American Economic Journal: Macroeconomics 2025, 17(1): 343–378
https://doi.org/10.1257/mac.20200149

Spatial Integration and Agricultural Productivity: Quantifying the Impact of New Roads†

By Tasso Adamopoulos*

I study the effects of Ethiopia's 1997–2014 road expansion program on agricultural productivity and structural change by combining a quantitative spatial sectoral framework with novel district-level panel data on agricultural production and geocoded transport costs. In the model, the spatial heterogeneity of transport costs affects the distribution of production and mobile inputs across locations and sectors, and the allocation of land across crops within locations. Varying transport costs to their new actual levels, the model delivers substantial structural change, a rise in agricultural productivity one-tenth of the data, and a pattern of productivity gains across districts consistent with the data. (*JEL* H54, O13, O18, Q11, Q13, R42, R53)

The process of development is accompanied by a process of structural change, whereby economic activity shifts from agriculture to the rest of the economy. Agricultural productivity plays a key role in this transition, particularly for countries at early stages of economic development.[1] A key challenge is understanding what the fundamental drivers of structural change and agricultural productivity are, and quantifying them.

This paper focuses on transportation infrastructure as a driver of structural change. A characteristic of low-income countries is that they have poor transport infrastructure and high internal transportation costs,[2] which operate as a major impediment for farmers in accessing both domestic and international markets. Further, the spatial heterogeneity of transportation costs within developing countries can have implications for what crops are produced, where they are produced, and with what inputs, all of which can impact agricultural productivity and induce structural change.

* Department of Economics, York University (email: aadamo@yorku.ca). Virgiliu Midrigan was coeditor for this article. This paper supersedes a previous version circulated under the title "Spatial Integration, Agricultural Productivity, and Development: A Quantitative Analysis of Ethiopia's Road Expansion Program." The GIS consultant Mohammed Nuru as well as Stein Monteiro provided superb research assistance. I thank Simon Alder, Andrew Dickens, Doug Gollin, Berthold Herrendorf, Gustavo Ventura, Pramilla Krishnan, and Theodore Papageorgiou for helpful comments. I have also benefited from comments at York University, the SED in St. Louis, the LAEF Conference on Structural Change in Santa Barbara, and the National University of Singapore. I gratefully acknowledge financial support from the International Growth Centre, and thank Alemayehu Seyoum Taffesse for facilitating access to Ethiopia's Agricultural Sample Surveys. All potential errors are my own.

† Go to https://doi.org/10.1257/mac.20200149 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] There is a large literature that emphasizes the importance of agriculture for development and cross-country income differences—e.g., Schultz (1953); Gollin, Parente, and Rogerson (2002); Restuccia, Yang, and Zhu (2008); Caselli (2005).

[2] See, for example, Adamopoulos (2011).

I quantify the aggregate and local structural change and productivity effects of a major transport infrastructure intervention that altered the spatial distribution of transport costs. In particular, I study Ethiopia's comprehensive road expansion program over 1997–2014, which resulted in a major overhaul of the country's network in terms of both volume and quality. To quantify the gains from improved market access on agricultural productivity, I combine a novel district-level panel dataset with a spatial model of agricultural production and structural change. I assemble the district-level panel over 1996–2014 by overlaying agricultural production data with geocoded transport costs between agricultural production sites and crop markets. The spatial model of agricultural productivity features local land frictions and location-good-specific transport costs, which affect the distribution of production across locations, crop choice, the allocation of labor across locations and sectors, and intermediate input use. I find that the change in the distribution of actual transport costs leads to structural change, with the share of employment in agriculture dropping 5.5 percentage points and the aggregate real yield increasing by 14.7 percent, about one-tenth of the actual yield gain over 1996–2014. In addition, there is a restructuring of the agricultural sector, with a shift in the composition of crops produced toward cash crops, and an increase in average farm size. Similar to the data, I find that local yield gains exhibit a U-shaped pattern with respect to changes in transport costs. I attribute this pattern to the (mis)alignment of changes in absolute and comparative advantages implied by the transport cost changes, and the associated specialization of districts.

In the late 1990s, Ethiopia was a low-income country, with its economy heavily skewed toward agriculture and low agricultural productivity. At the same time, Ethiopia had very low road network density and high domestic transport costs.[3] These characteristics were shared by many other developing countries, particularly in sub-Saharan Africa. Once Ethiopia embarked on its road expansion program in 1997, the density and quality of its road network saw major improvements, and its agricultural productivity surged. This paper measures the contribution of road improvements to the surge in Ethiopia's productivity over 1996–2014, using micro-level data and a structural model.

The micro-level data allow me to construct a district-level panel over 1996–2014, consisting of an agricultural production component and a geographic component. The agricultural production component of the panel draws from repeated waves of household-level data from the Ethiopian Agricultural Sample Surveys, on the type and quantity of crops produced, land allocations by crop, and input use. For the geographic component, I estimate travel times between district centers and crop markets, using detailed geographic information system (GIS) data on the road infrastructure network at each point in time and high-resolution data on the topography of the terrain that has to be traveled to reach the relevant market. I find that in 1996 transport costs are on average very high and exhibit substantial spatial heterogeneity within Ethiopia. By 2014, there has been a considerable drop in the level—36 percent, on average—and the dispersion of transport costs.

[3]Based on data from Adamopoulos (2011) and the World Bank's World Development Indicators.

I first use the panel data to provide empirical motivation for the effect of improvements in market access. I employ a two-way fixed effects difference-in-differences design that leverages the staggered rollout of the road development program and the heterogeneous exposure to road development across districts in Ethiopia. The specification includes district and year fixed effects. The identification strategy compares two districts that are identical in time-invariant characteristics but differ in the amount and timing of additional market access. I find evidence that districts with improved access to grain markets experienced significant increases in productivity, with higher grain yields, more fertilizer use, and increased specialization in grains. The baseline difference-in-differences estimate indicates that a 1 percentage point increase in market access resulted in a 0.1 percentage point increase in the average yield over all grains.

Motivated by the empirical evidence on the relative local effects of improved market access, I develop a spatial equilibrium model of structural change to quantify, from a macroeconomic perspective, the aggregate-level general equilibrium impact of the road development program. I build into the model empirically relevant channels at the district level, such as intermediate input use and crop choice, as well as the economy-wide mechanisms of the reallocation of agricultural production across space according to comparative advantage and the reallocation of labor across sectors. The model allows me to assess not only the broader overall effects of changes in the actual entire distribution of transport costs across districts in Ethiopia but also to run additional counterfactual experiments and quantify the role of the different model mechanisms.

In particular, I develop an equilibrium model of structural change featuring an urban center and multiple rural agricultural production locations. Each rural location can produce a food crop for domestic consumption or a cash crop for the export market. Consumers in the urban location have nonhomothetic preferences over the consumption of food and the nonagricultural goods produced in the urban center. Shipments of crops to the urban center for consumption or export are subject to domestic, crop-location-specific transportation costs. Transport costs also raise the cost of disbursing imported intermediate inputs from the urban center to the multiple rural locations. The food farming technology also requires labor. The model features frictions to the allocation of land across crops, within locations. Changes in the distribution of good- and location-specific transportation costs alone reallocate food production across locations and alter the allocation of land across crops within locations as well as the distribution of intermediate input use and labor across locations, inducing structural change and productivity effects at the aggregate and local level.

To isolate the effects of transport cost changes over 1996–2014 on productivity, my quantitative approach involves three steps. First, I calibrate the spatial production structure of the model to aggregate and district-level data for the Ethiopian economy for 1996, before the road program began. The district-specific land distortions are identified from the observed allocation of land across crops within districts. To assess the appropriateness of the calibrated model, I show that its predictions in terms of relative food yield gains, once perturbed with the change in food transport costs alone, are in line—in sign and magnitude—with the empirically estimated ones from the data, using the same difference-in-differences specification. Next, keeping all else equal, I feed into the model exogenously the entire distribution of the observed changes in transportation costs across goods and locations, implied by my data on the actual changes in the volume and quality of the road network in Ethiopia as of 2014. I then compare the equilibrium changes implied by the model with all transport cost changes to the actual level changes in the data over the period 1996–2014, in terms of both aggregate metrics and spatial distributional patterns across districts. I note that a location in the model directly maps into a district in the data. As a result, I do not have to rely on parametric distributions for transportation costs or productivities.

The model implies substantial structural change, with a drop in the employment share in agriculture by 5.5 percentage points and an increase in the aggregate yield by 14.7 percent. This number is 20 percent higher if the direct resource savings from lower transport costs are taken into account. To appreciate the magnitude of these gains, I note that they account for about 10 percent of the overall yield gain experienced by Ethiopia over the period 1996–2014. These changes result in a substantial increase in agricultural value added per worker, by 23.4 percent. In terms of the mechanism, as transport costs fall overall, food production is increasingly undertaken by relatively more productive rural districts with a corresponding reallocation of labor across space. Given that the demand for food is inelastic, this allows for an overall shift of land to cash crops and an overall shift of labor toward nonagriculture, with an associated increase in average farm size. These changes encapsulate the structural transformation of the economy induced by the transport cost changes and, given the size of the agricultural sector, imply substantial gains in aggregate income, with real GDP per worker increasing 22 percent. Quantitatively, the two key mechanisms of the reallocation of production across space/crops and the reallocation of labor across space/sectors contribute roughly equally to the magnitude of the agricultural productivity effects, with the intermediate input use channel playing a smaller role.

In terms of local outcomes, I find that the distribution of the gains is uneven across districts. The model delivers a U-shaped pattern of district-level yield gains with respect to food transport cost changes across districts, a relationship that is also present in the data. In the model, among districts that are completely specialized in food crops, the biggest gains are experienced by those that observe the largest drops in their transport costs. For these districts, changes in the level of their food transport costs and their relative food-to-cash transport costs are strongly aligned. Among districts that produce both crops (incompletely specialized), the largest gains are experienced by those with the smallest change in the level of their food transport costs. For these districts, while the level of their food transport costs falls, their relative food-to-cash transport costs tend to increase.

Structural change, the shift from agriculture to manufacturing and services, is a ubiquitous feature of the process of development (Herrendorf, Rogerson, and Valentinyi 2014). There is a large literature in growth and development that studies mechanisms generating a structural transformation of the economy—e.g., Kongsamut, Rebelo, and Xie (2001); Ngai and Pissarides (2007); Acemoglu and Guerrieri (2008); Boppart (2014); and Comin, Lashkari, and Mestieri (2021). However, this literature does not specifically focus on the role of transport infrastructure, as I do here. A related literature shows that agriculture plays a key role in understanding productivity disparities across countries (Gollin, Parente, and Rogerson 2002; Restuccia, Yang, and Zhu 2008; Caselli 2005).[4]

A recent literature in macroeconomics shows that internal transport costs matter at the aggregate level for development and the sectoral composition of the economy: Adamopoulos (2011); Herrendorf, Schmitz, and Teixeira (2012); and Gollin and Rogerson (2014). This paper contributes to this literature by overlaying microdata on farm production and detailed geocoded market access data to evaluate the impact of a particular road expansion program.

