Paving the Way for Development:

The Impact of Road Infrastructure on Agricultural Production and Household Wealth in the Democratic Republic of Congo

John Ulimwengu, Jose Funes, Derek Headey, and Liang You

International Food Policy Research Institute

2033 K Street NW, Washington DC 20006-1002 USA

Abstract

Given its vast land resources and favorable water supply, the Democratic Republic of Congo's (DRC) natural agricultural potential is immense. However, the economic potential of the sector is handicapped by one of the most dilapidated transport systems in the developing world (World Bank, 2006). Road investments are therefore a high priority in the government's investment plans, and those of its major donors. Whilst these are encouraging signs, very little is known about how the existing road network constrains agricultural and rural development, and how these new road investments would address these constraints. To inform this issue the present paper primarily employs GIS-based data to assess the impact of market access on agricultural and rural development (ARD). Compared to existing work, however, the paper makes a number of innovations to improve and extend the generic techniques used to estimate the importance of market access for ARD. First, the DRC road network data is augmented with survey-based data from Minten and Kyle (1999) on agricultural transport times to calculate improved "market access" measures for the DRC. Second, we follow Dorosh et al (2009) in estimating the long run relationship between market access and agricultural production, although we also investigate the relationship with household wealth. Finally, we run simulations of how proposed infrastructure investments would affect market access, and how market access would in turn affect agricultural production and household wealth.

Keywords: Infrastructure, market access, road and river transport, agricultural production, poverty.

1. Introduction

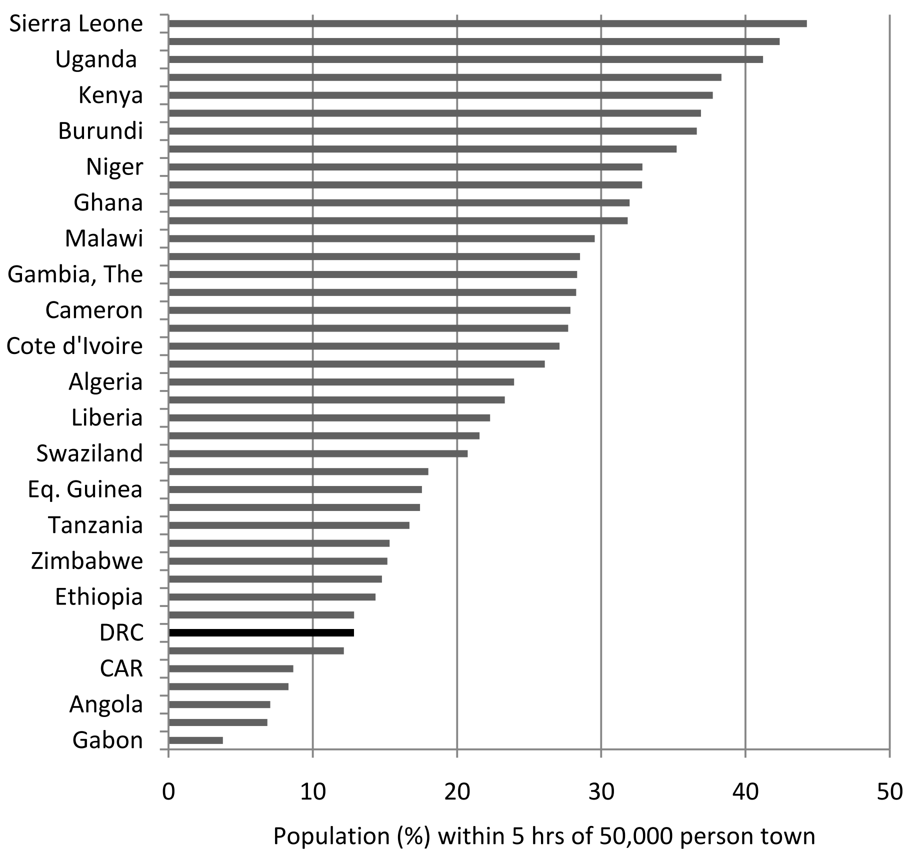
The agricultural potential in the Democratic Republic of Congo (DRC) is immense. By one 'back of the envelope' calculation, if yields in the DRC's 80 million hectares of arable land were to catch up to the global technological frontier, the country could feed around one third of the world's population.[1] But sheer biophysical potential is not the same as economic potential. Decades of conflict, corruption and economic mismanagement have severely weakened the socioeconomic base of the country. Between 1960 and 2001, the economic experienced the largest economic decline in the world (less than -3% per annum), and the vast agricultural sector - which employs over three-quarters of the population - has suffered particularly badly, especially in recent years. Agricultural exports declined from 40% of all exports in 1960 to only 10% in 2000, and the food surplus per person declined by an astonishing 30% between 1975 and 2000. Unsurprisingly, around two thirds of the country lives on less than $1 per day, 70% face food insecurity of some sort while 16 million people suffer from chronic malnutrition, yields are a minuscule fraction of their potential, and the country imports around one-quarters of its cereal consumption (Appendix A). In short, the DRC is a severely depressed economy in which the vast majority of the population survives in a subsistence agricultural economy.

Despite being the third largest country in Africa and one of the poorest, the question of how to reverse decades of economic stagnation in the DRC is one that the research community has scarcely touched upon. Whilst we know that agriculture is important, even in mineral-rich economies (e.g. Indonesia, Chile, Nigeria), achieving agricultural growth requires a range of investments *in* agriculture (R&D, extension services, irrigation projects, input distribution policies, etc), but also investments *for* agriculture. In the case of the DRC, we argue that it is actually an investment for agriculture - rural roads – that is currently the binding constraint on agricultural growth. Our reasoning is quite simple. First, a range of research has demonstrated that roads are extremely important for agricultural development (see Van de Walle, 2002), and that weak transport infrastructure is an especially severe constraint across much of Africa. Second, transport infrastructure in the DRC is particularly weak (Minten and Kyle 1999; World Bank 2006). Figure 1

[1] Eric Tollens, professor and agronomist at the Catholic University of Leuven, quoted by § CO, the magazine of the Belgian development cooperation, No. 4, p 32, 33 La Voix du Congo.

shows the percentage of the population estimated to be within 5 hours drive to a 50,000 person town. DRC has one of the lowest 'market access' scores in Africa, and we will demonstrate below that these estimates almost certainly overestimate market access on the ground. For one thing, many roads in DRC are roads in name only, and survey evidence suggests that transport times are also increased by around 40% in the wet season, which in the DRC lasts for close to six months.

Figure 1. Market access in Sub-Saharan Africa



Source: Authors calculations.

Finally, as our title suggests, rural roads are somewhat unique in terms of their capacity to literally pave the way for other investments, such as schools, health services, and security services (AITD & UNESCAP, 2000; Fan, 2008). In agriculture, better roads can drastically reduce the cost of inputs such as fertilizers, seeds, and extension services (Gregory and Bumb, 2008; Ahmed and Hossain 1990; Dercon et al, 2008). On the output side better roads increase the scope of profitable trade, which in turn encourages on-farm investments to raising agricultural production (Binswanger et al 1993; Khachatryan et al, 2005). This in turn should raise rural incomes, lower food prices (and hence raise disposable income in urban areas), reduce spatial disparity in food prices, and reduce dependence on food imports. Hence, better rural roads increase net returns to other worthy investments in both the farm and non-farm sectors.

The good news is that we are by no means alone in identifying infrastructure as a severe bottleneck on the DRC's development and on agricultural growth in particular. A recent World Bank review attributes the decapitalization of the DRC's agricultural sector to the collapse of the country's infrastructure network, and identifies infrastructure investments as one of the four critical policy goals for the sector. The DRC government and its donors have likewise identified infrastructure as a priority sector. The World Bank and British governments have signed a five year accord for the rehabilitation and upgrading of 1,800 kilometers of high priority roads. These emergency projects already made it possible to open 4,200 km of roads, and will thus make it possible to cover more than 40% of the 15,000 km priority roads in the DRC. Finally, China is now becoming a major international investor in China. Whilst the financial crisis and political tensions with traditional donors have lead to delays in the negotiations between the DRC and China, the ambitions of the partnership constituted one of the largest infrastructure investments in African history, including around 5,800 km of road rehabilitations and an equally long railway networks.

But although these investments in principle address a binding constraint on the DRC's economy, they also involve risks. First, debt-funded investments need to generate high returns in order to offset the debt burden. Second, infrastructure may be a generic solution to the DRC's problem, but the spatial allocation of infrastructure investments might significantly determine their broader socioeconomic impact. Africa as a whole has a checkered history in which infrastructure investments have primarily served extractive industries rather than agriculture. Roads and railways which link mines to ports, or even capital city to capital city, could potentially bypass major agricultural production zones and the population centers they might service.

Thus, while we have strong priors that infrastructure is important for Congolese agriculture, there remain a number of ill-informed issues which this paper tries to address. First, we try to identify the magnitudes of the various channels by which the existing infrastructure network impacts on agricultural development and broader economic welfare in the DRC. Second, we simulate the impacts of alternative infrastructure investments on these economic outcomes.

The methods by which we do so build on existing techniques, although we extend and adapt these techniques in several ways (Section 2). First, we follow the burgeoning 'GIS literature' in estimating market access based on imposing simple travel time assumptions on geo-referenced maps of the DRC road network, as in Figure 1. However, because these assumptions are derived from generic travel-time assumptions rather the DRC-specific assumptions, we adapt the estimates to DRC's circumstances using survey-based travel-time estimates from Minten and Kyle (1999). We then re-estimate the likely impact of the DRC's planned infrastructure investments on market access across the country, as well as other scenarios such as a 'transport corridor' investment strategy versus a 'feeder road' strategy. Section 2 also outlines our methodologies for estimating crop production potential, actual crop production, and population density.

With these variables, a baseline market access scenario and several alternative investment scenarios, we then turn to the question of what relationship market access has on economic welfare. To begin with we econometrically estimate the impact of market access on crop production, following Dorosh et al (2008) (Section 3). For a second dimension, we then use a recent Demographic Health Survey (DHS) for the DRC to estimate the relationship between a proxy for market access (travel-time to health services) and a wealth poverty (Section 4). Finally, these various elasticities between market access and welfare outcomes are then used to simulate the impacts of the alternative investment strategies highlighted above (Section 5). Section 6 concludes.

2. Methods

In this section we outline the methods use to construct geospatial dataset that includes crop production, a measure of market access that effectively links population distribution with transport infrastructure and terrain characteristics, and a measure of agroclimatic crop suitaiblity that account for the biophysical potential of physical areas in terms of soil and climatic suitability. We then discuss our econometric strategy for establishing the relationships between these variables.

2.1 Estimating agroclimatic crop suitability

Different crops have different thermal, moisture, and soil requirements, particularly under rainfed conditions. The Food and Agriculture Organisation (FAO) with the collaboration of the International Institute for Applied Systems Analysis (IIASA), has developed the Agro-ecological Zones (AEZ) methodology on the basis of an inventory of land resources and evaluation of biophysical limitations and potentials. The AEZ methodology provides a standardized framework for the characterization of climate, soil, and terrain conditions relevant to agricultural production. Crop modeling and environmental matching procedures are used to identify crop-specific limitations of prevailing climate, soil, and terrain resources, under different levels of inputs and management conditions. This methodology also provides maximum potential and agronomically attainable crop yields and suitable crop areas for basic land resources units (usually grid-cells in the recent digital databases) (Fischer et al 2001; FAO 2003).

