the only difference that the dependent variable is now price levels in markets, rather than price differences across markets (and, consequently also one set of seasonality and trend variables is included). Since the PSM technique cannot adequately capture the seasonality in price levels, we have adjusted the price series in the PSM estimation for seasonality, using the exact same specification as in the OLS estimation (but of course excluding the impact variable). Estimation results, reported in Table 6, are shown results for all markets (column (1)), for typical source markets (column (2)) and for typical destination markets (column (3)).

Table 6 Impact of mobile phones on maize prices in source and destination markets

Dependent variable: real maize prices ($p_{j,t}$)			
Estimation technique: OLS			
	(1)	(2)	(3)
cell phone dummy	-924.7***	-666.8*	-818.4**
	(292.6)	(310.9)	(348.4)
R^2	0.7961	0.8144	0.7844
no. of observations	4693	2247	2446
Dependent variable: real maize prices ($p_{j,t}$), seasonally adjusted			
Estimation technique: PSM / KM			
	(1)	(2)	(3)
ATT	-903.6*** (140.5)	-532.5*** (187.9)	-943.0*** (162.6)
ATU	-116.7	-311.3	-850.3
ATE	-572.7	-496.8	-893.4
treated, on support	930	910	94
treated, off support	2692	810	1808
untreated, on support	675	175	108
untreated, off support	396	352	436
no. of observations	4693	2247	2446

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index. OLS estimations include year-month and market dummies, and seasonality and trends. The outcome variable in the PSM /KM estimation is seasonally adjusted in an equivalent way. In case of OLS: robust standard errors in brackets below the coefficient are clustered by market. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$. Column 1 shows the estimation with prices of both source and destination markets, column 2 with prices of typical source markets and column 3 with prices of typical destination markets. The propensity

[27] The current exercise does not allow to identify if maize growers benefit from mobile phones. For that purpose we need to investigate farm-gate prices vis-à-vis market prices.

score estimation, common support, and standardized bias of covariates of the PSM are not reported but available from the author on request.

The estimations indicate statistically significant reduction in prices, both in source and destination markets, and combined, and also for both estimation techniques. The results are rather sensitive for different ways of attributing markets to either the source or destination: we have searched for a configuration that followed our priors (see Appendix, Table A3) and yields mutually consistent outcomes. Despite the substantial differences between the size of the impact – a certain difference is needed for consistency with a reduction in price dispersion – we could not reject equality between source and destination coefficients. On the basis of this evidence we should conclude that the benefits of the improved efficiency are not concentrated on either the consumption or the production side.

7. Summary and conclusion

This study investigates empirically the impact of the mobile phone roll-out in Mozambique on price dispersion and transport costs. Estimations suggest a 4.5% to 7% decrease in price dispersion, indicating an improvement in the efficiency of maize markets as a result of the introduction of mobile phones. The reduction in transport cost is larger: for different specifications this reduction is, evaluated at the average maize price, around 7%-9% points larger reduction. The larger impact on trade costs suggests that a part of the benefits of the introduction of mobile phones translates into increased rents for traders, next to improved arbitrage and efficiency of maize markets. Combined estimation further supports this claim. Finally, the evidence suggest that the reduction in price dispersion (or, euquivalently, the improved markets efficiency) comes about primarily in the form of lower prices in destination markets, in the form of higher source prices. Hence, the benefits of improved market efficiency accrues mainly to consumers and market prices in source areas appear not to be affected. The retail market prices used in this empirical work are not adequate to investigate if maize growers have benefited from mobile phones. The evidence does support, however, a modest supply response from producers. Robustness of impacts is verified by checking the parallel trend assumption underlying the diff-in-diff approach, and by employing propensity score matching to control for possible selection bias. The plausibility of several alternative explanations and potential threats is discussed.

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.

Aker, J.C., 2010, 'Information for Markets Near and Far: Mobile Phones and Agricultural Markets in Niger', *American Economic Journal: Applied Economics*, 2, 3, 46-59.

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

Aker, J.C. and C. Ksoll, 2016, 'Can mobile phones improve agricultural outcomes? Evidence from a randomized experiment in Niger', *Food Policy*, 60, 44-51.

Alfani, F., C. Azarri, M. d'Errico and V. Molini, 2012, 'Poverty in Mozambique: New Evidence from Recent Household Surveys', The World Bank, Policy Research Working Paper 6217.

Anderson, J. and E. van Wincoop, 2004, 'Trade Costs', *Journal of Economic Literature*, 42, 3 691-751.

van den Boom, B., 2010, Analysis of poverty in Mozambique, household poverty status, child malnutrition and other indicators 1996/97, 2002/03, 2008/09, study commissioned by the Royal Dutch Embassy in Maputo, mimeo.

Dominguez-Torres, C. and C.Briceño-Garmendia, 2011, 'Mozambique's Infrastructure: A Continental Perspective', AICD report, World Bank, Washington.

Donovan, C., 2008, 'Agricultural Statistics in Mozambique: Institutional Organization and Performance', report was produced as a background paper for the World Bank on agricultural statistics, under the direction of Richard Harris.

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.

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

Haggblade, S., T. Jayne, D. Tschirley and S. Longabaugh, 2008, ‘Potential for Intra-Regional Maize Trade in Southern Africa: n Analysis for Zambia at the Sub-National Level’, ReSAKKS, working paper no 35.

International Telecommunication Union (ITU), 2016, www.itu.int/ (accessed December 2016).

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

Kaminski, J., L. Christiaensen and C.L.Gilbert, 2016, ‘Seasonality in Local Food Prices and Consumption: Evidence from Tanzania’, *Oxford Economic Papers*, 68, 3, 736-757.

Keyser, J.C. and P. Arlindo, 2007, ‘Mozambique Country Case Study’, background paper for the Competitive Commercial Agriculture in Africa Study (CCAA).

Muto, M. and T. Yamano, 2009, ‘The Impact of Mobile Phone Coverage Expansion on Market Participation: Panel Data Evidence from Uganda’, *World Development*, 37,12, 1887-1896.

Overa, R, 2006, ‘Networks, Distance and Trust: Telecommunications Development and Changing Trading Practices in Ghana’, *World Development*, 34, 7, 1301-1315.

Tadesse, G. and G. Bahiigwa, 2015, ‘Mobile Phones and Farmers’ Marketing Decisions in Ethiopia’, *World Development*, 68, 296-307.

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, Directorate of Economics, Republic of Mozambique.

Tostão, E and B. Wade Brorsen, 2005, 'Spatial price efficiency in Mozambique's post-reform maize markets', *Agricultural Economics*, 33, 2, 205–214

