WPS7729

Policy Research Working Paper 7729

Roads and Rural Development in Sub-Saharan Africa

Claudia N. Berg
Brian Blankespoor
Harris Selod

WORLD BANK GROUP
Development Research Group
Environment and Energy Team
June 2016

Abstract

This paper assesses the relation between access to markets and cultivated land in Sub-Saharan Africa. Making use of a geo-referenced panel over three decades (1970–2005) during which the road network was significantly improved, the analysis finds a modest but significant positive association between increased market accessibility and local cropland expansion. It also finds that cropland expansion, in turn, is associated with a small but significant increase in local gross domestic product. These results are suggestive of agricultural activities that develop at the extensive margin, which are mostly to serve local demand, but are not indicative of commercial agriculture that serves external markets.

This paper is a product of the Environment and Energy Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at hselod@worldbank.org, bblankespoor@worldank.org and cberg@imf.org.

Roads and Rural Development in Sub-Saharan Africa [1]

Claudia N. Berg, Brian Blankespoor, and Harris Selod [2]

Keywords: Transportation, Market access, Cropland expansion, Africa

JEL: R4, O1, Q15

[1] The authors thank Rabah Arezki, Punam Chuhan-Pole, Richard Damania, Uwe Deichmann, Siobhan Murray, Sandrine Mesplé-Somps and Gilles Spielvogel for earlier discussions related to this paper, and Rémi Jedwab and Adam Storeygard for making the roads data they have constructed available for calculation of accessibility indexes, as well as Rose Choi for research assistance. Funding from DFID under the World Bank's Strategic Research Program 'Transport Policies for Sustainable and Inclusive Growth' is gratefully acknowledged. Corresponding author: Harris Selod (email: hselod@worldbank.org, address: 1818 H Street, NW, Washington, D.C. 20433).
[2] Berg: Research Department Commodities Unit, International Monetary Fund. Blankespoor and Selod: Development Research Group, The World Bank.

1. Introduction

Due to climatic and soil conditions, Sub-Sub-Saharan Africa has enormous potential for agriculture. Over the past four decades, although the region's total cropped land is estimated to have increased 37 percent, reaching 211 million hectares (see Table A1 in the Appendix), this vast increase falls short of the available land suitable for cultivation in the region with estimates between approximately 70 and 200 million hectares depending on the definition (Deininger et al., 2011).[1] Despite Sub-Sub-Saharan African countries being mostly agrarian, agriculture in the region compares poorly to other regions of the world given the low yields, widespread subsistence farming, and the heavy reliance on food imports. Many explanations have been put forward to explain Africa's backlog in agriculture, including the low adoption of modern technologies, poor access to credit, insecure property rights and lack of access to markets due to poor transport infrastructure. Lack of transport infrastructure, in particular, is hindering the development of agriculture in Sub-Sub-Saharan Africa (see Ali et al., 2015; Berg, Deichmann, Liu, and Selod, *forthcoming*). Deininger et al. (2011) report that more than half of the untapped potential for cultivation in the region is located more than 6 hours away from a major market.

The objective of this study is to investigate the relation between road investments and rural development in Sub-Sub-Saharan Africa. Over the past four decades, there has been a very significant extension of the paved road network in Africa, which increased from 77,800 km in 1970 to 185,000 km in 2005 (see Table A2). The network, however, remains of poor quality and of insufficient extent and density. Foster and Briceño-Garmendia (2010) report an average road density of 137 km/100 km^2 in Sub-Sub-Saharan Africa, which is far below the 211 km/100 km^2 in other comparable low-income countries. Transport in Sub-Sub-Saharan Africa is also very costly to use, not only because of slow travel times due to the poor state of roads, but because of 'non-physical' transport costs related to delays, poor competition in the transport industry leading to higher prices, and corruption (Raballand, Macchi, and Petracco, 2010). The literature on transport and development shows that roads can have a very important impact on agriculture and rural development more generally. The main channel is by improving access to markets for agricultural produce. Better roads can boost commercial agriculture, participation to markets, and the adoption of modern techniques (see Kyeyamwa et al., 2008, on Uganda; Minten, Koru, and Stifel, 2013, on Ethiopia; Damania et al., *forthcoming*, on Nigeria). It can also result in the growth of the non-agricultural sector (see Ali et al., 2015, on Nigeria). In Mali, however, although better road connections seem to increase local employment in the agricultural and service sectors in rural areas, it reduces employment in the manufacturing sector (Blankespoor, Mesplé-Somps, Selod and Spielvogel, 2016). The evidence above suggests that the development of one sector may occur at the expense of another; some activities may relocate following transport investments, or road investments may be necessary but not sufficient to generate structural transformation. The impact on poverty reduction from rural roads, however, is unambiguously positive in the existing empirical literature. Roads in Tanzania reduce incentives to migrate out from rural areas (Gachassin, 2013). The channel can be through higher incomes as in Nepal and Madagascar (Jacoby, 2000, and Jacoby and Minten, 2009). All of the above suggest a potentially high return from road investments in the rural areas of Sub-Sub-Saharan Africa. Unfortunately, road investments have been insufficient in the region, even in places with high agricultural potential (Blimpo, Harding, and Wantchekon, 2013; Wantchekon and Stanig, 2015). This lack of investment can be explained by funding difficulties amplified by corruption (Collier, Kirchberger, and Söderbom, 2015).

In this paper, we investigate how increased market access stemming from road network improvement in Sub-Sub-Saharan Africa increases the amount of cultivated land over a long period (1970-2005). This period covers a significant improvement in the nascent road network in the region as well as important demographic spatial changes associated with population increases and urbanization. Although agricultural impacts can occur at both the extensive and the intensive margins, due to data constraints, we focus on the former only (cropland area). In the context of Sub-Saharan Africa, this approach is reasonable given stagnating agricultural yields and the patterns of extensive rather than intensive margin growth (Deininger et al., 2011, Rakotoarisoa et al., 2011). To explore the relationship between roads and cropland area, we bring together four geo-referenced panel datasets on roads, cropland, 'local GDP' and urban population, all defined at a small geographic level.

The closest papers in the literature are Jedwab and Storeygard (2016) and Blankespoor et al. (2016), which similarly study the impact of market access using the same panel road data. Jedwab and Storeygard (2016) estimate the impact of a change in market access on urbanization in 39 countries. Blankespoor et al. (2016) focus on the specific case of Mali to investigate the impact of changes in market access on the dynamics of population and sectors of employment. In the present paper, we focus on the impact on agricultural land use for the whole of Sub-Saharan Africa.

The remainder of this paper is organised as follows. In Section 2, we present our empirical approach, introducing our measure of market access and presenting our identification strategy. This is then followed by a section in which we describe our data sources and construction of variables, and provide the relevant descriptive statistics for the analysis. Section 4 presents the results. Section 5 concludes.

2. Empirical framework

2.1. Measuring access to markets

Following the standard approach in the literature, we calculate domestic market access for a given location as a function of the weighted sum of the populations of all other locations, with a weight that decreases with transport time.[2] Formally, we define market access in a location *i* at time *t* ($MA_{i,t}$):[3]

$$MA_{i,t} = 1 + \sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta} \tag{1}$$

where $P_{j,t}$ is the population in location *j* at time *t*, $\tau_{ij,t}$ is the travel time between locations *i* and *j* at time *t*, and θ is a trade elasticity parameter.[4]

As our statistical analysis will also focus on changes in market access (see below), we calculate the change in the logarithm of the market access index between dates *t* and *t*-1:

$$\Delta \ln MA_{i,t} = \ln(1 + \sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta}) - \ln(1 + \sum_{j \neq i} P_{j,t-1} \tau_{ij,t-1}^{-\theta}) \tag{2}$$

Following Blankespoor et al. (2016), this change can be decomposed to account for changes in the extension and quality of the road network as well as for changes in the spatial distribution of the population. We use the following formula:

$$\begin{aligned} \Delta \ln MA_{i,t} = {} & \left(\ln(1 + \textstyle\sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta}) - \ln(1 + \sum_{j \neq i} P_{j,t} \tau_{ij,t-1}^{-\theta})\right) \\ & + \left(\ln(1 + \textstyle\sum_{j \neq i} P_{j,t} \tau_{ij,t-1}^{-\theta}) - \ln(1 + \sum_{j \neq i} P_{j,t-1} \tau_{ij,t-1}^{-\theta})\right) \\ = {} & \beta_r \Delta_{road} \ln MA_{i,t} + \beta_p \Delta_{pop.} \ln MA_{i,t} \end{aligned} \tag{3}$$

where $\beta_r \Delta_{road} \ln MA_{i,t}$ represents the change in travel time between time t and time *t-1*, holding population constant for the population distribution at time t, and $\beta_p \Delta_{pop.} \ln MA_{i,t}$ represents the change in population between time t and t-1 holding travel time constant for the state of the network at time t-1. Following Donaldson (forthcoming), we use an elasticity of trade, $\theta,$ equal to 3.8. As Jedwab and Storeygard (2016) and Blankespoor et al. (2016), we also use alternative values of the trade elasticity as robustness checks.

