Working paper

International Growth Centre

Spatial integration, agricultural productivity, and development

A quantitative analysis of Ethiopia's road expansion program

Tasso Adamopoulos

August 2018

When citing this paper, please use the title and the following reference number:
F-32404-ETH-1

DIRECTED BY

LSE

UNIVERSITY OF OXFORD

FUNDED BY

UKaid from the British people

Spatial Integration, Agricultural Productivity, and Development: A Quantitative Analysis of Ethiopia's Road Expansion Program [†]

Tasso Adamopoulos
York University

August 2018

Abstract

I study the effects of Ethiopia's 1996-2014 road expansion program on aggregate and local agricultural productivity and development outcomes. I combine a quantitative spatial framework with a novel district-level panel data set on agricultural production and transport costs. I estimate transport costs between district centers and domestic crop markets accounting for the volume and quality of the road network, and the topography of the terrain. The model features multiple rural locations, where delivering crops to market, as well as accessing intermediate inputs is subject to location-good-specific transport costs. The spatial heterogeneity of transport costs affects the distribution of production and mobile inputs across locations, and the allocation of land across crops within locations. I calibrate the model to the 1996 spatial agricultural production structure of Ethiopia, and then change transport costs to their 2014 levels. The model implies a substantial increase of 13.6% in the aggregate real yield, which rises by 20% with the direct resource savings from lower transport costs. These gains account for about 10% of the overall yield gain in the data over 1996-2014. The model also delivers a U-shaped pattern of yield gains across districts with respect to transport costs, similar to the one observed in the data.

[†]I gratefully acknowledge financial support from the IGC. I am grateful to Alemayehu Seyoum Taffesse and the Central Statistical Agency for sharing the data from the Agricultural Sample Surveys. The GIS consultant Mohammed Nuru, and Stein Monteiro provided superb research assistance. I thank Andrew Dickens, Doug Gollin, Pramilla Krishnan, and Theodore Papageorgiou for helpful comments. All errors are my own. Contact: Tasso Adamopoulos, aadamo@yorku.ca.

1 Introduction

In the late 1990s Ethiopia was a low income country, with its economy heavily skewed towards agriculture, an employment share in agriculture of over 85%, and its agricultural productivity at 55% of its 1960s level, in real terms.[1] At the same time, Ethiopia had one of the lowest road network densities and motor vehicle usages in the world, and high domestic transport costs.[2] These characteristics were shared by many other developing countries, particularly in Sub-Saharan Africa.

Starting in 1997, Ethiopia embarked on a comprehensive program of road and highway expansion, in terms of both volume and quality of network, with the explicit goal of increasing the connectivity of rural communities to markets. The volume of the entire road network increased 3-fold, while the volume of the rural road network increased 4.7-fold over 1996-2014. The proportion of asphalt roads in good condition increased from 17% in 1997 to 73% in 2010. The proportion of rural roads in good condition increased from 21% to 53% over the same period. Since then real agricultural productivity has not only rebounded, but surpassed its 1960s levels. Given the importance of agricultural productivity for poverty reduction, the process of development and the accompanying process of structural change, it is important to understand what factors affect agricultural productivity and to quantify their contribution. In this paper, I study quantitatively the effects of Ethiopia's road expansion program on aggregate and local agricultural productivity outcomes, as well as its development process.

To quantify the gains from improved market access opportunities on agricultural productivity I use a quantitative spatial model and micro-level data from Ethiopia. In particular, I first construct a novel district-level (*woreda*) panel data set, over 1996-2014, that overlays agricultural production data with geo-coded transport costs associated with the road network expansions. The agricultural production component of the panel draws from repeated waves of household-level data from the Ethiopian Agricultural Sample Surveys, on the type and quantity of crops produced, land allocations

[1] Based on data from the Groningen Growth and Development (GGDC) 10-sector database.

[2] Based on data from Adamopoulos (2011), and World Development Indicators, World Bank.

by crop, as well as input use. I estimate the geo-coded transport costs between district centers and crop markets, using detailed GIS information on the road infrastructure network at each point in time and high resolution data on the topography of the terrain that has to be travelled to reach the relevant market.

I then develop a simple spatial equilibrium model featuring an urban center and multiple rural agricultural production locations. Each rural location can produce a food crop for domestic consumption, or a cash crop for the export market. Consumers in the urban location have non-homothetic preferences over the consumption of food, and non-agricultural goods produced in the urban center. Shipments of crops to the urban center for consumption or export are subject to domestic crop-location-specific transportation costs. Transport costs also raise the cost of disbursing imported intermediate inputs from the urban centre to the multiple rural locations. The food farming technology also requires labor, as does non-agricultural production in the urban center. Rural locations are heterogeneous along three dimensions: the total amount of agricultural land; the productivities with which they produce crops; and the crop-specific transportation costs they face in delivering crops to markets (and accessing intermediate inputs). Changes in the distribution of good- and location-specific transportation costs reallocate food production across locations, alter the allocation of land across crops within locations, as well as the distribution of intermediate input use and labor across locations. The model has implications for both aggregate and local outcomes, associated with changes in transport costs.

To isolate the effects of transport cost changes over 1996-2014 on productivity, my empirical approach involves three steps. First, I calibrate the spatial production structure of the model to aggregate and micro-level economic and geographic transport cost data for the Ethiopian economy for 1996, before the comprehensive road infrastructure program began. Second, keeping all else equal, I feed into the model exogenously only the actual changes in transportation costs, implied by my data on the actual changes in the volume and quality of the road network in Ethiopia. Third, I compare the equilibrium changes implied by the model with only transport cost changes to the actual changes in the data over the period 1996-2014, in terms of both aggregate statistics and spatial distributional patterns across districts. The quantitative experiment described essentially answers the following question: What would the aggregate real yield have been in 1996, if the road infrastructure was that of 2014 rather than 1996?

To implement this empirical approach I combine the above structural model with the district-level panel data. The spatial unit of observation in the data, corresponding to a location in the model, is a woreda, a fairly small spatial unit. The model requires district-level information on the total amount of arable land, the transportation costs, and the productivity of food and cash crops. I take the amount of arable land per district (in hectares) directly from the AgSS data, and I back out the crop-specific transportation costs directly from the travel times to the nearest grain market and Addis Ababa, estimated in the geographic component of the analysis. Then I use the structure of the model to find the vector of crop-specific productivity terms such that the equilibrium of the model before the infrastructure expansion, exactly matches total real output per hectare for cereals and the allocation of land (in hectares) across the two crops, in each district (woreda).

With the model calibrated to the 1996 spatial agricultural production structure of the Ethiopian economy, I then change transport costs to their 2014 levels. The model implies a substantial increase of 13.6% in the aggregate economy-wide real yield. This number is 20% higher if the direct resource savings from lower transport costs are taken into account. To appreciate the magnitude of these gains, I note that they account for about 10% of the overall yield gain experienced by Ethiopia over the period 1996-2014. In terms of the mechanism, in the model, as transport costs fall overall, food production is increasingly undertaken by relatively more productive rural districts. Given that the demand for food is inelastic, this allows for an overall shift to cash crops in the economy. The labor required for food production falls, generating a structural shift towards non-agriculture, with an associated increase in average farm size. The increase in the land share to cash crops, the drop in the agricultural labor share, and the increase in the average farm size produced by the model, under the 2014 transportation costs changes alone are in the same neighborhood as the ones observed in the data over the period 1996-2014.

I terms of local outcomes, I find that the distribution of the gains is uneven across districts. The model, under the 1996-2014 changes in transportation costs delivers a U-shaped pattern of woreda-level yield gains with respect to transport costs across districts, where the non-linearity is linked to the variation in relative food to cash crop transport costs. I show this U-shaped relationship is present not only in the model but also in the data over the period 1996-2014.

While the model is stylized it is rich enough to capture key aspects of the spatial structure of agricultural production in developing economies. In particular, it allows for "within" location choices, such as crop and intermediate input choices, as well as "across" location production reallocation. Its simplicity allows me to treat each woreda in the data as the unit of observation. As a result, I do not have to rely on parametric assumptions about the distributions from which transportation costs and productivities are drawn from, for each crop-woreda pair. Instead, transportation costs, before and after, are estimated from geographic measures of travel times from GIS software, and productivities by crop for each rural location are backed out from the model by matching woreda-level targets in the data.

These results suggest that better access to markets, afforded by the expansion of the road network have had a sizable return in terms of aggregate productivity and development outcomes. In addition, the analysis shows that the gains from these changes are not uniform across localities. In particular, the districts that experience the largest yield gains are not necessarily only those that experience the largest drops in the *level* of their transport cost. There are other factors driving heterogeneity in the responses of localities even after controlling for transport costs, such as relative transport costs across crops, and the relative productivities. The implication is that one should not expect a uniform response to lowering transport costs across the board, in the face of inherent heterogeneity.

The importance of agriculture for development has been emphasized in the earlier development literature, e.g. Schultz (1953), and in a more recent quantitative macroeconomics literature, which shows that agriculture plays a key role in understanding the large productivity disparities across countries, Gollin et al. (2002), Restuccia et al. (2008), Caselli (2005). Developing countries are much more unproductive in agriculture than in non-agriculture when compared to developed countries, and in addition employ most of their labor in agriculture. An important challenge for policy and academic research alike is to understand why agricultural productivity is so low in developing countries. There are several recent contributions in the macro-development literature that study this question, among others Lagakos and Waugh (2013), Adamopoulos and Restuccia (2014), Gollin et al. (2014), Tombe (2015), Donovan (2016), Adamopoulos and Restuccia (2018). This paper contributes to this literature by studying a distinct factor, the importance of farm connectivity to markets.[3]

A recent literature in macroeconomics shows that internal transport costs matter for development and the sectoral composition of the economy: Adamopoulos (2011), Herrendorf et al. (2012), Gollin and Rogerson (2014). This paper contributes to this literature by overlaying micro-data on farm production and detailed geo-coded market access data, to evaluate the impact of a particular road expansion program.

This paper relates to a large literature studying the economic impacts of transport infrastructure investments, in the form of roads, highways or railroads. One strand of the literature uses general equilibrium trade or economic geography models to measure the effects of transport infrastructure projects, e.g., Donaldson (2018), Donaldson and Hornbeck (2016), Allen and Arkolakis (2014), Alder (2018), Asturias et al. (2016) among others. A more recent literature studies the welfare impact of changes in the transportation network in a general equilibrium setting, Allen and Arkolakis (2016), Fajgelbaum and Schaal (2017), Felbermayr and Tarasov (2015). None of these papers however focus on agriculture per se. A related literature estimates local effects of transport infrastructure expansion, e.g., Banerjee et al. (2012), Faber (2014) Baum-Snow et al. (2017), Storeygard (2016),

[3] Adamopoulos and Restuccia (2014) emphasize idiosyncratic policies that affect resource allocation in agriculture. Idiosyncratic transport costs can be viewed through these lens as one particular policy that affects the spatial allocation of resources.

and most closely related to this paper in context Asher and Novosad (2018). A key characteristic of this literature is the use of appropriate identification strategies to address the potential endogeneity of the placement of the relevant transport infrastructure, and estimate its causal effects. Brooks et al. (2017) use a hybrid approach to show the effect of improved market access in rural Nicaragua.

This paper is most closely related to two notable papers, Costinot and Donaldson (2016) and Sotelo (2018), who also employ multi-region spatial frameworks that link domestic trade frictions with agricultural productivity, and welfare, when factors are allocated on the basis of comparative advantage. In addition, Sotelo (2018) examines the effects of counterfactual changes in the infrastructure policy in Peru. Besides the country contexts, models and calibration approaches being different, in addition I evaluate the effects of the actual changes in the road network of Ethiopia over time.

The paper proceeds as follows. Section 2 describes the road network data, and how they are used to estimate the geo-coded transport costs. In Section 3 I discuss the agricultural production data, and panel assembling. The spatial framework is developed in Section 4. I calibrate the model to aggregate, and district-level moments from the Ethiopian data in Section 5. Section 6 reports the aggregate and distributional effects from the quantitative experiments. I conclude in Section 7

2 Roads Data

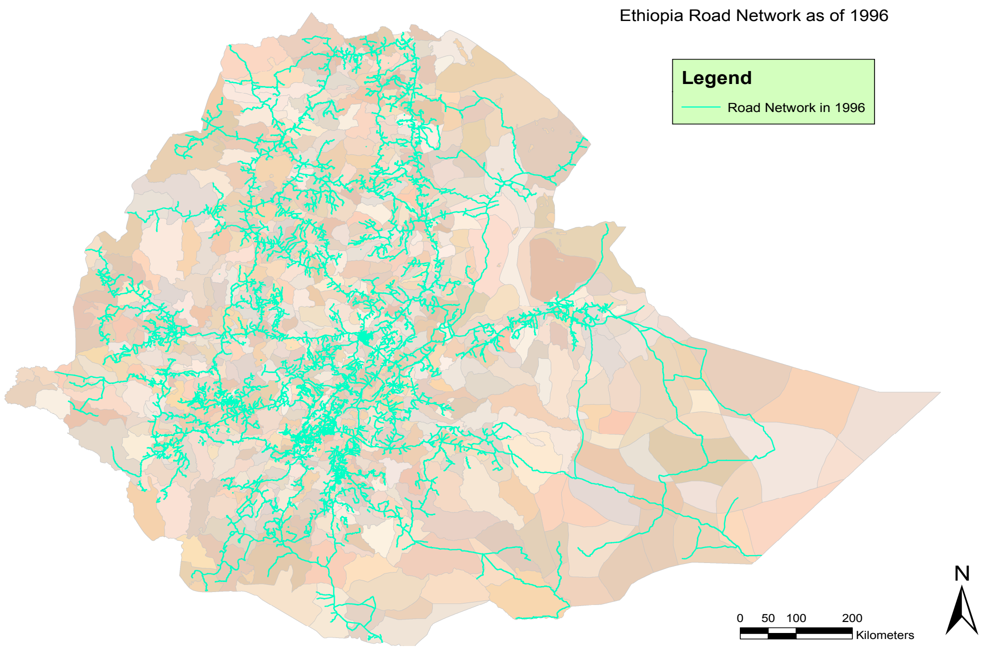
Over the last two decades Ethiopia has embarked on an extensive road development program, as a pilar of its growth strategy. Starting in 1997, through the implementation of successive Road Sector Development Programmes there has been substantial improvement in the volume and distribution of the road network, as well as the conditions of the existing roads. The Universal Rural Road Access Program (URRAP), under the country's recent Growth and Transformation Plan (2010-2015), aimed to extend roads that would connect all small administrative units (*kebele*) in rural areas to all weather roads.