In this paper we measure potential yields for each of three production systems defined in the FAO/IIASA suitability datasets: Irrigated – high input (we simply call it "irrigated"), Rainfed – high input, Rainfed – low input. Then for each of the three input levels, we define our land suitability by crop based on four classes: very suitable, suitable, moderately suitable, and marginally suitable. Finally, the potential yield is calculated as the area-weighted average of the above four suitable classes (FAO 1981; FAO 2003).[2] To summarize, the agroclimatic crop suitability of a geographical area is a function of three factors: (1) the production system; (2) the crop mix; and (3) the suitability of the land for that crop mix. An important point to note is that factors (2) and (3) are essentially directly observed from location-specific data, whereas the production system (1) is not. In the DRC we know that there is very little use of irrigation or modern inputs: FAO data for the pre-civil war period of the 1990s suggest that there was about 0.2 tractors per 1000 agricultural workers, $0.2 worth of modern fertilizers per worker per year, and that just over 0.1% of the land area was irrigated. Hence the most plausible measure of agroclimatic crop suitability is one based on the low input-rainfed technology.

[2] Some crops have many types, such as highland and lowland maize germplasm, sub-divided by maturity class. In such a case the single "maize" crop surface is a composite in which each pixel would use the best variety most suitable for the location.

2.2 Estimating the spatial distribution of crop production in the DRC

In order to evaluate food security, technology potential and the environmental impacts of production in a strategic and regional context, the International Food Policy Research Institute (IFPRI) has been developing a spatial allocation model (SPAM) for generating highly disaggregated, crop-specific production data by a triangulation of any and all relevant background and partial information. This includes national or sub-national crop production statistics, satellite data on land cover, maps of irrigated areas, biophysical crop suitability assessments, population density, secondary data on irrigation and rainfed production systems, cropping intensity, and crop prices. This information is compiled and integrated to generate "prior" estimates of the spatial distribution of individual crops. Priors are then submitted into an optimization model that uses cross-entropy principles and area and production accounting constraints to simultaneously allocate crops into the individual "pixels" of a GIS database. The result for each pixel (notionally of any size, but typically from 25 to 100 square km) is the area and production of each crop produced, split by the shares grown under irrigated, high-input rainfed, low-input rainfed conditions (each with distinct yield levels).

First tested in Latin America, the spatial allocation model was then used to generate spatial distributions of crop area and production for 20 major crops in Sub-Sahara Africa. These 20 crops are: wheat, rice, maize, barley, millet, sorghum, potato, sweet potato, cassava and yams, plantain and banana, soybean, dry beans, other pulse, sugar cane, sugar beets, coffee, cotton, other fibres, groundnuts, and other oil crops. For the DRC, we included the latest (circa 2005) district-level area and production for the following crops: cassava, bean, paddy rice, plantain, sweet potato, millet and potato.

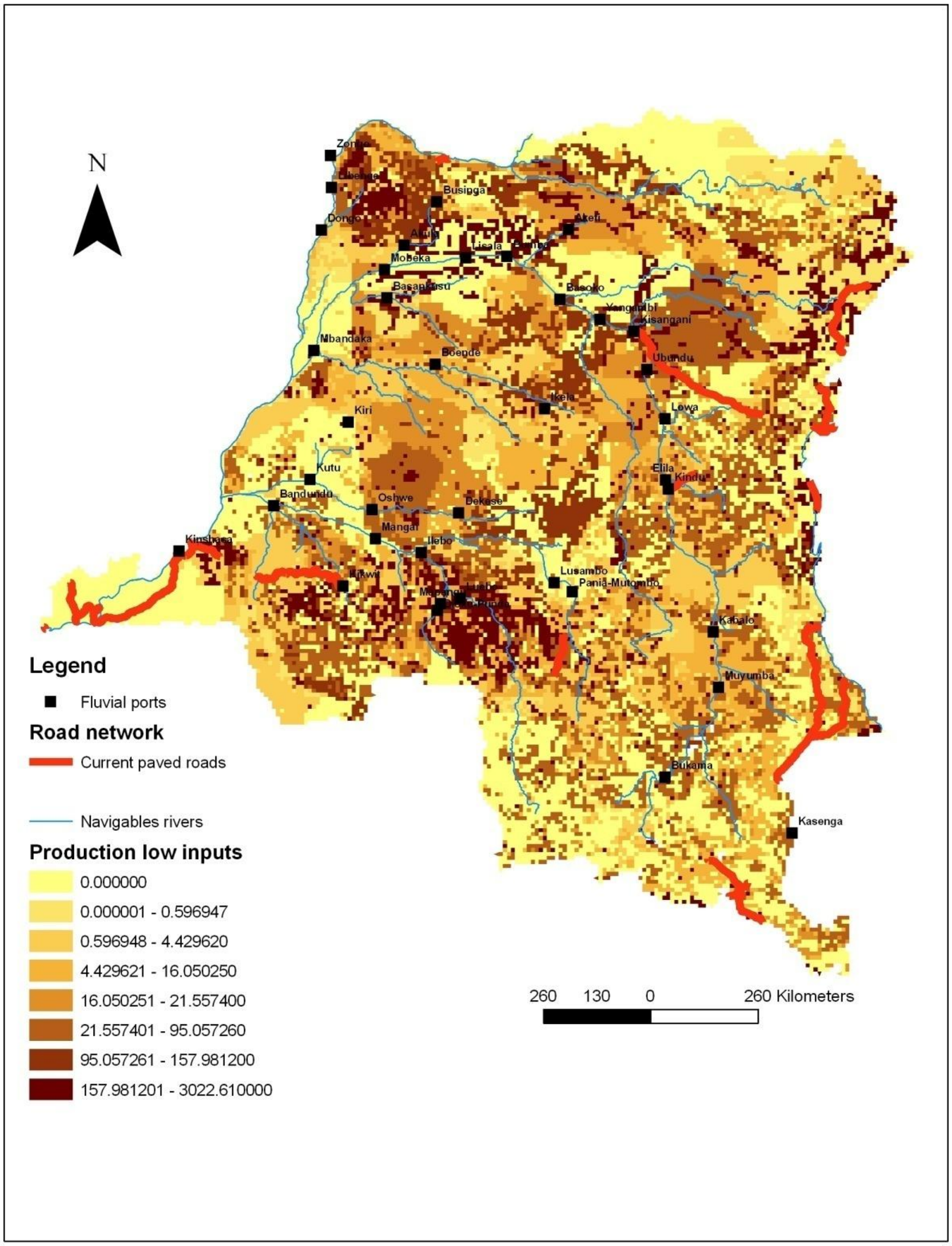
Here we only briefly and informally describe the spatial allocation methodology. A more detailed description of the technique is presented in Appendix A, while still more complete descriptions of the data sources and the detailed model can be found in You et al. (2007), and You, Wood and Wood-Sichra (2009). As noted above, the spatial crop allocation problem is defined in a cross entropy framework (You and Wood, 2006) in which all real-value parameters are first transformed into a corresponding probability form. The objective function of this spatial allocation model is the cross entropy of area shares and their prior, which are subject to aeries of adding up constraints for crop areas, land cover image, crop suitability information, aggregation constraints between subnational units, irrigation potential, and a simple adding up constraint for crop shares.

Obviously an informed prior is very important for the success of the model. We create the prior based upon available evidence on prices, yields, crop suitability, and population density. For those geopolitical units without area statistics, we simply merge them together and obtain the total area for that merged unit by subtracting the sum of available subnational areas from national total. After this pre-allocation, we calculate the prior by normalizing the allocated areas over the whole country. To convert the allocated crop areas into production, we need to consider both the broader production systems and the spatial variation within the systems. We first calculate an average potential yield within subnational unit, then estimate actual crop yields for each pixel in the different production systems. Finally, the production of crop j in production system l, and pixel i, $Prod_{ijl}$, is calculated as multiplication of crop area (A), cropping intensity, and the estimated yield:

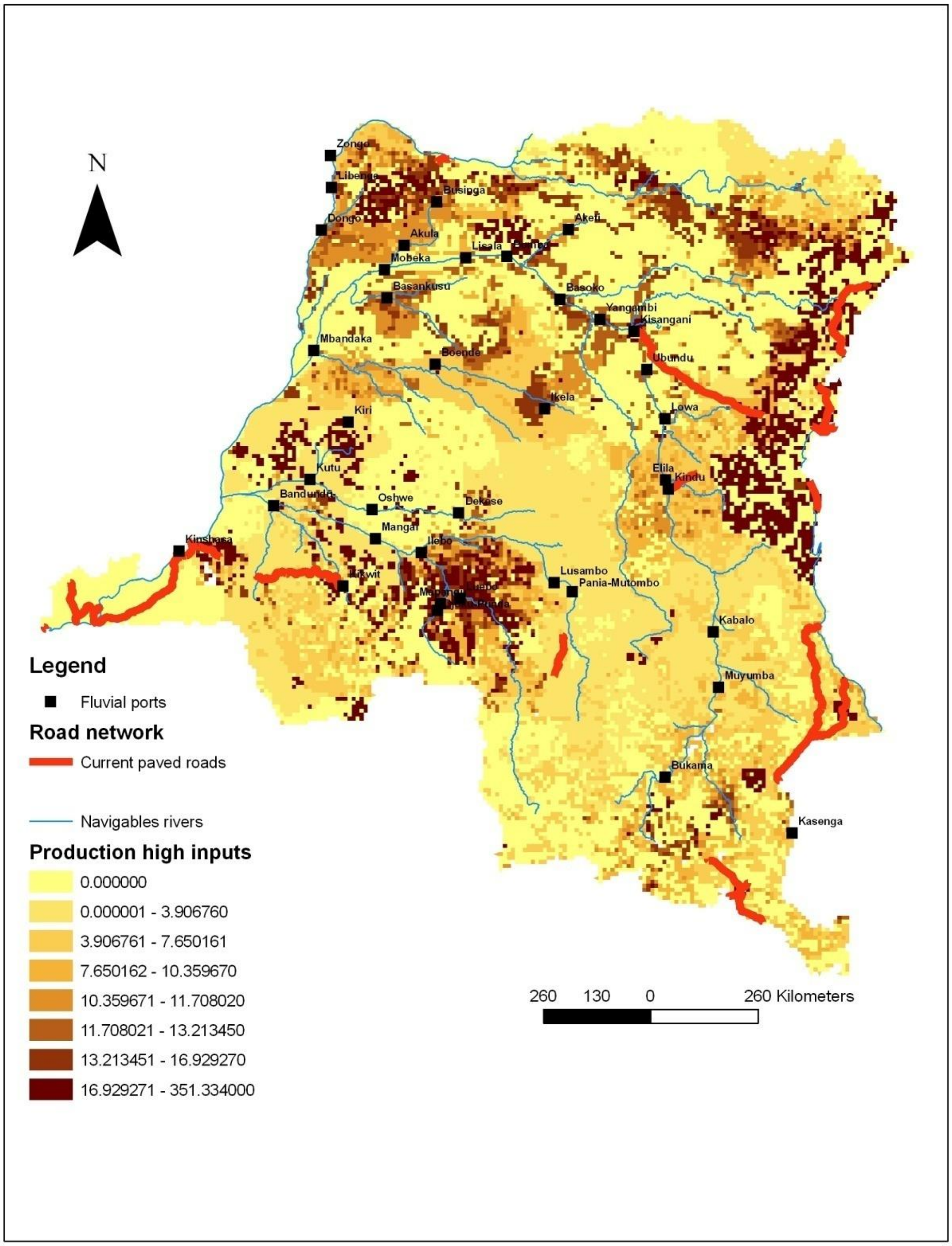
$$\mathrm{Pr}\,od_{ijl} = (A_{ijl} \times CroppingIntensity_j) \times Y_{ijl} \quad (1)$$

We run the modified spatial allocation model country by country. A post-processing program would take the results from the model and calculate both the harvest areas and productions by pixels. Figure 3.1 shows the crop distribution maps for cereal crops and roots and tubers. These are the 5x5 minutes (about 9x9 km^2 on the equator) crop distribution maps. Similar maps are also generated for other major crops, covering over 90% of total crop land in SSA. In addition to these area distribution maps, the model results include production and harvested area distribution maps as well the sub-crop type maps split by production input levels (irrigated, high-input rainfed, low-input rainfed and subsistence). Maps 1 and 2 present the spatial distribution of production potential for both low and high inputs scenarios while actual production is reported in Map 3.

