Journal of Applied Economics

ISSN: 1514-0326 (Print) 1667-6726 (Online) Journal homepage: https://www.tandfonline.com/loi/recs20

# Port rail connectivity and agricultural production: evidence from a large sample of farmers in Ethiopia

Atsushi Iimi, Haileysus Adamtei, James Markland & Eyasu Tsehaye

**To cite this article:** Atsushi Iimi, Haileysus Adamtei, James Markland & Eyasu Tsehaye (2019) Port rail connectivity and agricultural production: evidence from a large sample of farmers in Ethiopia, Journal of Applied Economics, 22:1, 152-173, DOI: 10.1080/15140326.2019.1591814

**To link to this article:** https://doi.org/10.1080/15140326.2019.1591814



Published online: 27 Mar 2019.

JOURNAL OF APPLIED ECONOMICS
2019, VOL. 22, NO. 1, 152–173
https://doi.org/10.1080/15140326.2019.1591814

Routledge
Taylor & Francis Group

OPEN ACCESS  Check for updates

# Port rail connectivity and agricultural production: evidence from a large sample of farmers in Ethiopia

Atsushi Iimi, Haileysus Adamtei, James Markland and Eyasu Tsehaye

Transport Global Practice, World Bank Group, Washington D.C., USA

**ABSTRACT**

Agriculture important in Africa, employing a large share of the labor force and earning foreign exchange. Transport connectivity has long been a crucial constraint in the region. In theory, railways have the advantage of shipping bulky freight, such as fertilizer, at low costs. However, in many African countries, railways were in virtual bankruptcy in 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. The paper takes advantage of the historical event that a major rail line connecting the country to Port Djibouti was abandoned during the 2000s. With the fixed effects and instrumental variable techniques combined, an agricultural production function is estimated. It is found that deteriorated transport accessibility to the port had a significantly negative impact. The use of fertilizer particularly decreased with increased transport costs.

**ARTICLE HISTORY**
Received 8 November 2017
Accepted 9 February 2019

**KEYWORDS**
Agriculture production; transport infrastructure

**JEL CLASSIFICATION**
Q12; R40; C26; C23

## 1. Introduction

Agriculture is an important economic sector in Africa. In many countries, more than half of the total population is still engaged in agricultural production. Agriculture and agribusiness 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% of the recommended amount of fertilizer, although the farmers' fertilizer consumption has been increasing gradually (Xu, Zhengfei Guan, & Jayne, 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.[1] 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% (Gwilliam, 2011). The literature shows that regardless of mode, improved transport infrastructure can provide

**CONTACT** Atsushi Iimi aiimi@worldbank.org World Bank, 1818 H Street NW, Washington, DC 20433, USA

[1]See Spielman et al. (2012) for a detailed discussion on various institutional constraints to promote the use of advanced inputs in Ethiopia.

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, & Koolwal, 2009). Donaldson (2018) 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, historically, rail transportation has played an important role in shipping bulky freight at low costs. Since the later 1890s, many European powers developed rail links to connect fertile and resource-rich inland areas to seaports (e.g., Amin, Willetts, & Matheson, 1986; Jedwab & Moradi, 2012). In theory, rail transport still has the advantage of 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 non-operational. 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 the 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, Duflo, & Qian, 2012; Datta, 2012), it is a challenge to estimate an unbiased impact of large-scale infrastructures, 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 a positive correlation between transport investment and economic outcomes. By nature, a 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 new spatial data, namely road network data that the Ethiopian Roads Authority developed. It will be shown that port connectivity, in particular, varies across locations and deteriorated considerably in the late 2000s, when a main regional rail line 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., Mekonnen, Jeffrey, & Greg, 2013; Rashid, Nigussie, Nicholas, & Gezahengn, 2013; Spielman, Mekonnen, & Alemu, 2012; Taffesse, Dorosh, & Gemessa, 2012), 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 some policy implications. Section V discusses robustness in the transport sector. Then, Section VI concludes.2.

## 2. 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% of the workforce, accounting for 80% 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 the expansion of land area cultivated. Since the early 2000s, the growth factor has been changed to increases in yield supported by gradual intensification of input use, such as improved seeds, pesticide, and fertilizer (see Taffesse et al. (2012) and Spielman et al. (2012) for a detailed discussion). The current crop productivity in Ethiopia is generally favorably compared to its neighboring countries.

Still, the use of such advanced inputs is still limited in Ethiopia. Despite various government's promotions, fertilizer was applied to less than 40% of cereal acreage, and pesticides to about 20% of the land. The use of improved seeds and irrigation is even more limited to several percents (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). The location also seems to matter. Two regions, Oromia and Amhara, where agricultural production is concentrated, account for 70% of total fertilizer use (Rashid et al., 2013).

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

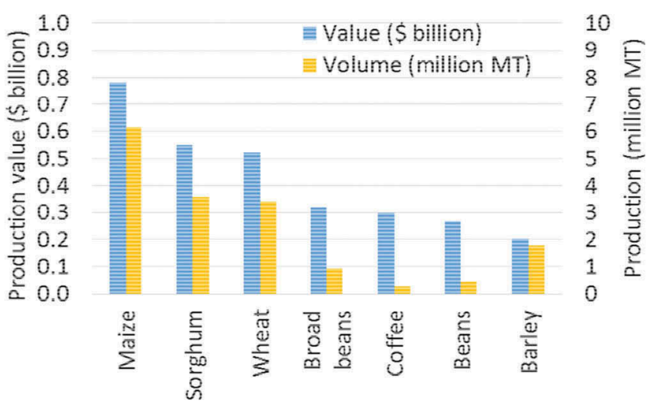


Figure 1. Major crop production in Ethiopia, 2012.