These efforts have had a substantial impact on the extent and quality of the road network in Ethiopia. The volume of the total network increased almost 3-fold, from 24,970 kilometers in 1997 to 69,951 kilometers in 2014. However, the volume increase in the rural road network has been 4.7-fold (from 9,100 in 1997 to 43,094 kilometers in 2014). The federal road network volume increased from 15,870 km to 26,857 km over 1997-2014 (1.7-fold increase). According to data from the Ethiopian Roads Authority, which keeps track of the changes in the network, the road density (including community roads) over 1997-2010 increased from 24 kilometers per 1000 squared kilometers to 136.6 kilometers per 1000 squared kilometers, and from 0.49 kilometers per 1000 people to 1.83 kilometers per 1000 people. In terms of qualitative indicators, the proportion of asphalt roads in good condition increased from 17% in 1997 to 73% in 2010. The proportion of rural roads in good condition increased from 21% to 53% over the same period.

To assess the effect of Ethiopia's road infrastructure expansion program I use detailed GIS data on the universe of roads in Ethiopia, starting in 1996, just before the program began. In particular, the road network data in vector form are obtained from the *Ethiopian Roads Authority* (ERA) for highways and regional roads, and the *Regional Roads Authorities* for regional roads. This data is obtained biennially for the period 1996-2014, specifically covering the years 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014. The road network data provide information not only on the volume but also on the quality of every link in the network, each year. The data come with information on road class and surface type (e.g., whether a particular road is a highway or a town road, and if a town road whether dirt, or asphalt etc.), year of construction, as well as year of upgrading or rehabilitation.

Figure 1 shows the road network before the comprehensive road expansion program. Figure 2 provides a map of Ethiopia's entire road network as of 2014, indicating both the new links in the network (blue) as well as the links of the pre-1996 network that have been rehabilitated or upgraded by 2014. A casual inspection of the two maps shows a substantial expansion in the volume and quality of the network, especially with respect to feeder roads and roads reach rural dispersed communities.

Figure 1: Roads in Ethiopia - *Before* the Program



Note: With data from the Ethiopian Road Authority (ERA).

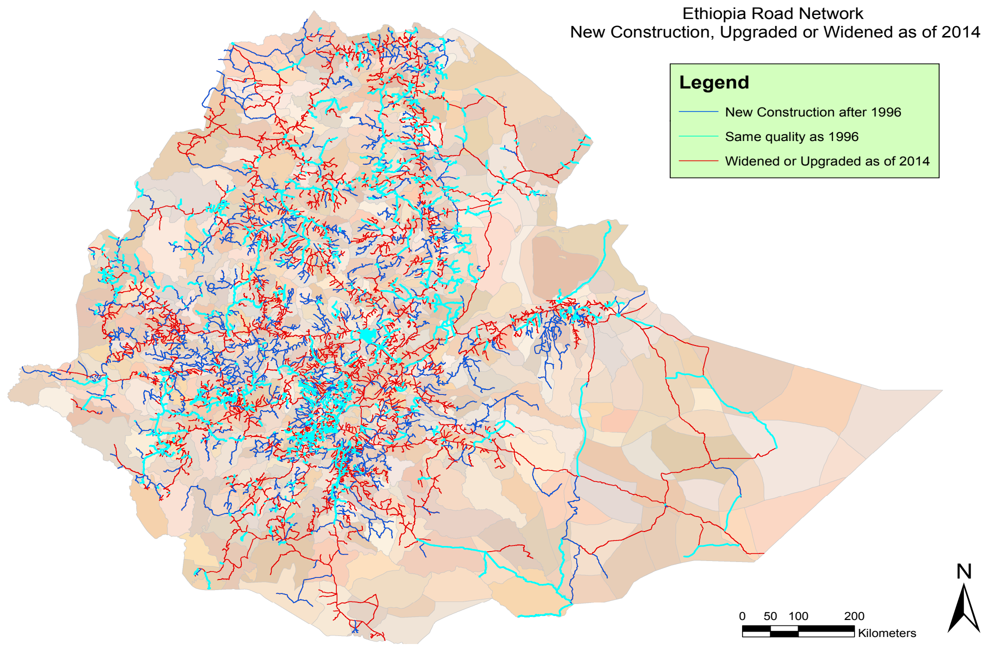
This roads data is the main ingredient going into the estimation of the geo-coded transportation costs, outlined below.

2.1 Geo-coded Transportation Costs

The goal is to estimate geo-coded transportation costs from agricultural production sites to agricultural markets. The spatial unit of observation is taken to be a district (woreda).[4] The main measure of transportation costs I use in the analysis is the travel time in minutes between the district centroids and the nearest destination crop markets. For food crops (cereals), the possible destinations where output can be disbursed are taken to be Ethiopia's 33 major wholesale grain markets (obtained from the Ethiopian Grain Trading Enterprise), which are spread throughout

[4]Ethiopia is subdivided in ascending order of disaggregation into regions, zones, woredas, and kebeles.

Figure 2: Roads in Ethiopia - *After* the Program



Note: With data from the Ethiopian Road Authority (ERA). The network links in blue represent newly constructed roads after 1996. The network links in red represent rehabilitation or quality upgrade of pre-1996 network links.

Ethiopia. The food crop travel time for each district is the travel time to the nearest grain market. For cash crops, that are primarily destined for exporting via the capital, the destination market for computing the domestic transportation cost is Addis Ababa.

To estimate a panel of travel times from woredas to destination crop markets I overlay the universe of the actual road network data by year described above, with high resolution geographic data on elevation and land use, along with the GPS coordinates of the woreda centroids and the destination crop markets. The layer of geographic data on land use and elevation is used to obtain as precise geo-coded estimates of travel time as possible by taking into account the topography of the terrain that has to be travelled to reach the relevant market. This captures, the type of land a farmer would have to travel on foot or animal drawn cart before reaching the road, but also accounts for the fact that travel speeds are different on steep roads than on flat surfaced roads.

In order to implement this methodology the entire extent of Ethiopia is formatted in a high resolution grid where the size of a cell (or pixel) in the grid is 250m $\times$ 250m. All the data are at this fine level of disaggregation in raster files for ArcGIS. The travel time (cost) in each cell depends on whether there is a road or not, what type of road there is if a road exists, the type of terrain within the cell if no road exists, and finally the topography (slope) of the terrain. Using Dijkstra's algorithm, I then determine the optimal route for each district center to each destination grain market as the least-accumulative-cost path. The nearest grain market is the one with the lowest cumulative-cost along the set of optimal routes. The measure of geo-coded transport cost for a district is the travel time along the optimal route to the nearest grain market. I note, that the nearest route market is not held fixed over time but is allowed to change in the algorithm. If a different market becomes the nearest one after a given improvement in the road network that involves a particular Woreda, the computed travel time will be the one to the new nearest market. Note that this measure of travel time varies over time as the extent and quality of the network expands.[5]

Figure 3 shows the average travel time in minutes, from all woreda centers in Ehiopia to their nearest grain market, over the period 1996-2014, following the above methodology. Average travel times to grain markets decreased gradually over time, dropping from 474 minutes (or 7.9 hours) in 1996 to 317.9 minutes (or 5.3 hours), corresponding to a drop of 33%. While the travel times for the entire country dropped considerably over time their level still remained high by 2014. The expansion of the road network in Ethiopia not only reduced average geographic transport costs but also reduced the dispersion of travel costs across woredas. In 1996, 17% of woredas were within 2 hours from a major grain market, while by 2014 this number had increased to 30%.

[5]The travel time from the centroid better reflects the reality that farmers face, as the center of a woreda does not necessarily fall where a town is located. Nevertheless, I also consider alternative measures of transport costs: the distance from the Woreda centroid to the nearest market through the existing road network; the travel time from the centroid to the nearest market without accounting for terrain and land use; the average distance that can be traveled within an hour from the Woreda centroid given the road network (service coverage analysis); the travel time from the Woreda capital to the nearest grain market accounting for topography and land use; the travel time to nearest town with population 20, 50, 100, 250 thousand in turn; the travel time to the nearest port. These measures are all highly correlated.

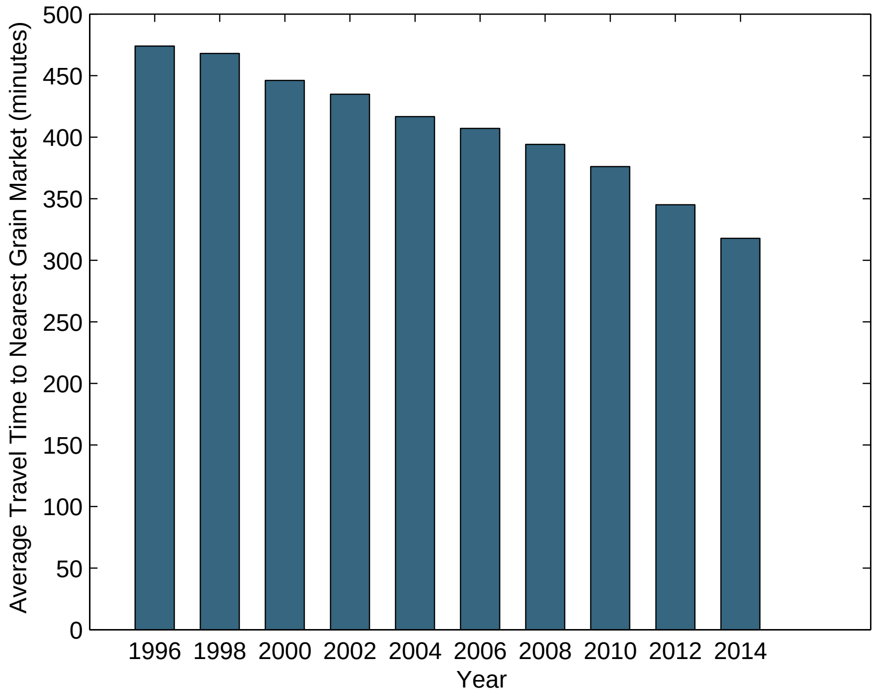


Figure 3: Average Travel Time to Nearest Grain Market Over Time

3 Agricultural Production Data

I use household-level data from the Ethiopian Agricultural Sample Survey (AgSS), a nationally representative annual survey data collected by the Central Statistical Agency (CSA). The data contain information at the field level (a household typically has more than one field) on what crops are produced, what quantity is produced, how much of the land is allocated to the production of the crop, whether fertilizer is used and how much, whether the field is irrigated, and whether it uses improved seed. The data I use cover the period from 1995/96 to 2014/15.[6] Given that the AgSS data do not necessarily follow the same households over time and do not contain GPS information on the location of individual households I conduct the analysis at the woreda-level, the lowest level of spatial disaggregation for which a reliable panel could be constructed.[7] See Warner et al. (2015)

[6]The exceptions are the years 1997/98, 1998/99, 2001/02, and 2002/03 for which data are not available.

[7]Ethiopia is subdivided, in ascending order of disaggregation, into regions, zones, woredas (districts), and kebele (farmer associations).

for a discussion of the challenges involved in assembling a more disaggregate panel.

An issue that arises in merging the AgSS household-level data over the long number of years required for my purposes is that there was redistricting of zones and Woredas over time. To address discrepancies of woreda identifiers that arise from redistricting I homogenize the coding across all years using the 2007 IPUMS zonal and woreda boundaries and identifiers. While the AgSS waves from 2003/04 and on abide by the IPUMS coding, the earlier years to not. The earlier years were cross checked against IPUMS coding using the names of the weredas and zones. This process allows me to match a total of 428 woredas between the earlier and later period.

The quantitative analysis focuses on comparing the period before the comprehensive infrastructure program begins (1997) to the end of the period (2014) of the study. In order to have a more representative sample of household observations per woreda, and to ameliorate any potential noisiness of the household-level data, I pool household data from three years for the earlier period (1995/96, 1996/97, and 1999/00) and three years for the later period (2012/13, 2013/14, and 2014/15).[8]

The above process allows me to obtain a woreda-level panel on agricultural production, land allocations across crops, and input use. The measure of agricultural productivity I use at the woreda-level is the real yield or land productivity, measured as real output per hectare. To construct a real measure of yield over a basket of crops, I aggregate using as common set of prices across woredas, the average prices for each crop over the period 2004-07 in Ethiopia (in local currency units), obtained from the Food and Agricultural Organization (FAOSTAT).

The crops with available output and land data in 1996 and 2014 are all the cereals (barley, maize, millet, oats, rice, sorghum, teff, wheat), and legumes (such as chick peas, dry beans), seeds (such as linseed, sesame, sunflower), spices (such as cardamon, nutmeg), fruit (such as mangoes, papayas, pineapples), vegetables (such as chillies and peppers, garlic, kale), godere, enset, sugar cane, avocados. While coffee has output data at the end of the panel, it does not have output data at

[8] I have also experimented with using just the initial and final year in the AgSS sample, and the results are not significantly different.

the beginning of the panel. Given that the relative price of coffee is high in the price data from FAOSTAT, including coffee only in the later years in the panel, would inflate productivity gains. As a result I exclude coffee in the rest of the analysis.

