Port Rail Connectivity and Agricultural Production

Evidence from a Large Sample of Farmers in Ethiopia

Atsushi Iimi
Haileyesus Adamtei
James Markland
Eyasu Tsehaye

World Bank Group
Transport and ICT Global Practice Group
June 2017
Abstract

Agriculture remains an important economic sector in Africa, employing a large share of the labor force and earning foreign exchange. Among others, transport connectivity has long been a crucial constraint in Africa. In theory, railways have a particularly important role to play in shipping freight and passengers at low cost. However, most African railways were in virtual bankruptcy by the 1990s. Using a large sample of data comprised of more than 190,000 households over eight years in Ethiopia, the paper estimates the impacts of rail transport on agricultural production. Methodologically, the paper takes advantage of the historical event that a major rail line connecting the country to the regional hub, the Port of Djibouti, was abandoned in the 2000s. With spatially highly disaggregated fixed effects and instrumental variables incorporated, an agricultural production function is estimated. The elasticity with respect to port connectivity is estimated at 0.276. The use of fertilizer is also found to increase with transport cost reduction, supporting the fact that a large amount of fertilizer is imported to Ethiopia.

This paper is a product of the Transport and ICT Global Practice Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The authors may be contacted at aiimi@worldbank.org.
Port Rail Connectivity and Agricultural Production: Evidence from a Large Sample of Farmers in Ethiopia

Atsushi Iimi, Haileyesus Adamtei, James Markland, Eyasu Tsehaye

Transport & ICT Global Practice
The World Bank Group

Key words: Agriculture production; transport infrastructure.
JEL classification: Q12, R40, C26, C23.

* Corresponding author.
I. INTRODUCTION

Agriculture remains an important economic sector in Africa. In many countries, more than half of the total population is still engaged in agricultural production. Agriculture and agrobusiness are currently estimated to contribute to US$31 billion or nearly half of the GDP of the region. This is projected to continue growing to US$1 trillion by 2030 (World Bank 2013). However, Africa’s agriculture remains small-scale farming with few advanced inputs used. In Zambia, farmers use 25 kg of nitrogen per hectare, about 20 percent of the recommended amount of fertilizer, although the farmers’ fertilizer consumption has been increasing gradually (Xu et al., 2009). In Mali, most farmers still do not use irrigation, although even small-scale irrigation can increase productivity and household income dramatically (Dillon, 2011). As a result, the vast majority of farmers live below the poverty line in Africa.

Among others, transport connectivity has long been a crucial constraint. In Africa, rural accessibility—measured by the proportion of the rural residents who live within a 2-km walking distance from an all-weather road—is estimated at less than 30 percent (Gwilliam 2011). The literature shows that regardless of mode, improved transport infrastructure can provide better market access to farmers, stimulating agricultural production. In Bangladesh, improved rural roads lowered agricultural input prices and increased output prices, thereby boosting agricultural production (Khandker, Bakht and Koolwal, 2009). Donaldson (2016) uses historical panel data in colonial India and shows that the railroad network developed during the colonial era reduced transport costs to boost agricultural trade.

In Africa, railways have a particularly important role to play in shipping freight and passengers at low cost. Since the later 1890s, many European powers developed rail links to connect fertile and resource-rich inland areas to seaports (e.g., Amin et al. 1986; Jedwab and

---

1 See Spielman et al. (2012) for detailed discussion on various institutional constraints to promote the use of advanced inputs in Ethiopia.
Moradi, 2012). In theory, rail transport still has the advantage for long-haul freight shipment. However, most African railways were in virtual bankruptcy by 1990 (e.g., Olievschi 2013). Despite some localized improvements in recent years, they remain largely nonoperational. This must have significant negative impacts on African economies, particularly landlocked countries. In the region, there are 15 landlocked countries.

For landlocked countries, transport and trade costs are prohibitively high in Africa. For instance, for Malawi, which is a landlocked country in Southern Africa, the cost of importing a 20-foot container of goods is US$2,895. In addition, it takes about 39 days. Both are unfavorably compared with a regional gateway country, Tanzania (US$1,615 and 26 days, respectively). Similarly, for Ethiopia, another landlocked country, the importing cost is as high as US$2,960. The cost for its neighboring country, Djibouti, with a regional hub port is one-third of it (US$910). Thus, having good access to regional ports is critical to maintain competitiveness of the economy. Rail transport is often an important option for long-haul freight shipments from and to the ports.

The current paper aims at estimating the impact of railway connectivity on agricultural production in Ethiopia, a landlocked country in East Africa. As well known in the literature (e.g., Banerjee et al. 2012; Datta, 2012), it is a challenge to estimate an unbiased impact of large-scale infrastructure, such as railroads, highways and ports. Endogeneity is a matter of main concern. While transport investment is believed to raise the productivity of the economy, transport infrastructure is also often placed where economic productivity is inherently high. Thus, regardless of its real impact, there is likely to be positive correlation between transport investment and economic outcomes. By nature, randomized control trial is barely possible in the transport infrastructure context. It is a typical network industry.

The current paper’s contributions are twofold: First, it generates historical transport connectivity estimates in Ethiopia, using the Ethiopian Roads Authority’s road network data. It will be shown that port connectivity, in particular, varies across locations and deteriorated considerably in the late 2000s, when a main regional railway ceased operating. Second, the
paper provides an unbiased impact of port connectivity on agricultural production, using a very large sample of data, comprised of over 190,000 households from 2003 to 2010. Despite the relatively rich literature on Ethiopian agriculture (e.g., Spielman et al. 2012; Taffesse et al. 2012; Mekonnen et al. 2013; Rachid et al. 2013), there are few studies that explicitly examine the impacts of regional transport connectivity. The endogeneity problem is addressed by combining the highly disaggregated location-specific fixed-effects and new instrumental variables (IV).

The remaining sections are organized as follows: Section II provides a brief overview of recent developments in the agriculture and transport sectors in Ethiopia. Section III discusses our empirical strategy and describes our data. Section IV presents the main estimation results, and Section V discusses robustness and some policy implications in the transport sector. Then, Section VI concludes.

