TI 2019-047/V
Tinbergen Institute Discussion Paper

Mobile Phones and Mozambique Traders: What is the Size of Reduced Search Costs and Who Benefits?

Wouter Zant[1]

[1] Vrije Universiteit Amsterdam; Tinbergen Institute

Mobile Phones and Mozambique Traders:

What is the Size of Reduced Search Costs and Who Benefits?

Wouter Zant*

Abstract

We investigate to what extent the roll-out of the mobile phone network in Mozambique reduced transport costs and search costs, and thereby decreased spatial price dispersion and improved market efficiency. Estimations are based on data of transport costs of maize grain and maize market prices. The mobile phone rollout explains a 10%-13% reduction in maize price dispersion. Around half of this reduction is associated with search costs related to transport, the other half with other search costs, for example for the collection of maize in source markets. Search costs are substantial and also a substantial component of total transport costs. Benefits of increased market efficiency are biased towards 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 thank Jenny Aker, Taryn Dinkelman, Benedito Cunguara and conference and seminar participants at UCLA in Berkeley (ABCA2017), Lissabon (EEA-ESEM2017) and the VU Amsterdam for useful comments on previous versions. All remaining errors are my own.

Introduction

Price information is a major requirement for the efficient operation of agricultural markets as it drives the behavior of traders. In sub-Sahara Africa, access to price information used to be costly due to long distances, poor transport and communication infrastructure and elementary developed marketing institutions. Traditionally, information on maize prices across markets in sub-Sahara Africa is collected by traders travelling to markets, through word-of-mouth and through personal and professional networks. In the Mozambique case – the country of our case study – fairly reliable information on agricultural prices and markets is collected and disseminated – already for a number of decades – on a weekly basis by Sistema de Informação de Mercados Agrícolas (SIMA). The introduction of mobile phones in the late 1990s has drastically changed access to information. 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. The newly available mobile phone technology allows traders to assess maize prices in many distant markets instantaneously, efficiently, at low costs and customized to personal needs. Improved information has lowered search costs, leading to a reduction of transport costs and to a reduction of price dispersion across markets. This paper investigates if and to what extent the decrease in the costs of information due to the introduction of mobile phones, has reduced search and transport costs and has improved the efficient operation of markets. In particular we estimate the impact of mobile phones in Mozambique on grain transport costs and on the spatial dispersion of maize 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 spatial 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). Even without price impacts, these efficiency gains involve substantial welfare improvements. Muto and Yamano (2009) investigate marketing costs of maize and bananas during the introduction of mobile phones in Uganda, using household survey data for 2003 and 2005. They 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 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 between year variation in cow peas prices is reduced.

The current study contributes to this literature by replicating the estimations of the impact of mobile phones on spatial price dispersion for Mozambique using prices for maize, a crop that is key to food security. Similar to previous work we show that spatial price differences decreases with the introduction of mobile phones. Unlike previous work, the current study investigates, explains and quantifies the different impact of mobile phones on price dispersion and transport costs, quantifies relative size of search costs in transport and distance related transport costs, and distinguishes collection and transport related search costs, by combining spatial price differences and transport costs data. 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 to what extent producers or consumers capture the benefits of improved market efficiency. For the impact estimations we use data on maize grain transport costs by itinerary and weekly recorded data of retail market prices of white maize grain, respectively for the period 2001-2010, and 1997-2007 (source: SIMA). These core data are complemented with data on distance between markets, population, rainfall, fuel prices, and consumer prices. With the exception of mobile phone rollout and rainfall data, all data are obtained from public domain sources. For the identification of the impact of mobile phones on dispersion of agricultural prices, we use the rollout of the mobile phone infrastructure. A difference-in-difference approach (DiD) with time and trade-pair fixed effects is applied to estimate impacts. Trade-pairs with and without mobile phone technology are shown to satisfy the common trend assumption, both in spatial price differences and in transport costs. Since the roll-out is unlikely to be random, we address possible selection bias by complementing the impact estimations with propensity score matching. We find that the introduction of mobile phones in Mozambique has reduced maize price dispersion by 10%-13%. Around half of this reduction is associated with transport related search costs, the other half with collection of maize related search costs. The benefits of Increased efficiency of maize markets are further shown to be biased towards consumer markets.

The rest of this paper proceeds as follows. Section 1 presents the background on maize marketing, maize trade and maize prices in Mozambique, and discusses the introduction of mobile phones in Mozambique. Section 2 sets out the conceptual framework, discusses data and data sources, and elaborates the empirical strategy. Section 3 presents the impact estimations and robustness checks. Section 4 measures benefits in source and destination markets. Section 5 highlights potential threats and alternative explanations. Section 6 presents the summary and conclusion.

1. 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

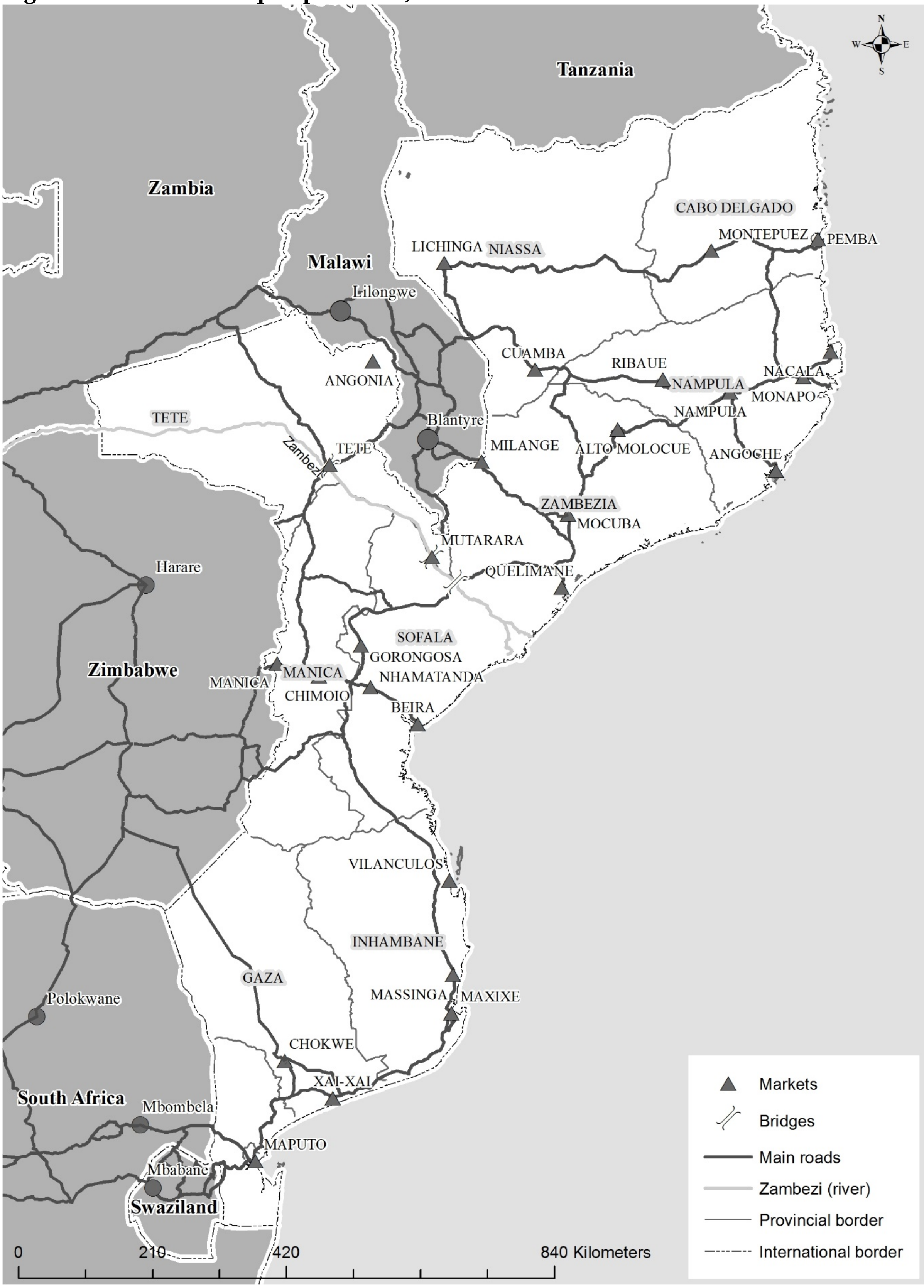
[1] Staples in Mozambique are maize, rice, cassava, wheat, sorghum, millet, sweet potatoes beans and groundnuts.

[2] The empirical work in this paper is based on data for the period from the end of the 1990s to 2007, and with a few extensions to 2010. This explains relevance and justifies reference to slightly older policy reports and articles.

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 A6, Table A3).

Domestic production of maize is concentrated in the central and northern part of Mozambique (see Figure 1). 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 A6). Most agricultural production in Mozambique is rain-fed. Extreme weather like drought and flooding cause additional fluctuations 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). Tropical storms in early 2019 also devastated the 2018-2019 crop in large parts of the country. Due to widespread subsistence farming only a limited share of production (around 30% of total production) is traded on the market. 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. 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 Figure 1).

Figure 1 Mozambique: provinces, markets and roads



Source: VU SPINlab

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, partly underscoring the large size of the country (Dominguez-Torres and Briceño-Garmendia, 2011). From the early 1990s onwards the percentage of roads in good or fair condition has increased from 30% to 83%, which is above the average of other Sub-Saharan low-income countries. However, accessibility to rural areas 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 originates from agriculture. Moreover, the condition of the rural network – 40% of the rural roads is in poor condition – stands 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 condition, but secondary, tertiary and rural roads are in poor condition, and especially during the rainy season many of these roads cannot be used. In summary, the infrastructure of trunk roads in Mozambique connects cities and major towns reasonably well and is no impediment to domestic trade and to the free flow of agricultural produce between these locations.

Trade in maize grain – the standard white maize grain quality[4] – takes place throughout Mozambique. However, the Zambezi river (see Figure 1) creates a natural barrier to domestic

[3] All numbers on road infrastructure sourced from Dominguez-Torres and Briceño-Garmendia, 2011.
[4] White maize grain is produced, consumed and traded throughout Mozambique and is the dominant type of maize. Since it traded from north to south and from east to west, and without denying possible quality differences, it is reasonable to assume that white maize grain is a homogenous product throughout Mozambique and over time. Homogeneity over time and across locations is an assumption in the conceptual framework.

trade[5]: 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[6]. 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 northwest, 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 which are recorded for itineraries where trade of maize grain takes place, and which are used in the current study (source: SIMA), confirm these stylized facts (see also Appendix, Table A3).

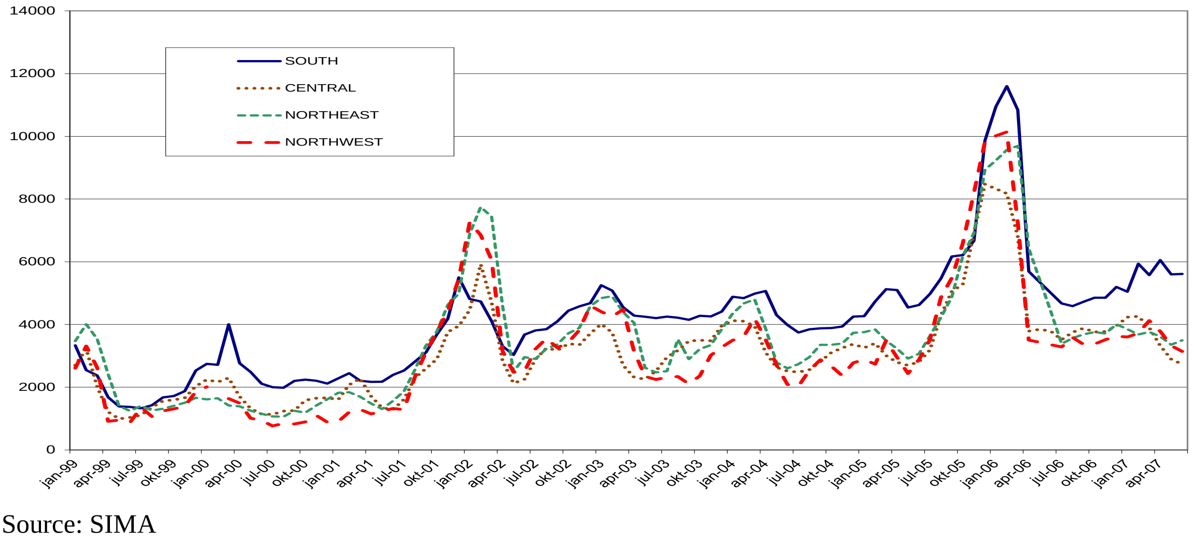
The trading sector consists of itinerant traders, large scale assemblers, wholesale traders, millers and retailers. Retailers and millers are at the end of the value chain and are primarily involved in earning returns by value addition rather than earning returns on trade and transport. Wholesale traders take an intermediate position: they buy from assemblers in source regions and supply to mills of various sizes in urban areas. This activity may entail gains from price differences between geographically dispersed markets, but is likely to have a large

[5] 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, 2019a).