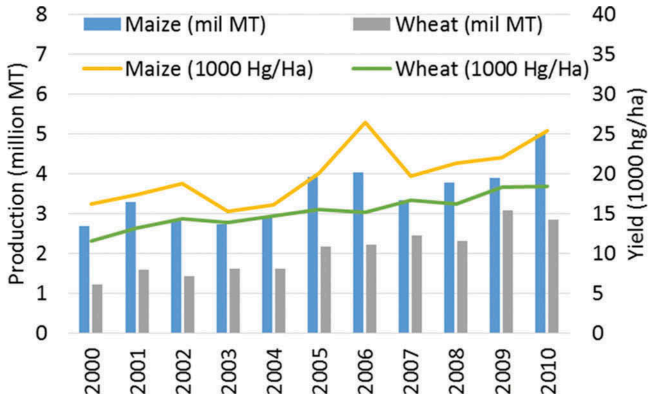


**Figure 2.** Maize and wheat yields in Ethiopia. Source: FAOSTAT. Source: FAOSTAT.

**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 | 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)

bottleneck to the Ethiopian economy. More than 95% 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 twentieth century, the Ethio-Djibouti Railways has been 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.

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% of the total rail freight traffic. Other commodities were fuel and other bulk cargo.

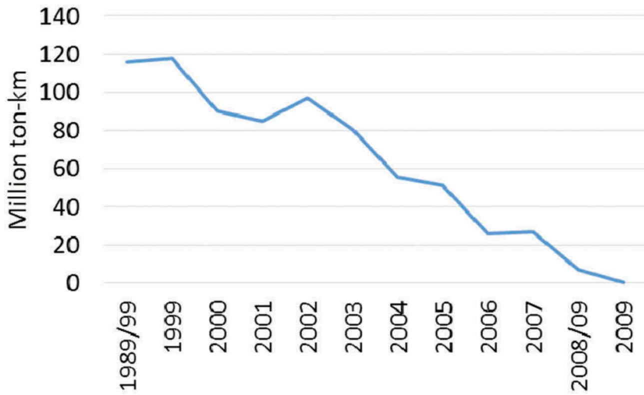


**Figure 3.** Rail freight transport (million ton-km). Source: Ethiopia CSA.

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 a detailed discussion). Meanwhile, the national use of fertilizer did not change dramatically but was gradually increased 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).

From the transport economics point of view, the shutdown of the Ethio-Djibouti railway must have caused significant costs to the economy. Railways normally have a 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 a typical vehicle operating costs (VOCs) on roads in Ethiopia. Based on the traditional

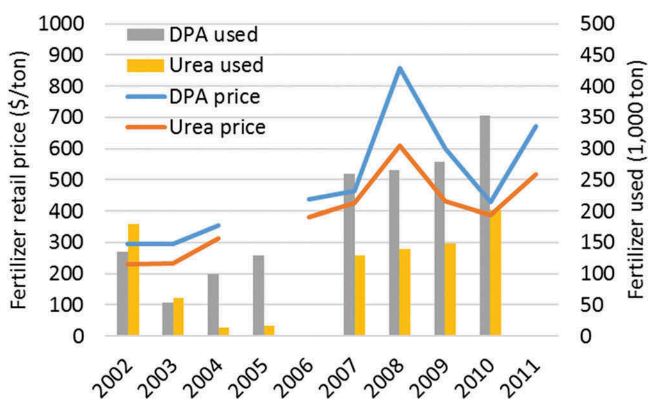


**Figure 4.** Fertilizer use and retail prices in Ethiopia. Sources: FAOSTAT and Rashid et al. (2013).

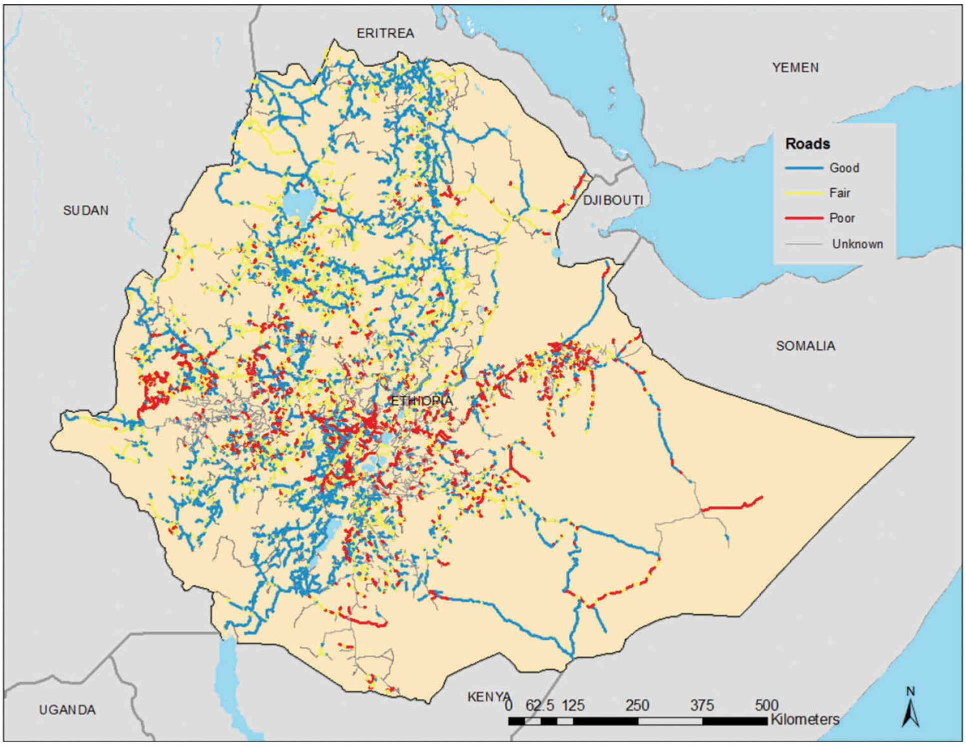


**Figure 5.** Road network in Ethiopia. Source: World Bank (2016) based on Ethiopia Road Authority data.

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% of fertilizer farmgate prices (Rashid et al., 2013). In Africa, timely delivery and application of right fertilizer are often identified as an important determinant of agricultural productivity. A recent field trial in Zambia, for instance, shows that crop yield can be improved by 6583%, depending on the composition of macronutrients in fertilizers (Chirwa, Mrema, Mtakwa, Kaaya, & Lungu, 2017).

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% 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% are in good condition (Figure 5).

