Connectivity and Local Opportunity

Luke Heath Milsom*

30th March, 2026

Abstract

This paper studies the distributional impact of road construction on educational opportunity in spatial equilibrium. I show that roads have two effects: they enable moves to opportunity and move opportunity itself. To capture both channels, I embed a quantitative spatial model with education choice into a place-effects framework. I estimate the model using digitised historical road maps in Benin, Cameroon, and Mali, combining a movers design with a novel instrument that isolates indirect variation in connectivity. I find that road building since 1970 increased the causal effect of place on primary school completion by 12.5% on average. The first channel, moving to opportunity, explains 84% of the average effect. However, the second channel, moving opportunity, is sufficiently large to leave 15% of locations worse off in aggregate.

JEL: R12, R42, O18, H54
Keywords: Inequality, Opportunity, Roads, Spatial, Market Access

*KU Leuven and the IFS, luke.heathmilsom@kuleuven.be. I thank Johannes Abeler, Abigail Adams, Sam Altmann, Sarah Clifford, Ian Crawford, Binta Zahra Diop, Dave Donaldson, Christian Dustmann, Thibault Fally, Cecile Gaubert, Douglas Gollin, Marco Gonzalez-Navarro, Nina Guyon, Vernon Henderson, Shihang Hou, Allan Hsiao, Vatsal Khandelwal, Hamish Low, Niclas Moneke, Sanghamitra Mukherjee, Jessica Pan, Barbara Petrongolo, Simon Quinn, Itzhak Rasooly, Ferdinand Rauch, Isabelle Roland, Elizabeth Sadoulet, Paul Schaudt, Giulio Schinaia, Evan Soltas, Adam Storeygard, Gabriel Ulyssea, Tony Venables, Verena Wiedemann, Hannah Zillessen, and participants at various seminars, conferences, and workshops, for useful comments and conversations.

1 Introduction

Where you grow up is an important determinant of the education you complete. Indeed, local returns to education depend on access to jobs that reward schooling. I focus on the low- and low-middle-income country setting of Benin, Cameroon, and Mali — where spatial differences in returns are starkest (Bryan et al., 2020; Gollin et al., 2014; Gollin et al., 2021; Lagakos, 2020) but we know the least about how to affect them. In this paper, I argue that road building shapes the spatial distribution of educational opportunity by changing access to, and the geography of, such jobs. Reducing spatial frictions directly enables moves to opportunity by lowering travel costs. However, as households and firms endogenously react in spatial equilibrium, the underlying geography of high-return locations itself shifts. Through this second channel, opportunity could move away from a given location: Reducing spatial frictions need not be Pareto-improving.

To understand how road building shapes the geographic distribution of opportunity, I first build a quantitative spatial economics model embedded within the canonical place-effects framework. This model delivers a reduced-form relationship between education decisions and market access that I can then take to the data using historical road maps I digitised. To estimate this relationship, I employ a dual identification strategy using plausible exogenous variation in market access to overcome the endogeneity of road placement and a movers design to separately identify place-effects from sorting. Finally, using the reduced form results to recover model parameters, I then leverage the full structure of the model to estimate the aggregate effects of road construction and quantify mechanisms. By shutting down endogenous responses, I can decompose each location-specific aggregate effect of road building into that due to greater moves to opportunity and that due to the underlying geography of opportunity moving.

I estimate the reduced-form impact of changing connectivity on local educational opportunity, using data from historical road maps that I digitise, and national censuses in Benin, Cameroon, and Mali. By studying all three countries in Sub-Saharan Africa with sufficient data, I can show external validity across road networks and test mechanisms in different settings. Local educational opportunity is defined following the place-effects literature as the causal effect of growing up in a given location on the probability of completing primary school (Alesina et al., 2021; Chetty and Hendren, 2018b; Milsom, 2023).[1] I then estimate the theory-derived log-linear relationship between primary school completion and

[1] I focus on primary school completion for three main reasons. First, measures of income commonly used in high-income countries are less appropriate in this setting, where the majority of workers are not waged, and subsistence agriculture is common. Second, primary school completion is the most salient margin of education in my setting, with only 7% of my overall sample going on to complete secondary school. Third, primary school completion is widely available in data sources with sufficient geographic granularity and size to be amenable to my analysis.

local market access by combining a movers design (Chetty and Hendren, 2018a; Chetty and Hendren, 2018b; Chiovelli et al., 2025) with plausibly exogenous variation in market access using a novel instrument. This instrument isolates the indirect component of changes in market access due to the impact of roads rippling through the network. Simulation and graphical evidence show this strategy performs well when roads are placed endogenously. I find that growing up in a one standard deviation higher market access location increases the probability of completing primary school by 7 percentage points (12%) on average.[2]

To understand the aggregate effect of road building, I leverage the full structure of the model. The model builds on Hsiao (2024), Allen et al. (2020), and Chetty and Hendren (2018a) by endogenising place-effects in spatial equilibrium, allowing firms and workers, as well as wages and prices, to adjust in response to a change in trade and migration costs. Solving the model in changes and recovering parameters from the reduced-form estimates allows me to perform counterfactual analysis. I find that road building since 1970, on average, increased local opportunity by 12.5%. However, this average hides considerable heterogeneity: 15% of locations saw negative aggregate effects, and a fifth of locations saw increases in excess of 20%. These exercises show that road building is an effective policy lever in moving the spatial distribution of opportunity in Benin, Cameroon, and Mali, but also one where location-specific heterogeneity is important, and effects need not be Pareto improving.

I then decompose the location-specific effect of roads into two channels: the direct effect of enabling moves to opportunity and the spatial equilibrium force of opportunity moving. The direct effect of improving roads explains 84% of the average impact and is positive everywhere because lower travel costs make a wider range of locations worth moving to. Without endogenous equilibrium responses, reducing spatial frictions is always Pareto improving. The spatial-equilibrium force of opportunity moving is, on average, also positive, as fewer frictions allow economic activity to allocate more efficiently across space. However, this channel also generates losers as activity concentrates in some locations, and others are left behind. Further decomposing the spatial-equilibrium channel, I find that labour market adjustment alone is negative. In-migration to high-wage locations depresses wages, flattening the distribution. However, goods market adjustment offsets this, as higher-population locations also attract higher demand. Although these forces are somewhat balanced on average, they bite to differing degrees at the individual location level. The 10th percentile of the isolated endogenous spatial equilibrium response is -5.8pp, and the 90th percentile is +8.8pp. Many locations with little direct exposure to road building experience large effects through

[2] I develop this instrumental variables strategy because neither placebo roads (Donaldson, 2018), nor plausible incidental connections (Banerjee et al., 2020) are available in my setting. The reduced form result is robust to reasonable variations in specification, only using within-family variation in exposure to locations, and controlling for endogenous exposure to exogenous shocks (Borusyak and Hull, 2023).

endogenous responses alone. For a sizable minority, the moving distribution of opportunity, rather than moving to opportunity, is the dominant force.

The main contribution of this paper is to demonstrate that road building alters the geographic distribution of educational opportunity and that spatial equilibrium responses can leave some locations worse off. In showing this I contribute to the literature studying the geography of opportunity by treating place effects as equilibrium objects rather than parameters to be estimated (Chetty and Hendren, 2018a; Chetty and Hendren, 2018b; Deutscher, 2020; Laliberté, 2021; Van Maarseveen, 2025; Alesina et al., 2021; Rojas Ampuero, 2022; Chetty et al., 2016; Chiovelli et al., 2025). Relative to Chyn and Daruich (2025), and Eckert and Kleineberg (2024) who consider place effects in a general equilibrium setting, I study the central policy lever of road building in a low and low-middle income country setting where spatial misallocation is most severe (Bryan et al., 2020; Gollin et al., 2014; Lagakos, 2020). I also contribute to the literature on education and roads in low- and middle-income countries, which has focused on reduced-form effects of connections on education (Adukia et al., 2020; Asher and Novosad, 2020; Mukherjee, 2012), and work on the general-equilibrium effects of education investments, which has not considered transport infrastructure (Hsiao, 2024; Khanna, 2023; Fujimoto et al., 2025; Desmet et al., 2025). I combine these approaches, and in doing so can show not only that roads affect the impact of place on education, but also how spatial equilibrium responses mediate that effect.

Methodologically, I contribute in two ways. First, I build a quantitative spatial economics model with education choice embedded within a place-effects framework (Redding and Rossi-Hansberg, 2017; Allen et al., 2020) that delivers a market-access sufficient statistic (Donaldson and Hornbeck, 2016), and allows costly migration and trade (Morten and Oliveira, 2024; Bryan and Morten, 2019; Adamopoulos, 2025; Tsivanidis, 2026; Sotelo, 2020). Relative to Hsiao (2024), who models education choice in a spatial setting focusing on the supply effects of new schools, I endogenise the geographic distribution of returns to education in spatial equilibrium. Second, to identify reduced-form impacts and recover model parameters, I develop a novel identification strategy that isolates indirect changes in market access. This contributes to the literature on identifying the causal effects of transport infrastructure by developing an approach that performs well in simulations, builds upon the far-away variation approach (Jedwab and Storeygard, 2021; Donaldson and Hornbeck, 2016), and can be used when planned & historical routes or incidental connections are not available (Donaldson, 2018; Baum-Snow, 2007; Baum-Snow et al., 2017; Duranton and Turner, 2012; Banerjee et al., 2020; Faber, 2014; Ghani et al., 2016; Michaels, 2008; Fenske et al., 2023).

The rest of this paper proceeds as follows: section 2 describes the setting and data, section 3 elucidates the model, section 4 explains the identification approach and presents the reduced form regression results, section 5 quantifies the impact of road building since

1970 and considers mechanisms, and finally section 6 concludes.

2 Setting & data: Benin, Cameroon, & Mali 1970-2020

Benin, Cameroon, and Mali form an excellent setting to study how connectivity shapes spatial inequality, as well as each country being an important setting in which to study these questions in and of themselves. This is for three main reasons. First, each displays substantial historical variation in local connectivity due to road building, providing sufficient variation to estimate causal effects. However, each country also has significantly different road networks and trajectories — allowing me to test the mechanisms in different settings. Second, low baseline paved road coverage and high projected urbanisation and population growth make road infrastructure investment an imminent policy priority.

Third, Benin, Cameroon, and Mali also have the advantage of being the only countries in sub-Saharan Africa with data sufficiently rich to identify the impact of changes in connectivity on the causal effect of place on primary education completion. To estimate causal place effects using a movers design with cross-sectional data, it is necessary to observe where an individual is when the census was taken, their previous location, their birth location, and how long they have resided in their current location. To uncover causal place effects, previous research (Chetty and Hendren, 2018a; Laliberté, 2021; Deutscher, 2020) has mainly relied on administrative data that is not typically publicly available. This data is necessary for these studies to observe full migration histories and to match child and parent outcomes over long time frames. Such rich data is not needed to estimate place effects in the censuses I use; it is possible to discern migration histories from cross-sectional evidence. In addition to migration information, I require data on primary education completion,[3] and age, at a granular and consistent geographic level over multiple cross-sectional waves. Only three countries in Sub-Saharan Africa — Benin, Cameroon, and Mali fit the stringent data requirements.

2.1 Historical context

The histories of road development in the three countries share a common arc, with informative differences. In all three cases, modern road corridors trace their origins to colonial-era infrastructure built for military control and resource extraction — the French in Benin and Mali, the Germans and subsequently the British and French in Cameroon (Tchindjang et al.,

[3] In Benin, Cameroon, and Mali, education is compulsory for the first six years of schooling between the ages of 6 and 11/12, which covers primary school. Benin has a 6-4-3 education system; those who state they attend Islamic school are dropped from the sample. See Appendix L for details. Cameroon has a dual education system with a French-speaking 6-4-3 structure and a minority English-speaking 7-5-2 structure in some parts of the country; this is mapped onto the 6-4-3 system in my data. Any such location-specific differences will be taken into account in my analysis. Mali follows a 6-3-3 system during the period I study.

2005). Although in Benin, road infrastructure has antecedents in precolonial state formation. For example, the Kingdom of Dahomey maintained a formalised "Royal Road" that linked the Atlantic port of Whydah (Ouidah) to the royal capital at Abomey (Alpern, 1999). After independence in 1960, Benin and Mali entered a period of political instability that constrained long-horizon transport investment: military coups reshaped government in Benin (1972) and Mali (1968), while Cameroon's institutional environment was disrupted by the shift from a federal to a unitary state in 1972. In all three, World Bank financing became the dominant external driver of road construction from the 1970s onward (World Bank, 1982; World Bank, 1970; International Development Association, 1981), and each country eventually established a dedicated Road Fund — Cameroon by the late 1990s, Mali in July 2002, and Benin under parallel mechanisms — reflecting a shared institutional response to chronic deferred maintenance problems. A common inflection point is the shift from network expansion toward maintenance of existing stock, occurring at different moments between the mid-1980s and early 2000s in each country, typically following macroeconomic stress and structural adjustment.

Despite this shared arc, there are significant differences between each country. First, the supranational institutional environment driving road investment differs sharply between Cameroon and the other two countries. Benin and Mali are both members of the West African Economic and Monetary Union (UEMOA), which in 2001 adopted a Community Action Program for Infrastructure and Road Transport (PACITR), defining priority interstate corridors for upgrading across member states. Additionally, Benin's position as a coastal transit country adds a further layer to its road-building history that is absent from Mali and Cameroon. On the other hand, Cameroon, as a Central African Economic and Monetary Community (CEMAC) member, faced a different regional corridor logic oriented toward connecting the landlocked economies of Chad and the Central African Republic to the port of Douala, which accounts for a substantial bulk of Central African transit traffic. This means that the external drivers of road building in Cameroon are not only institutionally distinct from those in Benin and Mali, but geographically oriented in a different direction — toward Central Africa rather than intra-West African integration.

Additionally, Mali's landlocked status naturally changes its road-building logic. Mali's external trade has historically depended on long overland corridors to coastal ports — predominantly the Abidjan corridor, which handled the large majority of Mali's trade until instability in Côte d'Ivoire in the early 2000s forced a reorientation toward Dakar and Tema. While the Niger River provides seasonal navigability across roughly 1,800 km of the country, it functions as a partial and unreliable substitute for roads, confined to the high-water season of approximately June through December in normal rainfall years. Road-building decisions in Mali have consequently been shaped heavily by the transit corridor logic of connecting

Bamako to coastal ports, with an additional overlay — distinct from the other two countries — of large areas of desert in the north of the country.

Appendix section A provides more detail on the historical context and its relevance for road building in each country.

2.2 Census data

I use rich and comprehensive census data to deduce variation in local opportunity. I use data on 8 million observations from 1976 to 2013, across 163 localities from every available census from Benin, Cameroon, and Mali.[4] In the censuses described, I can geo-locate individuals at the second administrative unit level. These localities have a median population of 267,000 across all samples. Benin's Communes have a median population of 103,000 with an inter-quartile range of 71,000 to 173,000. Mali's circles have a median population of 308,000 with an inter-quartile range of 197,000 to 520,000. Cameroon's departments have a median population of 456,000 with an inter-quartile range of 225,000 to 907,000. A broadly comparable geographical unit in the US would be commuter zones. In Benin, I use censuses from 1992, 2002, and 2013, encompassing over 2.2 million individual-level observations in 77 consistent localities (Communes). In Cameroon, I use censuses from 1976, 1987, and 2005, covering over 3.4 million individual-level observations in 39 consistent localities (Departments). Finally, in Mali, I use censuses from 1998 and 2009, giving over 2.4 million individual-level observations in 48 consistent localities (Circles). In total, this gives me a sample of over 8 million individual-level observations across 163 localities and 444 locality-year cells. A benefit of using this data is that it is freely and publicly available from IPUMS International.

I use the causal effect of growing up in a given location on the probability of completing primary school as my main measure of local opportunity for three main reasons. First, due to the large informal sector, it is unclear whether later life income, as used in other contexts, is appropriate here. Additionally, data on primary completion rates are available at a fine geographic level over time and are less likely to suffer from significant measurement error. Second, primary completion rates are correlated with opportunity more broadly defined in later life. In this setting, individuals who have completed primary school are less likely to work in agriculture, have better housing quality, and higher incomes (Psacharopoulos and Patrinos, 2018). Finally, primary schooling is the most salient margin of education. In my sample, about a third of individuals have completed primary school, but only 7% have completed secondary school. Thus, although defensible as a relevant and general measure of

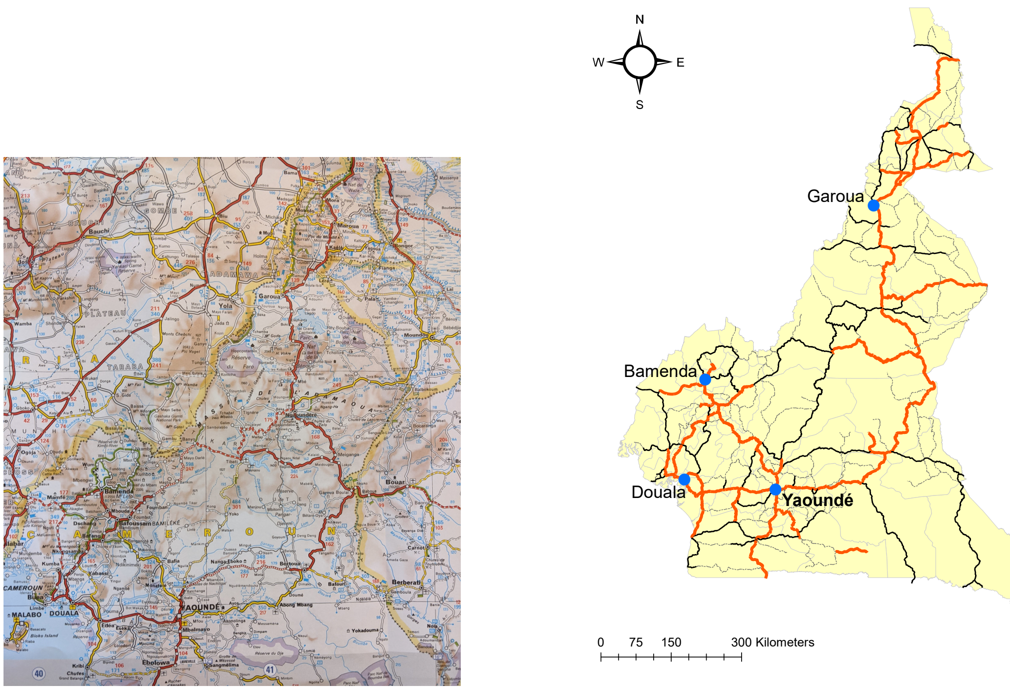
[4] Data is accessed from IPUMS (2020) and consists of 10% samples, with thanks to the National Institute of Statistics and Economic Analysis in Benin, the Central Bureau of Census and Population Studies in Cameroon and the National Directorate of Statistics and Informatics in Mali, who provided the underlying data.

local opportunity, primary education completion cannot speak to all potential dimensions, and so opportunity in this paper should be taken to mean local *educational* opportunity.