Here I focus on the set of Woredas (428 in total) with available AgSS data in both the beginning of the period (around 1996) and the end of the period (around 2014). Over the woredas in the balanced panel, the average yield over all crops across woredas increased 4.4-fold, implying an annual average growth rate of 9.7%. Over the same period the yield over grain crops increased 2.5-fold, with an annual average growth rate of 5.9%. This is remarkable growth in real agricultural productivity by any standard. Note, that this growth is not due to price changes since crops have been aggregated using a common set of prices, purging the effects of any possible inflation in prices. While productivity growth has been ubiquitous across virtually all woredas the productivity gains have not been shared equally. Figure 4 shows the histogram of log- growth rates in aggregate real output per hectare across woredas. As is clear from the figure, although almost all growth rates are positive there is wide dispersion across woredas.

Next, I merge the agricultural productivity data from the AgSS with the geo-coded transport cost data, summarized by the travel times from each woreda centroid to the nearest major grain market, and Addis Ababa. Table 1 presents summary statistics for the estimated geo-coded transport costs. There are two points to note. First, average transport costs from woreda centroids to grain markets dropped from 345 minutes in 1996 to 220 minutes in 2014, a -36% change. The transport costs from woreda centroids to Addis Ababa are higher in level, both in 1996 and 2014, dropping by -24% over the period. Second, the dispersion of transport costs across Woredas dropped, implying better accessibility to grain markets for more Woredas. For example, the share of Woredas within two hours of a major grain market increased from 0.20 in 1996 to 0.34 in 2014, for grain markets. In the case of Addis Ababa, this share started from a low level of 0.03, and increased to 0.06 over the period.

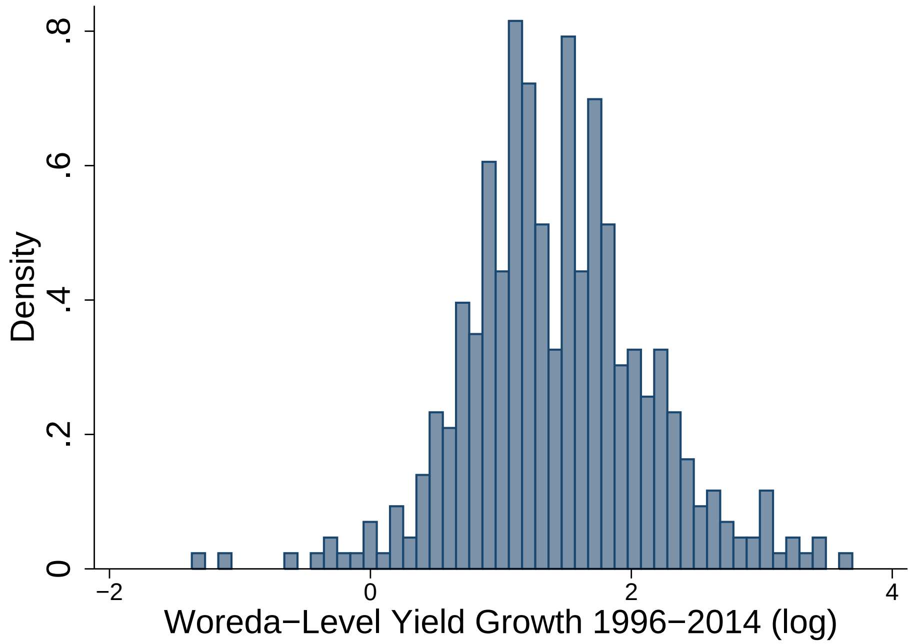


Figure 4: Dispersion of Real Output per Hectare Growth Rates Across Woredas

4 A Spatial Model of Agricultural Production

4.1 Environment

Consider a spatial economy with an urban center and a finite number of J rural locations, indexed by $j \in \mathcal{J} \equiv \{1, 2, ...J\}$. The economy produces two agricultural goods, a food crop f, and a cash crop s, as well as a non-agricultural good n. Agricultural production takes place only in the rural locations, while non-agricultural production takes place only in the urban center. Each rural location can produce either of the two crops. The outputs of the same crop across locations are perfect substitutes for each other. The food crop is used only for domestic consumption in the urban center, while the cash crop is fully exported through the urban center.[9] It is assumed that there is unlimited demand abroad at the international price of the cash crop. The non-agricultural

[9]Qualitatively the results would not change if instead the cash crop was partially consumed domestically. However, because domestic consumption of cash crops is small relative to the domestic consumption of food, as well as the export amount of cash crops, I simplify the model along this dimension.

Table 1: Summary Statistics of Estimated Transport Costs (Travel Time in Min)

	To Nearest Grain Market		To Addis Ababa	
	1996	2014	1996	2014
Mean travel time in min	345.4	220.2	575.1	438.9
Median travel time in min	241.1	165.0	550.9	419.0
Std Dev –log	0.90	0.84	0.60	0.61
90/10 ratio	22.4	19.4	7.6	7.7
75/25 ratio	8.9	7.6	4.0	4.2
Fraction of Woredas < 2 hours	0.20	0.34	0.03	0.06

Source: Author calculations.

good is used only for consumption in the urban center.

Preferences There is a representative household in the urban center with preferences over food and non-agricultural goods,

$$u(c_f, c_n) = \begin{cases} \bar{f} + \log(c_n), & \text{if } c_f \geq \bar{f}. \\ c_f, & \text{if } c_f < \bar{f}. \end{cases}$$

where $\bar{f}$ is the minimum consumption requirement of food, and c_n is the consumption of the non-agricultural good. These non-homothetic preferences capture Engel's law, whereby when income is low it is fully allocated to the consumption of food but as income rises and that level of food consumption is achieved the remaining income is allocated to the consumption of the non-agricultural good. The representative household is endowed with total amount of labor N, that is inelastically supplied to the market. The representative household also owns the productive land in the different locations L_j. Production of each crop in each location j is undertaken by a representative farm. The farms are also owned by the household, and therefore any profits they make accrue to the household as income.

Production of food crop The food crop in each location j is produced using land, labor and imported intermediate inputs, according to a decreasing returns to scale technology,

$$y_{fj} = \left[z_{fj}^{1-\gamma}\left(n_j^{\alpha}\ell_{fj}^{1-\alpha}\right)^{\gamma}\right]^{\theta} x_j^{1-\theta} \tag{1}$$

where y_{fj} is output of the food crop, z_{fj} is food crop productivity, and n_j, ℓ_{fj}, x_j are labor, land, and intermediate inputs respectively used in the production of food in location j. $(1-\theta)$ determines the elasticity of final output with respect to intermediate inputs. The object in brackets raised to θ is the production function net of intermediate inputs, with parameter $\gamma < 1$ regulating the extent of returns to scale. Parameter α captures the importance of labor relative to land. Note that decreasing returns to scale imply incomplete specialization and thus the food crop will be produced by every location j.

Production of cash crop The cash crop in each location j is produced according to a constant returns to scale technology that is linear in land,

$$y_{sj} = z_{sj}\ell_{sj}$$

where z_{sj}, y_{sj}, ℓ_{sj} are productivity, output and land under the cash crop technology.

Production of non-agricultural good The non-agricultural good is produced by a representative firm in the urban location according to constant returns to scale technology that is linear in labor,

$$Y_n = AN_n$$

where A is non-agricultural (labor) productivity and N_n is the amount of labor allocated to non-agricultural production.

The total amount of land in location j can be allocated to the production of food or cash crops within that location. Labor is used only in the production of food crops within each location and it is perfectly mobile across all rural locations and the urban center. The non-agricultural good is the numeraire with its price normalized to one. Let p_f be the relative consumer price of the food crop in the urban location, which is endogenous. Note that because food produced in one location is a perfect substitute for food produced in another location, in equilibrium the consumer prices of food in the urban center from different locations will have to be the same and equal to p_f.

This small open economy imports all the non-agricultural intermediate inputs from abroad, which are assumed to be inelastically supplied in the international market, and in exchange exports the cash crop it produces. Given that the cash crop is fully exported and the non-agricultural intermediate inputs are fully imported their international prices p_s^*, and p_x^* respectively are taken as given (small open economy assumption). While p_x^* is the price of intermediate inputs upon landing in the urban center, the local prices of intermediate inputs in the different rural locations will differ according to their location-specific transportation costs for delivering intermediate inputs. Similarly, the farm-gate price that farms receive for the cash crop will be lower than the international price by their crop-location-specific transport costs.

Transportation Technology Delivery of crops from each rural location to the urban center for consumption (food crop) or export (cash crop), as well as the delivery of imported intermediate inputs from the urban center to the rural locations is subject to origin-good-specific transportation costs of the iceberg form. In particular, to sell 1 unit of crop $i \in \{f, s\}$ to the urban center, farms in location j have to ship $\tau_{ij} \geq 1$ units of the crop. Similarly, in order for one unit of imported intermediate inputs to arrive in rural location j, τ_{xj} units have to be shipped. Given that the consumer price of food has to be the same in the urban center regardless of origin, the transport technology implies that the farm-gate producer prices of food will differ across locations at origin according to the transport costs involved in delivering their output to the market, p_f/τ_{fj}. Similarly the farm-gate price of cash crops will be p_s^*/τ_{sj} and the farm-gate price of imported intermediate inputs $p_x^* \tau_{xj}$ in location j. In other words, transport costs reduce the price farms receive for their goods, and raise the prices they pay for their intermediate inputs.

Market Structure All domestic goods and factor markets are perfectly competitive. The economy-wide market clearing condition for food is,

$$\bar{f} = \sum_{j=1}^{J} c_{fj} \tag{2}$$

where $c_{fj} = \frac{y_{fj}}{\tau_{fj}}$ is the amount of food (consumption) delivered to destination in the urban centre originating from location j, and y_{fj} is the amount of the food crop produced and shipped from rural location j. Note that while the only source of demand for food from any location is the consumers of the city centre the amount of consumption is not equal to the amount of food produced in each rural location, since part of the output "melts" in transit. In other words, the difference between production and consumption of food from each location differs by the transport costs from each location to the urban centre. So c_{fj} is also the amount of net output of the food crop from location j. Within each location there is barrier μ to the allocation of land between cash and food crops, such that the rental price of land under food crops is a fraction of the rental price of land under cash crops,

$$q_{fj} = (1 - \mu)\, q_{sj} \tag{3}$$

where q_{ij} is the rental price of land under crop i in location j. The barrier μ is introduced for quantitative purposes, in order to match the ratio of the aggregate yield in cash relative to food crops in the data. The market clearing condition for land in location j is,

$$\ell_{fj} + \ell_{cj} = L_j \tag{4}$$

The labor market clearing condition requires that the total amount of labor used in all rural locations and the urban location is equal to the total amount of labor in the economy,

$$N_a + N_n = N$$

where N_a is the total amount of labor devoted to agricultural production across all rural locations,

$$N_a = \sum_{j=1}^{J} n_j$$

Since the non-agricultural good is produced and consumed in the urban center, the market clearing condition is,

$$Y_n = c_n$$

The entire amount of cash crop production from each location j is shipped to the urban center for export, with the export value upon arrival at the urban center,

$$ex_j = p_s^* \frac{y_{sj}}{\tau_{sj}}$$

All intermediate inputs are imported, with a value upon reaching their destination in each rural location j

$$im_j = p_x^* \tau_{xj} x_j$$

The small open economy's total exports are $EX = \sum_j ex_j$ and imports are $IM = \sum_j im_j$. The economy's net exports (trade balance) are then given by, $NX = EX - IM$

To sum, rural locations are heterogeneous with respect to: (a) crop-location-specific productivities $\{z_{fj}, z_{sj}\}$; (b) the total amount of productive land L_j; and (c) the vector of location-good-specific transportation costs $\{\tau_{fj}, \tau_{sj}, \tau_{xj}\}$.

4.2 Competitive Equilibrium

The profit maximization problem of the food crop farm in rural location j is given by,

$$\max_{\{n_j,\ell_{fj},x_j\}}\left\{\frac{p_f}{\tau_{fj}}\left[z_{fj}^{1-\gamma}\left(n_j^{\alpha}\ell_{fj}^{1-\alpha}\right)^{\gamma}\right]^{\theta}x_j^{1-\theta}-w_jn_j-q_{fj}\ell_{fj}-p_{xj}x_j\right\}$$

subject to the constraint that the total land allocated to food crop production in a given location cannot exceed the total amount of land in that location, $\ell_{fj}\leq L_j$. w_j and q_{fj} are the wage rate and the rental rate of land in location j. Standard non-linear optimization techniques can be used to solve this problem numerically for every location, given a relative price for food p_f. At an interior optimum, the first order conditions to this problem imply,

$$\alpha\gamma\theta\frac{p_f}{\tau_{fj}}\frac{y_{fj}}{n_j}=w_j \tag{5}$$

$$(1-\alpha)\,\gamma\theta\frac{p_f}{\tau_{fj}}\frac{y_{fj}}{\ell_{fj}}=q_{fj} \tag{6}$$

$$(1-\theta)\,\frac{p_f}{\tau_{fj}}\frac{y_{fj}}{x_j}=p_{xj} \tag{7}$$

where final output can be re-written as,

$$y_{fj}=z_{fj}^{1-\gamma}\left(\frac{n_j}{\ell_{fj}}\right)^{\alpha\gamma}\ell_{fj}^{\gamma}\left(\frac{x_j}{y_{fj}}\right)^{\frac{1-\theta}{\theta}} \tag{8}$$

Equations (5) and (6) imply that the labor-land ratio employed in food production in each location j is given by,

$$\frac{n_j}{\ell_{fj}}=\frac{1-\alpha}{\alpha}\frac{q_{fj}}{w_j} \tag{9}$$

Equation (7) implies that the intensity with which food crop farms apply intermediate inputs depends on the elasticity of final output with respect to intermediate inputs and the relative cost of intermediate inputs to the producer price of food,

$$\frac{x_j}{y_{fj}} = (1-\theta)\,\frac{p_f}{\tau_{fj} p_{xj}} \tag{10}$$

The cash crop farm in each location j solves a simple problem,

$$\max_{\ell_{sj}} \left\{ p_s^* \frac{z_{sj}}{\tau_{sj}} \ell_{sj} - q_{sj} \ell_{sj} \right\}$$

where the first order condition pins down the rental price of land in each location j,

$$q_{sj} = p_s^* \frac{z_{sj}}{\tau_{sj}} \tag{11}$$

The profit maximization problem of the non-agricultural firm in the urban center is,

$$\max_{N_n} \{A N_n - w N_n\}$$

where w is the wage rate. The first order condition implies that the wage rate is pinned down by non-agricultural productivity $w = A$. Given that labor is perfectly mobile across the urban and all rural locations the wage rate in each rural location will be equal to this wage rate, $w_j = w = A$.