This paper relates to a large literature studying the economic impacts of transport infrastructure investments in the form of roads, highways, or railroads. One strand of the literature uses general equilibrium trade or economic geography models to measure the effects of transport infrastructure projects—e.g., Donaldson (2018); Donaldson and Hornbeck (2016); Allen and Arkolakis (2014); Alder (2023); Asturias, García-Santana, and Ramos (2019); and Jaworski, Kitchens, and Nigai (2023), among others. A more recent literature studies the welfare impact of changes in the transportation network in a general equilibrium setting (Allen and Arkolakis 2022; Fajgelbaum and Schaal 2020; Felbermayr and Tarasov 2022). None of these papers, however, focus on agriculture or structural change per se. A related empirical literature estimates local effects of transport infrastructure expansion—e.g., Banerjee, Duflo, and Qian (2012); Faber (2014); Baum-Snow et al. (2017); Storeygard (2016); and, most closely related to this paper in context, Shamdasani (2021) and Asher and Novosad (2020), who find, respectively, that rural roads in India improve agricultural technologies and lead to a reallocation of workers out of agriculture. A key characteristic of this empirical literature is the use of credible identification strategies to address the potential endogeneity of the placement of the relevant transport infrastructure and estimate its causal effects. I also empirically estimate the effects of new roads but examine their aggregate impact using a structural model as well.

A recent micro-to-macro development literature uses general equilibrium structural models to study the macroeconomic effects of well-identified micro-level interventions—e.g., Fried and Lagakos (2022); Lagakos, Mobarak, and Waugh (2023); Buera, Kaboski, and Shin (2021); and, notably for this paper, Brooks and Donovan (2020), who study the role of new footbridges in integrating rural communities to labor markets in Nicaragua. I study the relative local effects and the aggregate macroeconomic impact of a large-scale, economy-wide road development program.

This paper is most closely related to two notable papers, Costinot and Donaldson (2016) and Sotelo (2020), who also employ multiregion spatial frameworks that

[4] The importance of agriculture for development has been emphasized in the earlier development literature—e.g., Schultz (1953). Developing countries are much more unproductive in agriculture than in nonagriculture when compared to developed countries and, in addition, employ most of their labor in agriculture. An important challenge for policy and academic research alike is to understand why agricultural productivity is so low in developing countries. There are several recent contributions in the macrodevelopment literature that study this question—among many others, Lagakos and Waugh (2013); Adamopoulos and Restuccia (2014); Gollin, Lagakos, and Waugh (2014); Tombe (2015); Donovan (2021); and Adamopoulos and Restuccia (2022). This paper contributes to this literature by studying a distinct factor: the importance of farm connectivity to markets.

link domestic trade frictions with agricultural productivity and welfare when factors are allocated on the basis of comparative advantage. In addition, Sotelo (2020) examines the effects of counterfactual changes in the infrastructure policy in Peru. The key difference from this literature is that I quantify the effects of new roads on the structural transformation out of agriculture. To do this, I depart from the literature by developing a model of structural change with multiple locations, featuring nonhomothetic preferences. This allows me to speak to channels that the macrodevelopment literature has emphasized, such as the reallocation of labor across sectors, average farm size, input intensity, and sectoral productivity. In addition, I allow for frictions to the allocation of land in local markets, which have been shown to be important for low-income countries—e.g., Adamopoulos and Restuccia (2014)—and consider an extension with barriers to the mobility of labor across space.

The paper proceeds as follows. Section I outlines Ethiopia's road development program. Section II describes the assembly of the panel dataset and empirically estimates the productivity effects of roads. The spatial framework is developed in Section III. I calibrate the model to aggregate and district-level moments from the Ethiopian data in Section IV. Section V reports the aggregate and distributional effects from the quantitative experiments. I conclude in Section VI.

I. Ethiopia's Road Development Program

For more than two decades, Ethiopia embarked on an extensive road development program as a pillar of its growth strategy. Starting in 1997, through the implementation of successive phases of the Road Sector Development Programme (RSDP), there has been substantial improvement in the volume and distribution of the road network, as well as in the conditions of the existing roads. The first phase, RSDP I, covered the period 1997–2002; the second, RSDP II, covered 2002–2007; RSDP III covered 2007–2010; and RSDP IV covered the period 2010–2015. The RSDPs consisted of constructing new roads and rehabilitating/upgrading existing roads. The road development program was comprehensive and covered federal, regional, rural, and district roads. The first three phases primarily emphasized federal and regional roads. The more recent Universal Rural Road Access Program was a major component of the RSDP IV and emphasized rural and district roads, aiming to connect all lower administrative units in rural areas to all-weather roads. While the bulk of the overall financing came from the Ethiopian government, the World Bank was a major partner, with other governmental and nongovernmental donors also contributing.

These efforts had a substantial impact on the extent and quality of the road network in Ethiopia. The volume of the total network increased almost threefold, from 24,970 kilometers in 1997 to 69,951 kilometers in 2014. However, the volume increase in the rural road network has been 4.7-fold (from 9,100 in 1997 to 43,094 kilometers in 2014). The road density (including community roads) over 1997–2016 increased from 24.1 kilometers per 1,000 squared kilometers to 102.8 kilometers per 1,000 squared kilometers and from 0.46 kilometers per 1,000 people to 1.23 kilometers per 1,000 people. In terms of qualitative indicators, the proportion of asphalt roads in good condition increased from 17 percent in 1997 to 73 percent in

2016. The proportion of rural roads in good condition increased from 21 percent to 55 percent over the same period (ERA 2016). This has also had a large impact on traffic and mobility. The vehicle kilometers of traffic increased from 3.8 million in 1997 to 18.9 million by 2015.

The main federal authority responsible for new construction and rehabilitation projects throughout all the RSDP phases was the Ethiopian Roads Authority (ERA). Regional authorities usually proposed potential road projects that were then evaluated against the ERA's broad criteria/guidelines. In terms of project selection, there is no single well-defined blueprint that was followed to determine which projects were undertaken. Key determinants for earmarking road construction projects were access to large isolated rural populations, improvement of the overall efficiency of the network, and upgrading high-traffic roads and those in poor condition. While economic considerations were taken into account, according to the ERA criteria, established areas, areas with economic potential, and underprivileged areas were prioritized (ERA 2016). The particular route of an earmarked road project was then determined, taking into account engineering, topographical, and environmental considerations as well as the costs and budgetary constraints of the ERA (ERA 2013). A more detailed description of the broad criteria used for the preliminary selection of road projects is provided in online Appendix A.

II. Data and Empirical Analysis

To assess the effect of Ethiopia's road infrastructure expansion program, I construct a panel of agricultural production data and effective travel times from agricultural production sites to agricultural markets, estimated from the observed road network over time. In this section, I briefly describe the data and how I use it to construct the panel. A more detailed description of the data assembly is provided in online Appendix B. Before moving on to the model, I use the panel data to empirically estimate the effect of the improved travel times on agricultural productivity, employing a difference-in-differences framework.

A. *Roads Data*

I use administrative GIS data on the universe of roads in Ethiopia, starting in 1996, just before the program began, until 2014, biennially (ERA 1996–2014).[5] The road network data provide information not only on the volume but also on the quality of every link in the network, each year. The data come with information on road class and surface type (e.g., whether a particular road is a highway or a town road, and whether a town road is dirt, asphalt, etc.), year of construction, as well as year of upgrading or rehabilitation.

Panel A in Figure 1 shows the road network in 1996, before the comprehensive road expansion program. Panel B in Figure 1 provides a map of Ethiopia's entire road network as of 2014, indicating both the new links in the network (blue) as well

[5] The road network data in vector form are obtained from the ERA for highways and regional roads.

as the links of the pre-1996 network that have been rehabilitated or upgraded by 2014. A casual inspection of the two maps shows a substantial expansion in the volume and quality of the network over the period 1996–2014, especially with respect to feeder roads and roads reaching rural dispersed communities. These road data are the main ingredient going into the estimation of the geocoded transportation costs, outlined below.

B. *Geocoded Transportation Costs*

The goal is to estimate geocoded transportation costs from agricultural production sites to agricultural markets. The spatial unit of observation is taken to be a district, or "woreda."[6] The measure of transportation costs I use in the analysis is the travel time in minutes between the district centroids and the nearest destination crop markets. For food crops (cereals), the possible destinations where output can be disbursed are taken to be Ethiopia's 33 major wholesale grain markets (obtained from the Ethiopian Grain Trading Enterprise), which are spread throughout Ethiopia. The food crop travel time for each district is the travel time to the nearest grain market. For cash crops, which are primarily destined for exporting via the capital, the destination market for estimating the domestic travel time is Addis Ababa.

To estimate a panel of travel times from districts to destination crop markets, I overlay the universe of the actual road network data by year described above, with high-resolution geographic data on elevation and land use, along with the GPS coordinates of the district centroids and the destination crop markets. The layer of geographic data on land use and elevation is used to obtain as precise geocoded estimates of travel time as possible by taking into account the topography of the terrain that has to be traveled to reach the relevant market. Online Appendix B.1 describes in more detail how the geographic data are used to estimate the panel of travel times.

Table 1 presents summary statistics for the estimated geocoded transport costs. There are two points to note. First, average effective travel times from district centroids to grain markets dropped from 345 minutes in 1996 to 220 minutes in 2014, a −36 percent change. The travel times from district centroids to Addis Ababa are higher in level, in both 1996 and 2014, dropping by 24 percent over the period. Second, the dispersion of transport costs across districts dropped, implying better accessibility to markets for more districts. The share of districts within four hours of a major grain market increased from 0.50 in 1996 to 0.71 in 2014. The share of districts within four hours of the capital of Addis Ababa started from the lower level of 0.11 and increased to 0.20.

C. *Agricultural Production Data*

I use household-level data from the Ethiopian Agricultural Sample Survey (AgSS) over the period from 1995/96 to 2014/15. I use these data to construct a district-level panel on agricultural production, land allocations across crops, and

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

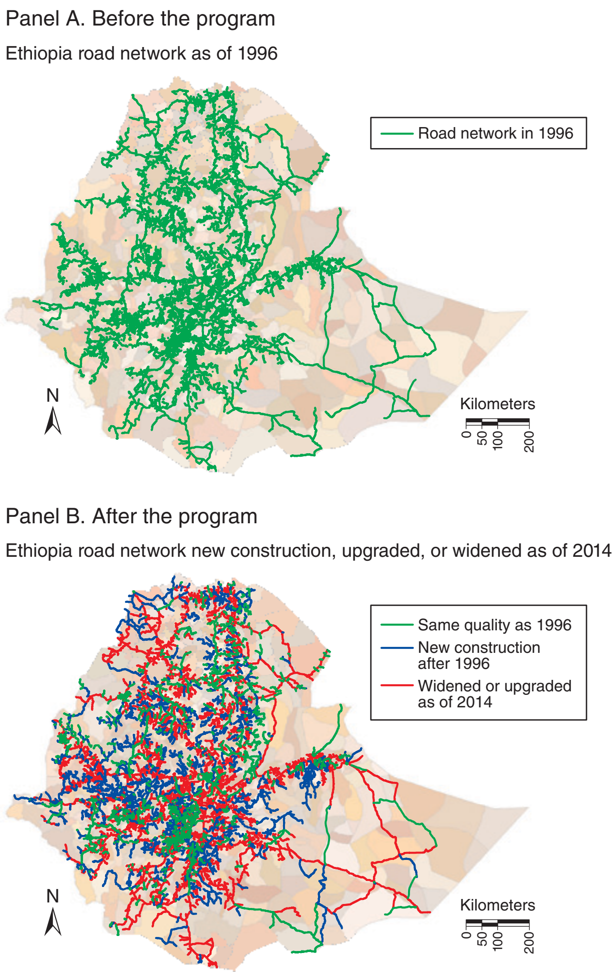

Figure 1. Roads in Ethiopia

Notes: With data from the ERA. In panel B, the network links in blue represent newly constructed roads after 1996; the network links in red represent rehabilitation or quality upgrade of pre-1996 network links.

intermediate input use (fertilizer). The district, or woreda, is the lowest level of spatial disaggregation for which a reliable panel could be constructed. In online Appendix B.2 I detail how I assemble the district-panel, merging the AgSS data over a long number of years and harmonizing district codes.