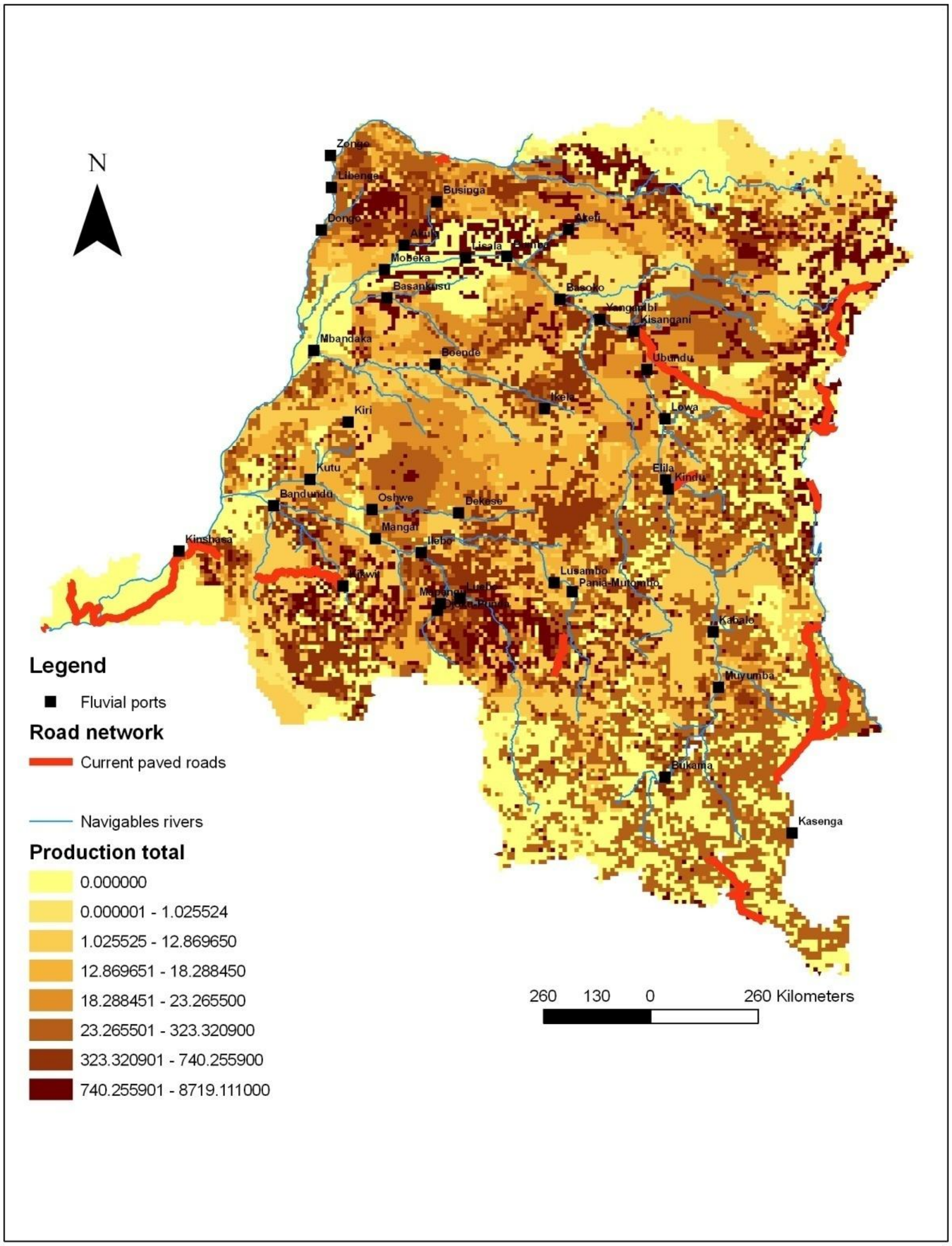
Map 1. Potential crop production in the DRC in a low input scenario (1000s of dollars)



Map 2. Potential crop production in the DRC in a high input scenario (1000s of dollars)



Map 3. Estimated total crop production in the DRC



2.3 Population data in the DRC and local market access

To identify the nearest city and its population size, we used the Global Rural-Urban Mapping Project (GRUMP) population data from the Center for International Earth Science Information Network (CIESIN).[3] These population counts for the year 2000 were adjusted to match UN totals. We then combined friction grids and the locations of cities with different sizes, and calculated travel time to nearest town/city of (i) 50,000 population or more, (ii) 100,000 population or more, and (iii) 200,000 population of more. For the details, see Thomas (2007).[4]

In addition to defining markets on the basis of town or city size as we do below – e.g. 50,000 or 100,000 person towns – we also follow Dorosh et al. (2008) in considering local market size, since this may also influence crop production. There is no consensus on defining the boundary/size of local market (or market potential measure), but a standard method is to use a distance decay model and calculate population aggregates decayed over distance.[5] Thus, local market size is calculated as:

(2) $$\text{local market size}_{\text{i}} = \sum_{k} w_{k,i}\, pop_k$$

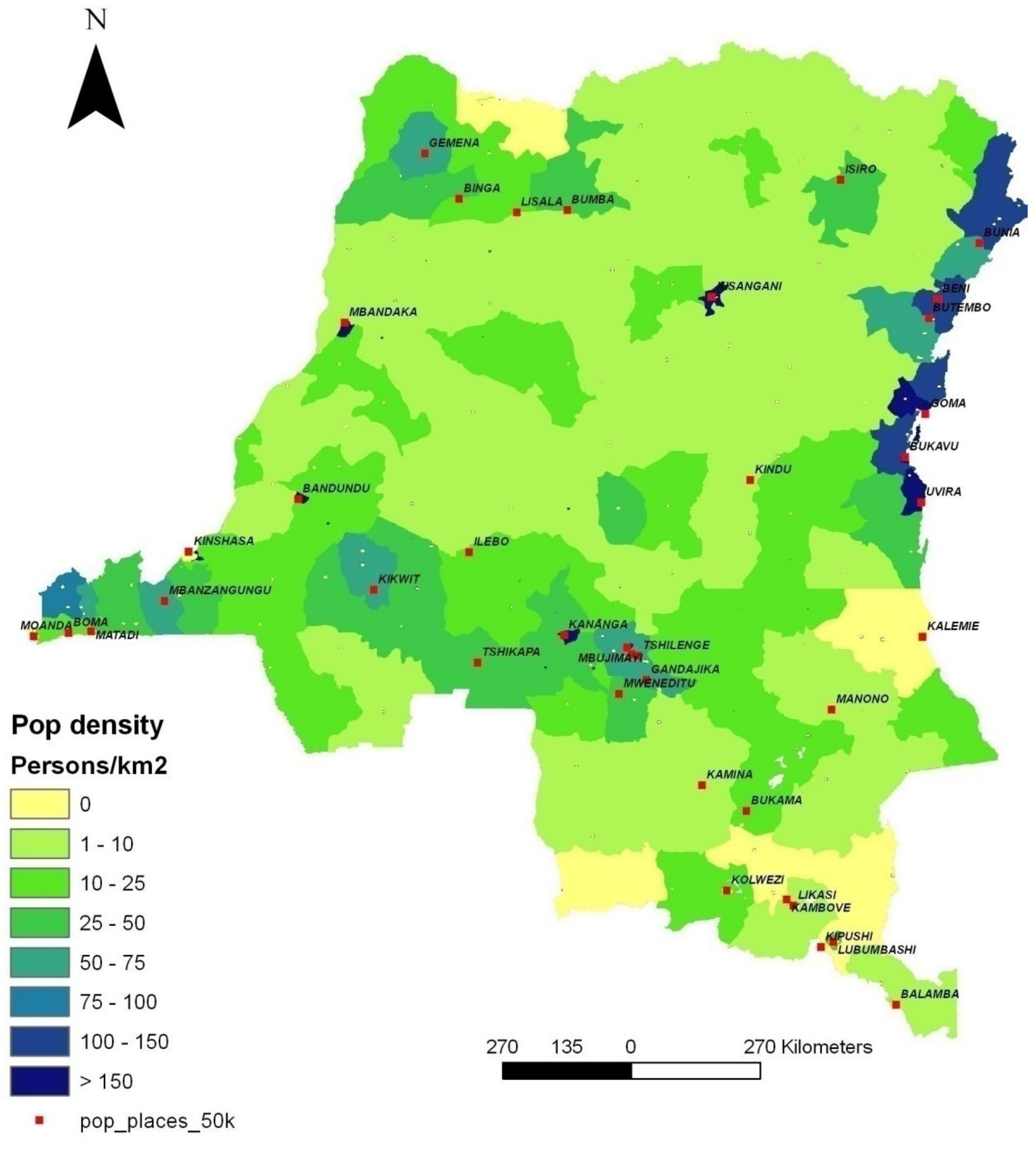
where pop_k is the population aggregate in neighboring area k and the distance weight is $\text{w}_{\text{k,j}} = 1/(d_{k,i})^{\gamma}$ where $d_{k,i}$ is the Euclidean distance between k and i in kilometers and γ is an arbitrary decay parameter. Following Dorosh et al we use two proxy variables: (i) a population count in its own pixel, and (ii) a distance-weighted population aggregate in neighboring areas within a 100km radius (excluding its own population). We divide these areas into 6 subgroups (radius between 1-2 km, 2-5km, 5-10km, 10-20km, 20-50km, and 50-100km) as listed in Annex Table 4. The input data are from the GRUMP population counts in year 2000 at 1km resolution (see Map 4).

3 Specifically, it is the Gridded Population of the World, version 3, with Urban Reallocation (GPW-UR).

4 Details of the calculations for Mozambique are given in Dorosh and Schmidt (2008).

5 See Deichmann (1997) for a review of the issues related to this methodology.

Map 4. Population density in the DRC



2.4 Estimating access to markets

Lack of access to both input and output markets has been identified as a significant constraint on agricultural development in sub-Saharan Africa and elsewhere. In our modeling exercises we computed travel times to major cities[6], airports, fluvial (river) and maritime ports. In each case accessibility was computed using the cost distance function from ESRI,[7] which is defined as the time needed to travel from a pixel to the nearest location of interest.

Modeling accessibility required the creation of a friction surface, which represents the time needed to cross each pixel. Both speeds on and off roads are affected by the friction surface which is integrated by various input layers such as the transport network, land cover, urban areas, slope, water bodies, international boundaries, and elevation. The first layer we consider is the elevation and slope since these are factors that affect both on- and off-road speeds, and hence the majority of other infrastructure layers. In effect, then, these factors are used as multiplying factors over the entire friction layer, as per Van Wagtendok and Benedict (1980):

(2) $$v = v_0 e^{-ks}$$

where v = off-road foot-based velocity over the sloping terrain; v0 = the base speed of travel over flat terrain; s = slope gradient (metres per metre); k = a factor which defines the effect of slope on travel speed.

For DRC we assume a base speed of 5km/hr with k set to 3.0 and constant for uphill and downhill travel. The velocities over the slope grid were computed and then converted into a friction factor by dividing the base speed by the slope speed. This was then used as a multiplier against the other friction components.