[6] 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!).

component of value added through collecting, sorting, quality control and distribution. The key agents in Mozambique 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). Consequently, 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, and maize trade is less common long beyond the post-harvest months. 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. However, it is likely that working capital is a constraint to business. To our knowledge there is no information on actual trade flows of white maize grain or on the number of traders actively involved in maize trade in Mozambique.

Figure 2 Maize price by region (nominal prices, January 1999-July 2007)



Source: SIMA

Maize prices over time (see Figure 2) 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 A3 and A4) 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 A4). Deficit urban areas, generally, have higher price levels and lower price volatility compared to surplus rural areas (see Appendix Figure A5). This pattern aligns with simple models with fixed per kg transaction costs (Fafchamps and Vargas-Hill, 2008). Seasonality in maize prices also generates –and this is particularly relevant for this research – seasonality in spatial price differences between source and destination markets (see Zant, 2019a). The seasonal fluctuation in spatial price differences suggests both a period during which arbitrage is most profitable and an apparent lack of traders or trade flows to fully exploit this.

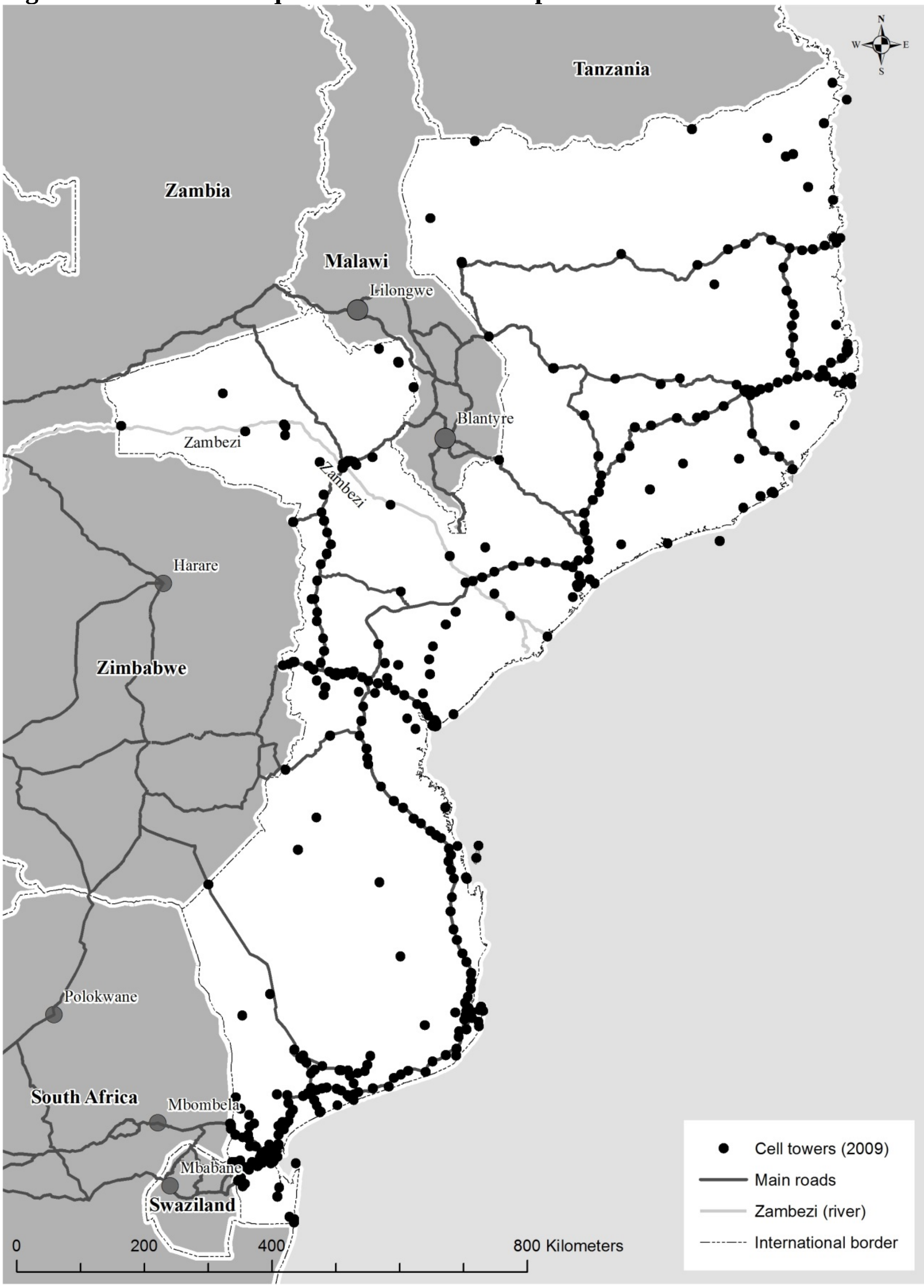
Mobile phone rollout

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 years following the introduction 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 and high per capita income, combined with low construction and maintenance costs for cell phone towers. Visual inspection of the roll-out map (see Figure 3) 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 per km2 land area and per head of the 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. Since the determinants of rollout are key to understanding selection bias in rollout, we elaborate more formally on these variables in the context of the propensity score estimation.

In the 2000s average mobile phone network density[7] in Mozambique as whole increased 5 to 6 fold[8]. 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. A modest share according to western standards, 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 widespread promotion of the pre-payment system which solved the cashing problem – a major 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[9], it is likely that use and access to mobile phone services is biased against the poor.

[7] 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.
[8] Figure 3 shows the 2009 network, the final year of our mobile phone network data.
[9] 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).

Figure 3 Mozambique: network of mobile phone towers in 2009



Source: VU SPINlab

2. Conceptual Framework, Data and Empirical Strategy

Conceptual framework

The mechanism underlying the impact of mobile phone services on price dispersion and transport costs is associated with the efficiency of traders in agricultural commodities in collecting merchandise at source locations and finding attractive arbitrage opportunities, and the efficiency of drivers and transporters in organizing transporting activities. Traders in agricultural commodities monitor prices of agricultural prices in various markets, both source and destination, 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[10]. 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 helps in optimizing trade decisions.

Transporters earn an income from selling transport services. Transported merchandise could be any merchandise, but in the current developing country context concerns transport of agricultural commodities. Similar to the case of the commodity trader, transporters monitor 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. Unlike the situation for the trader / arbitrageur there is no publicly accessible source of information (like SIMA) that records and disseminates information on potential freight. Consequently

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

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 reductions in transport costs 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.

The objective of the current paper is empirical: therefore, it is beyond the scope of this paper to develop a complete model of price formation and spatial arbitrage in agricultural markets. Instead we develop a few elements of a framework to guide the empirical estimations (partly copied from and inspired by Fafchamps and Vargas-Hill, 2008). As the turn-around time in developing country domestic trade is typically short (less than a month, see for example Fafchamps, Gabre-Madhin and Minten, 2005), and to keep the framework simple, we ignore price risk. The starting point is arbitrage and transaction costs under perfect competition. The difference between price at source and price at destination reflect search costs, transport costs, storage costs and processing costs of traders. In the current set-up we also ignore storage and processing costs and focus on search and transport costs. A standard arbitrage condition under perfect competition requires:

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

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*. We identify two types of marketing tasks, or specialized agents in the marketing chain. The first concerns traders who search for maize by traveling around in rural areas, looking for attractive maize offers on local markets or finding farmers with surplus maize for sale, with the eventual goal to sell this maize in distant high price consumer markets. The second concerns transporters who are responsible for the logistics of the physical long-distance transport from source markets to destination markets. Hence, we have formally:

(2) $$p_j - p_i = c_{ij} + \tau_{ij}$$

where c_{ij} is the (per unit) costs of traders associated with collecting maize from local markets and farmers, and τ_{ij} is the (per unit) cost associated with total cost of transport of moving maize grain from location i to location j. Next, we assume that both tasks – collecting maize in source markets and transport of maize grain – entail search costs: for the first task this is straightforward. In fact, for this task the only costs are search costs. For the second task, search involves identification of potential flows of merchandise, selection of most attractive transport jobs, planning of efficient collection of full truckloads and making arrangements for return cargo. The other major component of total transport costs are the cost associated with traveling from source to destination. We follow Fafchamps and Vargas-Hill (2005) for the introduction of search costs in this framework. Assume that for each quantity q of maize collection costs are cq, where these costs are proportional to the time spent on search. The number of traders is N_c, where the subscript refers to the collection task. Assume that more traders increase search costs of collecting maize from farmers. Let the probability of finding a farmer with maize for sale (or a seller on a local market) be $1/N_c$ per unit of time and the search cost per a unit of time is θ, than the per unit cost of collecting q is $c = \theta N_c$. The search costs component for transporters is constructed likewise. Now costs of each quantity q of maize transported is τq. Since total transport costs are determined both by search and by travel, the total per unit cost of transporting q is:

(3) $$\tau = \varphi N_\tau + \gamma d,$$

where d is the distance between source and destination, and γ is the per km per kg travel cost. We may re-write the arbitrage equation as:

(4) $$p_j - p_i = [(\theta N_c)_{ij} + (\varphi N_\tau + \gamma d)_{ij}]$$

And if collectors and transporters perform both tasks and, hence, are the same ($N_c = N_\tau = N$):

(5) $$p_j - p_i = [((\theta+\varphi)N)_{ij} + \gamma d_{ij}]$$

From equation (5) and equation (3) we learn that distance related transport cost appears identically and equally sized, in per unit transport costs and spatial price differences. We exploit this property when we confront the conceptual framework with the data.

Solving equation (5) for N generates[11]:

(6) $$N = [(p_j - p_i) - \gamma d_{ij}] / (\theta+\varphi)$$

Hence, an improvement in search technology, for example through the introduction of mobile phones, reduces search costs for both collecting maize and transporting maize ($\theta+\varphi$), and – given prices on source and destination markets and travel costs – will attract more traders to the trading business. Alternatively, if the number of traders is fixed (for example, in the short run, or because of working capital constraints) and under competitive conditions, this could decrease the price difference between destination and source. The prevalence of large unexploited arbitrage opportunities (see Zant, 2019a) is an indication that the number of traders is not adjusting swiftly.

Data and data sources

The data on the rollout of mobile phone infrastructure, sourced from the Ministry of Transport and Communication of Mozambique[12], contain 547 names of locations of mobile phone towers, their corresponding latitude and longitude coordinates and the first year of operation.

[11] Expected trader profit is $\pi = (p_j - p_i)q - [((\theta+\varphi)N)_{ij} + \gamma d_{ij}]q$; and under free entry the number of traders increases to the point where $\pi = 0$; Inserting this and solving for N yields equation (6).

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

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 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[13]. 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 confirm the effectiveness of the SIMA price information[14]. 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

[13] 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.

