Roads, Competition and the Informal Sector

Elena Perra[a,1], Marco Sanfilippo[b,2], Asha Sundaram[c,3]

[a] *Department of Economics, University of Florence,*
[b] *Department of Economics and Statistics, University of Turin, and Collegio Carlo Alberto,*
[c] *Department of Economics, Univerity of Auckland,*

Abstract

We examine the impact of competition arising from better connectivity to markets on formal and informal firms. Combining geolocalized information on road improvements under a large infrastructure investment programme with data on manufacturing firms in Ethiopia, we show that an increase in competition is associated with higher labour productivity, capital-intensity, investment in physical capital and wages in the formal sector. On the contrary, there is no associated increase in labour productivity or wages in the informal sector. In fact, increased competition is associated with lower capital-intensity and investment, a shift in composition towards workers without primary education and a lower likelihood of operating in the informal sector. We thus highlight that the benefits of infrastructure improvement programmes may not accrue uniformly in the economy.

Keywords: Infrastructure, Roads, Market Access, Ethiopia, GIS, Productivity

[1] elena.perra@unitn.it
[2] marco.sanfilippo@unito.it
[3] a.sundaram@auckland.ac.nz

1. Introduction

Recent empirical work has emphasized the important role of the informal sector in job creation and structural transformation in developing countries (La Porta and Shleifer, 2014). In the manufacturing sector, though formal firms contribute more to productivity growth, a majority of firms are informal and employ a larger share of the workforce (Diao *et al.*, 2021; Kruse *et al.*, 2021). In spite of the prominent role played by informal firms in generating jobs, few analyses have looked at how policy reforms or other external shocks shape the composition of the manufacturing sector in terms of formal and informal firms.

This is also true of the literature on the impacts of large-scale infrastructure development programmes. Studies have looked at manufacturing as a whole, but not at differential effects across formal and informal manufacturing. The exception is a study by Chatterjee *et al.*, 2021 for India, who find that the benefits of public investments in infrastructure did not accrue to informal firms, while they were productivity-enhancing for formal firms. This paper is an attempt to fill this gap. We investigate the effect of increased competition from an improvement in road connectivity due to an extensive infrastructure development project in Ethiopia. Our analysis focuses on formal and informal manufacturing firms. This focus is important for various reasons. While formal sector firms can respond to competition by investing in better technologies, R&D and increasing efficiency De Loecker and Goldberg, 2013; Topalova and Khandelwal, 2011, informal firms lack such capacity. Factors such as inadequate access to credit and information and a low level of education among informal entrepreneurs (La Porta and Shleifer, 2014) can hamper their ability to respond to increased competition from firms in connected areas. Furthermore, the increase in competition is likely to impact firm selection as in Melitz, 2003, as the least productive firms exit. Dynamics of firm adjustment may differ between the formal and informal sectors, given the informal sector's role in providing a means of survival in developing countries.

To examine the impacts of competition from an increase in connectivity, we combine granular, geolocalized, information on road improvements under the Ethiopian Road Sector Development Programme (RSDP) with firm level data from the formal and informal sectors of Ethiopia. We utilize the Small Scale Industries Survey, covering small and informal firms, and the Survey on Large and Medium Manufacturing Firms, which provides data on the formal sector. We explore a range of firm outcomes, such as the likelihood of operating informally, choice of technique, investment in physical capital, labour productivity and the composition of workers given their level of education. Ethiopia is an excellent case for various reasons. First, the RSDP was a massive road improvement programme commencing in July 1997 to improve connectivity in the country. New roads were built and existing roads upgraded in quality, generating variation across time and space in improvements in the road network and reductions in travel time (Fiorini *et al.*, 2021). Reductions in travel time occurred not just because the road network expanded, but also because roads were upgraded and better quality roads (such as paved, relative to gravel) allow greater speeds. Second, roads dominate transport in Ethiopia, which is characterized by an almost complete lack of infrastructure

substitutes[4]. The Ethiopian road network accounts for 90-95% of total inter-urban freight (Worku, 2011). Therefore, road improvements and the resulting expansion of the transport network and reduction in travel time produce sizeable changes in trade cost for firms. Finally, the availability of granular data on road improvements and firm location, complemented with data on formal and informal firms allows us to isolate the effects of road improvements on firm outcomes in a quasi-experimental setting.

Our empirical analysis follows the existing literature (Donaldson and Hornbeck, 2016; Huang and Xiong, 2018; Alden, 2019; Jedwab and Storeygard, 2020; Fiorini *et al.*, 2021) and adopts the spirit of the market-access approach to construct a measure of competition from greater road connectivity. We begin by treating each Ethiopian district (*woreda* in the Ethiopian context) as a local market. For each district and industry of a firm in a given year, we construct a weighted average of the inverse of travel times to all other districts given the road network and travel speed (which depends on the quality of the road), where the weights are total production in the district and industry. Variation in this measure captures both variation in production in the firm's industry in connected markets and variation in travel times as roads are expanded and improved. It is thus a time-varying measure of changes to competition faced by the firm both in its local and in connected markets as the road network evolves. Borrowing from the literature (Donaldson, 2018; Huang and Xiong, 2018), we refer to this measure as consumer market-access (CMA) to convey the idea that it measures the availability of alternatives for consumers. Of course, as travel times decrease with the improving road network, firms will also have better access to consumers in connected markets (an improvement in the firm's market-access (FMA), following (Donaldson, 2018)). We account for this improved FMA with fixed effects that vary by district and year. Arguably, this effect is uniform across industries, so that the CMA measure primarily captures competition from producers in connected markets. We then relate CMA to firm outcomes such as productivity, capital-labour ratio, investment, wages and skill composition of the workforce.

Identification of the causal effects of competition from road connectivity improvements on firm outcomes encounters the canonical concern of endogeneity. We discuss each of the two components of our CMA measure: production and travel time. We argue that production in connected markets is largely exogenous to local firm outcomes, given that each district is a small economy relative to all the other Ethiopian districts. Travel time given the road network is more likely to be endogenous. As reported by Gebresilasse (2018), Moneke (2019), and Fiorini *et al.* (2021) for the Ethiopian case, road construction under the RSDP was potentially non-random, both in terms of timing and placement. It is plausible that the allocation decision of policymakers was driven by several considerations, which range from higher economic and social potential of particular districts to political motivations (Burgess *et al.*, 2015). For instance, using the same data for Ethiopia, Perra, 2022 provides evidence

[4]Ethiopia has no direct access to the sea, nor does its territory contain any transportation substitutes to roads, such as navigable rivers, canals or railroads, with the exception of a single railroad line to Djibouti, which was not functioning during our study period (Fiorini *et al.*, 2021).

that areas connected ethnically and politically to the ruling party received more roads, and roads of better quality. For this reason, our identification strategy relies on an instrumental variable approach that exploits road improvements occurring outside "exclusion areas", defined as the surroundings of the pre-RSDP (1969) Ethiopian road network. The rationale is that incremental investments in roads are expected to be endogenous, given their proximity to pre-existing road arteries and connections between towns.

We find that the impacts of competition, measured by CMA, vary substantially across formal and informal sector firms. Among formal firms, a one standard deviation increase in CMA corresponds to a 0.58% increase in labour productivity, an effect almost twice as big as the effect for the sample taken as a whole. For formal firms, we report evidence of increases in CMA being linked to improvements in capital-intensity, investment in physical infrastructure and wages. Among informal firms, we find no relationship between CMA and labour productivity. If anything, the relationship is weakly negative. Results show that an increase in CMA reduces the capital-labour ratio and investments among informal firms. In addition, exploiting information on the level of education of each individual worker within a firm from the SSIS, we find that an increase in CMA is associated with a larger share of workers without primary education and a smaller share with higher education. Furthermore, the likelihood of operating informally decreases with an increase in CMA.

These findings are consistent with a framework where greater competition is associated with exit of less productive formal firms, while surviving ones respond to competition by adopting more capital-intensive technologies and improving efficiency. Labour released from the formal sector is displaced to the informal sector, where firms see decreases in capital-intensity, investment, the share of educated workers and labour productivity. However, exit of less productive firms in the informal sector also means that surviving informal firms are of higher productivity, countering the negative effect to a certain extent. Overall, our results highlight that competition from better road connectivity due to road infrastructure improvements may disadvantage the informal sector, as it disciplines the formal sector.

This paper speaks to the literature looking at the benefits of public investments in infrastructure (Faber, 2014; Duranton and Turner, 2012) and at the literature studying the role of geography in influencing firm choices (Puga, 2009; Redding and Turner, 2015; Bernard *et al.*, 2012; Redding, 2020). We augment this literature by explicitly focusing on the informal sector and underlining the differential effects of such investments on informal firms, whose adaptation strategies and constraints differ from those of formal firms. Our study is also related to the literature that looks at exogenous trade shocks on informality Goldberg and Pavcnik (2003), though we look at reductions in intranational trade costs and an increase in domestic competition. The trade literature has revealed contrasting findings about the impacts of trade liberalization on informality Dix-Carneiro *et al.* (2021), McMillan (2019), and McCaig and Pavcnik (2018). Closest to our setting is the paper by Nataraj (2011), who looks at the impact of India's unilateral trade liberalization on productivity of formal and informal firms and underscores the importance of delving into the differences between them as they adjust

to market shocks.