II. OVERVIEW OF AGRICULTURE AND PORT ACCESS IN ETHIOPIA

In Ethiopia, agriculture remains among the most important economic sectors. It produces about one-third of GDP and employs 70 percent of the workforce, accounting for 80 percent of the country’s merchandise exports (Ethiopian Agricultural Transformation Agency, 2014). Ethiopia produces about US$6 billion in crops a year, including maize, sorghum and wheat (Figure 1). Coffee is one of the traditional export crops.

Agricultural production has been increasing gradually since the 1990s (Figure 2). The increase in the 1990s was largely attributed to expansion of land area cultivated. Since the early 2000s, the growth factor has been changed to increases in yield (see Taffesse et al. (2012) and Spielman et al. (2012) for detailed discussion). The current crop productivity in Ethiopia is generally favorably compared to its neighboring countries.
Still, the use of advanced inputs, such as fertilizer, pesticide and improved seeds, remains limited in Ethiopia despite various government programs toward promoting the use of advanced inputs. Fertilizer was applied to less than 40 percent of cereal acreage, and pesticides to about 20 percent of land. The use of improved seeds and irrigation is limited to several percent on average (Table 1). Of particular note, fertilizer use per hectare declined in the mid-2000s (Table 2). This was explained by the poor quality and high costs of inputs and the unavailability or delayed delivery of preferred inputs (Taffesse et al. 2012). Location also seems to matter. Two regions, Oromia and Amhara, where agricultural production is concentrated, account for 70 percent of total fertilizer use (Rashid et al. 2013).

### Table 1. Use of advanced inputs for cereals (% of cultivated area)

|          | Fertilizer | Improved seeds | Pesticide | Irrigation |
|----------|------------|----------------|-----------|------------|
|          | 1997/98    | 2001/02        | 2007/08   | 1997/98    | 2001/02 | 2007/08 | 1997/98 | 2001/02 | 2007/08 |
| Maize    | 18.0       | 45.7           | 32.8      | 5.2        | 12.5 | 19.5 | 1.3 | 1.9 | 2.9 | 1.1 | 3.2 | 2.2 |
| Sorghum  | 2.9        | 16.9           | 3.1       | 0.2        | 0.4 | 0.1 | 3.1 | 1.7 | 5.4 | 0.4 | 1.1 | 1.2 |
| Wheat    | 57.0       | 56.7           | 62.1      | 5.6        | 2.0 | 2.9 | 31.3 | 28.1 | 43.6 | 0.3 | 0.4 | 0.5 |
| Barley   | 34.4       | 39.6           | 30.5      | 0.1        | 0.4 | 0.6 | 9.6 | 9.1 | 20.7 | 0.6 | 0.8 | 1.2 |

Source: Taffesse et al. (2012)

### Table 2. Application of fertilizer (kg per ha)

|          | 1997/98 | 2001/02 | 2007/08 |
|----------|---------|---------|---------|
| Maize    | 25      | 28      | 54      |
| Sorghum  | 4       | 1       | 3       |
| Wheat    | 75      | 56      | 85      |
| Barley   | 33      | 20      | 30      |

Source: Taffesse et al. (2012)
The main hypothesis of the paper is that agricultural production would have been affected by the deterioration of port accessibility because of the partial and full shutdown of the Ethio-Djibouti railway in 2007 and in 2009. Port access is a critical bottleneck to the Ethiopian economy. More than 95 percent of total exports and imports pass through the Port of Djibouti, about 800 km away from the capital city, Addis Ababa. Since its completion in the early 20th century, the Ethio-Djibouti Railways has been a major transport means for Ethiopia to access the global market. Until the 1990s, it carried more than 100 million ton-km of freight (Figure 3). By 2007, however, the rail operations were ceased between Addis Ababa and Dire Dawa, mainly because of insufficiency of financial resources and resultant lack of maintenance. The level of rail services deteriorated further. By 2009, the whole rail line ceased operating.

![Figure 3. Rail freight transport (million ton-km)](source: Ethiopia CSA)

The agriculture sector used to be dependent on rail transportation in both input and output terms. Ethiopia imports almost all fertilizer and other advanced inputs and exports a significant amount of agricultural commodities, such as coffee. According to the Ethio-Djibouti rail operation data, historically, about 50,000 tons of coffee and vegetables were exported and about 5,000 tons of fertilizer were imported by rail. These accounted for 40 percent of the total rail freight traffic. Other commodities were fuel and other bulk cargo.

Although the rail traffic was already a small fraction of agricultural exports and fertilizer imports, transport costs are in theory determined by modal competition. The existence or
partial existence of cheap rail transportation should have impacted on market road transport costs. Interestingly, the fertilizer retail prices increased significantly in 2008 (Figure 4), though there are many governmental policies that disturb fertilizer prices in Ethiopia, such as fixed bank interest rates on fertilizer and no storage costs allowed by cooperatives (see Rashid et al. (2013) for detailed discussion). Meanwhile, the national use of fertilizer did not change dramatically but was gradually increasing during the 2000s. Thus, the price spike was not caused by the supply-demand balance. Of course, there are many other possible reasons. Our identification strategy aims at controlling for these factors (see below).

Figure 4. Fertilizer use and retail prices in Ethiopia

From the transport economics point of view, this must have caused significant costs to the economy. Railways normally have comparative advantage for long-haul freight transportation. The freight rates used to be 0.25 to 0.31 Ethiopian birrs (ETB) or 2.9 to 3.5 U.S. cents per ton-km, much lower than typical vehicle operating costs (VOCs) on roads in Ethiopia. Based on the traditional highway engineering model (HDM4), the VOCs vary from 8 to 10 U.S. cents per ton-km, depending on road conditions. Thus, the cease of the rail operations is considered to have more than doubled transport costs of all imports and exports. Agricultural input prices were certainly negatively affected. In Ethiopia, transport costs account for 64-80 percent of fertilizer farmgate prices (Rashid et al. 2013).
The impact of the rail abandonment may partially have been mitigated by rapid road improvements. Road transport is a natural alternative connecting Port of Djibouti and Ethiopia. Ethiopia has a road network composed of 99,522 km of roads, of which 12,640 km or about 17 percent are paved. In recent years, the government has been making significant efforts to develop and maintain its road network. About 26,000 km of federal roads are generally well-maintained. In 2010, the Universal Rural Road Access Program was embarked upon, aimed at connecting all communities (kebeles) by all-weather roads. Still, rural roads are still largely in poor condition: Only about 30 percent are in good condition (Figure 5).