[14] 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/).

de milho branco), recorded for 27 markets[15], from January 1999 to December 2007. White maize grain is the dominant quality of maize produced, traded and 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 correlated with the season and with occasional droughts and reported by SIMA staff to be due to a lack of transactions[16]. There could be a concern that the missing observations are correlated with mobile phone status. Intuitively, this is unlikely: missing observations are due to seasonality and have nothing to do with mobile phone rollout. Also formal tests 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 as maize prices (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[17] and by the weight of the bags transported. Total transport costs are

[15] 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. See Figure 1 for the locations of these markets in Mozambique.

[16] The weekly SIMA bulletins are available for all weeks. Also, prices of other crops are, generally, available during the market weeks in which maize prices are missing. Both facts further support and strengthen the claim that the price survey was run every market week, but that maize was occasionally unavailable in the market.

[17] Differences in volumes for different crops justify separately recorded per kg transport costs.

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).

Then, a number of miscellaneous variables from different sources are used. Distance in kilometres, both road distance and Euclidian distance ("as the crow flies"), and traveling time in hours, are taken from GoogleMaps, accessed at the time of writing the first version of this study (2016)[18]. 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[19]. We use these data to capture shocks on the supply side due to flooding or drought. Data on population by city and district are from three censuses (August 1997, September 2007, July 2016), published by Instituto Nacional de Estatistica Moçambique. Monthly population series are constructed by interpolation. Population is used as an approximation of (relative) demand in destination locations, and population density in source districts to approximate the ease of finding farmers with surplus maize. 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, and used as covariates in the estimations. 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.

[18] Note that changes in road infrastructure during the period of study (1999-2007) are absorbed by time fixed effects in the difference-in-difference estimation (see *Empirical strategy*).

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

Empirical strategy

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

(7) $$y_{jkt} = \beta_0 + \beta_1\, cell_{jkt} + \boldsymbol{X}_{jkt}\gamma + \eta_{jk} + \theta_t + \varepsilon_{jkt}$$

where y_{jkt} is the spatial price difference or the transport costs between markets k and j, $cell_{jkt}$ 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 $\boldsymbol{X}_{jkt}$ represents variables that influence price dispersion and transport costs, such as drought and flooding in sources markets, fuel prices reflecting trade costs and differences in demand in destination markets due to population size and income. Parameters η_{jk} and θ_t represent market pair and time fixed effects, and ε_{jkt} 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.

Next, we introduce a minor adjustment to the above difference-in-difference specification. We approximate the trade-pair fixed effect with road distance, assuming that that road distance between market pairs fully captures market pair fixed effects, and thereby maintains the DiD character of the estimation. Simultaneously, it yields a coefficient with a straightforward economic interpretation – per kilometre per kg travel cost – that nicely fits the conceptual framework: The cross-sectional variation in road distance helps to measure the contribution of road distance to transport costs and price dispersion, to measure the size of distance-related costs of transport vis-à-vis search costs, to match transport costs estimations with spatial price dispersion estimations and to disentangle search costs for transport and search costs for collecting maize. In formula, we have:

(8) $$y_{jkt} = \beta_0 + \beta_1\, cell_{jkt} + \beta_2\, roaddistance_{jk} + \boldsymbol{X}_{jkt}\gamma + \theta_t + \varepsilon_{jkt}$$

where $roaddistance_{jk}$ is the distance by road between source k and destination j in kilometres. Equation (8) is the basic specification for estimations.

As spatial price dispersion is partly determined by transport costs (see equation (2)), we could further exploit the spatial price difference and transport cost data. We could regress price dispersion adjusted for per kg transport costs on the cell phone intervention variable, jointly with fixed effects. Implementing this estimation automatically matches the different estimation samples. In formula (and for convenience omitting trends and seasonality by source and destination) this is:

(9) $$[(p_{kt}-p_{jt}) - transport\ costs_{jkt}] = \beta_0 + \beta_1\ cell_{jkt} + \eta_{jk} + \theta_t + \varepsilon_{jkt}$$

A positive significant impact (β_1) offers another estimate of the size of non-transport related search costs, or search costs associated with collection, and should be similar in size to the difference of the cell phone coefficients of specification (8), estimated, respectively, for spatial price difference and transport costs.

To capture geographically diverging price, quality, technology and network developments over time, we have included source and destination specific trends to the DiD specification. Likewise, the strong and geographically diverging seasonality in maize prices (see previous section and Appendix) is controlled for by including source and destination specific monthly dummies in the estimations. Estimations confirm that including source and destination specific trends and seasonality substantially improve the performance of the estimations (not reported, available from the author).

Finally, once a reduction in spatial price differences by a reduction is search costs is supported by the evidence, it is informative to investigate how mobile phones impact on price levels in source markets and in destination markets. In other words, is the reduction in spatial price differences associated with a (relative) increase in source market prices, or a (relative) decrease in destination markets prices, or both. We estimate the following equations:

(10) $p_{j,t} = \beta_0 + \beta_1\, cell_{j,t} + \boldsymbol{X}_{j,t}\gamma + \eta_j + \theta_t + \varepsilon_{j,t}$

(11) $p_{k,t} = \beta_0 + \beta_1\, cell_{k,t} + \boldsymbol{X}_{k,t}\gamma + \eta_k + \theta_t + \varepsilon_{k,t}$

where subscripts j and k now refer to source and destination markets.

Shocks in demand, supply and trade may affect measured impact and, consequently, we should control for these factors by including covariates ($\boldsymbol{X}_{jkt}$). We have experimented with a variety of variables: the size of population by city (market) to account for differences in demand in source and destination markets, or the sum of these to account for the likelihood of trade; various transformations of rainfall, reflecting drought and flooding, as key determinants of supply shocks, in view of the predominantly rain-fed nature of agriculture[20], where the influence of these rainfall shocks 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.

Empirical issues: sample

To begin with, we measure price dispersion as the positive maize price difference between market j and k in period t, in formula $p_{k,t} - p_{j,t}$ if $p_{k,t} > p_{j,t}$, for pairs of markets for which j≠k, and assume that the higher price pertains to destination markets and the lower price to source markets. The good part of this step is the resulting impressive and huge sample of more than 40,000 observations of spatial price differences that unambiguously tracks (reversals in) arbitrage returns (see Appendix, Table A1). The bad part is that only a subsample of these spatial price differences reflect transport costs and search costs[21]. How to find this subset of

[20] Drought is specified as the (log) of a threshold rainfall relative to actual rainfall, conditional on below threshold rainfall levels. 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. 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.

[21] No trade in spatial price equilibrium arises if spatial price differences are too low to cover trade costs, or if spatial price differences are higher than needed to cover trade costs but reflect (other) constraints to trade. In the literature the parity bound model offers a technique to distinguish these regimes (see Baulch, 1997; Zant, 2013).

observations? This is where the transport cost data are helpful: transport costs are only recorded if trade and transport truly has taken place. Hence, as a first step we have matched the sample of the spatial price differences to the sample of transport cost data, by trade pair and year[22]. Next, since the transport cost data are fragmented and only available for a limited number of trade routes and periods, and as a result small in number, we are keen to strengthen the statistical power of the spatial price difference estimations by adding additional observations to the matched sample.

To realise a larger sample we exploit two key observations[23]: 1. The presence of typical source markets and typical destination markets and 2. The observation that long distance trade is restricted to a few months after harvest. We use the number of records of source and destination markets in transport cost data to identify typical source and destination markets. In this context we immediately add that there are only a few markets that are strictly source markets or strictly destination markets. Most markets play both roles: they are source vis-à-vis some markets, and destination vis-à-vis other markets[24]. To complicate things further: roles may switch, both within and between years. As a result markets identified as either (strict) source or (strict) destination are only crude approximations. To stay on the safe side we have therefore limited their number[25]. Nevertheless, decisions on these source and destination markets are, generally, in line with complementary information on long run values of per capita production, the availability of data on growers' prices, population size by market and location

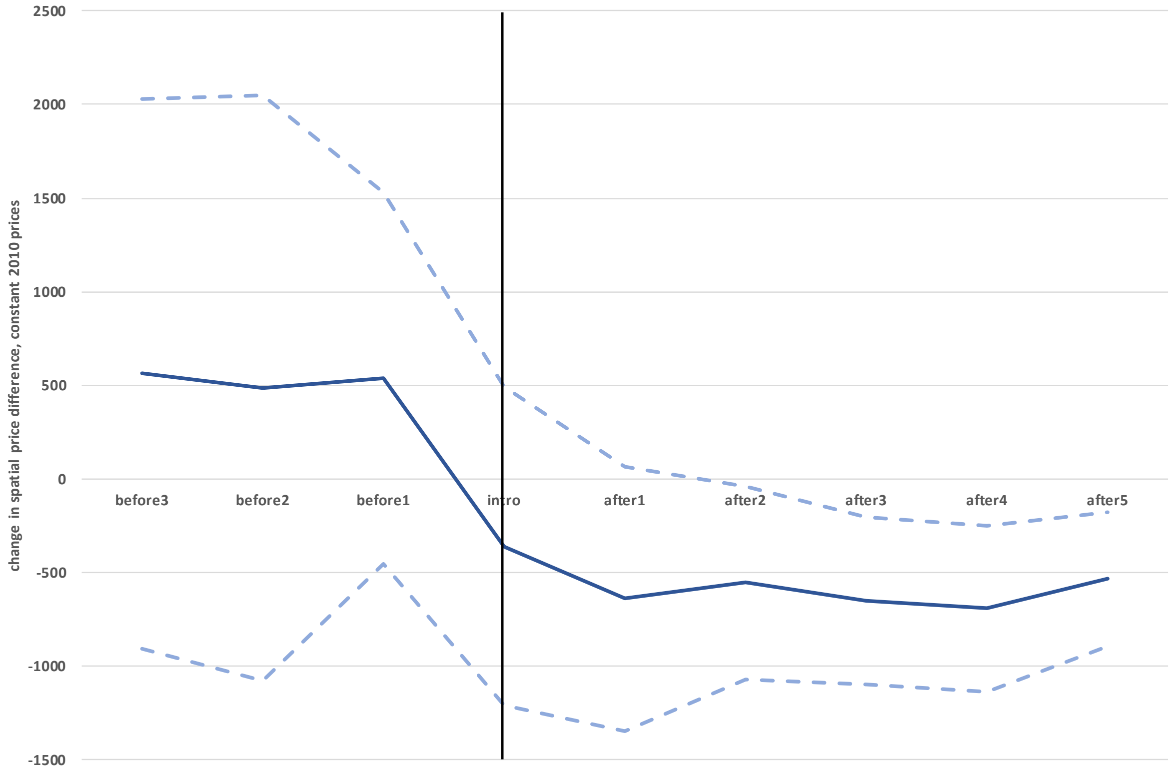
[22] The [market pair–year] observations in the transport cost data are not claimed to exhaustively cover [market pair–date] combinations with actual trade. The constructed sample is only a fraction of the (relevant) sample of [market pair–date] combinations with actual trade.
[23] This procedure is not exactly advanced scientific inference, but it is certainly not arbitrary: we exploit the stylized facts that are observed in the data and that are known from the regular seasonal pattern of trade and arbitrage returns (see also Zant, 2019a).
[24] Examples are Tete, Nampula and Chimoio: maize in the Tete market is sourced from Angonia and Chimoio, and further transported to Chimoio, Maxixe, Maputo, and Massinga. Likewise, maize in the Nampula market is sourced from Alto Molocue and Mocuba, and further transported to Maxixe, Maputo, Xai-Xai and Beira.
[25] Source markets: Alto Molocue, Angonia, Gorongosa, Lichinga, Manica, Mocuba, Montepuez, Nhamatanda and Ribaue; Destination markets: Beira, Massinga, Maputo, Maxixe, Nacala, Pemba, Quelimane, and Xai-Xai.