Finally, we contribute to a small but growing strand of evidence on the implications of the RSDP in Africa. While previous work has investigated the impact of roads on firm productivity (Fiorini *et al.*, 2021; Shiferaw *et al.*, 2015), no evidence has so far been available on the informal sector. In this respect, our findings are in line with the work by Diao *et al.* (2021), who find that the productivity benefits of global integration accrue disproportionately to formal firms at the top of the distribution, with gains concentrated in productivity than in employment. We thus emphasize the tension between inclusive employment growth and enhancements in productivity.

The remainder of the paper is organised as follows: Section 2 outlines a conceptual framework, Section 3 provides overview on the context of the Ethiopian RSDP; Section 4 describes the data employed in the analysis; Section 5 outlines the empirical approach adopted for this study; Section 6 reports the results, with robustness checks shown in Section 7. Finally, Section 8 concludes.

2. Conceptual Framework

In this section, we draw upon the vast literature on trade liberalization and its impacts on firms to outline a conceptual framework for our analysis. The goal of our research is to identify the impacts of an increase in competition from firms in connected markets on the formal and informal sectors. Specifically, we appeal to Melitz, 2003, Topalova and Khandelwal, 2011 and De Loecker and Goldberg, 2013. We posit that firms are heterogenous in their productivity. Entry into the formal sector requires payment of a fixed cost, following on literature that documents costs of registration, setting up a tax identity and licensing costs to operate in the formal sector. Similarly, export to connected markets entails a fixed cost of hiring intermediaries and setting up transportation routes.

Thus, it is only firms with productivity high enough to pay the fixed cost of entry to the formal sector that operate formally, and only firms with productivity high enough to pay the fixed cost of servicing connected markets that outside their local area. As in Melitz, 2003, an increase in competition from firms in connected markets results in contraction of output and forces the least productive firms in the formal sector to exit (some of these exiting firms may eventually end up in the informal sector). Additionally, as argued in Topalova and Khandelwal, 2011; De Loecker and Goldberg, 2013, competition is associated with investments in better technology, R&D and increases in efficiency at the firm level, which translates to higher labour productivity in the presence of capital in the model.

Turning to the informal sector, we first note that not only are there fundamental differences between the formal and informal sectors, but the two sectors also interact with each other. With increased competition, informal firms contract and less productive informal firms exit. Though more productive informal firms may be hampered in making investments to improve

technology and efficiency, we expect aggregate productivity of surviving firms in the informal sector to be higher through the exit channel. An additional channel that yields the opposite relationship is the interaction between the formal and informal sectors.

As less productive formal firms exit, workers at these firms are unlikely to be absorbed in the more productive firms, who respond to competition by adopting more sophisticated technology and capital-intensive techniques of production (Diao *et al.*, 2021). This is particularly true of less educated workers, since technology is intensive in high-skilled labour. In a setting where the informal sector is not only a source of employment but also of survival (in the absence of social welfare), informal sector firms absorb these workers, becoming "labour absorbing" firms (Diao *et al.*, 2021). Thus, through this channel, we would expect to see a decrease in the capital-intensity and labour productivity, and an increase in the share of less educated workers in informal sector firms. The relationship between competition and labour productivity in the informal sector would thus depend on the dominant channel. In fact, we expect to find no relationship if the two channels completely counteract each other.

3. The Road Sector Development Programme

The Road Sector Development Programme (herafter, RSDP), is part of a broader action, involving reform processes targeted to fighting poverty, boosting the economy and favouring structural transformation. The RSDP is an investment programme implemented in Ethiopia since 1997 and still ongoing, with the objectives of rehabilitating existing Ethiopian roads and constructing new networks. It has been estimated that Ethiopian Road Network has increased from 26550 km in 1997 to 113066 km in 2016, while the proportion of the country's rehabilitated roads has increased from 22% to 72%. Therefore, road density per 100 sq. km has risen significantly from 21.1 km in 1997 to 102.8 km in 2016 (Ethiopian Road Authority, 2016; World Bank, 2021). The main authorities in charge of its implementation were the Ethiopian Roads Authority (ERA) and the Regional Roads Authorities (RRAs). This large-scale development project has attracted particular interest by the academic world, with the objective of analysing the impact of infrastructural investments on agricultural productivity, businesses' performance and structural transformation processes (Adamopoulos, 2019; Shiferaw *et al.*, 2015; Fiorini *et al.*, 2021; Fiorini and Sanfilippo, 2019; Moneke, 2019). The RSDP has been a key promoter of increasing agricultural productivity, sped up business activity and stimulated structural transformation, directly from the agricultural to the service sector.

One of the main goal of the Ethiopian government is to reach the level of a lower middle-income country by 2025. In order to implement the "Ethiopian Vision", various reforms have taken place to call for sustainable growth processes, involving different spheres of development. These reforms range from poverty alleviation programmes and the commercialization of agriculture to the canalization of resources for the transformation of the private sector, alongside a strengthening of the public sector to promote good governance practices, boosting women and youth empowerment, and ensuring their participation in the democratic processes (Ministry of Finance and Economic Development, 2010). In order to accelerate these

transformations, road infrastructure has been considered as a crucial trigger to revitalize the country's long-term sustainable growth prospects (Worku, 2011; MoFED, 2006). Indeed, the lack of infrastructural networks and the consequently high transport costs have been identified as two of the major hurdles to growth for most developing countries, but particularly so for the African continent (Esfahani and Ramire, 2003). Studies have shown how high transport costs pose important barriers to labour supply and hinder significant market opportunities (Franklin, 2018; Atkin and Donaldson, 2015). Therefore, efficient infrastructural facilities are considered essential to generate virtuous economic and social processes (Donaldson, 2018; Storeygard, 2016; Alder, 2015). Numerous studies have indicated that spending in infrastructure is one of the main channels on which governments can rely to promote growth, international economic attractiveness, poverty alleviation and affect several other dimensions of local development (Banerjee *et al.*, 2020; Berg *et al.*, 2017; Redding and Turner, 2015).

The RSDP is considered the largest infrastructural investment project ever implemented by the Ethiopian government and one of the most ambitious of the entire region (Shiferaw *et al.*, 2015). Its construction has required a significant investment in terms of foreign currency, with an estimated cost of its implementation during this fourteen years of around US$7.08 billion (Worku, 2011; Shiferaw *et al.*, 2015). The five-year plan of the RSDP has been implemented thorough annual action plans, closely supervised and influenced by the government (Shiferaw *et al.*, 2015). Although the ERA has assigned different criteria for the road upgrading projects, it is not clear which specific variables the ERA employs to operationalize them (Worku, 2011; Shiferaw *et al.*, 2015)[5].

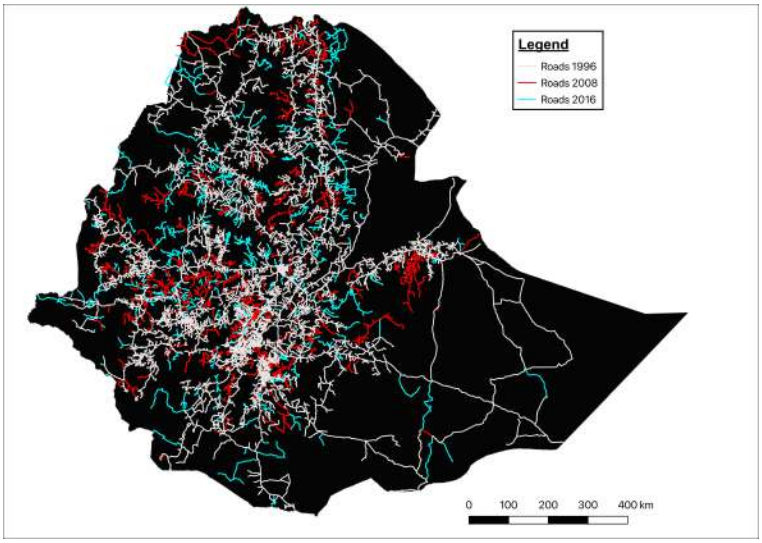

Figure 1: Visualization of RSDP's improvements in road connections. Source: Authors' calculations on RSDP data

[5]With the subsequent potential problems connected with ethnic/political favouritism mechanisms.

Improvements due to the implementation of the RSDP are documented in Figure 1, which shows the remarkable advancement in terms of road network from 1996 (i.e. the state of the road sector at the baseline) and two successive periods, 2008 and 2016. The registered average travel time to major economic centers declined from 385.6 to 322.5 hours for the period 1996-2008 (Shiferaw *et al.*, 2015). The improvements in multidimensional aspects of road accessibility indicators are reported in Table 1, which shows a general increase in the proportion of roads in good and serviceable conditions (ERA, 2009).

Indicator	1997	2009
Proportion of asphalt roads in good condition	17%	70%
Proportion of gravel roads in good condition	25%	54%
Proportion of rural roads in good condition	21%	50%
Proportion of total road network in good condition	22%	54%
Road Density/1000 sq.km	21.1km	42.6km
Road Density/1000 Population	0.46km	0.57km

Table 1: Improvements in Road Infrastructure. Source: ERA (2009)

The improvements in road quality are captured by the registered drop in travel time needed to cross each segment of road. The data on road quality improvement are aggregated in the speed matrix proposed by ERA (2009) and reported in Table 2, which describes the average travel speed as a function of the particular road surface.

	Average Travel Speed	
Pavement Type	Before Upgrading	After Upgrading
Asphalt Roads	50 km/h	70 km/h
Major Gravel Roads	35 km/h	50 km/h
Minor Gravel Roads	25 km/h	45 km/h
Earth Surfaced Roads	20 km/h	30 km/h

Table 2: ERA Travel Speed Matrix

4. Data

4.1. Firm-level Data

We combine two sources of microdata covering the whole manufacturing sector in Ethiopia.