2.2. Econometric approach

In the analysis, we explore the links between roads and rural development by assessing the impact of accessibility on cropland area and on a local measure of economic activity. We first present the regressions of cropland area. For this, we estimate alternative regressions, starting with one of cropland area on the market access index, in levels:

$$\ln C_{i,t} = \alpha \ln MA_{i,t-1} + \boldsymbol{X}'_{i,t}\, \boldsymbol{\gamma} + \boldsymbol{Z}'_{i,t-1} \zeta + \boldsymbol{D}'_i\, \xi + \varepsilon_{i,t} \tag{4}$$

where $C_{i,t}$ is cropland area in grid cell i at time t, $MA_{i,t-1}$ is the lagged natural logarithm of the market access indicator calculated with Formula (1), $\boldsymbol{X}_{i,t}$ is a vector of control variables at time t, $\boldsymbol{Z}_{i,t-1}$ is a vector of control variables at time t-1, $\boldsymbol{D}_i$ is a vector of time-invariant dummies, and $\varepsilon_{i,t}$ is the error term. The coefficient of interest is α, the elasticity of cropland with respect to our market access index. We lag the market access index to address concerns of reverse causation where areas put into cultivation may influence the measure of market access by attracting population or road investments. Controls include the average annual rainfall over the previous decade and its square, the lagged population density, and the distance to the nearest major port. The annual rainfall accounts for time-variant climatic conditions influencing agriculture. It is a key determinant of cultivation in Sub-Saharan Africa where irrigation is scarce. The lagged population density is included to control for the local level of urbanization. Distance to the nearest major port provides a measure of external market access.[5] To control for any remaining unobserved heterogeneity, we also include country fixed effects, time dummies, and the interaction between the two. This specification is estimated under Ordinary Least Squares (OLS) as well as under a Fixed Effect (FE) regression (for which the invariant controls are removed). The FE regression better addresses endogeneity issues through the inclusion of fixed effects accounting for unobserved heterogeneity at the very small grid cell level.

A second approach consists in regressing the change in the natural logarithm of cropland area on the change in the natural logarithm of the market access index. This measure has the advantage of being decomposable according to Formula (2). We have:

$$\Delta \ln C_{i,t} = \beta \Delta \ln MA_{i,t} + \beta_s S_{i,t-1} + \beta_{ms} \Delta \ln MA_{i,t} \times S_{i,t-1} + \boldsymbol{X}'_{i,t}\, \boldsymbol{\chi} + \boldsymbol{Z}'_{i,t-1} \boldsymbol{\zeta} + \boldsymbol{D}'_i\, \boldsymbol{\xi} + \eta_{i,t} \quad (5)$$

where $\Delta \ln C_{i,t} = \ln C_{i,t} - \ln C_{i,t-1}$ is the change in the natural logarithm of cropland area, $\Delta \ln MA_{i,t}$ is the change in the natural logarithm of the market access index as in Formula (2), $S_{i,t-1}$ is a dummy variable indicating that the cropland area in the cell shrank between dates *t*-1 and *t*-2 ('shrinking cropland dummy'), $\Delta \ln MA_{i,t} \times S_{i,t-1}$ is the interaction term between the two. The other controls are the same as before, except for the inclusion of the logarithm of the cropland area at time *t*-1 to account for the initial level of cultivation, and $\eta_{i,t}$ is the error term. The shrinking land dummy is included to account for desertification (a serious problem in Sub-Saharan Africa) and for land conversion to other uses.

A third approach is to run the same regression as (5) but substituting the change in the logarithm of the market access index with its decomposition provided by Formula (3) and introducing the corresponding interaction terms. We have:

$$\begin{aligned}\Delta \ln C_{i,t} = {} & \beta_r \Delta_r \ln MA_{i,t} + \beta_p \Delta_p \ln MA_{i,t} \\ & + \beta_s S_{i,t-1} + \beta_{rs} \Delta_{road} \ln MA_{i,t} \times S_{i,t-1} + \beta_{ps} \Delta_{pop.} \ln MA_{i,t} \times S_{i,t-1} \\ & + \boldsymbol{X}'_{i,t}\, \boldsymbol{\chi} + \boldsymbol{Z}'_{i,t-1} \boldsymbol{\zeta} + \boldsymbol{D}'_i\, \boldsymbol{\xi} + \varphi_{i,t}\end{aligned} \tag{6}$$

where $\Delta_{road} \ln MA_{i,t}$ and $\Delta_{pop.} \ln MA_{i,t}$ are the respective roads and population components in the decomposition of the natural logarithm of the market access index defined in Formula (3), $\Delta_{road} \ln MA_{i,t} \times S_{i,t-1}$ and $\Delta_{pop.} \ln MA_{i,t} \times S_{i,t-1}$ are the respective interaction terms between these components and the shrinking cropland dummy, and $\varphi_{i,t}$ is the error term. All other controls are the same as in Regression (5).

Finally, we also investigate the relation between cropland area and our measure of local activity (local GDP) with the following regression estimated in levels:

$$\ln G_{it} = \gamma_c \ln C_{i,t-1} + \gamma_m \ln MA_{i,t-1} + \boldsymbol{X}'_{i,t}\, \boldsymbol{\chi} + \boldsymbol{Z}'_{i,t-1} \boldsymbol{\zeta} + \boldsymbol{D}'_i\, \boldsymbol{\xi} + \omega_{i,t} \tag{7}$$

where $\ln C_{i,t-1}$ is natural logarithm of cropland area at date *t*-1, $\ln MA_{i,t-1}$ is the natural logarithm of the market access index at date *t*-1, and ω_{it} is the error term. The other controls are the same as in Regression (4). As we investigate whether cropland expansion is associated with local economic development, our parameter of interest is γ_c. Observe here that we consider market access only as a control as it may exhibit some endogeneity given that our local GDP variable is constructed with a spatial allocation model that takes into account the presence of roads.[6]

3. Data

The above regressions require the use of geo-referenced panel data sets at a common spatial unit, which we present in this section.

For our main explained variable, cropland area, we use the HYDE 3.1 panel dataset (Goldewijk, Beusen, van Drecht and de Vos, 2011), which provides a measure of cultivated land (excluding urban areas and pasture land) in areas of 5 decimal degrees (approximately 10 by 10 km, or 10,000 hectares) for the whole world. We use these data for the subset of 290,416 cells in Sub-Saharan Africa and for the years 1970, 1980, 1990, 2000 and 2005 for which we have road data available. Cropland area in the region increased by 57 million hectares, reaching a total of about 211.5 million hectares in 2005 (see Table A1 in the Appendix). The annual growth rate of cropland area steadily increased over the period from 0.61 percent in the 1970-1980 decade to 1.68 percent in the 2000-2005 period. Map A1 in the Appendix shows the actual spatial distribution of cropland for 1970 and 2005. As can be seem from this map, the increase in cropland area over the period occurred throughout the Guinea Savannah zone that covers a large part of Sub-Saharan Africa (World Bank, 2009). The increase is particularly noticeable in West Africa (in particular for Nigeria, Niger, Ghana, Senegal and Burkina Faso), Central/East Africa (the countries surrounding Lake Victoria, as well as Sudan, Ethiopia and Kenya) and the Southern part of the continent (especially South Africa).

The other regressand (in Regression 7) is a measure of economic activity. For this, we used the local GDP measure constructed by UNEP (United Nations Environment Program) and the World Bank and aggregated it to our spatial unit for the above mentioned years.[7]

Our main variable of interest is the natural logarithm of the market access index defined in Formula (1). It makes uses of the road panel dataset constructed by Jedwab and Storeygard (2016), which we modified to be used with our spatial unit and the dates for data availability in HYDE 3.1. Table A2 in the Appendix presents the evolution of the categories over time for the period 1970-2005. This table clearly provides evidence that the period has significant increases in paved and improved roads, but the majority of roads remain of poor quality (dirt roads / other). Map A2 represents the same changes on a map of Sub-Saharan Africa for 1970 and 2005. The construction of the market access index also makes use of a regional panel database of urban populations in Sub-Saharan Africa—that we assemble using the City Population data—for the same dates and spatial unit as in the HYDE 3.1 data (see Appendices 4 and 5). Map A3 illustrates the increase in urban population in the region between 1970 and 2005.