Household income consists of labor income, the total return to land from all rural locations and the profits from producing the food crop in each rural location,

$$I = wN + \sum_j (q_j L_j) + \sum_j \pi_j$$

where $\pi_j = \theta\,(1-\gamma)\,\frac{p_f}{\tau_{fj}} y_{fj}$ are the food farm profits in location j.

Given the nature of the preferences the consumer will consume an amount of food $c_f = \bar{f}$ and allocate the residual income to the consumption of non-agricultural goods.

Solving for equilibrium From wage equalization across locations and sectors,

$$w_j = w = A$$

Net land rent equalization across crops within locations (3) implies,

$$q_{fj} = (1-\mu)\,\varphi_{sj}$$

and using (6) it can be shown that, at an interior optimum, the land input demand in food production is,

$$\ell_{fj} = \left[(1-\alpha)\,\gamma\theta p_f \frac{\varphi_{fj}}{(1-\mu)\,\varphi_{cj}}\right]^{\frac{1}{1-\gamma}} \left(\frac{n_j}{\ell_{fj}}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{x_j}{y_{fj}}\right)^{\frac{1-\theta}{\theta(1-\gamma)}} \tag{12}$$

where I define $\varphi_{fj} \equiv \frac{z_{fj}^{1-\gamma}}{\tau_{fj}}$ and $\varphi_{sj} \equiv \frac{p_s^* z_{sj}}{\tau_{sj}}$ as the "effective" productivity terms, that depend not only on actual productivity but also on the iceberg transportation costs.

In equilibrium, given that the world price of imported intermediate inputs in the urban centre is p_x^*, their price in rural location j is augmented by the corresponding transportation cost, $p_{xj} = p_x^* \tau_{xj}$. Given that consumers in the urban center want to consume a fixed amount of food $\bar{f}$ the market clearing condition for food (2) can be used to solve for the equilibrium relative price of food p_f. When the land allocation in every location is an interior optimum, (2) along with (8) and (12) can be used to solve analytically for the equilibrium relative consumer price of food in the urban center,

$$p_f = \frac{\bar{f}^{\frac{\theta(1-\gamma)}{1-\theta(1-\gamma)}}}{\left[\frac{(1-\alpha)}{(1-\mu)}\gamma\theta\right]^{\frac{\gamma\theta}{1-\theta(1-\gamma)}} (1-\theta)^{\frac{1-\theta}{1-\theta(1-\gamma)}}} \frac{1}{\left\{\sum_j \left(\frac{\varphi_{fj}}{\varphi_{cj}}\right)^{\frac{1}{1-\gamma}} \left(\frac{n_j}{\ell_{fj}}\right)^{\frac{\alpha\gamma}{1-\gamma}} \left(\frac{1}{\tau_{fj} p_{xj}}\right)^{\frac{1-\theta}{\theta(1-\gamma)}}\right\}^{\frac{\theta(1-\gamma)}{1-\theta(1-\gamma)}}} \tag{13}$$

With these prices all the equilibrium allocations at the local and the economy-wide level can be determined. The model pins down: the allocation of land across crops within locations; the distribution of agricultural production across space on the rural side of the economy; and the distribution of labor across space (allocation of agricultural labor across rural locations) and sectors (allocation of labor between agriculture and non-agriculture). In equilibrium, the economy-wide aggregate measures of interest and their distribution across space are affected by the overall level of transportation costs in the economy, their variation across goods, and the spatial dispersion of these transport costs.

The allocation of food production across rural locations is largely determined by the comparative advantage of locations in food production, where comparative advantage is determined not only by relative productivity between locations but also relative transport costs. To see this consider the simple case where land is the only factor of production, i.e., $\alpha = 0$ and $\theta = 1$. In this case the relative land allocation to food across locations j and k is determined by relative effective productivities across the two locations,

$$\frac{\ell_{fj}}{\ell_{fk}} = \left(\frac{\varphi_{fj}/\varphi_{sj}}{\varphi_{fk}/\varphi_{sk}}\right)^{\frac{1}{1-\gamma}}$$

This says that the relatively more "productive" location j is in food crops relative to cash crops in comparison to location k, the relatively more land will be devoted to food crops in location j relative to location k, and relatively more of food consumption will come from j relative to k. Note, however that effective productivity for crop i in location j, φ_{ij}, is determined not only by actual productivity z_{ij} but also by transportation costs τ_{ij}. Locations that face relatively high transport costs in producing food relative to cash crops will allocate less of their land to the production of food crops.

Transport costs affect also the labor-land ratio in food production across locations,

$$\frac{n_j/\ell_{fj}}{n_k/\ell_{fk}} = \frac{\varphi_{sj}}{\varphi_{sk}} = \frac{z_{sj}}{z_{sk}}\frac{\tau_{sk}}{\tau_{sj}}$$

Also, in equilibrium the intensity with which farmers use intermediate inputs depends on the transport costs that farmers have to pay for delivering their crops to markets τ_{fj} and the transport cost involved in having intermediate inputs delivered to their farm from the urban centre τ_{xj},

$$\frac{x_j/y_{fj}}{x_k/y_{fk}} = \frac{\tau_{fk}\tau_{xk}}{\tau_{fj}\tau_{xj}}$$

Finally note that average farm size for each rural location is the total amount of land in that location over the total number of (food) farm workers,

$$AFS_j = \frac{L_j}{n_j}$$

In the next section the model economy is calibrated to 1996 woreda-level and aggregate data for Ethiopia, and then the effect of transport infrastructure is assessed through the model by changing only the transportation costs in each woreda and for each good to their 2014 level.

5 Calibration

The spatial unit of observation of a "rural location" in the model is a *woreda* or district in the Ethiopian data. This is the most disaggregate level for which a reliable panel of economic (AgSS) and geographic data could be constructed. The strategy is to calibrate the benchmark economy to the Ethiopian district-level and aggregate data for 1996, just before the comprehensive transport infrastructure program was initiated.

I calibrate the model to match the spatial agricultural production structure of the Ethiopian economy, and in particular the structure over the set of woredas with complete geographic and AgSS data in my data set. The parameters that need to be determined are: (a) the $J \times 2$ matrix of crop-specific productivities across the different locations $\{z_{fj}, z_{sj}\}_{j=1}^{J}$; (b) the $J \times 3$ matrix of iceberg transportation costs for each of the crops, as well as the intermediate inputs between the different rural locations and the urban center $\{\tau_{fj}, \tau_{sj}, \tau_{xj}\}_{j=1}^{J}$; (c) the $J \times 1$ vector of total agricultural land for each location $\{L_j\}_{j=1}^{J}$; (d) non-agricultural productivity in the urban location A; (e) food crop technology parameters (γ, α, θ); (f) preference parameter $\bar{f}$; (g) the barrier to allocating land to cash crops μ.

My calibration approach does not rely on parametric assumptions about the distributions from which transportation costs (τ_{ij}) and productivities (z_{ij}) are drawn for each crop-location pair. Instead, transportation costs before and after are estimated from geographic measures of travel times from GIS software, and productivities by crop for each rural location are backed out from the model by matching woreda-level targets in the data. This procedure is described in detail below.

In the data $J = 402$, which includes the woredas for which agricultural production data form the AgSS and transport cost data are available in both 1996 and 2014. The food crop in the model corresponds to cereals in the data, which account for about 84% of the land allocation overall in the economy. Cash crops are taken to include all other crops. The beginning and end of the period are 1996 and 2014 using the pooled data for each, as described in 3. The world prices of cash crops p_s^*, and intermediate inputs p_x^* are normalized to one, as they do not vary in the quantitative experiments.

Agricultural land by location The total amount of land for each rural location $\{L_j\}_{j=1}^{J}$ is taken directly from the data to be the sum of agricultural land allocated to any crop, food or cash, across all households for that woreda in the 1996 AgSS data.

Total labor The total amount of labor in the economy N is taken directly from aggregate data for the Ethiopian economy in 1996, from the Groningen Growth and Development Centre (GGDC) 10-sector database (Timmer et al., 2015).

Transportation costs by location and good τ_{ij} In the benchmark economy, the computation of transportation costs for the two crops and intermediate inputs are estimated from travel times from woreda centroids to destination markets within Ethiopia, through the existing road infrastructure network, measured from the geographic analysis. When estimating travel times for food crops (cereals) the travel times used are those from woreda centroids to the nearest grain market. Note that in the model there is only one agricultural market in the urban center while in the data there are multiple. I use the travel time to the closest grain market as the measure of travel time to the central market in the model. While these regional grain markets are appropriate for estimating food transport costs they are unlikely to be a good approximation for the costs incurred for selling cash crops and purchasing intermediate inputs. Given that cash crops are primarily exported, and exports run through the capital of Addis Ababa, the transportation cost for cash crops is estimated from the travel time from a woreda centroid to Addis Ababa. Given that the distribution of intermediate inputs is centralized, the travel time between the woreda centroid and Addis Ababa is also used as a proxy for the transportation cost of intermediate inputs. Note that the model has iceberg transport costs, which use up resources, while the data involve travel times. As a result the travel times or "time" costs, associated with transportation, have to be mapped into iceberg transport costs or "goods" costs. To map travel times to iceberg transport costs I posit a transport cost function of the following form,

$$\tau_{ij} = 1 + \psi_i \cdot (tt_{ij})^{\eta}$$

where tt_{ij} is the travel time (in minutes) for good i from rural location j to the market. The parameter η captures the sensitivity of transport costs with respect to travel time, and with $\eta < 1$ the transport cost - travel time relationship is concave. $\psi_i > 0$ is a scale parameter that controls the units, in particular, regulating how far from one the implied transport costs are (with $\psi_i = 0$ there are no iceberg costs and $\tau_{ij} = 1$). Next, I explain how I calibrate the parameters of the transport cost function. Combes and Lafourcade (2005) estimate generalized transport costs for road transportation across French districts and report an elasticity of their transport cost measure with respect to real road distance of 0.8 in 1998. This implies that transport costs rise in a concave fashion with distance. While Combes and Lafourcade (2005) do not report the elasticity with respect to real time, the correlation of real time and real distance in their data is 0.986 in 1998 (Table 5). I impose the same concavity on my transport cost measure with respect to travel time by setting $\eta = 0.8$. Then given the value for η I calibrate the units parameter ψ_i for crops so that the total amount of resources devoted to transport as a share of consumer value of output in the model matches the share of transportation costs in the sales value of food in the data for Ethiopia in 1996. Based on a survey of grain wholesale traders across grain markets in Ethiopia for 1996, Gabre-Madhin (2001) shows that for grain "exporting" regions 26% of the sale price is accounted for by marketing costs of various kinds and the profit margin of the transporter. Direct transport costs, including road stops, during the transportation of grains accounted for 58% of the overall marketing costs, implying 13.2% of the final sale price is transport. This provides a lower bound on the transport cost share. However, given that some of the other marketing costs such as handling, sacking, storage, commission of brokers, travel cost of transporter, and profit of the transport company can arguably be attributed to "transportation," I target a transport cost share of the final sale price of 18%, which is between the lower bound of 13% and the upper bound of 26%. This implies $\psi_i = 0.00258$ for cereals, which I use for both food and cash crops in the model. The transport cost share of the delivered farm-gate price of fertilizer is higher. Minten et al. (2013) using data from Northwestern Ethiopia show that transportation costs, accounting for "last mile" costs, can raise the effective price of chemical fertilizer by up to 50%. I set $\psi_x = 0.0041$, which implies a share of transport costs in the farm-gate cost of intermediate inputs of 36%.

Food crop technology parameters (γ, α, θ) In the food crop production function γ and α regulate the extent of decreasing returns to scale and the income share split between land and labor respectively in the non-intermediate input part of the production function in (1). $(1 - \theta)$ is the elasticity of final output with respect to intermediate inputs. I calibrate α and γ to factor income shares and θ to the intermediate input intensity. As the model abstracts from capital, the capital income share will be captured in the $(1-\gamma)$ share. Given that these are technological parameters, and factor shares for the Ethiopian economy may be distorted, I calibrate them to the United States. Calibrating to the United States implies $\theta = 0.62$ to match an intermediate input cost share in the value of final output of 38% (Valentinyi and Herrendorf, 2008; Prasada Rao, 1993). Matching a capital income share in value added of 1/3 implies $\gamma = 2/3$ and matching a land income share of 0.18 implies $\alpha = 0.73$ Valentinyi and Herrendorf (2008). The resulting income share for labor in value added is 49%.[10]

Barrier to land reallocation The parameter μ that deters the reallocation of land towards cash crops is calibrated so that at the economy-wide level the ratio of average yield for cash crops to average yield for food crops is equal to the value for this statistic from AgSS data in 1996. The cash-food yield ratio is 1.33 in the data.