The first is the Large and Medium Manufacturing industry Survey (LMMS), an annual census of more structured firms published by the Central Statistical Agency (CSA). Data cover all firms with at least 10 persons engaged and that use electricity in their production process. These firms are required to respond to this census every year; therefore, this source include the universe of large and medium firms in the manufacturing sector. The census records provide information on the characteristics of each establishment, as well as detailed information on the size and composition of the workforce and on the location of each firm. Firms also provide details on sales values and quantity produced for the domestic and international market for each product, as well as information on raw materials, both domestic and imported, employed at the firm level for the production processes, and their share on total firm expenditure. Manufacturing industries are defined at the 4-digit level according to the ISIC Rev. 3 classification, while products are recorded following an internal classification by the CSA.

The second dataset is the Survey of Small-scale manufacturing Industries (SSIS). We combine all existing waves of the SSIS, covering the years 2001, 2004, 2007, 2010, 2013. This is a survey that covers small (i.e., those engaging less than 10 persons) and mainly informal firms operating in the manufacturing sector. The sample is single-stage stratified, considering six main industries (textiles and garments, metal work, wood work, leather and leather products, other manufacturing sectors and the grain mills industry), sampled in similar proportions across regions. Due to the lack of a proper sample frame, it is not necessarily representative of the sector but provides considerable information on the activities of smaller firms, which comprise the majority of firms in the country. Table A.1 reports precise figures for the years in which the SSIS and the census were run simultaneously. On average and consistently over time, small and informal firms represent the large majority of all manufacturing establishments, approximately half of total manufacturing employment, but a much smaller share in terms of total production, wage bill and capital expenditures[6].

We combine the two datasets (LMMS and SSIS), obtaining information at the firm, industry and *woreda* level. Note that, since the data included in the SSIS are based on a similar questionnaire than the census, we can compare most of the indicators without incurring in problems connected with measurement errors.

Based on the pooled data, we define informal firms as those who do not keep a book of account and have less than 10 employees[7]. This operationalization of informality is the closest one to the official definition that the CSA gives of an informal firm, i.e. one: "*that do not keep*

[6] See Table A.2 in the Appendix for an overview of the role played by the small business sector in Ethiopia.
[7] In Section 7.1 we also investigate alternative definitions of informality as a robustness check.

complete books of accounts; mainly engaged in market oriented production; do not register the enterprise and its employees; and have a very limited number of persons engaged (less than ten person) in the enterprise; have no license" (Siba, 2015).

From Table 3, coherently with expectations, an informal firm appears to be on average smaller, younger, less capital intensive and less export/import oriented with respect to a formal firm. This is reflected in Figure 2, which shows the productivity distributions of formal and informal firms. The wider right tail of the formal firms' distribution confirms their higher productivity level with reference to informal ones, which are concentrated in the left hand side of the distribution. In terms of their spatial collocation, Figure A.1 in Appendix, shows that they appear to be equally represented on the Ethiopian territory, with an higher concentration of both types of firms in the areas surrounding Addis Abeba[8].

	Informal	Formal
Variable	Mean	Mean
Labour Productivity	99.2	414.9
Capital	39.4	191.7
Employees	1.2	2.4
Wage_{pc}	7.3	28.0
Exports	1.9	3779.9
Imports	70.9	31285.1
Age	5.9	9.4
Assets	103.6	12030.5

Notes: Values in thousands ETB, except for Employees and Age

Table 3: Summary Statistics

[8]In Figure A.2 in the Appendix, it is reported a visual representation of the productivity of formal and informal firms.

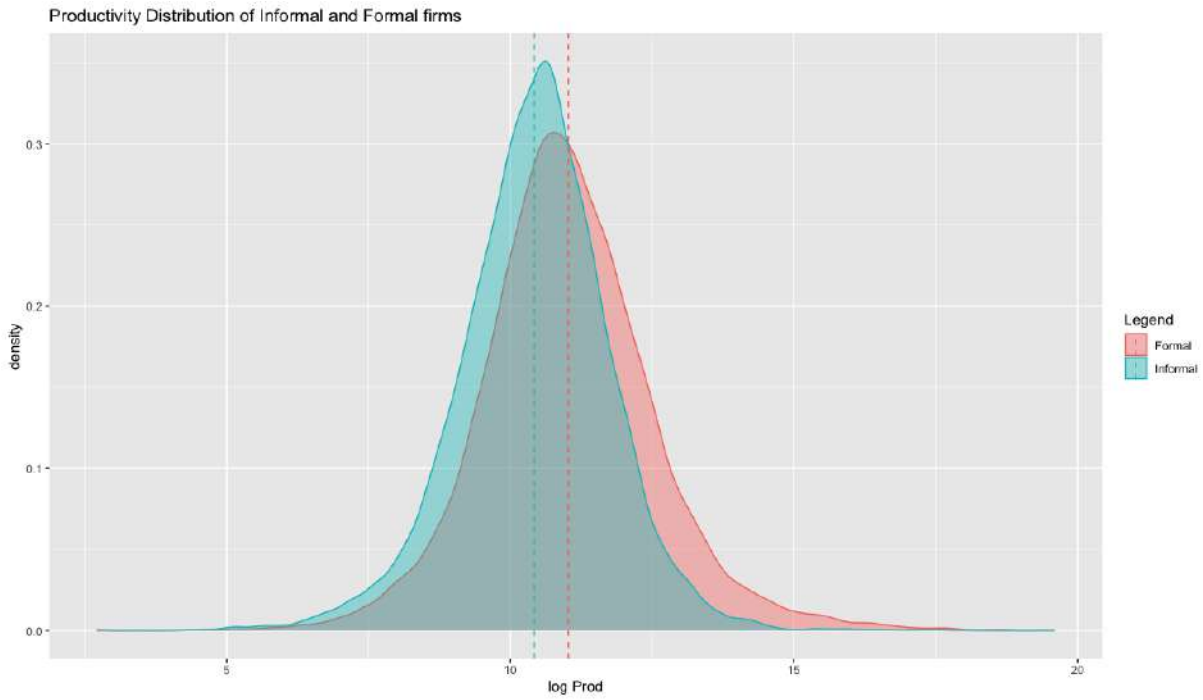

Figure 2: Productivity Distributions. Authors' calculations.

4.2. Roads

This paper employs rich geolocalized data on the Road Sector Development programme (RSDP), which spans a period going from 1996 to 2016. This database consists of time series shapefiles of the Ethiopian road network, describing the incremental improvements in terms of road surface (earth surface, minor gravel, major gravel and asphalt) and travel time needed to cross each road segment. Moreover, it provides details about whether a road-segment is categorised as not-rehabilitated, rehabilitated or completely new. This exhaustive and detailed data represents a significant source of information, with respect to the usually limited availability of time series data on transport infrastructures in low-income developing countries. The availability of information on the improvements in road surface allow to calculate the enhancements in terms of the average travel time needed to cross each segment of road, in accordance with the speed matrix reported by the ERA, and employed also by Shiferaw *et al.* (2015), Jedwab and Storeygard (2020), and Fiorini *et al.* (2021).

In the rest of the paper we employ a market access approach based on Donaldson and Hornbeck (2016) and Jedwab and Storeygard (2020):

$$MarketAccess_{xt} = \sum_{d \neq x} Weight_{dt} * \tau_{xdt}^{-\theta} \tag{1}$$

This indicator measures the minimum distance between the centroid of district x (origin) and the centroid of district d (destination) given the road and speed at year t; while $Weight_{dt}$ is a

proxy of district-level economic activity. The minimum distance in hours τ, is calculated employing the Dijkstra's algorithm; θ is the elasticity measuring how trade volumes fall as travel time increases. Empirical papers using a market access approach have resorted to different values of θ, usually in the range of 1 (the market potential approach originally proposed by Harris (1954)) to about 10 (Donaldson and Hornbeck, 2016). For our analysis we rely on a value of θ equal to 3.8. This is the same value adopted by Jedwab and Storeygard (2020) in their paper looking at the effects of road improvements on urbanization in Africa. Jedwab and Storeygard (2020) obtained this value using the estimated cost-distance elasticity for Nigeria and Ethiopia from Atkin and Donaldson (2015), which is 3 times larger the one found by Duranton *et al.* (2014) for the US[9]. In the robustness checks, reported in Section 7.2 we show that results are robust to different values of θ.

Market access is a powerful indicator able to account for different dimensions of the role of road improvements, including: (i) its capacity to account for the falling cost of transporting goods, allowing firms to sell their products to bigger markets in the country; and (ii) its capacity of increasing competition between firms. To some extent, an increase in market access can be viewed as a domestic shock, similar for several aspects to an international trade liberalization process. In the remaining of the paper we focus on the latter dimension, the Consumer Market Access (CMA). CMA accounts for how proximity to competitors based elsewhere in the country changes with improvements in the road network (Huang and Xiong, 2018). The definition of CMA is based on a modified version of equation 1, in which changes in travel time *tau* are weighted using the level of total production in each wereda-industry pair jx at time t:

$$CMA_{jxt} = \sum_{d \neq x} Production_{jdt} * \tau_{jxdt}^{-\theta}$$

Figure 3 visualizes average levels of CMA for formal and informal firms, separately.