The district-level measure of agricultural productivity I emphasize in the data is the real yield or land productivity, measured as real output per hectare.[7] To construct a real measure of yield over a basket of crops, I aggregate using, as a common set of

[7] The AgSS data do not report the amount of labor to compute labor productivity at the district level.

Table 1—Summary Statistics of Estimated Travel Time in Minutes

	To nearest grain market		To Addis Ababa	
	1996	2014	1996	2014
Mean travel time in minutes	345.0	220.1	574.5	438.5
Median travel time in minutes	240.7	164.7	550.2	418.9
Fraction of districts $<$ 4 hours	0.499	0.715	0.114	0.203

Notes: Author calculations based on estimated district-level geocoded transportation cost data. Summary statistics are reported for the balanced panel of 403 districts for which both agricultural production (AgSS) data and estimated transport cost data are available. The first two columns report statistics for estimated travel times from the district centroids to the nearest grain markets. The last two columns report statistics for the estimated travel times from the district centroids to the capital of Addis Ababa.

prices across districts, the average prices for each crop over the period 2004–2007 in Ethiopia (in local currency units), obtained from the Food and Agricultural Organization Statistical Database (FAOSTAT).

Across the districts in the merged production and travel times panel, the average yield over all crops across districts increased 4.4-fold, implying an annual average growth rate of 9.7 percent. Over the same period, the yield over grain crops increased 2.5-fold, with an annual average growth rate of 5.9 percent. While productivity growth has been ubiquitous across virtually all districts, the productivity gains have not been shared equally (see online Appendix B.2).

D. *Difference-in-Differences Estimates*

To provide empirical motivation for the model, I first estimate the local effect of improved access to grain markets on outcomes of interest, employing a two-way fixed effects difference-in-differences design, with continuous treatment and staggered timing. The treatment is continuous because, while market access improved for all districts by 2014, the extent of the improvement was heterogeneous across districts and occurred in a nondiscrete fashion. The staggered timing allows for variation in the timing of treatment, given that in practice the road development program was rolled out gradually, changing market access in different time periods for different districts. I obtain the difference-in-differences estimates of the road development program using the following specification:

$$y_{it} = \alpha + \delta\big(\Delta Access_i \cdot NewRoad_{it}\big) + \eta_i + \varphi_t + \varepsilon_{it}, \tag{1}$$

over years $t \in \{1996, 2004, 2006, 2008, 2010, 2012\}$, where y_{it} is the logarithm of the productivity outcome for district i in year t.[8] The variable $\Delta Access_i$ is the district-specific absolute overall change in log travel time to the nearest grain market, weighted by the initial specialization in grains. I measure a district's initial

[8] The agricultural data corresponding to each year of new roads t are the pooled data for the two years after the placement of roads—e.g., for the "2004" new roads, the agricultural data are for the period 2005–2006. The road data used for the last year are for 2012, for their impact to take effect on the agricultural pooled data by 2014.

specialization in grains—the extent of exposure, by its 1996 share of land in grains production. The market access interaction term $\Delta Access_i$ captures the "dose" of the treatment across districts, given that the extent of market access improvements and the penetration of these improvements were nonuniform across districts. Higher $\Delta Access_i$ captures more exposure to additional and/or improved roads. The variable $NewRoad_{it}$ is a dummy for whether a district is exposed to a new road in a particular time period, as captured by the drop in the travel time relative to the initial period. Once a district i experiences a drop in its travel time in period t, this variable takes the value of 1 thereafter—i.e., once a district becomes treated, it stays treated. The coefficient of interest is δ, which captures the average response across districts to being exposed to new and better roads. It measures the differential change in productivity after the implementation of the road development program attributable to an additional 1 percentage point increase in market access. φ_t capture time fixed effects, which control for aggregate time-varying factors affecting outcomes in all districts in Ethiopia in the same way; η_i capture district fixed effects, which control for both observed and unobserved district characteristics with time-invariant effects on outcomes, allowing to exploit within-district variation over time.

The key identifying assumption underlying the difference-in-differences specification in (1) is that changes in productivity for districts with smaller changes in market access provide a good counterfactual for the changes in productivity that would have taken place in districts with larger changes in market access had they been exposed to the same changes in market access. Note that this requires not only the standard parallel trends assumption in difference-in-differences designs but also implicitly that the average treatment effect on the treated would be the same for districts that were assigned different doses of the treatment had the dose been the same.[9]

Table 2 reports the results from the difference-in-differences analysis. Columns 1 to 3 show the results of improved access to grain markets for districts on different outcome variables (in logarithmic form): the average yield over all grains in column 1, the number of fields in a district that use fertilizer in column 2, and the extent of specialization in grains given by the share of output they account for in column 3. The results suggest that improved market access, afforded through the road development program, had a significant positive effect on the outcomes of districts exposed to the new roads, with higher yields, more extensive fertilizer use, and more specialization in grain crops. The coefficient on the interaction term δ captures the treatment effect. The results indicate that an additional 1 percentage point increase in market access (increase in dose of the treatment) leads on average to a 0.10 percentage point increase in the yield. The effects on fertilizer use and food crop specialization are somewhat larger in magnitude (0.16 and 0.15, respectively). The effects on the yield and the extent of specialization are statistically significant

[9] In online Appendix C.1 I show that districts that experienced large improvements in market access over 1996–2014 do not differ systematically from other districts, pretreatment in 1996, in terms of distance to grain markets or Addis Ababa, nor in terms of the outcome variables of yield, fertilizer use, and extent of specialization in grains. Further, I show that the timing of a district's treatment with new roads is not correlated with the initial distance from grain markets. The distribution of first-time treatment across districts is also not systematically related to the distance from Addis Ababa, except perhaps for the fact that the farthest districts get treated a bit later.

Table 2—Estimated Effects of New Roads on Productivity

	Dependent variable (in logs):		
	Yield	*Fertilizer use*	*Specialization*
	(1)	(2)	(3)
$(\Delta Access \cdot NewRoad_{it})$	0.097	0.160	0.151
	(2.13)	(1.83)	(1.98)
Intercept	7.423	4.815	−0.011
	(45.71)	(15.66)	(−0.04)
District fixed effects	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes
Observations	2,299	2,274	2,299
Adjusted R^2	0.93	0.74	0.58

Notes: All columns contain estimates from OLS regressions of district log outcomes on the product of treatment amount (extent of market access) and a time-varying dummy for whether the district is treated in a particular period, including time fixed effects and district fixed effects in each case. The outcome variable is the average yield over all grains in column 1, the number of fields in the district that use fertilizer in column 2, and the share of grain output in total output in column 3. The sample is a balanced panel of districts with six different time periods: 1996, 2004, 2006, 2008, 2010, and 2012. t-statistics are in parentheses.

at the 5 percent level, while the effect on fertilizer use is significant at the 10 percent level.[10] In online Appendix C.2 I show that these results are robust to a measure of market access that includes only the improvements in travel time and does not weigh them by their penetration through initial specialization.

The main empirical challenge in identifying the causal effects of new roads on productivity is the potential endogenous placement of roads. The assignment of new roads to districts—its timing as well as its intensity—may not be random but may depend on the economic prospects of districts or their political connections with upper-level decision-making governmental authorities.

In the case of Ethiopia, while economic considerations are taken into account, according to the broad ERA criteria outlined in Section I, established areas, areas with economic potential, and underprivileged areas are prioritized. On net, it is not clear whether there is bias in a particular direction. In terms of political factors, even if there are districts with political connections to regional authorities, which projects get approved and their timing eventually is determined centrally by the ERA based on strategic considerations and the feasibility of the projects. Further, the timing and number of new roads are beyond the control of districts themselves, lessening concerns about selection into the treatment of improved roads. In addition, the district fixed effects in specification (1) account to some extent for permanent differences across districts.

[10] As a cautionary note, a recent literature shows that the coefficient on the interaction term in two way fixed effects research designs with heterogeneous effects across units and over time confound more than the average effect of the treatment on the treated, and may be biased, e.g., Callaway, Goodman-Bacon, and Sant'Anna 2021; De Chaisemartin and d'Haultfoeuille 2022. Overall, however these results do indicate that, while productivity increased everywhere after the construction of the new roads, the effect was larger in districts that experienced more market access.

While the difference-in-differences estimates are informative about the road development program, even if the identification is persuading, they only capture the relative effects in one district relative to another district. This analysis does not capture level effects or aggregate general equilibrium effects, resulting from the reallocation of production across crops and space or the reallocation of labor across sectors. I now turn to a structural model to examine the overall aggregate and spatial distributional effects of the road development program, accounting for general equilibrium implications across locations and sectors.

III. A Spatial Model of Agriculture and Development

I develop a spatial equilibrium two-sector model of agriculture and nonagriculture to assess the effects of changes in transportation costs on structural change, agricultural productivity, and development at both the aggregate and local levels. Overall, the model pins down the allocation of land across crops within locations, the distribution of agricultural production across space on the rural side of the economy, and the distribution of labor across rural locations and sectors. In equilibrium, the economy-wide aggregate measures of interest and their distribution across space are affected by the overall level of transportation costs in the economy, their variation across goods, and the spatial dispersion of these transport costs. I consider the role of transport costs in the presence of local frictions to the allocation of land.

A. *Environment*

Consider a spatial economy with an urban center and a finite number of J rural locations, indexed by $j \in \mathcal{J} \equiv \{1, 2, \ldots, J\}$. The economy produces two agricultural goods: a food crop f and a cash crop s. In addition, the economy produces a nonagricultural good n. Agricultural production takes place only in the rural locations, while nonagricultural production takes place only in the urban center. Each rural location can produce either of the two crops. The outputs of the same crop across locations are perfect substitutes for each other. The food crop is used only for domestic consumption in the urban center, while the cash crop is fully exported through the urban center.[11] It is assumed that there is unlimited demand abroad at the international price for the cash crop. The nonagricultural good is used only for consumption in the urban center.

Preferences.—There is a representative household in the urban center with preferences over food and nonagricultural goods,

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

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

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

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

$$y_{fj} = \left[z_{fj}^{1-\gamma}\left(n_j^{\alpha}\ell_{fj}^{1-\alpha}\right)^{\gamma}\right]^{\theta} x_j^{1-\theta}, \tag{2}$$

where y_{fj} is output of the food crop; z_{fj} is food crop productivity; and n_j, ℓ_{fj}, and x_j are labor, land, and intermediate inputs, respectively, used in the production of food in location j. In equation (2), $(1-\theta)$ determines the elasticity of final output with respect to intermediate inputs. The object in brackets raised to θ is the production function net of intermediate inputs, with parameter $\gamma < 1$ regulating the extent of returns to scale. Parameter $\alpha < 1$ captures the importance of labor relative to land in food production. Note that decreasing returns to scale imply incomplete specialization, and thus the food crop will be produced, at least partly, by every location j.

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

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

where z_{sj}, y_{sj}, and ℓ_{sj} are productivity, output, and land under the cash crop technology, respectively. The presence of the cash crop technology allows for an alternative use of land, outside of food, in rural locations.[12] The decreasing returns to scale in food production and the linearity of the cash crop technology allow to capture the stylized feature of the data that rural locations can be completely specialized in food but not in cash crops.

Production of Nonagricultural Good.—The nonagricultural good is produced by a representative firm in the urban location according to a constant returns to scale technology that is linear in labor,

$$Y_n = AN_n,$$

[12] In online Appendix E.2 I show that the results are robust to including intermediate inputs in the production of cash crops under constant returns to scale.

where A is nonagricultural productivity and N_n is the amount of labor allocated to nonagricultural production.