2.3 The changing geography of connectivity: Digitising historical Michelin road maps

Road data comes from historical Michelin maps, which I have digitised from the following years: 2019, 2012, 2003, 1986, 1976, and 1969. In these maps, it is possible to consistently classify roads into highways, paved roads, improved roads (laterite or gravel), and dirt roads. This classification provides a full description of the (main inter-city) roads over time since 1969 in each country. The ability to distinguish road type is of particular importance, as much of the variation in connectivity, especially in later years, comes from upgrading roads rather than building new ones. Figure 1 gives an example of this process. Panel 1a shows an image of the raw Michelin map of Cameroon in 2019, and panel 1b shows the digitised version.

Figure 1 Digitizing Michelin road maps — Cameroon 2019



(a) Michelin map (b) Digitized map

Notes: This figure shows in panel 1a the original Michelin road map of Cameroon in 2019 and in panel 1b the digitised version. In panel 1b, thick red lines are paved roads, dark black lines are improved roads, and grey lines are dirt tracks. Note that in panel 1a the colour of roads denotes their importance/ frequency of use for long-range trucking and does not necessarily reflect their size or quality, which is denoted instead by the thickness of outlines.

Michelin maps are themselves constructed using four main sources (Jedwab and Storeygard, 2021): the previous Michelin map, government road maps, local information from

Michelin tyre stores across Africa, and finally direct correspondence from users. It is generally thought that this process leads to consistent and accurate road mapping over time. Indeed, the success of Michelin maps relied on them being a trustworthy source of information, and so Michelin had a vested interest in providing accurate maps. However, it's still very possible that not every change is noted, and even if a change is noted, it could only be included with some lag. Additionally, although I can use variation in road upgrading, this can only be observed when a road changes categories. That is, road maintenance or changes that would not count as upgrades across categories (such as pothole filling) are not captured. This may be a particular issue in the more recent years, as it is expected that a greater proportion of road spending reflects this unobserved variation.

3 A model of place effects in spatial equilibrium

In this section, I develop a quantitative spatial economics model of education choice that embeds the place-effects framework of Chetty and Hendren (2018a) into a spatial equilibrium with endogenous wages and prices, building on Hsiao (2024) and Allen et al. (2020). The model serves three purposes. First, it delivers market access as the sufficient statistic linking road infrastructure to local educational opportunity, providing a principled way to measure the impact of any road, given its position within the entire network and the pre-existing distribution of economic activity. Second, it motivates the reduced-form estimating equation that I take to the data in Section 4, which also allows me to recover structural parameters. Third, it provides the structure needed to recover the aggregate effects of road building and decompose them into the "moving to opportunity" and "opportunity moving" channels in Section 5.

I take as my point of departure the place-effects literature, where outcomes, Y_k, for individuals k depends on factors specific to k and those specific to where k grew up, denote by $i(k, a)$, the location where k was at age a. In this paper, I am interested in the outcome Y_k, a dummy variable that indicates whether or not an individual has completed primary school. Following Chetty and Hendren (2018a) and others, I assume that place and individual effects are additively separable. Given this, I can decompose outcomes into that due to place effects and that due to individual effects as in equation 1.

$$Y_k = \sum_a \mu_{i(k,a)} + \theta_k \tag{1}$$

Place effects are denoted by μ_i where i is a given location, and θ_k captures place-invariant individual-specific effects. In this section, I provide a structural interpretation for this expression in the case where the outcome of interest is schooling. To do this I develop a quantitative spatial economics model of education and location choice. This model follows Hsiao (2024), extending his framework by endogenising education demand in spatial equilibrium.

Timing and dynamics.

Individuals are born in a given period t in an origin location i. They will stay and be educated in i before choosing a location j to move to in adulthood.[5] Individuals receive education shocks ξ_k and a vector of location-specific skill shocks $\{\varepsilon_{jk}\}$. They choose education e_k after observing their education shock, but before observing their skill shocks, they then observe skill shocks before choosing their migration location. After migrating, each migrant has one child who will be educated in their parents' migration location.

This framework displays cross-generation dynamics because the previous period's migration patterns determine the next period's young who become educated and themselves choose where to migrate. However, within an individual, all choices happen within the same period, avoiding the need to consider dynamic choices. I will also make the simplifying assumption that individuals do not consider the utility of their unborn child when considering education choices or migration decisions. The dynamics in this framework are therefore those of spatial equilibrium and will be captured entirely by population movements.

Education.

Individuals choose e_k to maximise expected future utility during work, subject to education costs.

$$e_k^* = \arg\max_e \{\bar{v}_{it}(e) - c_{it}(e, \xi_k)\} \tag{2}$$

Where $\bar{v}_{it}(e)$ is the future expected utility, and $c_{it}(e, \xi_k)$ is the cost of education in location i in period t and is subject to the education shock ξ_k. I parameterise costs as a linear function of education $c_{it}(e, \xi_k) = e \cdot \tau_{it}^{educ} \cdot \xi_k$. Individuals move in adulthood to the location that maximises their utility subject to location-specific shocks ε_{jk}, therefore: $\bar{v}_{it}(e) = \mathbb{E}[\max_j v_{ijt}(e, \varepsilon_{jk})|e]$.

Utility.

Individuals with education e gain utility from working and living in a given location j in period t, which comprises of three components. First, individuals gain utility from local amenities a_{jt}, second from real wage income w_{jt}/P_{jt}, and thirdly are subject to iceberg migration costs $(\tau_{ijt}^m)^{-1}$. Real wage income is given by $w_{jt}/P_{jt} = \frac{r_{jt}}{P_{jt}} h_{jt}$ where r_{jt}/P_{jt} is the

[5]Therefore, in this simple version of the theory, I restrict equation 1 to the single-location setting.

real wage per unit of human capital and h_{jt} is the quantity of human capital acquired. This is a non-linear function of education completed $h_{jt} = e^{\eta}\varepsilon_j$, where ε_j are the location-specific skills shocks that will be resolved before choosing location and are distributed as a type two extreme value distribution with parameter θ. The utility of an individual born in i, moving to j can therefore be written as follows.

$$v_{ijt} = a_{jt} \cdot \left(\frac{r_{jt}}{P_{jt}} e^{\eta} \varepsilon_j \right) \cdot \left(\tau^m_{ijt} \right)^{-1} \tag{3}$$

Using this equation and leveraging properties of type two extreme value distributions, I can therefore write an expression for expected utility.

$$\bar{v}_{it}(e) = \mathbb{E}[\max_j v_{ijt}(e, \varepsilon_j)|e] = e^{\eta} \sum_j \pi_{ijt} \frac{a_{jt} r_{jt}}{P_{jt} \tau^m_{ijt}} \mathbb{E}[\varepsilon_j|j] \tag{4}$$

Where π_{ijt} is the probability that an individual moves to j from i in period t. By considering the migration choice problem, this object can be characterised.

<u>Migration</u>.
Individuals choose location j to migrate to, taking education e as given. As ε_j has a type two extreme value, the utility representation given in 3 results in the familiar gravity formulation for migration probability π_{ijt}.

$$\pi_{ijt} = \frac{\left(a_{jt} \frac{r_{jt}}{P_{jt}} (\tau^m_{ijt})^{-1} \right)^{\theta}}{\sum_q \left(a_{qt} \frac{r_{qt}}{P_{qt}} (\tau^m_{iqt})^{-1} \right)^{\theta}} \tag{5}$$

Note that using properties of the type two extreme value distribution, I can also write $\mathbb{E}[\varepsilon_j|j] = \gamma \pi_{ijt}^{-1/\theta}$, where $\gamma = \Gamma(1 - \frac{1}{\theta})$ and $\Gamma(\cdot)$ is the Gamma function. Putting this together with equation 5, equation 4 can be simplified to the following.

$$\bar{v}_{it}(e) = \gamma e^{\eta} \left(\sum_j \left(a_{jt} \frac{r_{jt}}{P_{jt}} (\tau^m_{ijt})^{-1} \right)^{\theta} \right)^{\frac{1}{\theta}} = \gamma e^{\eta} MA_{it}^{\frac{1}{\theta}} \tag{6}$$

Where market access is defined as the local availability of high utility (high amenity and or real wage) locations $\mathrm{MA}_{it} = \sum_j \left(a_{jt} \frac{r_{jt}}{P_{jt}} (\tau^m_{ijt})^{-1} \right)^{\theta}$. The benefits of education are therefore tied to the local availability of high returns to education. Local changes to connectivity will make it easier for individuals to move to high-return locations, and this is well captured within this framework. However, changing connectivity may also alter the distributions of local returns itself. To capture this force, of opportunity moving, I endogenise local wages and prices.

Endogenous local real wages.

Firms produce location-specific differentiated goods under within-location perfect competition but are subject to iceberg trade costs denoted by τ^t_{ijt}. Individuals have CES preferences over local varieties with an elasticity of substitution σ. Human capital is the only factor of production. Perfect competition implies that price equals marginal cost, $p_{jt} = r_{jt}/B_{jt}$ where B_{jt} is j, t-specific productivity. CES preferences and iceberg trade costs imply a familiar gravity equation in trade, which in turn gives the usual price aggregator $P^{1-\sigma}_{jt} = \sum_i(\tau^t_{ijt} r_{it}/B_{it})^{1-\sigma} = TMA_{jt}$. Similarly, assuming symmetric trade costs, given human capital, and assuming goods markets clear, I have that $Y_{jt} = r_{jt}H_{jt} = p^{1-\sigma}_{jt} TMA_{jt}$. Solving the implied system I find $r_{jt} = (H^{-1}_{jt} TMA_{jt})^{1/\sigma} B^{(\sigma-1)/\sigma}_{jt}$. Where H_j is the total local stock of human capital in a given location.

Solving education choice.

Using the above derivations, the choice problem of an individual, k, can be solved to find the following.

$$\ln(e_k) = \alpha + \frac{1}{\theta(1-\eta)} \ln(\text{MA}_{it}) - \frac{1}{1-\eta} \ln(\tau^{educ}_{it}) - \frac{1}{1-\eta} \ln(\xi_k) \tag{7}$$

Where $\alpha = \frac{1}{1-\eta} \ln(\gamma\eta)$. This equation encapsulates the place effects model. It can be rewritten in the form given in equation 1 by noting that place-effects are captured by: $\mu_{it} = \frac{1}{\theta(1-\eta)} \ln(\text{MA}_{it}) - \frac{1}{1-\eta} \ln(\tau^{educ}_{it})$. Individual-specific effects are given by the final term, $\theta_k = -\frac{1}{1-\eta} \ln(\xi_k)$. Those with higher education skill shocks (lower ξ_k) will complete more education irrespective of local conditions. Those who grow up in locations with higher market access or lower costs of education will complete more education, irrespective of their skill shock.

The sensitivity of educational outcomes to changes in market access depends on two parameters θ and η. The parameter $\theta > 0$ captures the dispersion of location-specific skill shocks. Larger θ implies less dispersed location-specific skill, and so higher migration sensitivity to changes in local (real) wages. Therefore, a higher θ implies a lower sensitivity of education choices to market access as the impact of any location-specific changes is smoothed out to a greater degree through increased migration. The parameter $\eta \in (0, 1)$ captures the concavity of the human capital production function. An η closer to 1 implies slower decreasing returns to education and thus greater sensitivity of education completion to incentives to educate, including changes to market access. With only variation in market access, these two channels cannot be separately identified, and instead only the bundle $\beta = (\theta(1-\eta))^{-1}$ can be identified. However, this combination is still informative about the reduced form effect of changing connectedness on local opportunity — and it is to identifying this relationship that I now turn.

4 Reduced form: The impact of market access on local opportunity

There are two key challenges associated with taking equation 7 to the data. First, measurement: I need to be able to measure a location's market access and how this changes over time. Second, identification. As τ_{it}^{educ} and ξ_k are not observed, but likely correlated with market access, a pair of strategies needs to be developed to allow β to be identified. Approximating e_k with a primary school completion indicator Y_k, I can write equation 7 in a regression form as:

$$Y_k = \alpha + \beta \cdot \ln\left(\mathrm{MA}_{it}\right) + v_k \tag{8}$$

Where the error term is given by $v_k = -\frac{1}{1-\eta}\ln(\tau_{it}^{educ}) - \frac{1}{1-\eta}\ln(\xi_k)$. This formalisation allows a clear elucidation of the two key empirical challenges. First, individuals may select into better locations, causing a correlation between $\ln(\mathrm{MA}_{it})$ and $\ln(\xi_k)$. Second, locations with higher market access may also be conducive to education for endogenous reasons, causing a correlation between $\ln(\mathrm{MA}_{it})$ and $\ln(\tau_{it}^{educ})$, where note τ_{it}^{educ} can be interpreted as all unobserved factors relating to the cost (or benefit) of education in i in period t.

4.1 Measuring market access

The market access of a given location increases in the distance weighted power-sum of opportunities available in said location $\mathrm{MA}_{it} = \sum_j \left(a_{jt}\frac{r_{jt}}{P_{jt}}(\tau_{ijt}^{m})^{-1}\right)^{\theta}$. However none of a_{jt}, r_{jt}, P_{jt} or τ_{ijt} are observed in the data. To overcome this, a location's market access can be written in a more convenient way. Note that $L_{jt} = \sum_i \pi_{ijt}L_{it} = \left(a_{jt}\frac{r_{jt}}{P_{jt}}\right)^{\theta}\sum_i(\tau_{ijt}^{m})^{-\theta}L_{it}\mathrm{MA}_{it}^{-1}$. Denote $\widetilde{\mathrm{MA}}_{it} = \sum_j(\tau_{ijt}^{m})^{-\theta}L_{jt}\mathrm{MA}_{jt}^{-1}$. As shown by Donaldson and Hornbeck (2016) the only solution for this series of equations is that $\mathrm{MA}_{it} = a\widetilde{\mathrm{MA}}_{it}$ for some constant a. Thus, up to scale, $\mathrm{MA}_{it} = \sum_j(\tau_{ijt}^{m})^{-\theta}L_{jt}\mathrm{MA}_{jt}^{-1}$. This expression can be taken to the data as L_{jt} is observed. To calculate market access only $T_{ijt} = (\tau_{ijt}^{m})^{-\theta}$ needs to be estimated.

4.1.1 Calculating time-varying location-to-location travel times

I parameterise the iceberg travel costs as depending on i to j travel times over the road network prevailing at time t, denoted as t_{ijt}. These times are estimated using the prevailing road network at time t using Dijkstra's algorithm. Data on the road network is generated from historical road maps that I digitise, and travel times over the four road categories are taken from Jedwab and Storeygard (2021). The online appendix section D details how t_{ijt} is calculated. I then parameterise these costs in a log-linear fashion such that $\ln(T_{ijt}) = -\theta \cdot \ln(\tau^m_{ijt}) = -\tilde{\theta} \cdot \ln(t_{ijt})$, where $\tilde{\theta} = b \cdot \theta$. Therefore, to calculate market access, it only remains to estimate $\tilde{\theta}$.

4.1.2 Estimating $\tilde{\theta}$

We can estimate $\tilde{\theta}$ using a standard gravity migration equation coupled with data from the censuses on locality-to-locality moves in each period. From the theory the number of individuals moving from i to j in period t can be written as $M_{ijt} = \pi_{ijt} L_{it}$ given the previously derived expression for π_{ijt} and including fixed effects results in the following framework given in equation 9. Where α_{it} and τ_{jt} denote origin-time and destination-time fixed effects respectively.

$$\ln(M_{ijt}) = -\tilde{\theta} \cdot \ln(t_{ijt}) + \alpha_{it} + \tau_{jt} + \varepsilon_{ijt} \tag{9}$$

This theory-consistent gravity equation can be estimated using pseudo Poisson maximum likelihood (Silva and Tenreyro, 2006; Yotov et al., 2016) to find $\hat{\tilde{\theta}} = \hat{\theta} \cdot \hat{b}$. By including this rich set of fixed effects, this specification only uses variation in travel times stemming from bilateral travel costs. For example, a common threat to identification would be that as a location becomes more attractive, planners are more likely to build better connections to this location. Conversely, planners might build connections to a previously flagging location in order to galvanise it. In either case, such behaviour would be captured by the fixed effects in this specification.