[9] Duranton *et al.* (2014) cost-distance elasticity for the US is 1.27, meaning that $1.27 * 3 \approx 3.8$

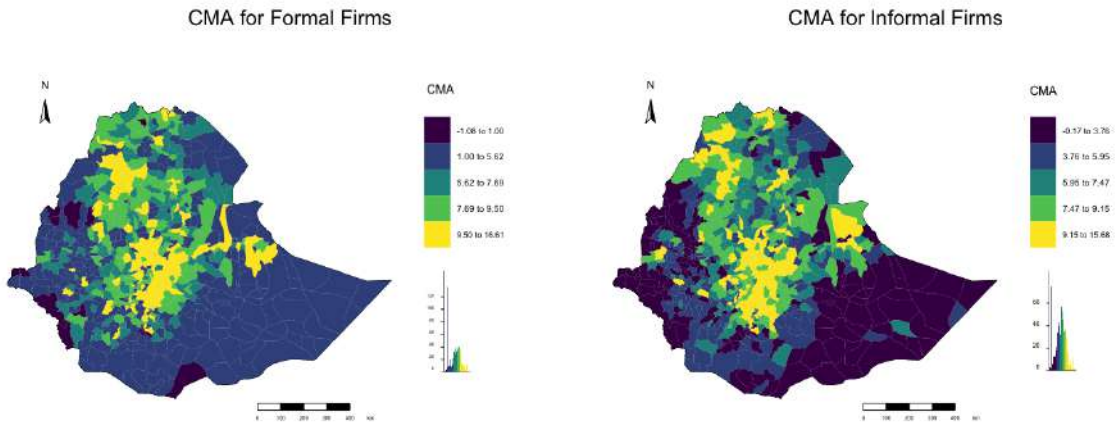

Figure 3: Visualisation of Consumer Market Access (CMA). Source: Authors' calculations on SSIS and LMMS data.

5. Identification Strategy

Our empirical analysis is based on the following equation:

$$Y_{ijxt} = \beta_1 CMA_{jxt} + \beta_2 X_{it} + \theta_{jx} + \mu_{rt} + \epsilon_{ijxt} \tag{2}$$

which estimates the effects of CMA on (formal and informal) firms' labour productivity, according to the definition provided above. Since we employ a repeated cross-section of firms, our identification strategy exploits changes within markets over time. For this reason we add industry-wereda, jx fixed effects. Wereda-year, xt, fixed effects are added to account for time specific changes occurring within each district over time, and are important to remove factors, such as changes in economic activity, that can affect the relation of interest. X includes firm specific controls, that we limit to the age of the firm and to a dummy variable accounting for the fact that a firm was originally surveyed in the SSIS or in the LMMS. Standard errors are clustered at the district level.

This approach needs to deal with endogeneity issues regarding the placement of transport infrastructure. Indeed, the choice of where to build infrastructure is not exogenous, since the random assignment of route placements is implausible. It is reasonable to assume that planners may have decided to allocate investments with specific goals – e.g., where high growth was expected or in specific peripheral counties connecting target nodes (Asher and Novosad, 2020; Duflo and Pande, 2007). Moreover, these domestic shocks in trade costs may indirectly affect locations that have not been promptly connected, such as those that are in between (Chandra and Thompson, 2000). Indeed, these shocks automatically translate into shocks in income, which are for their nature spatially correlated (Alden, 2019). The subsequent construction of the market access indicators is mechanically endogenous as well, since growth in a district may affect growth in neighboring districts, which in turn affect their neighbors with a positive feedback mechanism. A further challenge is brought forward by the possibility of incurring in omitted variable bias. Indeed, market access may be altered both by sudden population changes of the locations of trading partners, by changes in the road network connecting them, or by natural impediments and other unmeasured factors, which may influence the allocation decisions of road investments. Moreover, roads could be built in advance to benefit from the economic growth prospects of neighbouring cities or nearby economic hubs, or vice versa to avoid possible economic stagnation (Jedwab and Storeygard, 2020). Finally a further concern may arise from the will of policymakers to attract larger shares of infrastructure investments, which could be correlated across locations (Fiorini *et al.*, 2021). In the specific context of the RSDP in Ethiopia, such anticipation effects should be mitigated by the structure of the RSDP, which is linked to a five-year investment plan model. In order to deal with such potential concerns, we propose an instrumental variable approach which combines the insights brought forward by Moneke (2019), Gebresilasse (2018), Fiorini *et al.* (2021) and Jedwab and Storeygard (2020). This technique consists in the creation of alternative road networks, which are used as instruments to the market access measures of equation (2), by

exploiting the variation in bilateral costs (i.e. travel time) generated by these synthetic roads.

By digitizing the CIA's map of the Ethiopian territory, United States Central Intelligence Agency (CIA) (1969)[10], we apply a strategy similar to the one adopted by Jedwab and Storeygard (2020) and Fiorini *et al.* (2021). We identify a 50 km buffer following the trajectories of the digitised roads of 1969. These exclusion areas are employed in the construction of the synthetic road network as the areas within which our road network remains exactly as it was before the RSDP, while outside the exclusion zones, it changes according to the RSDP improvements. "Freezing" roads in the areas inside the buffer at the state of the art of 1996 enables us to exclude from the analysis all the enhancements of the RSDP that are more likely to be affected by endogenous drivers. Indeed, it is reasonable to assume that the exclusion zones reflect the principal Ethiopian thoroughfares, with the highest level of attractiveness for long-term investments. As shown by Bertazzini (2022), there is a long-run effect of transport networks and concentration of economic activity in Ethiopia.[11].

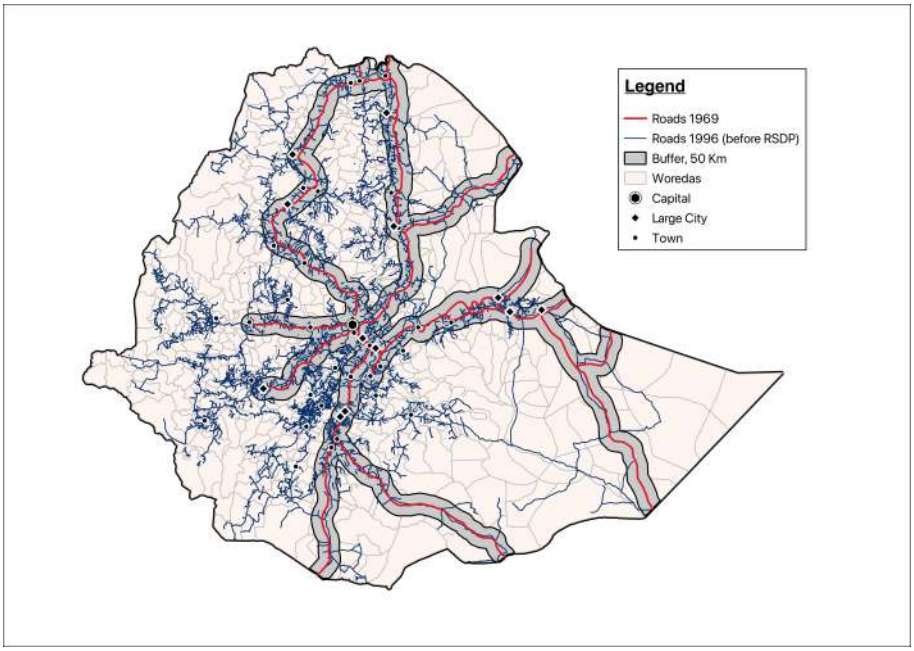

Figure 4: Buffer of 50 Km around roads of 1969. Authors' calculations.

[10]Reported in Figure A.3 in the Appendix

[11]Bertazzini (2022) shows how proximity to colonial roads, and therefore lower transport costs, generate high level of attractiveness for economic activity until the 1960s. In turn, this generates a positive feedback mechanism, driving investments in areas with higher advantages in terms of increasing returns of scale, meaning zones with higher level of economic development. During the Italian occupation 1935-1941, the Italian road programme built a total of 7000 km of roads, 3450 km of which were tarred (Baker, 1974), which where primarily designed to serve military purposes. By the time of the Liberation until 1951 colonial roads were kept operational, but only after 1960 other major road arteries were built. In particular, Imperial Highway Authority launched a major series of infrastructure projects which from 1951 to 1968 led to the construction and improvement of 7304 km of highway, with the goal of reaching all the main cities and towns linked to the capital and improving the access to the Lakes Region and the coffee-producing areas (Baker, 1974).

Therefore, by exploiting the buffer around the roads of 1969, we are confident to take into particular consideration the principal historical Ethiopian lanes, which are those that connect the main city centers and economic hubs of the country. In this way, we are able to account for endogenous local road improvements, but also for those improvements targeted at connecting the principal Ethiopian economic hubs and cities. As shown in Figure 4, the 50 km buffer includes all the major cities centers with a population larger than 50,000 people in 1994, and involves most of the Ethiopian cities. Differently from Jedwab and Storeygard (2020), we follow the strategy proposed by Fiorini *et al.* (2021) and we compute the bilateral distances between an origin centroid x of woreda of x, and a destination centroid d of woreda d, without taking into account road changes in the exclusion zone. We decide to follow this strategy since the RSDP project is a national infrastructural plan, implying that it is essential to consider endogenous political factors and spatial autocorrelation across Ethiopian regions and woredas.

6. Results

6.1. Formal and informal firms productivity - OLS and 2SLS

In Table 4 we use ordinary least squares (OLS) to estimate the relationship between labour productivity at the firm level and consumer market access (CMA), which represents competition and varies at the woreda, industry and year level. As hypothesized in our conceptual framework, we expect a positive relationship between competition and labour productivity in the formal sector, while the relationship for the informal sector depends on the dominant channel between firm exit (which raises the aggregate productivity of surviving firms) and labour absorption as the least productive formal firms exit (which depresses firm labour productivity if capital is held constant). All regressions include woreda-time and industry fixed effects. Columns (1) and (2) report results for the whole sample of informal and formal firms. Columns (3)-(4) and (5)-(6) report results for the informal and formal sector firms respectively. While columns (1), (3) and (5) do not include control variables, columns (2), (4) and (6) include controls for firm age and size.