**Figure 5. Road network in Ethiopia**

![Road network in Ethiopia](image)

Source: World Bank (2016) based on Ethiopia Road Authority data.

**III. EMPIRICAL MODEL AND DATA**

To examine the impacts of port connectivity and other factors, a conventional production function is considered (see, for instance, Gyimah-Brempong (1987), Bravo-Ortega and Lederman (2004), and Dorosh et al. (2012)): 
\[ \ln V_{ijt} = \beta_0 + \ln X_{ijt}' \beta_X + Z_{ijt}' \beta_Z + \varepsilon_{ijt} \]  

(1)

where \( V \) is the total value of crops produced by household \( i \) in location \( j \) at time \( t \).
Agricultural production is assumed to depend on the amounts of inputs used, \( X \), and household characteristics, \( Z \). \( \varepsilon \) is an idiosyncratic error. As in the literature, five major production factors are included: labor (denoted by \( L \)), cultivated rainfed land area (\( H \)), fertilizer (\( F \)), improved seeds (\( S \)), and irrigated land area (\( R \)). Transport connectivity, \( TR \), is also included as another production factor. The logarithms are taken for all these variables.

As often discussed in agricultural economics, one empirical issue is that many input variables are likely to be zero in developing countries. In Ethiopia, only one-third of farmers apply fertilizer. The use of improved seeds and irrigation is far more limited. A traditional approach is to add a small positive number to avoid taking the logarithm of zeros. This is an accepted practice, but it may cause a significant bias. Particularly, in our case, the vast majority of households do not use advanced inputs. Thus, Battese’s (1997) specification is incorporated:

\[ \ln V_{ijt} = \beta_0 + \sum_k \beta_k \ln x_{k,ijt}^* + \sum_k \delta_k D_{k,ijt} + Z_{ijt}' \beta_Z + \varepsilon_{ijt} \]  

(2)

where

\[ D_{k,ijt} = \begin{cases} 1 & \text{if } x_{k,ijt} = 0 \\
0 & \text{if } x_{k,ijt} > 0 \end{cases} \text{ and } x_{k,ijt}^* = \max(x_{k,ijt}, D_{k,ijt}) \]  

(3)

\( x_{k,ijt} \) is the amount of input \( k \). By adding the dummy variables for zero inputs \( D_k \), their zero-inflated effects will be removed. Since there are three truncated variables: fertilizer, improved seeds and irrigated land, the equation to be estimated is:

\[ \ln V_{ijt} = \beta_0 + \beta_L \ln L_{ijt} + \beta_H \ln H_{ijt} + \beta_F \ln F_{ijt}^* + \delta_F D_{F,ijt} + \beta_S \ln S_{ijt}^* + \delta_S D_{S,ijt} + \beta_R \ln R_{ijt}^* + \delta_R D_{R,ijt} + \beta_{TR} \ln TR_{ijt} + Z_{ijt}' \beta_Z + c_j + c_i + \varepsilon_{ijt} \]  

(4)
where $c_j$ and $c_t$ are district- and time-specific fixed effects. As will be discussed below, our data are not panel but cross-section data from different points of time. The current study uses eight rounds of the Ethiopian Agriculture Sample Surveys, in which interviewed households may or may not be the same across years. The original data do not allow to identify individual farmers or their locations. But the locational data are available at the district or woreda level, which is detailed enough to examine the impacts of transport connectivity. Ethiopia has 11 regions, which are divided into 72 zones. There are 567 districts in the country (Figure 6). All 195,000 farmers are located in one of the districts.

**Figure 6. Districts and major transport networks in Ethiopia**

The district-specific fixed effects, $c_j$, have a particularly important role to control for time-invariant local characteristics, such as agro-climatic potential, which is normally taken into account in the empirical agricultural literature. It is technically possible to include some

---

2 The numbers of administrative units may not be consistent with the current classification in the country, because our empirical data come from the agricultural censuses for the past 8 years, during which administrative boundaries have been changed. Some of the data cannot be matched. Therefore, the analysis uses the boundary data that fit the data best, and some woredas were omitted from the analysis.
measurements, such as the Spatial Production Allocation Model (SPAM) which provides agricultural productivity data at a highly disaggregated level (approximately 10 x 10 km resolution). However, such potential variables do not vary much over time, especially during our 8-year sample period. Our preferred specification is to use the district-specific fixed effects, while omitting agro-climatic potential variables. Empirically, the time-invariant district effects are found to be significant and cannot be removed from the equation.

For transport connectivity, $TR$, the current paper is focused on connectivity to port—in this case, Djibouti. It is defined by the lowest transport cost from each district to Djibouti. This is calculated by spatial software based on underlying unit costs of rail and road transportation. Unlike the existing literature, our transport variable is a multimodal measurement. This is advantageous because in reality, transport users always select the best means among available alternatives. For instance, it is estimated to have cost US$72.9 per ton to ship 1 ton of goods from Addis Ababa to Djibouti in 2007. The cost is divided into two parts: US$55 on road transport from Addis Ababa to Dire Dawa and US$18 on rail transport from Dire Dawa to Djibouti, because rail transport costs were lower at the latter section.