## 3. 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, Wang, You, and Schmidt (2012)):

$$\ln V_{ijt} = \beta_0 + \ln X'_{ijt}\beta_X + Z'_{ijt}\beta_Z + \varepsilon_{ijt}, \tag{1}$$

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. As usual, the logarithms are taken for all these variables to mitigate their skewness of the distribution.

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.[2] This is an accepted practice, but it may cause significant bias. Particularly, in our case, the vast majority of households do not use advanced inputs. Thus, Battese's (1997) specification is incorporated:

$$\ln V_{ijt} = \beta_0 + \sum\nolimits_k \beta_k \ln x^*_{k,ijt} + \sum\nolimits_k \delta_k D_{k,ijt} + Z'_{ijt}\beta_Z + \varepsilon_{ijt}, \tag{2}$$

where

$$D_{k,ijt} = \begin{cases} 1 & \text{if } x_{k,ijt} = 0 \\ 0 & \text{if } x_{k,ijt} > 0 \text{ and } x^*_{k,ijt} = \max\left(x_{k,ijt}, D_{k,ijt}\right) \end{cases} \tag{3}$$

$x_{k,ijt}$ 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:

$$\begin{aligned} \ln V_{ijt} = \beta_0 + \beta_L \ln L_{ijt} + \beta_H \ln H_{ijt} + \beta_F \ln F^*_{ijt} + \delta_F D_{F,ijt} + \beta_S \ln S^*_{ijt} + \delta_S D_{S,ijt} \\ + \beta_R \ln R^*_{ijt} + \delta_R D_{R,ijt} + \beta_{TR} \ln TR_{jt} + Z_{ijt'}\beta_Z + c_j + c_t + \varepsilon_{ijt}, \end{aligned} \tag{4}$$

where $c_j$ and $c_t$ are district- and time-specific fixed effects. As will be discussed below, our data are not the 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 the years. The original data do not allow the identification of 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).[3] All 195,000 farmers are located in one of the districts.

[2]In the following analysis, the small positive number approach is also used to examine the robustness of our results.

[3]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.


**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 measurements, such as the Spatial Production Allocation Model (SPAM) which provides agricultural productivity data at a highly disaggregated level (approximately 10 × 10 km resolution).[4] However, such potential variables do not vary much over time, especially during our relatively short sample period (8 years). Therefore, given the advantage of having eight rounds of cross-sectional data, 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 to control various local unobservables, such as climatic and topographic conditions, and proximity to local markets.

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.[5] This is calculated by spatial software (ArcGIS) 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.

[4]See You, Wood, and Wood-Sichra (2009).
[5]Each district's location is spatially defined by its centroid.

As an identification strategy, the current paper takes advantage of a historical event that Ethiopia has experienced for our data period: 2003–10.[6] 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-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. In principle, road rehabilitation and maintenance works can 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% 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% 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 Arica (e.g., between Nairobi and Mombasa), the profit margins can reach 60%. 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% markup taken into account. In the latter case, rail and road transport costs are conceptually comparable.

Given the unit costs of road and rail transportation, an optimal route is identified to minimize the total transport cost from a given location to the port of Djibouti.[7] Spatial software, ArcGIS, is used to compute this. The estimated costs vary over the years, because of the phased-out rail services, implementation of a number of road projects and possible deterioration of road condition due to lack of maintenance (Figure 7). The underlying road and rail network data are different from one year to another. As shown in the figure, the changes in transport costs vary substantially across locations. However, there appeared to be a tendency that greater changes were experienced in the south than in the north. This is primarily because of the alignment of the rail line passing through the southern regions. The northern areas have more direct access to

[6]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.

[7]The figures depict transport costs with approximately 10 × 10 km resolution. For the empirical regression, transport costs are calculated for each of the 567 districts.

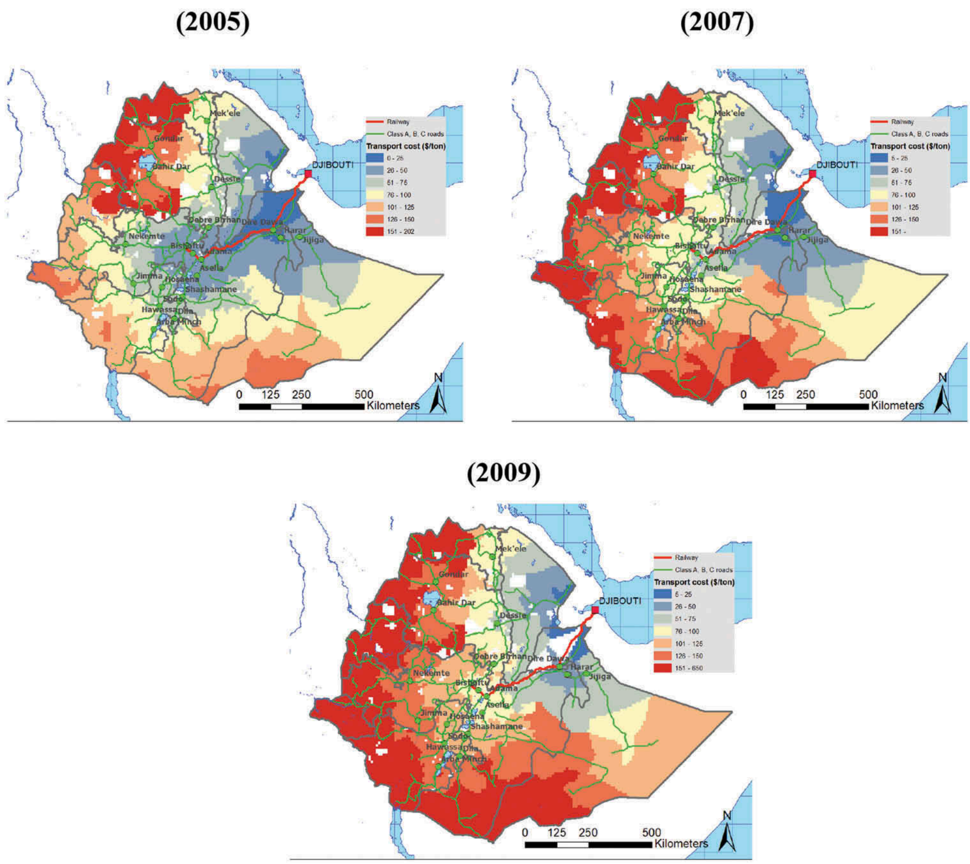


**Figure 7.** Transport costs to the Port of Djibouti (US$/ton). Source: Authors' calculation based on ERA data.

the port by road, e.g., Highway No 1 and 2, connecting Addis Ababa to the north and the northeast.[8]

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., Dercon, Gilligan, Hoddinott, & Woldehanna, 2009; Khandker et al., 2009; Qin & Zhang, 2016), 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. In reality, however, there may be some granularity at which people can benefit, rather than having access or no access.[9]

[8]This is an important geographic feather, which will be used to construct a binary variable to represent the treatment or beneficiary group in the following section.