Table 4: Results OLS: Labour Productivity

	WHOLE SAMPLE		INFORMAL		FORMAL	
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity	(5) Labor Productivity	(6) Labor Productivity
CMA	0.0600**	0.0587***	-0.0100	-0.0140	0.134***	0.135***
	(0.0265)	(0.0212)	(0.0227)	(0.0229)	(0.0297)	(0.0267)
Constant	10.50***	11.11***	10.79***	11.28***	10.27***	10.50***
	(0.253)	(0.257)	(0.207)	(0.223)	(0.311)	(0.290)
Observations	23,350	23,232	13,772	13,762	8,668	8,563
R-squared	0.405	0.430	0.457	0.460	0.344	0.362
District-Industry FE	Y	Y	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

We find that the coefficients of interest are positive and significant for firms in the formal sector. This is consistent with the idea that increased competition is associated with exit of less productive formal firms, resulting in greater aggregate productivity. In addition, competition can spur firms to invest in better technology, become more efficient and increase productivity. The relationship between competition and labour productivity is negative for firms in the informal sector, but the coefficient is not statistically significant.

Note that we expect aggregate productivity in the informal sector to increase through the exit channel. However, as firms contract in response to competition, less productive formal firms exit and surviving firms adopt more sophisticated technology, informal sector firms absorb these workers (more so the less educated ones), becoming "labour absorbing". Hence, through this channel, we would expect to see a decrease in the capital-intensity and labour productivity of informal sector firms. Our results indicate that this channel dominates, though weakly.

Broadly, our results are in line with the findings of Chatterjee *et al.* (2021) for India, we find different impacts of the increase in connectivity and resulting competition for informal and formal sector firms.

Table 5: Results IV: Labour Productivity

	WHOLE SAMPLE		INFORMAL		FORMAL	
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity	(5) Labor Productivity	(6) Labor Productivity
CMA	0.0622**	0.0600***	-0.0115	-0.0150	0.137***	0.137***
	(0.0266)	(0.0217)	(0.0245)	(0.0247)	(0.0301)	(0.0271)
Observations	23,350	23,232	13,772	13,762	8,668	8,563
R-squared	0.001	0.043	0.000	0.006	0.006	0.033
District-Industry FE	Y	Y	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
F-test	3241	3224	3154	3157	2527	2472

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Next, we provide results based on an instrumental variable approach using two stage least squares (2SLS). Second stage results are reported in Table 5, by exploiting instrumental variables constructed as discussed in Section 5. In order to isolate the exogenous improvements to the road network, we remove all upgrades in road construction defined in the exclusion zones around the roads present in 1969 in the construction of our IV. The first stage regression is reported in Table A.4 and shows that our instrument is a strong predictor of CMA. Similarly, the first stage F-statistic appears to be strong, confirming the relevance of our instrumental variable. It is also reassuring to see that results of the reduced-form (reported in Table A.5 in the Appendix) remain consistent. Table 5 shows that our results are consistent with the OLS results reported earlier. They show that one standard deviation increase in CMA corresponds to a 0.58% increase in productivity of formal firms and a 0.253% increase for the whole sample, in accordance with Table 4. This reassures us that our baseline results in Table 4 are not contaminated by endogeneity bias.

6.2. Formal and informal firms productivity - Heterogenous Effects

In Table 6 and 7, we explore heterogenous effects across districts with above and below median share of informal firms by number (columns (1) and (2)) and size measured by employment (columns (3) and (4)) for informal and formal firms respectively. The idea is to explore whether competition from better connectivity due to road improvements impacts labour productivity in the informal and formal sectors differently when the relative size of the informal sector in the area is high. In such areas, greater movement of workers into the informal sector as formal firms respond to competition is likely to crowd the informal sector, lowering labour productivity.

Indeed, we find in Table 6 that in areas with above median share of informal firms, CMA is associated with lower labour productivity and the coefficient is now statistically significant. This is not the case in areas with below median share of informal firms. From Table 7, we find that this heterogeneous impact is unique to the informal sector. In the formal sector, the increase in labour productivity with competition is largely uniform across areas with above and below median labour productivity. Thus, results in Table 6 and 7 lend further support to the channel whereby competition resulting from better road connectivity exerts differential effects across the formal and informal sectors, with negative productivity effects concentrated in the informal sector.

Table 6: Heterogenous Effects - Informal Sector: Labor Productivity

	Above Median	Below Median	Above Median	Below Median
	(1)	(2)	(3)	(4)
VARIABLES	Labor Productivity	Labor Productivity	Labor Productivity	Labor Productivity
CMA	-0.0746**	0.0175	-0.0684*	0.0144
	(0.0358)	(0.0371)	(0.0376)	(0.0375)
Observations	5,602	7,868	5,298	8,168
R-squared	0.003	0.007	0.003	0.007
District-Industry FE	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
F-test	1320	1674	1695	1872

Robust standard errors in parentheses
(1) & (2) Share of Informal Firms
(3) & (4) Share of Informal Size
*** p<0.01, ** p<0.05, * p<0.1

Table 7: Heterogenous Effects - Formal Sector: Labor Productivity

	Above Median	Below Median	Above Median	Below Median
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity
CMA	0.197***	0.231***	0.162***	0.223***
	(0.0465)	(0.0608)	(0.0473)	(0.0551)
Observations	4,917	3,479	4,338	4,121
R-squared	0.031	0.037	0.025	0.036
District-Industry FE	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
F-test	9064	2171	10798	2099

Robust standard errors in parentheses
(1) & (2) Share of Informal Firms
(3) & (4) Share of Informal Size
*** p<0.01, ** p<0.05, * p<0.1

6.3. Mechanisms

In this section, we explore the channels through which competition arising from better connectivity due to improvements in the road network impact labour productivity in the formal and informal sectors. The first channel is one of exit, whereby we expect less productive formal firms to exit potentially into the informal sector. We also expect less productive informal firms to exit the market. Furthermore, given that informal sector firms tend to be less productive than formal firms, we expect greater exit in the informal sector. In other words, broadly speaking, we expect an increase in competition to be associated with a formalization of firms. This means a negative relationship between CMA and the likelihood of being informal.

We explore this empirically following Nataraj (2011). We construct a binary variable that takes a value of one if the firm is in the informal sector and zero otherwise. We then estimate the relationship between CMA and this indicator of informality using the instrumental variable estimation strategy. Results are presented in Table 8. Column (1) includes no control variables, while columns (2) and (3) introduce some firm controls sequentially. Results show that the coefficient on CMA is negative and statistically significant, as anticipated. Thus, there is evidence that greater competition resulting from better connectivity to other intranational markets brought about by road improvements is associated with a lower likelihood of a firm operating in the informal sector. Differently from Nataraj (2011) and Goldberg and Pavcnik (2003), our coefficients of interest are significant, reinforcing the argument that the increase in connectivity is at the expense of the informal sector, strengthening the exit mechanisms on the left hand side of the firm productivity distribution.

Next, we explore the idea that the informal sector is "labour absorbing", or that it absorbs

Table 8: Effects of CMA on Informality

VARIABLES	(1) Informal	(2) Informal	(3) Informal
CMA	-0.00891	-0.00979*	-0.0118*
	(0.00828)	(0.00537)	(0.00655)
Observations	38,522	38,373	23,232
R-squared	0.000	0.199	0.249
District-Industry FE	Y	Y	Y
District-year FE	Y	Y	Y
Controls	N	Y	Y
F-test	3480	3467	3224

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

workers quitting formal firms that exit. If this channel were in operation, we would see an increase in the capital-labour ratio among formal firms and a decrease in the capital-labour ratio among informal firms[12]. In Table 9, we explore these outcome in an instrumental variables regression. Along with capital intensity, we report as well results on firms' investments and wages[13]. Columns (1)-(3), (4)-(6) and (7)-(9) report results for the whole sample, and for the sample of informal and formal firms, respectively.

As hypothesized, we find that for the formal sector, CMA is positively related to the capital-labour ratio and investment. This suggests that formal sector firms are increasingly capital-intensive and competition is associated with increased investment, in line with improvements in technology. These findings are corroborated by the positive effects on wages, as shown in column (9). We find no evidence for these positive effects in the informal sector. In fact, we find that greater competition from better connectivity is associated with a lower capital-labour ratio and lower investment. These results echo the idea that as labour is released from exiting formal firms due to competition, it is potentially absorbed in informal sector firms that become less capital-intensive.

We probe this channel further using detailed data on worker education levels in informal firms. We explore how changes in CMA affect the composition of the workforce by education. We can do this only for the sub-sample of informal firms covered by the SSIS, since this data set includes a module in which firms report worker characteristics such as level of education of individuals engaged in the firm. Table 10 reports estimates linking CMA to the total number

[12]Capital intensity is measured as the ration of the book value of fixed assets at the beginning of the year over total employment.

[13]Investment is measured by the size of expenditures in fixed assets during a fiscal year. Wages are per capita, calculated dividing the total wage bill of a firm by its total number of employees

of persons engaged (this includes both employees and working owners) and the share of persons engaged with (a) no education; (b) primary education; and (c) secondary education and above[14]. Results show that an increase in competition from higher CMA is associated with an increase in share of less educated workers (workers without primary education) in informal firms. This finding complements our earlier results showing an increase in investment and capital-intensity among formal firms in response to competition. Adoption of technology is likely to be intensive in high-skilled labour, potentially shifting worker composition towards more educated workers in the formal sector. We would expect to see a corresponding shift towards less educated workers among informal firms, as seen in column (2) of Table 10.