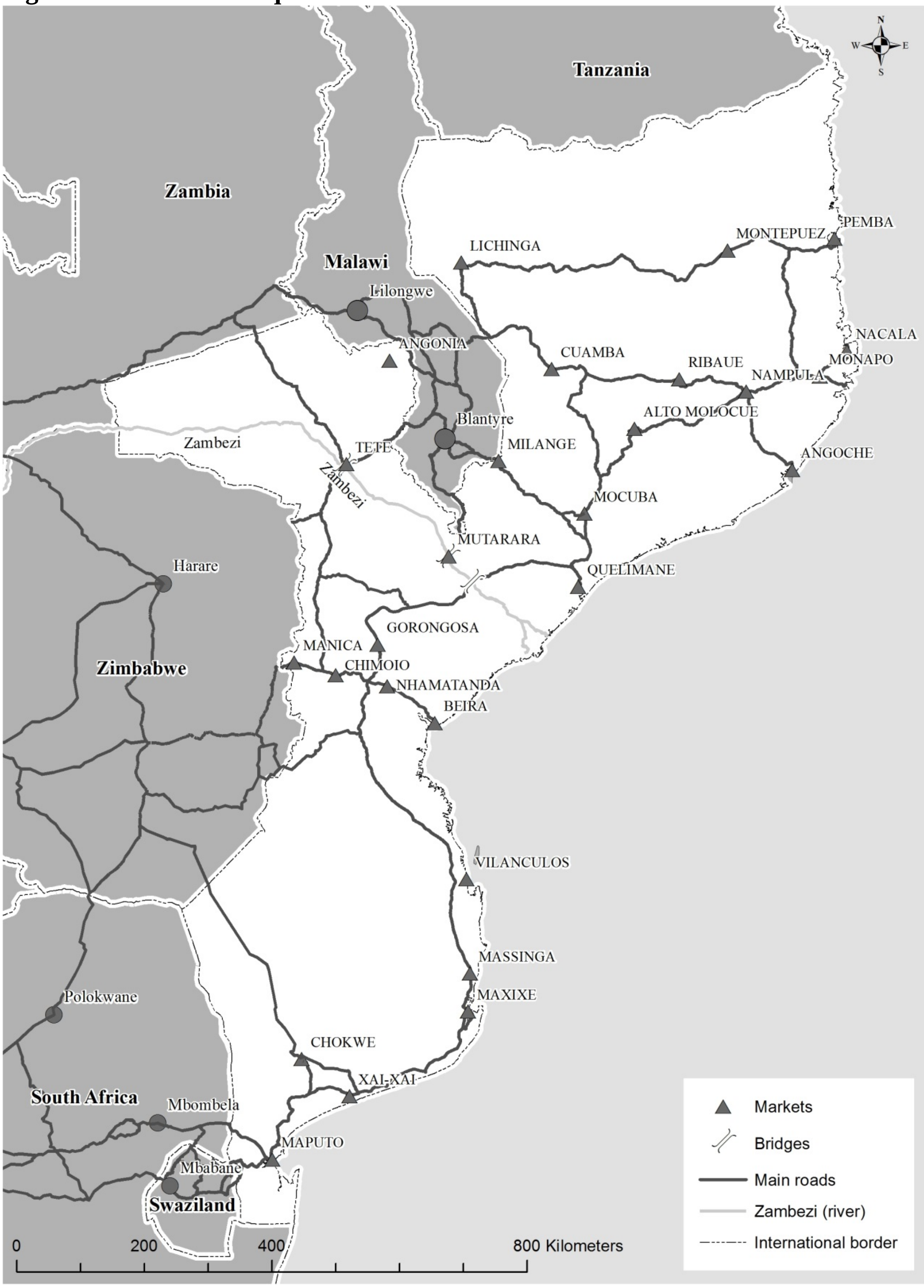
Zant, W., 2017, 'Bridges', Tinbergen Discussion Paper 17-117/V, Tinbergen Institute, Amsterdam.

Acknowledgements

Data on the roll-out of the mobile phone network, sourced from the Ministry Of Transport and Communication of Mozambique, were kindly made available by Jenny Aker, and FEWSNET data on rainfall were kindly made available by Benedito Cunguara. I am grateful to Hans Quené for skilful assistance in compiling the price and transport cost data, and to Jasper Dekkers (VU SPINlab) for constructing maps. For useful discussions and helpful comments and suggestions, I thank Benedito Cunguara, Taryn Dinkelman and seminar participants in Berkeley (ABCA2017), Lissabon (EEA-ESEM2017) and Amsterdam (Vrije Universiteit). All remaining errors are my own.