As an identification strategy, the current paper takes advantage of a historical event that Ethiopia has experienced for our data period: 2003-10. As discussed in the previous section, port connectivity was significantly deteriorated by the cease of the Ethio-Djibouti rail operations. The rail link between Addis Ababa and Dire Dawa was only operational until 2007. The Dire Dawa-Djibouti section was available until 2009. Rail costs are calculated based on the actual fares and adjusted by inflation. In real terms, the average unit rail cost from Addis Ababa to Djibouti declined from 3.5 U.S. cents per ton-km in 2003 to 3.2 U.S. cents in 2006. The real cost from Dire Dawa to Djibouti declined from 4.3 U.S. cents per ton-

---

3 See You et al. (2009).
4 Each district’s location is spatially defined by its centroid.
5 Changes in the rail sector are still continuing. In 2013, the Dire Dawa – Djibouti section was reopened. At the same time, a new rail line from Djibouti to Addis Ababa is also under construction.
km in 2003 to 3.3 U.S. cents per ton-km in 2008. Rail costs were generally much lower than road transport costs.

For road transportation, vehicle operating costs (VOCs) are estimated based on the Ethiopia Road Agency’s road network data. The estimated costs vary significantly across locations and over years (Figure 7). In principle, road rehabilitation and maintenance works will reduce the roughness of road surface, lowering VOCs. On the other hand, the road conditions deteriorate without proper maintenance, resulting in higher VOCs. For instance, if a paved road is in good condition (e.g., IRI=2), the unit VOC for a light truck is estimated at 7.1 U.S. cents per ton-km in Africa. If the road condition is poor (e.g., IRI=6), the unit VOC is 8.1 U.S. cents, 15 percent higher. For unpaved roads, it exceeds 10-15 U.S. cents per ton-km, depending on roughness. VOCs are also affected by other underlying parameters, such as fuel costs and time costs. Gasoline prices in Ethiopia, for instance, increased from US$0.52 per liter in 2003 to US$0.91 per liter in 2010, a more than 40 percent increase in real terms.

VOCs are a good proxy of market transport prices that people or firms have to pay. In theory, with free market entry, the market prices should converge on VOCs. However, this may not always be the case at least in the short run. Teravaninthorn and Raballand (2009) show that Africa’s average transport prices (6-11 U.S. cents) are relatively high compared to other regions. This is because of the poor quality of the road network and the lack of competition in the trucking industry. In East Africa (e.g., between Nairobi and Mombasa), the profit margins can reach 60 percent. To check the robustness of the results, two cases are considered in the following analysis: (i) VOCs are used as they are, and (ii) road transport costs are adjusted with a 60 percent markup taken into account. In the latter case, rail and road transport costs are conceptually comparable.
Given all these unit costs, the minimum transport costs are calculated from a given location to the port for each year.\(^6\) Therefore, as shown in the figure, the transport costs vary substantially across locations and across years.

**Figure 7. Transport costs to the Port of Djibouti (US$/ton)**

(2005) ![Map of Djibouti 2005 Transport Costs](image)

(2007) ![Map of Djibouti 2007 Transport Costs](image)

(2009) ![Map of Djibouti 2009 Transport Costs](image)

Source: Authors’ calculation based on ERA data.

Unlike other existing studies, it is defined as a continuous variable and thus can represent granularity of local connectivity. This is another advantage of our transport variable. In the existing literature (e.g., Khandker et al. 2009; Dercon et al. 2009; Qin and Zhang, 2016),

\(^6\) The figures depict transport costs with approximately 10 x 10 km resolution. For the empirical regression, transport costs are calculated for each of the 567 districts.
transport connectivity is often defined in a dichotomous fashion, merely indicating, for example, whether a particular type of road passes through a certain area, or whether road work was recently implemented in a particular area. Such approaches do not allow to control for the situation where different farmers at different locations would benefit differently from changing transport connectivity. Transport connectivity is often complex and nonlinear. Our transport variable has the advantage to measure such complexity.

The most important empirical challenge to estimate Equation (4) is endogeneity of the transport variable, $TR$. This is essentially because transport infrastructure placement must of necessity be endogenous (e.g., Banerjee et al. 2012; Datta, 2012; Jedwab and Moradi, 2012). While transport infrastructure is essential to increase economic productivity, governments tend to invest more in transport infrastructure where productivity is inherently high. Therefore, ordinary least squares (OLS) is likely to generate biased estimates.

To address this, the current paper combines two approaches. First, as discussed above, highly disaggregated district-specific fixed effects are included. This is expected to control for unobserved time-invariant location-specific effects. Unlike the existing panel analyses (e.g., Dercon et al. 2009; Khandker and Koolwal, 2011), our data are not panel but spatially highly disaggregated. Thus, this quasi-panel treatment helps to eliminate the endogeneity bias to a large extent.

Second, the instrumental variable technique is also used. Following Banerjee et al. (2012), which examine the causal effect of historical rail network development in China, two instruments are constructed: One is the straight-line distance from each district to the rail line that was operational at $t$. This is denoted by $RAIL_{jt}$ and clearly relevant to transport connectivity to the port at that time but may not be directly related to agricultural productivity. Another instrument is the difference in elevation between each district and its closest operational rail station, denoted by $ELEV_{jt}$. Again, the idea behind it is the same. Available rail stations changed over time. If the difference in elevation is greater, the transport costs are likely to be high, holding everything else constant. The road condition
tends to be poor in hilly and mountainous areas, raising VOCs. In addition, fuel consumption also increases with hilliness of the terrain.

Summary statistics are shown in Table 3. Our primary data source is the annual Agricultural Sample Surveys in Ethiopia. The analysis uses the data for 2003-10. Every year the survey covers approximately 30,000 households all over the nation. The surveys are nationally representative. About 80 percent of households surveyed live in three regions: SNNP, Oromia and Amhara (Figure 8). This is consistent with the national population distribution.