Put together, we can conceive of a story for the informal sector. Though the exit of less productive informal firms leads to greater productivity among surviving firms, workers who lose their jobs are absorbed into these firms without commensurate investments in capital. This means that they employ a lower capital-labour ratio and fewer educated workers associated with lower labour productivity.

Table 9: Capital Intensity and Investment

	WHOLE SAMPLE			INFORMAL			FORMAL		
VARIABLES	(1) Capital Intensity	(2) Investment	(3) Wage per capita	(4) Capital Intensity	(5) Investment	(6) Wage per capita	(7) Capital Intensity	(8) Investment	(9) Wage per capita
CMA	0.0503**	0.106**	0.0397***	-0.0494**	-0.0953*	0.00394	0.148***	0.294***	0.0463***
	(0.0205)	(0.0488)	(0.0111)	(0.0224)	(0.0526)	(0.0128)	(0.0361)	(0.0838)	(0.0172)
Obs.	33,569	38,373	25,072	22,492	26,406	14,831	9,840	10,858	9,071
R-sq.	0.116	0.008	0.084	0.075	0.005	0.012	0.086	0.011	0.068
Dist-Ind FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Dist-year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Controls	Y	Y	Y	Y	Y	Y	Y	Y	Y
F-test	3472	3467	2775	1913	1856	1220	2972	3286	3221

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

[14]Primary education corresponds to grades 1 to 8 in the Ethiopian system.

Table 10: Effects of CMA on Informality

VARIABLES	(1) Engaged Persons	(2) Share of workers without EDU	(3) Share of workers with primary EDU	(4) Share of workers with secondary and above
CMA	0.0106	0.00901***	-0.00234	-0.00967*
	(0.0278)	(0.00269)	(0.00511)	(0.00557)
Observations	25,818	25,818	25,818	25,818
R-squared	0.006	0.001	0.000	0.000
District-Industry FE	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
F-test	1696	1696	1696	1696

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7. Robustness checks

In this section, we test robustness of our baseline findings to a different definition of informality, varying trade elasticity numbers and an alternate instrumental variable. As we show in the following sections, our results remain qualitatively robust.

7.1. Alternate Definition of Informality

In order to assess the reliability of our results, we perform the previous analysis to check if results are sensitive to different definitions of informality. In this exercise, we consider the following two alternative definitions: (1) firms that do not keep books of accounts, meaning those not registered, independently of their size; and (2) firms split according to the specific survey, i.e. in the SSIS and the LMMS. The results are presented in Table 11 and appear to be consistent and in line with the baseline results in Table 4.

Table 11: Alternate definition of informality: Labor Productivity

	Book of Account		Survey	
	Formal	Informal	Formal	Informal
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity
CMA	0.131***	-0.00752	0.184***	-0.00423
	(0.0346)	(0.0234)	(0.0361)	(0.0249)
Observations	7,788	17,781	14,513	5,186
R-squared	0.011	0.003	0.003	0.013
District-Industry FE	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y
Controls	Y	Y	Y	Y
F-test	2509	3182	2980	2222

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7.2. Testing Different Trade Elasticities

In this Section, we test the sensitivity of our results to different values of θ, the distance decay parameter used to define CMA. θ captures the non-linear impact of distance on trade, and through its value is likely to be context specific, it is normally estimated in a range going from 1 to 10. Hence, we replicate our results using values of θ equal to 1, 3.124 and 8.22. A value of 1 corresponds to the canonical definition of market potential, as provided originally by Harris, 1954. We got to a value of 3.124 if we replicate our baseline scenario using the trade-travel time differential between Ethiopia and the US, that is estimated to be approximately 2.46 times by Atkin and Donaldson (2015). Last, 8.22 is the elasticity estimated in (Donaldson and Hornbeck, 2016). From Table 12, we find that independently of the value of θ, results appear consistent and with the baseline, suggesting that infrastructural investments are associated with higher productivity in the formal sector and not statistically significantly related to productivity in the informal sector.

Table 12: Robustness to Varying Trade Elasticity

	WHOLE SAMPLE		INFORMAL		FORMAL	
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity	(5) Labor Productivity	(6) Labor Productivity
			θ =1			
CMA	0.129***	0.120***	-0.00890	-0.0130	0.218***	0.217***
	(0.0426)	(0.0349)	(0.0316)	(0.0319)	(0.0436)	(0.0399)
R-squared	0.003	0.045	0.000	0.006	0.010	0.037
F-test	173390	173640	247158	248600	95781	95688
			θ =3.12			
CMA	0.0833***	0.0791***	-0.00939	-0.0132	0.165***	0.164***
	(0.0314)	(0.0254)	(0.0276)	(0.0279)	(0.0337)	(0.0304)
R-squared	0.002	0.044	0.000	0.006	0.008	0.035
F-test	4663	4632	4928	4933	3774	3689
			θ =8.22			
CMA	0.00372	0.00659	-0.0103	-0.0119	0.0297**	0.0333**
	(0.0106)	(0.00946)	(0.0117)	(0.0118)	(0.0135)	(0.0129)
R-squared	0.000	0.042	0.000	0.006	0.001	0.028
F-test	1256	1243	880.1	880.9	884.6	868.7
Observations	23,350	23,232	13,772	13,762	8,668	8,563
District-Industry FE	Y	Y	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

7.3. Using Different Instruments

Finally, Table 13, we present results using an alternate instrument. Following the insights of Moneke (2019) and Gebresilasse (2018), we construct an alternate instrumental variable. First, we create an alternative road network based on the Euclidean distance to the pre-existing Italian colonial road network, that have been digitised from Gli Annali dell'Africa Italiana (1937-1943). Next, we artificially distribute the entire length of roads under the RSDP to districts, with the constraint of connecting them to the road network by the end of the sample period using an artificial regional budgeting algorithm. We thus generate a time-varying instrument which derives its exogenous variation from the straight line distance to Italy's digitized colonial road map. From Table 12 we find that our results remain qualitatively similar. In fact, we now observe that the relationship between CMA and labour productivity in the informal sector is negative. This indicates that the channel whereby labour absorption in the informal sector depresses labour productivity dominates the channel whereby the more productive firms survive, in line with our findings in the section exploring mechanisms.

Table 13: Different IV: Labor Productivity

	WHOLE SAMPLE		INFORMAL		FORMAL	
VARIABLES	(1) Labor Productivity	(2) Labor Productivity	(3) Labor Productivity	(4) Labor Productivity	(5) Labor Productivity	(6) Labor Productivity
CMA	0.0135	0.0264	-0.0852*	-0.0904*	0.105**	0.130***
	(0.0364)	(0.0316)	(0.0484)	(0.0493)	(0.0466)	(0.0425)
Observations	23,350	23,232	13,772	13,762	8,668	8,563
R-squared	0.000	0.043	-0.002	0.004	0.006	0.033
District-Industry FE	Y	Y	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y
F-test	136.1	135.6	51.91	51.42	93.66	94.12

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

8. Conclusions

This study connects two areas of high priority for development policies: the informal sector, which is a pervasive feature of the developing world and plays a central role in countries' structural transformation; and infrastructure investments, which have been considered as crucial drivers of economic growth. Understanding the responses of the formal and informal sectors to large infrastructure projects is critical for policy. Although governmental provision of public goods should be seen as productivity-enhancing, it is plausible that these benefits vary significantly across formal and informal firms. Ignoring such heterogeneity in the response of firms to road infrastructure improvements can create distortions in resource allocation.

We find that increased competition resulting from a decrease in intranational trade cost as connectivity improves from road improvements is associated with higher productivity and wages in the formal sector. This is contrary to what we find for the informal sector, where labour productivity, if anything, decreases with greater competition. We also find opposite effects of an increase in competition on capital-labour ratios and investments, with positive (negative) effects for the formal (informal) sector. Greater competition is also associated with a lower likelihood of a firm operating informally and of informal firms employing a greater share of workers with less than primary education. We thus highlight that the benefits of road infrastructure improvements accrue unevenly across formal and informal sectors in developing countries.

Our paper can be seen as a first attempt to shed light on the differential impacts of improved infrastructure on the performance of the formal and informal sectors in developing countries. We call for further data on firms in the informal sector, especially of a panel nature, to further probe the dynamics of firm adjustments to large infrastructure programmes.

References

Adamopoulos, Tasso (2019). *Spatial Integration, Agricultural Productivity, and Development: A Quantitative Analysis of Ethiopia's Road Expansion Program*. 2019 Meeting Papers 86. Society for Economic Dynamics. URL: https://ideas.repec.org/p/red/sed019/86.html.

Alden, Chris (2019). *China's Evolving Military Strategy in Africa*. URL: https://chinaafricaproject.com/podcast-military-china-africa-security-strategy-chris-alden/ (visited on 09/12/2019).

Alder, Simon (2015). *Chinese Roads in India: The Effect of Transport Infrastructure on Economic Development*. 2015 Meeting Papers 1447. Society for Economic Dynamics. URL: https://ideas.repec.org/p/red/sed015/1447.html.

Asher, Sam and Novosad, Paul (2020). "Rural Roads and Local Economic Development". In: *American Economic Review* 110.3, pp. 797–823. DOI: 10.1257/aer.20180268. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20180268.