Goods Prices.—The nonagricultural good is the numeraire with its price normalized to one. Let p_f be the relative consumer price of the food crop in the urban location, which is endogenous. Note that because food produced in one location is a perfect substitute for food produced in another location, in equilibrium the consumer prices of food in the urban center from different locations will have to be the same and equal to p_f. This small open economy imports all the intermediate inputs from abroad, which are assumed to be inelastically supplied in the international market, and in exchange exports the cash crop it produces. Given that the cash crop is fully exported and the intermediate inputs are fully imported, their international prices p_s^*, and p_x^*, respectively, are taken as given. To the extent that there are differences in transportation costs faced by locations in delivering their crops to market, these will show up as differences in the farm-gate prices for food and cash crops. Similarly, while p_x^* is the price of intermediate inputs upon landing in the urban center, the local prices of intermediate inputs in the different rural locations will differ according to their location-specific transportation costs for delivering intermediate inputs.

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

Land Market Frictions.—The total amount of land in location j, L_j, can be allocated to the production of food or cash crops within that location. The market clearing condition for land in location j is

$$\ell_{fj} + \ell_{sj} = L_j. \tag{3}$$

I allow for frictions to the allocation of land within local markets. In particular, within each location there is a location-specific tax μ_j to the allocation of land between cash and food crops, such that the rental price of land under food crops is a fraction of the rental price of land under cash crops,

$$q_{fj} = (1 - \mu_j)\, q_{sj}, \tag{4}$$

where q_{ij} is the rental price of land under crop i in location j. The revenue from the taxes is redistributed to consumers in a lump-sum fashion. The tax μ_j is a catch-all for distortions to the allocation of land, which have been shown to be important in developing countries (Adamopoulos and Restuccia 2014).

Labor Markets.—Labor is used only in the production of food crops within each location, and it is perfectly mobile across all rural locations and the urban center.[13] The labor market clearing condition requires that the total amount of labor used in all rural locations and the urban location is equal to the total amount of labor in the economy,

$$N_a + N_n = N,$$

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

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

Goods Markets.—The economy-wide market clearing condition for food is

$$c_f = \sum_{j=1}^{J} c_{fj}, \tag{5}$$

where $c_{fj} = y_{fj}/\tau_{fj}$ is the amount of food (consumption) delivered to destination in the urban center originating from location j, and y_{fj} is the amount of the food crop produced and shipped from rural location j. Note that while the only source of demand for food from any location is the consumers of the city center, the amount of consumption is not equal to the amount of food produced in each rural location, since part of the output "melts" in transit. So c_{fj} is also the amount of net output of the food crop from location j.

Since the nonagricultural good is produced and consumed in the urban center, the market clearing condition is

$$Y_n = c_n N.$$

The entire amount of cash crop production from each location j is shipped to the urban center for export, with the export value upon arrival at the urban center being $ex_j = p_s^* y_{sj}/\tau_{sj}$. All intermediate inputs are imported, with a value upon reaching their destination in each rural location j of $im_j = p_x^* \tau_{xj} x_j$. The small open economy's total exports are $EX = \sum_j ex_j$ and imports are $IM = \sum_j im_j$. The economy's net exports are then given by $NX = EX - IM$.

To summarize, rural locations are heterogeneous with respect to (i) crop-location-specific productivities $\{z_{fj}, z_{sj}\}$, (ii) the total amount of productive

[13] In online Appendix E.2, in addition to the land frictions, I also allow for distortions to the allocation of labor across locations.

land L_j, (iii) local land frictions μ_j, and (iv) the vector of location-good-specific transportation costs $\{\tau_{fj}, \tau_{sj}, \tau_{xj}\}$.

Definition of Equilibrium.—A competitive equilibrium is a set of prices $\{p_f, w, (q_{sj}, q_{fj}, w_j)_{j=1}^J\}$, an allocation for each food crop farm $\{y_{fj}, \ell_{fj}, n_j, x_j\}$ and each cash crop farm $\{y_{sj}, \ell_{sj}\}$ in location j, an allocation for the nonagricultural firm $\{Y_n, N_n\}$, and a consumption allocation $\{c_f, c_n\}$, such that (i) the consumption allocation for urban consumers $\{c_f, c_n\}$ maximizes their utility subject to their budget constraint, given prices and the local land allocation frictions $(\mu_j)_{j=1}^J$; (ii) the production allocation for each food crop farm in location j, $\{y_{fj}, \ell_{fj}, n_j, x_j\}$, maximizes profits given prices, transportation costs $\{\tau_{fj}, \tau_{xj}\}$, and land L_j; (iii) the production allocation for each cash crop farm in location j, $\{y_{sj}, \ell_{sj}\}$, maximizes profits given prices and transportation cost τ_{sj}; (iv) the nonagricultural production allocation $\{Y_n, N_n\}$ maximizes the profits of the nonagricultural representative firm, given prices; and (v) the markets for labor, land, food crops, and nonagricultural goods clear.

B. *Analysis*

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

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

subject to the constraint that the total land allocated to food crop production in a given location cannot exceed the total amount of land in that location, $\ell_{fj} \leq L_j$. The location j wage rate is w_j.[14]

For a given price of food p_f, if the land constraint is binding in j, the optimal choice of land involves a corner solution $\ell_{fj} = L_j$. In this case, rural location j is completely specialized in the production of food. If the solution to the above problem for location j is at an interior optimum, then

$$\ell_{fj} = \left[p_f \gamma \theta \varphi_{fj} \left(\frac{x_j}{y_{fj}} \right)^{\frac{1-\theta}{\theta}} \right]^{\frac{1}{1-\gamma}} \left(\frac{1-\alpha}{q_{fj}} \right)^{\frac{1-\alpha\gamma}{1-\gamma}} \left(\frac{\alpha}{w_j} \right)^{\frac{\alpha\gamma}{1-\gamma}} < L_j, \tag{6}$$

and the location is incompletely specialized in the production of food—i.e., produces cash crops as well. I denote by $\varphi_{fj} = z_{fj}^{1-\gamma} / \tau_{fj}$ the "effective" productivity of location j in food crops, adjusting for transport costs.

For every location j, regardless of the extent of specialization, the intensity with which food crop farms apply intermediate inputs depends on the relative cost of intermediate inputs to the producer price of food,

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

[14] Standard nonlinear optimization techniques can be used to solve this problem numerically for every location, given a relative price for food p_f.

and food farm labor demand is a function of the food land input in that location,

$$n_j = \left[\left(\frac{\alpha}{w_j}\right)\gamma\theta p_f \varphi_{fj} \ell_{fj}^{(1-\alpha)\gamma}\left(\frac{x_j}{y_{fj}}\right)^{\frac{1-\theta}{\theta}}\right]^{\frac{1}{1-\alpha\gamma}}. \tag{8}$$

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

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

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

$$q_{sj} = \varphi_{sj}, \tag{9}$$

with $\varphi_{sj} \equiv p_s^* \frac{z_{sj}}{\tau_{sj}}$ being "effective" productivity in cash crop production, accounting for iceberg transportation costs, and being inclusive of the fixed international price of cash crops.

The profit maximization problem of the nonagricultural firm in the urban center is

$$\max_{N_n}\{AN_n - wN_n\},$$

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

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

$$I = wN + \sum_j(\hat{q}_{fj}\ell_{fj}) + \sum_j(1-\mu_j)\,q_{sj}(L_j - \ell_{fj}) + \sum_j \pi_j + TR,$$

where $\hat{q}_{fj}$ is the adjusted rental cost of food land in location j, which is equal to q_{fj} for incompletely specialized districts (where the land constraint is not binding) and greater than q_{fj} by the marginal benefit of additional food land in districts completely specialized in food (binding land constraint). The total revenue from the taxes to the allocation of land within districts, $TR = \sum_j \mu_j q_{sj}\ell_{sj}$, is redistributed lump-sum to consumers. Given the nature of the preferences, the household will consume an amount of food per member $c_f = \bar{f}$ and allocate the residual income to the consumption of nonagricultural goods.

The relative price of food crops in the urban center, p_f, must clear the market for food, (5). Consumers in the urban center consume a fixed amount of food per person, $\bar{f}$. Each rural location produces food, y_{fj}, which, upon delivery to the urban center, is $c_{fj} = y_{fj}/\tau_{fj}$, due to the incurred transport costs. Then the market clearing condition for food crops that implicitly determines p_f is

$$N \cdot \bar{f} = \sum_{j\in\mathcal{S}} c_{fj} + \sum_{j\notin\mathcal{S}} c_{fj}, \tag{10}$$

where $\mathcal{S}$ is the set of locations completely specialized in food crop production, $\mathcal{S} = \{j \in J : \ell_{fj} = L_j\}$.

Spatial Distribution of Food Production.—When some locations are completely specialized and others are incompletely specialized, the relative price of food and the spatial distribution of production cannot be determined analytically. However, for illustrative purposes we can obtain an analytical solution for the relative price of food and the land allocation if all locations were assumed to be incompletely specialized—i.e., there was an interior solution for every j. In this case it can be shown that the equilibrium land allocation in food production in each location j is

$$\ell_{fj} = \Lambda \frac{\left[\frac{\varphi_{fj}}{\varphi_{sj}^{1-\alpha\gamma}}\right]^{\frac{1}{1-\gamma}} \left[\frac{1}{1-\mu_j}\right]^{\frac{1-\alpha\gamma}{1-\gamma}} \left[\frac{1}{\tau_{fj}\tau_{xj}}\right]^{\frac{1-\theta}{\theta(1-\gamma)}}}{\left[\sum_{k=1}^{J} \left[\frac{\varphi_{fk}}{\varphi_{sk}^{\gamma(1-\alpha)}}\right]^{\frac{1}{1-\gamma}} \left[\frac{1}{1-\mu_k}\right]^{\frac{(1-\alpha)\gamma}{1-\gamma}} \left[\frac{1}{\tau_{fk}\tau_{xk}}\right]^{\frac{1-\theta}{\theta(1-\gamma)}}\right]^{\frac{1}{1-\theta(1-\gamma)}}}, \tag{11}$$

where Λ is a constant that summarizes parameters of the model. The amount of the economy's food that a location produces depends on the relative effective productivity and the (inverse) levels of transportation costs and land frictions of that location relative to those of all other locations. To understand the spatial distribution of food production implied by the model, consider the ratio of equilibrium land allocations between any two incompletely specialized locations j and k,

$$\frac{\ell_{fj}}{\ell_{fk}} = \left(\frac{\varphi_{fj}/\varphi_{sj}^{1-\alpha\gamma}}{\varphi_{fk}/\varphi_{sk}^{1-\alpha\gamma}}\right)^{\frac{1}{1-\gamma}} \left(\frac{1-\mu_k}{1-\mu_j}\right)^{\frac{1-\alpha\gamma}{1-\gamma}} \left(\frac{\tau_{fk}\tau_{xk}}{\tau_{fj}\tau_{xj}}\right)^{\frac{1-\theta}{\theta(1-\gamma)}}. \tag{12}$$

According to the first term in equation (12), if location j is relatively more productive in food crops than cash crops in comparison to location k, j will produce relatively more food compared to k. In other words, comparative advantage in effective productivity matters, which captures comparative advantage in relative actual productivities (z_{ij}) as well as comparative advantage in (inverse) relative transport costs (τ_{ij}) and relative land frictions (μ_j). Locations that face relatively high transport costs in producing food relative to cash crops will allocate less of their land to the production of food crops. The third term in equation (12) indicates that the inverse ratio in the levels of food and intermediate good transport costs is also relevant. The channel through which the levels of transport costs matter is the intermediate input use, as higher transport costs deter the use of intermediate inputs. Other key productivity measures are discussed in online Appendix F.