[9]For robustness check purposes, in the following section, a binary dummy variable is also generated based on the traditional pipeline approach. The results do not change.

The most important empirical challenge to estimate Equation (4) is endogeneity of the transport variable, *TR* (e.g., Banerjee et al., 2012; Datta, 2012; Jedwab & 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 deal with this problem, one approach is the regression discontinuity design, which distinguishes the treatment and control groups based on an exogenous rating. For instance, Asher and Novosad (2018) use the population threshold for rural road programs. Unfortunately, however, there is no clear objective mechanism or framework to prioritize road or rail investment in Ethiopia. Thus, 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 & Koolwal, 2011), our data are not the 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, three instruments are constructed: One is the straight-line distance from the nearest historic port. An antique map by Bowen (1747) is used to identify the ports or landing sites that already existed in the horn of Africa before the mid-18th century, which include Zeila, Berbera, Mogadishu, Brava and Jubo (currently, called Kismayo). Naturally, transport infrastructure has been developed from these ports toward inland areas, therefore, affecting the development of the current transport network, but may not necessarily be influencing current agricultural production. This seems to be particularly true when the current complexity of political economy in the horn of Africa. To allow the variable to vary over time, fuel prices are multiplied, following the same technique as Storeygard (2016). The international oil price index comes from the IMF Commodity Price database. The intuition behind this instrument ($PORT_{jt}$) is that transport costs would likely be higher if the historic port is far and fuel prices are high. It is less likely to be correlated to the recent agricultural activities, but the validity as an instrument will be examined with our data.

The second instrument is the straight line 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% 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$) | *V* | 195,637 | 18.34 | 23.66 | 0.0002 | 117.65 |
| Transport cost to Djibouti ($/ton) | *TR* | 195,637 | 100.33 | 41.37 | 8.04 | 196.06 |
| Household size | *L* | 195,637 | 5.09 | 2.26 | 1 | 16 |
| Cultivated rainfed land area (ha) | *H* | 195,637 | 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* | 52,706 | 38.03 | 28.75 | 0.001 | 108.00 |
| Irrigated land area (ha) | *R* | 13,684 | 0.12 | 0.21 | 0.00004 | 3.81 |
| Dummy variable for households receiving extension services | *EXT* | 195,637 | 0.15 | 0.36 | 0 | 1 |
| Dummy variable for male head | *MAL* | 195,637 | 0.81 | 0.39 | 0 | 1 |
| Household head's highest grade completed | *EDU* | 195,637 | 2.51 | 2.81 | 0 | 25 |
| Share of own land (0 to 1) | *OWN* | 195,637 | 0.88 | 0.25 | 0 | 1 |
| Straight-distance to the operational rail line | *RAIL* | 195,637 | 425 | 243 | 0.06 | 1,057 |
| Difference in elevation from the nearest operational rail station | *ELEV* | 195,637 | 719 | 980 | −1,677 | 3,916 |
| Straight-distance to the nearest historic port, multiplied by the international oil price index | *PORT* | 195,637 | 77.10 | 36.79 | 10.32 | 193.31 |
| Transport cost to large city with population of 100,000 or more ($/ton) | *TR_city* | 195,637 | 19.61 | 12.91 | 0.17 | 83.81 |
| Total value of crops that could be potentially grown ($ million per ha) | *A* | 153,275 | 6.82 | 4.75 | 0.34 | 28.10 |

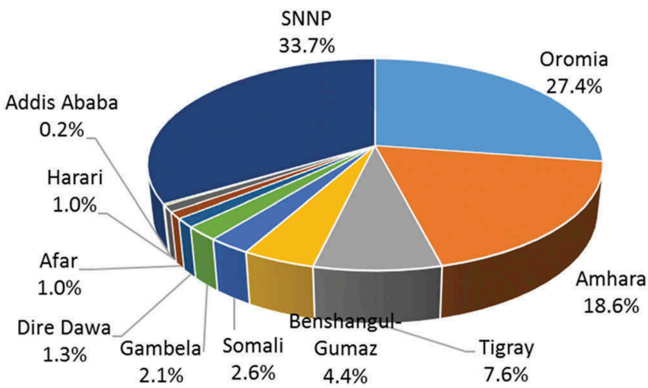


**Figure 8.** Surveyed households by region.

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% 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

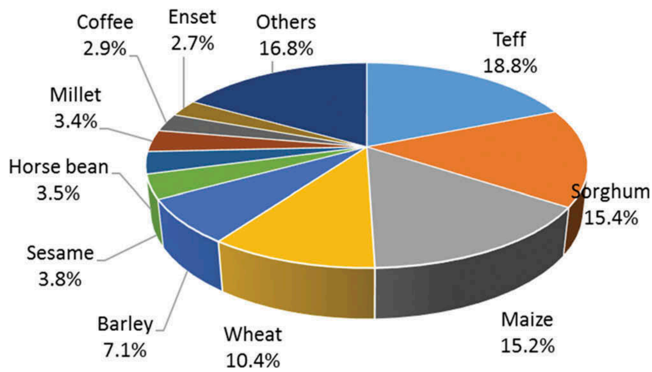


**Figure 9.** Cultivated land by crop in the sample data.

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., the volume of fertilizer used) and outputs (e.g., the quantity of maize harvested) are aggregated at the household level, and the household size is used as a proxy of labor inputted.[10]

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 uses fertilizer. Even if used, the volume used is minimal, on average 38.1 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.