When calculating the multiplier for elevation, we assume that elevations lower than 2000 meters have no effect on travel speed. For elevations above 2000, the following speed factor is applied

(3) $$f = 0.15e^{0.0007E}$$

[6] Major cities include access to Kinshasa, cities with equal or more than 50,000, 100,000 and more and 200,000 and more.

[7] For more details about the cost distance algorithm refer to: http://webhelp.esri.com/arcgisdesktop/9.2/index.cfm?TopicName=How%20Cost%20functions%20work (accessed on 02/21/09)

where f = the friction factor and E = elevation in meters.[8]

Finally, we consider travel times by transport type. Normally the approach here assumes travel times by transport type that are common across countries, such that highly detailed maps of transport routes (including road surfaces) suffice to give a good approximation of travel times on the ground (e.g. Nelson 2008; Dorosh et al, 2008). Hence, up-to-date maps are certainly highly important, and we have gone to considerable effort to update our information on road categories (176,000 km), rail networks (1,300 km), and river networks (23,000 km) in the DRC, as well as additional targets such as ports, maritime ports and national and international airports (fluvial ports are particularly important for the DRC as the Congo river and its branches are an important transport route for much of the population).

However, it is not at all clear that even these updated maps give a sufficiently accurate picture of the situation on the ground. As Minten and Kyle (1999; hereafter MK) note about the DRC:

> *Most of the road network is in bad condition, with important sections almost impassable and access to some interior areas severely curtailed. Rural roads are maintained by local authorities who have neither the resources nor the organizational capacity to carry out the task.*

In other words, a road might look 'normal' from a satellite picture on transport map, but in reality be "almost impassable". To minimize this error – which could potentially bias our results – we use transport survey data collected by MK for the early 1990s to incorporate travel times into our market access estimates that more clearly reflect the realities on the ground in DRC. Among other things, the MK survey asked agricultural traders about where they imported food from and how long the journey took. For each journey MK also distinguished between travel times on paved and unpaved roads. From that data we can obviously derive travel speeds by road type. An additional and very context-specific insight from the MK study is that the DRC's lengthy and intense wet season increases travel times by as much as 40%.

But whilst we consider the incorporation of MK's survey data into our estimates a significant improvement over our generic "cross-country" estimates, we must still acknowledge that significant measurement errors undoubtedly remain, as well as the possibility that we still underestimate travel

[8] To perform different market access scenarios and to speed up data processing, we used python-geoprocessing scripting language to run run geoprocessing operations and automate processes that range from setting geoprocessing environments, reclassifying variables, extracting attributes, and performing advance spatial analyses.

times in the DRC. First, the MK survey was conducted in the early 1990s, so their data is not very up-to-date. It is possible that this is not a major problem. Because of its economic stagnation and political turmoil, the DRC has not yet witnessed any major investments in infrastructure that would significantly improve travel times. If anything, roads are probably in worse condition now that they were in the 1990s, when they were already in terrible shape. A second problem is that the MK survey only considered travel times to Kinshasa, so most of their data only yields information on travel times in the west of the country. Given that the war in the east (North and South Kivu) may have lead to especially rapid deterioration of the road network, it is possible that we underestimate travel times in these parts. Still, all in all, we consider the incorporation of the MK a significant improvement.

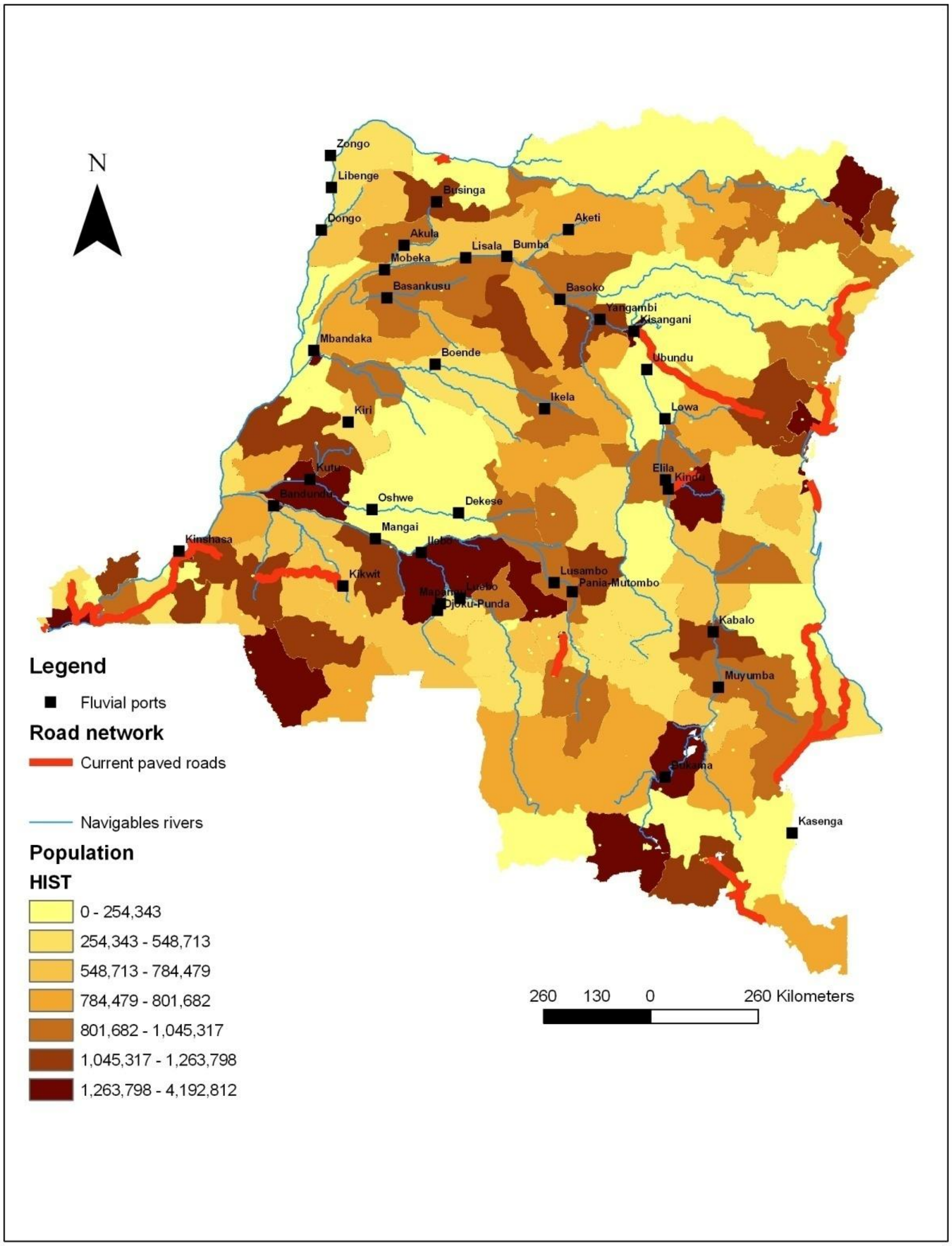
Table 1 shows assumed velocities by transport type for the dry and wet season, while Maps 5 and 6 show the transport network and the resulting estimates of market access in the DRC. Map 5 shows that the vast majority of land area in the country has very poor market access.

Table 1. Assumed travel times by transport type

Transport type	Velocity km/hr Dry season	Wet season*	Incorporates information from MK's survey?
Paved	80	46	Yes
Four wheel drive	30	17	Yes
Loose gravel	25	14	Yes
Trail	3	2	
Ferry crossing	5	3	
Rail-train	10	10	
Rivers	10	8	Yes

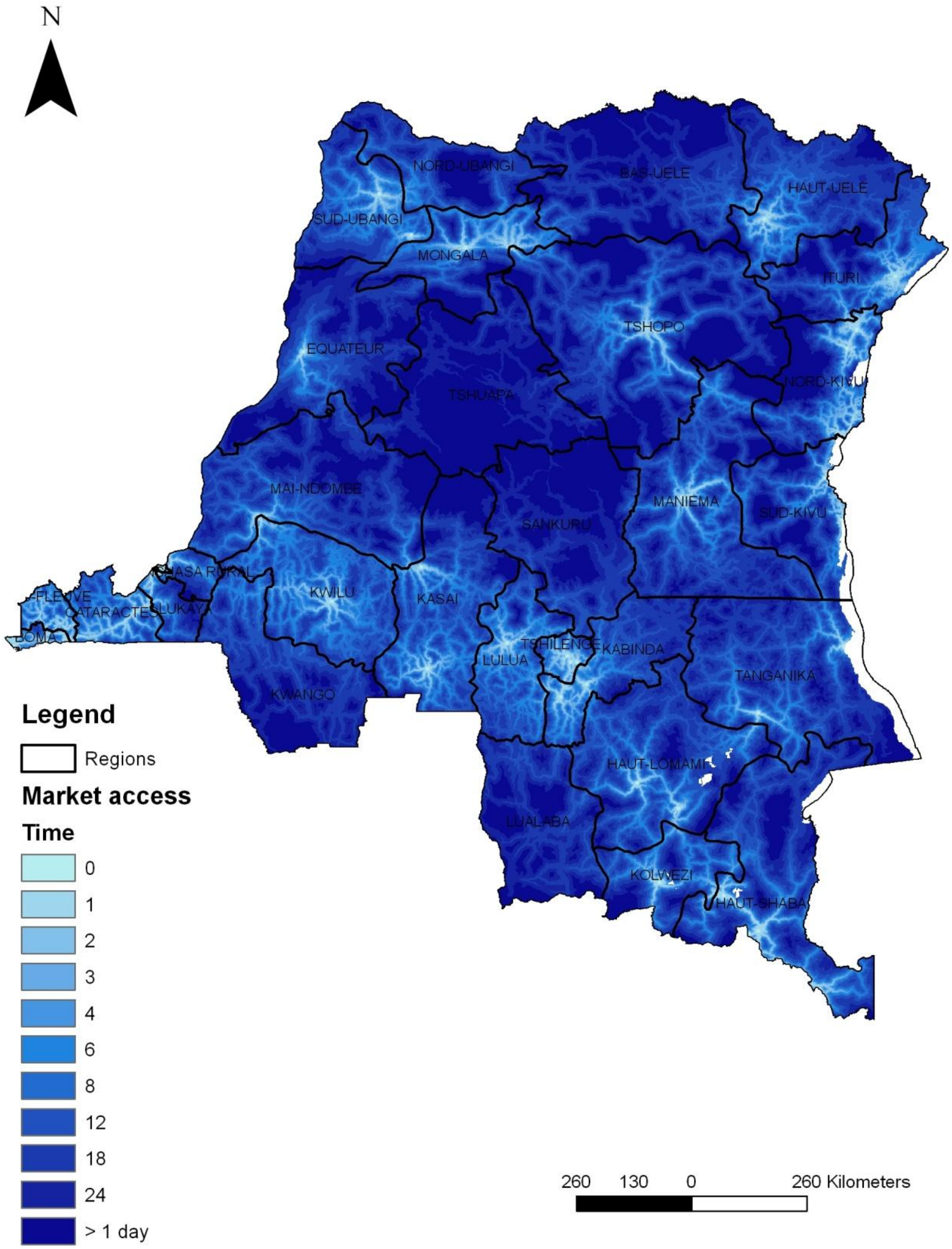
Notes: Speeds are partly based on existing assumptions (e.g. Nelson 2008), partly on anecdotal evidence for the DRC, and partly on MK's survey-based estimates of differences in travel times between paved and unpaved roads, and dry and wet seasons. *Wet season travel time differentials are also based on MK.

Map 5. Paved roads, river networks and population density in the DRC



Notes: Constructed by the authors from existing data sources. See text for details.

Map 6. Estimates of travel times to 50,000 person towns.