(see Appendix, Table A3), and price volatility (see Appendix, Figure A5), and closely correspond with the stylized facts from the Mozambique maize marketing reports (see e.g. Abdula, 2005; Tschirley et al., 2006, FEWSNET, 2010). Additionally, we restrict the additional observations to particular months: most farm households sell maize directly after harvest, during a restricted time span, often not longer than three to five months. Evidence from household surveys indicates that more than 80% of all maize grain transactions take place during five consecutive months, from June to October. Price data suggest a two months earlier start of the period in which trade takes place (see Appendix, Figure A3). Restricting added observations to a short after harvest period aims to capture observations that represent regular trade flows and to avoid observations that reflect uncommon arbitrage returns and potential trade reversals.

Empirical issues: Testing equality of pre-intervention trends of treated and non-treated

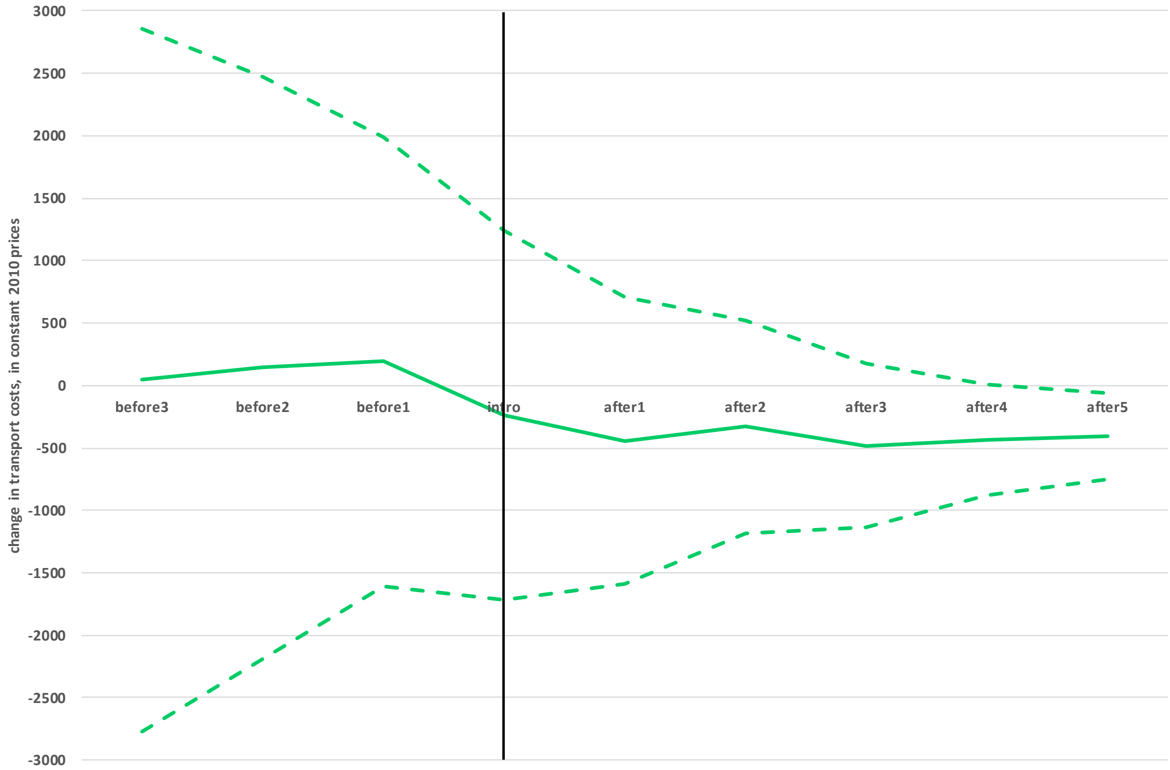
The DiD approach requires that pre-intervention outcomes of intervention and control groups have a common trend. It is most popular to show graphically that this parallel trend assumption is satisfied. Jointly with the parallel trend, the graphical evidence also reveals the dynamic path of impact, showing whether impacts are stable and persistent, and decrease or increase over time.

Figure 4 Testing for a common trend in the pre-introduction period: price difference



Note: the dotted lines indicate 95% confidence intervals.

Figure 5 Testing for a common trend in the pre-introduction period: transport costs



Note: the dotted lines indicate 95% confidence intervals.

Since all market-pairs in our sample obtain access to mobile phone facilities in the course of the rollout, there is no strict distinction between intervention and control groups. To construct the required information for a common trend test we estimate a slightly adjusted basic DiD specification: the impact variable is substituted for a set of annual indicator variables reflecting the number of years before and after the introduction of mobile phones. Hence, if d(year0) is an indicator variable with a value of 1 in the year of introduction and zero elsewhere, than d(year-1) is an indicator variable with a value of 1 one year before introduction and zero elsewhere, etc. If the pre-introduction trends of price dispersion or transport costs between with and without mobile phones are the same, then the pre-introduction coefficients should be insignificant: the difference in differences is not significantly different between the two groups in the pre-treatment period (see Autor, 2003, for an application of this test).

The results of the common trend test, shown in Figure 4 and Figure 5, confirm that both spatial price differences and per kg transport costs are on a common trend before the introduction of mobile phones: all "before" coefficients are not statistically different from zero. The Figures further support a statistically significant negative impact after the introduction of mobile phones. Impacts also appear to be stable over time. Finally the figures suggest that the impact on price dispersion is approximately twice as large as the impact on transport costs.

Empirical issues: addressing potential selection bias

For the estimation of equation (8) and (9) 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 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.

3. Empirical Estimation and Robustness Checks

Impacts on spatial price differences and transport costs

We first report estimations of a basic DiD specification that includes time fixed effects and road distance – the market pair fixed effect – but ignores covariates (see *Empirical strategy,* equation (8)). Additionally, the specification includes seasonality and trend, by source and destination. Column (1) in Table 1 reports the estimation based on a sample that matches market pair and year of the transport cost data. Columns (2) to (4) extend the sample of column (1) with trade pairs with typical destination markets (2), typical source markets (3), and the combination of these (4). When adding market pair observations with typical destination (source) markets (column (2) and (3)), we jointly excluded these markets as source (destination) markets. Following standard practice (see Bertrand et al., 2004) we report robust standard errors, clustered by market pair.

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

dependent variable: positive maize price difference between markets ($p_{kt}-p_{jt}$, $p_{kt}>p_{jt}$) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-828.2***	-767.5***	-785.0***	-813.1***
	(243.8)	(172.7)	(114.1)	(171.1)
road distance	1.654***	1.075***	0.403***	1.279***
	(0.504)	(0.250)	(0.083)	(0.392)
adj R^2	0.653	0.630	0.648	0.643
no. of observations	3686	5832	5794	4659
mean dependent variable:				
before intervention	2053.9 (451)	2127.3 (775)	2009.9 (960)	2154.9 (689)
after intervention	2261.3 (3235)	2141.7 (5057)	2127.6 (4834)	2276.7 (3970)

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Prices are converted to constant 2010 prices using the national consumer price index. All estimations include time fixed effects, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1): the sample is matched by market pair and year to the available transport cost observations; in column (2) market pairs with typical terminal markets as destination market are added to the sample of column (1); in column (3) market pairs with typical production areas as source markets are added to the sample of column (1), and column (4) combines (2) and (3). Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The results in Table 1 show a statistically significant reduction in price dispersion between markets as a result of the introduction of mobile phones, in all estimations. Expressed as a percentage of average pre-mobile-phone prices the reduction in spatial price differences has, at the margin, a size of 10%-13%[26]. The size of the reduction is (statistically) similar, but slightly lower when more observations are included to the estimation sample. The coefficient of road distance is also statistically significant in all estimations, but also decreasing in size when more observations are added to the sample. The estimated coefficients reflect per kg and per km transport cost and thereby shed some light on the relative size of pure transport cost vis-à-vis search costs: if we consider a trade transaction between market pairs that are 1000 km apart (for example Montepuez – Nacala: 413; Nhamatanda – Maputo: 1107; Alto Molocue – Maputo: 1822km), search costs are 57%-60% of trade costs. Their relative size obviously gets larger when distance travelled is shorter.

[26] Evaluated at the average pre-mobile-phone spatial price *differences* – reflecting the change that traders experience – the reduction is much larger and has, at the margin, a size of 37%-42%.

Table 2 Impact of mobile phones on transport costs: basic specification

dependent variable: transport costs of maize grain per kg (tc_{jkt}) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-438.5***	-426.7**	-396.1**	-399.8**
	(142.7)	(215.8)	(174.6)	(193.9)
road distance	1.556***	1.767	1.009*	1.528**
	(0.359)	(3.969)	(0.561)	(0.679)
R^2	0.816	0.792	0.812	0.858
no. of observations	1113	763	761	774
mean dependent variable:				
before intervention	2115.8 (284)	2368.8 (179)	2174.3 (245)	2280.0 (192)
after intervention	1665.7 (829)	1916.5 (584)	1200.6 (516)	1596.0 (582)

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Costs are converted to constant 2010 prices using the national consumer price index. All estimations include time fixed effects, 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, column (3) on a subset of pairs with typical production areas as source markets, and column (4) on a combination of (2) and (3). Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

We proceed with estimating the impact of mobile phone introduction on transports costs, reported in Table 2. The specification follows the spatial price difference estimations: In all estimations we have included time fixed effects and seasonality and trend, by source and destination. As a corollary to the spatial price difference estimations we investigate if outcomes are robust to restricting samples to typical source markets, typical destination markets and the combinations of these (Table 2, columns (2) to (4))[27].

Using all available data (column (1)), the impact of mobile phones on transport costs is statistically significant at the 1% level. The reduction in transport costs has, at the margin, a size of around 15%-21%, evaluated at the average pre-mobile phone per kg transport cost. If samples are adjusted along the same lines as Table 1, the size of the reduction is (statistically) similar, although less significant, mainly due to limited statistical power. The coefficient of road distance is also statistically significant and reasonably stable. Comparing Table 1 and Table 2 estimations we observe that the road distance coefficient in the spatial price difference estimation corresponds (statistically) with this coefficient in the transport costs estimation. This

[27] Note that, unlike the spatial price difference estimations, the samples of the transport cost estimations decrease due to these restrictions.

is a comforting result, since our conceptual framework requires that these coefficients are the same. The results further suggest how reduction of transport related search costs compares with reduction of collection of maize related search costs. The former is taken straight from Table 2, while the latter can be calculated by subtracting these from Table 1 impact estimates (which contains both): around 39% to 55% of the total search costs is associated with transport, and the remaining part of search costs is associated with the collection of maize in source areas.

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 DiD-OLS specification of Table 1 and Table 2 with covariates included. To control for a variety of supply, trade and demand effects we include in the estimations last season rainfall (rainfall from October to March), fuel prices, and population at source and destination. Estimations, reported in Table 3 and 4, further confirm previous results: impacts are statistically significant, with a size that is larger (in absolute terms) in the case of spatial price differences.

Table 3 Impact of mobile phones on dispersion of maize prices: DiD-OLS with covariates

dependent variable: positive maize price difference between markets ($p_{kt}–p_{jt}$, $p_{kt}>p_{jt}$) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-791.9^{***}	-770.7^{***}	-753.5^{***}	-774.8^{***}
	(236.7)	(173.7)	(134.4)	(193.6)
road distance	1.790^{***}	1.066^{***}	0.390^{***}	1.328^{***}
	(0.460)	(0.253)	(0.083)	(0.394)
seasonal rainfall, lagged	-1.946^{***}	-0.688	-1.577^{***}	-1.933^{***}
	(0.731)	(0.497)	(0.463)	(0.517)
population market pair	-0.011	0.009	0.004	0.004
	(0.032)	(0.019)	(0.020)	(0.021)
real diesel price [a]	yes	yes	yes	yes
adj R^2	0.666	0.639	0.656	0.664
no. of observations	3686	5832	5794	4659

Note to table: see Table 1

[a] Diesel prices are interacted with source markets.