**Table 3. Summary statistics**

|                                | Abb. | Obs.  | Mean  | Std. Dev. | Min   | Max   |
|--------------------------------|------|-------|-------|-----------|-------|-------|
| Total value of crops produced by household (US$) | V    | 195,623 | 18.34 | 23.66     | 0.0002 | 117.65 |
| Transport cost to Djibouti ($/ton) | TR   | 195,623 | 100.33 | 41.37     | 8.04   | 196.06 |
| Household size                   | L    | 195,623 | 5.09  | 2.26      | 1      | 16    |
| Cultivated rainfed land area (ha) | H    | 195,623 | 0.79  | 0.70      | 0.0001 | 4.31  |
| Improved seeds used (kg)         | S    | 14,968 | 3,818 | 9,121     | 0.002  | 150,000 |
| Fertilizer used (kg)             | F    | 52,704 | 38.03 | 28.75     | 0.001  | 108.00 |
| Irrigated land area (ha)         | R    | 13,676 | 0.12  | 0.21      | 0.0004 | 3.81  |
| Dummy variable for households receiving extension services | EXT  | 195,623 | 0.15  | 0.36      | 0      | 1     |
| Dummy variable for male head     | MAL  | 195,623 | 0.81  | 0.39      | 0      | 1     |
| Household head's highest grade completed | EDU | 195,623 | 2.51  | 2.81      | 1      | 25    |
| Share of own land (0 to 1)       | OWN  | 195,623 | 0.88  | 0.25      | 0      | 1     |
| Straight-distance to the operational rail line | RAIL | 195,623 | 425   | 243       | 0.06   | 1,057 |
| Difference in elevation from the nearest operational rail station | ELEV | 195,623 | 719   | 980       | -1,677 | 3,916 |

The surveys include all kinds of agricultural crops grown in Ethiopia, but major food crops, such as teff, sorghum, maize and wheat, account for 60 percent of that total cultivated land in the sample data (Figure 9). Our output variable is the total value of crops produced by each household. The original data provide the information on input and output quantities for each crop at the plot level. To aggregate the data at the household level, the FAO producer prices are used. Note that they are not consumer prices. In Ethiopia, not many farmers are actually involved in market transactions. The same 2010 prices are used. Therefore, the possible noise of fluctuating domestic or international commodity prices is eliminated.
The whole family members are assumed to be engaged in agricultural production. The average household size is 5.1. In the Agricultural Sample Surveys, no data are available that show how many family members are actually working on a particular type of crop or a particular plot. This is the main reason why households, not plots, are used as the unit of analysis in the current paper. All inputs (e.g., volume of fertilizer used) and outputs (e.g., quantity of maize harvested) are aggregated at the household level, and the household size is used as a proxy of labor inputted.7

As discussed, most farmers are small-holders in Ethiopia. The average acreage is 0.8 ha. The use of advanced inputs is generally limited. About one-quarter of farmers use fertilizer. Even if used, the volume used is minimal, on average 37.9 kg per household or 62.6 kg per ha. Improved seeds and irrigation are more limited. Hence, it is empirically important to deal with these zero-input variables.

---

7 Ideally, the number of household members who engage in agricultural activities should be used. Unfortunately, the AG census data do not tell us such information. In general, however, the number of all household members is a good proxy of the agricultural labor force in each household. The share of economically active members does not change much in a particular area. In addition, our model includes district-specific fixed effects. So, at that level, our labor variable is a good proxy.
As usual, basic household characteristics are also included, such as the sex of household heads and their educational attainment. The share of own land is also included. About 88 percent of land cultivated is owned by farmers themselves. The rest is borrowed land.

**IV. MAIN ESTIMATION RESULTS**

The instrumental variable (IV) regression is performed with the Battese’s specification incorporated. The results are shown in Table 4. Regarding endogeniety, the IV result is broadly similar to the OLS estimate. However, there is a marked difference in several coefficients. In the OLS estimation, for instance, the coefficient of \( TR \) is negative as expected but not statistically significant. According to the IV regression, the coefficient is negative and significant. Formally, the hypothesis that the transport variable is exogenous can easily be rejected. The conventional exogeneity test statistic is estimated at 398.83. Thus, the IV result is consistent, while the OLS is likely to be biased.

The validity of the two instruments constructed is also confirmed with the data. The overidentifying restriction test statistic is estimated at 1.422, which is less than any conventional threshold. Hence, our instruments are valid. As expected, both instruments have positive coefficients in the first-stage regression (Table 5). The coefficient of the distance to the nearest rail line (\( RAIL \)) is estimated at 0.309 with a standard error of 0.002. The coefficient of the difference in elevation from the nearest station (\( ELEV \)) is also significantly positive at 0.032. Therefore, as expected, transport costs increased as the Ethio-Djibouti rail operations deteriorated in the 2000s. Recall that our transport variable, \( TR \), measures not only rail but also road connectivity to the port. Thus, the evidence is interpreted to mean that while the rail infrastructure was abandoned and the road network has been improved, the latter development could not compensate for the former negative effect.

Regarding the impact of the deteriorated port connectivity, it is found significant in the IV estimation. The coefficient of \( \ln TR \) is -0.276, which is statistically significant (Table 4). Thus, if the transport cost to the port is reduced by 10 percent, the agricultural production value
would increase by 2.7 percent. This is lower than the findings of earlier studies in Africa. Dorosh et al. (2012) estimate the elasticity of crop production with respect to transport connectivity at -1.2 to -4.8. Decron et al. (2009) show that Ethiopian household consumption increases by 16 percent if access to all-weather road is granted. Although their specification is not directly comparable, the underlying change in transport connectivity seems to have been drastic, from 36 percent to 67 percent in the surveyed villages. This makes our estimated elasticity more comparable.