Notes: Constructed by the authors. See text for details.

3. Market access and agricultural production: What are the links in the DRC?

3.1 Conceptual Framework and Model

In assessing the implications of location and investments in transport costs on crop production and productivity in sub-Saharan Africa, we follow Dorosh et al (2008) in adopting a conceptual framework in which transport investments affect both the supply and demand for crop production. On the supply side, the production of crop *j* under production system *l* in location (pixel) *i* depends on the agronomic potential p_j, under the production system *l* in location *i*, and unobserved location-specific variables (Ω_i) such as output and factor prices, and available technology. Demand for a crop produced in location *i* depends on the size of the local market surrounding location *i*, which is in turn determined by population, distribution of per capita incomes and the trade regime (especially whether the domestic market is integrated with the international market).

The hypothesis to be tested is that better transport connectivity increases crop production (or productivity) after controlling for other factors. The effects of better transportation are assumed to take place through a reduction in transport costs of goods and services which raise producer prices of crops (depending on the elasticity of demand as well as supply). Reduced transport costs also lower the costs and profitability of supplying modern inputs such as fertilizers, seeds, extension services and other technologies (Ahmed and Hossain, 1990). However, because the DRC agricultural economy currently uses scarcely any of these modern inputs, we suspect that any positive association between market access and agricultural production primarily reflects the impacts of access to output markets for agricultural produce, rather than input markets. Were government policies to simultaneously invest in infrastructure and the adoption of modern inputs, it is probable that the impacts of infrastructure investments would be higher in the long run, although there are many factors in addition to transport costs that explain why African farmers do not adopt modern technologies. Indeed, the evidence from the Dorosh et al (2009) study is that the elasticity between market access and adoption of high input technologies (admittedly a crude proxy) is fairly low, between 0.02 and 0.09 (Table 8 in Dorosh et al, 2009).

Another impact of great market access on agricultural production is through effecting the composition of agricultural production. As lower transport costs result in a greater percentage reduction in the price of perishable and bulky items such as vegetables, the profitability of these items increases relative to non-perishable crops (the von Thunen hypothesis). Indeed, Minten and Kyle (1999) found that this von Thunen effect was very important in the DRC:

> *The more perishable and the higher value the products fruits, vegetables, cassava roots, cassava chikwangue, cassava leaves, tomatoes, pimento, the less distance they are transported. The basic less perishable staples (cassava chips, peanuts, maize) come from further away. The average distance they are transported is 337, 373, and 323 km, respectively. Compared to the vegetables 107 km., they come from three times as far. The von Thunen effect is also illustrated by the smaller standard deviation in distance traveled for the individual products compared to the standard deviation of the average. Only cassava chips and maize are characterized by a higher standard deviation indicating their omnipresence as a cash crop.*

Finally, where the transport cost reduction is large enough and widespread enough, there are potential general equilibrium effects on both the rural and urban non-farm sectors, wages and overall incomes, as well as 'non-economic' factors such as political stability and law enforcement. For example, increased agricultural trade boost demand for transport service services in the urban and rural non-farm economy. Transport times can also reduce the costs of migration (temporary or permanent). Finally, lower transport times reduce the costs of other investments and services, such as schooling, health, extension services and so on. While we cannot tease out which of these channels are most important (CGE modeling is better suited to that objective) we will approximately estimate the relationship between market and broader welfare measures from the 2007 Demographic Health Survey (DHS).

Turning first to the impacts of market access on crop production, we closely follow the basic model used by Dorosh et al (2008), which is a reduced-form crop production function:

(4) $\text{Crop production}_{ijl} = f(\text{agronomic potential}_{ijl}, \text{local market size}_{i}, \text{market access}_{i}, \Omega)$

Whilst the measurement of these variables is discussed above, the theoretical rational for the model is that these variables capture both supply-side factors such as agronomic potential and access to input markets (although these are not yet important in the DRC), as well as demand-side factors relating to access to local markets as well as major towns or cities. With regard to the latter we consider a 50,000 person town a sufficiently sizeable market, although we experiment with urban agglomerations of other sizes as well.

As for the econometric issues that arrive with such a model, there are several. First, it is necessary to correct for the bias in the regression estimates arising because the dependent variable (crop production/productivity) is left-censored data (i.e. by definition, their values are never less than a certain value, in this case, zero). To overcome this potential bias, we estimate the equations using a Tobit (censored regression) model and drop areas (pixels) that are unsuitable for agricultural production from our regression.

Secondly, there are potentially endogeneity or parameter heterogeneity issues. For example, omitted factors may determine both market access and agricultural production. Dorosh et al (2008) use the example of a road that may have been initially constructed primarily to connect a mining area to a port. Since that example is particularly relevant to the DRC it behooves us to reconsider the possibility that mining towns induce a bias. Specifically, Dorosh et al (2008) are concerned that the mining production could simultaneously increase market access and stimulate agricultural demand, thus driving up the elasticity between the two. In our view, if the mining population stimulates demand in the normal channels this is not a problem because travel times to markets are supposed to capture these demand effects. But if if mining towns represent unusual circumstances – e.g. unusually high incomes – this could at least create a parameter heterogeneity problem: mining towns stimulate higher local agricultural production than non-mining towns/cities.

Whilst this heterogeneity issue is interesting, we suspect it is not a major problem for several reasons. First, mining is not highly labor-intensive, so its impact on food demand is not especially large. Second, we do not find much evidence that mining provinces are significantly wealthier than non-mining provinces (see below). Third, we control for provincial fixed effects.

An equally important endogeneity issue relates to how well we observe agricultural potential. Roads are not randomly distributed. Instead road networks are normally designed so as to cater to larger populations. Since population density is in turn a function of agricultural output or potential, it is possible that omitting agricultural potential would lead to an overestimating of the impact of market access on agricultural production. While we do include a measure of agricultural potential, the same problem could also result if our measure is insufficiently accurate.

Finally, many of the arguments above point to complex interactions between the explanatory variables. To consider such interactions we specify more general non-linear models with quadratic and interaction terms.

3.2 Descriptive statistics

Table 2 presents descriptive statistics aimed at demonstrating some basic results for the key variables of interest. Pixel sizes are roughly 1 square kilometer, so the total sample for the regressions is very large – roughly 25,000 – however we only use about 15,000 pixels in the regressions many pixels do have crop production values. The value of crop production varies in value between zero and US$ 184.2.

Initially we were interested in testing a range of market access variables, but mutlicollinearity proved to be serious problem (see correlation in Table 3). However, the two most important market access variables for agricultural production in the DRC are access to cities and access to fluvial ports. MK (1999), for example, find that about two thirds of agricultural trade from the hinterland to Kinshasa is by road, and the other third by river. Moreover, as we saw in the previous section, nearly all of the DRC's 50K-plus cities are located on navigable rivers. The good news is that the correlation between travel time to a 50K-plus town and travel time to a fluvial port are not so highly correlated that mutlicollinearity becomes overly serious (r=0.61). The only other variable that is highly correlated with travel time to a city (50K or 100K) is the population of the pixel (r=-0.46), indicating that population density decreases with isolation from cities, as expected.

Table 2. Descriptive statistics

Variable	Obs	Mean	Std.	Max	Min
Crop production ($1000s)	24,955	18,129	78.7	184.2	0.0
Potential crop production ($1000s)	24,955	18,129	15600.0	15,900.0	0.0
Travel time to 50K town (minutes)	24,955	18,129	998.9	624.7	0.0
Travel time to 100K town (minutes)	24,955	18,129	1,084.7	643.3	0.0
Travel time to fluvial port (minutes)	24,955	18,129	1,049.3	641.73	5
Population	24,955	18,129	2,203.7	12341.4	0

Table 3. Correlation matrix of explanatory variables

	Travel time - 50K	Travel time - 100K	Travel time - port	Potential production	Population
Travel time - 50K	1.00				
Travel time - 100K	0.93	1.00			
Travel time - port	0.61	0.54	1.00		
Potential production	-0.01	0.01	0.11	1.00	
Population	-0.46	-0.46	-0.25	-0.07	1.00

Table 4 looks at these relationships in more detail by breaking up travel time to a 50K city by deciles (column 1) while column 2 shows the average travel times in the dry season for each decile. What is most astonishing is the absolute size of travel times. Even the second and third deciles involve travel times of well over 5 hours (a common benchmark for proximity), while the lower five deciles involve travel times from half a day to an extreme 1.5 days to reach a 50,000 person town. Column 4 also shows that these are not small populations living in isolation. The bottom five quintiles contain about 25% of the total population, and involve travel times of half a day or more to 50k person towns. As for agricultural production (Columns 5 and 6) most of this takes place in the less isolated regions. About 62% of production value takes place in the first 4 travel time quintiles. Finally, column 7 shows production as a percentage of potential production (based on the crop suitability measure described above). This ratio is very low (5% or less) for all degrees of isolation, but also declines almost monotonically with isolation, suggesting lack of market access is a significant constraint on the fuller utilization of the DRC's agricultural potential. Based on these basic descriptive statistics, we do expect a reasonably strong correlation between market access and agricultural production.

Table 4. Travel Time, Population and Crop Production in the DRC

1. Travel time decile	2. Travel time (DRC average)	3. Travel time (African average)	4. Percentage of population	5. Production ($1000)	6. Production (% total)	7. Production (% potential)
1	5.1	3.3	41.4	444.6	19.5	5.4
2	7.5	6.7	14.0	401.9	17.7	1.7
3	9.6	9.1	9.9	322.5	14.2	2.7
4	11.7	11.4	7.1	249.4	11.0	0.3
5	13.8	13.7	6.5	179.0	7.9	1.7
6	16.1	16.2	5.6	186.8	8.2	0.4
7	18.5	18.9	4.7	155.7	6.8	0.2
8	21.4	22.3	4.3	130.8	5.7	0.3
9	25.1	27.0	3.4	118.3	5.2	0.1
10	31.5	39.5	2.9	87.7	3.9	0.1

In Table 5, we present estimation results of the supply/demand crop production outlined in equation (4) in Section 2. For all regressions we use the Tobit regressor to address the censoring of values, although because both dependent and independent variables are in logs, the censoring is not especially important. Also, all regressions include territorial fixed effects. These territories are the smallest subnational units and number about 150 (some drop out because of limited observations). We also experimented with district fixed effects (of which there are about 30) and provincial fixed effects (about 15). These made no substantial differences to the results, although we tended to find that parameter heterogeneity was more of an issue with these more aggregated fixed effects. In other words, interactions between travel time and crop potential and travel time and population became significant when we stopped using the more disaggregated territorial effects. Those results are not reported here but are available upon request.

Turning to the results in Table 5, we first specified a simple log-linear that is quite similar to the models specified by Dorosh et al. (2008), with the only difference being that only include pixel-population rather than neighboring populations. This was because mutlicollinearity between the pixel population and the local (squared 100 km) population was very high in our sample, so much so that it precluded us specifying both variables (however, in other results reported below we address this issue through other means). Regression 5.1 indicates that the elasticity between travel time to a 50K-plus city and agricultural production is highly significant and equal to about -0.44, indicating that a 1% reduction in travel time would increase agricultural production by almost 0.5%. This is reasonably large, although the elasticity is still much lower than the analogous elasticities reported by Dorosh et al. for all of sub-Saharan Africa and sub-regions. The elasticity for agricultural potential is also quite low (0.18) although this is not unsurprising in a country where agricultural production is highly depressed.

However, in regression 5.2 we depart from Dorosh et al. by specifying a quadratic term for pixel-population was highly significant, indicating that the effect of population size on production was generally negative, but that the impact declined as population size increased. It is somewhat difficult to know what impact this is picking up though. Population size could reflect local market access, but it also picks up the size of the labor force (which ought to make the coefficient positive), or available land per capita (average farm size). It could also be that the coefficient is negative because highly dense areas are largely nonagricultural. For these reasons we do not focus much attention on the population term, although in further regressions reported below we experiment with regressions against population per capita.