Any remaining threats to identification must operate at the bilateral level. One first-order concern of this type is that bilateral cultural ties between two locations may both influence migration and the strength of the road connection between these places. In this case cultural ties, denoted by ct_{ijt}, would be an omitted variable potentially biasing $\hat{\tilde{\theta}}$. Using data from Weidmann et al. (2010), I can control directly for the group similarity between locations.

More generally, one may be concerned that a planner builds better connections between two locations in order to facilitate existing or encourage new migration between these two places. If this were the case, changes in travel time due to road building and upgrading may be for reasons endogenous to migration rates. Moreover, preexisting strong connections may be due to great social or economic ties, which may also encourage greater bilateral migration. To overcome these concerns, I first consider estimating the gravity equation in the cross-section only, and second, consider using straight-line “as the crow flies” bilateral centroid distances as opposed to the estimated travel costs, which use the actual road network.

Table 1 shows the results from estimating equation 9 using the approaches described above. These estimates show remarkable stability across specifications. Due to this consistency, the first column is taken as the preferred specification, leading to an estimate of $-\tilde{\theta} = -1.516$. Few estimates of this parameter are available in the literature from a developing country. Perhaps most informative Morten and Oliveira (2024) estimates an elasticity of -1.698 using meso-to-meso migration over a five-year period in Brazil. The estimating sample and context are very different between these two settings, but it is nevertheless reassuring that both estimates are similar.

Table 1 Estimating the Migration Elasticity

	(1) PPML	(2) PPML	(3) OLS	(4) OLS	(5) PPML	(6) PPML	(7) PPML
Log(Travel time)	-1.516***	-1.494***	-1.418***	-1.417***	-1.482***	-1.510***	
	(0.0379)	(0.0388)	(0.0203)	(0.0203)	(0.0388)	(0.0460)	
Group Similarity					0.106	0.124	0.0140
					(0.110)	(0.119)	(0.128)
Log(Dist. Crow)							-1.244***
							(0.0411)
Destination-time FE	X	X	X	X	X	X	X
Origin-time FE	X	X	X	X	X	X	X
Sample	Panel	Panel	Panel	Panel	Panel	X-section	X-section
Winsorised		1 & 99		1 & 99			
N	26234	26234	13616	13616	26234	9450	9450

Notes: This table shows the results from estimating equation 9 using various specifications. All specifications use destination-time and origin-time fixed effects with standard errors clustered at the origin-destination pair level. Column one uses PPML, column two also uses PPML, and winsorizes migration flows at the 1st and 99th percentile. Columns three and four instead use a log-linear OLS specification. Column five again uses PPML controlling for group similarity. Columns six and seven move to the cross-section, taking the first available year of data; in both of these columns, group similarity is controlled for. In column seven, instead of using travel times estimated from the actual road network, I use as-the-crow-flies centroid-to-centroid distances.

With estimates of $\tilde{\theta}$, in hand using bilateral travel times I can then calculate each location's market access: $\text{MA}_{it} = \sum_j t_{ijt}^{-1.516} L_{jt} \text{MA}_{jt}^{-1}$. This is a series of non-linear equations across locations, which can be efficiently solved using an iterative procedure and has a unique up-to-scale solution (Donaldson and Hornbeck, 2016). Table B.1 in the online appendix shows that estimates of $\tilde{\theta}$ are stable across countries and years.

4.2 Identifying the reduced form effect of market access on local opportunity

To identify the causal effect of market access on local opportunity, I combine a movers design with a market access instrumental variables strategy that leverages plausibly exogenous sources of variation.

4.2.1 Movers Design

The movers' design uses variation in exposure to different levels of market access by comparing individuals who moved to high market access locations earlier in their childhood to those who moved later and so were less exposed. This allows selection into high-quality (high market access) locations, but requires that selection does not vary systematically with the age at which the child moves. This is the standard assumption in this literature, that previous work has interrogated (Milsom, 2023). However, due to its importance for identification, I perform various robustness and validation exercises probing it in section 4.4.

To operationalise the movers' design, I restrict my sample to one-time movers who are between the ages of 14 and 18 at the time of being surveyed, but could have moved at any age prior. In my sample, 81% of individuals in this age range have not moved, 13% have moved once, and 6% have moved more than once. I then make the simplifying assumption of a linear dose effect following earlier work in this setting (Milsom, 2023; Alesina et al., 2021). Under this assumption, what matters for an individual is their exposure to market access during their childhood. I compare children in the same cohort who move from the same origin location to destination locations with different levels of market access at different ages. Leveraging only variation in the age-at-move, I can then estimate the causal effect of spending an additional year of childhood in a one percent higher market access location on the probability of completing primary school.

We can write equation 1 for a sample of one time movers who are aged between 14 and 18 at the time of the census as $Y_k = \frac{1}{14}\left(m_k * \mu_{o(k),c(k)} + (14 - m_k) * \mu_{d(k),c(k)}\right) + \theta_k$. Where $o(k)$ is individual k's origin location, $d(k)$ is their destination location, m_k is the age at move, and $c(k)$ is k's cohort. From now on, I will suppress implicit dependence on k. I can include origin by cohort fixed effects and write place effects in terms of differences to find $Y_k = \left(1 - \frac{m_k}{14}\right) \cdot \Delta\mu_{odc} + \alpha_{oc} + \theta_k$. Following the theory discussion, I then write the location effect for a given location i as $\mu_{ic} = \beta \cdot \ln(\text{MA})_{ic} + v_{ic}$, where v_{ic} captures unobserved aspects of place effects. Substituting this into the above, I find $Y_k = \beta\left(1 - \frac{m_k}{14}\right)\Delta\ln(\text{MA}_{odc}) + \alpha_{oc} + \tau_m + \varepsilon_k$, where $\varepsilon_k = \left(1 - \frac{m_k}{14}\right)\Delta v_{odc} + \theta_k$ is unobserved. The conditioning on $\Delta\ln(\text{MA}_{odc})$, including age-at-move fixed effects, and finally allowing effects to differ before vs after moving at age 13, gives me the specification in 10. Following Milsom (2023), and Chetty and Hendren

(2018a), I allow differing intercept and slope coefficients for those who move before age 14 and those who move after (inclusive). Those who move after 14 have (almost all) already completed primary school,[6] and so market access should not affect their outcomes. The coefficient of interest is β_1, which can be interpreted as the impact of spending an additional year of childhood in a 100% higher market access location on the probability of completing primary school.

$$\begin{aligned} Y_k = \mathbb{1}_{[m_k \leq 13]} \cdot (\alpha_1 + \beta_1 \cdot (14 - m_k)) \cdot \Delta \ln(\text{MA}_{odc}) + \\ \mathbb{1}_{[m_k > 13]} \cdot (\alpha_2 + \beta_2 \cdot (14 - m_k)) \cdot \Delta \ln(\text{MA}_{odc}) + \\ \tau_m + \alpha_{oc} + \varepsilon_k \end{aligned} \tag{10}$$

By specifying the error term ($\varepsilon_k = \left(1 - \frac{m_k}{14}\right) \Delta v_{odc} + \theta_k$), the two potential sources of bias can be seen clearly. First, correlation between $m_k \Delta \ln(\text{MA}_{odc})$ and θ_k will cause bias. This occurs exactly when systematically "better" or "worse" individuals sort into higher or lower market access locations systematically earlier or later. That is, as with normal mover designs, sorting into better locations is permissible, but not that this happens systematically differently at different ages-at-moves. The assumption maintained here is weaker than the usual, as here I only require no such selection into locations on one characteristic — market access. The second potential source of bias is correlation between $\Delta \ln(\text{MA}_{odc})$ and Δv_{odc}. This will occur if higher market access locations are better on some other unobserved dimension, i.e., if there is a correlation between $\ln(\text{MA}_{lc})$ and v_{lc}. This seems possible in the case of road building, as roads are not exogenously placed. The planner could build roads to service booming locations or to galvanise flagging ones, and market access terms will inherit this endogeneity.

4.2.2 Not-on-least-cost-path variation

To overcome the endogeneity of road placement and therefore market access, I require an additional identification strategy and leverage the construction of market access terms to consider only not-on-least-cost-path variation.

Previous identification strategies designed to overcome the endogeneity of road placement include using placebo lines from planned but unbuilt routes (Donaldson, 2018; Okoye et al., 2019), using straight line or least-cost path spanning tree instruments[7] (Moneke, 2020; Michaels, 2008; Ghani et al., 2016; Faber, 2014), or leveraging *far-away* variation in road

[6] Figure C.2 in the appendix shows that most who go on to complete primary school have done so by age 14.

[7] Locations that just happen to lie between two cities that are being connected by a road may be plausibly described as exogenous. This is also known as the *inconsequential units* IV.

changes (Donaldson and Hornbeck, 2016; Jedwab and Storeygard, 2021). In my setting, it is difficult to see how the first two approaches can be implemented. First, I don't have data on unbuilt but planned placebo lines. Second, there is no clear set of locations that are being connected, and in addition, a significant proportion of the variation in travel times comes from road upgrading rather than the building of entirely new roads, in this setting, it's unclear how localities are "incidentally" connected.

The third strategy, leveraging far-away variation in roads, is appealing but suffers from a number of known drawbacks. First, it's unclear how far "far-away" should be; and although researchers can present many distances, it is ultimately an ad-hoc choice. Second, and more fundamentally, variation due to large projects which may be far away but for endogenous reasons, or relatively far away connections that are built to ease transport to, or encourage trade to a given location, remain threats to identification.

In this paper, I propose a novel identification strategy that builds upon the far-away variation approach by considering not-on-least-cost-path variation. The not-on-least-cost-path variation approach only uses changes in a locality i's market access that stems from indirect changes to all other locations' market access freezing migration costs along the least cost path from i to all other locations.[8] This approach has two intuitive explanations. First, one can consider it as the same as using far-away variation, but whereas far-away variation defines distance over Euclidean space, not-on-least-cost-path defines distance over network space, where a "distance" of one refers to one-degree removed indirect variation. Under this interpretation, the network structure is taken seriously and resolves the ad-hoc nature of what "far-away" may mean by appealing to the theory. Secondly, one can think of not-on-least-cost-path variation as approximating the decision-making process of the policymaker building roads, and using the residual variation. If a central planner builds roads to or from i in order to directly improve its connectivity, this variation is not used, and instead the residual variation in market access is leveraged. If, for example, a planner builds schools (affecting τ_{it}^{educ}) and roads in a location, this is exactly the variation that this instrument will not leverage.

To formalise the above discussion, consider a generic market access variable $\text{MA}_{it} = \sum_j (\tau_{ijt}^m)^{-\theta} L_{jt}(\text{MA}_{jt})^{-1}$. Following Donaldson (2018), I first remove own-location market access (therefore sum over other locations $j \neq i$) and freeze the market size variable at the initial level (L_{j0}) as these objects are co-determined with the outcome variable. Then I can use the not-on-least-cost-path variation (freeze the least cost path to the initial value τ_{ij0}^m)

[8] This approach is not to be confused with one variation of the incidental connection approach which uses an estimated least cost route over an estimated cost surface to instrument for actual connections. I don't take this approach in this paper for the same reasons that incidental connection approaches, in general, are unsuitable, and in addition, because a significant proportion of the variation comes from within road quality, and it is unclear how to leverage this intensive margin dimension using this approach.

by constructing the following instrument.

$$\mathrm{MA}_{it}^{NLCP} = \sum_{j \neq i} (\tau_{ij0}^{m})^{-\theta} L_{j0} \left(\mathrm{MA}_{jt}^{-i}\right)^{-1}$$

Where $\mathrm{MA}_{jt}^{-i} = \sum_{k \neq i} (\tau_{jkt}^{m})^{-\theta} L_{k0} \left(\mathrm{MA}_{kt}^{-i}\right)^{-1}$. This object freezes the least-cost-path from i to j, by fixing the cost matrix at period zero (τ_{ij0}^{m}). The within-summation MA_{jt}^{-i} terms are calculated by summing over all locations not including i to further purge direct effects. Changes in MA_{it}^{NLCP} will therefore only arise due to changes in other locations' market access MA_{jt}^{-i} (for $j \neq i$) due to road-building — not the direct effect of decreasing transport costs between i and j.

The instrument can be re-written as a shift-share instrument where the shares are given by $s_{ij} = (\tau_{ij0}^{m})^{-\theta} L_{j0}$ and the shifts are given by $g_{jt} = \left(\mathrm{MA}_{jt}^{-i}\right)^{-1}$, then I can write $\mathrm{MA}_{it}^{IV} = \sum_{j \neq i} s_{ij} \cdot g_{jt}$.[9] The identifying assumption is that, conditional on fixed effects and controls (including region×year), these destination network shocks are orthogonal to unobserved determinants of outcomes in i, except through their impact on i's market access. This helps clarify the main threats to identification. First, baseline exposure may be endogenous. To overcome this, I follow the approach of Borusyak and Hull (2023) and re-centre the instruments; details of this procedure are in sub-section 4.4. Second, non-modelled spillovers. Other shocks in j could affect i proportionally to s_{ij}, such as political economy shifts, or other public infrastructure correlated with roads. Similarly, regional policy combining road building and other projects would cause correlation across local shocks and i's outcomes. These threats are dealt with by (1) combining the not-on-least-cost-path approach with the far-away variation approach to remove variation local to i, and (2) including region by year fixed effects in our analysis. I also directly control for related concerns surrounding clientelism.

To validate the reasonableness of the arguments made in this section, I provide some Monte-Carlo simulation evidence presented in table 2. I create a simulated network of over 1,000 nodes. Each node lies on the unit square and belongs to one square quadrant, denoted as a region. I randomly generate an undirected graph over these nodes, where each location is directly connected to its six nearest neighbours, and there is a small probability of a

[9] This formulation allows us to clarify where the variation comes from by decomposing changes in market access into the components given below.

$$d\ln(\mathrm{MA}_i) = \sum_{j \neq i} K_{ij} \cdot d\ln\left((\tau_{ijt}^{m})^{-\theta}\right) - \sum_{j \neq i} K_{ij} \cdot d\ln(\mathrm{MA}_j)$$

Where $K_{ij} = s_{ij}\mathrm{MA}_j^{-1}$. The first term of the right-hand side gives the direct effect of changes to the network on market access in i, whereas the second term gives the indirect network-mediated effect. It is this second term that is captured by the not-on-least-cost-path instrument, which effectively purges market access of the direct effect.

random long-link connection between any two randomly chosen nodes. Over this "road" network, I then calculate the travel time between each pair of nodes along connected links as $\omega_{ij} = ||s_i - s_j|| + 0.05$ where s_i are the two-dimensional coordinates of location i. This is converted into a travel cost using $T_{ij} = d_{ij}^{-2}$, where d_{ij} is the shortest path from i to j along the network. This creates the baseline network. Each node is also given a random baseline population $L_i \sim LogNormal(0, 0.5^2)$ and the sum of population is normalised to one.

To this I construct a road-building shock on a random subset of edges according to two data-generating processes. The first incorporates endogenous road upgrading. Let location i have some unobserved characteristic $u_i \overset{iid}{\sim} \mathcal{N}(0, 1)$. Then each baseline edge (a, b) is upgraded with probability increasing in u_a and u_b and equal to $\mathbb{P}(\text{upgrade } (a, b)) = 0.18\frac{p_a+p_b}{2}$ where $p_i = \Lambda\left(1.8\frac{u_i-\bar{u}}{\sigma_u}\right)$ and $\Lambda(x) = \frac{1}{1+e^{-x}}$. If a road segment is upgraded, its travel time is multiplied by 0.6. The shortest travel times between each pair of nodes are then calculated over this new, randomly upgraded, network. The second data-generating process follows the first, but additionally introduces some regional building program where for some region r, if the start and end point of the road both fall within r, there is an additional probability that the road is upgraded given by 0.33.

In each location under each scenario, I then calculate market access terms and market access using the not-on-least-cost-path instrument, using far-away variation (only variation beyond the 20th percentile of distances), and both. Let $\varepsilon_i \overset{iid}{\sim} \mathcal{N}(0, 0.25^2)$ be a random shock. I then construct two data-generating processes for the location outcome y_i following the two processes for road upgrades discussed above.

$$\text{DGP-A:} \qquad y_i = 1 \cdot \log(\text{MA}_i) + 2 \cdot u_i + \varepsilon_i$$

$$\text{DGP-B:} \qquad y_i = 1 \cdot \log(\text{MA}_i) + 2 \cdot u_i + 2 \cdot \mathbb{1}_{[r(i)=r]} + \varepsilon_i$$

The true coefficient on log market access is therefore one. Using these data-generating processes, I regress y_i on log market access using OLS, using OLS plus region fixed effects, using the far-away IV, using the not-on-least-cost-path IV, using the NLCP and far-away IV, and using the NLCP IV and region fixed effects. Table 2 shows the results averaged over 250 simulations.

Panel A of table 2 shows the performance of the various estimation strategies when the endogeneity comes from correlated local characteristics. The OLS coefficient is very biased (150%), and including regional fixed effects does not attenuate this bias. Focusing on far-away variation reduces the bias to 50%, but does not eliminate it. However, using a not-on-least-cost-path approach significantly reduces the bias to around 8%. Panel B then turns to the second data-generating process, which includes a regional development program. Here, naive OLS is hopelessly biased, although including regional fixed effects does reduce this bias it remains at 133%. The far-away variation approach is not helpful here, nor is only using the not-on-least-cost-path approach — both have bias similar to naive OLS. Including both not-on-least-cost-path and far-away variation deals with both sources of endogeneity fairly well and exhibits a bias of 33%. However, the most successful approach is simply to condition on region fixed effects when using the not-on-least-cost-path instrument, which has a bias of 9%. In sum, the simulation results show that the not-on-least-cost-path approach is the most successful in dealing with the non-random placement of roads. However, it is not without caveats, and conditioning on regional fixed effects is important when dealing with regionally correlated development programs.