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

IV. Calibration

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

The parameters that need to be determined in order to calibrate the model to match the spatial agricultural production structure of the Ethiopian economy are (i) the $J \times 2$ matrix of crop-specific productivities across the different locations $\left\{z_{fj}, z_{sj}\right\}_{j=1}^{J}$; (ii) the $J \times 3$ matrix of iceberg transportation costs for each of the crops, as well as the intermediate inputs between the different rural locations and the urban center $\left\{\tau_{fj}, \tau_{sj}, \tau_{xj}\right\}_{j=1}^{J}$; (iii) the $J \times 1$ vector of local land frictions $\left\{\mu_j\right\}_{j=1}^{J}$; (iv) the $J \times 1$ vector of total agricultural land for each location $\left\{L_j\right\}_{j=1}^{J}$; (v) nonagricultural productivity in the urban location A; (vi) food crop technology parameters (γ, α, θ); and (vii) the preference parameter $\bar{f}$.

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

In the data, $J = 403$, which includes the districts for which agricultural production data and transport cost data are available. The food crop in the model corresponds to cereals in the data, which account for 84 percent of the land allocation overall in the economy. Cash crops are taken to include all other crops. The beginning and end of the period are 1996 and 2014, using the pooled data for each period, as described in Section IIC. The world prices of cash crops p_s^* and intermediate inputs p_x^* are normalized to one, as they do not vary in the quantitative experiments.

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

Total Labor.—The total amount of labor in the economy N is taken directly from aggregate data for the Ethiopian economy in 1996, from the Groningen Growth and Development Center (GGDC) ten-sector database (Timmer, de Vries, and de Vries 2015).

Transportation Costs by Location and Good τ_{ij}.—In the benchmark economy, transportation costs for the two crops and intermediate inputs are estimated from travel times from district centroids to destination markets within Ethiopia, through the existing road infrastructure network, measured from the geographic analysis for 1996. When estimating transportation costs for food crops (cereals), the travel times used are those from district centroids to the nearest grain market. Note that in the model there is only one agricultural market in the urban center, while in the data there are multiple. I use the travel time to the closest grain market as the measure of travel time to the central market in the model. While the regional grain markets in Ethiopia are appropriate for estimating food transport costs, they are unlikely to be a good approximation for the costs incurred for selling cash crops and purchasing intermediate inputs. Given that cash crops are primarily exported and that exports run through the capital of Addis Ababa, the transportation cost for cash crops is estimated from the travel time from a district centroid to Addis Ababa. Given that the distribution of intermediate inputs is centralized, the travel time between the district centroid and Addis Ababa is also used for intermediate inputs. Note that the model has iceberg transport costs, which use up resources, while the data involve travel times. To map travel times to iceberg transport costs I posit a transport cost function of the following form:

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

where tt_{ij} is the travel time (in minutes) for good i from rural location j to the market. The parameter η captures the sensitivity of transport costs with respect to travel time, and with $\eta < 1$ the transport cost–travel time relationship is concave. $\psi_i > 0$ is a scale parameter that controls the units—in particular, regulating how far from one the implied transport costs are (with $\psi_i = 0$, there are no iceberg costs, and $\tau_{ij} = 1$). Next, I explain how I calibrate the parameters of the transport cost function. I estimate the elasticity parameter η using data on price gaps for grains between regional wholesale markets and the capital of Addis Ababa, along with my GIS effective travel times between these markets. The grain prices are from the Ethiopian Grain Trade Exchange, which reports prices for internationally traded grains. Given that crops are traded internationally through the capital of Addis Ababa, the wholesale price gaps are a reasonable proxy for transport costs. Across all grains, with crop fixed effects, I estimate an elasticity parameter of price gaps with respect to an effective travel time of 0.79. This implies that transport costs rise in a concave fashion with distance. A detailed description of the data used and the estimation is provided in online Appendix D. I impose the same concavity on my transport cost measure with respect to travel time by setting $\eta = 0.79$. Then, given the value for η, I calibrate the unit parameter ψ_i for crops so that the total amount of resources devoted to transport as a share of consumer value of output in the model matches the share of transportation costs in the sales value of food in the data for Ethiopia in 1996. Based on a survey of grain wholesale traders across grain markets in Ethiopia for 1996, Gabre-Madhin (2001) shows that for grain "exporting" regions, 26 percent of the sale price is accounted for by marketing costs of various kinds and the profit margin of the transporter. Direct transport costs, including road stops, during the transportation of grains accounted for 58 percent of the overall marketing costs, implying that 13.2 percent of the final sale price is transport. This provides a lower bound on the transport cost share. However, given that some of the other marketing costs such as handling, sacking, storage, commission of brokers, travel cost of transporter, and profit of the transport company can arguably be attributed to "transportation," I target a transport cost share of the final sale price of 18 percent, which is between the lower bound of 13 percent and the upper bound of 26 percent. This implies $\psi_i = 0.00273$ for cereals, which I use for both food and cash crops in the model. The transport cost share of the delivered farm-gate price of fertilizer is higher. Minten, Koru, and Stifel (2013), using data from Northwestern Ethiopia, show that transportation costs, accounting for "last mile" costs, can raise the effective price of chemical fertilizer by up to 50 percent.[15] I set $\psi_x = 0.0043$, which implies a share of transport costs in the farm-gate cost of intermediate inputs of 36 percent to be conservative.

Food Crop Technology Parameters (γ, α, θ).—The elasticity of output with respect to intermediate inputs is chosen to match the intermediate input cost share in the value of final output. Based on aggregate estimates, in 2011 the value of nonagricultural intermediate inputs in gross output for the agricultural sector in Ethiopia was 13 percent (Kebede and Heshmati 2023), which implies $\theta = 0.87$.[16] In the food crop production function, γ and α regulate the extent of decreasing returns to scale and the income share split between land and labor, respectively, in the nonintermediate input part of the production function. While the literature has employed different approaches to discipline γ, using microdata and a dynamic panel approach, Manysheva (2022) estimates the returns to scale in agriculture for Tanzania to be 0.76. Given the similarity between Tanzania and Ethiopia in terms of stage of development, and given the structure of their agricultural sectors, I use this value in my benchmark analysis. Given the value of γ, I choose α to match a land income share in agriculture of 22 percent for Ethiopia (Kebede and Heshmati 2023). The resulting income share for labor in value added is 54 percent. In online Appendix E.1 I consider alternatives to the benchmark elasticity parameters for robustness.

Rural Productivity Parameters z_{ij}.—For each district j, the productivity terms of the food crop and the cash crop technologies $\{z_{fj}, z_{sj}\}$ are chosen to match the food and cash crop yields for that district, y_{fj}/ℓ_{fj} and y_{sj}/ℓ_{sj}, respectively. Given that all districts produce food crops, the actual yield for food crops is available for every district. However, for districts that produce only food crops (completely specialized), I use additional data from the Global Agro-Ecological Zones (GAEZ) project (GAEZ 2000) to estimate the cash crop yield.[17] In particular, I use potential yields by individual crop available at high spatial resolution from GAEZ to aggregate up and estimate potential yields for food and cash crops for each completely specialized district. I then use the ratio of cash/food crop potential yields across

[15] This is consistent with Aggarwal et al. (2022), who find high transport costs for fertilizer in rural Tanzania.

[16] The low use of intermediate inputs is consistent with earlier estimates from the Food and Agriculture Organization of the United Nations that found an intermediate input share of 10 percent (Prasada Rao 1993).

[17] See Adamopoulos and Restuccia (2022) for the use of GAEZ potential yields to assess agricultural productivity.

districts and the actual food yield to impute potential cash crop productivity. Real yields are computed across districts by aggregating individual crops with a common set of prices.

Identification of Local Land Frictions.—The location-specific frictions to land μ_j deter the reallocation of land toward cash crops within a location. I back out these frictions from the observed 1996 allocation of land across food and cash crops within each district as wedges in the marginal products of land between crops. For districts that produce both crops (incompletely specialized), the actual allocation of land across crops is used to identify the misallocation of land, conditional on crop productivities. For districts that are completely specialized in food, I use the actual food crop productivity and the estimated cash crop productivity to back out the extent of land frictions that would rationalize complete specialization in food production. The implied economy-wide cash/food yield ratio is 1.33 as in the data.

Urban Productivity A.—I normalize the consumer price of food to 1 in the benchmark economy. The nonagricultural productivity parameter A is calibrated to match in equilibrium a target for the share of labor in agriculture of 86 percent, based on aggregate data for the Ethiopian economy, from the GGDC ten-sector database (Timmer, de Vries, and de Vries 2015).

Food Consumption Requirement $\bar{f}$.—As I do not have access to consumption data for Ethiopia, the subsistence level of food consumption is pinned down by the equilibrium requirement that the total consumption of food by domestic consumers fully absorbs the total production of food from all districts, net of their respective transport costs. Given that food production data and transport costs are targeted explicitly in the calibration as described above, the market clearing condition for food (5) determines the value of the per capita subsistence food consumption term $\bar{f}$.

The economy-wide calibrated parameters of the model, which are common across all locations, along with their descriptions and targets, are provided in Table 3. Table 4 provides a description of the location-specific parameters that are mapped into actual district-level data, along with their data targets. The calibrated model does well in replicating aggregate and spatial features of the Ethiopian economy. The values of key variables of interest in the calibrated benchmark economy are provided in the first column of Table 6. However, the model also does well in matching district-level statistics that are not targeted in the calibration. Panel A in Figure 2 compares the spatial distribution across districts of food farm labor (share in total economy-wide labor engaged in food production) to the 1996 distribution of households engaged in cereal production across districts (as a share of the total households engaged in cereals).[18] There is a strong positive correlation between model and data for the food crop labor shares.

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

Table 3—Calibrated Common Parameters

Parameter	Description	Value	Data target
γ	Share of land and labor (food)	0.76	Returns to scale
α	Share of labor (food)	0.71	Land income share
θ	Noninterm. input share (food)	0.87	Interm. input intensity
η	Trans. costs elasticity	0.79	Estimated from price data
ψ_f, ψ_s	Trans. cost scale, crops	0.00273	Trans. share of crop price
ψ_x	Trans. cost scale, fertilizer	0.00430	Trans. share of fertilizer price
A	Urban non-agricul. prod.	645.6	Share of agricultural labor
N	Total number of workers	24,806	Total 1996 labor
$\bar{f}$	Subsistence food consumption	1,181.9	Per capita 1996 production

Table 4—Calibrated District-Specific Parameters

Parameter	Description	Data target (1996)
$\{L_j\}_{j=1}^J$	Total agricultural land	Total land, AgSS
$\{\tau_{fj}\}_{j=1}^J$	Food crop iceberg trans. cost	Travel time, nearest grain market
$\{\tau_{sj}\}_{j=1}^J$	Cash crop iceberg trans. cost	Travel time, Addis Ababa
$\{\tau_{xj}\}_{j=1}^J$	Inter. input iceberg trans. cost	Travel time, Addis Ababa
$\{z_{fj}, z_{sj}\}_{j=1}^J$	Productivity by district	Food, cash yields, AgSS and GAEZ
$\{\mu_j\}_{j=1}^J$	Land frictions by district	Food land share, AgSS

The AgSS data provide reliable information on whether any given field operated by a household uses fertilizer. In addition, the AgSS contains information on the amount of fertilizer applied. However, these data are more sparse and less reliable for time series comparisons. Using the AgSS data, I construct a district-level measure of intermediate input use as the share of all fields that have any amount of fertilizer applied to them. On average this has increased from 32 percent in 1996 to 52 percent in 2014, indicating a significant increase in the use of fertilizer. In the model, there are no fields, so an intermediate input intensity district-level measure can only be constructed as the share of intermediate inputs in final output in each district. The spatial distribution of district-level intermediate input intensities is not targeted in the calibration. Nevertheless, as panel B in Figure 2 shows, the intensive margin intermediate input measure from the benchmark economy in the model is strongly positively correlated with the extensive margin intermediate input intensity measure from the data, with a correlation coefficient of 0.48.