Rural productivity parameters z_{ij}; urban productivity A For each woreda j the productivity terms of the food crop and the cash crop technologies $\{z_{fj}, z_{cj}\}$ are chosen to match two targets for that woreda: (i) the land allocated to food production ℓ_{fj}, and (ii) the actual yield for food crops y_{fj}/ℓ_{fj} in that location (which is equivalent to targeting output y_{fj} since ℓ_{fj} is also targeted). The woreda-level total yield over food crops aggregates crops using a common set of prices. These targets along with the equilibrium equations are sufficient to recover all the variables of interest in the model. I outline the key steps here.

I first normalize the consumer price of food to 1 in the benchmark economy. Then from the food farm's first order condition with respect to land (6), the first order condition of the cash crop farm (11), and the net land rate equalization across crops within woredas (3) the "effective" productivity

[10]Using micro data for Malawi Restuccia and Santaeulalia-Llopis (2015) estimate factor income shares for labor, land, and capital of 0.419, 0.391, and 0.19 respectively. Given that Ethiopia is close to Malawi in terms of stage of development, these factor shares used to parameterize the non - intermediate input component of the food production function, would imply $\gamma = 0.81$ and $\alpha = 0.52$. According to Prasada Rao (1993) the intermediate input share for Ethiopia in the 1980s is 10%.

parameter for cash crops can be backed-out as,

$$\varphi_{sj} = \frac{(1-\alpha)\,\gamma\theta}{(1-\mu)}\frac{p_f}{\tau_{fj}}\frac{y_{fj}}{\ell_{fj}}$$

The "effective" productivity term for the food crop technology can be recovered residually from the food crop production function in each location (8),

$$\varphi_{fj} = \frac{y_{fj}/\tau_{fj}}{\ell_{fj}^{\gamma}\left(\frac{n_j}{\ell_{fj}}\right)^{\alpha\gamma}\left(\frac{x_j}{y_{fj}}\right)^{\frac{1-\theta}{\theta}}}$$

where I have used (9) and (10) to compute the labor-land ratio and the intermediate input intensity for each woreda. Note, that to do this requires a value for the non-agricultural productivity parameter A, which is calibrated to match a target for the share of labor in agriculture of 86% based on aggregate data for the Ethiopian economy, from the GGDC 10-sector database (Timmer et al., 2015). To see this note that from $N_a = \sum_j n_j$ we can solve for A as,

$$A = \frac{\alpha}{1-\alpha}\frac{1-\mu}{\hat{N}_a}\sum \varphi_{sj}\ell_{fj}$$

where $\hat{N}_a$ is the target for the overall employment in agriculture.

Food consumption requirement $\bar{f}$ Given that food production data and transport costs are targeted explicitly in the calibration as described above, the market clearing condition for food (2) pins down the value of the subsistence food consumption term $\bar{f}$ as the total production of food from all woredas net of transport costs.

The economy-wide calibrated parameters of the model, that are common across all locations, along with their descriptions are provided in Table 2. Table 3 provides a description of the location-specific parameters that are mapped into actual woreda-level data, along with their data targets. The calibrated model does well in replicating aggregate and spatial features of the Ethiopian economy.

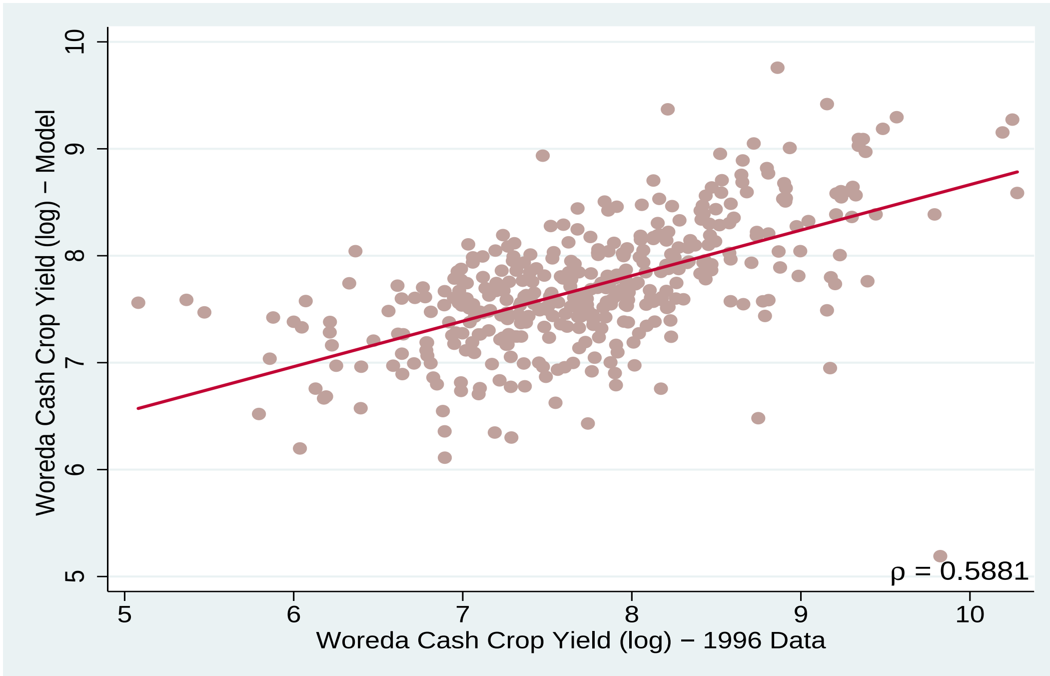
Table 2: Calibrated Common Parameters

Parameter	Description	Value
γ	share of land and labor	0.67
α	share of labor	0.73
θ	non-intermediate input share	0.62
μ	barrier to land reallocation	0.91
η	sensitivity of transport costs to travel time	0.8
ψ_f, ψ_s	transport cost scale parameter - crops	0.00258
ψ_x	transport cost scale parameter - fertilizer	0.00410
A	urban non-agricultural productivity	414.6
N	total number of workers	24806
$\bar{f}$	subsistence food consumption (000s)	29315

Table 3: Calibrated Woreda-specific Parameters

Parameter	Description	Target Data (1996)
$\{L_j\}_{j=1}^{J}$	total agricultural land	total land from AgSS Data
$\{\tau_{fj}\}_{j=1}^{J}$	food crop iceberg transport cost	travel time to nearest grain market
$\{\tau_{sj}\}_{j=1}^{J}$	cash crop iceberg transport cost	travel time to Addis Ababa
$\{\tau_{xj}\}_{j=1}^{J}$	inter. input iceberg transport cost	travel time to Addis Ababa
$\{z_{fj}, z_{sj}\}_{j=1}^{J}$	productivity parameters by woreda	food yield and land share from AgSS data

Figure 5: Cash Crop Yield by Woreda: Model vs. 1996 Data

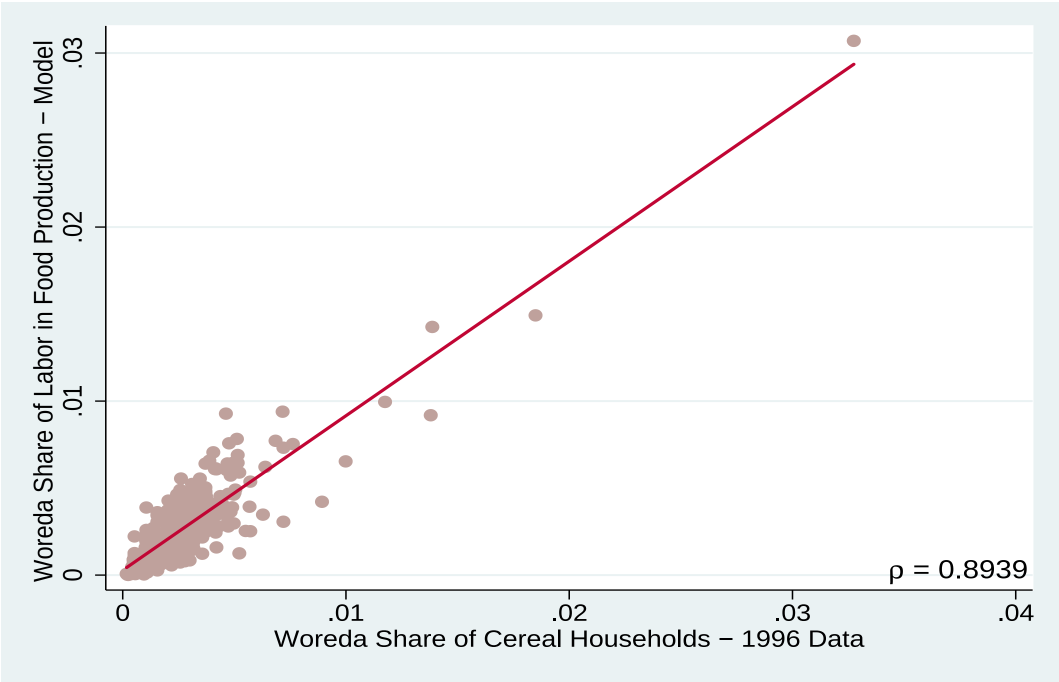


Notes: "Model" refers to the value in the calibrated benchmark economy. "1996 Data" refers to the woreda value from the 1996 pooled data from the AgSS (CSA) in Ethiopia.

The values of key variables of interest in the calibrated benchmark economy are provided in Table 4. However, the model also does well in matching woreda-level statistics that are not targeted in the calibration. Figure 5 compares the woreda-level yields for cash crops implied by the model against their counterparts in the 1996 pooled AgSS data, that were not explicitly targeted. Figure 6 compares the spatial distribution across woredas of food farm labor (share in total economy-wide labor engaged in food production) to the 1996 distribution of households engaged in cereal production across woredas (as a share of the total households engaged in cereals).[11] There is a strong positive correlation between model and data for both the cash crop yields and the food crop labor shares. The AgSS provide reliable information on whether any given field operated by each household uses fertilizer. In addition, the AgSS contains information on the amount of fertilizer applied. However, this data is more sparse and less reliable for time series comparisons. Using the

[11]The AgSS does not provide information on labor. I use the number of households engaged in cereal production as a proxy for food labor.

Figure 6: Food Farm Labor Share by Woreda: Model vs. 1996 Data



Notes: "Model" refers to the value in the calibrated benchmark economy. "1996 Data" refers to the woreda value from the 1996 pooled data from the AgSS (CSA) in Ethiopia.

AgSS data I construct a woreda-level measure of intermediate input use as the share of all fields that have any amount of fertilizer applied to them. On average this has increased from 32% in 1996 to 52% in 2014, indicating a significant increase in the use of fertilizer. In the model, there are no fields, so an intermediate input intensity woreda-level measure can only be constructed as the share of intermediate inputs in final output in each woreda. The spatial distribution of woreda-level intermediate input intensities is not targeted in the calibration. Nevertheless, as Figure 7 shows the intensive margin intermediate input measure from the benchmark economy in the model is strongly positively correlated with the extensive margin intermediate input intensity measure, with a correlation coefficient of 0.48.

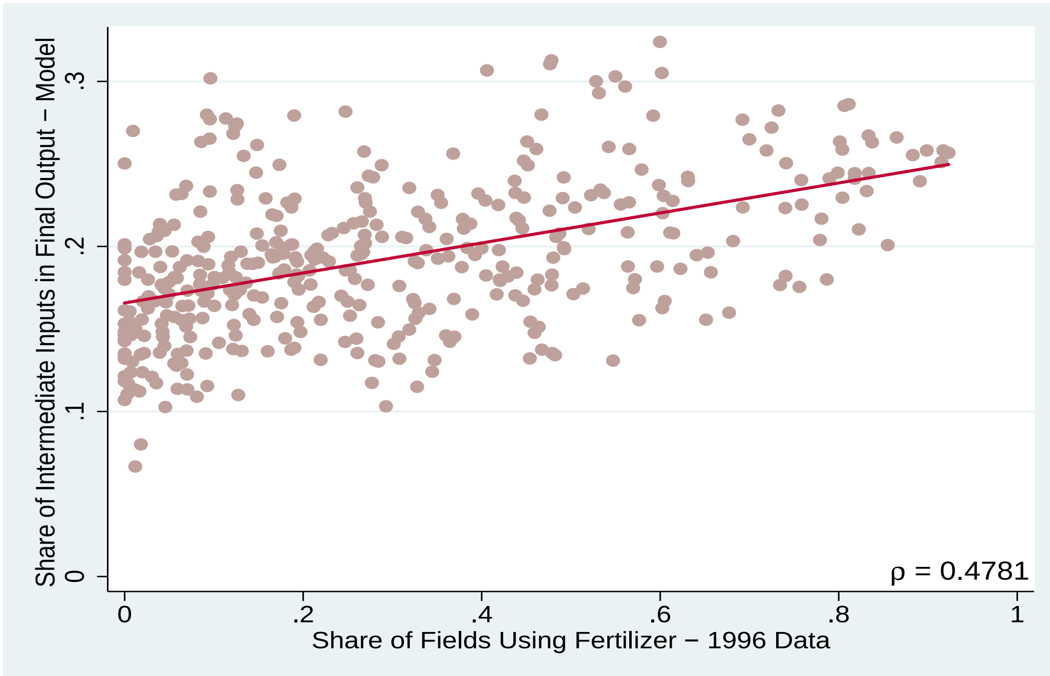
Table 4: Calibrated Benchmark Economy (BE)

Statistic	Description	Value in BE
Y_a/L	Real Aggregate Yield	1676.9
Y_f/L_f	Yield in Food Crops	1591.1
Y_s/L_s	Yield in Cash Crops	2122.8
C_a/L	Real Net Aggregate Yield	1338.4
C_f/L_f	Net Yield in Food Crops	1304.2
C_s/L_s	Net Yield in Cash Crops	1515.8
VA_a/L	Real Value Added Yield	1653.2
VA_a/N_a	Real Value Added per worker	1782.8
N_a/N	Share of Employment in Agriculture (%)	86.0
L_f/L	Total Share of land in food (%)	83.9
X_f/Y_f	Intermediate Input Intensity (%)	31.1
p_f	Consumer price of food	1.00
L/N_a	Average Farm Size	1.26
GDP/N	Real GDP per Worker	1357.4
	Micro-level Statistics	
	STD of log–Food Yield	0.59
	STD of log–Aggregate Yield	0.60
	CORR of log–(Food Yield, Trans. Costs)	-0.14
	CORR of log–(Aggregate Yield, Trans. Costs)	-0.14

6 Quantitative Experiment

The main experiment involves studying the effects from reducing geographic transport costs across all woredas from their actual 1996 levels to their actual 2014 levels. In order to isolate the effects of transport cost changes alone I keep all other parameters to their 1996 levels. In other words, I ask what would be the aggregate and spatial micro-level effects on the Ethiopian economy if the only change between 1996 and 2014 had been the change in the transportation network and the associated changes in transportation costs? I then compare these changes to the actual changes in the variables of interest that occurred in the data over the same period. The model allows me to assess directly the effects of transport cost changes irrespective of the other changes that may have

Figure 7: Intermediate Input Use by Woreda: Model vs. 1996 Data



Notes: "Model" refers to the value in the calibrated benchmark economy. "1996 Data" refers to the woreda value from the 1996 pooled data from the AgSS (CSA) in Ethiopia.

occurred over the same period and which may have also contributed to changes in the variables of interest.