Table 4. OLS and IV estimation results with the Battese’s specification

|        | OLS Coef. | OLS Std.Err. | IV Coef. | IV Std.Err. |
|--------|-----------|--------------|----------|-------------|
| lnTR   | -0.013    | (0.015)      | -0.276   | (0.019) *** |
| lnL    | 0.034     | (0.005) ***  | 0.033    | (0.005) *** |
| lnH    | 0.829     | (0.003) ***  | 0.829    | (0.003) *** |
| lnS    | -0.001    | (0.002)      | -0.001   | (0.002)     |
| lnF    | 0.010     | (0.003) ***  | 0.009    | (0.003) *** |
| lnR    | 0.176     | (0.005) ***  | 0.175    | (0.005) *** |
| Ds     | -0.173    | (0.011) ***  | -0.169   | (0.011) *** |
| Df     | 0.093     | (0.012) ***  | 0.092    | (0.012) *** |
| Dr     | -1.041    | (0.019) ***  | -1.040   | (0.019) *** |
| EXT    | 0.040     | (0.006) ***  | 0.037    | (0.006) *** |
| MAL    | 0.059     | (0.006) ***  | 0.059    | (0.006) *** |
| lnEDU  | 0.026     | (0.003) ***  | 0.026    | (0.003) *** |
| OWN    | 0.121     | (0.008) ***  | 0.122    | (0.008) *** |
| constant | 3.441    | (0.210) ***  | 4.350    | (0.214) *** |

Obs. 195,623 | 195,623
R-sq 0.772 | 0.772
Wald chi2 4068993
No. of dummy var.
District 450 | 450
Year 7 | 7
Exogeneity test chi2 398.83 ***
Overidentifying restriction test chi2 1.422

Note: The dependent variable is the log of the total value of crops produced by each household. Robust standard errors are shown in parentheses. *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.

For other explanatory variables, labor is found to be useful but less productive. The elasticity is only 0.03. This is a common characteristic of subsistence farming in Africa, which is
generally labor-intensive. Not many gains can be expected from adding more labor. Land is productive, and its large elasticity reflects that the agricultural growth during the sample period was largely explained by land expansion. Fertilizer and irrigation also have positive elasticities. These are broadly consistent with the earlier evidence (e.g., Taffesse et al. 2012; Spielman et al. 2012; Mekonnen et al. 2013). The impact of improved seeds remains unclear. The coefficient is nearly zero and insignificantly different from zero.

The extension service variable has a positive and significant coefficient of 0.037. This is also consistent with some earlier studies, such as Dercon et al. (2009). Receiving a visit of an extension service officer can contribute to agricultural production growth.

Table 5. First stage regression for IV estimation

|                  | Battese specification | Small positive value specification |
|------------------|-----------------------|-----------------------------------|
|                  | Coef. | Std.Err. | Coef. | Std.Err. |
| lnRAIL           | 0.309 (0.002) ***     | 0.309 (0.002) ***                 |
| ELEV             | 0.032 (0.001) ***     | 0.032 (0.001) ***                 |
| lnL              | 0.464 (0.386)         | 0.476 (0.385)                     |
| lnH              | 0.259 (0.197)         | 0.248 (0.197)                     |
| lnS              | 1.289 (0.197) ***     | -0.220 (0.079) ***                |
| lnF              | -0.647 (0.362) *      | -0.144 (0.077) *                  |
| lnR              | 0.964 (0.440) **      | 1.094 (0.391) ***                 |
| Ds               | 9.812 (1.284) ***     |                                   |
| Df               | -1.677 (1.219)        |                                   |
| Dr               | -4.779 (1.798) ***    |                                   |
| Ext              | -6.856 (0.689) ***    | -7.918 (0.669) ***                |
| Mal              | -0.054 (0.509)        | -0.052 (0.509)                    |
| lnEDU            | 0.840 (0.230) ***     | 0.833 (0.230) ***                 |
| Own              | 2.375 (0.806) ***     | 2.352 (0.806) ***                 |
| Constant         | 2.857 (0.005) ***     | 2.864 (0.004) ***                 |
| Obs.             | 195,623               | 195,623                           |
| R-sq             | 0.976                 | 0.976                             |
| F statistic      | 5.9E+10               | 5.91E+10                          |
| No. of dummy var.| District 450          | 450                               |
|                  | Year 7                | 7                                 |

Note: The dependent variable is the log of the transport cost, TR. Robust standard errors are shown in parentheses. For presentation purposes, the coefficients and standard errors are multiplied by 1,000. *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.
V. DISCUSSION

From the methodological point of view, one may be concerned whether the above results are robust against specification, especially, the way of treating zero-input variables. To examine this, the traditional specification using a small positive number (i.e., 0.01) is also estimated (Table 6). The results are broadly similar to the above. In the IV regression, the transport connectivity has a negative and significant impact on agricultural production. While labor and land are both productive, the latter is the most important production factor. The positive impact of irrigation does not change, but the fertilizer variable turned out to have a negative and significant coefficient, which is different from the above and counterintuitive. This might be a bias created by replacing all zeros with an artificial small positive number.

|         | OLS            |            | IV            |            |
|---------|----------------|------------|---------------|------------|
|         | Coef.          | Std.Err.   | Coef.         | Std.Err.   |
| lnTR    | -0.012 (0.015) | ***        | -0.273 (0.019) | ***        |
| lnL     | 0.037 (0.005)  | ***        | 0.036 (0.005) | ***        |
| lnH     | 0.829 (0.003)  | ***        | 0.830 (0.003) | ***        |
| lnS     | 0.017 (0.001)  | ***        | 0.017 (0.001) | ***        |
| lnF     | -0.007 (0.001) | ***        | -0.007 (0.001) | ***        |
| lnR     | 0.230 (0.004)  | ***        | 0.230 (0.004) | ***        |
| EXT     | 0.057 (0.006)  | ***        | 0.054 (0.006) | ***        |
| MAL     | 0.061 (0.006)  | ***        | 0.061 (0.006) | ***        |
| lnEDU   | 0.026 (0.003)  | ***        | 0.027 (0.003) | ***        |
| OWN     | 0.121 (0.008)  | ***        | 0.122 (0.008) | ***        |
| constant| 3.413 (0.210)  | ***        | 4.313 (0.214) | ***        |

Obs. 195,623
R-sq 0.772
Wald chi² 3688089
No. of dummy var. 450
District 7
Exogeneity test chi² 388.20 ***
Overidentifying restriction test chi² 0.484

Note: The dependent variable is the log of the total value of crops produced by each household. Robust standard errors are shown in parentheses. *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.
As discussed above, another issue is that transport costs based on vehicle operating costs may be the same as actual market transport prices. With a 60 percent markup assumed, the transport costs to the port were recalculated (TR*). Then, the OLS and IV regression models are re-estimated with Battese’s specification. The results are broadly unchanged (Table 7). The transport connectivity has a negative and significant impact on agricultural production. Even in the OLS model, the coefficient became negative. Other coefficients remain quite similar to the previous results.