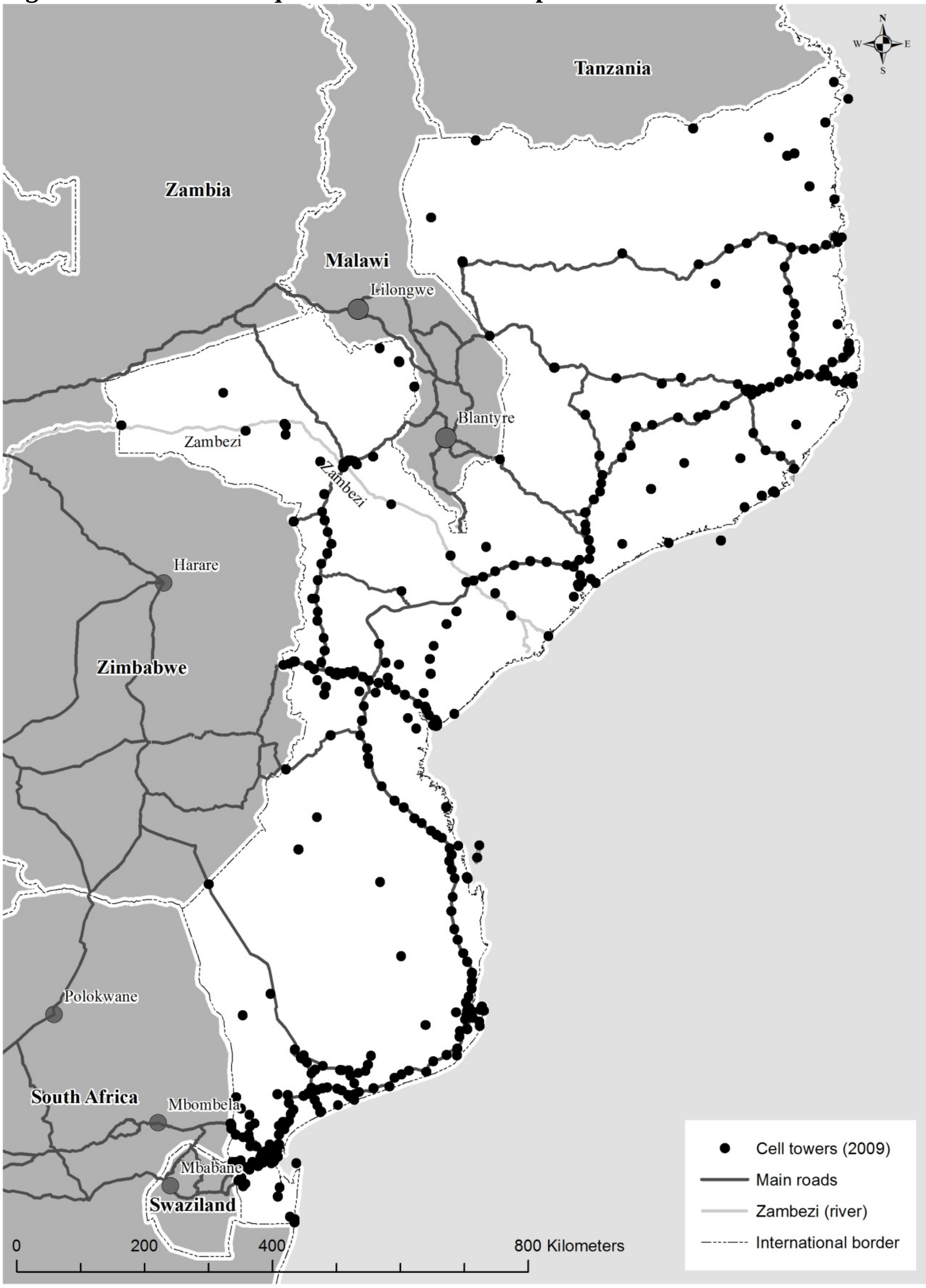
Appendix

Figure A1 Mozambique: markets and roads



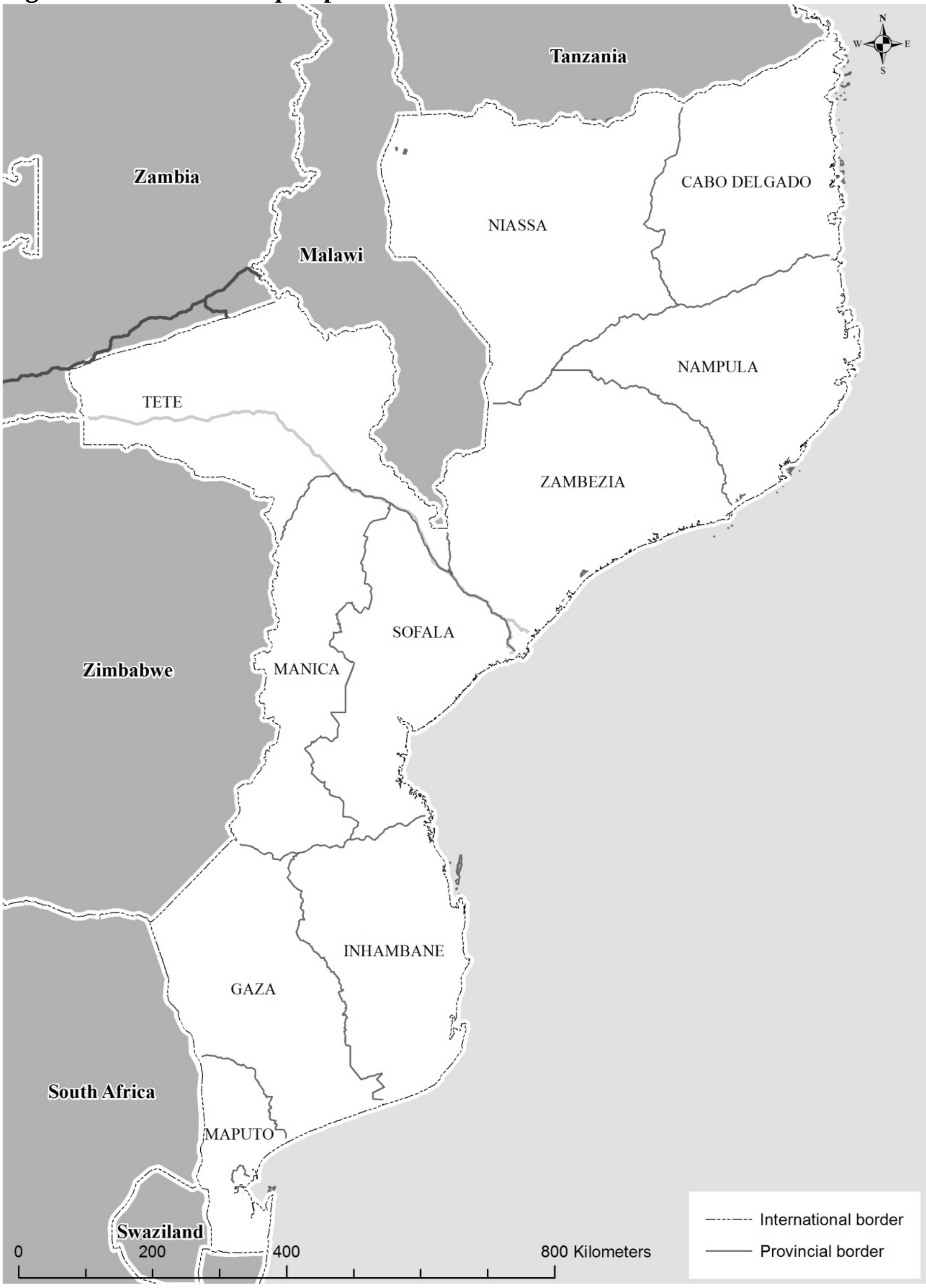
Source: VU SPINlab

Figure A2 Mozambique: network of mobile phone towers in 2009



Source: VU SPINlab

Figure A3 Mozambique: provinces



Source: VU SPINlab

Figure A4 Maize price by region (nominal prices, January 1999-July 2007)