Table 2 Not-on-least-cost path simulation evidence

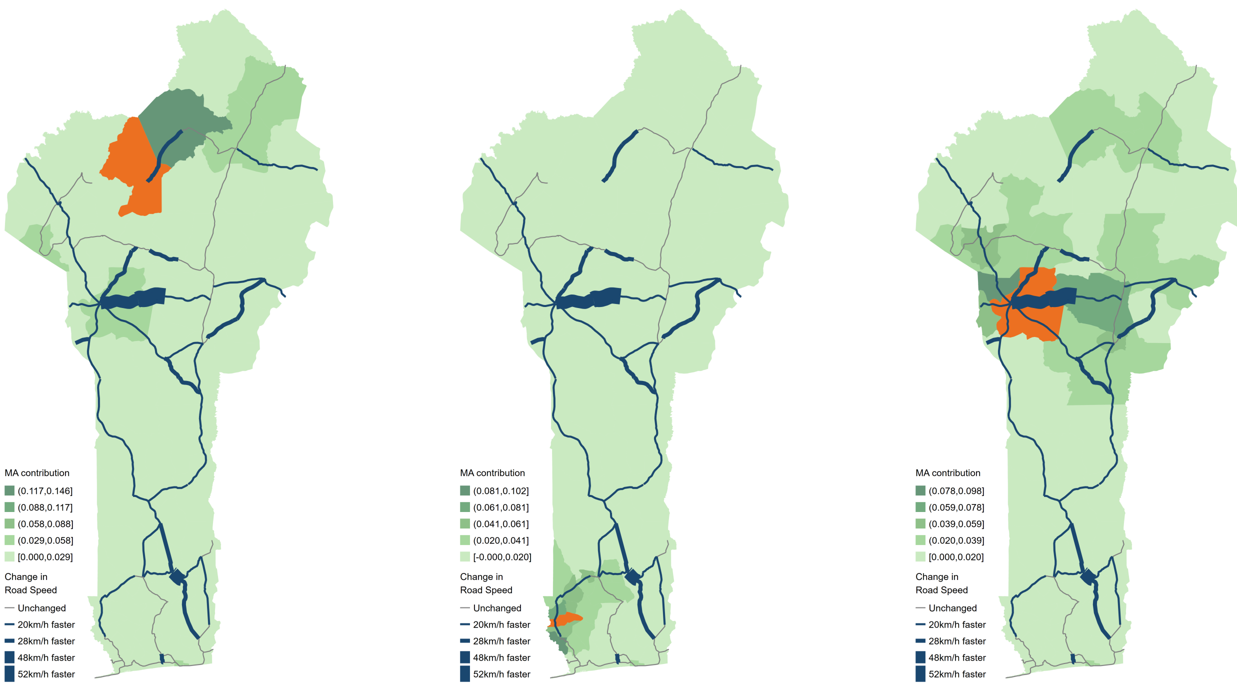
Estimator	Mean	SD	Bias
DGP A: Local Targeting			
OLS	2.554	0.428	1.554
OLS + Region FE	2.564	0.435	1.564
2SLS: Far-away IV	1.495	0.863	0.495
2SLS: NLCP IV	0.919	0.506	-0.081
2SLS: NLCP + Far-away variation IV	0.933	0.930	-0.067
2SLS: NLCP IV + Region FE	0.921	0.517	-0.079
DGP B: Local and Regional Targeting			
OLS	4.319	0.366	3.319
OLS + Region FE	2.328	0.396	1.328
2SLS: Far-away IV	4.101	0.859	3.101
2SLS: NLCP IV	-3.480	1.468	-4.480
2SLS: NLCP + Far-away variation IV	1.325	1.279	0.325
2SLS: NLCP IV + Region FE	0.909	0.505	-0.091

Notes: This table reports Monte Carlo evidence for the market-access instruments described in the text. Each simulation draws a network of 1,000 nodes on the unit square (nodes assigned to one of four regions), baseline populations, and road-upgrading shocks; outcomes are generated under two data-generating processes: (A) $y_i = \log(\mathrm{MA}_i) + 2u_i + \varepsilon_i$ (endogenous upgrading through u_i) and (B) $y_i = \log(\mathrm{MA}_i) + 2u_i + 2\mathbb{1}[r(i) = r] + \varepsilon_i$ (endogenous upgrading plus a region-level program). For each draw, I estimate the coefficient on $\log(\mathrm{MA}_i)$ using: OLS; OLS with region fixed effects; IV using "far-away" variation (constructing exposure using only destinations beyond the 20th percentile of network distances); IV using the not-on-least-cost-path (NLCP) instrument, i.e., holding bilateral travel costs at baseline least-cost paths (τ^m_{ij0}) and using shocks $g_{jt} = (\mathrm{MA}^{-i}_{jt})^{-1}$; IV combining NLCP and far-away restrictions; and NLCP-IV with region fixed effects. Entries are averages across 250 simulations. SD refers to the standard deviation across simulations.

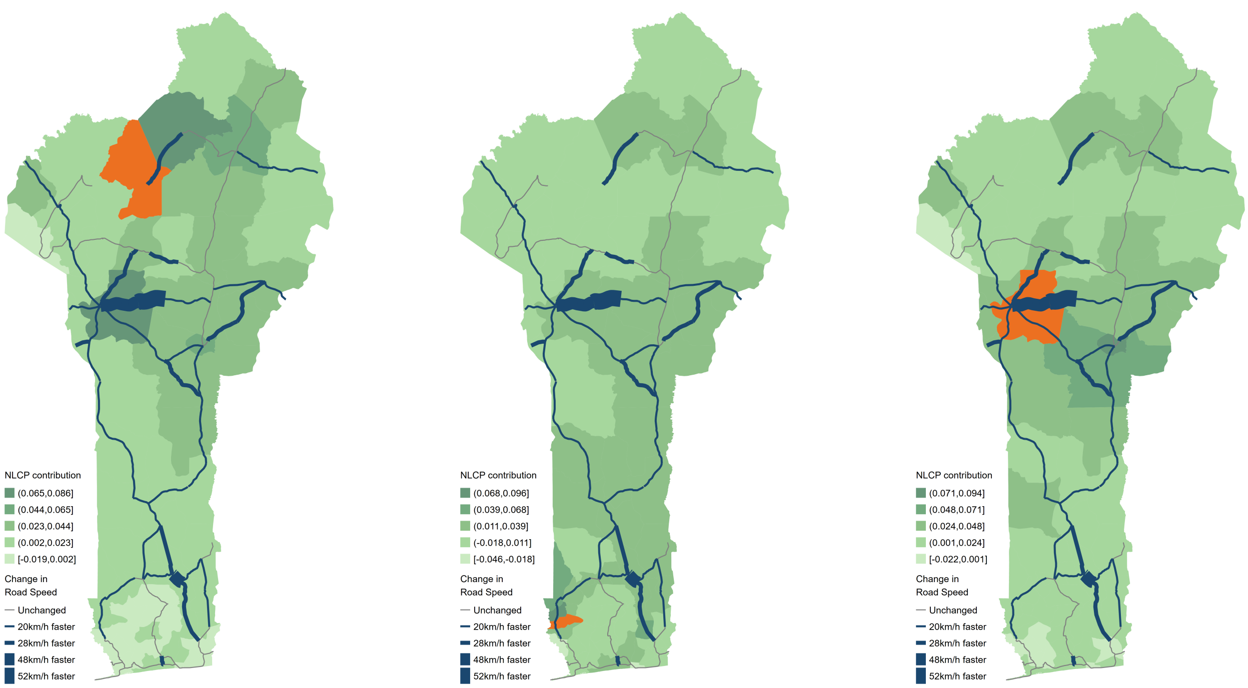
The simulation exercise culminating in the results shown in table 2 gives evidence to suggest that for a given, plausible, theoretical data-generating process, the not-on-least-cost-path instrument performs well. However, these theoretical data-generating processes may be far removed from reality. To better understand what variation the instrument picks up in the real-life data, we can consider some example cases. Figure 2 shows three such example cases in Benin. This figure shows, for a given focal location (highlighted in orange), which locations the long-difference (1970-2020) variation in market access or the market access instrument comes from. For a given focal location i, the (leave-origin-out) change in market access is given by $\Delta \mathrm{MA}_i = \sum_{j \neq i} (\tau^m_{ij2020})^{-\theta} L_{j2020} \mathrm{MA}^{-1}_{j2020} - \sum_{j \neq i} (\tau^m_{ij1970})^{-\theta} L_{j1970} \mathrm{MA}^{-1}_{j1970}$, each donor location j therefore contributes $\xi^i_j = \left((\tau^m_{ij2020})^{-\theta} L_{j2020} \mathrm{MA}^{-1}_{j2020} - (\tau^m_{ij1970})^{-\theta} L_{j1970} \mathrm{MA}^{-1}_{j1970} \right) (\Delta \mathrm{MA}_i)^{-1}$. In this first row of figures (a), (b), and (c), I show how variation in the raw (leave-origin-out) market access variation is proportioned across space, that is, I plot ξ^i_j for three given i. It is clear from these figures that locations nearby, or those that saw large decreases in travel costs to/from the origin, contribute more to variation. In the second row of figures, I plot the analogous quantity for the not-on-least-cost-path instrument, keeping the same focal location i in each column. Graphically, these sub-figures (d), (e), (f) show more diffuse sources of variation. The variation in the not-on-least-cost-path instrument is less likely to come from nearby locations or be influenced by changes in the connectivity between the focal location and other destinations.

Figure 2 Example real-life variation captures by MA_{it} and MA_{it}^{NLCP}.

(a) MA: Kérou (b) MA: Dogbo (c) MA: Djougou



(d) NLCP: Kérou (e) NLCP: Dogbo (f) NLCP: Djougou



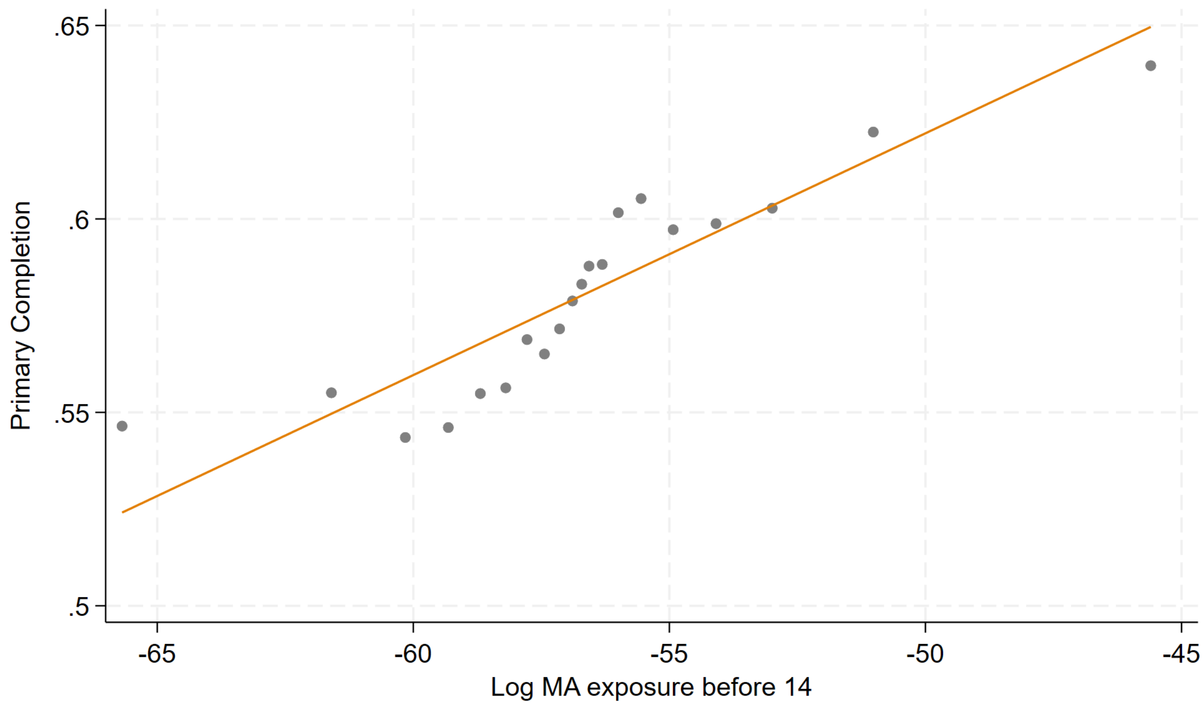
Notes: These maps show the proportion of variation in market access (top row) or the not-on-least-cost-path IV (second row) that comes from each sub-location within Benin for a given focal location that is highlighted in orange. Variation is calculated as the long difference 1970-2020. A darker green indicates a higher proportion of the overall variation is due to that area. Superimposed on each map are changes in the road network over this period. Thicker blue lines indicate larger changes, thin gray lines indicate no change.

Finally, regional or corridor-based road-building strategies adopted for reasons endogenous to local outcomes are the greatest remaining threats to identification when adopting the not-on-least-cost-path strategy. This is intuitive, as in such strategies, roads could be built between two locations to improve outcomes in a third unconnected but nearby area. This intuition is born out in the simulation results in DGP:B, where the NLCP IV alone is considerably biased. This simulation, however, also shows the solution. Including regional fixed effects removes the bias. However, in the simulation, the dimension on which the area-based policy was decided was the region, exactly what is controlled for. In real life, it is possible that regional policy does not neatly fit within administrative boundaries. To account for this, I comprehensively review all regional or corridor-based road-building strategies affecting Benin, Cameroon, and Mali. I identify 16 such major initiatives. In regressions, I can then control for a fixed effect for each such initiative and therefore only leverage within variation. As well as controlling for locality fixed effects, this approach allows me to control for time-varying initiatives in a principled manner. Appendix H in the online Appendix provides details on each identified initiative and shows that the results are unchanged by the inclusion of destination-locality level road-corridor initiative fixed effects (origin-locality by year fixed effects absorb all variation from the origin-location level). This is reassuring, as although some such initiatives may have been missed, the largest ones should be captured.

4.3 Results

Figure 3 shows the binscatter relationship between primary school completion rates and log market access exposure before age 14, controlling for origin by cohort and age at move fixed effects, but without following a movers design or tackling the endogeneity of market access. Nevertheless, this figure shows a strong positive and approximately linear association; individuals who were more exposed to higher market access during the first 13 years of their childhood were more likely to complete primary education.

Figure 3 Binscatter relationship without instrumenting



Notes: This figure shows the non-parametric relationship between market access exposure before 14 and primary completion in a bin-scatter figure, conditional on origin by cohort and age at move fixed effects.

Figure 4 shows the results from estimating equation 10 on a sample of one-time movers between the ages of 14 and 18. It displays coefficients $\widehat{\beta}_1$ and $\widehat{\beta}_2$ as well as their 95% confidence intervals. Each colour and marker shape corresponds to a different empirical strategy. Each specification includes origin location by year born by census year by age fixed effects and age-moved fixed effects, and standard errors are clustered at the origin-year born level. In blue circles, I plot the baseline results from regressing primary school completion on log market access. In red diamonds, I use the same specification, but instrument market access with the not-on-least-cost-path instrument. In green triangles, I then employ the mover's design described in equation 10. In yellow squares and purple hollow circles, I repeat this specification, instrumenting the change in market access with the not-on-least-cost path strategy and the not-on-least-cost-path plus far-away variation strategy.

Figure 4 Reduced form results

Notes: This figure shows the results from the reduced form analysis. All regressions are run on a sample of 14 to 18-year-olds and include origin location by year-born by census-year by age fixed effects, and age moved fixed effects. Standard errors are clustered at the origin by year of birth level, and indicated by horizontal lines. In blue with a circular marker, I show the baseline results of regressing average log market access over one's childhood on primary completion. In red with a diamond marker, I instrument market access using not-on-least-cost-path variation as discussed in the text (the first stage Kleibergen-Paap rank Wald F statistic is 8695). In green with triangular markers, I then report slope coefficients from estimating the movers design given in equation 10. These coefficients refer to movers before and after age fourteen as indicated on the x-axis. In yellow with square markers, I then combine the instrumental variables approach with the movers design (the first stage Kleibergen-Paap rank Wald F statistic is 322.6). Finally, in purple, I use a stricter instrument that combines the not-on-least-cost-path approach with that using far-away variation with a 50km cut off (the first stage Kleibergen-Paap rank Wald F statistic is 181.0).

Across all specifications, Figure 4 shows a clear positive impact of exposure to high market access on primary completion rates before 14, and a reassuringly null impact after. Coefficients can be interpreted as the percentage point impact of spending an additional year in a one percent higher market access location. Therefore, the overall impact for non-movers in a given location of a one percent increase in market access is a $0.005 \times 14 = 0.07$ percentage point increase in the probability of completing primary education. The standard deviation of log market access is about one, so moving to a one standard deviation higher market access location at birth increases the probability of completing primary school by 7 percentage points or 12% of the average primary completion rate, which is 58% in my estimating sample.

4.4 Robustness and validity checks

In this section, I discuss four robustness and validity checks of the above results. Details of these tests and results can be found in the online appendix section I.

First, the movers' design relies on the assumption that selection effects do not vary with the age at move. This cannot be tested in general; however, I can test whether selection on some observable characteristics varies with the age at move. To do this, I compare the "quality of move" measured as the change in location market access for those whose mothers have, and do not have, primary education, at different ages. If individuals with educated mothers systematically moved earlier (or later) to higher market access locations relative to those with non-educated mothers, this would be in violation of the identifying assumption. Figure I.1 in the online appendix section I plots these results and finds no difference between those with and those without educated mothers.