The control variables that we include in our regressions—average annual rainfall, population density, and distance to nearest major port—are from several sources. We processed monthly rainfall data from PREC/L (Chen, Xie, Janowiak, and Arkin, 2002), which we aggregated to the annual level, averaged over each decade, and aligned with the HYDE spatial unit. The population density was calculated dividing the number of people in a grid cell (available from HYDE) by the land area in that cell. Finally, we also use the Euclidean distance from each grid cell to the nearest major port.

Table A4 in the Appendix presents summary statistics of the variables that we use in the regressions. All values are at the cell level (roughly 10 km × 10 km, or 10,000 ha). Cropland area varies significantly across cells (from uncultivated to almost fully cultivated areas) and its mean increases over the period from 559 to 774 hectares. There are however, a significant number of cells in which cropland decreased, possibly due to the conversion to other land uses or desertification. The table also shows a steady increase in the market access index over the 35-year span of the study. This is due to both improvements in roads and changes in urban population (see Maps A2 and A3 in the Appendix). Changes in the population, in particular, contribute two to three times more than road improvements to variations in the logarithm of the market access index.

4. Analysis and results

We start with the estimation of a baseline Regression (4), which assesses the link (in levels), between market access and cropland area. The impact of market access on cropland is consistently estimated to be positive and significant under both OLS estimation (column 2) and the less biased FE approach (column 4) with controls. Under FE (our preferred specification), a doubling of the lagged market access index—something which would occur after 25 years at the current rate[8]—is associated, on average, with a 1.8 percent increase in the cropland area. This is a very modest effect in comparison with the annual increase in cropland of 1.6 percent occurring over the recent period (see Table A1 in the appendix). To put this into context, given the total cropland of 211 million hectares in Sub-Saharan Africa, a doubling of the market access index would only result in an overall expansion of 3.8 million hectares.

Table 1: Estimates of the impact of market access on cropland

	(1)	(2)	(3)	(4)
	OLS	OLS	FE	FE
Ln $MA_{t\text{-}1}$	0.185***	0.035***	-0.025***	0.018***
	(0.00)	(0.00)	(0.00)	(0.00)
Avg. rainfall $_t$		0.644***		-0.071***
		(0.00)		(0.01)
(*Avg. rainfall*$_t$)^2		-0.063***		0.002***
		(0.00)		(0.00)
Ln pop. density$_{t\text{-}1}$		1.241***		-0.708***
		(0.01)		(0.02)
Ln dist. to major port		-0.088***		
		(0.00)		
Constant	0.983***	0.497***	1.120***	1.351***
	(0.00)	(0.04)	(0.00)	(0.01)
Country × year dummies	No	Yes		
Observations	1,161,664	1,160,235	1,161,664	1,160,235
Number of groups			290,416	290,130
R-squared	0.073	0.482	0.000	0.010

Notes: This table presents estimates from OLS (columns 1 and 2) and fixed effect (columns 3 and 4) regressions of the natural logarithm of cropland area at time *t* on the natural logarithm of the lagged market access index (*Ln MA* $_{t\text{-}1}$). Controls included in the OLS regression (column 2): average rainfall over the period (*Avg. rainfall* $_t$) and its square, natural logarithm of lagged population density (*Ln pop. density* $_{t\text{-}1}$), natural logarithm of distance to nearest major port (*Ln dist. to major port*), country dummies, year dummies, and country × year dummies. Controls included in the FE regressions (column 4): period average rainfall and its square, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

We now turn to the estimation of the impact of the change in the market access on the change in cropland area (Regression 5). Under this specification, doubling of the market access index is associated with an increase in cropland area by 6.3 percent.[9] This is more than 3 times the impact estimated in levels under Regression (4). Note that in Regression (5), we have introduced a dummy variable for cells where cropland had been decreasing in the previous period. We find that for these cells the improved market access actually accelerates the reduction in cropland area.

Table 2: Estimates of the impact of the change in market access on the change in cropland area

	(1) OLS	(2) OLS	(3) FE	(4) FE
Δ Ln MA $_t$	0.019***	0.127***	0.033***	0.088***
	(0.00)	(0.00)	(0.00)	(0.00)
Shrinking cropland $_{t\text{-}1}$		-0.346***		-0.205***
		(0.00)		(0.00)
Δ Ln MA $_t$ × shrinking cropland $_{t\text{-}1}$		-0.336***		-0.272***
		(0.01)		(0.01)
Ln C $_{t\text{-}1}$		-0.064***		-0.328***
		(0.00)		(0.00)
Avg. rainfall $_t$		0.117***		-0.201***
		(0.00)		(0.01)
(Avg. rainfall $_t$)^2		-0.011***		0.016***
		(0.00)		(0.00)
Ln pop. density $_{t\text{-}1}$		-0.129***		-0.350***
		(0.00)		(0.01)
Ln dist. to major port		-0.002**		
		(0.00)		
Constant	0.012***	-0.164***	0.011***	0.825***
	(0.00)	(0.01)	(0.00)	(0.01)
Country × year dummies	No	Yes		
Observations	1,161,664	1,160,235	1,161,664	1,160,235
Number of groups			290,416	290,130
R-squared	0.000	0.164	0.000	0.141

Notes: This table presents estimates from OLS (columns 1 and 2) and fixed effect (columns 3 and 4) regressions of the change in the natural logarithm of the cropland area between years t and t – 1 on the change in the natural logarithm of market access between year t and t – 1 (*Δ Ln MA* $_t$). Controls included in the OLS regression (column 2): dummy variable indicating a decrease in cropland during the previous period (*Shrinking cropland* $_{t\text{-}1}$), interaction term between the change in the natural logarithm of market access and the dummy variable (*Δ Ln MA* $_t$ × *Shrinking cropland* $_{t\text{-}1}$), natural logarithm of lagged cropland area (*Ln* $C_{t\text{-}1}$), average precipitation over the period (*Avg. rainfall* $_t$) and its square, natural logarithm of lagged population density (*Ln pop. density* $_{t\text{-}1}$), natural logarithm of distance to nearest major port (*Ln dist. to major port*), country dummies, year dummies, and country × year dummies. Controls included in the FE regression (column 4): dummy variable indicating a decrease in cropland during the previous period, interaction term between the change in the natural logarithm of market access and the dummy variable, natural logarithm of lagged cropland area, natural logarithm of average precipitation over the period, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

Table 3 presents the results from Regression (6), assessing the impact of a change in market access on the change in cropland area, distinguishing between what is due to the change in roads and what is the due to the change in the population distribution. The impact is positive for both components, but most of the effect appears to be due to the change in population, not the change in roads.[10] This is qualitatively consistent with the findings of Blankespoor et al. (2016). We find here that an increase in market access due to changes in the road network has a much smaller effect on cropland expansion compared to an increase in market access due to changes in the population distribution. This is suggestive that, on average, the extensive margin in agricultural production could respond mostly to local population demand. This is consistent with possible barriers to trade beyond the local vicinity.

Although it was long taken for granted that, because of stagnating yields, production mostly increased at the extensive margin in Sub-Saharan Africa (Deininger et al., 2011), recent studies have found that increases have occurred through both the productive and intensive margins, starting in the 1990s (see Seck et al., 2013, and Goyal and Nash, *forthcoming*). To test whether this has an impact on the relation between market access and cropland we stratified the sample into two subperiods—1970-1990, during which yields were stagnating, and 1990-2005, during which some increases in yields have been observed—and re-estimated Regressions (5) and (6) for each subsample. Results (see Table A5 and A6 in the Appendix) show that the significant and positive relation remains. Interestingly, the coefficients on the change in total market access (Table A5) and on the roads and population components of the change in market access (Table A6) are greater for the most recent period, suggesting that the extensive margin response to market access has strengthened, if anything.

Table 3: Estimates of the impact of the change in market access on cropland area (roads and pop.)