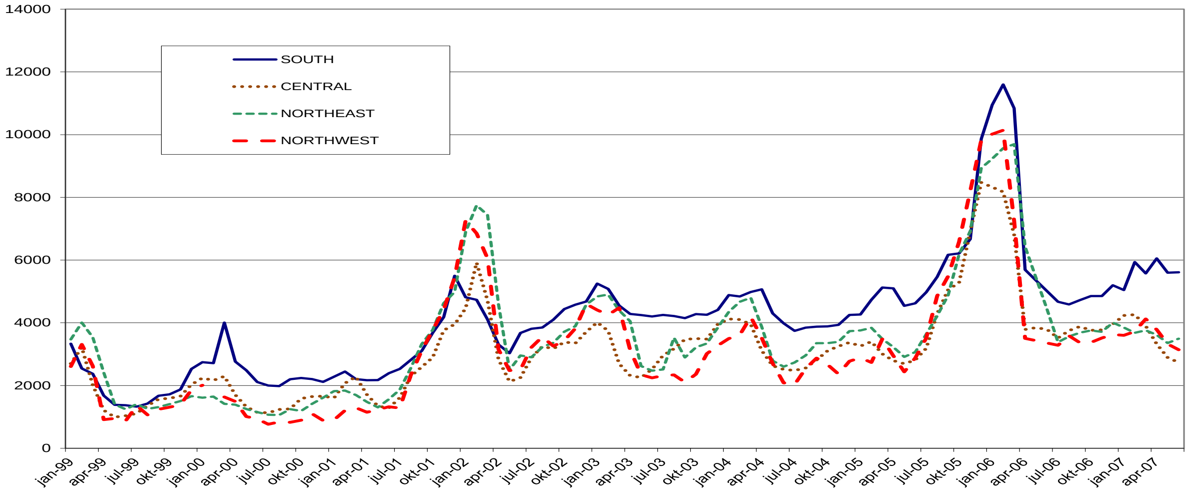


Figure A5 Prices in source and destination markets in the south

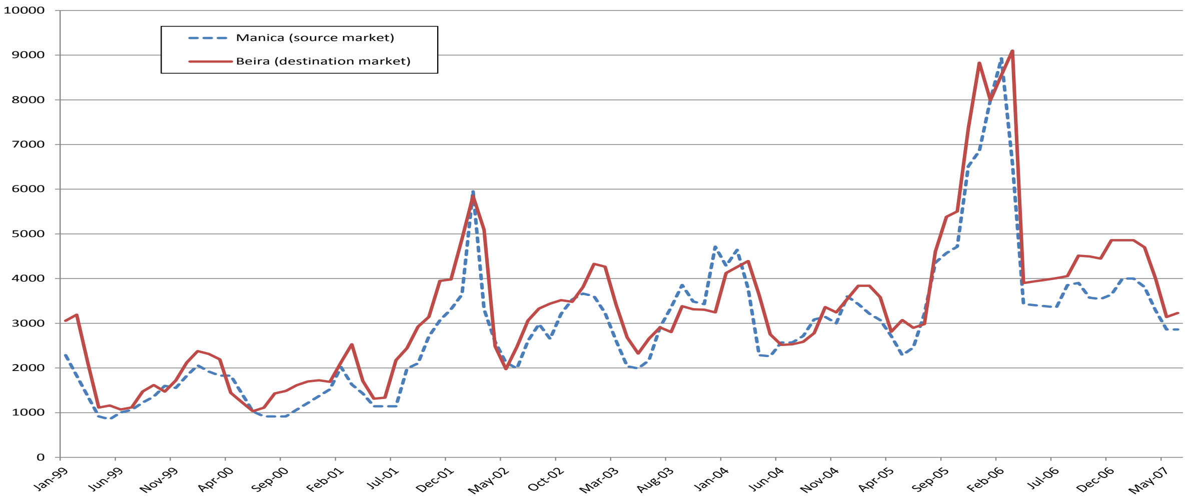
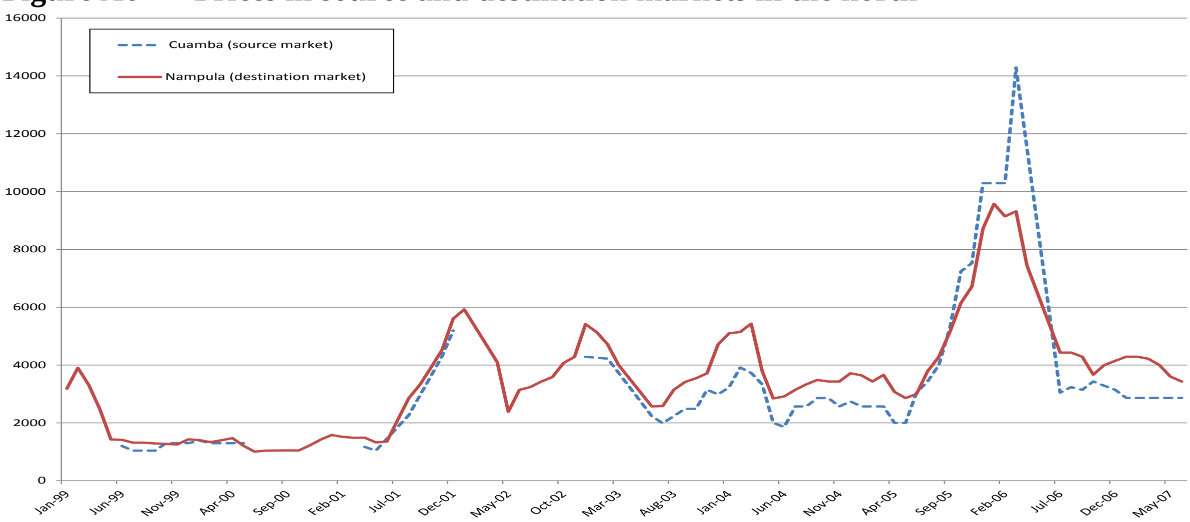
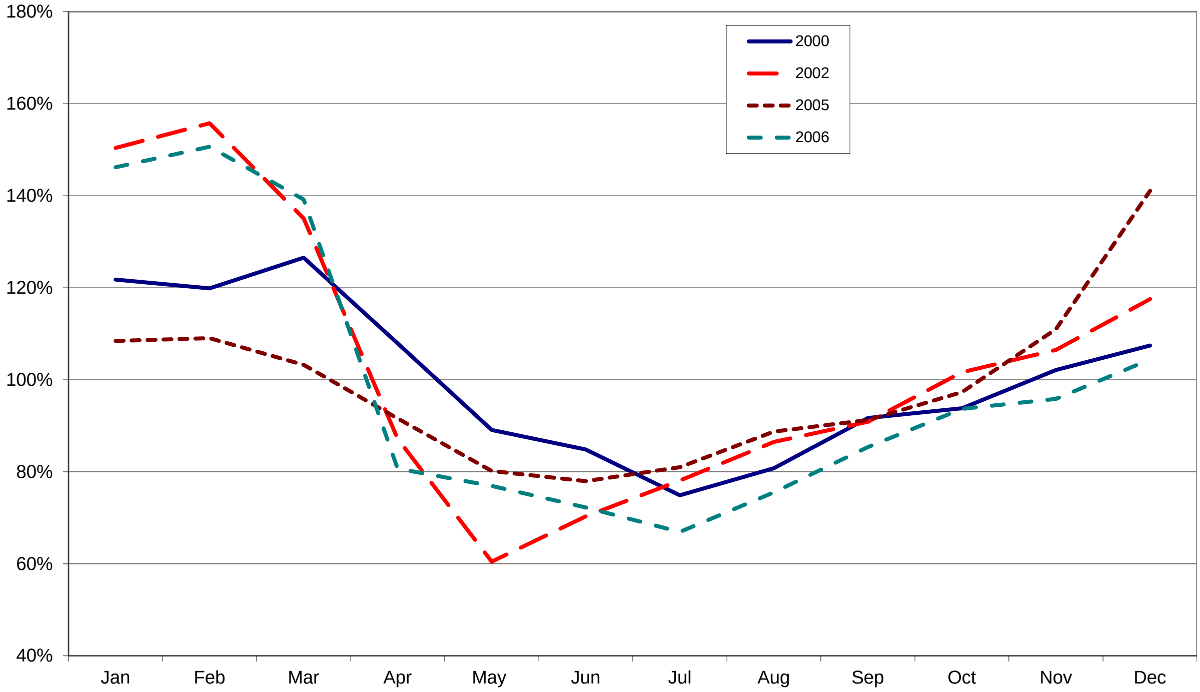


Figure A6 Prices in source and destination markets in the north



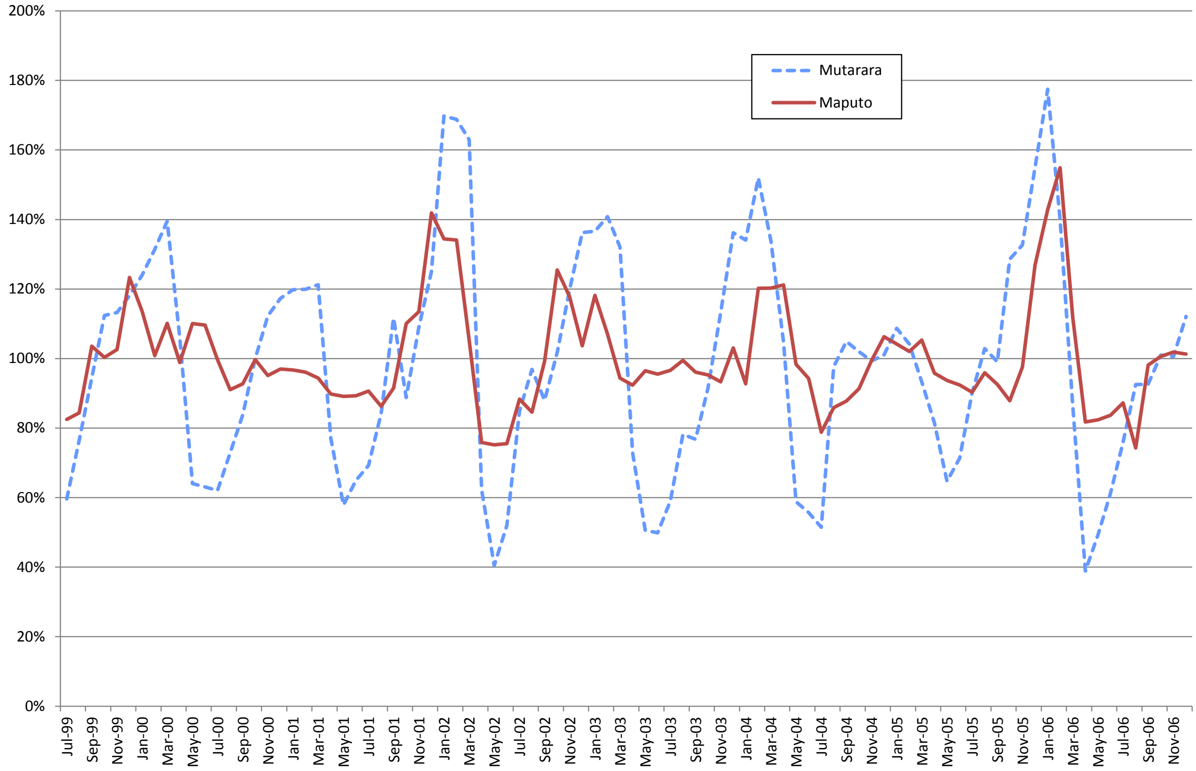
Source: SIMA (Figure A4-A6)