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% of land cultivated is owned by farmers themselves. The rest is borrowed land.

## 4. Main estimation result

The instrumental variable (IV) regression is performed with the Battese's specification incorporated. The results are shown in Table 4. Regarding endogeneity, the IV result is broadly similar to the OLS estimate. However, there is a marked difference in several coefficients. In the OLS estimation, 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 227.94. Thus, the IV result is consistent, while the OLS is likely to be biased.

The validity of the three instruments constructed is also confirmed with our data. The overidentifying restriction test statistic is small and less than any conventional threshold. Hence, our instruments are valid. In the first-stage regression, two instruments have

[10]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.

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

| | OLS | | IV | |
|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. |
| ln$TR$ | -0.013 | (0.015) | -1.139 | (0.093)*** |
| ln$L$ | 0.034 | (0.005)*** | -0.071 | (0.019)*** |
| ln$H$ | 0.829 | (0.003)*** | 0.912 | (0.013)*** |
| ln$S$ | -0.001 | (0.002) | 0.031 | (0.015)** |
| ln$F$ | 0.010 | (0.003)*** | 0.071 | (0.022)*** |
| ln$R$ | 0.176 | (0.005)*** | 0.270 | (0.008)*** |
| $D_S$ | -0.173 | (0.011)*** | -0.038 | (0.031) |
| $D_F$ | 0.093 | (0.012)*** | 0.088 | (0.064) |
| $D_R$ | -1.041 | (0.019)*** | -1.607 | (0.043)*** |
| *EXT* | 0.040 | (0.006)*** | -0.096 | (0.023)*** |
| MAL | 0.059 | (0.006)*** | -0.035 | (0.030) |
| ln*EDU* | 0.026 | (0.003)*** | -0.129 | (0.017)*** |
| *OWN* | 0.121 | (0.008)*** | 0.092 | (0.029)*** |
| constant | 3.441 | (0.210)*** | -234.413 | (47.118)*** |
| Obs. | 195,637 | | 195,637 | |
| R-sq | 0.772 | | … | |
| Wald chi2 | | | 398774.8 | |
| No. of dummy var. | | | | |
| District | 450 | | 450 | |
| Year | 7 | | 7 | |
| Exogeneity test chi2 | | | 227.94*** | |
| Overidentifying restriction test chi2 | | | 4.3E-08 | |

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% level, respectively.

positive coefficients as expected (Table 5). The coefficient of the distance to the historic ports multiplied by oil prices (*PORT*) is positive and significant, implying that transport costs tend to be higher where the ports are distant and fuel costs are high. The distance to the nearest operational rail line (*RAIL*) is also significantly positive. It is estimated at 0.337 with a standard error of 0.001. Therefore, as expected, transport costs increased as the Ethio-Djibouti rail operations deteriorated in the 2000s.

An unanticipated result is the coefficient of the difference in elevation from the nearest station (*ELEV*), which turns out to be significantly negative at 0.032. It may be correlated to and dominated by other instruments. However, the evidence may be able to be interpreted to mean that railways are difficult to develop in steep and mountainous terrain, and instead, the road network has been improved. The latter development may dominate the former. Recall that our transport variable, *TR*, measures not only rail but also road connectivity to the port.

Regarding the impact of the deteriorated port connectivity, it is found significant in the IV estimation. The coefficient of $\ln TR$ is $-1.139$, which is statistically significant (Table 4). Thus, if the transport cost to the port is reduced by 1%, the agricultural production value would increase by about the same percentage. This is broadly consistent with earlier studies in Africa. Dorosh et al. (2012) estimate the elasticity of crop production with respect to transport connectivity at $-1.2$ to $-4.8$. Dercon et al. (2009) show that Ethiopian household consumption increases by 16% if access to the all-weather road is granted.

For other explanatory variables, labor is found to be unproductive. The coefficient is negative. Although our labor variable may be overestimated because it is measured by the total number of household members, this is a common characteristic of subsistence

**Table 5.** First stage regression for IV estimation.

| | Battese specification | | Small positive value specification | |
|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. |
| ln*RAIL* | 0.337 | (0.001)*** | 0.306 | (0.002)*** |
| *ELEV* | -0.160 | (0.009)*** | -0.081 | (0.010)*** |
| *PORT* | 4.155 | (0.023)*** | 4.183 | (0.024)*** |
| ln*L* | 0.773 | (0.349)** | 0.704 | (0.363)* |
| ln*H* | 0.063 | (0.181) | 0.185 | (0.187) |
| ln*S* | 1.311 | (0.177)*** | -0.167 | (0.073)** |
| ln*F* | -0.165 | (0.331) | 0.025 | (0.073) |
| ln*R* | -0.031 | (0.400) | 0.821 | (0.369)** |
| $D_S$ | 9.601 | (1.160)*** | | |
| $D_F$ | -1.392 | (1.106) | | |
| $D_R$ | -0.377 | (1.601) | | |
| *EXT* | -8.217 | (0.620)*** | -10.074 | (0.624)*** |
| MAL | -0.026 | (0.456) | -0.074 | (0.477) |
| ln*EDU* | 0.430 | (0.206)** | 0.490 | (0.213)** |
| *OWN* | 0.937 | (0.734) | 0.846 | (0.770) |
| constant | 2.659 | (0.004)*** | 2.753 | (0.004)*** |
| Obs. | 195,637 | | 195,623 | |
| R-sq | 0.980 | | 0.979 | |
| F statistic | 733425 | | 2.8E+05 | |
| No. of dummy var. | | | | |
| District | 450 | | 450 | |
| Year | 7 | | 7 | |

The dependent variable is the log of the transport cost, *TR*. Robust standard errors are shown in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.

farming in Africa. It is generally labor-intensive. Not many gains can be expected from adding more labor. The land is productive, and its large elasticity reflects that the agricultural growth during the sample period was largely explained by land expansion. Fertilizer, improved seeds, and irrigation also have positive elasticities. These are broadly consistent with the earlier evidence (e.g., Mekonnen et al., 2013; Spielman et al., 2012; Taffesse et al., 2012).