In regression 5.3 we add a new variable – travel time to fluvial ports. As we saw in Map 1, Section 2, fluvial ports are extremely important in the DRC because the population has historically agglomerated on these rivers for the benefits that accrue in terms of trade, transport, and to a less extent, water supply. It turns out that add fluvial ports to the study was very important. In fact, adding this target significantly reduces the elasticity on market access to 50K-plus towns, from around -0.43 in regressions 5.1 and 5.2, to just -0.16 in regression 5.3. In contrast, the elasticity between travel time to fluvial ports and crop production is around 0.37. In regression 5.4 we drop travel time to 50K-plus cities to see whether fluvial ports access might simply be picking up the effect of 50K-plus towns. However, the coefficient on fluvial ports is substantially larger in regression 5.4 than the coefficients in regressions 5.1 and 5.2, so it appears that there is a genuinely large effect on production of access to fluvial ports. This is not surprising insofar as connecting a farmer to one river port obviously connects him/her to other river ports. Moreover, every river port in DRC directly connects to the country's largest city (Kinshasa) and the country's only international (maritime) port (Boma). In terms of sheer physical access to population centers, then, the river network has great potential. The question of the river network's trade potential is taken up in our concluding section.

Finally regressions 5.5 to 5.7 replicate regressions 5.1 to 5.3 with travel time to 100K-plus towns replacing travel time to 50K-plus towns. However, the results are materially the same, with just some slight reduction in elasticities.

Table 5. Estimating the impacts of road connectivity on crop production (log)

Regression No.	5.1	5.2	5.3	5.4	5.5	5.6	5.7
Estimation method	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit
Ln(travel time to 50K city)	-0.44***	-0.43***	-0.16***				
Ln(travel time fluvial port)			-0.37***	-0.51***			-0.43***
Ln(travel time to 100K city)					-0.43***	-0.41***	-0.10**
Ln(potential production, low inputs)	0.18***	0.18**	0.18***	0.18***	0.18***	0.18***	0.18***
Ln(population)	-0.05**	-0.43***	-0.45***	-0.50***	-0.05*	-0.43***	-0.46***
Ln(population), squared		0.027***	0.027***	0.033***		0.027***	0.030***
Total observations	15,122	15122	15122	15125	15125	15125	15125
Pseduo R-squared	0.084	0.084	0.085	0.085	0.083	0.083	0.085
Territorial fixed effects	Yes	Yes	yes	yes	yes	yes	yes

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6. Comparing across alternative fixed effects and samples

Source of elasticities >>	DRC	DRC	DRC	Dorosh – All SSA	Dorosh – West Africa
Fixed effects	Territories	Districts	Provinces	Countries	Countries
No. of fixed effects	150	30	15	42	5
Total observations	15,525	15,525	15,525	125,982	15,500
Ratio (%) of fixed effects to observations	0.97	0.19	0.10	0.03	0.03
Elasticities					
Travel time to 100K-plus city	-0.43***	-0.46***	-0.41***	-2.864***	-1.102***
Crop potential	0.18***	0.22***	0.20***	0.247***	0.406***

Finally, Table 6 compares our results to those of Dorosh et al (2008), in which authors run similar agricultural production regressions for sub-Saharan Africa as a whole (SSA) as well as West Africa, the region which is most similar to the DRC in terms of agro-climatic factors and crop mix. The first three columns also report results from alternative aggregations of fixed effects. In the DRC sample we do not find that using alternative aggregations of fixed effects makes any substantive difference to the results. However, the main finding in Table 6 is that the elasticities for market access and crop potential are much smaller in our DRC sample that they were in the full African or West African samples used by Dorosh et al. One concern is that the Dorosh et al study uses very limited country effects, which could conceivably lead to some upward bias in their results (for example, if fixed effects simultaneously account for both greater market access and greater production levels), but we have no way of confirming this, and we should also note that the disparities could well be real. Indeed, one problem we face in this study is that every element of the DRC economy –the infrastructure and agriculture sectors in particular – is so depressed that the elasticities in Tables 5 and 6 do not reveal the true potential of agriculture in the DRC (see our concluding section for more discussion of this issue.

Finally, in addition to the robustness tests involving fixed effects, we also engaged in one other potentially important robustness test. Instead of specifying total crop production as the dependent variable we specified total production per capita. Although the pixel-population are no doubt measured with considerable error (there has not been a census in the DRC since the 1980s), production per capita is a variable that ought to have a closer connect to rural welfare (i.e. incomes, food security) than total production, which is more important from a trade perspective. The results are reported in Appendix C. The per capita production regressions Appendix C also include a new explanatory variable - local population density – which is as a proxy for local market access in the Dorosh et al (2008) study. However, as in the Dorosh et al results for low-input African agriculture, we find that the elasticity of this variable is negative. We suspect that this is because higher local population densities may be capturing smaller farm sizes and the greater prevalence of nonfarm activities. Again we can only attached very limited importance to these results. The more important finding from the robustness tests in Appendix C is that the per capita elasticities for market access and crop potential are very similar to those reported in tables 5 and 6.

4. The relationship between isolation and poverty in the DRC

In this section we try to establish what the relationship is between access to markets and general poverty reduction. Travel time, or isolation, has been established as significant determinant of poverty reduction in a variety of studies, although estimates of the size of the impact do vary substantially. Kwon (2000) finds that a 1% increase in road investment is associated with 0.3% decrease in poverty incidence through direct impact on wage and employment in Indonesia (Kwon 2000). Jalan and Ravallion find that for every 1% increase in kilometers of roads per capita, household consumption rises by 0.08% in poor regions in China (Jalan and Ravallion 2002). Glewwe et al. (2000) conclude that rural communes in Vietnam with paved roads have a 67% higher probability of escaping poverty than those without. And several studies in the volume by Fan (2008) find that rural roads have a very high impact on poverty reduction in places as diverse as China, India and Uganda. Given the poor state of infrastructure in the DRC, we have a strong presumption that travel time is also an important determinant of Congolese poverty, although we also need to bear in mind that other weaknesses in the economy could reduce that advantages of proximities to towns and markets (e.g. poor public service delivery).

Ideally, we would also like to establish the impact that agricultural has on poverty reduction in the DRC, and the interactions between market access, agriculture and poverty. However, neither of the two substantial household surveys available to us – the Demographic Health Survey (DHS) and the Living Standards Measurement Survey (LSMS) – had agricultural components to them, so linking up agricultural production as a transmission mechanism for infrastucture's effect on poverty in the DRC is not yet possible. Nevertheless, the DHS is useful in that it has what we believe to be a good proxy for travel time to sizeable towns/cities, "travel time to the nearest health facility". Moreover, although the DHS is not principally an economic survey, it does contain an asset-based poverty index that has been tested, validated and strongly advocated by several leading development economists (Filmer and Pritchett, 2001; Sahn and Stifel, 2003). This index is constructed via principal components analysis of all the available asset variables in the DHS survey, all of which are listed in Table 7.

Table 7. Asset variables used in the construction of the DRC's DHS wealth index

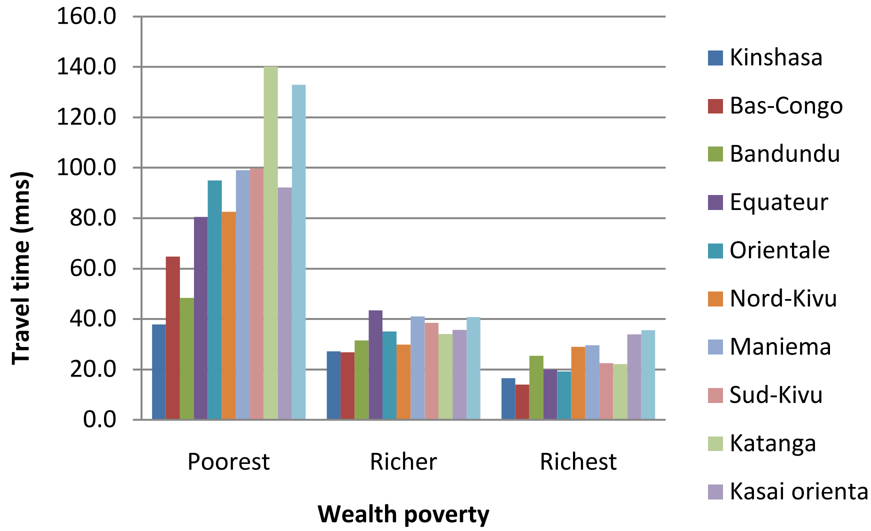
Source of drinking water	Share toilet with other households
Type of toilet facility	Type of cooking fuel
Has electricity	Have bednet for sleeping
Has radio	Has a mobile telephone
Has television	Has grill, heater
Has refrigerator	Has chair(s)
Has bicycle	Has bed(s)
Has motorcycle/scooter	Has lamp(s)
Has car/truck	Has stove, cooker
Main floor material	Has hoe(s)
Main roof material	Has sewing machine
Has telephone	Has canoe, dugout

Although we are confident that each of these measures provides a sufficiently accurate representation of their latent variables – isolation and poverty – there were some technical issues that required careful consideration. First, the asset-based poverty measure may be biased insofar as it could underestimate poverty in urban areas simply because some basic assets are easier to obtain in urban areas. For example, 43% of households in Kinshasa – which is an exceptionally poor city by international standards – own a mobile phone, and Kinshasa is the only province in the DRC with substantial electricity supply. Since Kinshasa in particular was a major concern in this regard and in several other regards, we chose to run our wealth regressions separately for each province. This appears to be a sensible choice as Kinshasa is something of an outlier in terms of the degree to which travel time is associated with wealth.

A second issue relates to market access proxy. Health facilities in the DRC are almost solely available in major towns, so it is quite likely that "travel time to nearest health facility" is a good proxy for travel time to nearest major town. Still we must acknowledge potential biases and general measurement error. In terms of biases, it is possible that health facilities are not only urban biased (which is what we assume anyway) but also biased to capital cities or mining towns, and so on. Arguably a more important bias is that access to a health facility influences poverty not through infrastructure or market access per se but through the health facilities themselves. Likewise access to a city may improve access to education, which in turn effects poverty. In order to more closely capture the effects of access to markets, we therefore run regressions which control for education and health outcomes, as well as other household characteristics such as age and marital marital status. When education and health are included in the regressions we call this the "market proximity effect", and when education and health are excluded we call this the "total proximity effect".

As for the results, Appendix D reports the full regression results, while the results in Figure 2 and Table 8 concentrate on isolation per se. Figure 2 shows relationship between travel time to health facilities and asset-based poverty within each province. Specifically we look at the poorest (5^{th}) quintile, the second quintile ("richer"), and the first quintile ("richest"). Figure 2 demonstrates that with the exception of Kinshasa and neighboring Bandundu province, the difference in travel times between the poorest and richest Congolese is substantial. In virtually all provinces, the poorest quintile have to travel at least twice as long to reach a health facility as the richest. On this basis "travel time" looks like a potentially powerful determinant of wealth.

Figure 2. The relationship between travel time to health facilities and asset-based poverty



In Table 8 we report results from more rigorous tests of this hypothesis for two sets of results: the "market proximity effect" in which we try to net out education and health impacts of location, and the "total proximity effect", which includes health and education effects. Beginning with the former, we find that even after controlling for education and health influences, travel time to "markets" has a large negative association with wealth. For all provinces the elasticity between wealth and travel time is significant at the 5% level or higher. However, the elasticities reported in column2 vary between roughly -0.06 in Kinshasa, Bandundu and Maniema, to almost -0.30 in North and South Kivu, and -0.44 in Katanga. This is potentially an important finding because it indicates that the impact of transport infrastructure on wealth could vary substantially by location. The high potential but rather conflict-torn Kivu provinces, for example, suggest high returns to improving market access, while access to cities is even more important in the mining hub of Katanga.