	(1) OLS	(2) OLS	(3) FE	(4) FE
Δ_{road} Ln MA $_t$	0.026***	0.016***	0.032***	0.007***
	(0.00)	(0.00)	(0.00)	(0.00)
$\Delta_{pop.}$ Ln MA $_t$	0.001	0.457***	0.042***	0.418***
	(0.00)	(0.00)	(0.01)	(0.01)
Shrinking cropland $_{t\text{-}1}$		-0.330***		-0.212***
		(0.00)		(0.00)
Δ_{road} Ln MA$_t$ × shrinking cropland $_{t\text{-}1}$		0.015***		0.021***
		(0.01)		(0.01)
$\Delta_{pop.}$ Ln MA$_t$ × shrinking cropland $_{t\text{-}1}$		-1.218***		-1.158***
		(0.02)		(0.01)
Ln cropland $_{t\text{-}1}$		-0.060***		-0.312***
		(0.00)		(0.00)
Avg. rainfall $_t$		0.113***		-0.203***
		(0.00)		(0.01)
(Avg. rainfall $_t$)^2		-0.010***		0.016***
		(0.00)		(0.00)
Ln pop. density $_{t\text{-}1}$		-0.129***		-0.280***
		(0.00)		(0.01)
Ln dist. to major port		-0.002*		
		(0.00)		
Constant	0.012***	-0.160***	0.010***	0.806***
	(0.00)	(0.01)	(0.00)	(0.01)
Country × year dummies	No	Yes		
Observations	1,161,664	1,160,235	1,161,664	1,160,235
Number of groups			290,416	290,130
R-squared	0.000	0.190	0.000	0.165

Notes: This table presents estimates from OLS (columns 1 and 2) and fixed effect (columns 3 and 4) regressions of the change in the natural logarithm of cropland area between year t and t – 1 on the change in the natural logarithm of market access due to roads between year t and t – 1 (Δ_{road} *Ln MA* $_t$) and due to population between year t and t – 1 ($\Delta_{pop.}$ *Ln MA* $_t$). Controls included in the OLS regression (column 2): dummy variable indicating a decrease in cropland during the previous period (Shrinking cropland $_{t\text{-}1}$), interaction term between the change in the natural logarithm of market access due to roads and the dummy variable (Δ_{road} *Ln MA* $_t$ × *Shrinking cropland* $_{t\text{-}1}$), interaction term between the change in the natural logarithm of market access due to population and the dummy variable ($\Delta_{pop.}$ *Ln MA* $_t$ × *Shrinking cropland* $_{t\text{-}1}$), natural logarithm of lagged cropland area (*Ln MA* $_{t\text{-}1}$), natural logarithm of average precipitation over the period (*Ln avg. rainfall* $_t$), natural logarithm of lagged population density (*Ln pop. density* $_{t\text{-}1}$), natural logarithm of distance to nearest major port (*Ln dist. to major port*), country dummies, year dummies, and country × year dummies. Controls included in the FE regressions (column 4): dummy variable indicating a decrease in cropland during the previous period, interaction term between the change in the natural logarithm of market access due to roads and the dummy variable, interaction term between the change in the natural logarithm of market access due to population and the dummy variable, natural logarithm of lagged cropland area, natural logarithm of average precipitation over the period, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

Our results presented in Tables 1-3 points toward a modest but significant relationship between improved market access in Sub-Saharan Africa and the land under cultivation. This however does not tell us the extent to which this association may involve structural transformation (through the shift towards commercial agriculture and non-agricultural jobs) and higher local incomes. In contrast to Blankespoor et al. (2016) who examine the issue of structural transformation induced by improved accessibility by looking at changes in the local structure of employment at the scale of a single country, this is not possible in our case because such data do not exist at the regional scale. Instead, we make use of a local GDP measure to investigate the links between cropland expansion and local economic activity.[11] The results from Regression (7) are reported in Table 4 for OLS and FE. Our preferred specifications, the FE regressions with controls (column 5 and 6), provide consistent estimates of the association between cropland expansion and local GDP irrespective of the inclusion or the exclusion of the market access control.[12] According to these estimates, a doubling of cropland area is associated with a modest increase of 2-3 percent in local GDP. This is suggestive of land used for subsistence rather than commercial agriculture and consistent with our previous result that cropland expansion mainly responds to local population growth.

Finally, note that as a robustness check, we re-ran all the above regressions with an alternative market access index calculated with a trade elasticity of θ=8.2 (as in Eaton and Kortum, 2002) instead of 3.8 (as in Donaldson, *forthcoming*). Under this more rapid decay function, the market access index becomes more dependent on the existence of roads and population in the local vicinity of the grid cell. Table A7 in the Appendix reports the estimated coefficients for the variable of interest in each one of the FE regressions previously reported in Tables 1-4. It appears that all our results are robust to this check.

Table 4: Regression of local GDP on cropland

	(1) OLS	(2) OLS	(3) OLS	(4) FE	(5) FE	(6) FE
Ln cropland $_{t\text{-}1}$	1.913***	0.956***	0.946***	0.024***	0.027***	0.024***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Ln MA $_{t\text{-}1}$			0.075***			0.160***
			(0.00)			(0.00)
Avg. rainfall $_t$		2.205***	2.191***		-0.734***	-0.707***
		(0.01)	(0.01)		(0.01)	(0.01)
(Avg. rainfall $_t$)^2		-0.226***	-0.225***		0.049***	0.048***
		(0.00)	(0.00)		(0.00)	(0.00)
Ln pop. density $_{t\text{-}1}$		2.045***	1.917***		1.086***	0.853***
		(0.01)	(0.01)		(0.01)	(0.01)
Ln dist. to major port		-0.578***	-0.556***			
		(0.00)	(0.00)			
Constant	6.054***	11.093***	10.841***	8.115***	9.370***	9.238***
	(0.00)	(0.06)	(0.06)	(0.00)	(0.02)	(0.02)
Country × year dummies	No	Yes	Yes			
Observations	1,161,628	1,160,235	1,160,235	1,161,628	1,160,235	1,160,235
Number of groups				290,407	290,130	290,130
R-squared	0.365	0.621	0.622	0.000	0.031	0.039

Notes: This table presents estimates from OLS (columns 1, 2 and 3) and fixed effect (columns 4, 5 and 6) regressions of the logarithm of GDP at time t on the natural logarithm of cropland area at time $t-1$ (*Ln cropland* $_{t\text{-}1}$). Controls included in the OLS regression (columns 2 and 3): natural logarithm of lagged cropland area (*Ln cropland* $_{t\text{-}1}$), natural logarithm of lagged market access (*Ln MA* $_{t\text{-}1}$), natural logarithm of average precipitation over the period (*Ln avg. rainfall* $_t$), natural logarithm of lagged population density (*Ln pop. density* $_{t\text{-}1}$), natural logarithm of distance to nearest major port (*Ln dist. to major port*), country dummies, year dummies, and country × year dummies. Controls included in the FE regressions (columns 5 and 6): logarithm of lagged cropland area, natural logarithm of lagged market access, natural logarithm of average precipitation over the period, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

5. Conclusion

This paper is an initial attempt to assess the effect of accessibility to transport on rural development in Sub-Saharan Africa. To the best of our knowledge, our paper is the first to investigate the relation between access to markets from road improvements and the spatial expansion of cultivated land using geo-referenced panel data. Our analysis was carried out for a 35-year span when roads were constructed and cropland area expanded in the region quantifies the link between the two. In accordance with the theory, we find suggestive evidence that improved market access leads to more land put into cultivation. The effect, however, is quite modest as a doubling of our measure of market access leads to an increase in cropland by approximately 2 to 6 percent. This, however, seems driven by increases in local population much more than by improved accessibility from better roads, which is suggestive of agriculture mainly responding to local demand. Complementary to this result, we also find that cropland expansion only marginally increases local GDP, which is consistent with the idea that a significant part of cropland expansion in Sub-Saharan Africa is not accompanied by the development of commercial agriculture and that trade of agricultural goods is to a certain extent limited to the immediate vicinity of production. In the wake of our study on the extensive margin of agriculture (cropland expansion), further research will be needed to assess whether market access has an impact on the intensive margin (yields), which, however, are known to have been stagnating in the region. More generally, research will also be needed to identify the enabling environment that could allow transport investment to better support rural development in the region.

---

[1] Based on calculations from Fischer and Shah (2010), Deininger et al. (2011) report figures for suitable land in non-forest areas under different population density criteria: 68 million hectares (respectively 201 million hectares) for areas of less than 5 people (respectively 25 people) per km^2 (see Table A2.6 page 165).