While the impacts of extension services and general education are inconsistent with our prior expectation, the land ownership is found to be important to improve agricultural production growth. This is consistent with some earlier studies, such as Melesse and Bulte (2015) in Ethiopia, though the opposite evidence can also be found in the literature (e.g., Jacoby & Minten, 2007; Sitko, Chamberlin, & Hichaambwa, 2014).

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:

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

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:

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

The latent variable of the input is denoted by $X^*$.

**Table 6.** IVTOBIT and TOBIT regression on fertilizer and seed use.

| Dependent var. | *F* | | *S* | | *S* | |
|---|---|---|---|---|---|---|
| | IVTOBIT | | IVTOBIT | | TOBIT | |
| Estimaton model | Coef. | Std.Err. | Coef. | Std.Err. | Coef. | Std.Err. |
| ln*TC* | -9.62 | (1.37)*** | -0.50 | (0.66) | -0.62 | (0.44) |
| *EXT* | 50.84 | (0.34)*** | 14.07 | (0.16)*** | 14.11 | (0.34)*** |
| MAL | 9.02 | (0.37)*** | 0.97 | (0.17)*** | 1.04 | (0.17)*** |
| ln*EDU* | 3.69 | (0.17)*** | 1.15 | (0.07)*** | 1.16 | (0.08)** |
| *OWN* | -11.82 | (0.53)*** | -0.24 | (0.25) | -0.23 | (0.23) |
| constant | 35.82 | (6.75)*** | -19.21 | (3.42)*** | -19.21 | (3.42)*** |
| Obs. | 195,637 | | 195,637 | | 195,637 | |
| Wald chi2 | 57604.5 | | 11270.2 | | | |
| Pseudo R2 | | | | | 0.127 | |
| No. of dummy var. | | | | | | |
| District | 450 | | 450 | | 450 | |
| Year | 7 | | 7 | | 7 | |
| Exogeneity test | 31.84*** | | 0.09 | | | |

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% level, respectively.

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 6). 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.

## 5. Discussion

From the methodological point of view, one may be concerned whether the above results are robust against the 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 7). 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 is unproductive, the land is productive and the most important production factor. The positive impacts of advanced inputs, including

Table 7. OLS and IV estimation results using the small positive number specification.

| | OLS | | IV | |
|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. |
| ln*TR* | -0.012 | (0.015) | -1.050 | (0.089)*** |
| ln*L* | 0.037 | (0.005)*** | -0.061 | (0.018)*** |
| ln*H* | 0.829 | (0.003)*** | 0.910 | (0.013)*** |
| ln*S* | 0.017 | (0.001)*** | 0.023 | (0.006)*** |
| ln*F* | -0.007 | (0.001)*** | 0.018 | (0.005)*** |
| ln*R* | 0.230 | (0.004)*** | 0.351 | (0.009)*** |
| *EXT* | 0.057 | (0.006)*** | -0.087 | (0.022)*** |
| MAL | 0.061 | (0.006)*** | -0.028 | (0.028) |
| ln*EDU* | 0.026 | (0.003)*** | -0.121 | (0.016)*** |
| *OWN* | 0.121 | (0.008)*** | 0.093 | (0.028)*** |
| constant | 3.413 | (0.210)*** | -223.743 | (45.317)*** |
| Obs. | 195,623 | | 195,623 | |
| R-sq | 0.772 | | … | |
| Wald chi2 | | | 401904.8 | |
| No. of dummy var. | | | | |
| District | 450 | | 450 | |
| Year | 7 | | 7 | |
| Exogeneity test chi2 | | | 213.87*** | |
| Overidentifying restriction test chi2 | | | 1.8E-06 | |

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% level, respectively.

fertilizer, improved seeds, and irrigation do not change, but the magnitude of the estimated coefficients may be underestimated. This might be a bias created by replacing all zeros with an artificial small positive number.

As discussed above, another issue is that transport costs based on vehicle operating costs may be the same as actual market transport service prices. With a 60% 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 8). 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.

One may also be concerned about omitted variables and exclusion restriction. Recall that our empirical specification includes a number of district-specific fixed-effects, which control for various unobserved local conditions. To see this, the same IV regression is performed excluding them.[11] The result is broadly unchanged (Table 9). Under this specification, it is possible to include some more independent variables representing local characteristics, such as land productivity. Agro-climatic crop yield, denoted by $A$, is used from SPAM.[12] Indeed, the crop suitability has a positive impact on actual production. Independently of that, transport costs still have a significantly negative impact. Note that unlike the main results shown above, the validity of IVs cannot be confirmed when the district-specific fixed-effects are excluded: The overidentifying restriction tests can be rejected. Therefore, there are potentially a lot of omitted variables, and the district-specific fixed-effects play an important role to control them.

[11]The zone-specific fixed-effects are still included. The zone is a higher administrative unit than the district in Ethiopia.
[12]In SPAM, some districts are estimated to have little agricultural productivity. Thus, the Battese's specification, i.e., $D_A$, is applied here as well.

**Table 8.** OLS and IV estimation results with transport service prices taken into account.

| | OLS | | IV | |
|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. |
| ln*TR** | -0.028 | (0.011)** | -0.197 | (0.014)*** |
| ln*L* | 0.034 | (0.005)*** | 0.033 | (0.005)*** |
| ln*H* | 0.829 | (0.003)*** | 0.830 | (0.003)*** |
| ln*S* | -0.001 | (0.002) | -0.001 | (0.002) |
| ln*F* | 0.010 | (0.003)*** | 0.010 | (0.003)*** |
| ln*R* | 0.176 | (0.005)*** | 0.175 | (0.005)*** |
| $D_S$ | -0.172 | (0.011)*** | -0.170 | (0.011)*** |
| $D_F$ | 0.093 | (0.012)*** | 0.093 | (0.012)*** |
| $D_R$ | -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)*** |
| ln*EDU* | 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. | 195,637 | | 195,637 | |
| R-sq | 0.772 | | 0.772 | |
| Wald chi2 | | | 396661 | |
| No. of dummy var. | | | | |
| District | 450 | | 450 | |
| Year | 7 | | 7 | |
| Exogeneity test chi2 | | | 234.37*** | |
| Overidentifying restriction test chi2 | | | 3.3E-06 | |

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% level, respectively.