Column 3 in Table 8 reports the estimated impact on wealth – in terms of standard deviations in the normalized wealth index – of reducing travel time to a market by 2 hours, based on separate linear regressions used to calculate marginal impacts. This seems a reasonable experiment because as we saw from Table 4, average travel times in the DRC are very large, so for many poor Congolese reducing a lengthy travel time (e.g. 10 hours) by 2 hours is still a relatively small adjustment.[9] Because Kinshasa is already so highly urbanized we ignored the unusually large impacts in this province. We find that in a few provinces the wealth impact of lower transport times is quite low (Bandundu, Equateur, Maniema, the Kasai provinces, and Orientale) while it is again much larger in Bas-Congo, Katanga and the Kivu provinces. In this last group reducing transport times by around 2 hours would lead to wealth increases of half a standard deviation in the normalized wealth index.

Finally, looking at the total proximity effect, in all cases except Bandundu we find that the total effect is indeed larger than the "markets" effect. In one instance – Equateur – we find that the elasticity doubles when education and health channels are excluded, but excluding that outlier the difference is only 10%. Put another way, most of the effect of proximity on wealth is not through access to urban education or health services, but through other channels, which we rather loosely label here as "markets".

[9] The standard deviation of the travel time to a health facility is about 1.4 hours, and the maximum travel time is 15 hours.

Table 8. The estimated impact of travel time on wealth across provinces

	Market proximity effect [b]		Total proximity effect [a]			
Region	Elasticity	Impact of 2 hr travel time reduction (std. deviations)	Elasticity	Impact of 2 hr travel time reduction (std. deviations)	Obs.	R^2
Bandundu	-0.06**	0.12	-0.05	0.21***	332	0.29
Bas-Congo	-0.21***	0.62	-0.23	0.65***	248	0.25
Equateur	-0.05**	0.06	-0.12	0.14***	310	0.26
Kasaï Occident	-0.12***	0.07	-0.13	0.09***	280	0.27
Kasaï Oriental	-0.15***	0.20	-0.16	0.23***	289	0.27
Katanga	-0.44***	0.38	-0.46	0.57***	311	0.43
Kinshasa	-0.07**	1.15	-0.09	0.94***	343	0.30
Maniema	-0.07**	0.09	-0.08	0.11***	293	0.23
Nord-Kivu	-0.29***	0.49	-0.33	0.58***	289	0.49
Orientale	-0.14***	0.18	-0.17	0.24***	272	0.29
Sud-Kivu	-0.25***	0.37	-0.33	0.47***	280	0.39
Average [c]	*-0.17*	*0.34*	*-0.19*	*0.38*		

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%. Elasticities and marginal impacts are computed from separate regressions, although the significance levels are approximately the same.
a. The total effect is the elasticity based on regressions which exclude education and health controls. b. The access to market effect is based on regressions which net out the education and health impacts of roads. c. the average excludes Kinshasa.

5. Conclusions

The question of how best to translate the DRC's enormous agricultural potential into an engine of economic growth and poverty reduction is a vitally important question, but one still significantly under-researched. This paper has argued that infrastructure is probably the most binding constraint on what is a highly dispersed and predominantly agrarian economy, but our principal goal was not to address the very general question of whether infrastructure is important for the DRC, but how important infrastructure is for agricultural production and poverty. As it turns out, our results also provide preliminary evidence that on the question of what type of infrastructure is important for agricultural production and trade in the DRC.

Given our strong priors about the importance of agriculture in the DRC, we unsurprisingly find highly significant and negative elasticities between travel times to sizeable cities (50 or 100K), although these elasticities are small relative to those of similar cross-country tests (Dorosh et al, 2008). Moreover, city access by itself is less important than access to cities *and* ports. Since this is potentially a very important finding insofar as it provides a partial answer to the 'What kind of infrastructure?' question posed above, it behooves us to consider the theoretical merits of this finding in more detail. There are several significant reasons why access to fluvial ports is so important in the DRC. First, an individual port significantly broadens the scope of the market beyond the port city itself, by connecting a farm to other cities through the DRC's vast river network. Indeed, the DRC river network – which extends about 23,000 km – is around twelve times longer than the DRC's paved road network (less than 2,000 km). Second, because of historical patterns of population settlement and the traditional advantages that river trade has in comparison to a very weak road network, well over half of the DRC's forty-odd 50K-plus cities lie on one of a navigable river. Third, practically all of the rivers in question not only flow into the largest market (Kinshasa) but also into the DRC's only international maritime port.

So given this apparent network scale advantage it is perhaps not surprising that access to fluvial ports has an even larger statistical association with agricultural production than city access alone. Nevertheless, the trade potential of the river network is limited by several factors. First, it is obviously not possible for a port on one river to access ports on unconnected rivers, so trade patterns obviously follow the natural course of the river, whereas road networks are far less constrained by such natural barriers. Second, river transport is very slow. According to MK's findings for average Kinshasa traders, "A complete cycle by road takes 4 days to travel, 3 days to gather and buy the products, and 2 days to sell them while a cycle on the river lasts much longer: 20 days on the river, 10 days for gathering, and 3 days for selling." Moreover, MK's survey revealed that although river transport appears to be about one third cheaper than road transport, losses for river transport are quite high, presumably because of the long duration of river journeys.

Hence river travel is only suitable for relatively non-perishable goods. However, one way of increasing the potential for agricultural trade on the DRC river networks is to promote agro-processing in river ports. This would not reduce the perishability of agricultural produce, it would also facilitate employment growth and nascent industrialization. Informal communications with DRC policymakers also suggest that there is considerable scope to reduce river travel times, perhaps by as much as 50%. It is beyond the scope of the paper to offer more rigorous evidence on how much weight the DRC government should put on river transport rehabilitation versus road rehabilitation, but several further points are worth mentioning. First, roads and rivers are symbiotic. Even our own results relate to road-based travel times to ports, so we are implicitly exploring this synergy. This should remind us that it is improving the efficiency of the infrastructure network as a whole that is important - improving the road and river networks by themselves and linking them up in better ways are vital means to achieving that goal.

A second finding in this paper is that the road and rail investment proposed by various donors will have quite a limited impact on market access for the agricultural sector. This is not entirely surprising. For one thing, the dispersion of the rural populatin means that feeder roads also have to be improved (see Dorosh and Schmidt's (2008) study on Mozambique). Second, most of the proposed road rehabilitations provide transport infrastructure for mining towns, and although this may facilitate some local agricultural trade, the broader impact on the DRC's vast agrarian economy is minimal. As we noted above, one of the reasons why road transport is currently relatively unattractive compared to river transport is that the road network as a whole is very weak, and to a great extent the road chain is only as strong as its weakest link. A second problem is that many of the areas with the highest agricultural potential, such as North and South Kivu are ignored by the proposed investments, even though these regions are a potential breadbasket. If adequate political stability can be achieved in these eastern provinces, road infrastructure there could open up considerable new opportunities for agricultural trade, especially with the relatively proximate mining regions of the south-east, which current import considerable quantities of food from Zambia.

Finally, although infrastructure is clearly important in the DRC, we have good reason to believe that we have probably underestimated the potential impacts that improved infrastructure could have on agricultural and rural development in the DRC. Unlike many other African countries, the DRC uses virtually no modern inputs, such as fertilizers or seeds. For this reason we believe that the estimated elasticities between production and market access only capture demand-side effects. However, if government policies can increase extension services and promote the adoption of modern inputs, then more vibrant input markets will also increase the returns to market access. Another problem is that even these demand-side effects will be unusually weak in the DRC at the present time because incomes are so low. The good news is that at such low incomes sustained income growth will largely be spent on food, thus stimulating demand and opening up trade opportunities. Another issue future research could explore is the impact that a lack of agro-processing has on the capacity for river trade especially, but also road-based trade. And last but not least, there is a high prevalence of unobservable obstacles to trade such as impassable roads and conflict in North Kivu. Such obstacles may mean that our estimates are subject to significant measurement error and even downward biases. All of these factors should remind us that although roads and other infrastructures do indeed "pave the way" for future developments, the returns to roads still heavily depend upon how they are used.

Appendix A. The state of agriculture in the DRC

Table A1 below presents some key indicators regarding the role of agriculture in the DRC. The DRC's Global Hunger Index is about equal to that of other sub-Saharan African countries, but the DRC is a large net agricultural importer and a large net food importer (about three times the African average). Around 27% of cereal consumption is based on imports. Agriculture is also clearly an important sector in a country with such a large rural population. Indeed, a 67% rural share probably understates the true share in rural areas.

Table A1. Key indicator of agricultural and nutritional status in the DRC

	DRC	Africa	LATAC	East Asia	South Asia	MENA
Global Hunger Index[a] (1-100)	25.1	24.4	8.9	14.0	24.8	7.8
Net agricultural exports[b] (% total imports)	-4.2	15.3	13.0	10.2	-2.0	-4.7
Net food exports[b] (% imports)	-7.3	-2.7	5.6	1.0	-1.7	-2.8
Cereal imports[c] (% cereal consumption)	27.0	37.5	44.2	16.4	10.0	49.4
GDP per capita[d] (2000 I$)	272	2,309	7,432	4,548	2,079	5,547
Rural population[d] (% total)	67.3	62.0	35.2	61.4	75.8	40.5

Notes: Sources for data are as follows: a. IFPRI; b. Aksoy and Dik-melik (2008); c. FAO (2008); d. World Bank (2008). Only low and middle income countries are included. LATAC is Latin America and the Caribbean, and MENA is the Middle East and North Africa. 'Africa' refers only to sub-Saharan Africa.

Appendix B – The Spatial Allocation Model (SPAM) for estimating crop production

We define our spatial crop allocation problem in a cross entropy framework (You and Wood, 2006). The first thing to do is to transform all real-value parameters into a corresponding probability form. We first need to convert the reported harvested area, $HarvestedArea_{jl}$ for each crop j at input level l into an equivalent physically cropped area, $CropArea_{jl}$, using cropping intensity.

(3.1) $$CropArea_{jl} = HarvestArea_{jl} / CroppingIntensity_{jl}$$

Let s_{ijl} be the area share allocated to pixel i and crop j at input level l with a certain country (say **X**). A_{ijl} is the area allocated to pixel i for crop j at input level l in country **X.** Therefore:

(3.2) $$s_{ijl} = \frac{A_{ijl}}{CropArea_{jl}}$$

Let π_{ijl} be the prior area shares we know by our best guess for pixel i and crop j at input level l in country **X**. The modified spatial allocation model can be written as follows:

(3.3) $$\underset{\{s_{ijl}\}}{MIN} \qquad CE(s_{ijl}, \pi_{ijl}) = \sum_i \sum_j \sum_l \ s_{ijl} \ln s_{ijl} - \sum_i \sum_j \sum_l \ s_{ijl} \ln \pi_{ijl}$$

subject to:

(3.4) $$\sum_i s_{ijl} = 1 \quad \forall j \, \forall l$$

(3.5) $$\sum_j \sum_l \ CropArea_{jl} \times s_{ijl} \le Avail_i \quad \forall i$$

(3.6) $$CropArea_{jl} \times s_{ijl} \le Suitable_{ijl} \quad \forall i \, \forall j \forall l$$

(3.7) $$\sum_{i \in k} \sum_l \ CropArea_{jl} \times s_{ijl} = SubCropArea_{jk} \quad \forall k \forall j \in J$$

(3.8) $$\sum_{l \in L} CropArea_{jl} \times s_{ijl} \le IRRArea_i \quad \forall i$$

(3.9) $$1 \ge s_{ijl} \ge 0 \quad \forall i, j, l$$

where:

i : $i = 1, 2, 3, \ldots,$ pixel identifier within the allocation unit, and

j: $j = 1, 2, 3, \ldots,$ crop identifier (such as maize, cassava, rice) within the allocation unit, and

l: l = *irrigated, rainfed-high input, rainfed-low input, subsistence,* management and input levels for crops

k: $k = 1, 2, 3, \ldots$, identifiers for sub-national geopolitical units

J: a set of those commodities for which sub-national production statistics exist

L: a set of those commodities which are partly irrigated within pixel i.