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

dependent variable: transport costs of maize grain per kg (tc_{jkt}) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-413.5***	-363.8**	-470.9*	-390.9**
	(128.5)	(160.8)	(239.9)	(192.6)
road distance	1.546***	1.877	1.150**	1.731**
	(0.356)	(3.993)	(0.571)	(0.764)
seasonal rainfall, lagged	0.258	0.908	-0.036	0.299
	(0.382)	(0.491)	(0.505)	(0.557)
population market pair	-0.066***	-0.052	-0.050**	-0.035
	(0.023)	(0.043)	(0.025)	(0.033)
real diesel price [a]	yes	yes	yes	yes
R^2	0.816	0.792	0.811	0.856
no. of observations	1113	763	761	774

Note to table: See Table 2.
[a] Diesel prices are interacted with source markets.

We proceed with estimating spatial price differences adjusted for per kg transport costs (see *Empirical strategy*, equation (9)). We match transport cost data, by trade pair and by year or quarter, with spatial price differences[28]. Using period averages of transport costs implicitly assumes that per kg transport costs do not change in the very short run, within a year or a quarter. This procedure may also control for a certain degree of noise in the transport cost data, possibly through over-reporting by traders[29].

Table 5 Impact of mobile phones on dispersion of maize prices: combining spatial price differences with per kg transport costs

dependent variable: positive maize price difference between markets ((p_{kt}–p_{jt})–tc_{jkt},p_{kt}>p_{jt}) minus transport costs of maize grain per kg (tc_{jkt}) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-476.5*	-556.7*	-527.3*	-491.6**
	(264.7)	(315.9)	(290.2)	(243.7)
adj R^2	0.583	0.591	0.692	0.702
no. of observations	2318	2250	1535	1452

Note to table: Maize price data (from January 1999 to December 2007) and transport cost data (August 2001 to December 2010), both sourced from SIMA, overlap from 2001 to 2007. Prices and costs are converted to constant 2010 prices using the national consumer price index. All estimations include time fixed effects, and seasonality and trends, by source and destination. Equations are estimated using OLS. Column (1) (and (2)): minimum and median per kg transport cost by market pair and *year*, are matched to spatial price differences; Column (3) (and (4)): minimum and median per kg transport cost by market pair and *quarter* are matched to spatial price differences; Robust standard errors in brackets below the coefficient are clustered by market-pairs. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

[28] We experimented with matching the average, the median and the minimum of per kg transport costs, per quarter or per year.
[29] There is some support for this assertion since – for a small part of the observations – per kg transport costs by trade-pair and date exceed the corresponding spatial price differences.

The estimation results reported in Table 5 are consistent with previously found estimates of the reduction in price dispersion and transport costs. Estimated Table 5 coefficients are approximately equal to Table 1 (or Table 3) impacts minus Table 2 (or Table 4) impacts.

Robustness checks: impacts with propensity score matching estimations

In order to address possible selection bias, we verify estimation outcomes by 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: 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 substantial delay. The degree in which markets are geographically embedded in a network of towns, cities and villages is likely to matter. The location of markets in the cell phone network will affect costs and potential demand, and is, hence, also likely to be an important determinant.

Following these considerations we estimate the propensity score with trends, population of source and destination markets, district population density, geographical location, distance to big cities and seaports, network density and climate. All other than binary variables and trends, are log transformed. The deliberate focus on trends, geographical factors and long term structural determinants in explaining the propensity score guarantees exogeneity to maize market outcomes. Also, these variables simultaneously influence assignment into treatment or control group and, through this channel, the outcome variable, and are themselves not affected by assignment into treatment or control. Results of the propensity score estimation are reported in the Appendix (Table A4 and A5). Coefficients of the covariates in the propensity score estimation have expected signs: positive for population size and network density, and negative for district population density and distance to big cities and seaports. Also signs of geographical variables align with expectations, with positive coefficients for market pairs south of the Zambezi. 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, at least in the spatial price difference estimations. 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[30].

Table 6 Impact of mobile phones on price dispersion: PSM, Kernel Matching

outcome variable: maize price difference between markets ($p_{jt} – p_{kt}$) in constant 2010 prices				
	(1)	(2)	(3)	(4)
ATT	-792.9** (313.9)	-838.6*** (144.9)	-785.1*** (154.1)	-783.7*** (166.1)
ATU	-574.8	-553.6	-650.4	-803.6
ATE	-771.9	-762.4	-749.5	-788.6
treated, on support	2132	1498	1993	1602
treated, off support	1122	4007	3832	3359
untreated, on support	227	547	715	524
untreated, off support	206	536	721	561
no. of observations	3687	6588	7261	6046

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Prices are converted to constant 2010 prices using the national consumer price index. Equations are estimated using Propensity Score Estimation with Kernel Matching (Epanechnikov kernel; bandwidth=0.06; see main text for details). Samples in column (1) to (4) correspond with the samples in Table 1 and Table 3, column (1) to (4). Standard errors are in brackets next to the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 7 Impact of mobile phones on transport costs: PSM, Kernel Matching

outcome variable: transport costs of maize grain per kg (tc_{jkt}) in constant 2010 prices				
	(1)	(2)	(3)	(4)
ATT	-542.3** (267.0)	-532.2** (233.2)	-527.9 (356.7)	-563.1 (464.6)
ATU	-149.5	283.6	-816.1	-390.5
ATE	-395.0	-193.5	-777.4	-485.4
treated, on support	295	131	16	50
treated, off support	501	415	360	178
untreated, on support	177	93	103	41
untreated, off support	66	76	109	118
no. of observations	1039	715	588	387

Note to table: Transport cost data are from August 2001 to December 2010 (source: SIMA). Costs are converted to constant 2010 prices using the national consumer price index. Equations are estimated using Propensity Score Estimation with Kernel Matching (Epanechnikov kernel; bandwidth=0.06; see main text for details). Samples in column (1) to (4) correspond with the samples in Table 2 and Table 4, column (1) to (4). Standard errors are in brackets next to the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

The PSM impact estimations by and large confirm the estimation results obtained through OLS difference-in-difference, reported in the previous tables. Impact coefficients (ATT) are slightly lower (in absolute terms) but well in the range of the corresponding OLD/DiD estimates. The transport cost estimation (Table 7) are performing less, most likely due to lack of statistical power.

[30] 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 region of common support between treatment and comparison group is shown graphically (see Appendix, Figure A7). 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 spatial price differences 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)[31]. The results of this exercise indicate that matching on the estimated propensity score balances the different covariates in the matched samples reasonably well (see Appendix, Table A5.1 and A5.2).

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 100 of the nearest controls, with replacement, combined with a caliper threshold, where the caliper takes values between 0.005 and 0.1. Replacement is justified because the distribution of the propensity score is 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 A7). 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.

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

Further, ordering is done randomly since estimations with NN matching are dependent on the ordering of the data. Estimations with NN matching generate similar results as with Kernel Matching, though with a lower accuracy (see Appendix, Table A6 and A7). The relative similarity of the estimations with different types of matching offers confidence about the robustness of the kernel matching procedure, and, hence, about the OLS-DiD impact estimates.

4. Who benefits from access to mobile phones?

A welfare analysis in order to assess which group – farmers, traders, transporters or consumers – benefits from access to mobile phones, would be most attractive but is beyond the domain of the available data. However, we can make a few steps in exploring this issue. In the previous sections we have found a statistically significant decrease of spatial price difference. This price difference can come about by a decrease of price levels in destination markets or an increase of price levels in source markets. We can measure if the mobile phones induced reduction of price dispersion is due to an increase of prices in source markets or a decrease of prices in destination markets. Technically benefits may be shared evenly, with an equally sized price level decrease at destination and a price level increase at source, or the change in price difference is fully attributable to either a price level decrease at destination, or fully to an increase at source[32]. If the entire decrease in price dispersion is due to a decrease of prices in destination markets, consumers of maize fully capture 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, markets at the producer side[33] capture the benefits of access to mobile phones. As the

[32] It is also feasible that the price level decrease (increase) at destination (source) is larger than the reduction in the spatial price difference. In that case the benefits are negative for markets at source (destination).
[33] The current exercise does not allow to identify if maize growers benefit from mobile phones. For that purpose we need to investigate farm-gate or producer prices vis-à-vis market prices (see Zant, 2018).

level of competition is likely to be higher in destination markets than in source markets, we expect a bias of benefits towards destination markets.

We investigate this empirically by estimating essentially the same specification as in the case of spatial price differences, 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 (see equation (10) and (11)). Also the market fixed effects, although included, are obviously not associated with the road distance between markets. As in the case of the spatial price difference estimations, we accurately match, by year and market, the sample of price level observations to the sample of transport costs observations (column (1) to (3)). Alternatively we have constructed samples on the basis of typical source and destination markets (column (4)).

Table 8 Impact of mobile phones on levels of maize prices in destination markets

dependent variable: maize prices in destination markets (p_{jt}) in constant 2010 prices				
	(1)	(2)	(3)	(4)
cell phone	-1515.3**	-1301.0**	-1092.8***	-1095.8***
	(604.0)	(509.9)	(321.5)	(188.7)
adj R^2	0.827	0.821	0.794	0.799
no. of observations	2827	2269	1119	1451
mean dependent variable:				
before intervention	7632 (132)	7597.7 (63)	7194.0 (35)	5206.6 (272)
after intervention	7481 (2695)	7571.6 (2206)	6525.3 (1084)	6734.9 (1179)

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Prices are converted to constant 2010 prices using the national consumer price index. All estimations include time and market fixed effects, and seasonality and trends, by markets. Equations are estimated using OLS. Column (1): the sample is matched by destination market and year to the available transport cost observations; column (2) as column (1) but requiring at least 3 transport cost observations per market and year; column (3), as (2) with only observations from April to September; column (4) typical destination markets with only observations from April to September. Robust standard errors in brackets below the coefficient are clustered by market. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 9 Impact of mobile phones on levels of maize prices in source markets

dependent variable: maize prices in source markets (p_{kt})				
	(1)	(2)	(3)	(4)
cell phone	-616.0**	-734.4	-458.9	688.5*
	(270.0)	(1080.1)	(484.1)	(353.5)
adj R^2	0.850	0.838	0.857	0.752
no. of observations	2032	1462	859	1084
mean dependent variable:				
before intervention	7352.2 (216)	7046.5 (165)	6558.8 (109)	3890.5 (251)
after intervention	6614.5 (1816)	6514.0 (1297)	5650.2 (750)	5035.2 (833)

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Prices are converted to constant 2010 prices using the national consumer price index. All estimations include time and market fixed effects, and seasonality and trends, by markets. Equations are estimated using OLS. Column (1): the sample is matched by source market and year to the available transport cost observations; column (2) as column (1) but requiring at least 3 transport cost observations per market and year; column (3), as (2) with only observations from April to September; column (4) typical source markets with only observations from April to September. Robust standard errors in brackets below the coefficient are clustered by market. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Estimation results, reported in Table 8 and Table 9, are shown for destination markets (Table 8) and for source markets (Table 9) separately. The estimations indicate a statistically significant reduction in prices in destination markets, for all four samples. The size of the reduction is large with a larger sample, and decreases and becomes more accurate if observations are restricted to typical trading months. These latter estimations come closer to the previously estimated size of the reduction in spatial price differences. The estimation results for source markets are mixed and volatile[34]: estimations based on matched samples have a negative coefficient and are statistically significant in only one (out of three) estimations. The estimation based on selected typical source markets (column (4)) generates a positive coefficient, statistically significant at the 10% level. We are inclined to believe this latter result, as it is approximately consistent with the spatial price difference estimations and the price level estimations for destination markets and there is no decrease in source market prices.

[34] Results tend to be sensitive to the inclusion of remote markets (Lichinga, Angonia).