[2] Ideally, the measure should account for the income of potential consumers, but this information is usually not available, hence the use of population numbers. Examples of papers using a similar market access measure include

Harris (1954), Hanson (2005), Emran and Shilpi (2012), Dorosh, Wang, You, and Schmidt (2012), Jedwab and Storeygard (2016), and Blankespoor et al. (2016).
[3] We use $1 + \sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta}$ instead of $\sum_{j \neq i} P_{j,t} \tau_{ij,t}^{-\theta}$ so as to be able to calculate the natural logarithm of the market access index even when the weighted sum of the populations is equal to zero, which can occur as we restrict the calculation of the market access index to travel times of six hours or less.
[4] We use time indexes t and t-1 to refer to the years in our data (1970, 1980, 1990, 2000 or 2005; see Section 3). Note that in Formula (1), we exclude the population of the locality and use travel times based on roads prior to t-1 (see Appendix), which addresses endogeneity concerns in the regressions.
[5] Major ports are defined as ports that include direct or trans-shipment capacity as measured in the AICD report (Foster and Briceño-Garmendia 2010).
[6] The problem is attenuated by the fact that we use the *lagged* value of the market access variable.
[7] The number of people residing in each location is determined from LandScan and UNEP/GRID-Geneva. The urban/rural dichotomy is based on a density threshold. These information are given at the 30 arc second (approximately 1 by 1 km). The spatial model to allocate national or subnational GDP across space uses subnational population data from LandScan within urban and rural strata and does not make any direct use of roads or cropland.
[8] The average market access index grows at 3.95 per cent annually.
[9] We have $2^{0.088}$-1≈0.063.
[10] The estimates show that a marginal increase in $\Delta_{road} \ln MA_{i,t}$ has a smaller impact than a comparable increase in $\Delta_{pop.} \ln MA_{i,t}$. In our sample, the standard deviation in $\Delta_{road} \ln MA_{i,t}$ is also small than the standard deviation in $\Delta_{pop.} \ln MA_{i,t}$, from which we conclude that changes in population had more impact on cropland expansion than road improvements.
[11] Any measure of local GDP based on remote sensing, such as ours, is subject to measurement error (Chuhan-Pole, Dabalen, and Land, 2015). Even so, it is the best measure of local economic activity currently available and so we present it as suggestive evidence.
[12] Because of possible endogeneity concerns due to the construction of the local GPD variable as a function of the surrounding population (see endnote 7 in Section 3), the market access variable may itself be endogenous. We therefore interpret it as a control included to attenuate omitted variable bias.

References

Ali, R., Barra, A.F., Berg, C.N., Damania, R., Nash, J., & Russ, J. (2015). *Highway to Success or Byways to Waste*. Washington D.C.: The World Bank, 189 pages.

Berg, C.N., Deichmann, U., Liu, Y., & Selod, H. (forthcoming). Transport Policies and Development. *Journal of Development Studies*.

Blankespoor, B., Mesplé-Somps, S., Selod, H. & Spielvogel, G. (2016). The Dynamics of Villages, Towns and Cities in Sub-Saharan Africa: The Case of Mali. Mimeo.

Blimpo, M., Harding, R. & Wantchekon, L. (2013). Public Investment in Rural Infrastructure: Some Political Economy consideration. *Journal of African Economies*, 22(2), 57-83

Brinkhoff, T. City Population. http://www.citypopulation.de/

Buys, P., Deichmann, U., & Wheeler, D. (2010). Road network upgrading and overland trade expansion in Sub-Saharan Africa. *Journal of African Economies*, 19(3), 399-432.

Chen, M., Xie, P., Janowiak, J.E. & Arkin, P.A., 2002. Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology,* 3(3)**,** pp. 249-266.

Chuhan-Pole, P., Dabalen, A., & Land., B. (2015) *Socioeconomic Impact of Mining on Local Communities in Africa*, Report No: ACS14623, Washington DC: The World Bank, 173 pages.

Collier, P., Kirchberger, M., & Söderbom, M. (2015) The Cost of Road Infrastructure in and Low Middle Income Countries. *World Bank Economic Review*, doi:10.1093/wber/lhv037.

Damania, R., Barra, A.F., Berg, C.N., Russ, J., Nash, J. & Ali, R. (*forthcoming*). Agricultural Technology Choice and Transport. *American Journal of Agricultural Economics*.

Deininger, K., Byerlee, D., Lindsay, J., Norton, A., Selod, H. & Stickler, M. (2011). *Rising Global Interest in Farmland: Can it Yield Sustainable and Equitable Benefits?* The World Bank.

Diao, X., Hazell, P., Resnick, D. & Thurlow, J. (2007). *The role of agriculture in development: Implications for Sub-Saharan Africa*, International Food Policy Research Institute, Research Report 153, 66 pages.

Donaldson, D. (forthcoming) Railroads and the Raj: Estimating the impact of transportation infrastructure. *American Economic Review*, in press.

Dorosh, P., Wang, H.G., You, L., & Schmidt, E. (2012). Road connectivity, population, and crop production in Sub-Saharan Africa. *Agricultural Economics.* 43. pp. 89-103.

Eaton, J. & Kortum, S. (2002). Technology, Geography and Trade. *Econometrica*, 70(5) 1741-1779.

Emran, M.S., & Shilpi, F. (2012). The extent of the market and stages of agricultural specialization. *Canadian Journal of Economics*, 45(3) 1125-1153.

Fischer, G. & Shah, M. (2010). Farmland Investments and Food Security: A Statistical Annex. Report prepared under a World Bank and International Institute for Applied Systems Analysis, Vienna.

Foster, V., & Briceño-Garmendia, C. (2010). *Africa's Infrastructure: A Time for Transformation*. The World Bank Africa Development Forum.

Gachassin, M. (2013). Should I Stay or Should I Go: The role of roads in Migration Decisions. *Economic Development and Cultural Change,* 22(5), 796-826.

Goyal, A., & Nash, J., (*forthcoming*). "Public Spending priorities for Transforming African Agriculture," World Bank AFR Regional Study.

Goldewijk, K., Beusen, A., Van Drecht, G. & De Vos, M., 2011. The HYDE 3.1 spatially explicit database of human-induced global land-use change over the past 12,000 years. *Global Ecology and Biogeography,* 20(1)**,** pp. 73-86.

Harris, C.D. (1954) The Market as a Factor in the Localization of Industry in the United States. *Annals of the Association of American Geographers*, 44, 315-348.

Hanson, G. (2005). Market Potential, Increasing Returns, and Geographic Concentration. *Journal of International Economics*, 67, 1-24.

Jacoby, H. (2000). Access to Markets and the Benefits of Rural Roads. *The Economic Journal*, 110(465), pp. 713-737.

Jacoby, H., &Minten, B. (2009). On measuring the benefits of lower transport costs. *Journal of Development Economics*, 89(1), pp. 28-38.

Jedwab, R., & Storeygard, A. (2016). The Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa. *mimeo*, presented at the GWU/World Bank 3rd Urbanization and Poverty Reduction Research Conference (February 1, 2016).

Kyeyamwa, H., Speelman, S., Huylenbroeck, G.V., Opuda-Asibo, J., & Verbeke, W. (2008). Rising offtake from cattle grazed on natural rangelands in Sub-Saharan Africa: a transaction cost economic approach. *Agricultural Economics*, 39, 63-72.

Minten, B., Koru, B., & Stifel, D. (2013). The last mile(s) in modern input distribution: Pricing, profitability, and adoption, *Agricultural Economics*, 44, 629-646.

Mu, R., & van de Walle, D. (2011) Rural Roads and Local Market Development in Vietnam. *The Journal of Development Studies*. 47(5), pp. 709-34.

Murphy, K., Shleifer, A., & Vishny, R. (1989). Industrialization and the Big Push. *Journal of Political Economy*. 97(5), 1003-26.

Nelson, A., & Deichmann, U. (2004). The African Population Database. United Nations Environment Program (UNEP) and the Center for International Earth Science Information Network (CIESIN), Columbia University; New York.

Raballand, G., Macchi, P., & Petracco, C. (2010). *Rural Road Investment Efficiency: Lessons from Burkina Faso, Cameroon, and Uganda*. Washington, D.C.: World Bank.

Rakotoarisoa, M., Lafrate, M., & Paschali, M. (2011). *Why has Africa Become a Net Food Importer? Explaining Africa Agricultural and Food Trade Deficit*. Rome: FAO, 85 pages.

Seck, P., Touré, A., Coulibaly, J., Diagne, A., & Wopereis, M. (2013) "Africa's Rice Economy Before and After the 2008 Rice Crisis," ch. 2 in *Realizing Africa's Rice Promise* (eds. Wopereis, M., Johnson, D., Ahmadi, N., Tollens, E., & Jalloh, A.) CAB International.

Wantchekon, L., & Stanig, P. (2015). The Curse of Good Soil? Land Fertility, Roads, and Rural Poverty in Africa. Mimeo.

World Bank (2009). *Awakening Africa's Sleeping Giant: Prospects for Commercial Agriculture in the Guinea Savannah Zone and Beyond.* Directions in Development; Agriculture and Development, 218 pages.

World Bank and UNEP (2010). Distributed global GDP dataset Sub-national GRP. *In* World Bank and UNEP (ed) (UNEP).

Appendix

1. Tables

Table A1: Total Cropland area in Sub-Saharan Africa (million ha), 1970-2005

	Total cropland area (in '000 hectares)	Annual growth rate (in percent)
2005	211,539	1.68
2000	194,658	0.94
1990	177,358	0.78
1980	164,089	0.61
1970	154,466	

Source: HYDE 3.1 (Goldewijk et al., 2011); calculations by authors.