The iceberg transport costs for 2014 are obtained from the same transport cost function as above,

$$\tau_{ij,2014} = 1 + \psi_i \cdot (tt_{ij,2014})^{\eta}$$

for $i \in \{f, s, x\}$, where the travel times for 2014, $tt_{ij,2014}$, are the ones estimated in Section 2 from the road infrastructure network available in 2014. The associated changes in transport costs change the connectivity of woredas with markets, and do so in a heterogeneous fashion, since the volume and quality of the road network did not expand to all woredas at the same rate. As a result there is a change in both the level and the dispersion of transport costs across woredas. Keeping all other parameters (including productivity) in all rural and urban locations to their benchmark economy levels, I feed the 2014 iceberg costs $\{\tau_{fj,2014}, \tau_{sj,2014}, \tau_{xj,2014}\}_{j=1}^{J}$ into the model and solve for the new equilibrium.

6.1 Aggregate Effects

The aggregate results for the new equilibrium are presented in the second column of Table 5, and the changes relative to the 1996 benchmark economy in the third column. The first column repeats the benchmark economy values for completeness. I note that in the new equilibrium I aggregate across crops and goods using a common set of prices before and after the change in transport costs, just as statistical agencies measure "real" changes. The common set of prices I use are the ones from the benchmark economy net of transport costs.

There are substantial aggregate effects when transport costs are reduced to their 2014 levels. Productivity increases substantially as captured by several measures in Table 5. The aggregate economy-wide real yield in agriculture, measured as the real value of final output per unit of land, increases by 13.6%. This is achieved through an increase in within-crop real yields, 11.4% for food and 5.6% for cash crops, as well as a reallocation of land from food crop production to cash crop production in the economy overall. The share of land in food production in the economy drops from almost 84% in 1996 to under 72% after the transport cost changes.

The within-crop increase in the yield is accomplished through the reallocation of production across space, according to their comparative advantage. In particular, there is a given amount of food that has to be produced to meet the subsistence needs of the population $\bar{f}$. In the model comparative advantage across woredas in determined not only by relatively actual TFP (z_{ij}) but also by relative transportation costs (τ_{ij}). A high productivity region that faces a high transport cost may have a low "effective" productivity z_{ij}/τ_{ij}, producing a small share of the economy's food. When transport costs fall, connectivity to markets increases across woredas. When transport frictions are lower and trade of goods is "freer" food production shifts to relatively higher TFP woredas, with the freed up land being reallocated to cash crop production.

So food production is undertaken increasingly by more productive woredas. Note that because of the decreasing returns to scale all woredas produce some food, which eliminates the case of complete specialization in food by any woreda. But because they are more productive they do not need to devote as many resources to producing food. Given that labor is used only in food production, when land allocated to food falls, the amount of labor needed to produce that same amount of food also goes down, being now reallocated to non-agricultural production in the urban centre. The share of labor in agriculture in the model drops from 86% in the benchmark economy to 76.3%, a drop of almost ten percentage points.

With the drop in the overall engagement in food production, the demand for imported intermediate inputs falls. The overall share of intermediate inputs in the economy increases by 1.2 percentage points. This is because the relative cost of intermediate inputs depends not only on the domestic transport costs of those inputs to woredas but also on the price of food. While transport costs for intermediate inputs fall, the relative price of food also falls by 9%.

Note that the iceberg transport costs are resource costs that show up in the model as "melting" of goods in transit. As a result part of the output of each crop (food and cash) constitute payments to the transportation sector. The net amount of output of crop i delivered to the urban location (for consumption in the case of food, and for export in the case of cash crops) is y_{ij}/τ_{ij}. The real net yield, that nets out transport costs could be argued is the more appropriate measure of real productivity, because it also takes into account the direct resource savings from lower transport costs. The real net yield in the model increases by almost 18% when transport costs drop to their 2014 levels. Again this is the result of within-crop net yield increases (16.9% for food crops and 12.6% for cash crops) as well as reallocation of land to cash crops at the economy level. The indirect productivity gains achieved through the mechanism of the model account for 78% of the overall gains ($\log(1.136)/\log(1.178)$), implying that the direct savings from lower transport costs are the residual 22%.

Real value added per unit of land overall in agriculture increases 16.5%. Real value added is computed using a common set of prices not only for crops but also for intermediate inputs. Real value added per worker in agriculture increases by 31% due to the substantial drop in agricultural labor. The economy-wide agricultural land to labor ratio, or average farm size, increases by 12.7%. A simple decomposition of the total aggregate gain in value added per worker, as shown in (14), reveals that the real value added yield gain accounts for 56% (= $\log(1.165)/\log(1.313)$), while the average farm size gains for the remaining 44% (= $\log(1.127)/\log(1.313)$).

$$\underbrace{\frac{VA_a}{N_a}}_{1.313} = \underbrace{\frac{VA_a}{L}}_{1.165} \cdot \underbrace{\frac{L}{N_a}}_{1.127}, \tag{14}$$

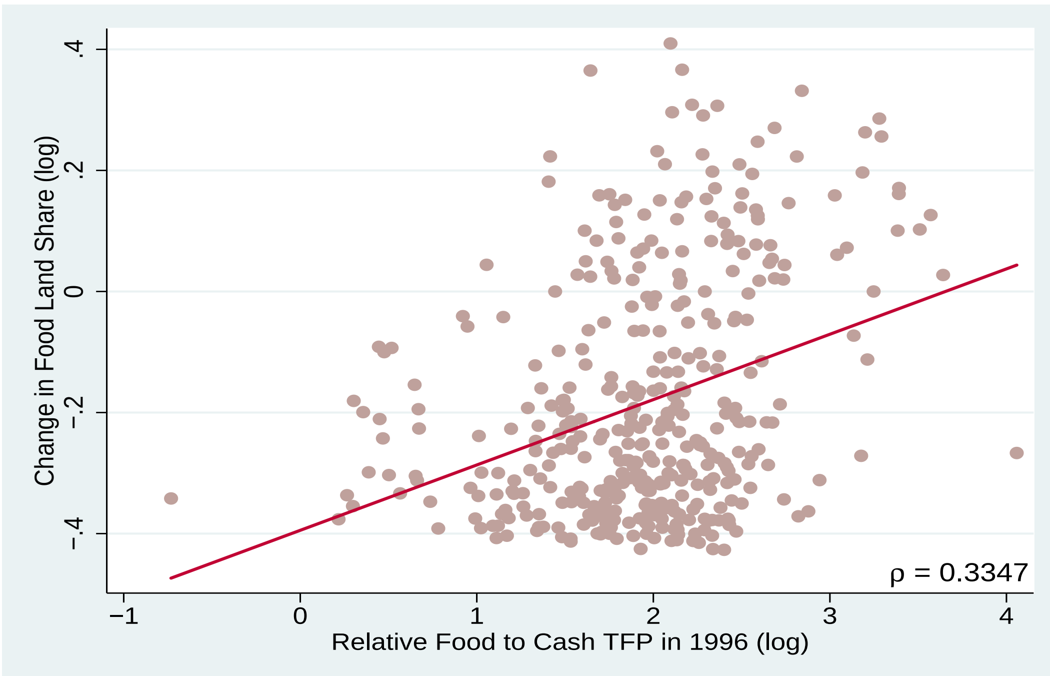
Real GDP per worker in the economy that also takes into account the output of the non-agricultural sector in the urban center increases by 28.7%. The reason the increase in GDP per worker is lower than the value added per worker in agriculture is that in the calibrated model, the level of the implied productivity of the non-agricultural sector is lower than in agriculture.

6.2 Spatial Distribution of Effects

While the economy-wide aggregate gains capture the overall effect of changes in transportation costs because they take into account the gains from the spatial reallocation of production, it is also important to understand the spatial patterns across woredas that the changes in transport costs impart. In this section I examine the spatial distributional consequences of the change in transport costs to their 2014 levels, following the expansion of the volume and quality of the road network across Ethiopia. The micro-level spatial patterns help shed light on the mechanism of the model.

In Figure 8 I plot the within-woreda change in the share of land allocated to food production, implied by the model after the fall in transport costs, against the ratio of productivities (TFP) between the food and cash crop technologies (both in logs). The positive relationship indicates that there was a reallocation of land towards food production, following the transport cost changes, in woredas that were relatively more productive in food crops in 1996. As woredas become more integrated with markets, and trade becomes freer, there is a reallocation of production according to comparative advantage.

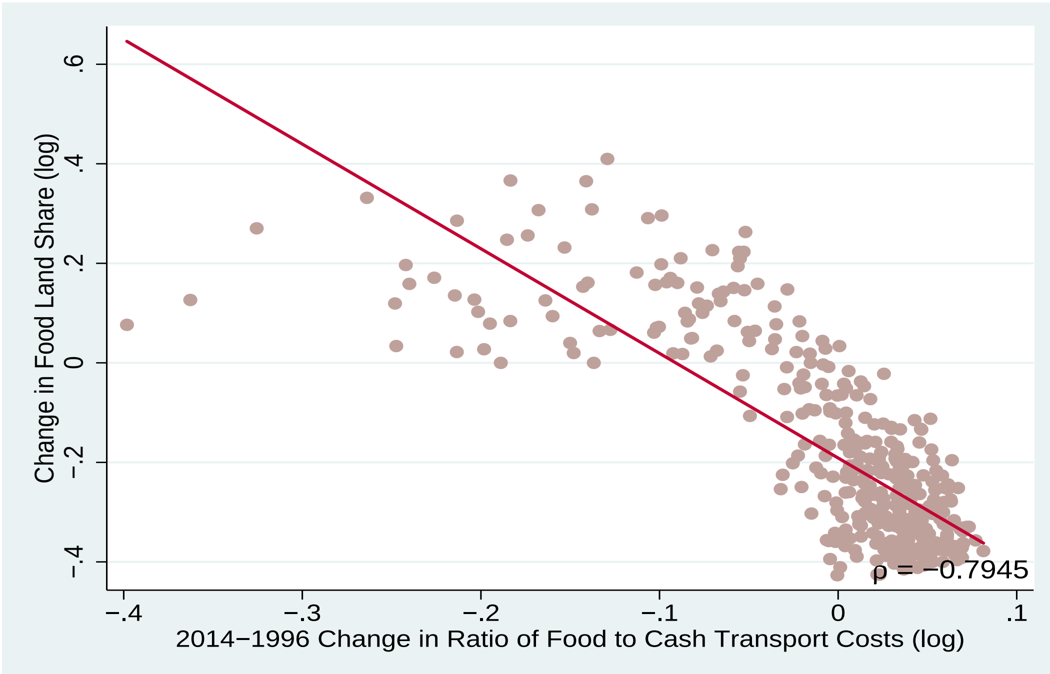
Figure 8: Change in Food Land Share (model) vs. Relative TFP in 1996



Notes: "Food Land Share" refers to the share of land within each woreda allocated to food production. The "Change" in this share refers to the change in the model after the reduction in transport costs. "Relative TFP" refers to the ratio of TFP in food production relative to TFP in cash crop production in the benchmark economy.