Table 10 Impact of mobile phones on levels of maize prices: constrained estimation

dependent variable: maize prices in both source and destination markets (p_t)				
	(1)	(2)	(3)	(4)
cell phone	-887.3***	-706.0***	-543.1***	-726.1***
in destination markets	(227.8)	(239.9)	(190.2)	(233.5)
cell phone	-87.3	94.0	256.9	73.9
in source markets	(227.8)	(239.9)	(190.2)	(233.5)
RMSE	1259.6	1122.6	980.6	854.7
no. of observations	4110	3146	1757	2328

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Prices are converted to constant 2010 prices using the national consumer price index. All estimations include time and market fixed effects, and seasonality and trends, by markets. Equations are estimated using OLS with a constraint on the cell phone coefficient (cell phone$_{\text{destination markets}}$ – cell phone$_{\text{source markets}}$ = –800). The samples in column (1) to (4) are combinations of the samples in the corresponding columns of the previous two tables. Robust standard errors in brackets below the coefficient are clustered by market. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

Consistency of the price level estimations with spatial price difference estimation requires that the change in price difference equals the change at destination markets minus the change at source markets ($\Delta(p_j - p_k) = \Delta p_j - \Delta p_k$). The reduction of the spatial price difference is estimated to be approximately –800 (see Table 1 and Table 3). Hence, to further explore the Table 8 and 9 results we have re-run the estimation of equation (10) and (11) in one equation, with two impact variables – one for source markets and one for destination markets – jointly with a constraint on the coefficients ($\Delta p_j - \Delta p_k = -800$). Results of these estimations, reported in Table 10, confirm a sizeable reduction of prices in destination markets and, if any, a much smaller and statistically insignificant increase of prices in source markets. Evaluated at mean prices before introduction of mobile phones this outcome corresponds with a price decrease of 7%-13% in destination markets and, at most, a price increase of around 0.5%-3% in source markets[35]. On the basis of this evidence we conclude that the benefits of the improved efficiency are biased towards the consumption side. The larger reduction in market prices on the consumption side is likely to reflect higher competition in destination markets or, alternatively, more bargaining power of traders in source markets relative to destination markets.

[35] Recall that prices in source markets are typically lower than prices in destination markets.

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’. At the very most it are long run trends and structural developments in the maize market that may play a – minimal – role in investing in mobile phone infrastructure. Given the huge and varying lags, this can be ignored and 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 increase trade flows and 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 development over the years, is unknown[36]. Nevertheless, we consider migration of traders on a large scale unlikely given market

[36] A possible increase in domestic trade flows could very well spill-over to international trade and such an increase in (international) trade will be reflected in bilateral trade with neighboring countries (Tanzania, Malawi, Zambia, Zimbabwe and South-Africa). We do have data of annual bilateral trade flows, notably annual total import and annual total export with neighboring countries (1995-2013). We use these series to run a simple difference-in-difference estimation, with trade partner and year fixed effects, and trade partner trends, in order to investigate if international trade flows (both import, export and their sum) are impacted by the introduction of mobile phones in Mozambique, with a potential impact starting in the years between 2000 and 2006. We were unable to detect a significant impact (results available from author).

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, in the first place, the short run response to increased information (changes in price dispersion and transport costs, with limited 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 improved transparency of market prices of inputs and outputs and lower trade costs, leading to higher productivity, higher farm gate prices and higher profitability. Consequently, maize growers may have an incentive to increase production which, in its turn, affects the maize market prices and spatial price differences. Elsewhere we have elaborated on the impact of mobile phones on farm gate prices (see Zant, 2019b): unfortunately we find no support for higher farm gate prices. Given this evidence it is unlikely that a supply response from maize growers related to the introduction of mobile phones affects markets outcomes and spatial price differences. Hence, impact estimates are also unlikely to be biased by supply response from maize growers.

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 spatial price differences. This implies that estimated impacts are biased downwards and should interpreted as lower bounds. Formally, survey data on consumption are needed to further verify possible changes in demand due to the introduction of mobile phone services. This is beyond the scope of the current work. We find some comfort in the fact that reported impact estimates are statistically significant, of substantial size, and represent, at the very least, lower bounds of impacts.

A final concern is about collusion: mobile phone services may enhance collusion between traders by facilitating communication and coordination. This may help traders in keeping prices low in source markets and high in destination markets. With an undoubtedly large number of mainly small-scale and informal traders, dispersed over a vast country and a multitude of itineraries, involving millions of trade transactions, it is difficult to believe that the measured impact is importantly affected by collusion. The evidence on changes in price levels and the analysis who benefits from improved efficiency indicates that traders have less power to affect prices in destination markets. If the larger power to affect prices in source markets is due to collusion or simply due to a smaller number of agents on either side of the (source) market is difficult to answer.

6. Summary and conclusion

This study investigates empirically the impact of the mobile phone roll-out in Mozambique on price dispersion and transport costs. Estimations indicate a 10% to 13% decrease in price dispersion, implying an improvement in the efficiency of maize markets as a result of the introduction of mobile phones. The reduction in price dispersion is associated with a reduction in search costs, of which approximately half is related to transport and the other half to other search costs like collecting maize in source markets. With a travelling distance between source and destination markets of 1000km (close to average long distance transport), distance related transport costs comprise around 60% of total transport costs, while the remaining part is transport-related search costs. The evidence further indicates that the reduction in price dispersion comes about primarily in the form of lower prices in destination markets and, at most, moderately higher prices in source markets: We find a reduction of prices in destination markets of 7%-13% and a price increase, if any, of around 0.5%-3% in source markets. Hence, the benefits of improved market efficiency accrues mainly to consumers, while market prices in source areas appear much less affected. The retail market prices used in this empirical work are not adequate to investigate if maize growers have benefited from mobile phones (see Zant, 2019b) for work in this direction). Robustness of impacts is verified by checking the parallel trend assumption underlying the DiD approach, and by employing propensity score matching to control for possible selection bias. Finally, the plausibility of alternative explanations and the relevance of potential threats to estimated impacts are 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.

Autor, D.H., 2003, ‘Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing’, *Journal of Labor Economics*, 2003, 21, 1.

Bertrand, M., E. Duflo and S. Mullainathan, 2004, ‘How much should we trust difference-in-differences estimates?’, *Quarterly Journal Of Economics*, 119, 1, 249-275.

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.

Baulch, B. 1997, ‘Transfer Costs, Spatial Arbitrage, and Testing for Food Market Integration’, *American Journal of Agricultural Economics,* 79, 2, 477-487.

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., E.Gabre-Madhin and B.Minten, 2005, ‘Increasing returns and market efficiency in agricultural trade’, *Journal of Development Economics*, 78, 406-442.

Fafchamps, M. and R. Vargas-Hill, 2008, ‘Price Transmission and Trader Entry in Domestic Commodity Markets’, *Economic Development and Cultural Change*, 56, 4, 729-766.

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 Dec. 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.

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

Zant, W., 2019a, ‘Bridges’, Tinbergen Discussion Paper 17-117/V, Tinbergen Institute, Amsterdam.

Zant, W., 2019b, ‘Mobile Phones and Mozambique Farmers: Less Asymmetric Information and More Trader Competition?’ Tinbergen Discussion Paper 18-055/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 Jenny Aker, 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 Prices in source and destination markets in the south

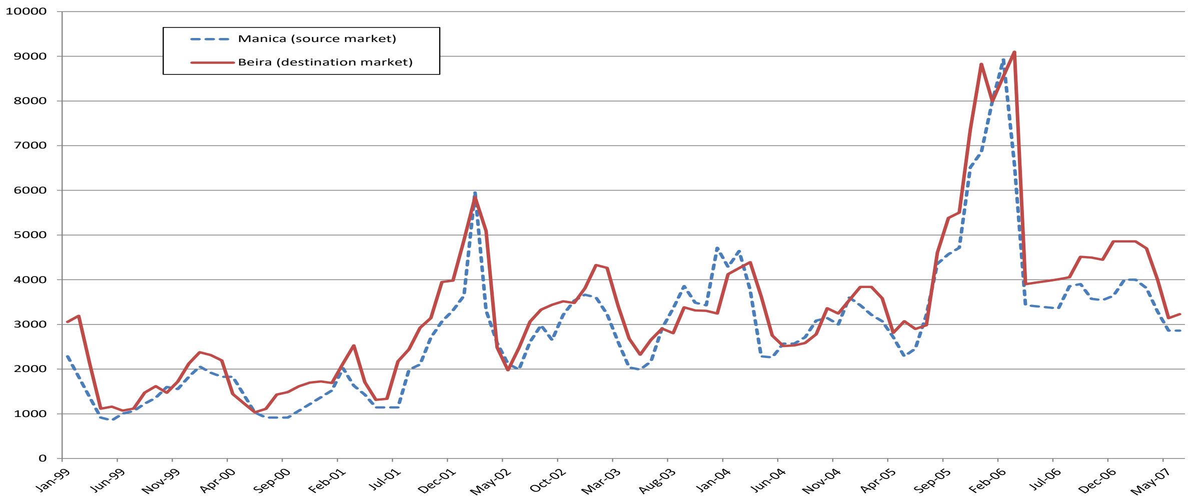


Figure A2 Prices in source and destination markets in the north

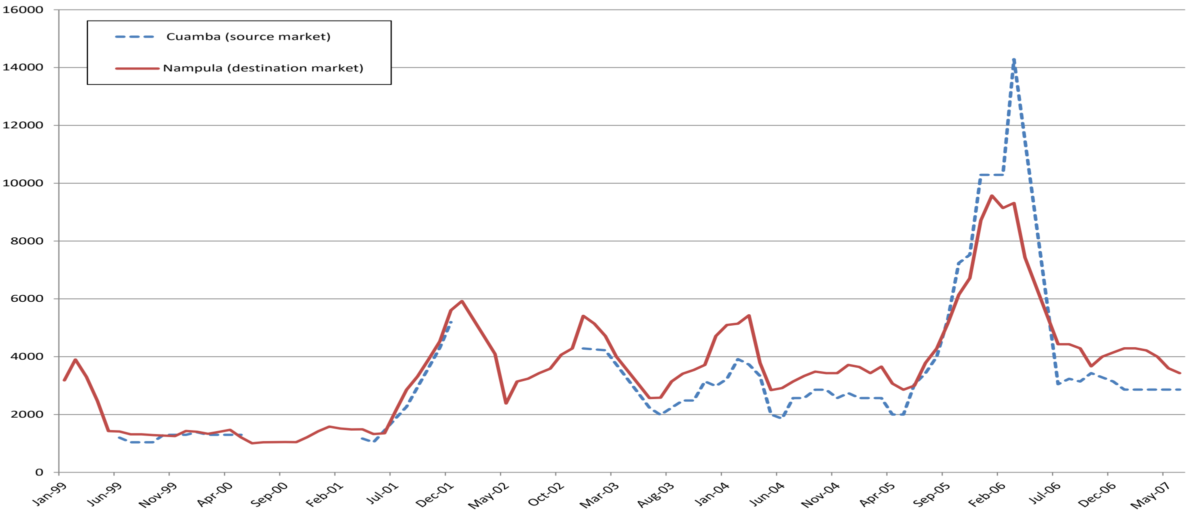
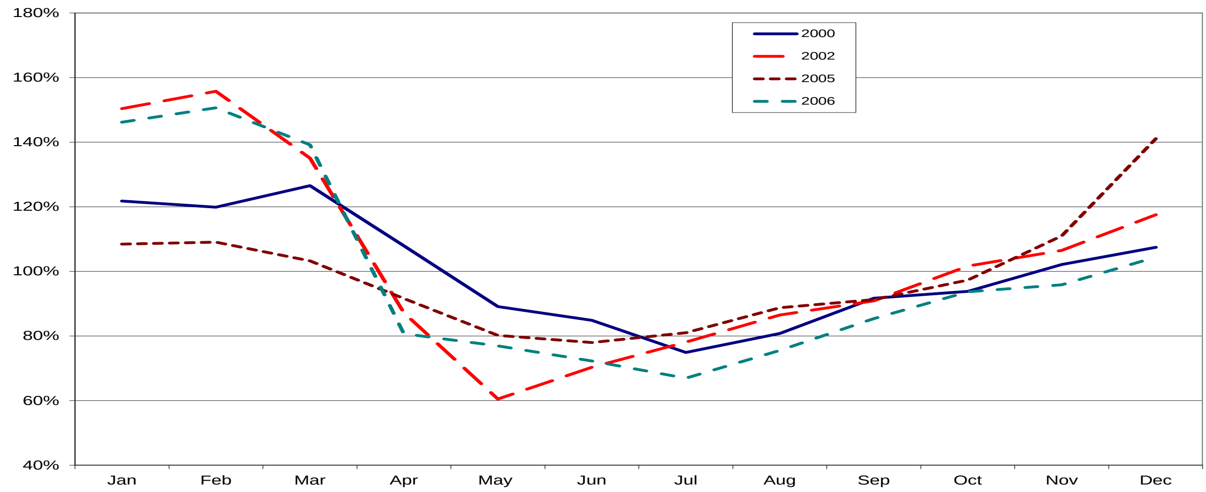