Table A2: Road network in Sub-Saharan Africa (thousand kilometers), 1970-2005

Variable	1970	1980	1990	2000	2005
Highways	-	1.3	2.6	3.2	3.1
Paved	77.8	120.4	168.7	180.1	182.4
Improved	142.0	153.7	152.0	154.8	154.8
Other	899.0	843.3	795.6	780.7	778.5

Note: These figures use the cross-sectional road network geometry of Buys, Deichmann and Wheeler (2010) updated by Jedwab and Storeygard (2016). A highway is a paved road with at least three lanes on each side. An improved road is laterite or gravel. The category 'other' include and any segment not identified as a highway, a paved road, or an improved road, especially dirt roads.
Source: Jedwab and Storeygard (2016). Calculations by authors.

Table A3: Average change in the natural logarithm of the market access index

	Average change in the natural logarithm of the market access index		
	Due to roads	Due to population	Overall
2000-2005	0.007	0.028	0.034
1990-2000	0.006	0.056	0.062
1980-1990	0.031	0.064	0.094
1970-1980	0.031	0.068	0.099

Note: Values in this table are averages of values at the cell level (approximately 10 km × 10 km).
Source: HYDE 3.1; calculation by authors.

Table A4: Summary statistics (cell level)

Variable	1970	1980	1990	2000	2005
Cropland (hectares)					
Mean	532	565	611	670	728
Min	0	0	0	0	0
Max	8,586	8,586	8,586	8,536	8,584
Local GDP (millions of constant 2000 USD)					
Mean	0.5	0.8	0.9	1.1	1.4
Min	0.0	0.0	0.0	0.0	0.0
Max	2,306.5	3,221.5	3,684.4	4,363.0	5,375.2
Ln MA					
Mean	0.52	0.61	0.71	0.77	0.80
Min	0.00	0.00	0.00	0.00	0.00
Max	11.77	12.28	12.88	13.30	13.49
Δ Ln MA					
Mean		0.10	0.09	0.06	0.03
Min		-6.64	-4.99	-4.17	-3.62
Max		5.98	6.74	6.03	7.46
Δ_{road} Ln MA					
Mean		0.03	0.03	0.01	0.01
Min		-6.08	-5.30	-4.47	-3.80
Max		5.60	6.71	5.91	7.38
$\Delta_{pop.}$ Ln MA					
Mean		0.07	0.06	0.06	0.03
Min		-6.65	-0.39	-0.64	-0.41
Max		2.41	2.22	1.88	0.96
Shrinking cropland dummy					
Mean		0.27	0.29	0.41	0.02
Min		0	0	0	0
Max		1	1	1	1
Avg. rainfall (mm)					
Mean	281	267	258	260	260
Min	1	0	0	1	0
Max	1271	1115	1143	1118	1165
Population density (people/km^2)					
Mean	0.15	0.20	0.27	0.36	0.41
Min	0.00	0.00	0.00	0.00	0.00
Max	688.36	1259.92	2088.86	2860.15	3171.12
Distance to major port (km)					
Mean	800	800	800	800	800
Min	0	0	0	0	0
Max	2,216	2,216	2,216	2,216	2,216

Notes: All values in this table are averages at the cell level (approximately 10 km × 10 km). *Cropland* stands for cropland area of the cell. *Ln MA* (respectively Δ *Ln MA*) is the natural logarithm of the market access index (respectively the change in the natural logarithm of the market access index over the period). Δ_{road} *Ln MA* (respectively $\Delta_{pop.}$ *Ln MA*) is the change in the logarithm of the market access index due to changes in roads (respectively changes in population) over the period, holding the population distribution constant at the initial date (respectively holding the road network constant at the final date). *Shrinking cropland dummy* takes value 1 if the cropland area in the cell decreased over the period. *Avg. rainfall* is the yearly average precipitation in the cell during the period. *Population density* is the ratio of population to the cell area. *Distance to major port* is the distance to the nearest major port listed in the Africa Infrastructure Country Diagnostic (see Foster and Briceño-Garmendia, 2010).

Source: HYDE 3.1 (Goldewijk et al., 2011), UNEP/World Bank, Chen et al. (2002), Jedwab and Storeygard (2016), and Nelson and Deichmann (2004). Calculations by authors.

Table A5: Estimates of the impact of the change in market access on the change in cropland area (FE regressions, by subperiod)

	(1) 1970-1990	(2) 1990-2005
Δ Ln MA $_t$	0.019***	0.094***
	(0.00)	(0.00)
Shrinking cropland $_{t\text{-}1}$	0.017***	-0.161***
	(0.00)	(0.00)
Δ Ln MA $_t$ × shrinking cropland $_{t\text{-}1}$	-0.060***	-0.250***
	(0.00)	(0.00)
Ln C $_{t\text{-}1}$	-0.416***	-0.490***
	(0.00)	(0.00)
Avg. rainfall $_t$	-0.100***	0.021***
	(0.00)	(0.01)
(Avg. rainfall $_t$)^2	0.004***	0.003***
	(0.00)	(0.00)
Ln pop. density $_{t\text{-}1}$	-0.056***	-0.339***
	(0.01)	(0.01)
Constant	0.709***	0.561***
	(0.00)	(0.01)
Observations	580,103	870,173
Number of groups	290,062	290,130
R-squared	0.210	0.272

Notes: This table presents estimates from fixed effect regressions of the change in the natural logarithm of the cropland area between years t and t – 1 on the change in the natural logarithm of market access between year t and t – 1 (*Δ Ln MA $_t$*) for the subperiods 1970-1990 (column 1) and 1990-2005 (column 2). Controls included in the regression: dummy variable indicating a decrease in cropland during the previous period, interaction term between the change in the natural logarithm of market access and the dummy variable, natural logarithm of lagged cropland area, natural logarithm of average precipitation over the period, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

Table A6: Estimates of the impact of the change in market access on cropland area (roads and pop., FE regressions by subperiod)

	(1) 1970-1990	(2) 1990-2005
Δ_{road} Ln MA $_t$	0.006***	0.006**
	(0.00)	(0.00)
$\Delta_{pop.}$ Ln MA $_t$	0.113***	0.491***
	(0.00)	(0.01)
Shrinking cropland $_{t-1}$	0.013***	-0.173***
	(0.00)	(0.00)
Δ_{road} Ln MA$_t$ × shrinking cropland $_{t-1}$	-0.012***	0.061***
	(0.00)	(0.00)
$\Delta_{pop.}$ Ln MA$_t$ × shrinking cropland $_{t-1}$	-0.281***	-1.279***
	(0.00)	(0.01)
Ln cropland $_{t-1}$	-0.405***	-0.468***
	(0.00)	(0.00)
Avg. rainfall $_t$	-0.103***	0.024***
	(0.00)	(0.01)
(Avg. rainfall $_t$)^2	0.004***	0.003***
	(0.00)	(0.00)
Ln pop. density $_{t-1}$	-0.035***	-0.310***
	(0.01)	(0.01)
Constant	0.702***	0.525***
	(0.00)	(0.01)
Observations	580,103	870,173
Number of groups	290,062	290,130
R-squared	0.230	0.292

Notes: This table presents estimates from fixed effect regressions of the change in the natural logarithm of cropland area between year t and t – 1 on the change in the natural logarithm of market access due to roads between year t and t – 1 (Δ_{road} *Ln MA* $_t$) and due to population between year t and t – 1 ($\Delta_{pop.}$ *Ln MA* $_t$) for the subperiods 1970-1990 (column 1) and 1990-2005 (column 2). Controls included in the regression: dummy variable indicating a decrease in cropland during the previous period, interaction term between the change in the natural logarithm of market access due to roads and the dummy variable, interaction term between the change in the natural logarithm of market access due to population and the dummy variable, natural logarithm of lagged cropland area, natural logarithm of average precipitation over the period, and natural logarithm of lagged population density. Robust standard errors are in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