Figure A7 Seasonality in maize prices, by year, selected years



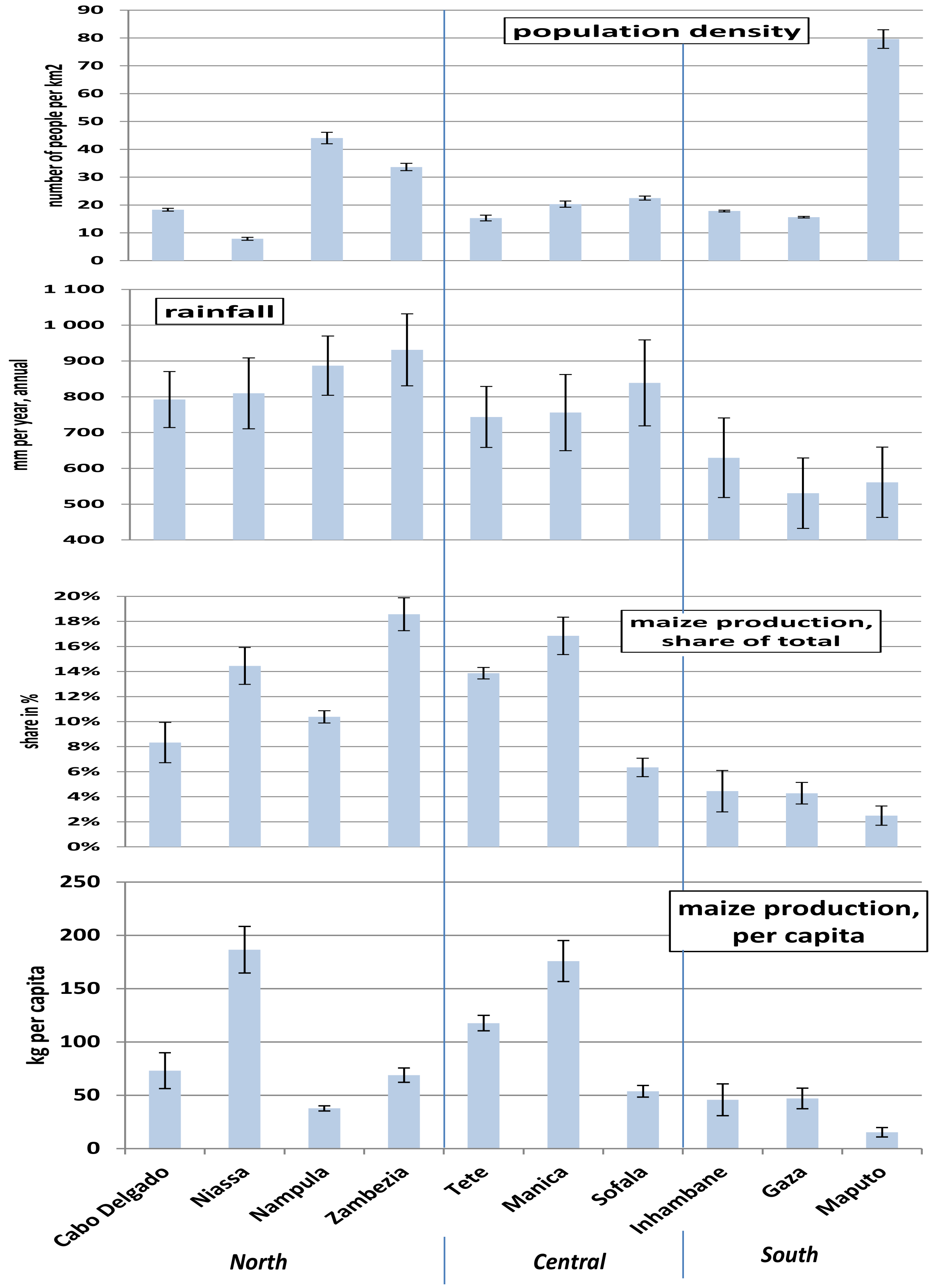
Source: (author's calculations based on data from) SIMA

Figure A8 Seasonality in maize prices, by market over time, source and destination



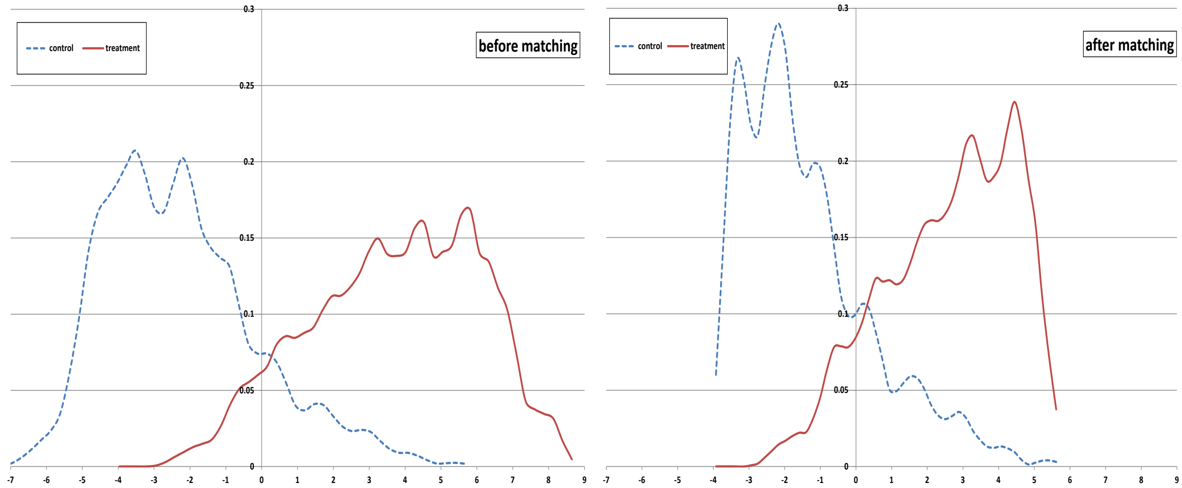
Source: (author's calculations based on data from) SIMA

Figure A8 Population density, rainfall and maize production by province, 1999-2007

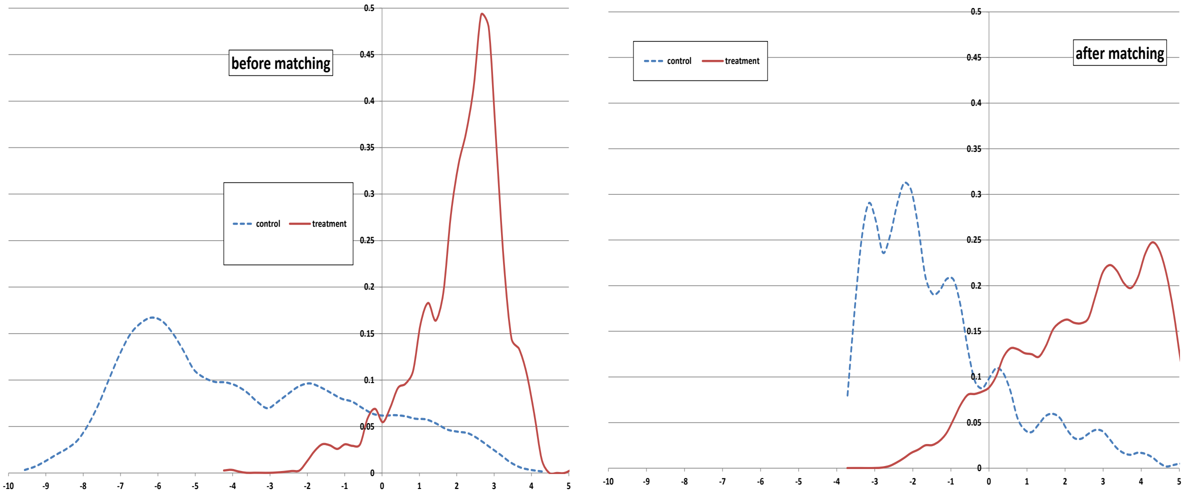


Source: (author's calculations based on data from) Instituto Nacional de Estatistica Moçambique, FEWSNET and Ministry of Agriculture, Early Warning Unit (Aviso Previo); The figure is based on aggregate (average) annual province data. Error bars show 95% confidence intervals. See the maps in this appendix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