Second, I include household fixed effects in equation 10, therefore only focusing on variation in age-at-move within a household across siblings. This specification obviates concerns around time-invariant household characteristics driving the observed relationship. Figure I.2 in online appendix I shows the results. When including household fixed effects, I find that $\widehat{\beta}_1$ is slightly attenuated, but is far from being significantly so in either an economic or statistical sense.

Third, I show the robustness of the main result to varying specifications. Figure I.3 shows that similar results are found if (i) I don't include origin by year born fixed effects but only year born fixed effects, (ii) I replace origin by year born fixed effects with origin fixed effects only, (iii) I remove age at move fixed effects, and (iv) I perform a PPML regression in levels.

Finally, I control for the possible non-random exposure to exogenous shocks baked into a market access type design Borusyak and Hull (2023). Intuitively, this problem arises when, under random road placement, some locations will still expect to have higher or lower market access due to, for example, their initial location in the network (central locations would mechanically see higher gains in market access). To overcome this, I follow Borusyak and Hull (2023) and specify a data-generating process for market access. I construct 250 random road networks starting from the baseline in 1970 and randomly upgrading roads until the actual country-year change in travel time has been reached. I then calculate each location's market access over each random network in each period and use the average of these terms, a location's *expected* market access. I then control for this expected market access and repeat the analysis shown in Figure 4. The online appendix section B shows the results from this exercise in figure B.1 — coefficients do not qualitatively or quantitatively change.

In addition to the above, one may have setting-specific empirical concerns. In the online appendix section J, I show that in this setting, clientelism is not a threat to identification

Burgess et al. (2015). In the online appendix section K, I show that top-coding due to primary completion rates nearing 100% is not a concern. Finally, in the online appendix section L, I show that the prevalence of Koranic schools or Medersas remains too low to be of empirical relevance in my setting.

5 Quantification and mechanisms

Section 4 recovered the reduced form impact of changes in connectivity on local opportunity. However, to recover the aggregate impact of road building since 1970 and decompose channels, I need to know the counterfactual market access under the scenario where no roads had been built. To recover this object, I will proceed in four steps. First, I derive a parsimonious non-linear system of equations in each location that determines all endogenous variables. Second, to avoid backing out location fundamentals, I solve this system in differences using exact hat algebra. Third, I identify the remaining model parameters. Fourth, I check for the existence and uniqueness of equilibrium and solve the system for the given counterfactual road network.

5.1 Solving the system

We can solve this system using exact-hat algebra (Dekle et al., 2008) into one of three endogenous variables in three equations in each location, where hats denote counterfactual changes. These three endogenous variables are (1) the stock of human capital in a given location $\hat{H}_{it}$, (2) market access $\widehat{\text{MA}}_{it}$ and (3) goods market access $\widehat{\text{TMA}}_{it}$. The resulting expressions are given below; details of the derivation can be found in the appendix section M.

$$\widehat{H}_i^{1+\frac{\theta}{\sigma}} = \widehat{\text{TMA}}_i^{\delta} \cdot \sum_j \omega_{ij} \cdot \left(\widehat{\tau}_{ij}^m\right)^{-\theta} \cdot \widehat{\text{MA}}_j^{\lambda}$$

$$\widehat{\text{TMA}}_i = \sum_j \rho_{ijt} \cdot \left(\widehat{\tau}_{ij}^t\right)^{1-\sigma} \cdot \widehat{\text{TMA}}_j^{\frac{1-\sigma}{\sigma}} \cdot \widehat{H}_j^{\frac{\sigma-1}{\sigma}}$$

$$\widehat{\text{MA}}_i = \sum_j \pi_{ij} \cdot \left(\widehat{\tau}_{ij}^m\right)^{-\theta} \cdot \widehat{\text{TMA}}_j^{\delta} \cdot \widehat{H}_j^{-\frac{\theta}{\sigma}}$$

Where $\lambda = \frac{\eta-\theta(1-\eta)}{\theta(1-\eta)}$ and $\delta = \frac{\theta(2\sigma-1)}{\sigma(\sigma-1)}$. The quantities $\omega_{ij}, \rho_{ij}, \pi_{ij}$ denote the initial shares of human capital, trade, and migration, which can be calculated. Given counterfactual changes in the road network and parameter estimates, this system can therefore be solved for the endogenous variables, including the counterfactual change in local opportunity $\hat{\mu}_i$. In the appendix section M, I derive conditions for the existence and uniqueness of the equilibrium of this system, following Allen et al. (2024). The parameter estimates used in the counterfactual exercise satisfy these conditions.

5.2 Identification of model parameters

We need to identify parameters η, σ, θ and trade and migration iceberg costs $T^t_{ijt} = (\tau^t_{ijt})^{\sigma-1}, T^m_{ijt} = (\tau^m_{ijt})^{\theta}$. In section 4 the quantity T^m_{ijt} was estimated, and a bundle of η and η was identified. Therefore, it only remains to identify η, σ, and θ separately as well as T^t_{ijt}.

5.2.1 Estimating trade costs T^t_{ijt}

As with T^m_{ijt} I can parameterise trade costs such that $T^t_{ijt} = (\tau^t_{ijt})^{\sigma-1} = t^{\tilde{\sigma}-1}_{ijt}$ where $\tilde{\sigma} = a \cdot \sigma$. Although data on migration is remarkably rich, covering a large time span and granular geography, allowing estimation of T^m_{ijt} with confidence — no analogous data exists for trade across space in this setting. As a result of this, I am forced to calibrate σ from the literature. I follow Morten and Oliveira (2024), who leverage variation in trade across Brazilian states. Morten and Oliveira (2024) calculates the elasticity of trade and migration costs to travel time. I take their estimated elasticity of trade to travel time (1.96) and rescale it by the ratio of my estimated migration costs to theirs (0.58) to account for differences in geography and environment. Following this method I find $\tilde{\sigma} = 1 + 1.96 \times 0.58 = 2.14$. Calibrating σ from Simonovska and Waugh (2014) to be $\sigma = 5$, I back out $a = 0.43$ and an elasticity of trade costs to travel times: $(-1.96 \times 0.58)/(1 - 5) = 0.28$.

5.2.2 Separately identifying θ and η

Following the strategy employed by Bryan and Morten (2019), I use model-implied regressions to separately identify θ and η. Note that the expected wage individuals from i who move to j attain is given by $\mathbb{E}[w_{jt}|i] = \gamma r_{jt} h_{it} \pi_{ijt}^{-1/\theta}$ an expression that motivates the following regression.

$$\ln(\bar{w}_{ijt}) = \alpha_{it} + \gamma_{jt} - \frac{1}{\theta} \ln(\pi_{ijt}) + v_{ijt} \tag{11}$$

This regression will therefore allow me to find estimates for θ which, coupled with our previous estimates of $\frac{1}{\theta(1-\eta)}$, will allow me to estimate η. However, estimates from equation 11 may be biased. As discussed in Bryan et al. (2014), any shock that affects the wages that individuals from i receive in j will also impact migration from i to j, for example, if demand for skills generated in i increases in j. To overcome this I use the instrumental variable approach suggested by Bryan and Morten (2019) and instrument $\ln(\pi_{ijt})$ with $\ln(\pi_{\neg ijt})$, where $\pi_{\neg ijt}$ is the proportion of people from other origins who migrate to j in period t. To take this strategy to the data I require information on local wages. This is found using DHS data on incomes and adopting an Engel curve approach following Young (2012), details can be found in the online appendix section F.

Table 3 shows the results from estimating equation 11 with various specifications. Table 3 also shows the implied values of $\widehat{\theta}$ and $\widehat{\eta}$, combining information with that from estimating equation 10 with corresponding standard errors calculated using the Delta method. Column two is our baseline specification where $\ln(\pi_{ijt})$ is instrumented as discussed above. I find the following parameter estimates $\theta = 22.64$ (2.214), $\eta = 0.434$ (0.170), $\sigma = 5$. This compares to Hsiao (2024) who finds $\theta = 20, \eta = 0.224$, and Bryan and Morten (2019) who find $\theta = 28$.

Table 3 Identifying θ and η

	(1)	(2)	(3)	(4)
$\ln(\pi)$	-0.0411***	-0.0442***	-0.0447***	-0.0473***
	(0.00356)	(0.00432)	(0.00425)	(0.00432)
θ	24.30	22.64	22.36	21.15
	(2.11)	(2.21)	(2.12)	(1.93)
η	0.47	0.43	0.43	0.39
	(0.16)	(0.17)	(0.17)	(0.18)
IV		X	X	X
F-stat		3247	3379	3349
FE	ot,dt	ot,dt	o,d,t	o,d
Obs	13999	13999	13999	13999

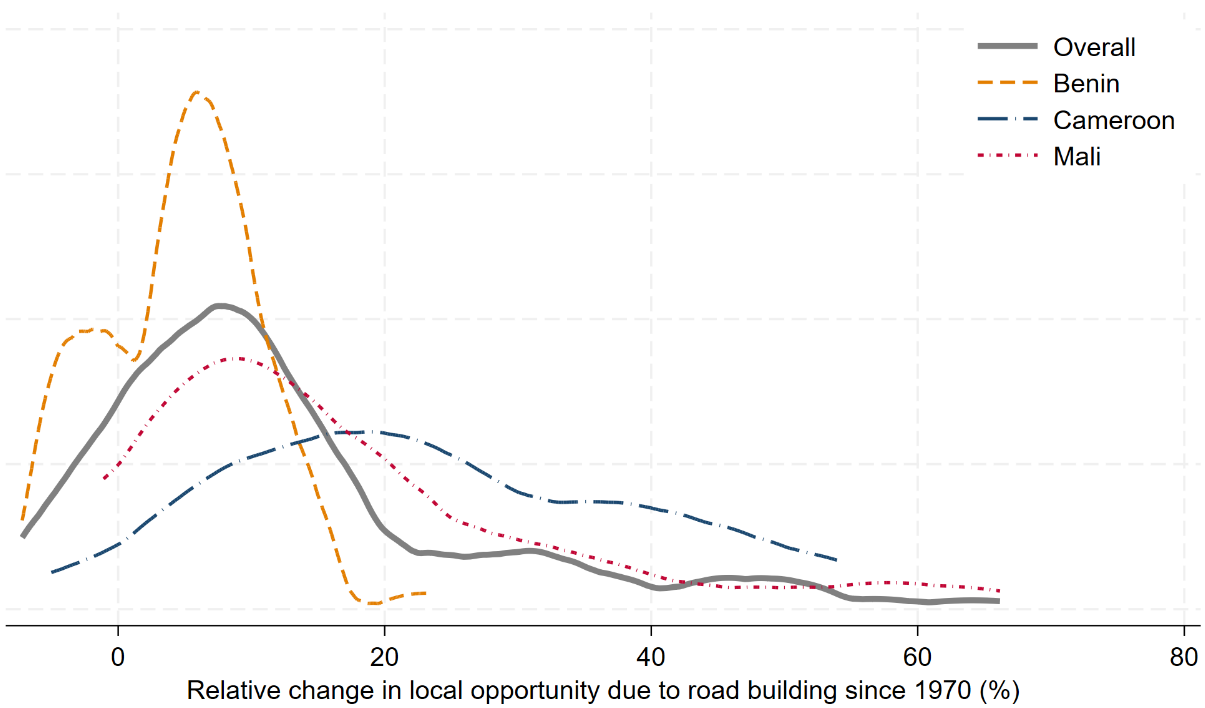
Notes: This table shows the results from estimating equation 11 and the implied values of $\widehat{\theta}$ and $\widehat{\eta}$ combining information with that from estimating equation 10 with corresponding standard errors calculated using the Delta method. Column (1) shows results with no instrumenting, including origin-time and destination-time fixed effects. Columns (2) to (4) show results instrumenting following Bryan and Morten (2019) with various fixed effects specifications. Column (2) includes origin-time and destination-time fixed effects and is our baseline specification, column (3) includes origin, destination, and time fixed effects, and finally column (4) only includes origin and destination fixed effects.

5.3 Quantification results

Using the full structure of the model, I can quantify the causal effect of road building between 1970 and 2020 on the distribution of local opportunity. I find that, on average, road building increased local opportunity. Due to road building since 1970, the causal effect of place on primary completion rates increased by, on average, 12.5%. This corresponds to a 0.24 percentage point higher average annualised growth rate. Over the same period across Benin, Cameroon, and Mali, the average annualised growth rate in primary education completion was 1.82 percent.[10]

The average hides considerable variation within-country across locations, with 15% of locations seeing aggregate negative effects. Figure 5 shows graphically how this average effect varies across locations. In each country, there is significant variation in the impact of road building since 1970 on local opportunity. In Cameroon, the distribution is relatively flat with a high average of 23%. Only three locations (8%) in Cameroon experience negative effects, whereas ten locations (26%) experience effects in excess of a 33% increase. Benin and Mali, on the other hand, have narrower distributions with means of 4.7% and 16.8%, respectively. In Benin, 19 locations (25%) saw negative impacts, whereas in Mali, 3 locations (6%) did. In Benin, the largest positive impact was an increase of 23%, whereas in Mali, the right tail extends to a 66% increase, and 17 locations (36%) experience an increase in opportunity of at least 33%.

Figure 5 Aggregate effects by country



Notes: This figure plots the empirical distribution over locations (weighted evenly) of the aggregate effect of road building since 1970 on local opportunity. In solid grey I plot the average over all countries, in dashed orange for Benin, in dot-long-dashed blue for Cameroon, and in dot-short-dashed red for Mali.

[10]This average weights each location equally. Data from the World Bank available https://genderdata.worldbank.org/en/indicator/se-prm-cmpt-zs?view=trend&geos=WLD_BEN_CMR_MLI.

5.4 Mechanisms: Moving to Opportunity or Moving Opportunity

Changes to the road network alter the spatial distribution of local opportunity through two main channels. First, opportunity will increase because it is now easier for individuals to move to areas with pre-existing high opportunity. Second, the spatial distribution of opportunity itself may change. That is, roads both allow moves to opportunity, and move opportunity. I can leverage the structure of the model to decompose aggregate effects into these two channels by shutting down endogenous equilibrium responses and recalculating the effect of road building since 1970.

The direct effect of removing frictions to moving to opportunity will be positive whenever connectivity is improved. Intuitively, this is because reducing spatial frictions increases the set of locations it may be worth moving to, conditional on any given set of preference draws — thus allowing weakly better moves. Allowing for economic activity to endogenously react will, on average, increase the benefits of road building, as reducing frictions allows a more efficient spatial allocation. However, it need not be the case that all locations gain. Activity will concentrate in some locations, and others will be left behind, potentially generating losers. The overall impact of road building on any given location will be a combination of these two effects. It is possible that the negative impact of opportunity moving away outweighs the positive impact of allowing greater moves to opportunity, generating absolute losers. In this manner, road building may not be Pareto improving.

Formally, the impact of changes in connectivity on local opportunity can be decomposed by appealing to equation 7 which states that the local proportional change in opportunity is equal to $\hat{e}_i = \widehat{\mathrm{MA}}_i^{\frac{1}{\theta(1-\eta)}}$ for location i. Using our derived expression for market access, $\widehat{\mathrm{MA}}_i = \sum_j \pi_{ij} \cdot \left(\widehat{\tau}_{ij}^m\right)^{-\theta} \cdot \widehat{\mathrm{TMA}}_j^{\delta} \cdot \widehat{H}_j^{-\frac{\theta}{\sigma}}$, the impact of changes in the road network on market access can therefore be split into the two channels.

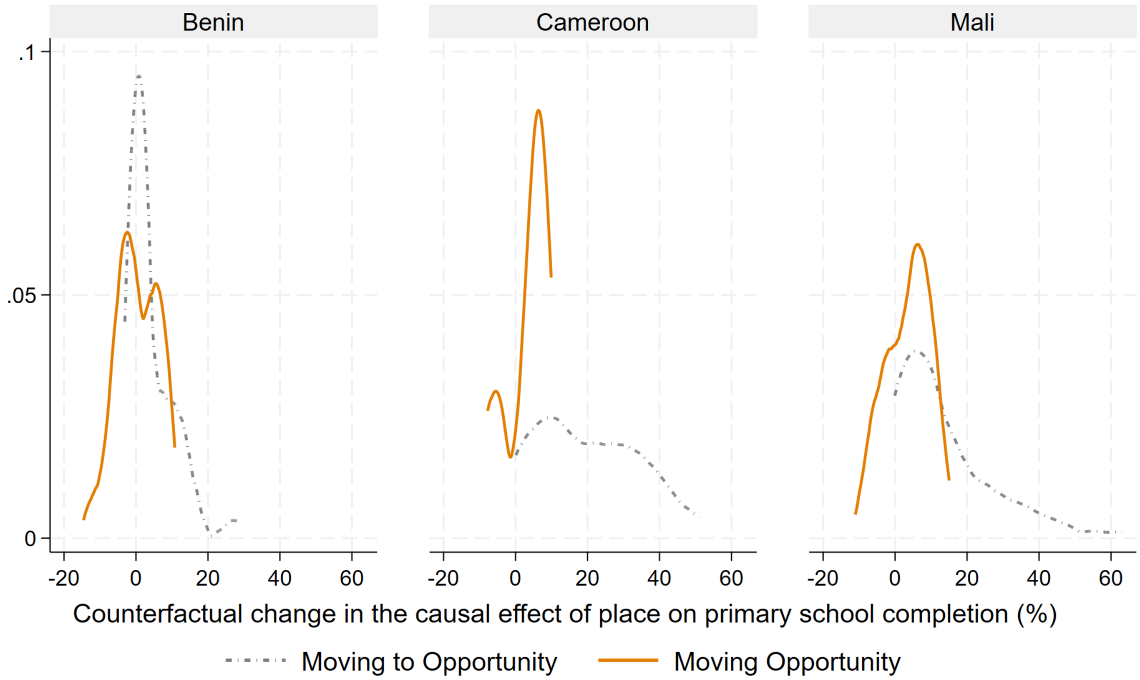
$$\begin{aligned}\hat{e}_i^{\theta(1-\eta)} - 1 = &\underbrace{\sum_j \pi_{ij}\left((\hat{\tau}_{ij}^m)^{-\theta} - 1\right)}_{\text{Moving to Opportunity}} \\ &+ \underbrace{\sum_j \pi_{ij}\left(\left(\widehat{\mathrm{TMA}}_j^{\delta} \cdot \widehat{H}_j^{-\frac{\theta}{\sigma}}\right) - 1\right) + \sum_j \pi_{ij}\left((\hat{\tau}_{ij}^m)^{-\theta} - 1\right)\left(\left(\widehat{\mathrm{TMA}}_j^{\delta} \cdot \widehat{H}_j^{-\frac{\theta}{\sigma}}\right) - 1\right)}_{\text{Moving Opportunity}}\end{aligned}$$

I can then identify the moving to opportunity (M2O) channel by shutting down equilibrium responses and thus setting $\widehat{\mathrm{TMA}}_j = \widehat{H}_j = 1$ and performing the following transformation $\mathrm{M2O}_i = \left(\sum_j \pi_{ij}\left((\hat{\tau}_{ij}^m)^{-\theta} - 1\right) + 1\right)^{1/\theta(1-\eta)}$. With this in hand, the moving opportunity channel can be identified as $\mathrm{MO}_i = \hat{e}_i - \mathrm{M2O}_i$.