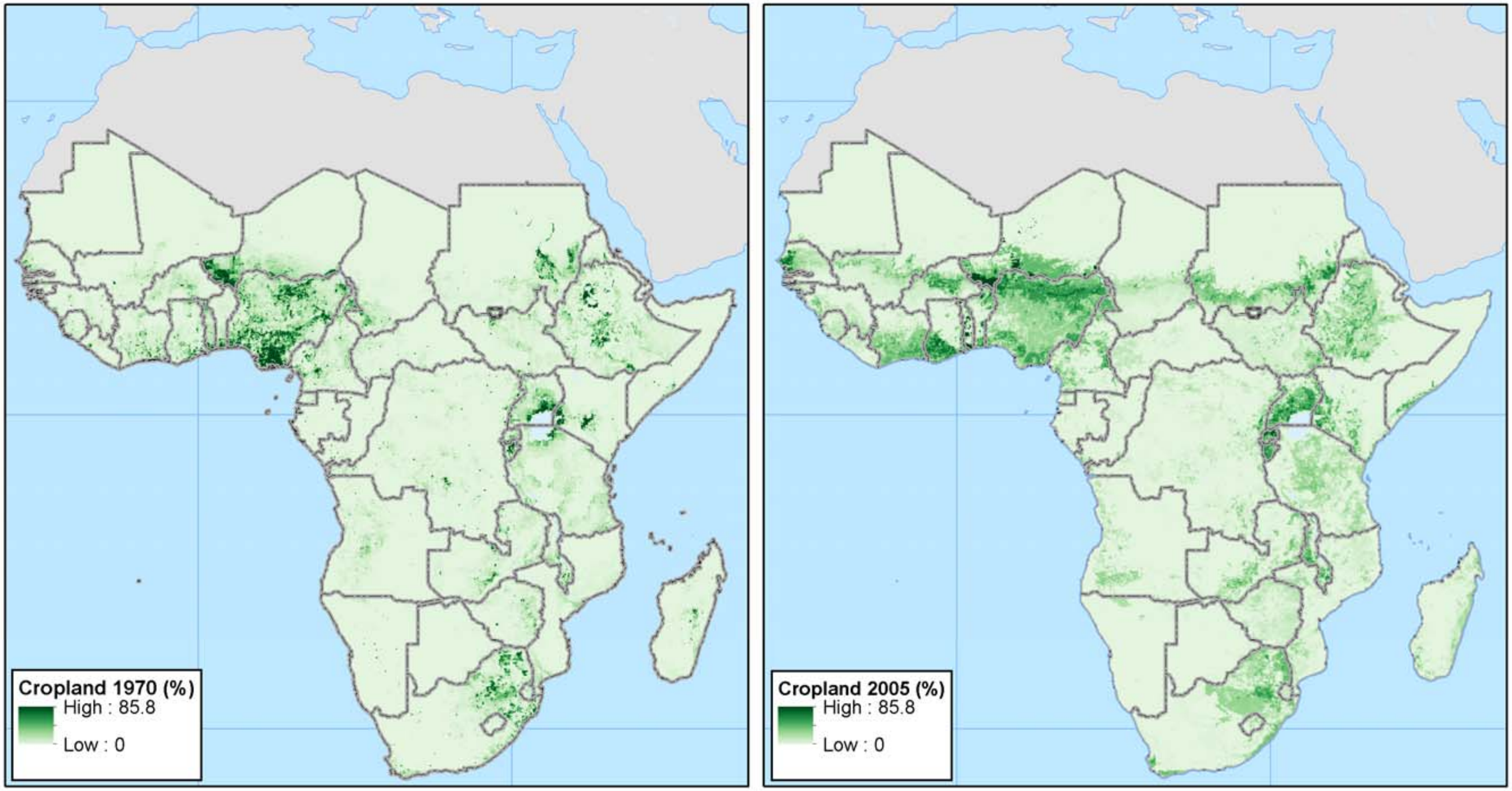
Table A7: Estimates of the effect of market access on cropland and on local GDP under different trade elasticity scenarios

Corresponding Table	Variable of interest	(1)	(2)
		$\theta = 3.8$	$\theta = 8.2$
Table 1	$Ln\ MA_{t\text{-}1}$	0.018*** (0.00)	0.018*** (0.00)
Table 2	$\Delta Ln\ MA_t$	0.088*** (0.00)	0.061*** (0.00)
Table 3	$\Delta_{road}\ Ln\ MA_t$	0.007*** (0.00)	0.015*** (0.00)
	$\Delta_{pop.}\ Ln\ MA_t$	0.418*** (0.01)	0.413*** (0.01)
Table 4	$Ln\ C_{t\text{-}1}$	0.024*** (0.00)	0.025*** (0.00)

Notes: Column (1) reports the coefficients of interest from the FE regressions with controls in Table 1-4 under our assumption of $\theta = 3.8$. Column (2) presents the coefficients from the same regression when the market access index is calculated with an alternative trade elasticity $\theta = 8.2$. Robust standard errors in parenthesis, where *** is significance at the 1% level, ** significance at the 5% level, and * significance at the 10% level.