Atkin, D. and Donaldson, D. (2015). "Who's Getting Globalized? The Size and Implications of Intra-National Trade Costs". In: *Kilts Center at Chicago Booth Marketing Data Center Paper Series*.

Baker, J. (1974). "Developments in Ethiopia's Road System". In: *Geography* 59.2, pp. 150–154. ISSN: 00167487. URL: http://www.jstor.org/stable/40568212.

Banerjee, Abhijit, Duflo, Esther, and Qian, Nancy (2020). "On the road: Access to transportation infrastructure and economic growth in China". In: *Journal of Development Economics* 145, p. 102442. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2020.102442. URL: https://www.sciencedirect.com/science/article/pii/S0304387820300171.

Berg, Claudia N. *et al.* (2017). "Transport Policies and Development". In: *Journal of Development Studies* 53.4, pp. 465–480. DOI: 10.1080/00220388.2016.119. URL: https://ideas.repec.org/a/taf/jdevst/v53y2017i4p465-480.html.

Bernard, Andrew B. *et al.* (2012). "The Empirics of Firm Heterogeneity and International Trade". In: *Annual Review of Economics* 4.1, pp. 283–313. DOI: 10.1146/annurev-economics-080511-110928. eprint: https://doi.org/10.1146/annurev-economics-080511-110928. URL: https://doi.org/10.1146/annurev-economics-080511-110928.

Bertazzini, Mattia (2022). "The long-term impact of Italian colonial roads in the Horn of Africa, 1935-2000". In: *Journal of Economic Geography, 22 (2022): 181-214* 22.1, pp. 181–214. DOI: https://doi.org/10.1093/jeg/lbaa017. eprint: https://doi.org/10.1093/jeg/lbaa017. URL: https://doi.org/10.1093/jeg/lbaa017.

Burgess, Robin *et al.* (2015). "The Value of Democracy: Evidence from Road Building in Kenya". In: *American Economic Review* 105.6, pp. 1817–51. DOI: 10.1257/aer.20131031.

Chandra, Amitabh and Thompson, Eric (2000). "Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system". In: *Regional Science and Urban Economics* 30.4, pp. 457–490. URL: https://EconPapers.repec.org/RePEc:eee:regeco:v:30:y:2000:i:4:p:457-490.

Chatterjee, Santanu, Lebesmuehlbacher, Thomas, and Narayanan, Abhinav (2021). "How productive is public investment? Evidence from formal and informal production in India". In: *Journal of Development Economics* 151, p. 102625. ISSN: 0304-3878. DOI: https://

doi.org/10.1016/j.jdeveco.2021.102625. URL: https://www.sciencedirect.com/science/article/pii/S030438782100002X.

De Loecker, Jan and Goldberg, Pinelopi Koujianou (2013). *Firm Performance in a Global Market*. Working Paper 19308. National Bureau of Economic Research. DOI: 10.3386/w19308. URL: http://www.nber.org/papers/w19308.

Diao, Xinshen *et al.* (2021). *Africa's Manufacturing Puzzle: Evidence from Tanzanian and Ethiopian Firms*. Working Paper 28344. National Bureau of Economic Research. DOI: 10.3386/w28344. URL: http://www.nber.org/papers/w28344.

Dix-Carneiro, Rafael *et al.* (2021). *Trade and Informality in the Presence of Labor Market Frictions and Regulations*. NBER Working Papers 28391. National Bureau of Economic Research, Inc. URL: https://ideas.repec.org/p/nbr/nberwo/28391.html.

Donaldson, Dave (2018). "Railroads of the Raj: Estimating the Impact of Transportation infrastructure". In: *American Economic Review* 108.4-5, pp. 899–934. DOI: 10.1257/aer.20101199. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20101199.

Donaldson, Dave and Hornbeck, Richard (2016). "Railroads and American Economic Growth: A 'Market Access' Approach". In: *The Quarterly Journal of Economics* 131.2, pp. 799–858.

Duflo, Esther and Pande, Rohini (2007). "Dams". In: *The Quarterly Journal of Economics* 122.2, pp. 601–646. URL: https://EconPapers.repec.org/RePEc:oup:qjecon:v:122:y:2007:i:2:p:601-646..

Duranton, Gilles, Morrow, Peter M., and Turner, Matthew A. (2014). "Roads and Trade: Evidence from the US". In: *The Review of Economic Studies* 81.2 (287), pp. 681–724. ISSN: 00346527, 1467937X. URL: http://www.jstor.org/stable/43551577.

Duranton, Gilles and Turner, Matthew A. (2012). "Urban Growth and Transportation". In: *The Review of Economic Studies* 79.4, pp. 1407–1440. ISSN: 00346527, 1467937X. URL: http://www.jstor.org/stable/23355076.

ERA (2009). *SDP Performance: 12 Years Later*. Tech. rep. Ethiopian Road Authority (ERA).

Esfahani, Hadi Salehi and Ramire, Maria Teresaz (2003). "Institutions, infrastructure, and economic growth". In: *Journal of Development Economics* 70.2, pp. 443–477. ISSN: 0304-3878. DOI: https://doi.org/10.1016/S0304-3878(02)00105-0. URL: https://www.sciencedirect.com/science/article/pii/S0304387802001050.

Ethiopian Road Authority (2016). *Road Sector Development Program 19 years performance assessment*. Tech. rep. Ministry of Transport, Ethiopian Road Authority (ERA). URL: http://www.era.gov.et/documents/10157/37f7faf1-4614-4abb-8be4-2146ffadfe27.

Faber, Benjamin (2014). "Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System". In: *The Review of Economic Studies* 81.3 (288), pp. 1046–1070. ISSN: 00346527, 1467937X. URL: http://www.jstor.org/stable/43551619.

Fiorini, Matteo and Sanfilippo, Marco (2019). "Roads and jobs in Ethiopia". In: DOI: https://doi.org/10.35188/UNU-WIDER/2019/752-1.

Fiorini, Matteo, Sanfilippo, Marco, and Sundaram, Asha (2021). "Trade liberalization, roads and firm productivity". In: *Journal of Development Economics* 153, p. 102712. ISSN: 0304-3878. DOI: https://doi.org/10.1016/j.jdeveco.2021.102712. URL: https://www.sciencedirect.com/science/article/pii/S0304387821000869.

Franklin, Simon (2018). "Location, Search Costs and Youth Unemployment: Experimental Evidence from Transport Subsidies". In: *Economic Journal* 128.614, pp. 2353–2379. DOI: 10.1111/ecoj.12509. URL: https://ideas.repec.org/a/wly/econjl/v128y2018i614p2353-2379.html.

Gebresilasse, Mesay M. (2018). "Rural Roads, Agricultural Extension, and Productivity". In: *Boston University* Job Market Paper.

Gli Annali dell'Africa Italiana (1937-1943). *Piccioli, Angelo*. Tech. rep. Mondadori.

Goldberg, Pinelopi and Pavcnik, Nina (2003). "The response of the informal sector to trade liberalization". In: *Journal of Development Economics* 72.2, pp. 463–496. URL: https://EconPapers.repec.org/RePEc:eee:deveco:v:72:y:2003:i:2:p:463-496.

Harris, Chauncy D. H (1954). "The Market as a Factor in the Localization of Industry in the United States". In: *Annals of the Association of American Geographers* 44.4, pp. 315–348. ISSN: 00045608, 14678306. URL: http://www.jstor.org/stable/2561395.

Huang, Yuxiao and Xiong, Wentao (2018). "Geographic Distribution of Firm Productivity and Production: A Market Access Approach". In: *Harvard PhD dissertation*.

Jedwab, Remi and Storeygard, Adam (2020). *The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010*. Working Paper 27670. National Bureau of Economic Research. DOI: 10.3386/w27670. URL: http://www.nber.org/papers/w27670.

Kruse, Hagen *et al.* (2021). "A manufacturing renaissance? Industrialization trends in the developing world". In: *UNU-WIDER* Working Papers. DOI: https://doi.org/10.35188/UNU-WIDER/2021/966-2.

La Porta, Rafael and Shleifer, Andrei (2014). "Informality and Development". In: *Journal of Economic Perspectives* 28.3, pp. 109–26. DOI: 10.1257/jep.28.3.109. URL: https://www.aeaweb.org/articles?id=10.1257/jep.28.3.109.

McCaig, Brian and Pavcnik, Nina (2018). "Export Markets and Labor Allocation in a Low-Income Country". In: *American Economic Review* 108.7, pp. 1899–1941. DOI: 10.1257/aer.20141096. URL: https://www.aeaweb.org/articles?id=10.1257/aer.20141096.

McMillan Margaret S.and McCaig, Brian (2019). *Trade Liberalization and Labor Market Adjustment in Botswana*. NBER Working Papers 26326. National Bureau of Economic Research, Inc. URL: https://ideas.repec.org/p/nbr/nberwo/26326.html.

Melitz, Marc (Nov. 2003). "The Impact of Trade on Intra-Industry Re-Allocation and Aggregate Industrial Productivity". In: *Econometrica* 71, pp. 1695–1725. DOI: 10.1111/1468-0262.00467.

Ministry of Finance and Economic Development (2010). *Growth and Transformation Plan 2010/11 -2014/15*. Tech. rep. Federal Democratic Republic of Ethiopia, Ministry or Finance and Economic Development. URL: http://extwprlegs1.fao.org/docs/pdf/eth144893.pdf.

MoFED (2006). *Ethiopia: Building on Progress, A Plan for Accelerated and Sustained Development to End Poverty (PASDEP)*. 2005/06-2009/10, Volume I, Addis Ababa. Ministry of Finance and Economic Development.