Figure A3 Seasonality in maize prices, by year, selected years



Source: (author's calculations based on data from) SIMA (Figure A1-A3)

Figure A4 Seasonality in maize prices, by market over time, source and destination

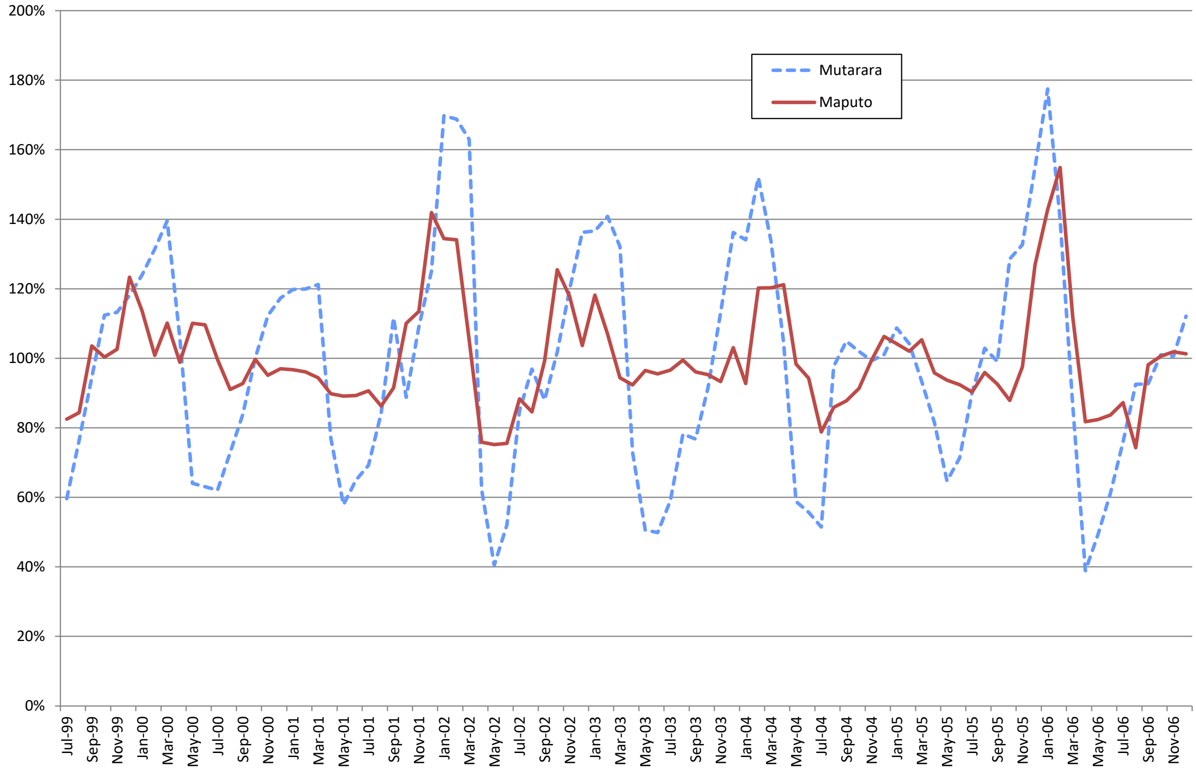
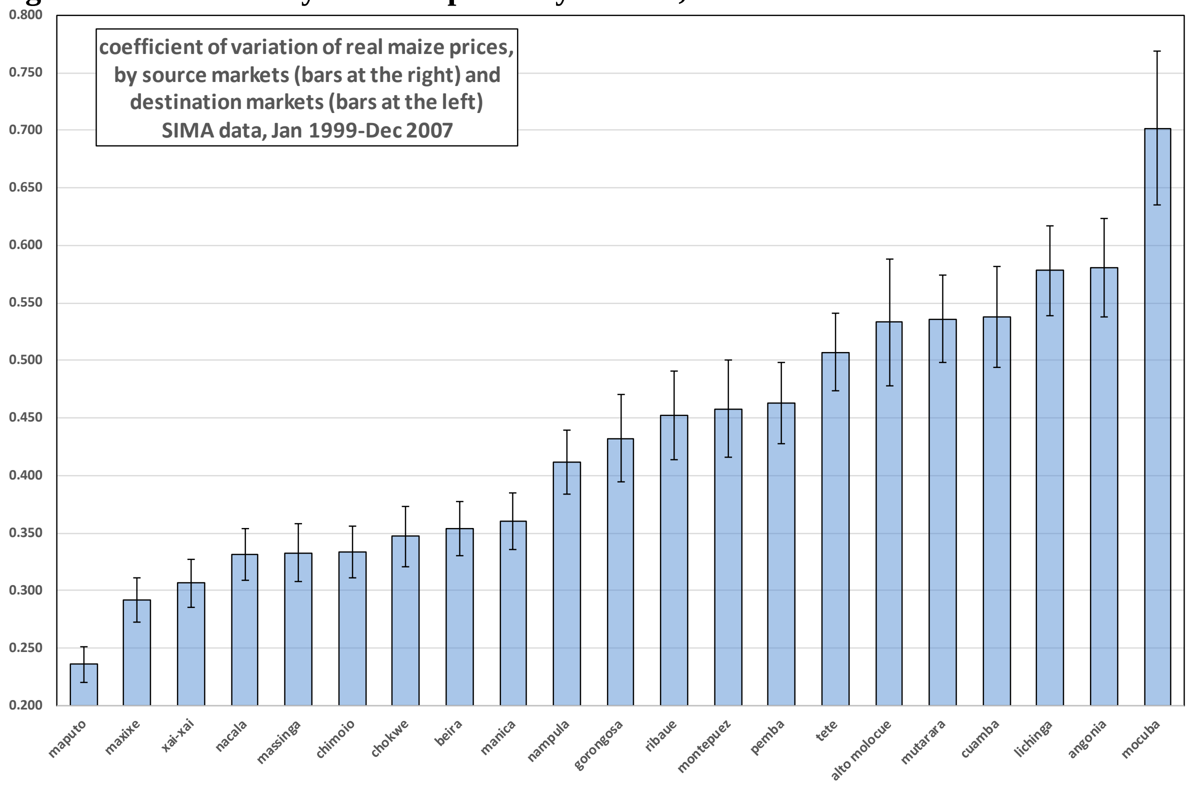
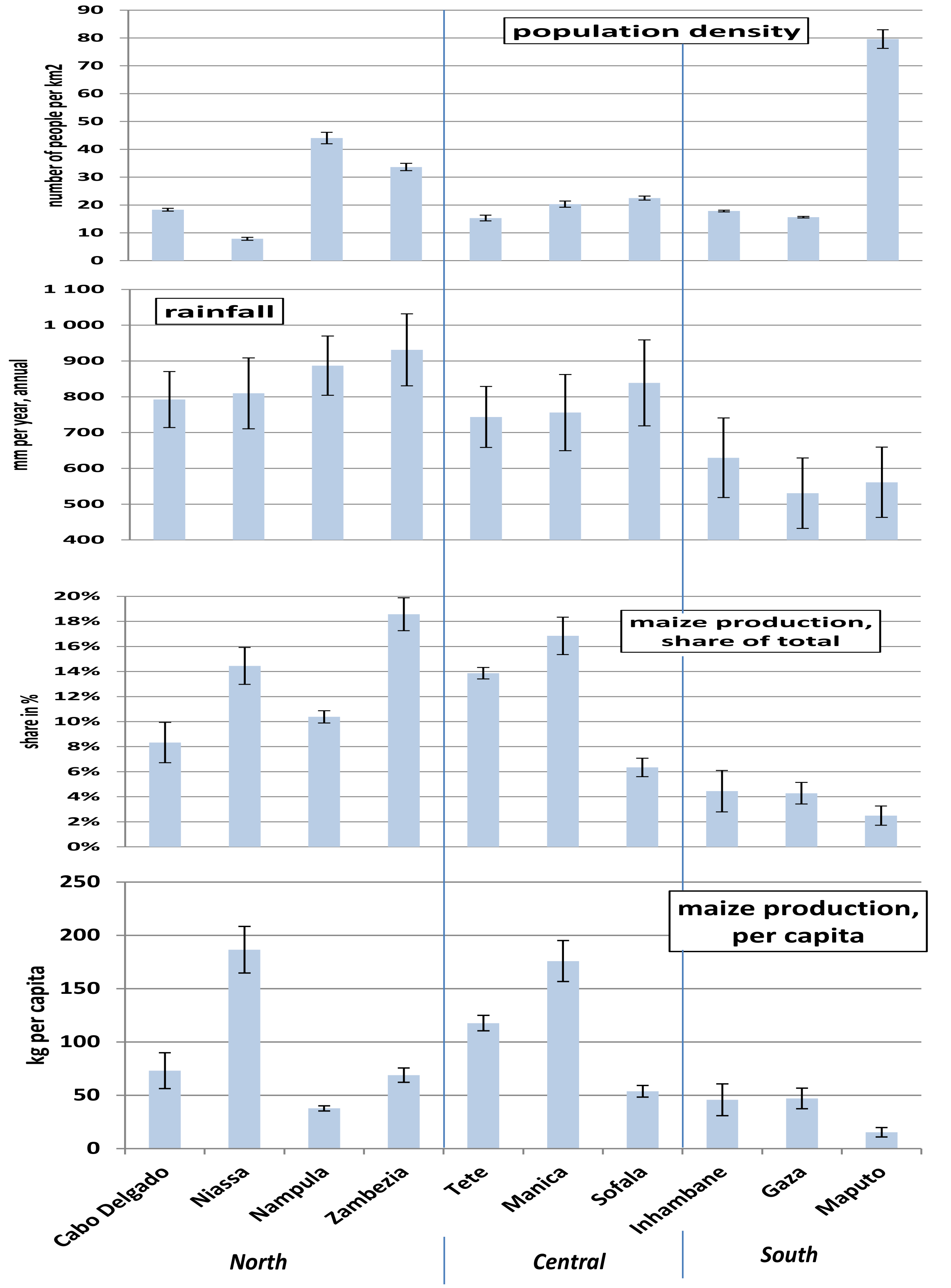


Figure A5 Volatility in maize prices by market, source and destination markets



Note: Error bars show 95% confidence intervals; Source: (author’s calculations based on data from) SIMA

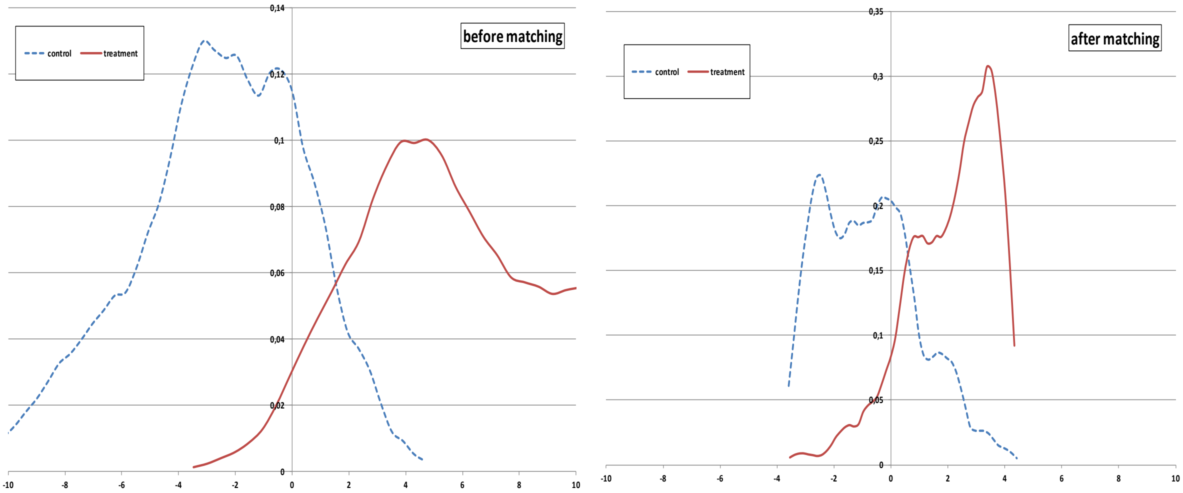
Figure A6 Population density, rainfall and maize production by province, 1999-2007



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 paper for the location of provinces.