A. *Comparing Model to Estimated Treatment Effects*

To assess the appropriateness of the model, here I compare the structural model–implied effects to the evidence from the difference-in-differences estimates in Section IID. To facilitate the comparison, I simulate in the model the equilibrium

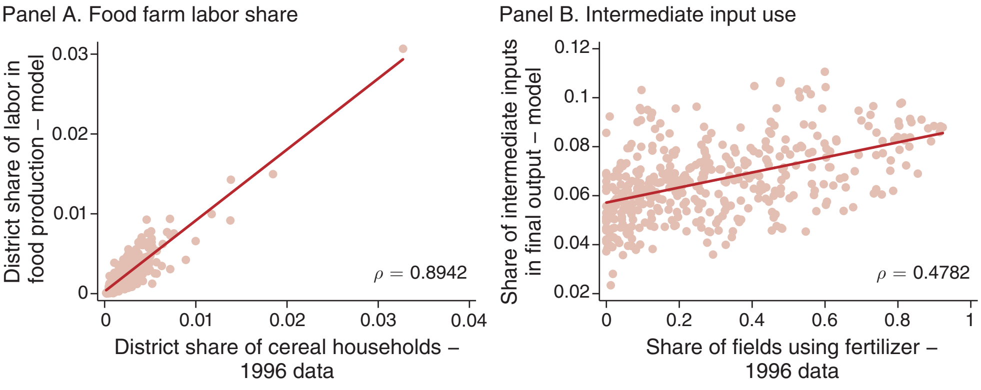

Figure 2. Cross-District Comparison of Model to 1996 Data

Notes: "Model" refers to the value in the calibrated benchmark economy. "1996 Data" refers to the district value from the 1996 pooled data from the Ethiopia Agricultural Sample Surveys.

effects of a change in food transport costs alone from their 1996 level to their 2014 level, keeping the cash and intermediate input transport costs constant. With the implied equilibrium model–generated data, before (calibrated) and after (new equilibrium) the change in food transport costs, I run the same difference-in-differences specification as with the actual panel data, with the same market access interaction term as in equation (1). I note that while, in the model, transport costs are of the iceberg form, in the estimation, I include the travel times exactly as in the empirical specification, to keep the treatment the same. I include year and district fixed effects with both the model-generated data and the actual data.

The first column in Table 5 shows the difference-in-differences yield elasticity estimate from the equilibrium model–generated data. For comparison, the second column repeats the estimated elasticity from the data. The resulting model "treatment effects" are consistent in sign and magnitude with the empirical results. Reassuringly, the elasticity of the yield with respect to market access in the model is 0.084, close to the estimated elasticity from the empirical difference-in-differences design of 0.097. The fact that 87 percent of the yield gain estimated from the data can be accounted for through the mechanisms of the calibrated model suggests that it is a reasonable framework in which to conduct other experiments and counterfactuals.

V. Quantitative Experiment

The experiment involves studying the effects from reducing geographic transport costs across all districts from their actual 1996 levels to their actual 2014 levels. In order to isolate the effects of transport cost changes alone, I keep all other parameters to their 1996 levels. In other words, I ask what the aggregate and spatial micro-level effects on the Ethiopian economy would be if the only change between 1996 and 2014 had been the change in the transportation network and the associated

Table 5—Comparing Model to Data Estimates

	Dependent variable (in logs): *Yield*	
	Model	Estimated
	(1)	(2)
$(\Delta Access \cdot NewRoad_{it})$	0.084 (10.59)	0.097 (2.13)
District FE	Yes	Yes
Year FE	Yes	Yes

Notes: Both columns contain estimates from OLS regressions of district log yield over all grains on the product of treatment amount (extent of market access) and a time-varying dummy for whether the district is treated in a particular period, including time fixed effects and district fixed effects in each case. The first column presents estimates from the model-generated data, the second column from the actual panel data. *t*-statistics are in parentheses.

changes in transportation costs? I then compare these changes to the actual changes in the variables of interest that occurred in the data over the same period. The model allows me to assess directly the effects of transport cost changes irrespective of the other changes that may have occurred over the same period and that may have also contributed to changes in the variables of interest.

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

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

for $i \in \{f, s, x\}$, where the travel times for 2014, $tt_{ij,2014}$, are the ones estimated in Section IIA from the road infrastructure network present in 2014. The associated changes in transport costs alter the connectivity of districts with markets, and do so in a heterogeneous fashion, since the volume and quality of the road network did not expand for all districts at the same rate. As a result, there is a change in both the level and the dispersion of good-specific transport costs across districts. Keeping all other parameters (including productivity) in all rural and urban locations to their benchmark economy levels, I feed the 2014 iceberg costs $\{\tau_{fj,2014}, \tau_{sj,2014}, \tau_{xj,2014}\}_{j=1}^{J}$ into the model and solve for the new equilibrium. The mean food (cash) net iceberg transport cost $(\tau - 1)$ drops by 29.7 (19.3) percent, while the median drops by 25.9 (19.4) percent.

A. *Aggregate Effects*

The aggregate outcomes in the new equilibrium, associated with the 2014 transport costs, are presented in the second column of Table 6, and the percentage changes relative to the 1996 benchmark economy (first column) are presented in the third column. I note that in the new equilibrium, I aggregate across crops and goods using a common set of prices before and after the change in transport costs, just as statistical agencies measure “real” changes. The common set of prices I use are the ones from the benchmark economy, net of transport costs.

Table 6—Aggregate Effects of Reducing Transport Costs to 2014 Levels

Statistic	Benchmark economy (BE)	2014 transport costs	Percentage change (%)
Real aggregate yield	1,674.64	1,920.98	14.7
Yield in food crops	1,591.11	1,746.42	9.8
Yield in cash crops	2,108.85	2,388.61	13.3
Real net aggregate yield	1,339.75	1,590.67	18.7
Net yield in food crops	1,304.38	1,502.28	15.2
Net yield in cash crops	1,523.59	1,827.45	19.9
Real value added yield	1,585.96	1,831.97	15.5
Real value added per worker	1,992.44	2,458.65	23.4
Share of employment in agriculture	0.86	0.81	−5.5
Total share of land in food	0.84	0.73	−11.0
Intermediate input intensity	0.11	0.11	0.7
Consumer price of food	1.00	0.94	−6.4
Average farm size	1.26	1.34	6.8
Real GDP per worker	1,384.23	1,688.73	22.0

Notes: The column "Benchmark economy (BE)" displays the values for each variable in the baseline calibrated economy. The column "2014 transport costs" displays the values of each variable when transport costs are reduced to their 2014 levels. The percentage changes in the counterfactual economy (with reduced transport costs) relative to the BE are in the last column. All variables, except for those reported in shares, are reported as the percentage change in the counterfactual relative to the BE. For variables reported in shares, the last column displays the difference between the pre– and post–transport costs change.

There are substantial aggregate effects when transport costs alone are reduced to their 2014 levels in terms of structural change and productivity, as captured by several metrics in Table 6. The share of labor in agriculture in the model drops from 86 percent in the benchmark economy to 80.5 percent, a drop of 5.5 percentage points. As a summary statistic of the economy-wide productivity effects in agriculture, value added per worker increases by 23.4 percent. The aggregate real yield in agriculture, measured as real value of final output per unit of land, increases by 14.7 percent. This is achieved through an increase in within-crop real yields, 9.8 percent for food and 13.3 percent for cash crops, as well as a reallocation of land from food crop production to cash crop production in the economy overall. The share of land in food production in the economy drops from almost 84 percent in 1996 to under 73 percent after the transport cost changes.

The driver of these substantial effects is the change in transport costs across goods and space through the mechanisms of the model. The within-crop increases in the yields are achieved through the spatial reallocation of production across districts according to changes in relative comparative advantage implied by the changes in transportation costs. To see this, recall that in the model, comparative advantage across districts is determined not only by relative actual TFP (z_{ij}) but also by relative transportation costs (τ_{ij})—i.e., by relative "effective" productivity. After the drop in transport costs, food production is undertaken increasingly by relatively more "productive" districts, and labor is reallocated across rural districts according to the redistribution in food production. Given the inelastic demand for food in the country, captured by the subsistence requirement $\bar{f}$, as the economy becomes more productive in producing food, districts do not need to devote as many resources to food production. The land that is freed up from food production is allocated to the production of cash crops. Given that labor is used only in food production, when land allocated to food falls, the amount of labor needed to produce that same amount of food also goes down, now being reallocated to nonagricultural production in the urban center. With the drop in the overall engagement in food production, the demand for imported intermediate inputs falls. However, the use of intermediate inputs on the intensive margin increases. The overall share of intermediate inputs in final output for the economy increases by less than 1 percentage point. This is because the relative cost of intermediate inputs depends not only on the domestic transport costs of delivering those inputs to districts but also on the relative price of food. While transport costs for intermediate inputs and food fall, the relative price of food also falls by 6 percent, counteracting part of the transport cost savings depending on the district.

Note that the iceberg transport costs are resource costs that show up in the model as goods "melting" in transit. As a result, part of the output of each crop (food and cash) constitutes payments to the transportation sector. The net amount of output of crop i delivered to the urban location (for consumption in the case of food, and for export in the case of cash crops) is y_{ij}/τ_{ij}. An alternative real productivity metric to consider is the real net yield, which nets out transport costs and thus takes into account the direct resource savings from lower transport costs. The real net yield in the model increases by 18.7 percent when transport costs drop to their 2014 levels. The indirect productivity gains achieved through the mechanisms of the model account for 80 percent of the overall gains $(\log(1.147)/\log(1.187))$, implying that the direct savings from lower transport costs are the residual 20 percent.

The key metric of value added per worker in agriculture takes into account not only changes in the real value added yield (which increases by 15.5 percent) but also the induced equilibrium changes in agricultural labor, described above. The economy-wide agricultural farm land per worker—or average farm size in the model—increases by 6.8 percent. A simple decomposition of the total aggregate gain in value added per worker reveals that the real value added yield gain accounts for 69 percent $(= \log(1.155)/\log(1.234))$, while the average farm size gains for the remaining 31 percent $(= \log(1.066)/\log(1.223))$.

$$\underbrace{\frac{VA_a}{N_a}}_{1.234} = \underbrace{\frac{VA_a}{L}}_{1.155} \cdot \underbrace{\frac{L}{N_a}}_{1.068}. \tag{13}$$

These findings indicate that lower transport costs not only raise productivity in farming but also lead to a restructuring of the agricultural sector, characterized by a shift of production toward more export oriented cash crops, lower employment in agriculture, and larger farm sizes.