Table 7. OLS and IV estimation results with alternative transport cost variable

|        | OLS          | IV          |
|--------|--------------|-------------|
|        | Coef. Std.Err. | Coef. Std.Err. |
| lnTR*  | -0.028 (0.011) ** | -0.197 (0.014) *** |
| lnL    | 0.034 (0.005) *** | 0.033 (0.005) *** |
| lnH    | 0.829 (0.003) *** | 0.830 (0.003) *** |
| lnS    | -0.001 (0.002) | -0.001 (0.002) |
| lnF    | 0.010 (0.003) *** | 0.010 (0.003) *** |
| lnR    | 0.176 (0.005) *** | 0.175 (0.005) *** |
| DS     | -0.172 (0.011) *** | -0.170 (0.011) *** |
| DF     | 0.093 (0.012) *** | 0.093 (0.012) *** |
| DR     | -1.041 (0.019) *** | -1.040 (0.019) *** |
| EXT    | 0.040 (0.006) *** | 0.037 (0.006) *** |
| MAL    | 0.059 (0.006) *** | 0.059 (0.006) *** |
| lnEDU  | 0.026 (0.003) *** | 0.026 (0.003) *** |
| OWN    | 0.121 (0.008) *** | 0.122 (0.008) *** |
| constant | 3.495 (0.208) *** | 4.085 (0.209) *** |
| Obs.   | 195,689      | 195,689     |
| R-sq   | 0.772        | 0.772       |
| Wald chi2 | 4070971     |             |
| No. of dummy var. | District 450 | 450 |
| Year   | 7            | 7           |
| Exogeneity test chi2 | 329.59 *** |             |
| Overidentifying restriction test chi2 | 3.187 * |             |

Note: The dependent variable is the log of the total value of crops produced by each household. Robust standard errors are shown in parentheses. *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.

From a policy point of view, it is of particular interest how improved (or aggravated) transport connectivity results in higher (or lower) agricultural production. To explore this
issue, two input variables, fertilizer and improved seeds, are regressed over the transport connectivity variable:

\[
X_{ijt} = \gamma_0 + \gamma_{TR} \ln TR_{ijt} + Z_{ijt}' \gamma_Z + c_j + c_t + u_{ijt} \quad \text{for } X = \{F, S\} \tag{5}
\]

An empirical issue is that the dependent variables (i.e., input variables, \(F\) and \(S\)) are censored: Not many farmers use fertilizer or improved seed. To deal with this problem, a conventional truncation mechanism is assumed:

\[
X_{ijt} = \begin{cases} 
X^*_{ijt} = \gamma_0 + \gamma_{TR} \ln TR_{ijt} + Z_{ijt}' \gamma_Z + c_j + c_t + u_{ijt} & \text{if } X^*_{ijt} > 0 \\
0 & \text{otherwise}
\end{cases} \tag{6}
\]

The latent variable of the input is denoted by \(X^*\).

The IV tobit regression is performed with the same instruments as above. The fertilizer use is found to be a possible major channel to explain agricultural growth in Ethiopia. In the regression on fertilizer, the coefficient of \(TR\) is estimated at -6.58, which is significant (Table 8). Thus, the fertilizer use increases with transport cost reduction. This is consistent with the earlier literature (e.g., Rashid et al. 2013). In Ethiopia, the vast majority of the fertilizer prices are attributed to transport costs from the port to regional cooperatives’ warehouses and farmers. The evidence is also similar to the finding by Qin and Zhang (2016), in which road access is found to promote farmers’ use of fertilizer in China.

On the other hand, the use of improved seeds is not necessarily explained by transport costs to the port. The exogeneity test cannot be rejected. Thus, the simple tobit model is also performed. The results are unchanged. The coefficients of \(TR\) are negative but not statistically significant. This may be because of the use of improved seeds remains much more limited than the use of fertilizer. As pointed out by Spielman et al. (2012), there are a number of institutional reasons for Ethiopia’s low adoption of improved seeds, such as lack
of private innovator capacity and lack of information on the proper variety choice. The current supply of improved seed is not timely and falls short of the increasing demand.

Table 8. IVTOBIT and TOBIT regression on fertilizer and seed use

| Dependent variable | Estimation model | Coef. | Std.Err. | Coef. | Std.Err. | Coef. | Std.Err. |
|--------------------|------------------|-------|----------|-------|----------|-------|----------|
|                    | IVTOBIT          |       |          | IVTOBIT |          | TOBIT  |          |
| lnTC               | -6.58 (1.34) *** | -0.53 (0.67) | -0.62 (0.44) |
| EXT                | 51.28 (0.34) *** | 14.11 (0.16) *** | 14.11 (0.34) *** |
| MAL                | 9.37 (0.37) *** | 1.04 (0.17) *** | 1.04 (0.17) *** |
| lnEDU              | 3.71 (0.17) *** | 1.16 (0.07) ** | 1.16 (0.08) ** |
| OWN                | -12.13 (0.53) *** | -0.23 (0.24) | -0.23 (0.23) |
| constant           | 5.30 (6.53)     | -21.36 (3.02) *** | 21.07 (2.97) *** |

Obs. 204,446     204,446     204,446
Wald chi2 58501   11576.8
Pseudo R2 0.127

Note: The dependent variables are the amounts of fertilizer and improved seeds used by each household, respectively. Standard errors are shown in parentheses. *, ** and *** indicate statistical significance at the 10, 5 and 1 percent level, respectively.

VI. CONCLUSION

Agriculture remains an important economic sector in Africa, employing a large share of the labor force and earning foreign exchange. Among others, transport connectivity has long been a crucial constraint in Africa. In theory, railways have a particularly important role to play in shipping freight and passengers at low cost. However, most African railways were in virtual bankruptcy by the 1990s.