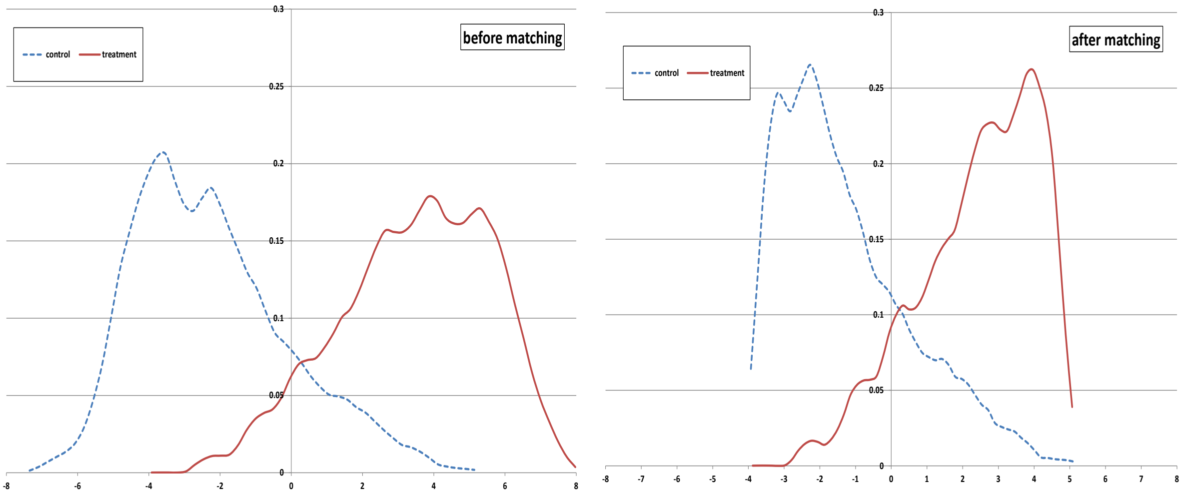
Figure A9 Common support between treatment and control group: price dispersion



Note: PSM, Kernel Matching, all observations (Table 3A, column 1)

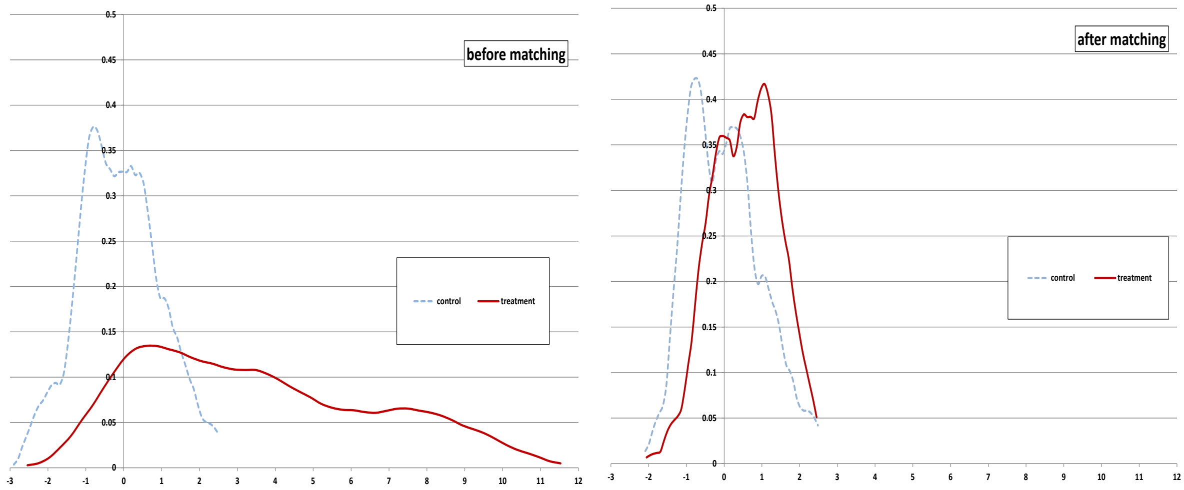


Note: PSM, Kernel Matching, typical destination markets (Table 3A, column 2)

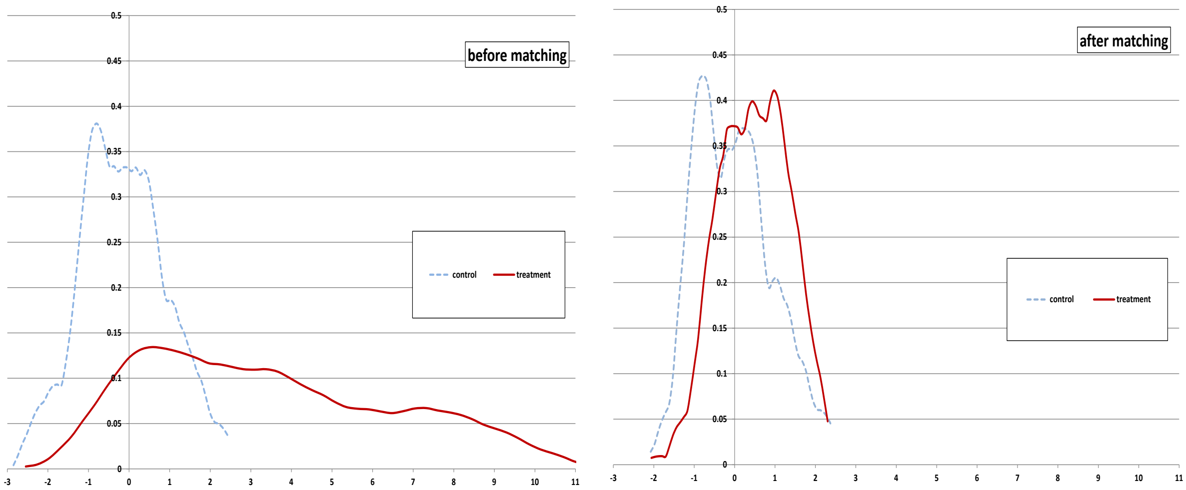


Note: PSM, Kernel Matching, typical source markets (Table 3A, column 3)

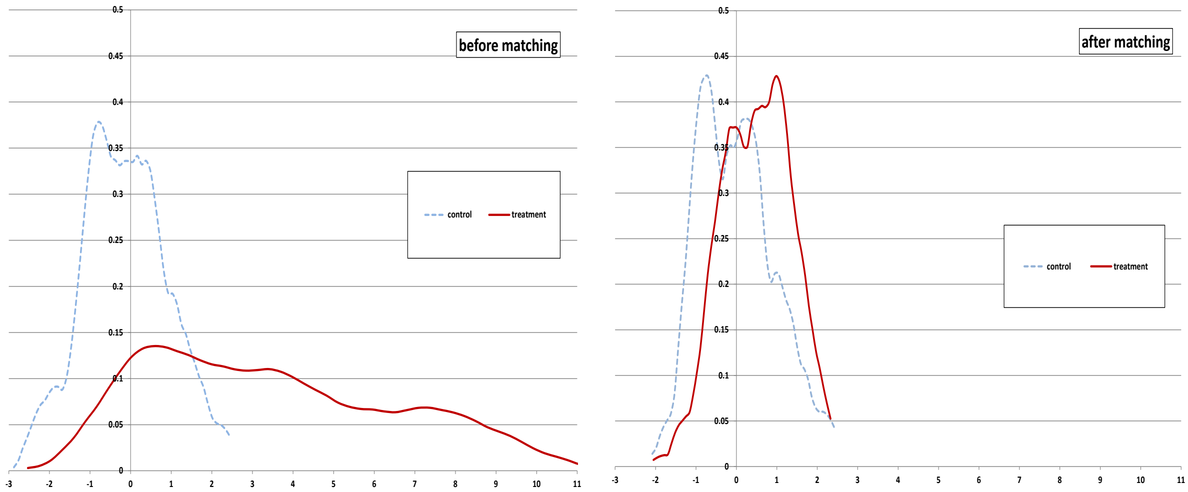
Figure A10 Common support between treatment and control group: transport costs



Note: PSM, Kernel Matching, all observations (Table 3B, column 1)



Note: PSM, Kernel Matching, only typical destination markets (Table 3B, column 2)



Note: PSM, Kernel Matching, only typical source markets (Table 3B, column 3)

Table A1 Number of observations and missings, by year

	price (p_j)		price difference (p_j-p_k)		transport costs (tc_{jk})	
	#	%	#	%	#	%
1999	753	53.6%	5386	29.5%		
2000	653	46.5%	4070	22.3%		
2001	651	46.4%	3961	21.7%	66	1.7%
2002	585	41.7%	3078	16.9%	326	2.8%
2003	681	48.5%	4124	22.6%	269	2.3%
2004	720	51.3%	4739	26.0%	71	0.6%
2005	747	53.2%	5076	27.8%	134	1.2%
2006	618	44.0%	4693	25.7%	87	0.8%
2007	820	58.4%	6702	36.7%	56	0.5%
2008					83	0.7%
2009					61	0.5%
2010					34	0.3%
all observations	6228		41829		1187	

Note to table: # indicates the number of available observations, and % indicates the share of observations in the total number of potential observations, under the assumption that the sample of markets or market pairs is representative. Hence, 100-% is the share of missings. In the case of price differences the table only reports the number of observations with a positive price difference.