Real GDP per worker in the economy, which also takes into account the output of the nonagricultural sector in the urban center, increases by 22 percent. This is a

Table 7—Effects of Different Channels

Model	Value added per worker in agriculture	Land share in food crops	Agricultural employment share	GDP per worker
All channels	23.4	72.8	80.5	22
No land reallocation	7.5	83.9	76.9	5.7
No interm. input change	19.4	74.4	81.6	19.9
No labor reallocation	8.1	79.1	86.0	12.8

Notes: The first and fourth columns display the percentage change in the counterfactual economy relative to the benchmark, the second and third columns the shares in the counterfactual economy, when transport costs are reduced to their 2014 levels. "All channels" refers to the full model. "No land reallocation" keeps the land allocation to 1996. "No intermediate input change" keeps the intermediate input intensity in each district to its 1996 level. "No labor reallocation" keeps the labor allocation across space and sectors to 1996.

substantial increase due to the heavy reliance of the economy on developments in the agricultural sector.[19]

Role of Key Mechanisms.—The model features three main channels: (i) the reallocation of production across districts and crops, (ii) the use of intermediate inputs, and (iii) the reallocation of labor across sectors and space. To assess the relative importance of each channel for overall agricultural productivity, I conduct a set of counterfactual experiments whereby I shut down each channel in turn but allow the other channels to operate. The results are presented in Table 7, where the first row repeats the results when all channels are active. I report four key outcome statistics when all transport costs are reduced to their 2014 levels: the change in value added per worker, the shares of land in food and labor in agriculture, and the change in GDP per worker. The "No land reallocation" counterfactual keeps the allocation of land across crops in all districts to its 1996 level. The "No intermediate input change" counterfactual keeps the intermediate input intensity to its 1996 level, while the "No labor reallocation" counterfactual keeps the allocation of labor across space and sectors to its 1996 level.

The results suggest that the distribution of production across space and the allocation of land across crops is the most important channel, as shutting it down reduces agricultural productivity by 65.6 percent relative to the baseline results. The overall reallocation of labor across space and sectors is almost equally important, as eliminating this channel reduces agricultural labor productivity by 63 percent. Note that the first keeps the land share in food to its benchmark 1996 level, while the second keeps the share of labor in agriculture. The intermediate input channel plays a smaller role.

[19] In online Appendix G, I show that these results are robust to including local labor markets.

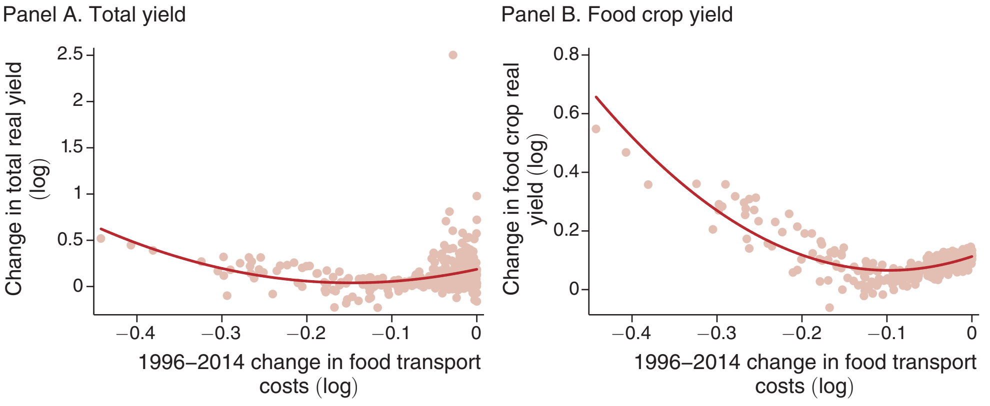

Figure 3. Model Changes in Yields with Transport Cost Changes

Notes: "Change in total yield" and "Change in food crop yield" refer to the change in the real overall yield (valued at a common set of prices) and food crop yield, respectively, in the model, after the reduction in transport costs. The x-axis represents the log-change of food transport costs over 1996–2014.

B. *Spatial Distribution of Effects*

While the economy-wide aggregate gains capture the overall effect of changes in transportation costs because they take into account the gains from the spatial reallocation of production, it is also important to understand the spatial patterns across districts that the changes in transport costs impart. The spatial distributional consequences of the transport cost changes further help shed light on the mechanisms of the model.

In Figure 3, panel A, I plot the change in the log–real total yield by district against the change in log–level of transport costs in food by district. There is a U-shaped pattern across districts between changes in yields and changes in the level of food transport costs. This implies that for one set of districts, the smaller the changes in the level of their food transport costs, the larger the gains (upward-sloping part of the U-curve), while, for another set of districts, the larger the changes in the level of food transport costs, the larger the gains (downward-sloping part of the U-curve). In order to understand what accounts for this relationship, note that the U-shaped pattern across districts is also present when examining the change in log–real yield in food crops alone against log–changes in the level of food transport costs, as illustrated in Figure 3, panel B. This is to be expected, as in the model the cash crop production technology is linear in land, and therefore there are no within-district changes to the real yield in cash crops. As a result, the U-shaped pattern of the food crop yield carries over to the district-level total yield.

At a proximate level, the reason for the U-shaped pattern is due to the degree of specialization in food crops: the yield growth–transport cost change relationship is negative among districts that are completely specialized in food production but is positive for the districts that are incompletely specialized in food (i.e., those that also produce cash crops).

For districts completely specialized in food, the larger the drop in food transport costs, the larger the increase in intermediate input intensity and the larger the increase in the labor-land ratio (as lower transport costs increase the relative cost of land). Lower transport costs allow these districts to better exploit the initial comparative advantage they have in food crop production. For districts that are incompletely specialized in food, the smaller the increase in the food transport cost, the larger the switch to cash crops and, by diminishing returns, the larger the increase in the food yield, accounting for the upward-sloping part of the U-curve. The total yield increases even more for these districts because land is allocated toward the relatively more productive cash crops.

The degree of specialization of a district depends on the extent to which changes in the level of food transport costs τ_{fj}, or "absolute advantage," are aligned with changes in the relative food-to-cash transport costs τ_{fj}/τ_{sj}, or "comparative advantage." While food transport costs drop for all districts, for districts that completely specialize in food, the drops are very large, and they are completely aligned with changes in relative food-to-cash transport costs—i.e., both the level and relative transport costs exhibit large drops and are virtually perfectly correlated. For districts that are incompletely specialized, the level of food transport costs (which decreases) and relative food-to-cash transport costs (which tend to increase for these districts) are misaligned, and they are more weakly correlated.

Changes in Spatial Inequality.—It is natural to ask, to what extent did the building of new roads mitigate spatial inequality across districts? Before the road expansion program in Ethiopia began, transport costs were high, particularly for isolated regions detached from the capital center of Addis Ababa. In Table 8, panel A, I order districts according to their distance (travel time) from Addis Ababa in 1996 and group them into quintiles of the distribution of distances from Addis Ababa. As seen in the first column, the closest districts (Q1) were on average 3.5 hours away from the capital in 1996, while the farthest districts (Q5) were more than 16.6 hours away on average. The second column shows the average increase in the total yield for the districts in each of the 1996 distance quintiles, implied by the model when all transport costs change to their actual 2014 levels. While all districts benefited from the drops in transport costs, the districts farthest away in 1996 (Q5) exhibited higher yield growth on average, 26.1 percent, than the closest districts in 1996 (Q1), which had 14.6 percent growth. These growth differences contributed to the partial convergence of the more distant districts from Addis Ababa to the nearby ones.

Table 8, panel B orders and groups districts into quintiles according to the level of their 1996 total yield, with Q1 being the lowest-productivity 20 percent in 1996 and Q5 the highest-productivity 20 percent. The first column displays the average total yields for each group in 1996. The second column shows the average yield growth for each group following the drops in transport costs to their 2014 levels. While the relationship is not monotonic, the least productive districts in 1996 (Q1) exhibited higher growth than the other districts.

In sum, there is some evidence of convergence and a dent in spatial inequality, with more distant and lower productivity districts benefiting more from the new infrastructure.

Table 8—Changes across Space and Productivity

Panel A. Ordered according to 1996 distance from Addis Ababa

Quintile	Travel time to Addis Ababa (min)	Model % change in yield (with transport cost changes)
Q1	210.7	14.6
Q2	405.3	11.4
Q3	548.3	16.3
Q4	712.0	17.5
Q5	994.2	26.1

Panel B. Ordered according to 1996 yield

Quintile	Total yield	Model % change in yield (with transport cost changes)
Q1	855	26.7
Q2	1,307	11.9
Q3	1,717	15.0
Q4	2,314	15.6
Q5	4,672	17.6

Notes: In panel A, districts are ordered according to their 1996 distance from Addis Ababa and grouped into quintiles of their distance distribution. In panel B, districts are ordered according to their total yield in 1996 and grouped into quintiles of the yield distribution. The second column shows the average yield growth rate across districts within each quintile.

C. *Comparison to Data Changes*

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

Table 9 compares the aggregate changes from the model (first column) to the ones in the data (second column) for key variables of interest. I focus here on the real yield as the measure of agricultural productivity rather than value added per worker since, due to lack of data, I cannot calculate the latter at the district level. The real aggregate yield increases 14.7 percent in the model, accounting for 9.2 percent $(\log(1.147)/\log(4.4))$ of the overall increase in the same metric in the data. If, however, the direct resource savings from the transport cost reductions are included, then the aggregate net yield increases by 18.7 percent, accounting for 11.5 percent of the overall yield gain in the micro-level agricultural production data. In other words, the model with only transport cost changes can account for about one-tenth of the yield gains in the data. The gross yield in food crops in the model accounts for 10 percent of the one observed in the data.

In terms of other statistics, the drop in the share of land allocated to food crops and the increase in average farm size are in the neighborhood of these changes in the data over the period 1996–2014. The model also accounts for less than half of the drop in the share of labor employed in agriculture. Finally, the model also generates an increase in GDP per worker that is about half of the actual change in the data.

Table 9—Comparison of Model and Data Changes (Aggregate Statistics)

Statistic	Changes due to transport cost reductions (%)	Changes in data over 1996–2014 (%)
Real gross aggregate yield	14.7	341.9
Real net aggregate yield	18.7	341.9
Real yield in food crops	9.8	153.9
Total share of land in food	−11.0	−8.5
Average farm size	6.8	6.2
Share of employment in agriculture	−5.5	−12.9
Real GDP per worker	22.0	67.1

Notes: The first column shows changes relative to the benchmark economy implied by the model when all transport costs are reduced to their 2014 levels. All the changes in the data are computed from the Ethiopian Agricultural Sample Surveys data over 1996–2014, with the exception of the "Share of employment in agriculture" and "Real GDP per worker" values, which are computed from the GGDC ten-sector database as changes over 1996–2011.

Next, I compare the spatial pattern of the yield gains produced by the model with the ones in the data. Figure 4 compares (log) changes in the total district-level yield in the micro data to (log) changes in the total district-level yield implied by the model against the actual (log) changes in the level of transport costs over the period 1996–2014. While in the data there is more noise and the magnitudes of the changes in the district-level yields are larger than those produced by the model with only changes in transport costs, the U-shaped pattern of the district-level gains with transport costs changes is present in both the data and the model.

VI. Conclusions

This paper has studied a particular episode of a large-scale infrastructure intervention undertaken in Ethiopia starting in 1997. To measure the effects of the road expansion program, I combined a quantitative spatial framework with novel panel data on agricultural production and transportation costs. I find that the changes in transport costs implied by the expansion of the road network have had a sizable impact on productivity and the structure of the agricultural sector in Ethiopia. The gains in real output per unit of land are about one-tenth of the overall gains observed in the data. I also find that "closer" markets contribute to a structural transformation of the agricultural sector, with more export-oriented cash crop production, fewer farmers, and higher average farm size as employment shifts to other sectors of the economy. These effects are sizable and in the neighborhood of what occurs in the data.