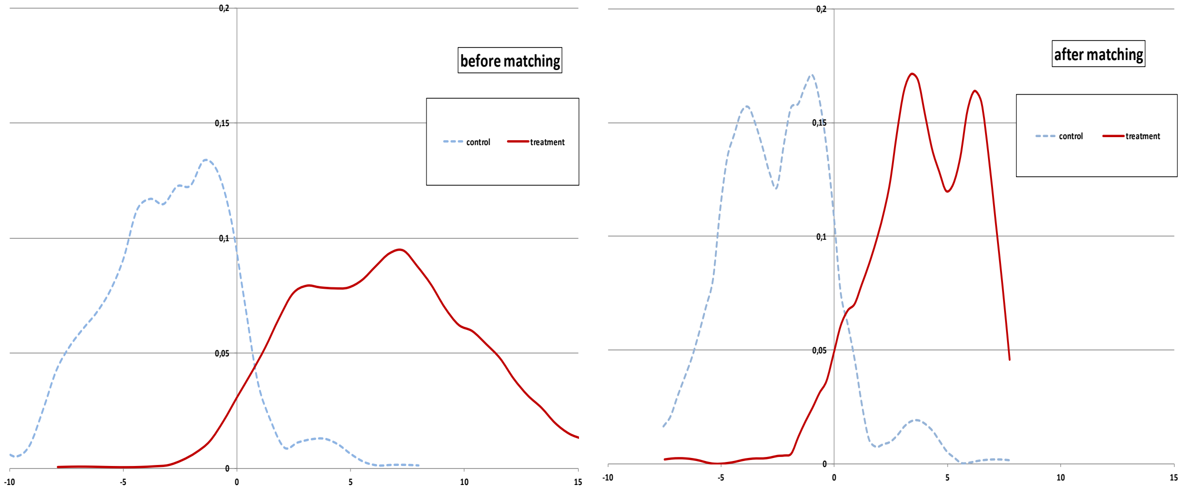
Figure A7 Common support between treatment and control group

price dispersion



Note: PSM, Kernel Matching for spatial price difference (Table 6, column 4, common support figures for estimation reported in column 1, 2, and 3 available from the author)

transport costs



Note: PSM, Kernel Matching for transport costs (Table 7, column 1; common support figures for estimation reported in column 2, 3 and 4 are similar and available from the author)

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. Hence, the share of missing observations is equal to 100 minus the number under %. 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	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 A4.1 First stage logistic estimation of propensity score: price dispersion sample

Dependent variable: probability of having access to mobile phone technology (cell phone)				
	(1)	(2)	(3)	(4)
trend		0.887**	-1.286***	0.637**
		(0.433)	(0.352)	(0.280)
trend, squared		0.120***	0.326***	0.125***
		(0.046)	(0.038)	(0.029)
trend (inverse)	-44.02***			
	(2.47)			
ln(distance to big city, source)	-1.185***	-0.224***	-0.420***	-0.135***
	(0.101)	(0.049)	(0.047)	(0.040)
ln(distance to big city, destination)		0.352***	0.424***	0.285***
		(0.054)	(0.043)	(0.037)
ln(district population density, source)	-3.266***	-2.005***	-1.506***	-1.278***
	(0.262)	(0.152)	(0.141)	(0.117)
ln(district population density, destination)	-0.923***	-1.474***	-0.675***	-0.706***
	(0.200)	(0.411)	(0.167)	(0.152)
ln(network density, source)		4.233***	6.052***	4.375***
		(0.334)	(0.340)	(0.257)
ln(network density, destination)		1.720**	0.854***	0.925***
		(0.781)	(0.324)	(0.288)
ln(city population, source)	1.393***	1.860***	1.687***	1.884***
	(0.156)	(0.112)	(0.102)	(0.090)
ln(city population, destination)	2.003***	1.901***	1.970***	1.591***
	(0.148)	(0.248)	(0.128)	(0.107)
north of Zambezi	1.977***	-0.301	-1.120***	-1.608***
	(0.476)	(0.394)	(0.247)	(0.219)
south of Zambezi	8.001***	2.864***	3.936***	2.456***
	(0.544)	(0.207)	(0.198)	(0.150)
pseudo R2	0.684	0.680	0.735	0.697
observations	3624	5166	5536	6432

Note to table: see Table 6 and Table 7 for sample specification of each column. Equations are estimated using a logit specification. Standard errors are in brackets below the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A4.2 First stage logistic estimation of propensity score: transport cost sample

Dependent variable: probability of having access to mobile phone technology (cell phone)				
	(1)	(2)	(3)	(4)
trend (inverse)	-15.56***	-12.88***	-20.52***	-17.80***
	(1.389)	(1.926)	(2.790)	(1.675)
ln(distance to big city, source)	-0.883***	-0.677***	8.695***	-1.375***
	(0.162)	(0.257)	(2.154)	(0.208)
ln(distance to big city, destination)	0.541***	1.161***	0.657***	0.399***
	(0.143)	(0.268)	(0.167)	(0.128)
ln(district population density, source)	-0.870***	-1.344***	-3.666***	-0.789***
	(0.282)	(0.401)	(0.653)	(0.294)
ln(district population density, destination)	-2.248**	-6.632***	-3.282***	-1.835**
	(0.929)	(1.802)	(1.008)	(0.845)
ln(network density, source)	3.443***	5.548***	16.74***	2.340***
	(0.846)	(1.427)	(3.647)	(0.849)
ln(network density, destination)	3.977**	14.954***	5.906***	2.911*
	(1.706)	(3.526)	(1.901)	(1.589)
ln(city population, source)	3.792***	4.535***	5.280***	3.925***
	(0.356)	(0.489)	(0.624)	(0.411)
ln(city population, destination)	3.457***	6.956***	4.794***	2.992***
	(0.650)	(1.272)	(0.749)	(0.559)
north of Zambezi	2.095***	2.492***	5.629***	2.568***
	(0.598)	(0.754)	(1.240)	(0.676)
south of Zambezi	5.890***	5.530***	7.490***	6.683***
	(0.758)	(0.870)	(1.354)	(0.880)
pseudo R2	0.739	0.773	0.835	0.749
observations	1126	962	737	1007

Note to table: see Table above

Table A5.1 Standardised Bias of Covariates, before and after matching price dispersion sample

sample	(1)		(2)		(3)		(4)	
before, after matching variables	before	after	before	after	before	after	before	after
trend			1.566	0.511	1.350	0.395	1.438	0.436
trend squared			1.475	0.466	1.313	0.352	1.378	0.396
trend inverse	-1.480	-1.181						
ln(distance to big city, s)	-0.238	-0.022	-0.319	-0.281	-0.339	-0.308	-0.356	-0.219
ln(distance to big city, d)			-0.249	0.061	-0.318	-0.023	-0.324	0.030
ln(distr. pop.density, s)	-0.596	-0.387	-0.365	-0.255	-0.371	-0.317	-0.230	-0.165
ln(distr.pop.density, d)	0.151	0.085	-0.053	-0.024	0.084	0.096	0.064	0.089
ln(network density, s)			0.752	0.320	0.980	0.342	0.804	0.381
ln(network density, d)			-0.154	-0.006	-0.028	0.116	-0.055	0.128
ln(city population, s)	0.539	0.107	0.675	0.414	0.744	0.374	0.788	0.303
ln(city population, d)	0.630	0.161	0.320	-0.011	0.428	0.034	0.422	-0.010
north of Zambezi	0.037	0.147	0.089	0.235	-0.453	-0.123	-0.431	-0.033
south of Zambezi	0.515	0.185	0.634	0.028	1.095	0.472	0.907	0.348

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 6 and 7).

Table A5.2 Standardised Bias of Covariates, before and after matching Transport cost sample

sample	(1)		(2)		(3)		(4)	
before, after matching variables	before	after	before	after	before	after	before	after
trend inverse	-1.207	-1.001	-1.175	-0.561	-1.625	-1.135	-1.238	-0.942
ln(distance to big city, s)	-0.485	-0.027	-0.575	-0.144	-0.149	0.111	-0.597	-0.214
ln(distance to big city, d)	-0.285	-0.188	-0.333	0.031	-0.255	-0.004	-0.449	-0.216
ln(distr. pop.density, s)	-0.091	-0.212	-0.086	-0.229	-0.198	-0.142	-0.133	-0.265
ln(distr.pop.density, d)	0.010	0.152	-0.017	0.070	0.123	-0.074	-0.061	0.106
ln(network density, s)	0.477	0.466	0.547	0.154	0.638	0.054	0.586	0.372
ln(network density, d)	0.007	0.235	-0.035	0.219	0.239	0.281	-0.102	0.224
ln(city population, s)	0.890	0.514	1.012	0.475	0.819	0.411	0.857	0.547
ln(city population, d)	0.318	0.096	0.344	-0.100	0.234	-0.152	0.410	0.085
north of Zambezi	0.503	0.138	-0.021	0.068	0.364	0.126	-0.002	0.161
south of Zambezi	-0.040	0.339	0.557	0.330	0.302	0.239	0.419	0.232

Note to table: see table above.

Table A6 Impact of mobile phones on price dispersion: PSM, Nearest Neighbour

outcome variable: real maize price difference between markets ($p_{jt} - p_{kt}$)				
	(1)	(2)	(3)	(4)
ATT	-838.9*** (242.2)	-782.9*** (247.0)	-811.1** (286.2)	-859.6** (309.7)
ATU	-502.7	-767.9	-912.5	-790.7
ATE	-814.1	-780.9	-831.1	-849.4
treated, on support	3254	5505	5825	4961
treated, off support	0	0	0	0
untreated, on support	260	853	1436	858
untreated, off support	173	230	0	227
no. of observations	3687	6588	7261	6046

Note to table: Maize price data are from January 1999 to December 2007 (source: SIMA). Equations are estimated using Propensity Score Matching with Nearest Neighbour (n=2 to n=100; and caliper: 0.005 to 0.1; 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 next to the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table A7 Impact of mobile phones on transport costs: PSM, Nearest Neighbour

outcome variable: real transport costs of maize grain per kg (tc_{jkt})				
	(1)	(2)	(3)	(4)
ATT	-542.3** (267.0)	-497.3 (330.9)	-414.0* (216.8)	-555.2 (392.4)
ATU	-149.5	-37.9	-1140.1	-545.2
ATE	-395.0	-398.8	-675.8	-551.1
treated, on support	295	546	376	228
treated, off support	501	0	0	0
untreated, on support	177	149	212	159
untreated, off support	66	20	0	0
no. of observations	1039	715	588	387

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=2 to n=100; and caliper: 0.005 to 0.1). Standard errors are in brackets next to the coefficient. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.