Figure 6 shows the distribution across locations of each channel in Benin, Cameroon, and Mali. The results confirm the intuition. The direct, moving to opportunity, effect is positive as long as roads improve, but heterogeneous across locations. Areas that saw large improvements in their connectivity, meaning that spatial frictions between them and attractive locations reduced considerably, saw the largest increases in opportunity. In all countries, some locations were almost entirely unaffected by direct improvements in their connectivity. This doesn't necessarily mean that road building didn't change the travel time from such a location to other places, but that it didn't meaningfully change the travel time to locations that one would want to travel to from such a location. For example, if location i is unambiguously the most attractive location, then irrespective of road building, the direct effect will always be insignificant.

The distribution of the equilibrium, moving opportunity, effect is shown in orange in figure 6. This effect is less heterogeneous than the direct effect as it is smoothed over space. As expected it is on average positive in each country with a mean of 0.34 in Benin, 3.34 in Cameroon, and 3.61 in Mali. However, in each country, spatial equilibrium effects are negative in multiple locations. The extent to which negative effects are generated differs across countries. In Benin and to a lesser extent Mali, the equilibrium adjustment is significantly more negative than in Cameroon.

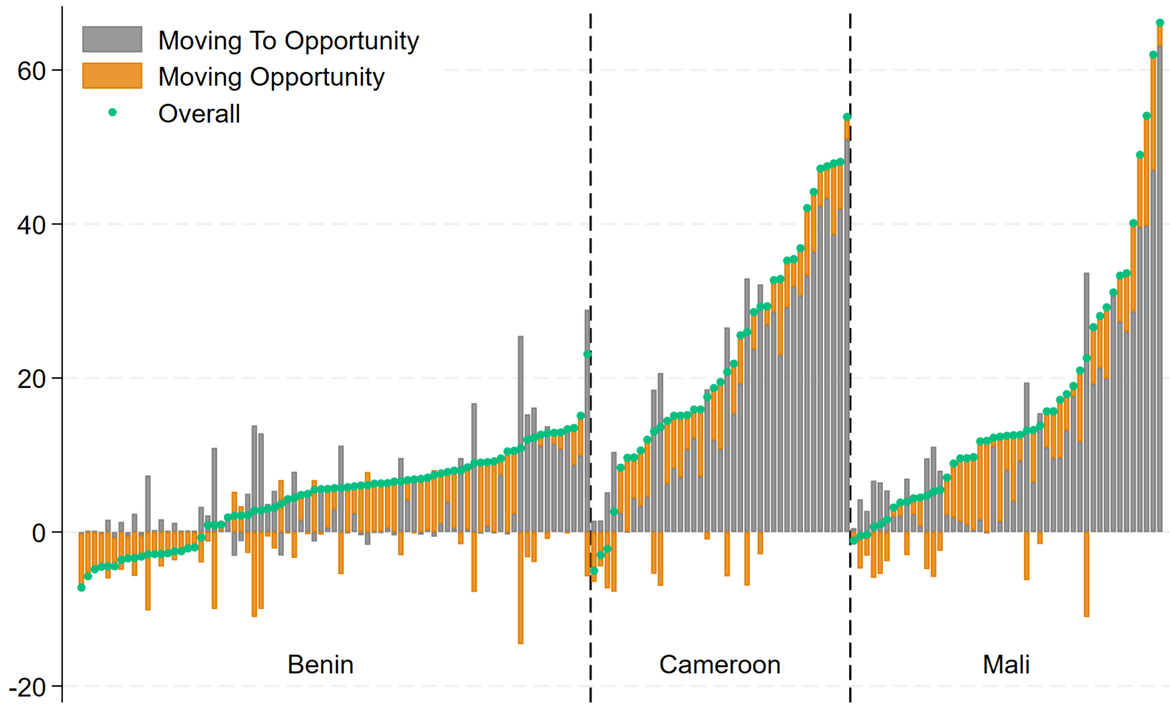
Figure 6 Mechanisms decomposition



Notes: This figure shows the distribution over locations of each constituent channel of the overall change in local opportunity. In orange, I show the indirect spatial equilibrium adjustments. In dot-dash gray, I show the direct moving to opportunity effect. The left-most panel shows the distribution for Benin, the middle panel for Cameroon, and the right-most panel for Mali. Effects are reported in percentage change.

Figures 7 and 8 turn to the location-level decomposition. Figure 7 shows a stacked bar graph decomposing each location-level aggregate effect (in green) into that due to moving to opportunity (gray bar) and that due to moving opportunity (orange bar). This figure highlights how equilibrium responses can be negative at the individual-location level, and even sufficiently so to cause locations to lose overall. It also highlights how the size of the equilibrium response is independent of the direct effect (with a correlation of 0.06), and that its magnitude differs across countries. In Benin, the average equilibrium response in absolute terms (4.8) is larger than the average direct effect (4.7), whereas in Cameroon and Mali it is much smaller.

Figure 7 Mechanisms decomposition

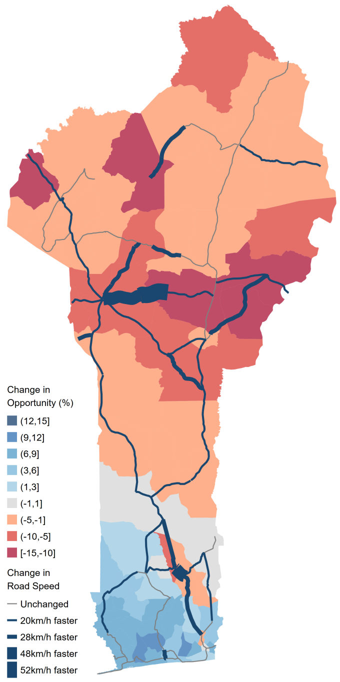


Notes: This figure shows a stacked bar graph over locations in Benin, Cameroon, and Mali. In green markers, I show the aggregate impact of road building since 1970. In gray bars, I show the direct effect of enabling moves to opportunity, and in orange bars, the indirect spatial equilibrium effect of opportunity moving. Effects are reported in percentage change.

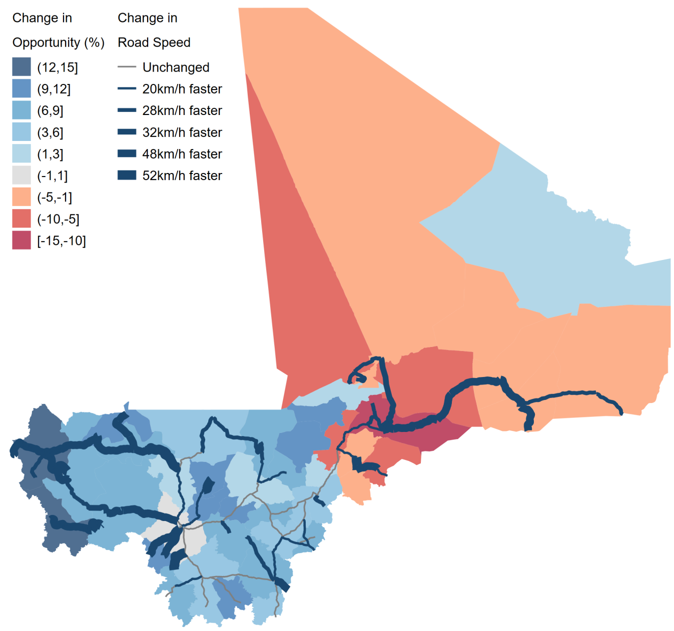
Finally, Figure 8 maps the geography of the isolated spatial equilibrium response of opportunity moving. These maps allow us to discern the geography of winners and losers from the equilibrium responses. In each country, the relatively more remote areas were more likely to see negative equilibrium responses. In Benin, this is particularly stark, with only the relatively richer and more densely populated coastal south of the country seeing gains. Figures B.2, B.3, and B.4 in the appendix similarly map the overall effects and the direct effects in each country.

Figure 8 Mapping the isolated GE effect of opportunity moving

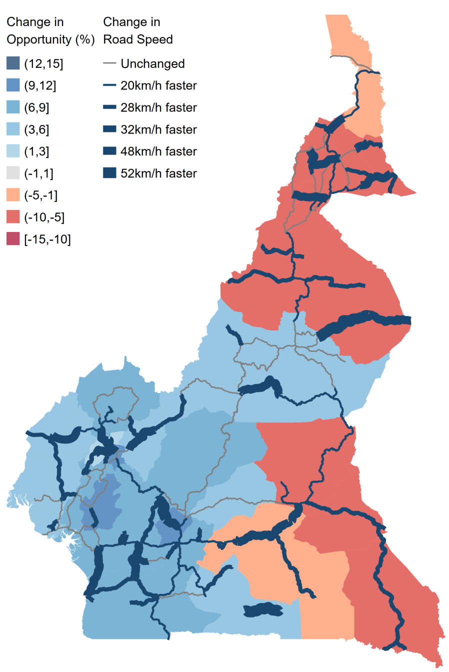
(a) Benin



(b) Mali



(c) Cameroon



Notes: These figures map the geographic distribution of the spatial equilibrium effects of opportunity moving in each of Benin, Cameroon, and Mali in response to road building in each country since 1970. Warmer colors indicate larger losses, cooler colors indicate larger gains. The change in roads between 1970 and 2020 is indicated on each map by solid dark blue lines of varying thickness. Thicker lines indicate larger increases in travel time. Effects are measured in percentage change.

5.4.1 Decomposing spatial equilibrium effects into labor supply and demand channels

The second channel, of spatial equilibrium adjustment, can be further decomposed into the endogenous reaction of the population and the endogenous reaction of goods trade. Endogenous migration will be negative, as more immigration to high-opportunity locations causes more competition for such opportunities. The endogenous reaction of demand, however, is likely to be positive as the same influx of individuals increases demand for goods and so the opportunities available. To do this I again appeal to equation 7 which can be split into four components as shown in the expression below.

$$
\begin{aligned}
\hat{e}_i^{\theta(1-\eta)} - 1 = & \underbrace{\sum_j \pi_{ij} \left((\hat{\tau}_{ij}^m)^{-\theta} - 1 \right)}_{\text{Direct}} \\
& + \underbrace{\sum_j \pi_{ij} \left(\widehat{H}_j^{-\frac{\theta}{\sigma}} - 1 \right) + \sum_j \pi_{ij} \left((\hat{\tau}_{ij}^m)^{-\theta} - 1 \right) \left(\widehat{H}_j^{-\frac{\theta}{\sigma}} - 1 \right)}_{\text{Labor Supply}} \\
& + \underbrace{\sum_j \pi_{ij} \left(\widehat{\mathrm{TMA}}_j^{\delta} - 1 \right) + \sum_j \pi_{ij} \left((\hat{\tau}_{ij}^m)^{-\theta} - 1 \right) \left(\widehat{\mathrm{TMA}}_j^{\delta} - 1 \right)}_{\text{Labor Demand}} \\
& + \underbrace{\sum_j \pi_{ij} \left(\widehat{\mathrm{TMA}}_j^{\delta} - 1 \right) \left(\widehat{H}_j^{-\frac{\theta}{\sigma}} - 1 \right) + \sum_j \pi_{ij} \left((\hat{\tau}_{ij}^m)^{-\theta} - 1 \right) \left(\widehat{\mathrm{TMA}}_j^{\delta} - 1 \right) \left(\widehat{H}_j^{-\frac{\theta}{\sigma}} - 1 \right)}_{\text{Labor Supply - Labor Demand interaction}}
\end{aligned}
$$

However, unlike the simple two-channel decomposition, these individual components can't be identified because $\widehat{H}_j$ and $\widehat{\mathrm{TMA}}_j$ depend on each other. Denote each component conditional on the other being fixed as follows $\widetilde{\widehat{H}}_j = \widehat{H}_j \mid \widehat{\mathrm{TMA}}_j = 1$, and $\widetilde{\widehat{\mathrm{TMA}}}_j = \widehat{\mathrm{TMA}}_j \mid \widehat{H}_j = 1$ and note that $\widetilde{\widehat{H}}_j \neq \widehat{H}_j$ and $\widetilde{\widehat{\mathrm{TMA}}}_j \neq \widehat{\mathrm{TMA}}_j$. Using this notation, I can identify four channels by iteratively shutting down dimensions of endogenous response.

Firstly, I can identify the direct labour supply effect by setting $\widehat{\mathrm{TMA}}_j = 1$ and recalculating the counterfactual. The difference between this estimated counterfactual and that with no equilibrium response is denoted by $\mathrm{DirectLS}_i$. By endogenising labor supply, I allow local opportunity to endogenously adjust to induced migration flows. Intuitively, higher inflows correspond to a local labour supply shock, which will crowd out opportunities by putting downward pressure on wages. This will cause wages to decrease in those areas where inflows are largest, which will be exactly those areas with preexisting higher opportunity. Conversely, local wages will rise in areas with preexisting low opportunities. However, the highest-value location is the only relevant location (conditional on individual location-specific skill shocks) when choosing where to migrate, so although this force may appear to be equilibrating, it will necessarily reduce opportunities available from each origin location. When combined with the direct effect, these two channels give the impact of road building on opportunities in frameworks that allow endogenous labour supply to impact local outcomes. The combined impact will be much more negative than the direct effect, and will normally imply that road building decreased opportunity in a given location.

Secondly, I can identify the *direct* labor demand effect by taking the analogous approach of setting $\widehat{H}_j = 1$, and calculating DirectLD_i. By endogenising labor demand, I allow local opportunity to adjust endogenously to changes in demand for labour caused by changes in the spatial distribution of economic activity. Reductions in transport costs cause trade costs to fall, altering spatial patterns of trade. On average, as costs have fallen, trade volumes will increase, causing wages to rise. However, prices will also adjust, and trade may be diverted from some locations that see a relatively shallower decline in transport costs; therefore, some locations may lose from this effect.

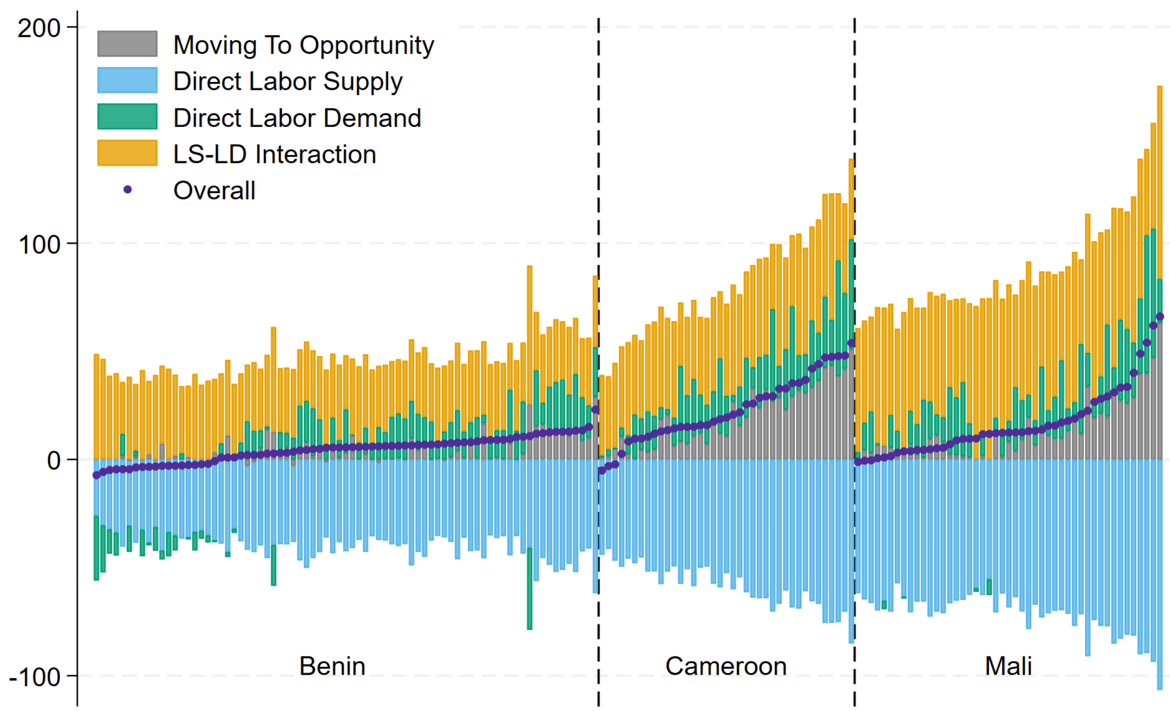
Thirdly, I can identify the labor supply-labor demand interaction effect that is only present when both margins are allowed to adjust endogenously. This effect will capture both the interaction term in the decomposition of the overall effects, and how the differences between $\widehat{H}_j$ and $\widetilde{\widehat{H}}_j$, and $\widehat{\text{TMA}}_j$ and $\widetilde{\widehat{\text{TMA}}}_j$, propagate through each channel. This effect can be identified as $\text{Int}_i = \text{MO}_i - \text{DirectLS}_i - \text{DirectLD}_i$. This effect captures the potentially complex interaction between endogenous migration and trade. In general, locations that received more population inflows will now see labour demand, and so wages rise. For this reason, the interaction effect is expected to be positive. Note that, in the case of autarky, the labour demand and labour supply effects will directly cancel, and the overall impact of roads will be equal to the direct effect.[11] Not allowing labour demand to endogenously respond would likely cause researchers to significantly underestimate the positive impact of road building on local opportunities, both through the direct and interaction effects.