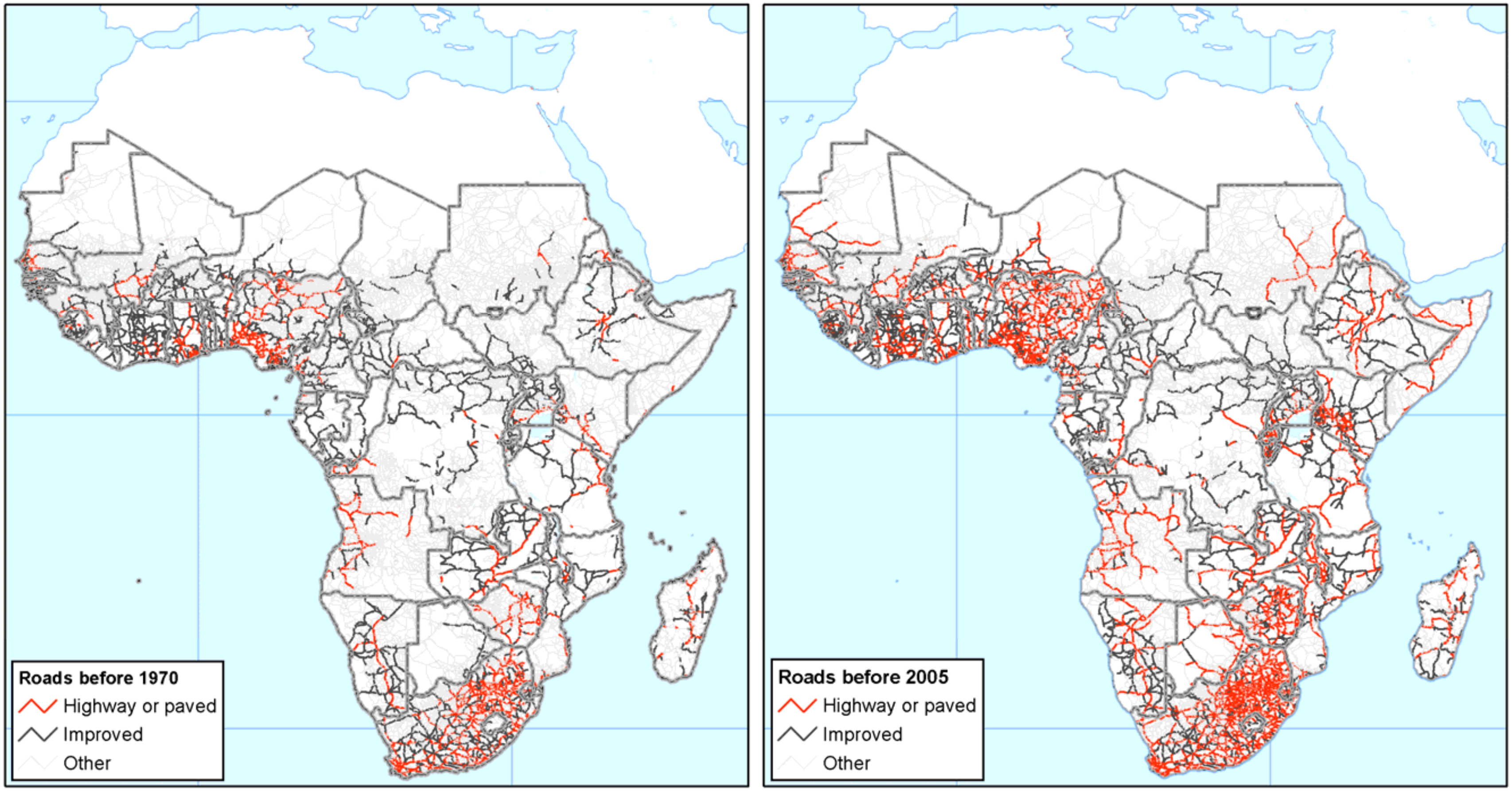
2. Maps

Map A1: Cropland in Sub-Saharan Africa, 1970 and 2005



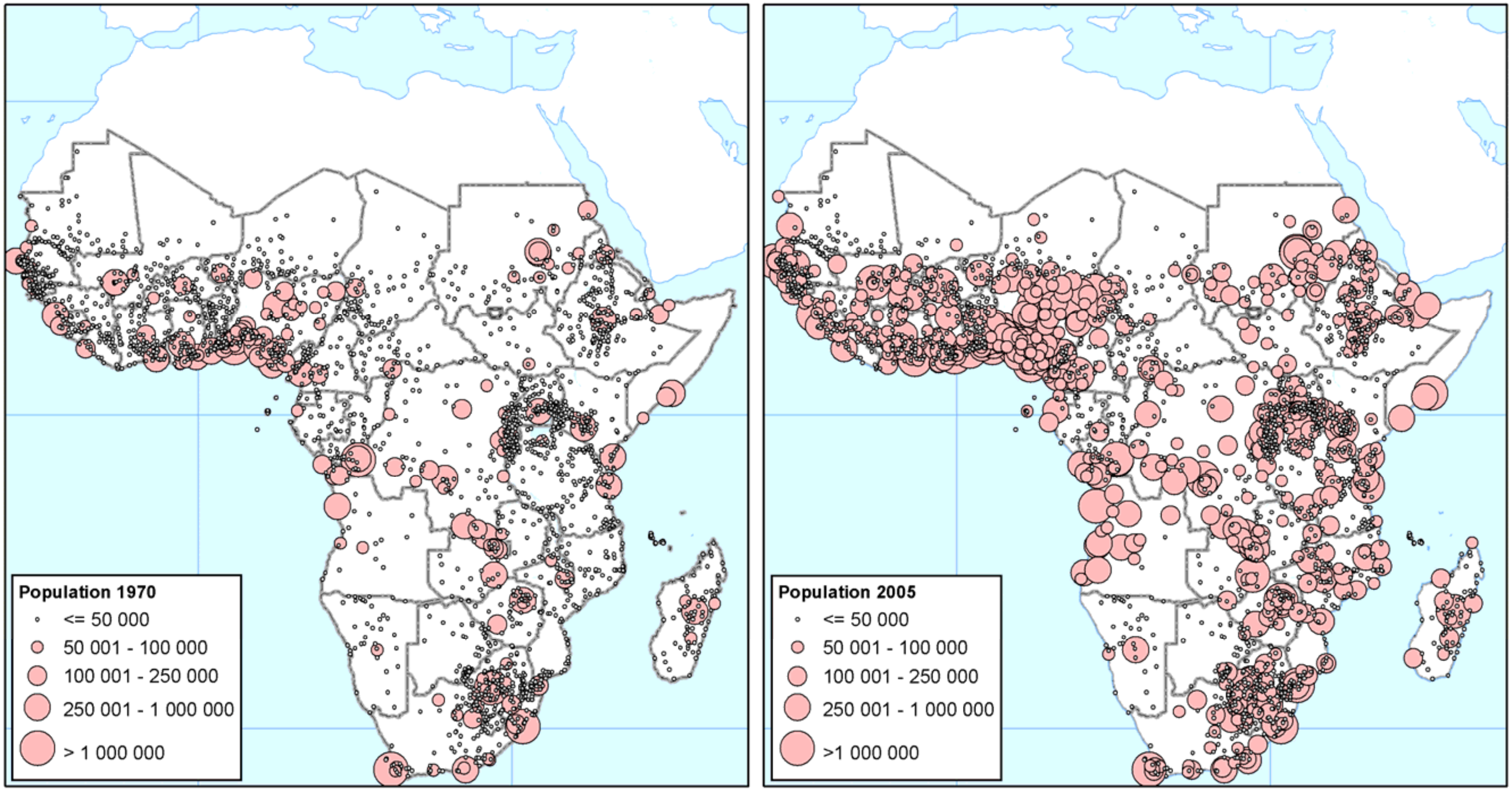
Source: HYDE 3.1 (Goldewijk et al., 2011) and World Bank.
Note: This map represents the percentage of a grid cell under cropland.

Map A2: The road network in Sub-Saharan Africa, 1970 and 2005



Source: Nelson and Deichmann (2004) and Jedwab and Storeygard (2016).
Note: These maps represent a cross-sectional road geometry derived from Buys, Deichmann and Wheeler 2010 and updated by Jedwab and Storeygard (2016). ‘Roads before 1970’ include roads identified on 1968 and 1969 maps. ‘Roads before 2005’ include roads identified on 2003 maps. A highway is a paved road with at least three lanes on each side. An improved road is laterite or gravel. The category ‘other’ include and any segment not identified as a highway, a paved road, or an improved road, especially dirt roads.

Map A3: Cities in Sub-Saharan Africa, 1970 and 2005



Source: Maps by authors. Data derived from Thomas Brinkhoff, City Population, http://www.citypopulation.de (see Appendix for details).

3. Construction of the urban population panel, 1969-2016

We constructed a panel of major urban population places for cities in Africa, using population data available on the website www.citypopulation.de and geo-referenced each of the cities using coordinate information from DeLorme (http://www.delorme.com) and the Global Rural-Urban Mapping Project (http://sedac.ciesin.columbia.edu/data/collection/grump-v1). For the cities which did not appear on either DeLorme or GRUMP, we used coordinates from www.Wikipedia.com or www.latlong.net. Eighteen cities, for which the coordinates could not be found, were dropped.[13] Since population figures were only available for dates when a census was carried out, populations for intermediary dates were estimated using the following exponential growth function:

$$P_t = P_0 \exp(gt)$$

where P_t is the population at time t, P_0 is the initial population at, g is the population growth rate, and t is the number of years elapsed. In the case of two points or more points in time for a given city, g was calculated from that data. For cases where only one observation was available, we used country-level urban population growth rates published in the World Urbanization Prospect (United Nations, 2012). We used these growth rates to fill in the population values for the years before the first available census and after the last available census.

4. Construction of the market access variable:

The market access (MA) index combines road and population data for the years for which the HYDE data is available (1970, 1980, 1990, 2000 and 2005). In order to compute the MA index, we combined the population data that we constructed for the same dates (see Appendix 4) with the roads panel from Jedwab and Storeygard (2016), which provides geo-referenced information on road categories.[14] We sequentially explain below the modifications that were made to the roads and to the urban population data in order to compute the MA index.

First, we modified the geometry of the Jedwab-Storeygard data to enable network analysis by examining the connection of small gaps in road segments throughout the road network, especially near cities. We then constructed a continent-wide road network for each HYDE year (for example, 1970, 1980, etc.) by considering the road segments and associated functional classes available for the most recent date *prior* to that year (for example, we used road information from 1969 in West Africa to build the 1970 continent-wide road network). For each HYDE year t, we then calculated travel times τ_{ijt} between any two pairs of nodes (i and j) on the reconstructed continent-wide road network.[15]

We also modified the geometry of the urban population data (see Appendix 4 above) by ensuring that each populated place was associated with the nearest node on the reconstructed continent-wide road network, yielding the population measure P_{jt} for each node j. It then became possible to calculate the market access index for all nodes according to Formula (1). We further created a smooth surface for the market access index originating from these nodes using Radial Basis Functions and recovered values for every cell. Finally, we discounted these value by the Euclidian distance of the corresponding cell to the nearest road.

---

[13] The 18 cities which are not geo-referenced are: Tchiamba Nzassi in Congo; Massiogo and Mbera in Mali; Canicado and Nhamayabue in Mozambique; Bouhidide, Diogountourou, Hassi Chegar, Lexeibe, Nbeikett Lehouach in Mauritania; Kamuhanda in Rwanda; Al Husayhisa, Al Huwattah, and Umm Shukah in Sudan; Torgbonbu in Sierra Leone; Mondi Forest in Swaziland; Chiwezi and Kibaigwa in Tanzania.

[14] Jedwab and Storeygard (2016) built their panel road database by digitizing historical Michelin maps to identify the road categories over time to the (static/cross-sectional) geometry from Nelson and Deichmann (2004).

[15] We have about 90,000 nodes. We used ESRI 10.1 Network Analyst to construct a Network dataset with the impedance of travel time by road segment. Along with Jedwab and Storeygard (2016). We assume the following speeds by road category: 80 km/h for a highway, 60 km/h for a paved road, 40 km/h for an improved road, 12 km/h for a dirt road, 6 km/h for the unknown category and 5 km/h in the absence of a road.

Appendix reference list

Dobson, J. E., Bright, E.A., Coleman, P.R., Durfee, R.C., & Worley, B.A. (2000). LandScan: A Global Population Database for Estimating Populations at Risk. *Photogrammetric Engineering & Remote Sensing*, 66(7): 849-857.

United Nations, Department of Economic and Social Affairs, Population Division (2012) World Urbanization Prospects: The 2011 Revision, CD-ROM Edition.