Using the case of Ethiopia, the paper recast light on the possible impacts of rail transport on agricultural production. It is a methodological challenge to estimate an unbiased impact of large-scale infrastructure, such as railroads, because of the well-known endogeneity of
infrastructure placement. The paper attempted to overcome the problem by using a large sample of data comprised of over 190,000 households over 8 years, and constructing valid instruments with a spatial technique applied. The paper took advantage of the historical event in Ethiopia that the Ethio-Djibouti Railways ceased operating from time to time during the 2000s.

It is shown that transport costs were increased substantially for the period of 2003-2010. This is largely because of a partial and full cease of the rail operations in 2007 and 2009, and partly because of increasing vehicle operating costs for road transport, mainly driven by increased fuel prices. Substantial road improvements have been made during the same period. But they do not seem to have been able to compensate for the negative impact of abandoned rail transport services.

The paper also found that the transport connectivity variable is likely endogenous. Thus, the OLS estimates are biased. In the IV regression, the impact of the port connectivity is found to be significant. The elasticity is estimated at 0.276. Thus, a 10 percent reduction in transport costs to the port would increase agricultural production by 2.7 percent. The use of fertilizer is also found to increase with transport cost reduction. This can be interpreted to mean that efficient port connectivity is particularly important to promote fertilizer use and increase crop production.
REFERENCES

Amin, Mohamed, Duncan Willetts, and Alastair Matheson. 1986. *Railway across the Equator: The Story of the East African Line*. The Bodley Head Ltd. London.

Banerjee, A., E, Duflo, and N. Qian. 2012. On the road: Access to transportation infrastructure and economic growth in China,” NBER Working Paper 17897, National Bureau of Economic Research, Washington, DC.

Bravo-Ortega, and Lederman. 2004. Agricultural productivity and its determinants: Revisiting international experiences. *Estudios de Economia*, Vol. 31(2), pp. 133-163.

Datta, Saugato. 2012. The impact of improved highways on Indian firms. *Journal of Development Economics*, Vol. 99(1), pp. 46-57.

Dercon, Stefan, Daniel Gilligan, John Hoddinott and Tassew Woldehanna. 2009. The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages, *American Journal of Agricultural Economics*, Vol. 91(4), pp. 1007-1021.

Dillon, Andrew. 2011. Do differences in the scale of irrigation projects generate different impacts on poverty and production? *Journal of Agricultural Economics*, Vol. 62(2), pp. 474-492.

Donaldson, Dave. 2016. Railroads and the Raj: The economic impact of transportation infrastructure. *American Economic Review*, forthcoming.

Dorosh, Wang, You, and Schmidt. 2012. Road connectivity, population, and crop production in Sub-Saharan Africa. *Agricultural Economics*, Vol. 43, pp. 89-103

Ethiopian Agricultural Transformation Agency. 2014. Annual Report 2013/14: Transforming Agriculture in Ethiopia.

Gwilliam, Ken. 2011. *Africa’s Transport Infrastructure: Mainstreaming Maintenance and Management*. The World Bank.

Gyimah-Brempong. 1987. Scale elasticities in Ghanaian cocoa production. *Applied Economics*, Vol. 19, pp. 1383-1390.
Jedwab, R. and A. Moradi. 2012. Colonial investments and long-term development in Africa: Evidence from Ghanaian railroads, unpublished paper, George Washington University; STICERD, London School of Economics; and University of Sussex.

Khandker, Shahidur, and Gayatri Koolwal. 2011. Estimating the long-term impacts of rural roads: A dynamic panel approach, Policy Research Working Paper No. 5867, The World Bank.

Khandker, Shahidur, Zaid Bakht, and Gayatri Koolwal. 2009. The poverty impact of rural roads: Evidence from Bangladesh. Economic Development and Cultural Change, Vol. 57(4), pp. 685-722.

Mekonnen, Dawit Kelemework, Jeffrey Dorfman, and Greg Fonsah, 2013. Productivity and Efficiency of Small Scale Agriculture in Ethiopia, Southern Agricultural Economics Association 2013 Annual Meeting, February 2-5, 2013, Orlando, Florida.

Olievschi, Vasile Nicolae. 2013. Railway Transport: Framework for improving railway sector performance in Sub-Saharan Africa. SSATP Working Paper No. 94, Africa Transport Policy Program.

Rashid, Shahidur, Nigussie Tefera, Nicholas Minot, Gezahengn Ayele. 2013. Fertilizer in Ethiopia. IFPRI Discussion Paper 01304, International Food Policy Research Institute.

Qin, Yu, and Xiaobo Zhang. 2016. The road to specialization in agricultural production: Evidence from rural China. World Development, Vol. 77, pp. 1-16.

Spielman, David, Dawit Mekonnen, and Dawit Alemu, 2012. Seed, fertilizer, and agricultural extension in Ethiopia, IFPRI book chapters, in: Food and agriculture in Ethiopia: Progress and policy challenges, chapter 4 International Food Policy Research Institute.

Taffesse, Alemayehu Seyoum; Dorosh, Paul A. and Gemessa, Sinafikeh Asrat. 2012. Crop production in Ethiopia: Regional patterns and trends. In Food and Agriculture in Ethiopia: Progress and Policy Challenges, ed. Paul A. Dorosh and Shahidur Rashid. Chapter 3 pp. 53-83. Philadelphia, PA: University of Pennsylvania Press.

Teravaninthorn, Supeer, and Gaël Raballand. (2009). Transport Prices and Costs in Africa: A Review of International Corridors. The World Bank.
Xu, Zhiying, Zhengfei Guan, T.S. Jayne, Roy Black. 2009. Factors influencing the profitability of fertilizer use on maize in Zambia. *Agricultural Economics*, Vol. 40, pp. 437-446.

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

World Bank. 2013. *Growing Africa: Unlocking the Potential of Agribusiness*. 