Figure 9 decomposes the total effect, denoted as purple markers, into each of these channels. In gray, I plot the direct effect of decreasing migration costs and allowing more moves to opportunity as before. In blue, I then isolate the direct labour supply effect. In line with our intuition, this is resoundingly negative as the previously high-opportunity locations see their wages fall. In green, I isolate the direct labour demand effect. This is, in the main, positive, but it will induce a reshuffling of demand that causes some locations to be negatively affected. Finally, in orange, I show the supply-demand interaction effect that is

[11] To see this note that $\widehat{\text{TMA}}_i = \sum_j \rho_{ij}(\hat{\tau}^t_{ij})^{1-\sigma}\widehat{\text{TMA}}_j^{\frac{1-\sigma}{\sigma}}\widehat{H}_j^{\frac{\sigma-1}{\sigma}}$, and therefore as under trade-Autarky $\rho_{ii} = 1,\ \rho_{ij} = 0\ \forall j \neq i$ this implies that $\widehat{\text{TMA}}_i = \widehat{H}_i^{\frac{\sigma-1}{2\sigma-1}}$. Substituting this into our expression for market access, I find that $\widehat{\text{MA}}_i = \sum_j \pi_{ij}(\hat{\tau}^m_{ij})^{-\theta}\left(\widehat{H}_j^{\frac{\sigma-1}{2\sigma-1}}\right)^{\delta}\widehat{H}_j^{\frac{-\theta}{\sigma}} = \sum_j \pi_{ij}(\hat{\tau}^m_{ij})^{-\theta}$.

only present in models that allow both margins to adjust. This is also positive, as areas that saw larger increases in labour supply see the largest increases in demand. Figure 9 shows that at the locality level, not accounting for endogenous responses would result in incorrect estimates of the impact of roads, and only accounting for endogenous labour supply or labour demand individually would also result in biased average estimates. However, the degree of bias varies across countries; in Cameroon, the direct effect is a fairly good approximation of the overall impact, whereas in Benin and Mali it is less so.

Figure 9 Mechanisms decomposition



Notes: This figure decomposes the full-equilibrium location-level effect of road building since 1970 on opportunity into the constituent components of the market-access decomposition. For each location, the purple marker reports the total effect from the full model (with endogenous migration and trade). Gray bars report the direct effect of lower migration costs computed by shutting down endogenous migration and trade responses (setting $\widehat{H}_i = 1$ and $\widehat{\mathrm{TMA}}_i = 1$). Blue bars report the direct labour-supply channel computed by allowing endogenous migration but holding trade fixed (setting $\widehat{\mathrm{TMA}}_i = 1$) and taking away from this the direct effect. Green bars report the direct labour-demand channel computed by allowing trade to adjust while holding migration fixed (setting $\widehat{H}_i = 1$) and taking away the direct effect. Orange bars report the interaction term that arises only when both margins adjust, defined as the residual required for the sum of components to equal the full-model total. All effects are expressed as percent changes in opportunity.

Tables 4, 5, and 6 explore heterogeneity in effects in more detail; they detail heterogeneity across each component to initial market access, initial population, and the size of the direct effect, respectively. Each table shows heterogeneity across the following components in corresponding columns: (1) the overall effect with endogenous labour supply and labour demand, (2) isolated direct effect (not present in table 6), (3) isolated direct labour supply effect, (4) isolated labour demand effect, (5) the interaction between the endogenous labour supply and labour demand effects, and finally (6) the overall general equilibrium effects. In row one each table shows the aggregate results over all countries (including country fixed effects with robust standard errors), and in rows two, three, and four, each table shows the effects for each country individually.

Table 4 shows heterogeneity with respect to (log) initial market access, measured in the first year in which population data is available. Overall, in column one, I find that locations with higher initial market access saw larger increases in opportunity due to road building since 1970, although this effect is not statistically significant in Mali. This average positive impact masks heterogeneity by component. The direct impact doesn't systematically vary by initial market access, whereas the endogenous response of labour supply is significantly more negative in areas with preexisting high market access. This is intuitive; high market access areas are likely to be those with higher real wages and therefore those that attract an influx of people when migration costs fall, which in turn pushes down wages. Conversely, the direct labour demand effect is increasing in initial market access, although this is only true in Benin and Cameroon. Column six shows the impact of log initial market access on the overall GE effect, which is positive.

Table 4 Heterogeneity by initial market access

	(1) Overall	(2) Direct	(3) Labor Supply	(4) Labor Demand	(5) LS-LD interaction	(6) GE Overall
Log Market Access	12.22***	2.711	-7.287***	19.97***	-3.178	9.508***
	(2.702)	(3.219)	(1.986)	(2.743)	(2.752)	(1.089)
Benin	10.10***	-1.362	-5.999***	27.80***	-10.34***	11.46***
	(1.312)	(2.107)	(1.591)	(2.598)	(1.988)	(1.245)
Cameroon	21.14***	8.255	-12.49**	32.37***	-6.988*	12.89***
	(7.717)	(8.197)	(5.118)	(5.494)	(3.900)	(2.070)
Mali	12.15	7.965	-7.349	-0.810	12.35**	4.188**
	(8.000)	(8.366)	(5.579)	(5.152)	(4.864)	(2.065)
Observations	163	163	163	163	163	163

Notes: This table reports heterogeneity in the model-implied percentage change in opportunity due to road building since 1970 as a function of initial market access (measured in the first year for which population data are available in each country). Each column corresponds to a component of the decomposition: (1) the total full-equilibrium effect; (2) the direct effect; (3) the labour-supply channel net of the direct effect; (4) the labour-demand channel net of the direct effect; (5) the interaction between endogenous labour supply and labour demand, net of the direct effect and each channel in isolation; and (6) the total general-equilibrium adjustment. Row 1 pools all countries and includes country fixed effects. Subsequent rows report country-specific estimates. The unit of observation is a location; all locations are weighted equally, and robust standard errors are reported in parentheses.

Table 5 then shows heterogeneity in the impact of road building since 1970 on opportunity by initial population, again measured in the first available year in each country. Column one shows that overall areas with higher population saw larger gains, and this is true in all three countries, although not statistically significant in Cameroon. This aggregate positive effect is, however, driven by different channels in each setting. In Benin, areas with a higher initial population saw a larger labour demand effect, mitigated but not nullified by the endogenous migration reaction of individuals. In Cameroon, the positive reaction comes mainly from the direct effect; areas with higher initial population are those areas where migration costs to higher-opportunity locations fell the most. Finally, in Mali, the correlation is driven by the direct effect as in Cameroon. However, in Mali, this direct effect is mitigated by the endogenous migration response of individuals, which itself is attenuated by the reaction of firms.

Table 5 Heterogeneity by initial population

	(1) Overall	(2) Direct	(3) Labor Supply	(4) Labor Demand	(5) LS-LD interaction	(6) GE Overall
Log Population	6.019***	5.812***	-4.335***	2.853	1.689	0.207
	(2.294)	(2.054)	(1.517)	(1.929)	(1.508)	(0.877)
Benin	3.866***	1.227	-1.972	8.140**	-3.529**	2.639**
	(1.354)	(1.586)	(1.576)	(3.174)	(1.366)	(1.328)
Cameroon	2.950	4.637	-2.062	0.498	-0.123	-1.687
	(3.949)	(3.225)	(2.264)	(3.096)	(1.464)	(1.545)
Mali	11.92**	12.06***	-9.519***	0.0270	9.355***	-0.137
	(4.661)	(4.245)	(2.837)	(3.370)	(2.643)	(1.373)
Observations	163	163	163	163	163	163

Notes: This table reports heterogeneity in the model-implied percentage change in opportunity due to road building since 1970 as a function of initial population (measured in the first year for which population data are available in each country). Each column corresponds to a component of the decomposition: (1) the total full-equilibrium effect; (2) the direct effect; (3) the labour-supply channel net of the direct effect; (4) the labour-demand channel net of the direct effect; (5) the interaction between endogenous labour supply and labour demand, net of the direct effect and each channel in isolation; and (6) the total general-equilibrium adjustment. Row 1 pools all countries and includes country fixed effects. Subsequent rows report country-specific estimates. The unit of observation is a location j; all locations are weighted equally, and robust standard errors are reported in parentheses.

Finally, in table 6, I consider how the overall effect and each component are correlated with the direct effect. Unsurprisingly, column one shows a positive correlation on average and across countries. By comparing coefficients across countries, it is clear that the direct effects are very predictive of the overall impact on average in Cameroon and Mali, but much less so in Benin. Column two shows that areas with the largest direct effect saw the largest negative impacts due to the endogenous reaction of labour supply — this is intuitive as the most positively affected areas see the largest in-migration and therefore negative impact on local wages. Analogously, the labour demand effect is more positive in these areas as their better connectedness brings in more trade demand.

Table 6 Heterogeneity by direct effect

	(1) Overall	(2) Labor Supply	(3) Labor Demand	(4) LS-LD interaction	(5) GE Overall
Direct Effect	0.965***	-0.707***	0.454***	0.217**	-0.0351
	(0.0482)	(0.0264)	(0.125)	(0.0899)	(0.0482)
Benin	0.534***	-0.755***	0.173	0.115	-0.466***
	(0.0894)	(0.129)	(0.411)	(0.226)	(0.0894)
Cameroon	1.014***	-0.682***	0.486***	0.210**	0.0142
	(0.0586)	(0.0398)	(0.130)	(0.0860)	(0.0586)
Mali	1.068***	-0.711***	0.522**	0.257	0.0675
	(0.0726)	(0.0235)	(0.208)	(0.163)	(0.0726)
Observations	163	163	163	163	163

Notes: This table reports heterogeneity in the model-implied percentage change in opportunity due to road building since 1970 as a function of the direct effect of road building, calculated by shutting down the endogenous labour supply and demand responses. Each column corresponds to a component of the decomposition: (1) the total full-equilibrium effect; (2) the labour-supply channel net of the direct effect; (3) the labour-demand channel net of the direct effect; (4) the interaction between endogenous labour supply and labour demand, net of the direct effect and each channel in isolation; and (5) the total general-equilibrium adjustment. Row 1 pools all countries and includes country fixed effects. Subsequent rows report country-specific estimates. The unit of observation is a location; all locations are weighted equally, and robust standard errors are reported in parentheses.

6 Conclusion

Opportunity is unevenly distributed across space. This paper asks whether investments in connectivity can change this geography of opportunity, defined as the causal effect of place on primary education completion. Combining a quantitative spatial economics model with a novel identification strategy and rich data from Benin, Cameroon, and Mali, I find that road building does alter local opportunity — both by encouraging moves to opportunity and by altering the underlying distribution of opportunity itself.

Road building since 1970 increased local opportunity by 12.5% on average, but this average hides considerable heterogeneity: 15% of locations experienced negative impacts. The direct effect of easing moves to opportunity explains 84% of the average impact and is positive everywhere. However, endogenous labour supply and demand responses are crucial to understanding location-specific heterogeneity and are the source of negative effects. That these results hold across three countries with distinct road networks, institutional environments, and supranational frameworks strengthens confidence in their generality.

Three implications follow. First, road building is a quantitatively important policy lever for shifting educational opportunity — it is not merely a trade or urbanisation tool. Second, distributional consequences are central. Improvements in one corridor affect the entire system through the endogenous adjustment of labour supply and demand, creating local winners and losers beyond those directly connected. Infrastructure investments should be evaluated as a system rather than project-by-project, since spatial equilibrium responses mean that the effect of any given road depends on the entire network. Third, as a result of this endogenous reshuffling, there can be absolute opportunity losers from road building. Reducing spatial frictions is not Pareto-improving, and road building is not a silver bullet.

References

Adamopoulos, Tasso (2025). "Spatial integration and agricultural productivity: Quantifying the impact of new roads". In: *American Economic Journal: Macroeconomics* 17.1, pp. 343–378.

Adukia, Anjali, Sam Asher, and Paul Novosad (2020). "Educational investment responses to economic opportunity: evidence from Indian road construction". In: *American Economic Journal: Applied Economics* 12.1, pp. 348–76.

Alesina, Alberto et al. (2021). "Intergenerational mobility in Africa". In: *Econometrica* 89.1, pp. 1–35.

Allen, Treb, Costas Arkolakis, and Xiangliang Li (2024). "On the equilibrium properties of spatial models". In: *American Economic Review: Insights* 6.4, pp. 472–489.

Allen, Treb, Costas Arkolakis, and Yuta Takahashi (2020). "Universal gravity". In: *Journal of Political Economy* 128.2, pp. 393–433.

Alpern, Stanley B. (1999). "Dahomey's Royal Road". In: *History in Africa* 26, pp. 11–24. DOI: 10.2307/3172135.

Asher, Sam and Paul Novosad (2020). "Rural roads and local economic development". In: *American economic review* 110.3, pp. 797–823.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian (2020). "On the road: Access to transportation infrastructure and economic growth in China". In: *Journal of Development Economics* 145, p. 102442.

Baum-Snow, Nathaniel (2007). "Did highways cause suburbanization?" In: *The quarterly journal of economics* 122.2, pp. 775–805.

Baum-Snow, Nathaniel et al. (2017). "Roads, railroads, and decentralization of Chinese cities". In: *Review of Economics and Statistics* 99.3, pp. 435–448.

Borusyak, Kirill and Peter Hull (2023). "Nonrandom exposure to exogenous shocks". In: *Econometrica* 91.6, pp. 2155–2185.

Bryan, Gharad, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2014). "Underinvestment in a profitable technology: The case of seasonal migration in Bangladesh". In: *Econometrica* 82.5, pp. 1671–1748.

Bryan, Gharad, Edward Glaeser, and Nick Tsivanidis (2020). “Cities in the developing world”. In: *Annual Review of Economics* 12.1, pp. 273–297.

Bryan, Gharad and Melanie Morten (2019). “The aggregate productivity effects of internal migration: Evidence from Indonesia”. In: *Journal of Political Economy* 127.5, pp. 2229–2268.

Burgess, Robin et al. (2015). “The value of democracy: evidence from road building in Kenya”. In: *American Economic Review* 105.6, pp. 1817–51.

Chetty, Raj and Nathaniel Hendren (2018a). “The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects”. In: *The Quarterly Journal of Economics* 133.3, pp. 1107–1162.

Chetty, Raj and Nathaniel Hendren (2018b). “The impacts of neighborhoods on intergenerational mobility II: County-level estimates”. In: *The Quarterly Journal of Economics* 133.3, pp. 1163–1228.

Chetty, Raj, Nathaniel Hendren, and Lawrence F Katz (2016). “The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment”. In: *American Economic Review* 106.4, pp. 855–902.

Chiovelli, Giorgio et al. (2025). *Civil War-Induced Displacement and Human Capital*. Tech. rep. Facultad de Ciencias Empresariales y Economia. Universidad de Montevideo.

Chyn, Eric and Diego Daruich (2025). “An equilibrium analysis of the effects of neighborhood-based interventions on children”. In: *American Economic Review* 115.12, pp. 4476–4522.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum (2008). “Global rebalancing with gravity: Measuring the burden of adjustment”. In: *IMF Staff Papers* 55.3, pp. 511–540.

Desmet, Klaus, Dávid Krisztián Nagy, and Esteban Rossi-Hansberg (2025). *Human Capital Accumulation Across Space*. Tech. rep. National Bureau of Economic Research.

Deutscher, Nathan (2020). “Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia”. In: *American Economic Journal: Applied Economics* 12.2, pp. 220–49.

Donaldson, Dave (2018). “Railroads of the Raj: Estimating the impact of transportation infrastructure”. In: *American Economic Review* 108.4-5, pp. 899–934.

Donaldson, Dave and Richard Hornbeck (2016). “Railroads and American economic growth: A “market access” approach”. In: *The Quarterly Journal of Economics* 131.2, pp. 799–858.

Duranton, Gilles and Matthew A Turner (2012). “Urban growth and transportation”. In: *Review of Economic Studies* 79.4, pp. 1407–1440.

Eckert, Fabian and Tatjana Kleineberg (2024). *The geography of opportunity: Education, work, and intergenerational mobility across us counties*. Tech. rep. Technical report, Working Paper.

Faber, Benjamin (2014). "Trade integration, market size, and industrialization: evidence from China's National Trunk Highway System". In: *Review of Economic Studies* 81.3, pp. 1046–1070.