Table A2 Missing observations: correlations with cell phone rollout

dependent variable: missing observations in price, price difference and transport cost data (binary)

	price		price difference		transport costs	
	(1)	(2)	(3)	(4)	(5)	(6)
cell phone dummy	0.0038	-0.0057	-0.0062	-0.0155	0.0187	0.0169
	(0.0132)	(0. 0637)	(0.0092)	(0.0095)	(0.0129)	(0.0134)
markets	yes	yes				
market pairs			yes	yes	yes	yes
trend by market	yes	yes				
trend by source			yes	yes	yes	yes
trend by destination			yes	yes	yes	yes
seasonality by market	no	yes				
seasonality by source			no	yes	no	yes
seasonality by destination			no	yes	no	yes
adj R^2	0.4295	0.4395	0.3405	0.3589	0.1176	0.1197
no. of observations	12177	12177	316602	316602	27146	27146

Note to table: Maize price data and transport cost data are respectively from January 1999 to December 2007 and from August 2001 to December 2010 (source: SIMA). Missing observations is a binary variable that takes the value 1 if an observation is missing and zero elsewhere, and under the assumption that the sample of markets or market pairs is representative. Equations are estimated using OLS. Robust standard errors in brackets below the coefficient are clustered by markets (1-2) and market-pairs (3-6). $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A3 Justifying priors: what are source and destination markets in maize trade?

markets	1	2	3	4	5	6
Pemba	73.0	0%	0.2%	2.9%	139	yes
Montepuez	73.0	21.4%	5.6%	0.0%	76	no
Lichinga	186.5	0%	2.4%	0.4%	142	no
Nacala	37.6	0.3%	0.0%	4.4%	206	yes
Monapo	37.6	29.4%	0.6%	0.0%	43	no
Angonia	117.6	40.0%	8.2%	0.0%	14	no
Cuamba	186.5	29.9%	0.4%	0.2%	79	no
Ribaue	37.6	21.4%	2.1%	0.2%	26	no
Nampula	37.6	0%	10.4%	10.8%	472	no
Alto Molocue	68.7	5.5%	22.5%	0.0%	42	no
Angoche	37.6	10.7%	0.1%	0.0%	90	yes
Milange	68.7	0%	0.5%	0.0%	30	no
Tete	117.6	0%	7.7%	10.5%	156	no
Mocuba	68.7	34.5%	3.4%	0.5%	169	no
Mutarara	68.7	35.3%	0.5%	0.1%	9	no
Quelimane	68.7	0%	0.2%	0.0%	193	yes
Gorongosa	175.8	36.7%	7.9%	0.4%	19	no
Manica	175.8	72.9%	3.8%	0.0%	36	no
Chimoio	175.8	84.4%	8.0%	3.4%	237	no
Nhamatanda	53.7	0%	12.4%	0.1%	26	no
Beira	53.7	0.3%	0.6%	13.5%	432	yes
Vilanculos	45.7	0%	0.0%	0.2%	37	yes
Massinga	45.7	13.2%	0.4%	4.2%	21	yes
Maxixe	45.7	0.3%	0.0%	10.2%	109	yes
Chokwe	47.0	1.4%	0.9%	0.7%	53	no
XaiXai	47.0	0%	0.7%	8.4%	116	yes
Maputo	15.3	0%	0.3%	20.1%	1095	yes

Note to Table: Column 1: per capita production in kg, 1999-2007, by province (source if > 65); 2: availability of producer price data, 1999-2009, weekly, by market. (source if > 15%); 3: source markets in transport cost data, 2001-2010, weekly, by market (source if >2%); 4: destination markets in transport cost data, 2001-2010, weekly, by market (destination if >2%); 5: population size in 2007, x1000, by market (destination if >100,000); 6: located on the coastline (destination if yes). Markets are ordered from north to south. Markets that align with most source (destination) market characteristics are printed in bold (italics).

Table A4 Using determinants of transport costs to verify and clean the data

dependent variable: nominal transport costs for bags of maize grain, different source and destination markets				
	(1)	(2)	(3)	(4)
road distance	53.3***	51.8***	49.6***	45.5***
	(5.5)	(5.4)	(6.9)	(6.5)
road quality	-651.1***	-643.8***	-913.2**	-745.0*
	(201.8)	(187.8)	(419.8)	(377.0)
weight of bags	208.5**	195.8**	426.4***	400.9***
	(89.4)	(85.5)	(80.2)	(87.0)
fuel price	247.9	-2811.0	3380.1***	yes[a]
	(548.6)	(1437.5)	(1083.9)	
consumer price index	1106.6	2986.1**	3967.8***	3603.6***
	(1067.9)	(1294.1)	(1362.8)	(1322.8)
trend	-477.1	-380.5***	-2595.3***	-2214.6**
	(611.8)	(627.8)	(882.9)	(827.6)
seasonality by source market (dummies)	no	no	yes	yes
year (dummies)	no	yes	yes	yes
R^2	0.4423	0.4636	0.6117	0.6438
sample period	8-01/12-10	8-01/12-10	8-01/12-10	8-01/12-10
number of observations	1135	1135	1135	1135

Note to table: See main text for source and construction of variables. Robust standard errors in brackets are clustered by market-pairs. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$. a. Fuel price interacted with source market.

Table A5 Impact of mobile phones on dispersion of maize prices: including source and destination specific trends and seasonality to the basic specification

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cell phone dummy	-228***	-109.8	-148*	-95.5	-135.1	-112.4	-151*	-153*	-191**
	(85)	(82.0)	(82.7)	(87.4)	(87.6)	(84.4)	(83.7)	(84.8)	(86.3)
trend by source	yes	yes	no	no	no	yes	yes	no	no
trend by destination	yes	no	yes	no	no	no	no	yes	yes
seasonality by source (dummies)	yes	no	no	yes	no	yes	no	yes	no
seasonality by destination (dummies)	yes	no	no	no	yes	no	yes	no	yes
market pairs (dummies)	yes	yes	yes	yes	yes	yes	yes	yes	yes
year-month (dummies)	yes	yes	yes	yes	yes	yes	yes	yes	yes
Adj R^2	0.498	0.436	0.447	0.443	0.463	0.448	0.468	0.460	0.478
No. of observations	39498	39498	39498	39498	39498	39498	39498	39498	39498

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. Equations are estimated using OLS. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A6 Impact of mobile phones on transport costs: including source and destination specific trends and seasonality to the basic specification

dependent variable: real transport costs of maize grain per ton-km

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
cell phone dummy	-674***	-474*	-396*	-149	-211	-371	-534**	-332	-392*
	(235)	(253)	(224)	(207)	(192)	(240)	(215)	(231)	(214)
trend by source	yes	yes	no	no	no	yes	yes	no	no
trend by destination	yes	no	yes	no	no	no	no	yes	yes
seasonality by source (dummies)	yes	no	no	yes	no	yes	no	yes	no
seasonality by destination (dummies)	yes	no	no	no	yes	no	yes	no	yes
market pairs (dummies)	yes	yes	yes	yes	yes	yes	yes	yes	yes
year-month (dummies)	yes	yes	yes	yes	yes	yes	yes	yes	yes
R^2	0.870	0.781	0.783	0.817	0.817	0.825	0.830	0.831	0.832
No. of observations	1090	1090	1090	1090	1090	1090	1090	1090	1090