$Avail_i$: total agricultural land in pixel i, which is equal to total agricultural area estimated from land cover satellite image as described in the previous section.

$Suitable_{ijl}$: the suitable area for crop j at input level l in pixel i, which comes form FAO/IIASA suitability surfaces as introduced in the previous section.

$IRRArea_i$; the irrigation area in pixel i from global map of irrigation

The objective function of the spatial allocation model is the cross entropy of area shares and their prior. Equation (3.4) is adding-up constraints for crop-specific areas. Equation (3.5) is land cover image constraint that the actual agricultural area in pixel i from satellite image is the upper limit for the area to be allocated to all crops. Equation (3.6) is the constraint that the allocated crop area cannot exceed what are suitable for the particular crop. Constraint (3.7) sets the sum of all allocated areas within those subnational units with existing statistical data to be equal to the corresponding subnational statistics. Constraint (3.8) includes the irrigation information: the sum of all allocated irrigated areas in any pixel must not exceed the area equipped for irrigation indicated in global map of irrigation (Siebert et al, 2001). The last equation, Equation (3.9) is basically the natural constraint of s_{ijl} as shares of total crop areas.

Obviously an informed prior (π_{ijl}) is very important for the success of the model. We create the prior based upon the available evidence. First for each pixel, we calculate the potential revenue as

$$(3.10) \qquad Rev_{ijl} = Price_j \times Pricevar_{ij} \times Yield_{jl} \times Suitability_{ijl} \times Suitable_{ijl}$$

where $Price_j$ and $Yield_{jl}$ are the price index and the average yield for crop j at input level l (yield only) for the allocation unit (countries in SSA), $Suitability_{ijl}$ is the suitability for crop j at input level l and pixel i, which is represented as proportion (value between 0 and 1) of the optimal yield. $Pricevar_{ij}$ is the price variability (value between 0 and 1) for crop *j* and pixel *i*. Currently we use the population density as an approximation for spatial price variation. Then we pre-allocate the available statistical crop areas (at various geopolitical scales) into pixel-level areas by simple weighting:

(3.11) $$Area_{ijl} = SubCropArea_{jk} \times Percent_{jl} \times \frac{\mathrm{Re}\,v_{ijl}}{\sum_{i \in k} \mathrm{Re}\,v_{ijl}} \quad \forall j \forall i \forall l$$

where $Areai_{jl}$ is the area pre-allocated to pixel *i* for crop *j* at level *l*, $Percent_{jl}$ is the area percentage of crop *j* at input level *l*. For those geopolitical units without area statistics, we simply merge them together and obtain the total area for that merged unit by subtracting the sum of available subnational areas from national total. After this pre-allocation, we calculate the prior by normalizing the allocated areas over the whole country.

(3.12) $$\pi_{ijl} = \frac{Area_{ijl}}{\sum_{i} Area_{ijl}} \quad \forall j \forall i \forall l$$

To convert the allocated crop areas into production, we need consider both the broader production systems and the spatial variation within the systems. We first calculate an average potential yield within SRUs, $\bar{Y}_{jl}$, for crop *j* in production system *l* using the allocated areas (A_{ijl}) as weight:

(3.13) $$\bar{Y}_{jl} = \frac{\sum_{i} Suitability_{ijl} \times A_{ijl}}{\sum_{i} A_{ijl}}$$

We then estimate the actual crop yield of crop *j* in production system *l* and pixel *i* (Y_{ijl}) as

(3.14) $$Y_{ijl} = \frac{Suitability_{ijl} \times Yield_{jl}}{\bar{Y}_{jl}}$$

where $Yield_{jl}$ is the statistical yield (from census data) for crop *j* in production system *l*. The production of crop *j* in production system *l*, and pixel *i*, $Prod_{ijl}$, could be calculated as the following:

(3.15) $$\mathrm{Pr}\,od_{ijl} = (A_{ijl} \times CroppingIntensity_{j}) \times Y_{ijl}$$

Appendix C. Results with the log of crop production per capita as the dependent variable

Regression No.	C1	C2	C3	C4	C5	C6	C7
Estimation method	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit	Tobit
Ln(travel time to 50K city)	-0.37***	-0.35***	-0.04				
Ln(travel time fluvial port)			-0.44***	-0.47***			0.013
Ln(travel time to 100K city)					-0.36***	-0.34***	-0.48***
Ln(potential production, low inputs)	0.18***	0.18***	0.18***	0.18***	0.18***	0.18***	0.18***
Ln(100 km^2 pop. density)	-0.55***	-1.25***	-1.34***	-1.32***	-0.54***	-1.27***	-1.38***
Ln(100 km^2 pop. density), squared		0.049***	0.056***	0.054***		0.051***	0.059***
Total observations	15,122	15122	15122	15136	15125	15125	15125
Pseduo R-squared	0.125	0.126	0.127	0.127	0.125	0.125	0.127
Territorial fixed effects	Yes	Yes	yes	yes	Yes	yes	yes

Notes: * significant at 10%; ** significant at 5%; *** significant at 1%.

References

AITD & UNESCAP, 2000. Evaluation of Infrastructural Interventions for Rural Poverty. Asian Institute of Transport Development and United Nations Economic and Social Commission for Asia and the Pacific.

http://www.unescap.org/ttdw/Publications/TPTS_pubs/pub_1959/rurpovfulltext.pdf

Ahmed, R., Hossain, M., 1990. Developmental Impact of Rural Infrastructure in Bangladesh, IFPRI, Research Report 83, Washington, DC.

Ali, Ifzal, and Ernesto M. Pernia. 2003. *Infrastructure* and Poverty Reduction – What is the Connection? ERD Policy Brief Series, Economic and Research Department, Number 13, Asian Development Bank

Binswanger, H.P., Khandker, S.R., Rosenzweig, M.R., 1993. How infrastructure and financial institutions affect agricultural output and investment in India. Journal of Development Economics 41, 337–366.

Center for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IFPRI); and World Resources Institute (WRI). 2000. Gridded Population of the World (GPW), Version 2. Palisades, NY: CIESIN, Columbia University. Available at http://sedac.ciesin.columbia.edu/plue/gpw.

Cirera, X, Arndt, C. 2008. Measuring the impact of road rehabilitation on spatial market efficiency in maize markets in Mozambique. Agricultural Economics 39 (2008) 17–28.

Dercon, S, Gilligan, D.O., Hoddinott, J., Woldehanna, T, 2008. The Impact of Agricultural Extension and Roads on Poverty and Consumption Growth in Fifteen Ethiopian Villages. IFPRI Discussion Paper 00840.

Dorosh, Paul, Schmidt, Emily. 2008. Mozambique Corridors: Implications of Investments in Feeder Roads. Unpublished manuscript. The World Bank, Washington DC.

Dorosh, Paul, Wang, Hyoung-Gun, You, Liang and Emily Schmidt, 2009. Crop Production and Road Connectivity in Sub-Saharan Africa: A Spatial Analysis. Unpublished manuscript. The World Bank, Washington DC.

Fan, Shenggen (2008). Public Expenditures, Growth, and Poverty: Lessons from Developing Countries Baltimore, MD: John Hopkins University Press.

Fischer, Gunther, M. Shah, H. Velthuizen, F. Nachtergaele, 2001. *Global Agro-ecological Assessment for Agriculture in the 21st Century*, International Institute for Applied Systems Analysis, Laxenburg, Austria

FAO (Food and Agriculture Organization). 1981. *Report of the Agro-Ecological Zones Project*, World Soil Resources Report No 48 (1-4). Rome: FAO.

FAO. 2003. *World agriculture towards 2015/2030: An FAO perspective*. FAO and EarthScan.

Filmer, D. and Pritchett, Lant (2001). Estimating wealth effects without expenditure data - or tears: an application to education enrollment in states of India. Demography, 38: 115-132.

Glewwe, P., M. Gragnolati, and H. Zaman. 2000. Who Gained from Vietnam's Boom in the 1990s? An Analysis of Poverty and Inequality Trends. World Bank Working Paper 2275, Washington, D.C.

Jacoby, H.G., 2000. Access to markets and the benefits of rural roads. Econ. J. 110, 717–737.

Jalan, J., and M. Ravallion, 2002. "Geographic Poverty Traps? A Micro Model of Consumption Growth in Rural China." Journal of Applied Econometrics 17(4):329-46.

Khachatryan, A., von Oppen, Matthias, Doluschitz, Reiner and Khachatryand N. 2005. Response of Plant Productivity to Improved Agricultural Markets in India: Application of an Advanced Econometric Cross-Section Time Series Analysis. Tropentag Conference on International Agricultural Research for Development, Stuttgart-Hohenheim, October 11-13, 2005

Kwon, E. K., 2000. "Infrastructure, Growth, and Poverty Reduction in Indonesia: A Cross-sectional Analysis." Asian Development Bank, Manila.

Minten, Bart, Kyle, Steven, 1999. The effect of distance and road quality on food collection, marketing margins, and traders' wages: evidence from the former Zaire. Journal of Development Economics. 60 (1999): 467–495.

Ruijsav, Arjan, Schweigmanb, Caspar, Lutz, Clemens, 2004. The impact of transport- and transaction-cost reductions on food markets in developing countries: evidence for tempered expectations for Burkina Faso. Agricultural Economics 31 (2004) 219-228

Sahn, David E. and Stifel, David C. (2003). Urban-Rural Inequality in Living Standards in Africa. Journal of African Economies, 12: 564-597.

Shannon, C. 1948. *A Mathematical Theory of Communication*, Bell System Technology Journal, 27(1948): 379-423.

Siebert, Stefan, P. Döll and J. Hoogeveen, 2001. *Global map of irrigated areas version 2.0*. Center for Environmental Systems Research, Univerity of Kassel, Germany / Food and Agriculture Organization of the United Nations, Rome, Italy

Van de Walle, D., 2002. Choosing rural road investments to help reduce poverty. World Development 30, 575–589.

Wood, Stanley, K. Sebastian, F. Nachtergaele, D. Nielsen, and A. Dai, 1999. Spatial Aspects of The Design and Targeting of Agricultural Development Strategies, *Environment and Production Technology Division Discussion Paper No. 44,* International Food Policy Research Institute, Washington D.C.

World Bank, 2006. Democratic Republic of the Congo Agricultural Sector Review. World Bank, Washington D.C.

You, L. and S. Wood. 2006. An entropy approach to spatial disaggregation of agricultural production. *Agricultural Systems*. Vol.90, Issues1-3, 329-347.

You, L., S. Wood, U. Wood-Sichra, J. Chamberlin. 2007. Generating plausible crop distribution maps for Sub-Sahara Africa using a spatial allocation model. *Information Development,* Vol.23, No.2/3, p.151-159

You, L., S. Wood, U. Wood-Sichra. 2009. Generating plausible crop distribution and performance maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural System* 99, Issues 2-3, p.126-140.