**Table 9.** OLS and IV estimation without district fixed effect and instruments.

| | IV | | IV | | OLS | |
|---|---|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. | Coef. | Std.Err. |
| ln*TR* | -0.078 | (0.015)*** | -0.081 | (0.016)*** | -0.089 | (0.106) |
| ln*A* | | | 0.051 | (0.004)*** | 0.052 | (0.027)* |
| $D_A$ | | | 0.072 | (0.009)*** | 0.087 | (0.066) |
| ln*L* | 0.009 | (0.005)* | 0.010 | (0.005)** | 0.008 | (0.015) |
| ln*H* | 0.831 | (0.003)*** | 0.830 | (0.003)*** | 0.831 | (0.010)*** |
| ln*S* | -0.002 | (0.002) | -0.002 | (0.002) | -0.001 | (0.006) |
| ln*F* | 0.011 | (0.004)*** | 0.011 | (0.004)*** | 0.011 | (0.011) |
| ln*R* | 0.179 | (0.005)*** | 0.178 | (0.005)*** | 0.177 | (0.023)*** |
| $D_S$ | -0.202 | (0.011)*** | -0.201 | (0.011)*** | -0.196 | (0.044)*** |
| $D_F$ | 0.125 | (0.012)*** | 0.123 | (0.012)*** | 0.127 | (0.031)*** |
| $D_R$ | -1.044 | (0.020)*** | -1.042 | (0.020)*** | -1.035 | (0.117)*** |
| *EXT* | 0.067 | (0.006)*** | 0.066 | (0.006)*** | 0.065 | (0.019)*** |
| MAL | 0.061 | (0.006)*** | 0.061 | (0.006)*** | 0.061 | (0.011)*** |
| ln*EDU* | 0.036 | (0.003)*** | 0.035 | (0.003)*** | 0.035 | (0.007)*** |
| *OWN* | 0.125 | (0.008)*** | 0.127 | (0.008)*** | 0.128 | (0.020)*** |
| *RAIL* | | | | | -0.026 | (0.044) |
| *ELEV* | | | | | 0.109 | (0.375) |
| constant | 3.019 | (0.168)*** | 2.949 | (0.169)*** | 3.010 | (0.522)*** |
| Obs. | 195,637 | | 195,637 | | 195,637 | |
| R-sq | 0.7459 | | 0.7461 | | 0.746 | |
| Wald chi2 | 577691 | | 578928 | | | |
| No. of dummy var. | | | | | | |
| Zone | 59 | | 59 | | 59 | |
| Year | 7 | | 7 | | 7 | |
| Exogeneity test chi2 | 54.284*** | | 42.216*** | | | |
| Overidentifying restriction test chi2 | 789.81*** | | 784.31*** | | | |

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

For exclusion restriction, even the instrumental variables can be added as independent variables. The OLS is performed. The coefficients of these instrumental variables are found insignificant. Our instruments do not seem to influence crop production, ln$V$. More formally, the $F$-statistic to test the hypothesis that these coefficients are zero is estimated at 0.22. Therefore, our instrumental variables can be excluded. Of course, this is consistent with our main results above, which passed the overidentifying restriction tests, which examine both whether the instruments are uncorrelated with the error term, and whether the excluded instruments are correctly excluded from the estimated equation. All the indications are that the IV estimation with the district-specific fixed-effects is preferred and most reliable.

Finally, the advantage of using a continuous variable to measure transport accessibility, i.e., $TR$, is abandoned. The treatment and control groups are defined by a binary number in an ordinary manner. Thus, the identification strategy is the traditional pipeline method, though actual beneficiaries (B) are reduced, not added. Based on the first-year data, households are classified into the treatment group ($D$) if transport costs including railways are lower than road only transport costs. Households living in the same treatment areas are classified into the treatment group in the following years. This is because if people did not benefit from rail transport in the first year, they would not benefit in the following years, either. Note that the rail transport availability continued diminishing over the survey period. It was never improved.

Again, the results support the impact of rail transportation. With the district-specific fixed-effects, the coefficient of $B$ is estimated at 0.713, which is significant: Rail beneficiaries produced more crops (Table 10). This is a preferred model. On the other hand, when the district fixed-effects are deleted, the treatment group dummy ($D$) can be

**Table 10.** IV estimation under pipeline approach.

| | IV | | IV | |
|---|---|---|---|---|
| | Coef. | Std.Err. | Coef. | Std.Err. |
| $B$ | 0.713 | (0.053)*** | 0.340 | (0.039)*** |
| $D$ | | | 0.146 | (0.162) |
| ln$L$ | -0.007 | (0.008) | 0.009 | (0.005)* |
| ln$H$ | 0.863 | (0.006)*** | 0.832 | (0.003)*** |
| ln$S$ | 0.011 | (0.006)* | -0.002 | (0.002) |
| ln$F$ | 0.030 | (0.009)*** | 0.010 | (0.004)*** |
| ln$R$ | 0.223 | (0.005)*** | 0.178 | (0.005)*** |
| $D_S$ | -0.123 | (0.016)*** | -0.203 | (0.012)*** |
| $D_F$ | 0.093 | (0.026)*** | 0.131 | (0.013)*** |
| $D_R$ | -1.324 | (0.025)*** | -1.044 | (0.020)*** |
| $EXT$ | -0.018 | (0.010)* | 0.061 | (0.006)*** |
| MAL | 0.023 | (0.012)* | 0.061 | (0.006)*** |
| ln$EDU$ | -0.038 | (0.007)*** | 0.037 | (0.003)*** |
| $OWN$ | 0.111 | (0.013)*** | 0.127 | (0.008)*** |
| constant | -86.969 | (17.515)*** | 2.268 | (0.213)*** |
| Obs. | 195,637 | | 195,637 | |
| R-sq | … | | 0.745 | |
| Wald chi2 | 483813 | | 577319.2 | |
| No. of dummy var. | | | | |
| District | 450 | | 59 | |
| Year | 7 | | 7 | |
| Exogeneity test chi2 | 164.6*** | | 44.51*** | |
| Overidentifying restriction test chi2 | 8.7E-08 | | 663.005*** | |

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% level, respectively.

included. The result is similar; however, the overidentifying restrict test is rejected. Still, the significance of the positive coefficient of $B$ is unchanged. However, as discussed above, the overidentifying restriction test can easily be rejected.