Moneke, Niclas (2019). "Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia". In: *London School of Economics (LSE)* Job Market Paper.

Nataraj, Shanthi (2011). "The impact of trade liberalization on productivity: Evidence from India's formal and informal manufacturing sectors". In: *Journal of International Economics* 85.2, pp. 292–301. URL: https://EconPapers.repec.org/RePEc:eee:inecon:v:85:y:2011:i:2:p:292-301.

Perra, Elena (2022). *Road to Division: Ethnic Favouritism in the Provision of Road Infrastructure in Ethiopia*. Working Papers - Economics. Universita' degli Studi di Firenze, Dipartimento di Scienze per l'Economia e l'Impresa.

Puga, Diego (2009). *The magnitude and causes of agglomeration economies*. Working Papers 2009-09. Instituto Madrileño de Estudios Avanzados (IMDEA) Ciencias Sociales. URL: https://ideas.repec.org/p/imd/wpaper/wp2009-09.html.

Redding, Stephen J. (2020). *Trade and Geography*. Working Paper 27821. National Bureau of Economic Research. DOI: 10.3386/w27821. URL: http://www.nber.org/papers/w27821.

Redding, Stephen and Turner, Matthew (2015). "Transportation Costs and the Spatial Organization of Economic Activity". In: vol. 5. Elsevier. Chap. Chapter 20, pp. 1339–1398. URL: https://EconPapers.repec.org/RePEc:eee:regchp:5-1339.

Shiferaw, Admasu *et al.* (2015). "Road Infrastructure and Enterprise Dynamics in Ethiopia". In: *The Journal of Development Studies* 51.11, pp. 1541–1558. DOI: 10.1080/00220388.2015.1056785. eprint: https://doi.org/10.1080/00220388.2015.1056785. URL: https://doi.org/10.1080/00220388.2015.1056785.

Siba, Eyerusalem (2015). "Returns to Physical Capital in Ethiopia: Comparative Analysis of Formal and Informal Firms". In: *World Development* 68.C, pp. 215–229. URL: https://EconPapers.repec.org/RePEc:eee:wdevel:v:68:y:2015:i:c:p:215-229.

Storeygard, Adam (2016). "Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa". In: *Review of Economic Studies* 83.3, pp. 1263–1295. URL: https://ideas.repec.org/a/oup/restud/v83y2016i3p1263-1295..html.

Topalova, Petia and Khandelwal, Amit (Aug. 2011). "Trade Liberalization and Firm Productivity: The Case of India". In: *The Review of Economics and Statistics* 93.3, pp. 995–1009. ISSN: 0034-6535. DOI: 10.1162/REST_a_00095. eprint: https://direct.mit.edu/rest/article-pdf/93/3/995/1919022/rest_a_00095.pdf. URL: https://doi.org/10.1162/REST_a_00095.

United States Central Intelligence Agency (CIA) (1969). *Map of Ethiopia*. Retrieved from the Library of Congress. Library of Congress Geography and Map Division Washington, D.C. 20540-4650 USA dcu. URL: https://www.loc.gov/item/97687601/.

Worku, Ibrahim (2011). "Road Sector Development and Economic Growth in Ethiopia". Addis Ababa, Ethiopia: Ethiopian Development Research Insitute.

World Bank (2021). *Ethiopia-Road Sector Development Program APL4*. URL: https://projects.worldbank.org/en/projects-operations/project-detail/P106872.

A. Summary Statistics

Here I report some other tables.

Table A.1: Data from the Small Scale Industries Survey (SSIS) and Census of large and medium manufacturing firms (CSA)

Year	N firms surveyed (SSIS)	N firms surveyed (Census)
2001	8,054	759
2004	4,299	990
2007	11,314	1327
2010	3,882	1880
2013	11,307	2391
Total	38,856	7346

Table A.2: Small Business Sector in Ethiopia

Year	% Total Firms	% Total Emp.	% Total Prod.	% Total Wages	% Total Cap.
2001	97.67%	50.96%	2.54%	10.49%	11.18%
2004	97.53%	55.41%	9.63%	14.99%	11.76%
2007	97.02%	52.09%	9.29%	16.47%	12.05%
2010	96.47%	52.18%	9.55%	17.04%	11.41%
2013	97.99%	62.95%	16.56%	19.00%	10.68%

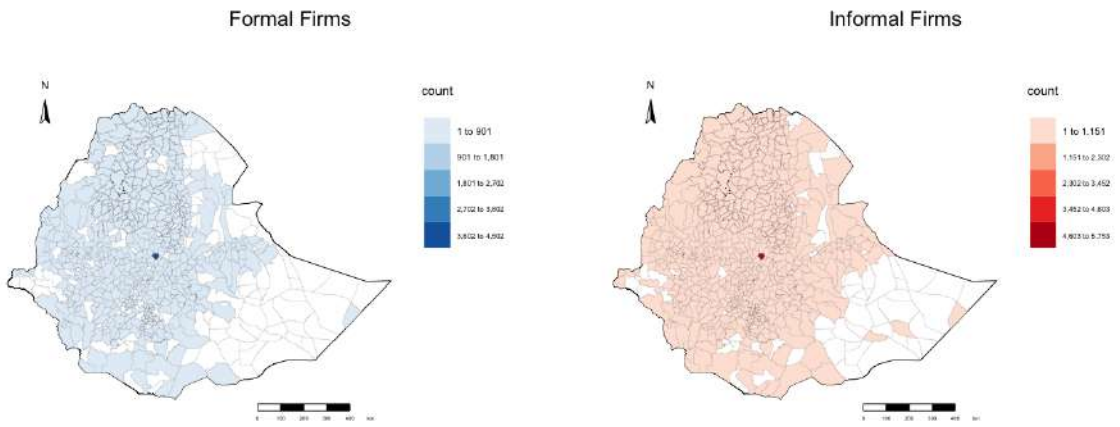

Figure A.1: Geographical Distribution of Formal and Informal firms on the Ethiopian territory. Authors' calculations.

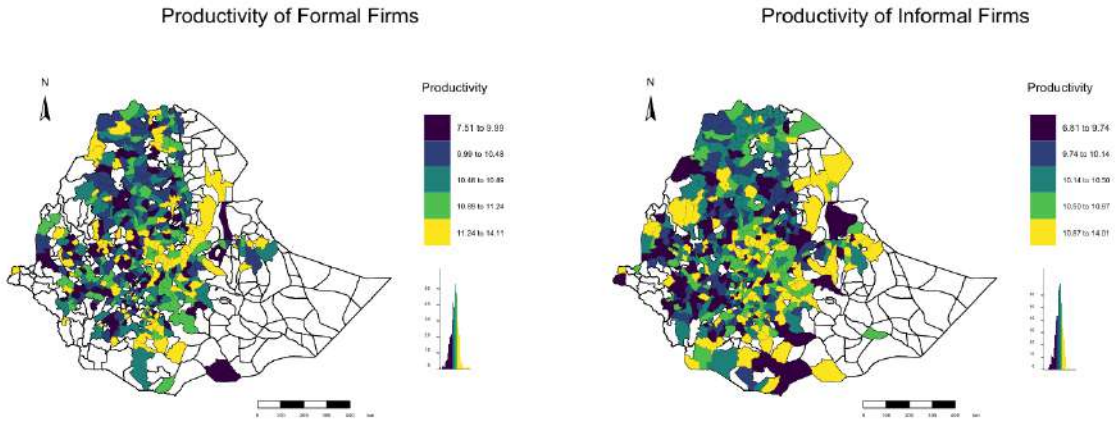

Figure A.2: Visualization of the Productivity of Formal and Informal firms. Authors' calculations.

Table A.3: CMA of Formal and Informal Firms

	Informal					Formal				
year	mean	sd	min	max	n	Mean	Sd	Min	Max	N
2001	5.40	4.32	-6.90	16.71	5564	7.14	4.86	-3.45	16.71	1666
2004	6.16	4.80	-6.74	16.44	3337	8.04	4.85	-5.83	16.45	1436
2007	7.85	3.73	-7.17	17.08	8034	9.05	3.82	-8.01	16.52	2997
2010	8.64	3.35	-1.84	18.14	2901	10.19	3.80	-2.99	17.94	1920
2013	9.73	3.11	-5.33	18.78	7659	10.40	3.55	-4.24	18.78	4325

Figure A.3: Ethiopian Road Network 1969

Table A.4: First Stages

	WHOLE SAMPLE	INFORMAL	FORMAL
	(1)	(2)	(3)
	CMA	CMA	CMA
CMA	0.964***	0.980***	0.956***
	(0.0169)	(0.0174)	(0.0190)
N	23350	13772	8668

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A.5: Reduced Form Specification: Labor Productivity

	WHOLE SAMPLE		INFORMAL		FORMAL	
	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	Labor Productivity	Labor Productivity	Labor Productivity	Labor Productivity	Labor Productivity	Labor Productivity
CMA_{IV}	0.0599**	0.0579***	-0.0112	-0.0147	0.131***	0.131***
	(0.0258)	(0.0210)	(0.0240)	(0.0242)	(0.0295)	(0.0264)
Constant	10.53***	11.15***	10.80***	11.28***	10.36***	10.60***
	(0.234)	(0.247)	(0.207)	(0.227)	(0.295)	(0.278)
Observations	23,350	23,232	13,772	13,762	8,668	8,563
R-squared	0.405	0.430	0.457	0.460	0.344	0.362
District-Industry FE	Y	Y	Y	Y	Y	Y
District-year FE	Y	Y	Y	Y	Y	Y
Controls	N	Y	N	Y	N	Y

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1