Land reallocations across crops respond not only to relative productivity but also to relative transport costs. In particular, for a given woreda, an increase in the transport cost for food relative to the transport cost for cash crops, would tend, all else equal, to reallocate land from food to cash crops in that woreda. In Figure 9 I plot the log-change in the woreda-level share of land engaged in food production against the log-change in the ratio of food to cash crop transport costs over the period 1996-2014. A positive value for the change in the transport cost ratio means that the transport cost of food increased relative to that of cash crops. The model predicts a strong negative

Figure 9: Change in Food Land Share (model) vs. Relative Transport Cost Changes



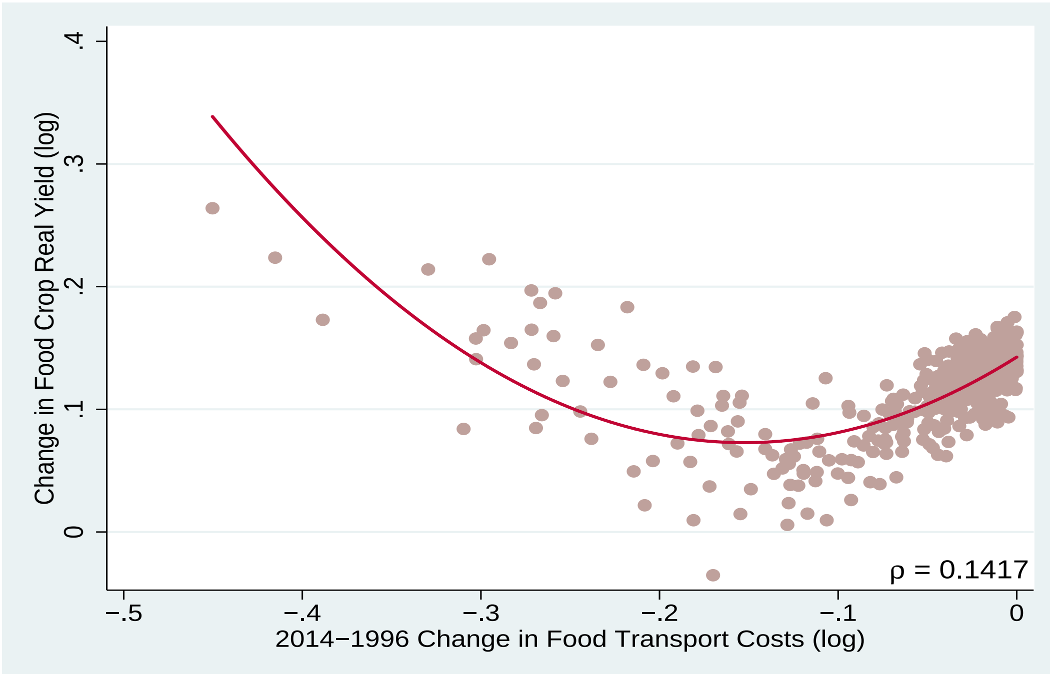
Notes: "Food Land Share" refers to the share of land within each woreda allocated to food production. The "Change" in this share refers to the change in the model after the reduction in transport costs. "Ratio of Food to Cash Transport Costs" refers to the woreda-level ratio. The x-axis involves the log-change of this ratio over 1996-2014.

relationship between changes in food land shares and changes in relative transport costs, with locations facing relatively higher food transport costs shifting away from food crop production. The largest increases in the food land share are in woredas where the relative transport costs dropped the most. In woredas where the relative food transport costs increased, there was a shift of land away from food production.

When the general level of transport costs decreases in the model, there is an increase in "effective productivity" in each woreda, which raises the rental price of land relative to labor (where the latter here is anchored by non-agricultural labor productivity). As land becomes relatively more expensive there is an increase in the labor-land ratio in food production. In addition, with lower transport costs for intermediate inputs, other things equal, there tends to be an increase in the intermediate input intensity under the food producing technology. These changes would tend to increase the food crop yield. As explained above however the allocation of land across crops within woredas depends on the food to cash crop relative transport costs. In woredas where the relative food transport costs also drop there tends to be an increase in the land allocated to food, whereas in woredas where relative food transport costs increase the land allocated to food tends to fall. Given that there are decreasing returns to land, in locations where land allocated to food production increases, the food crop yield would tend to fall, whereas in locations that shift away from food production the opposite will be true. As can be seen from equation (15) The overall effect on the food crop yield depends on the interaction of the decreasing returns to land effect with the labor-land ratio and intermediate input intensity effect. In locations that shift away from food production (those that experience increases of relative food transport costs or small drops) all the effects move in the same direction and the food crop yield would tend to increase in these locations. In locations with large drops in relative food transport costs, and substantial shifts of land to food production, the decreasing returns effect dominates the positive labor-land ratio and intermediate input use effects, and the food crop yield would decrease in these locations. The heterogeneous changes in transport costs and the heterogeneous responses of woredas imply a spatial U-shaped pattern of the food crop yield with transportation costs. This pattern is illustrated in Figure 10, that plots log changes in the real yield for food crops against log changes in the level of transport costs (for food). Note that because in the model the cash crop production technology is linear in land, there are no within-woreda changes to real yield in cash crops within woredas. As a result, the U-shaped pattern of the food crop yield carries over to the woreda-level aggregate yield. Figure 11 shows the U-shaped relationship between changes in the woreda-level aggregate yield and changes in transport costs.

$$\frac{y_{fj}}{\ell_{fj}} = z_{fj}^{1-\gamma} \left(\frac{n_j}{\ell_{fj}}\right)^{\alpha\gamma} \ell_{fj}^{\gamma-1} \left(\frac{x_j}{y_{fj}}\right)^{\frac{1-\theta}{\theta}} \tag{15}$$

Figure 10: Change in Food Crop Yield (model) vs. Transport Cost Changes



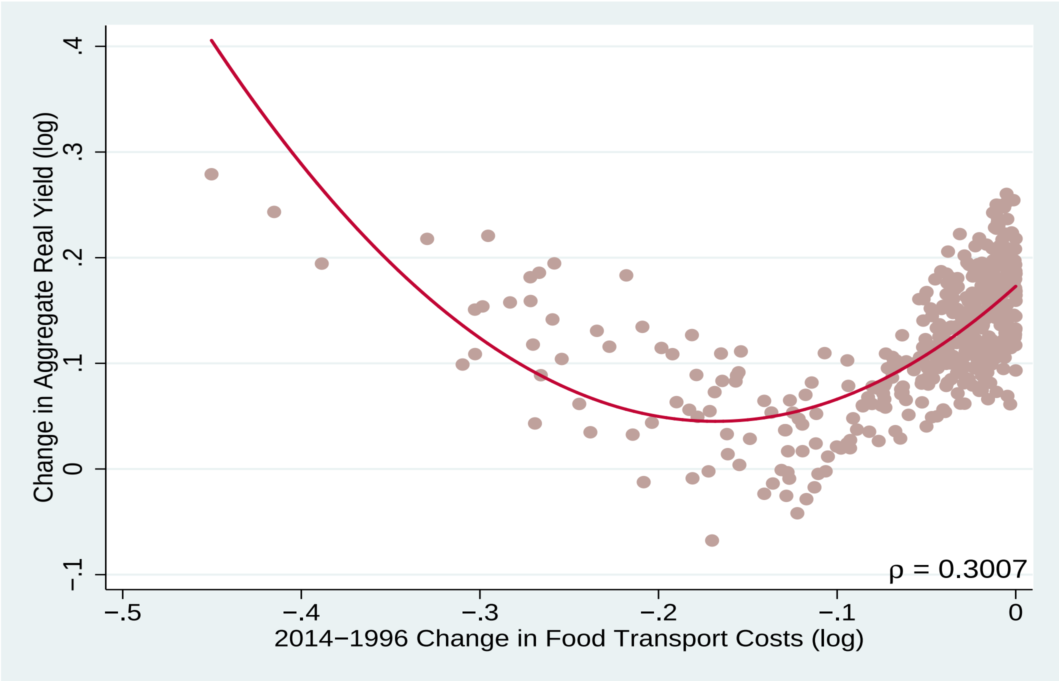
Notes: "Change in Food Crop Yield" refers to the change in the real economy-wide yield in food production in the model, after the reduction in transport costs. The x-axis involves the log-change of food transport costs over 1996-2014.

6.3 Comparison to Data Changes

In the above quantitative experiment the only object changed relative to the benchmark economy was the matrix of good-woreda-specific transportation costs. It is of interest to see how the changes in the allocation induced by the transport costs alone compare to the actual changes observed in the Ethiopian economy over the period 1996-2014.

Table 6 compares the aggregate changes from the model (first column) to the ones in the data (second column) for key variables of interest. The real aggregate yield of final output increases 13.6% in the model increases, accounting for 8.6% ($\log(1.136)/\log(4.419)$) of the overall increase in the data 341.9%. If however we include the direct resource savings from the transport cost reductions and consider the aggregate net yield, which increases by 17.8% then the yield gain in the model accounts for 11% of the overall yield gain in the AgSS data. In other words, the model

Figure 11: Change in Aggregate Yield (model) vs. Transport Cost Changes



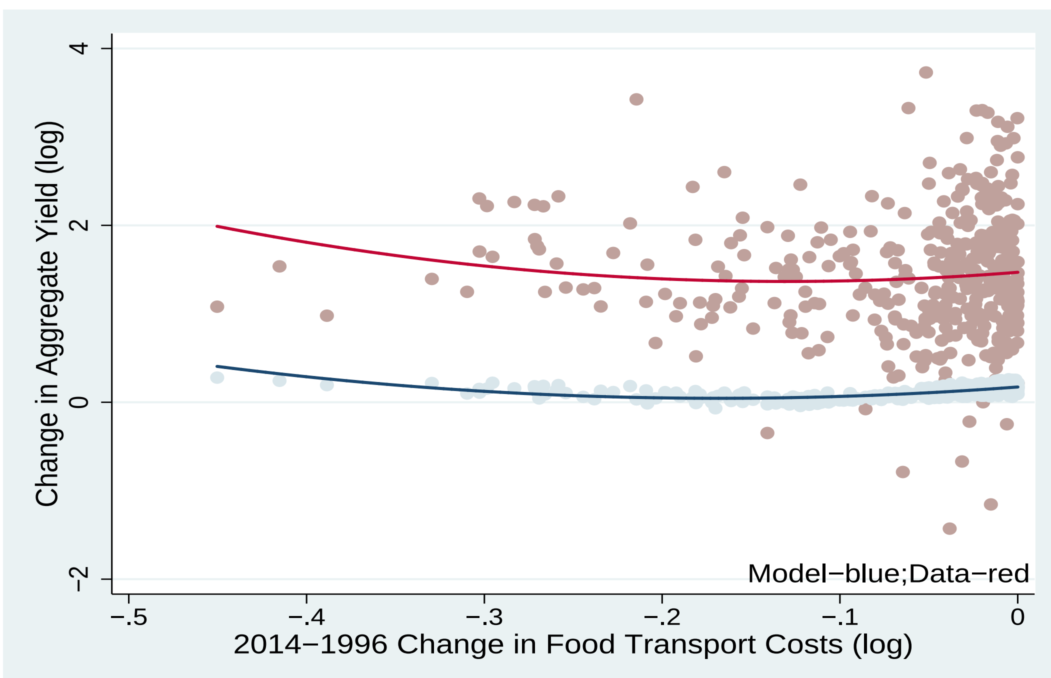
Notes: "Change in Aggregate Yield" refers to the change in the real economy-wide aggregate yield (valued at a common set of prices) in the model, after the reduction in transport costs. The x-axis involves the log-change of food transport costs over 1996-2014.

with only transport cost changes can account for about 1/10 of the yield gains in the data. The gross yield in food crops in the model accounts for 11.6% of the one observed in the AgSS data.

In terms of other statistics, the increase in the share of land allocated to cash crops and the drop in the share of labor in agriculture are in the neighborhood of these changes in the data over the period 1996-2014. Average farm size in the model increases almost double what it does in the. This "overshooting" is partly due to the fact that there is no labor in the cash crop technology, and thus as the economy shifts away from food, all the extra labor finds its way to non-agriculture, increasing average farms size. Finally, the model also generates an increase in GDP per worker that is about half of the actual change in the data.

Next, I compare the spatial pattern of the yield gains produced by the model with the ones in the data. Figure 12 compares (log) changes in the aggregate woreda-level yield in the AgSS data to (log) changes in the aggregate woreda-level yield implied by the model, against the actual changes in transport costs over the period 1996-2014. While in the data there is more noise and the magnitude of the changes in the woreda-level yields are higher than those produced by the model with only changes in transport costs, the U-shaped pattern of the woreda-level gains with transport costs changes is present in both the data and the model. In Figure 13 I show the same relationship but only for the woredas in which there was a decrease in the *relative* food to cash crop transport costs.[12] For these woredas there is a negative relationship between aggregate yield gains and transport cost reductions, implying that on average the woredas that gain the most are ones that experience the largest drops in their transport costs.

Figure 12: Model vs. Data: Aggregate Yield - Transport Cost Relationship (All woredas)

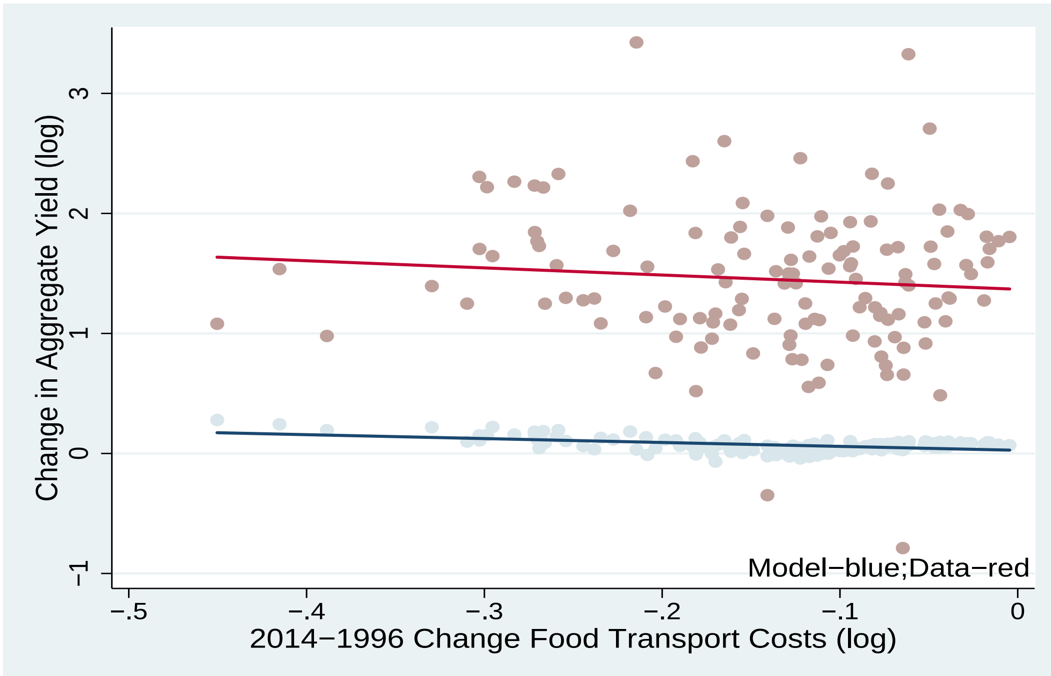


Notes: "Change in Aggregate Yield" refers to the change in the real economy-wide aggregate yield (valued at a common set of prices). In the model this is the change relative to the benchmark economy after the reduction in transport costs. In the AgSS data this is the actual change over 1996-2014. The x-axis involves the log-change of food transport costs over 1996-2014.