At the individual district level, the gains are not uniform. The model produces a U-shaped relationship between district-level gains and transport cost changes that is similar in nature to the corresponding one in the data. In particular, the districts that experience the largest yield gains are not necessarily only those that experience the largest drops in the *level* of their transport costs. There are other factors driving heterogeneity in the responses of localities even after controlling for transport costs, such as relative transport costs changes across crops and the relative initial productivities.

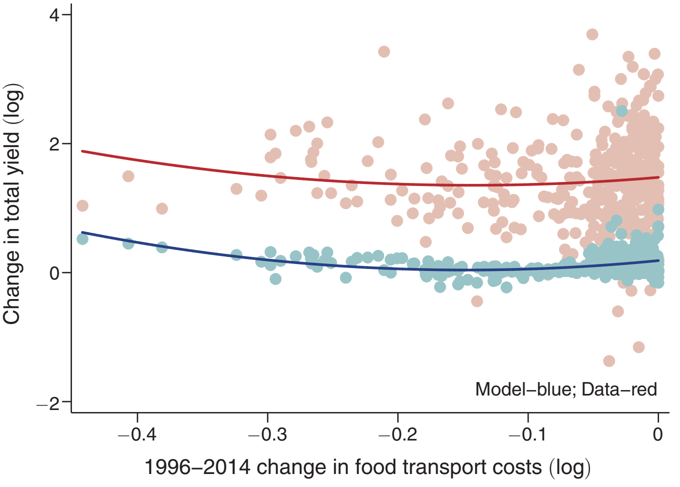

Figure 4. Model versus Data: Total Yield—Transport Cost Relationship (All Districts)

Notes: "Change in total yield" refers to the change in the real overall yield for each district (valued at a common set of prices). In the model this is the change relative to the benchmark economy after the reduction in transport costs. In the AgSS data this is the actual change over 1996–2014. The x-axis represents the log-change of food transport costs over 1996–2014.

The implication is that one should not expect a uniform response across regions to lowering transport costs across the board in the face of inherent heterogeneity. There is some evidence of spatial convergence across districts with relatively more distant and lower-productivity districts tending to exhibit higher productivity growth following the drop in transportation costs.

For an economy like Ethiopia that is heavily skewed toward agriculture, any productivity gains in this sector will translate to aggregate productivity benefits. This is a characteristic shared by many other developing countries, particularly in sub-Saharan Africa. I note that while the drops in transport costs have been large, Ethiopia started from a very high base, and their level as well as dispersion still remain high. The implication of the analysis here is that further investments in infrastructure expansion can have real productivity benefits for the economy. Finally, I note that the analysis has focused on quantifying the effects of transport infrastructure improvements for a given (1996) distribution of crop-specific TFPs across districts. If changes in the distribution of transport costs induce changes in the district-crop-specific TFPs—through, for example, the adoption of modern agricultural technologies and mechanization or the adoption of high-yield varieties—then the gains from transport infrastructure improvements can be potentially larger. A quantification of such gains would be a fruitful avenue for future research.

REFERENCES

Acemoglu, Daron, and Veronica Guerrieri. 2008. "Capital Deepening and Nonbalanced Economic Growth." *Journal of Political Economy* 116 (3): 467–98.

Adamopoulos, Tasso. 2011. "Transportation Costs, Agricultural Productivity, and Cross-Country Income Differences." *International Economic Review* 52 (2): 489–521.

Adamopoulos, Tasso. 2025. *Data and Code for: "Spatial Integration and Agricultural Productivity: Quantifying the Impact of New Roads."* Nashville, TN: American Economic Association; distributed by Inter-university Consortium for Political and Social Research, Ann Arbor, MI. https://doi.org/10.3886/E195703V1.

Adamopoulos, Tasso, and Diego Restuccia. 2014. "The Size Distribution of Farms and International Productivity Differences." *American Economic Review* 104 (6): 1667–97.

Adamopoulos, Tasso, and Diego Restuccia. 2022. "Geography and Agricultural Productivity: Cross-Country Evidence from Micro Plot-Level Data." *Review of Economic Studies* 89 (4): 1629–53.

Aggarwal, Shilpa, Brian Giera, Dahyeon Jeong, Jonathan Robinson, and Alan Spearot. 2022. "Market Access, Trade Costs, and Technology Adoption: Evidence from Northern Tanzania." *Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01263.

Alder, Simon. 2023. "Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development." Unpublished.

Allen, Treb, and Costas Arkolakis. 2014. "Trade and the Topography of the Spatial Economy." *Quarterly Journal of Economics* 129 (3): 1085–1140.

Allen, Treb, and Costas Arkolakis. 2022. "The Welfare Effects of Transportation Infrastructure Improvements." *Review of Economic Studies* 89 (6): 2911–57.

Asher, Sam, and Paul Novosad. 2020. "Rural Roads and Local Economic Development." *American Economic Review* 110 (3): 797–823.

Asturias, Jose, Manuel García-Santana, and Roberto Ramos. 2019. "Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India." *Journal of the European Economic Association* 17 (6): 1881–1940.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2012. "On the Road: Access to Transportation Infrastructure and Economic Growth in China." NBER Working Paper 17897.

Baum-Snow, Nathaniel, Loren Brandt, J. Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. "Roads, Railroads, and Decentralization of Chinese Cities." *Review of Economics and Statistics* 99 (3): 435–48.

Boppart, Timo. 2014. "Structural Change and the Kaldor Facts in a Growth Model with Relative Price Effects and Non-Gorman Preferences." *Econometrica* 82 (6): 2167–96.

Brooks, Wyatt, and Kevin Donovan. 2020. "Eliminating Uncertainty in Market Access: The Impact of New Bridges in Rural Nicaragua." *Econometrica* 88 (5): 1965–97.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin. 2021. "The Macroeconomics of Microfinance." *Review of Economic Studies* 88 (1): 126–61.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment." *arXiv:2107.02637*.

Caselli, Francesco. 2005. "Accounting for Cross-Country Income Differences." *Handbook of Economic Growth* 1 (a): 679–741.

Comin, Diego, Danial Lashkari, and Martí Mestieri. 2021. "Structural Change with Long-Run Income and Price Effects." *Econometrica* 89 (1): 311–74.

Costinot, Arnaud, and Dave Donaldson. 2016. "How Large are the Gains from Economic Integration? Theory and Evidence from US Agriculture, 1880-1997." NBER Working Paper 22946.

De Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2022. "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey." NBER Working Paper 29691.

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

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" approach." *Quarterly Journal of Economics* 131 (2): 799–858.

Donovan, Kevin. 2021. "The Equilibrium Impact of Agricultural Risk on Intermediate Inputs and Aggregate Productivity." *Review of Economic Studies* 88 (5): 2275–2307.

ERA (Ethiopian Roads Authority). 1996–2014. "GIS Road Network, 1996–2014."

ERA. 2013. *Route Selection Manual*. Addis Ababa: Government of Ethiopia.

ERA. 2016. *Road Sector Development Program: 19 Years Performance Assessment*. Addis Ababa: Government of Ethiopia.

Faber, Benjamin. 2014. "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System." *Review of Economic Studies* 81 (3): 1046–70.

Fajgelbaum, Pablo D., and Edouard Schaal. 2020. "Optimal Transport Networks in Spatial Equilibrium." *Econometrica* 88 (4): 1411–52.

Felbermayr, Gabriel J., and Alexander Tarasov. 2022. "Trade and the Spatial Distribution of Transport Infrastructure." *Journal of Urban Economics* 130: 103473.

GAEZ (Global Agro-Ecological Zones). 2000. "Global Agro-Ecological Zones, Version 3.0." Food and Agricultural Organization and International Institute for Applied Systems Analysis. https://www.gaez.iiasa.ac.at/ (accessed March 11, 2018).

Fried, Stephie, and David Lagakos. 2022. "Electricity and Firm Productivity: A General-Equilibrium Approach." *American Economic Journal: Macroeconomics* 15 (4) 67–103.

Gabre-Madhin, Eleni Z. 2001. *Market Institutions, Transaction Costs, and Social Capital in the Ethiopian Grain Market*. Washington, DC: International Food Policy Research Institute.

Gollin, Douglas, and Richard Rogerson. 2014. "Productivity, Transport Costs and Subsistence Agriculture." *Journal of Development Economics* 107: 38–48.

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

Gollin, Douglas, Stephen Parente, and Richard Rogerson. 2002. "The Role of Agriculture in Development." *American Economic Review* 92 (2): 160–164.

Herrendorf, Berthold, Richard Rogerson, and Akos Valentinyi. 2014. "Growth and Structural Transformation." *Handbook of Economic Growth* 2: 855–941.

Herrendorf, Berthold, James A. Schmitz Jr., and Arilton Teixeira. 2012. "The Role of Transportation in US Economic Development: 1840–1860." *International Economic Review* 53 (3): 693–716.

Jaworski, Taylor, Carl Kitchens, and Sergey Nigai. 2023. "Highways and Globalization." *International Economic Review* 64 (4): 1615–48.

Kebede, Selamawit, and Almas Heshmati. 2023. "Sectoral Linkage in the Ethiopian Economy: A Social Accounting Matrix Multiplier Analysis." Unpublished.

Kongsamut, Piyabha, Sergio Rebelo, and Danyang Xie. 2001. "Beyond Balanced Growth." *Review of Economic Studies* 68 (4): 869–82.

Lagakos, David, and Michael E. Waugh. 2013. "Selection, Agriculture, and Cross-Country Productivity Differences." *American Economic Review* 103 (2): 948–80.

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E. Waugh. 2023. "The Welfare Effects of Encouraging Rural–Urban Migration." *Econometrica* 91 (3): 803–37.

Manysheva, Kristina. 2022. "Land Property Rights, Financial Frictions, and Resource Allocation in Developing Countries." Unpublished.

Minten, Bart, Bethlehem Koru, and David Stifel. 2013. "The Last Mile(s) in Modern Input Distribution: Pricing, Profitability, and Adoption." *Agricultural Economics* 44 (6): 629–646.

Ngai, L. Rachel, and Christopher A. Pissarides. 2007. "Structural Change in a Multisector Model of Growth." *American Economic Review* 97 (1): 429–443.

Prasada Rao, D. S. 1993. "International Comparisons of Agricultural Output and Productivity." FAO Economic and Social Development Paper 112.

Restuccia, Diego, Dennis Tao Yang, and Xiaodong Zhu. 2008. "Agriculture and Aggregate Productivity: A Quantitative Cross-Country Analysis." *Journal of Monetary Economics* 55 (2): 234–50.

Schultz, Theodore William. 1953. *The Economic Organisation of Agriculture*. New York: McGraw-Hill.

Shamdasani, Yogita. 2021. "Rural Road Infrastructure & Agricultural Production: Evidence from India." *Journal of Development Economics* 152: 102686.

Sotelo, Sebastian. 2020. "Domestic Trade Frictions and Agriculture." *Journal of Political Economy* 128 (7): 2690–2738.

Storeygard, Adam. 2016. "Farther on Down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa." *Review of Economic Studies* 83 (3): 1263–95.

Timmer, Marcel P., Gaaitzen J. de Vries, and Klaas de Vries. 2015. "Patterns of Structural Change in Developing Countries." In *Routledge Handbook of Industry and Development*, edited by John Weiss and Michael Tribe, 65–83. Abingdon, UK: Routledge.

Tombe, Trevor. 2015. "The Missing Food Problem: Trade, Agriculture, and International Productivity Differences." *American Economic Journal: Macroeconomics* 7 (3): 226–58.