## 6. 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 a light on the possible impacts of rail transport on agricultural production. It is a methodological challenge to estimate an unbiased impact of large-scale infrastructures, 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 1.139. Thus, a 1% reduction in transport costs to the port would increase agricultural production by about the same percentage. 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.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Notes on contributors

*Atsushi Iimi* is senior economist of Transport Global Practice of the World Bank. He holds a Ph. D. in economics from Brown University. Before joining the Bank, he also worked at IMF and JICA/OEFC, Japan. His research interests include industrial organization, spatial analysis and transport economics. His research on these topics has been published in scholarly journals, such as the *Review of Industrial Economics*, *Journal of Urban Economics*, *Journal of Applied Economics*, the *Development Economies*, and *IMF Staff Papers*.

*Haileysus Adamtei* is senior highway engineer of Transport Global Practice of the World Bank.

*James Markland* is senior transport specialist specialist of Transport Global Practice of the World Bank.

*Eyasu Tsehaye* is Transport Consultant of the World Bank.

## References

Amin, M., Willetts, D., & Matheson, A. (1986). *Railway across the equator: The story of the east african line*. London: The Bodley Head Ltd.

Asher, S. E., & Novosad, P. 2018. Rural roads and local economic development. Policy Research Working Paper No. 8466. World Bank.

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

Battese, G. (1997). A note on the estimation of Cobb-Douglas production functions when some explanatory variables have zero values. *Journal of Agricultural Economics*, *48*(2), 250–252.

Bowen, E. (1747). A new and accurate map of Nubia and Abissinia, together with all the kingdoms tributary.

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

Chirwa, M., Mrema, J. P., Mtakwa, P. W., Kaaya, A., & Lungu, O. I. (2017). Yield response of groundnut (Arachis hypogaea L.) to boron, calcium, nitrogen, phosphorus and potassium fertilizer application. *International Journal of Soil Science*, *12*(1), 18–24.

Datta, S. (2012). The impact of improved highways on Indian firms. *Journal of Development Economics*, *99*(1), 46–57.

Dercon, S., Gilligan, D., Hoddinott, J., & Woldehanna, T. (2009). The impact of agricultural extension and roads on poverty and consumption growth in fifteen Ethiopian villages. *American Journal of Agricultural Economics*, *91*(4), 1007–1021.

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

Donaldson, D. (2018). Railroads and the Raj: The economic impact of transportation infrastructure. *American Economic Review, 108*, (4–5), 899–934.

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

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

Gwilliam, K. (2011). *Africa's transport infrastructure: Mainstreaming maintenance and management*. Washington, DC: The World Bank.

Gyimah-Brempong. (1987). Scale elasticities in Ghanaian cocoa production. *Applied Economics*, *19*, 1383–1390.

Jacoby, H., & Minten, B. (2007). Is land titling in Sub-Saharan Africa cost-effective? Evidence from Madagascar. *World Bank Economic Review*, *21*(3), 461–485.

Jedwab, R., & Moradi, A. 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, S., Bakht, Z., & Koolwal, G. (2009). The poverty impact of rural roads: Evidence from Bangladesh. *Economic Development and Cultural Change*, *57*(4), 685–722.

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

Mekonnen, D. K., Jeffrey, D., & Greg, F., 2013. Productivity and efficiency of small scale agriculture in Ethiopia, Southern Agricultural Economics Association 2013 Annual Meeting, February 2-5, 2013, Orlando, Florida.

Melesse, M., & Bulte, E. (2015). Does land registration and certification boost farm productivity? Evidence from Ethiopia. *Agricultural Economics*, *46*, 757–768.
Olievschi, V. N. 2013. Railway Transport: Framework for improving railway sector performance in Sub-Saharan Africa. SSATP Working Paper No. 94, Africa Transport Policy Program.
Qin, Y., & Zhang, X. (2016). The road to specialization in agricultural production: Evidence from rural China. *World Development*, *77*, 1–16.
Rashid, S., Nigussie, T., Nicholas, M., & Gezahengn, A. 2013. Fertilizer in Ethiopia. IFPRI Discussion Paper 01304, International Food Policy Research Institute.
Sitko, N., Chamberlin, J., & Hichaambwa, M. (2014). Does smallholder land titling facilitate agricultural growth?: An analysis of the determinants and effects of smallholder land titling in Zambia. *World Development*, *64*, 791–802.
Spielman, D., Mekonnen, D., & Alemu, D. (2012). Seed, fertilizer, and agricultural extension in Ethiopia. In P. Dorosh and S. Rashid (Eds.), *Food and agriculture in Ethiopia: Progress and policy challenges* (pp. 84–122. Philadelphia: University of Pennsylvania Press.
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.
Taffesse, A. S., Dorosh, P. A., & Gemessa, S. A. (2012). Crop production in Ethiopia: Regional patterns and trends. In P. A. Dorosh & S. Rashid Chapter 3 Eds., *Food and agriculture in Ethiopia: Progress and policy challenges* (pp. 53–83). Philadelphia: University of Pennsylvania Press.
Teravaninthorn, S., & Raballand, G. (2009). *Transport prices and costs in Africa: A review of international corridors*. Washington, DC: The World Bank.
World Bank. (2013). *Growing Africa: Unlocking the potential of agribusiness*. Washington, DC: Author.
World Bank. (2016). *Measuring Rural* Access: Using New Technologies. Washington DC: The World Bank.
Xu, Z., Zhengfei Guan, T. S., & Jayne, R. B. (2009). Factors influencing the profitability of fertilizer use on maize in Zambia. *Agricultural Economics*, *40*, 437–446.
You, L., Wood, S., & Wood-Sichra, U. (2009). Generating plausible crop distribution and performance maps for Sub-Saharan Africa using a spatially disaggregated data fusion and optimization approach. *Agricultural System*, *99*(2–3), 126–140.