Figure 14 shows that the level of the aggregate yield by woreda in the model after the drop in

[12]Note that food transport costs fall in all woredas but relative food to cash crop transport costs do not.

Figure 13: Model vs. Data: Aggregate Yield - Transport Cost Relationship (Largest drop Woredas)



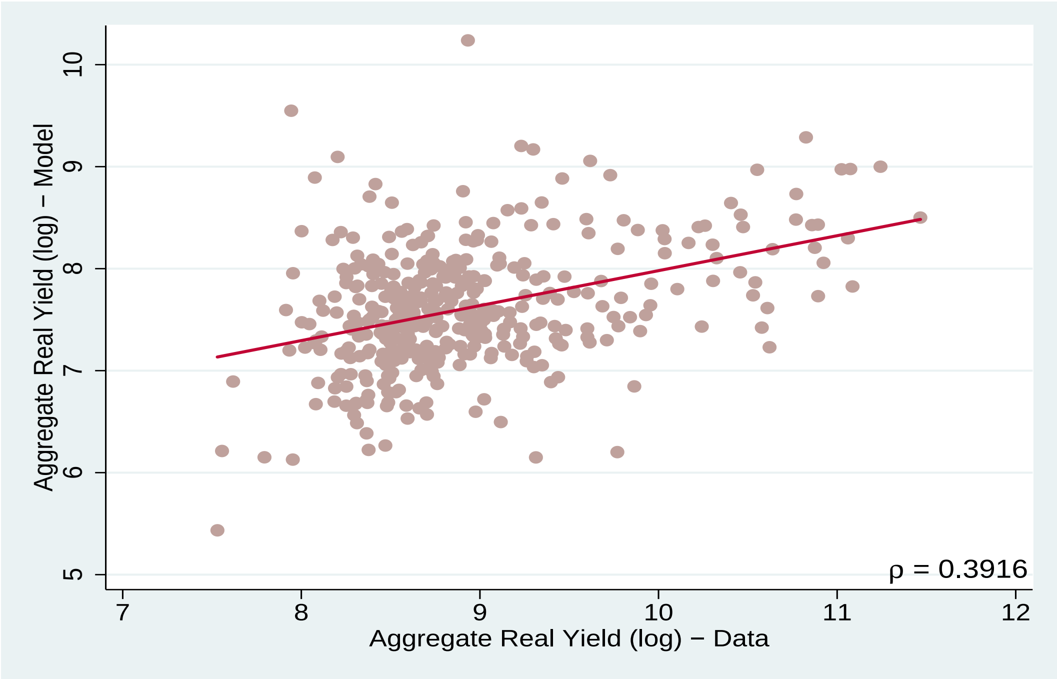
Notes: "Change in Aggregate Yield" refers to the change in the real economy-wide aggregate yield (valued at a common set of prices). In the model this is the change relative to the benchmark economy after the reduction in transport costs. In the AgSS data this is the actual change over 1996-2014. The x-axis involves the log-change of food transport costs over 1996-2014.

transport costs is highly correlated with the aggregate yield in the 2014 AgSS data (in logs).

7 Conclusions

This paper has studied a particular episode of a large-scale infrastructure project, undertaken in Ethiopia starting in 1997. To measure the effects of the road expansion program I combined a quantitative spatial framework with novel panel data on agricultural production and transportation costs. I find that the changes in transport costs implied by the expansion of the road network have had a sizable impact on productivity and the structure of the agricultural sector in Ethiopia. The gains in real output per worker are about 1/10 of the overall gains observed in the data. I also find that with "closer" markets there is an overall shift of agricultural production from food to

Figure 14: Model vs. Data: Aggregate Yield (log)



Notes: "Aggregate Yield" is the real economy-wide aggregate yield (valued at a common set of prices). In the model this is the value after the reduction in transport costs. In the AgSS data this is the actual value in 2014.

cash crops, and a drop in the share of labor engaged in crop production. These effects are sizable and in the neighborhood of what occurs in the data. However, at the individual district level the gains are not uniform. The model produces a U-shaped relationship between district-level gains and transport cost changes, that is similar in nature to the corresponding one in the data. This is not only because changes in relative transport costs are uneven across districts, but also because districts are inherently different along other dimensions. Finally, I note that while the drops in transport costs have been large, Ethiopia started from a very high base, and their level still remains high. The implication of the analysis here is that, further investments in infrastructure expansion, can have real productivity benefits for the economy.

References

Adamopoulos, T. (2011). Transportation costs, agricultural productivity, and cross-country income differences. *International Economic Review*, 52(2):489–521.

Adamopoulos, T. and Restuccia, D. (2014). The size distribution of farms and international productivity differences. *American Economic Review*, 104(6):1667–97.

Adamopoulos, T. and Restuccia, D. (2018). Geography and agricultural productivity: Cross-country evidence from micro plot-level data. Technical report, National Bureau of Economic Research.

Alder, S. (2018). Chinese roads in india: The effect of transport infrastructure on economic development. Unpublished manuscript, Univ. North Carolina, Chapel Hill.

Allen, T. and Arkolakis, C. (2014). Trade and the topography of the spatial economy. *The Quarterly Journal of Economics*, 129(3):1085–1140.

Allen, T. and Arkolakis, C. (2016). The welfare effects of transportation infrastructure improvements. *Manuscript, Dartmouth and Yale*.

Asher, S. and Novosad, P. (2018). Rural roads and local economic development.

Asturias, J., García-Santana, M., and Ramos Magdaleno, R. (2016). Competition and the welfare gains from transportation infrastructure: Evidence from the golden quadrilateral of india.

Banerjee, A., Duflo, E., and Qian, N. (2012). On the road: Access to transportation infrastructure and economic growth in china. Technical report, National Bureau of Economic Research.

Baum-Snow, N., Brandt, L., Henderson, J. V., Turner, M. A., and Zhang, Q. (2017). Roads, railroads, and decentralization of chinese cities. *Review of Economics and Statistics*, 99(3):435–448.

Brooks, W., Donovan, K., et al. (2017). Eliminating uncertainty in market access: The impact of new bridges in rural nicaragua. In *Understanding Productivity Growth in Agriculture*. University of Chicago Press.

Caselli, F. (2005). Accounting for cross-country income differences. *Handbook of economic growth*, 1:679–741.

Combes, P.-P. and Lafourcade, M. (2005). Transport costs: Measures, determinants, and regional policy implications for france. *Journal of Economic Geography*, 5(3):319–349.

Costinot, A. and Donaldson, D. (2016). How large are the gains from economic integration? theory and evidence from us agriculture, 1880-1997. Technical report, National Bureau of Economic Research.

Donaldson, D. (2018). Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and american economic growth: A market access approach. *The Quarterly Journal of Economics*, 131(2):799–858.

Donovan, K. (2016). Agricultural risk, intermediate inputs, and cross-country productivity differences. *Unpublished Working Paper, University of Notre Dame.*

Faber, B. (2014). Trade integration, market size, and industrialization: evidence from china's national trunk highway system. *Review of Economic Studies*, 81(3):1046–1070.

Fajgelbaum, P. D. and Schaal, E. (2017). Optimal transport networks in spatial equilibrium. Technical report, National Bureau of Economic Research.

Felbermayr, G. J. and Tarasov, A. (2015). Trade and the spatial distribution of transport infrastructure.

Gabre-Madhin, E. Z. (2001). Market institutions, transaction costs, and social capital in the ethiopian grain market. Technical report, Research Report 124, International Food Policy Research Institute (IFPRI), Washington D.C.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). Agricultural Productivity Differences across Countries. *American Economic Review*, 104(5):165–70.

Gollin, D., Parente, S., and Rogerson, R. (2002). The Role of Agriculture in Development. *American Economic Review*, 92(2):160–164.

Gollin, D. and Rogerson, R. (2014). Productivity, transport costs and subsistence agriculture. *Journal of Development Economics*, 107:38–48.

Herrendorf, B., Schmitz, Jr, J. A., and Teixeira, A. (2012). The role of transportation in us economic development: 1840–1860. *International Economic Review*, 53(3):693–716.

Lagakos, D. and Waugh, M. E. (2013). Selection, agriculture, and cross-country productivity differences. *The American Economic Review*, 103(2):948–980.

Minten, B., Koru, B., and Stifel, D. (2013). The last mile (s) in modern input distribution: Pricing, profitability, and adoption. *Agricultural Economics*, 44(6):629–646.

Prasada Rao, D. (1993). International comparisons of agricultural output and productivity. *FAO Economic and Social Development Paper*, (112).

Restuccia, D. and Santaeulalia-Llopis, R. (2015). Land misallocation and productivity. Unpublished manuscript, University of Toronto.

Restuccia, D., Yang, D. T., and Zhu, X. (2008). Agriculture and aggregate productivity: A quantitative cross-country analysis. *Journal of Monetary Economics*, 55(2):234–250.

Schultz, T. W. (1953). *The economic organisation of agriculture*. McGraw-Hill.

Sotelo, S. (2018). Domestic Trade Frictions and Agriculture,. Unpublished manuscript, University of Michigan.

Storeygard, A. (2016). Farther on down the road: transport costs, trade and urban growth in sub-saharan africa. *The Review of economic studies*, 83(3):1263–1295.

Timmer, M., de Vries, G. J., and de Vries, K. (2015). Patterns of structural change in developing countries. In *Routledge handbook of industry and development*, pages 79–97. Routledge.

Tombe, T. (2015). The missing food problem: Trade, agriculture, and international productivity differences. *American Economic Journal: Macroeconomics*, 7(3):226–58.

Valentinyi, A. and Herrendorf, B. (2008). Measuring Factor Income Shares at the Sector Level. *Review of Economic Dynamics*, 11(4):820–835.

Warner, J., Stehulak, T., and Kasa, L. (2015). Woreda-level crop production rankings in ethiopia: A pooled data approach. *Addis Ababa, Ethiopia*.

Table 5: Effects of Reducing Transport Costs to 2014 Levels

Statistic	Benchmark Economy BE	2014 Transport Costs	Percentage Change (%)
Aggregate Statistics			
Real Aggregate Yield (Y_a/L)	1676.9	1904.5	13.6
Yield in Food Crops (Y_f/L_f)	1591.1	1772.0	11.4
Yield in Cash Crops (Y_s/L_s)	2122.8	2240.6	5.6
Real Net Aggregate Yield (C_a/L)	1338.4	1576.2	17.8
Net Yield in Food Crops (C_f/L_f)	1304.2	1525.0	16.9
Net Yield in Cash Crops (C_s/L_s)	1515.8	1706.1	12.6
Real Value Added Yield (VA_a/L)	1653.2	1419.1	16.5
Real Value Added per worker (VA_a/N_a)	1782.8	2341.2	31.3
Share of Employment in Agriculture (N_a/N) (%)	86.0	76.3	-9.7
Total Share of land in food (L_f/L) (%)	83.9	71.7	-12.1
Intermediate Input Intensity (X_f/Y_f) (%)	31.1	32.3	1.2
Consumer price of food (p_f)	1.00	0.91	-9.0
Average Farm Size (L/N_a)	1.26	1.42	12.7
Real GDP per Worker (GDP/N)	1357.4	1055.0	28.7
Woreda-level Statistics			
STD of log–Food Yield	0.59	0.60	–
STD of log–Aggregate Yield	0.60	0.60	–
CORR of log–(Food Yield, Trans. Costs)	-0.14	-0.10	–
CORR of log–(Aggregate Yield, Trans. Costs)	-0.14	-0.13	–

Notes: The column "Benchmark Economy (BE)" displays the values of the each variable in the baseline calibrated economy. The column "2014 Transport Costs" displays the values of each variable when transport costs are reduced to their 2014 levels in the benchmark economy. The percentage changes in the counterfactual economy (with reduced transport costs) relative to the benchmark are in the last column. All aggregate variables, except for those reported in percentages, are reported as ratio of the counterfactual to the benchmark economy. For variables reported in percentages, the last column displays the difference between the pre- and post- transport costs change. "Woreda-level statistics" are reported in levels, and no percentage changes are reported.

Table 6: Comparison of Model and Data Changes (Aggregate Statistics)

Statistic	Changes due to Transport Cost Reductions (%)	Changes in Data Over 1996-2014 (%)
Real Aggregate Yield	13.6	341.9
Yield in Food Crops	11.4	153.9
Yield in Cash Crops	5.6	830.6
Total Share of land in food (change in %)	-12.1	-8.5
Average Farm Size	12.7	6.2
Share of Employment in Agriculture (change in %)	-9.7	- 12.6
Real GDP per Worker	28.7	67.1

Notes: The first column shows changes relative to the benchmark economy, implied by the model, when all transport costs are reduced to their 2014 levels. All the changes in the data are computed from the AgSS (CSA) data over 1996-2014, with the exception of the "Share of Employment in Agriculture" and "Real GDP per Worker" values which are computed from the GGDC 10-sector database as changes over 1996-2011.

The International Growth Centre (IGC) aims to promote sustainable growth in developing countries by providing demand-led policy advice based on frontier research.

Find out more about our work on our website
www.theigc.org

For media or communications enquiries, please contact
mail@theigc.org

Subscribe to our newsletter and topic updates
www.theigc.org/newsletter

Follow us on Twitter
@the_igc

Contact us
International Growth Centre,
London School of Economic and Political Science,
Houghton Street,
London WC2A 2AE

IGC
International Growth Centre

DIRECTED BY
LSE
UNIVERSITY OF OXFORD

FUNDED BY
UKaid from the British people

Designed by soapbox.co.uk