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight. Equations are estimated using OLS. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A7 Impact of mobile phones on dispersion of maize prices: including market pair trends

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$			
	(1)	(2)	(3)
cell phone dummy	-221***	-310***	-471***
	(89.6)	(113)	(126)
trend by source	yes	yes	yes
trend by destination	yes	yes	yes
trend by market pair	**yes**	**yes**	**yes**
seasonality by source (dummies)	yes	yes	yes
seasonality by destination (dummies)	yes	yes	yes
market pairs (dummies)	yes	yes	yes
year-month (dummies)	yes	yes	yes
Adj R^2	0.512	0.526	0.509
No. of observations	39498	25568	20723

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. Equations are estimated using OLS. Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A8 Impact of mobile phones on dispersion of maize prices: basic specification, with logarithmic transformation of dependent variable

dependent variable: ln[real positive maize price difference between markets, per km $((p_{j,t} - p_{k,t})/d_{jk})$]			
	(1)	(2)	(3)
cell phone dummy	-0.164***	-0.213***	-0.291***
	(0.0465)	(0.0596)	(0.0655)
Adj R^2	0.433	0.439	0.422
No. of observations	39498	25568	20723

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A9 Impact of mobile phones on transport costs: basic specification, with logarithmic transformation of dependent variable

dependent variable: ln[real transport costs of maize grain per ton-km]			
	(1)	(2)	(3)
cell phone dummy	-0.150**	-0.170***	-0.157**
	(0.0610)	(0.0599)	(0.0773)
R^2	0.857	0.878	0.878
no. of observations	1142	995	809

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance and bag weight. All estimations include year-month and market pair dummies, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): full sample ; column (2) is based on a subset of market pairs with typical terminal markets as destination market, and column (3) on a subset of pairs with typical production areas as source markets. Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A10 First stage logistic estimation of propensity score price dispersion sample

Dependent variable: probability of having access to mobile phone technology (cell phone dummy)			
Sample	1	2	3
variables \ est. technique	logit	logit	logit
trend	1.229*** (0.014)	1.210*** (0.015)	1.122*** (0.015)
ln(population, pair)	0.968*** (0.023)	0.857*** (0.025)	0.980*** (0.026)
ln(poverty, pair)	-1.702*** (0.155)	-1.116*** (0.184)	-3.041*** (0.197)
ln(distance to big city, pair)	-0.049** (0.020)	-0.057*** (0.021)	0.088*** (0.027)
pseudo R2	0.6370	0.6330	0.6280
observations	39498	31265	29296

Note to table: Column (1): full sample ; (2) is based on a subset of market pairs that excludes typical producer areas / assembly markets as destination market, while column (3) is based on a subset of pairs that excludes typical terminal markets as source markets. Standard errors are in brackets below the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Transport cost sample

Dependent variable: probability of having access to mobile phone technology (cell phone dummy)			
Sample	1	2	3
variables \ est.technique	logit	logit	logit
trend	1.185*** (0.101)	1.165*** (0.102)	1.173*** (0.100)
ln(population, pair)	0.659*** (0.122)	0.627*** (0.122)	0.644*** (0.124)
ln(distance to big city, pair)	-0.219** (0.093)	-0.245** (0.095)	-0.220** (0.093)
pseudo R2	0.3621	0.3618	0.3616
observations	1090	1035	1065

Note to table: see Table above

Table A11 Standardised Bias of Covariates, before and after matching price dispersion sample

Sample	1		2		3	
before and after matching	before	after	before	after	before	after
Variables						
trend	2.543	1.920	2.571	1.916	2.532	1.865
ln(population)	0.350	-0.036	0.363	-0.040	0.358	-0.039
ln(poverty)	-1.161	-0.869	-1.152	-0.851	-1.239	-0.883
ln(distance to big city)	-0.087	0.072	-0.068	0.093	-0.009	0.104

Note to table: 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 (Table 3A and 3B).

Transport cost sample

sample	1		2		3	
before and after matching	before	after	before	after	before	after
variables						
trend	1.458	0.358	1.450	0.302	1.465	0.325
ln(population)	0.609	0.355	0.625	0.359	0.598	0.331
ln(distance to big city)	-0.467	-0.201	-0.500	-0.232	-0.468	-0.196

Note to table: see table above.

Table A12 Impact of mobile phones on price dispersion: PSM, Nearest Neighbour

dependent variable: real positive maize price difference between markets, per km $((p_{j,t} – p_{k,t})/d_{jk})$

	(1)	(2)	(3)
ATT	-147.8** (62.1)	-189.7** (87.4)	-151.3** (65.5)
ATU	464.8	79.9	220.1
ATE	71.0	-115.6	-13.6
treated, on support	12097	3617	8629
treated, off support	2195	10970	1398
untreated, on support	6723	1370	5088
untreated, off support	7696	8772	5385
no. of observations	28711	24729	20500

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Nominal series are deflated with the consumer price index, and divided by road distance. Equations are estimated using Propensity Score Matching with Nearest Neighbour (N=20 and caliper:0.001; data restricted to 1999-2005, see main text for further details). Column (1): full sample ; (2) is based on a subset of market pairs that excludes typical producer areas / assembly markets as destination market, while column (3) is based on a subset of pairs that excludes typical terminal markets as source markets. Standard errors are in brackets below the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A13 Impact of mobile phones on transport costs: PSM, Nearest Neighbour

dependent variable: real transport costs of maize grain per ton-km $(tc_{ijk,t}/(d_{jk}.w_{i,t}))$

	(1)	(2)	(3)
ATT	-586.4*** (226.3)	-535.2*** (113.0)	-626.2*** (226.7)
ATU	-548.7	-382.1	-596.9
ATE	-570.6	-472.8	-613.8
treated, on support	312	317	304
treated, off support	498	447	485
untreated, on support	225	218	224
untreated, off support	55	53	52
no. of observations	1090	1035	1065

Note to table: See also table above. Transport cost data are from August 2001 to December 2010 (source: SIMA). Nominal series deflated with the consumer price index, and divided by road distance and bag weight. Equations are estimated using Propensity Score Matching with Nearest Neighbour (N=5 and caliper:0.02).

Table A14 Impact of mobile phones on maize production and maize area

Dependent variable: ln(aggregate maize area by province), $\ln(A_{j,t})$		
	(1)	(2)
number of mobile phone towers	0.004511***	0.003260***
	(0.00136)	(0.00094)
province (dummy)	yes	yes
year (dummy)	yes	yes
sample period	from 1993/94 to 2013/14	from 2002 to 2014
R^2	0.9434	0.8754
no. of observations	173	67
Dependent variable: ln(aggregate maize production by province), $\ln(Q_{j,t})$		
number of mobile phone towers	0.012262***	0.00905***
	(0.00323)	(0.00220)
province (dummy)	yes	yes
year (dummy)	yes	yes
sample period	from 1993/94 to 2011/12	from 2002 to 2014
R^2	0.9255	0.9326
no. of observations	132	80

Note to table: Aggregate maize area by province and aggregate maize production by province are annual data, respectively from 1993/94 to 2006/07 (column 1; source: Ministry of Agriculture, Early Warning Unit (Aviso Previo)[28]), and from 2002 to 2014 (column 2; Source: Trabalho de Inquérito Agrícola (TIA), taken from Anuário de Estatistica Agararia, 2002-2011 and 2012-2014). The former series are national aggregates, the latter series pertain to aggregates over small and medium sized farms. In these latter series (TIA series) various years of observations are missing (8 years available). The variable *number of mobile phone towers* is simply a count of the number of mobile phone towers, by year and province. All estimations include year and province dummies. The longer period estimations (column 1) also include trends by province. Equations are estimated using OLS. Robust standard errors in brackets below the coefficient are clustered by province. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[28] These series are taken from Kayser and Arlindo, 2007.