Fenske, James, Namrata Kala, and Jinlin Wei (2023). "Railways and cities in India". In: *Journal of Development Economics* 161, p. 103038.

Fujimoto, Junichi, David Lagakos, and Mitchell VanVuren (2025). *Macroeconomic Effects of 'Free'Secondary Schooling in the Developing World*. Tech. rep. National Bureau of Economic Research.

Ghani, Ejaz, Arti Grover Goswami, and William R Kerr (2016). "Highway to success: The impact of the Golden Quadrilateral project for the location and performance of Indian manufacturing". In: *The Economic Journal* 126.591, pp. 317–357.

Gollin, Douglas, Martina Kirchberger, and David Lagakos (2021). "Do urban wage premia reflect lower amenities? Evidence from Africa". In: *Journal of Urban Economics* 121, p. 103301.

Gollin, Douglas, David Lagakos, and Michael E Waugh (2014). "The agricultural productivity gap". In: *The Quarterly Journal of Economics* 129.2, pp. 939–993.

Hsiao, Allan (2024). *Educational Investment in Spatial Equilibrium: Evidence from Indonesia*.

International Development Association (1981). "Mali – Road Maintenance Project: Credit 1104 – Credit Agreement – Conformed". In.

IPUMS (2020). "Integrated Public Use Microdata Series, International: Version 7.2 [dataset]". In.

Jedwab, Remi and Adam Storeygard (2021). "The average and heterogeneous effects of transportation investments: Evidence from Sub-Saharan Africa 1960-2010". In: *Journal of the European Economic Association*.

Khanna, Gaurav (2023). "Large-scale education reform in general equilibrium: Regression discontinuity evidence from India". In: *Journal of Political Economy* 131.2, pp. 549–591.

Lagakos, David (2020). "Urban-rural gaps in the developing world: Does internal migration offer opportunities?" In: *Journal of Economic perspectives* 34.3, pp. 174–92.

Laliberté, Jean-William P (2021). "Long-term contextual effects in education: Schools and neighborhoods". In: *American Economic Journal: Economic Policy*.

Michaels, Guy (2008). "The effect of trade on the demand for skill: Evidence from the interstate highway system". In: *The Review of Economics and Statistics* 90.4, pp. 683–701.

Milsom, Luke Heath (2023). "Spatial inequality of opportunity in West Africa". In: *Journal of Public Economics* 225, p. 104944.

Moneke, Niclas (2020). *Can Big Push Infrastructure Unlock Development? Evidence from Ethiopia*. Tech. rep. Working paper.

Morten, Melanie and Jaqueline Oliveira (2024). “The effects of roads on trade and migration: Evidence from a planned capital city”. In: *American Economic Journal: Applied Economics* 16.2, pp. 389–421.

Mukherjee, Mukta (2012). “Do better roads increase school enrollment? Evidence from a unique road policy in India”. In.

Okoye, Dozie, Roland Pongou, and Tite Yokossi (2019). “New technology, better economy? The heterogeneous impact of colonial railroads in Nigeria”. In: *Journal of Development Economics* 140, pp. 320–354.

Psacharopoulos, George and Harry Anthony Patrinos (2018). “Returns to investment in education: a decennial review of the global literature”. In: *Education Economics* 26.5, pp. 445–458.

Redding, Stephen J and Esteban Rossi-Hansberg (2017). “Quantitative spatial economics”. In: *Annual Review of Economics* 9, pp. 21–58.

Rojas Ampuero, Fernanda Catalina (2022). “Sent Away: The Long-Term Effects of Slum Clearance on Children and Families”. PhD thesis. UCLA.

Silva, JMC Santos and Silvana Tenreyro (2006). “The log of gravity”. In: *The Review of Economics and statistics* 88.4, pp. 641–658.

Simonovska, Ina and Michael E Waugh (2014). “The elasticity of trade: Estimates and evidence”. In: *Journal of international Economics* 92.1, pp. 34–50.

Sotelo, Sebastian (2020). “Domestic trade frictions and agriculture”. In: *Journal of Political Economy* 128.7, pp. 2690–2738.

Tchindjang, Mesmin et al. (2005). “Mapping of Main Artery Network and Colonial Porterage in Cameroon (1900–1940)”. In: *ICC 2005: Proceedings of the 22nd International Cartographic Conference — Mapping Approaches into a Changing World*. A Coruña, Spain: International Cartographic Association (ICA).

Tsivanidis, Nick (2026). “Evaluating the impact of urban transit infrastructure: Evidence from Bogota’s Transmilenio”. In: *American Economic Review*.

Van Maarseveen, Raoul (2025). “Urbanization and educational attainment: Evidence from Africa”. In: *Journal of Development Economics*, p. 103654.

Weidmann, Nils B, Jan Ketil Rød, and Lars-Erik Cederman (2010). “Representing ethnic groups in space: A new dataset”. In: *Journal of Peace Research* 47.4, pp. 491–499.

World Bank (1970). *Dahomey – Highway Maintenance and Engineering Project: Credit 0215 – Credit Agreement – Conformed*.

World Bank (1982). *Project Performance Audit Report: Cameroon Second and Third Highway Projects*.

Yotov, Yoto V et al. (2016). *An advanced guide to trade policy analysis: The structural gravity model*. World Trade Organization Geneva.

Young, Alwyn (2012). “The African growth miracle”. In: *Journal of Political Economy* 120.4, pp. 696–739.

ONLINE APPENDIX

A Historical context

A.1 Benin

Road infrastructure in the territory of present-day Benin has deep antecedents in precolonial state formation. For example, the Kingdom of Dahomey maintained a formalized "Royal Road" that linked the Atlantic port of Whydah (Ouidah) to the royal capital at Abomey, supporting royal mobility and administrative authority along a core political axis (Alpern, 1999). From the late nineteenth century, road development was reshaped by French colonial rule. In this period, transport priorities reoriented toward coastal access and administrative control. After independence (1960), Benin's early postcolonial period was marked by severe political instability (multiple coups between 1963 and 1972), which constrained long-horizon public investment planning, including in transport infrastructure.

Benin's post-1970 road policy was somewhat hampered by political and macroeconomic shocks: the 1972 coup brought Mathieu Kérékou to power, followed by an explicit Marxist-Leninist turn from 1974 and the proclamation of the People's Republic of Benin in 1975. In parallel, external financing from the World Bank came online e.g., the World Bank's Highway Maintenance and Engineering Project (World Bank, 1970). By the late 1980s–early 1990s, economic stress and reform pressures culminated in Benin's "National Conference" opening on 19 February 1990, widely treated as a watershed in Francophone Africa's democratic transitions and closely intertwined with adjustment-era economic liberalization. Over this period, the mounting costs of maintaining the existing infrastructure increasingly bite (World Bank, 1994).

From the 2000s, Benin's road strategy continues to shift from network expansion toward preservation of the existing network, particularly on the port–corridor system anchored by Cotonou. For example, the World Bank's Transport Infrastructure Rehabilitation and Maintenance Project explicitly "signaled a shift" toward maintenance of existing road and port assets and protecting past investments. However, this framing coexisted with large upgrading programs such as the AfDB-supported Djougou–N'dali improvement project and related works in northern Benin.

A.2 Cameroon

Prior to motor-road systems, mobility depended heavily on footpaths, porterage, and waterways. Recent work has traced modern road corridors in Cameroon to the colonial era, when transport infrastructure served territorial control and resource extraction (Tchindjang et al., 2005). These "main artery" corridors built during German colonial rule, from circa 1900—initially used for conquest, porterage, and forced labour logistics—later became the backbone of Cameroon's contemporary trunk road system. Cameroon's decolonization then produced a spatially and institutionally complex situation: following a UN-supervised plebiscite in February 1961, southern British Cameroons united with the newly independent Republic of Cameroon to form a federal state.

Post-1970 road policy developed alongside major constitutional and economic changes. The federal system was abandoned with the move to a unitary state in 1972, which redefined the governance environment for national trunk-road planning and maintenance responsibilities. In the same broad period, the World Bank's Second and Third Highway Projects (World Bank, 1982) targeted key trunk corridors—including the Transcameroon route and the Douala–Foumban road. However, Cameroon's mid-1980s commodity-price shock and associated recession, followed by IMF-supported structural adjustment, reduced budgets to road building and complicated routine and periodic maintenance. By 1990, chronic maintenance shortfalls led to the creation of a dedicated Road Fund under Law, which became operational by the late 1990s.

A.3 Mali

Mali's connectivity long predates modern roads. Precolonial economic integration relied on the Niger River valley and trans-Saharan caravan routes, with Timbuktu and related nodes historically positioned as trading posts on caravan corridors and as major centers of Islamic learning and commerce. Much like in Benin and Cameroon, under colonial rule (French) the development of the transport network shifted towards military, control, and extraction priorities. After independence (1960), Mali entered a relatively volatile political period, with a military takeover in 1968 shaping state capacity and investment planning—including transport.

By the 1980s, World Bank operations explicitly framed objectives around protecting a "priority" road network culminating in projects that combined rehabilitation with reforms to the construction and maintenance sectors. Maintenance became a standalone policy object by the 1980s: the Road Maintenance Project (International Development Association, 1981a), for example, reflects a recognition that deferred maintenance was becoming a serious problem. Politically, Mali also experienced major regime change with the 1991 overthrow of military rule and the restoration of a civilian government in 1992.

The Malian Road Fund became operational in July 2002. It created maintenance resources through axle-load charges, a fuel levy, toll fees on selected segments, and budgetary allocations. These reforms, however, unfolded under security and governance shocks: dissatisfaction with the government's handling of the northern conflict precipitated the 2012 coup, followed by sustained warfare and insurgency pressures that have repeatedly displaced policy attention and disrupted infrastructure delivery and upkeep.

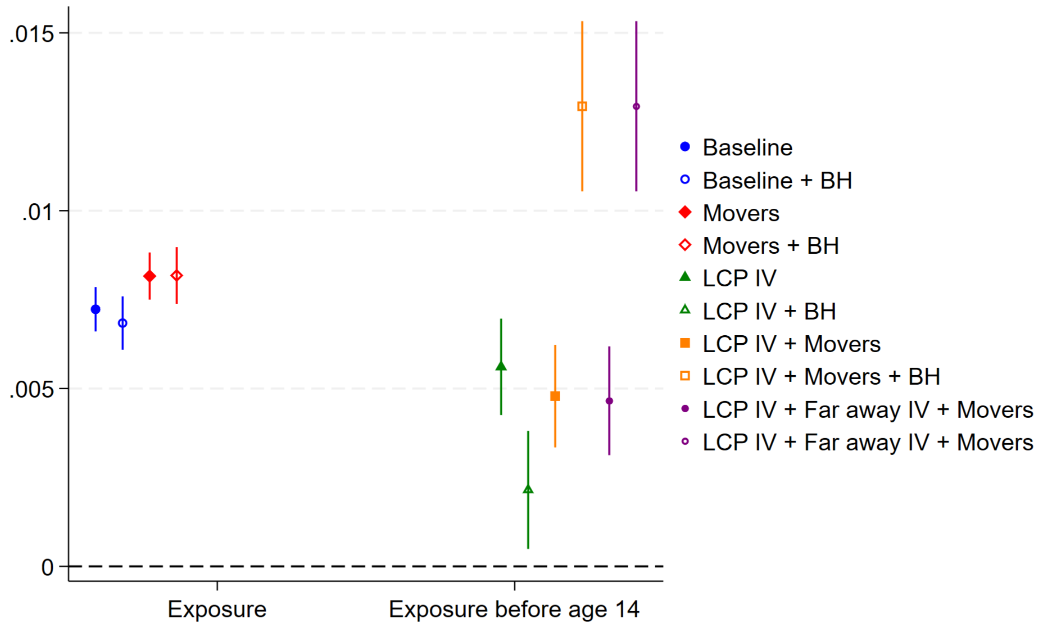
B Additional tables and figures

Table B.1 Estimating $\tilde{\theta}$ for each country in each year

Country	Year	Estimate	Standard Error
Benin	1992	-1.357822	.0552952
Benin	2002	-1.430456	.0672103
Benin	2013	-1.5607	.0950492
Cameroon	1976	-1.069	.0733561
Cameroon	1987	-1.091219	.0716346
Cameroon	2005	-1.165362	.0720341
Mali	1998	-1.282255	.0534792
Mali	2009	-1.316857	.0442214

Notes: This table shows the estimated elasticity of migration with respect to travel time at the country-year level, that is of estimating the migration gravity equation separately for each sample.

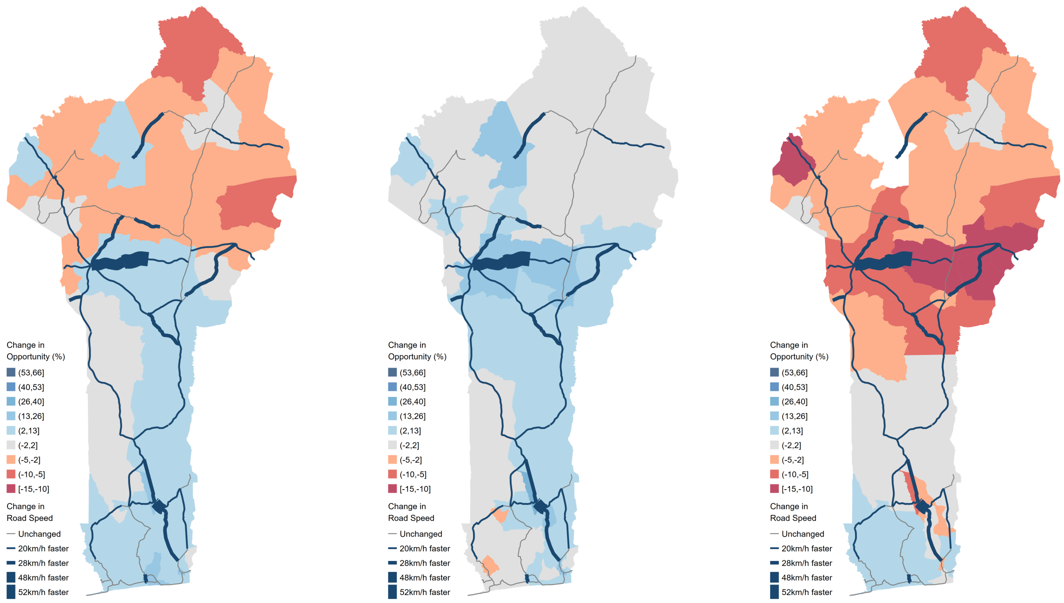
Figure B.1 Main results controlling for expected market access



Notes: This figure shows the robustness of the main result to controlling for expected market access. This procedure is described in detail in the main text.

Figure B.2 Benin Counterfactual Decomposition

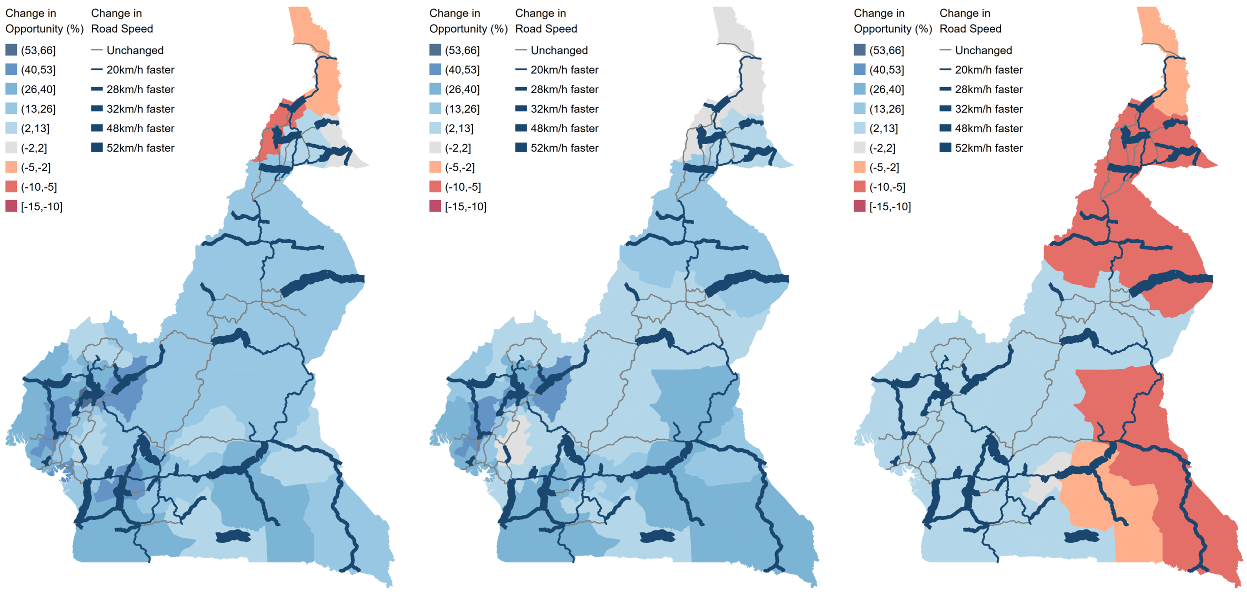
(a) Overall (b) Direct (c) GE effect



Notes: These maps show the spatial distribution across Benin's communes of the model-implied percentage change in opportunity due to road building since 1970, decomposed into its constituent components. Panel (a) reports the total full-equilibrium effect. Panel (b) reports the direct effect of lower migration costs, obtained by shutting down endogenous migration and trade responses. Panel (c) reports the net general-equilibrium adjustment, defined as the difference between the total effect and the direct effect. Shading indicates the percentage change in opportunity according to the legend; superimposed lines show changes in road speed since 1970, with thicker and darker lines indicating larger improvements.

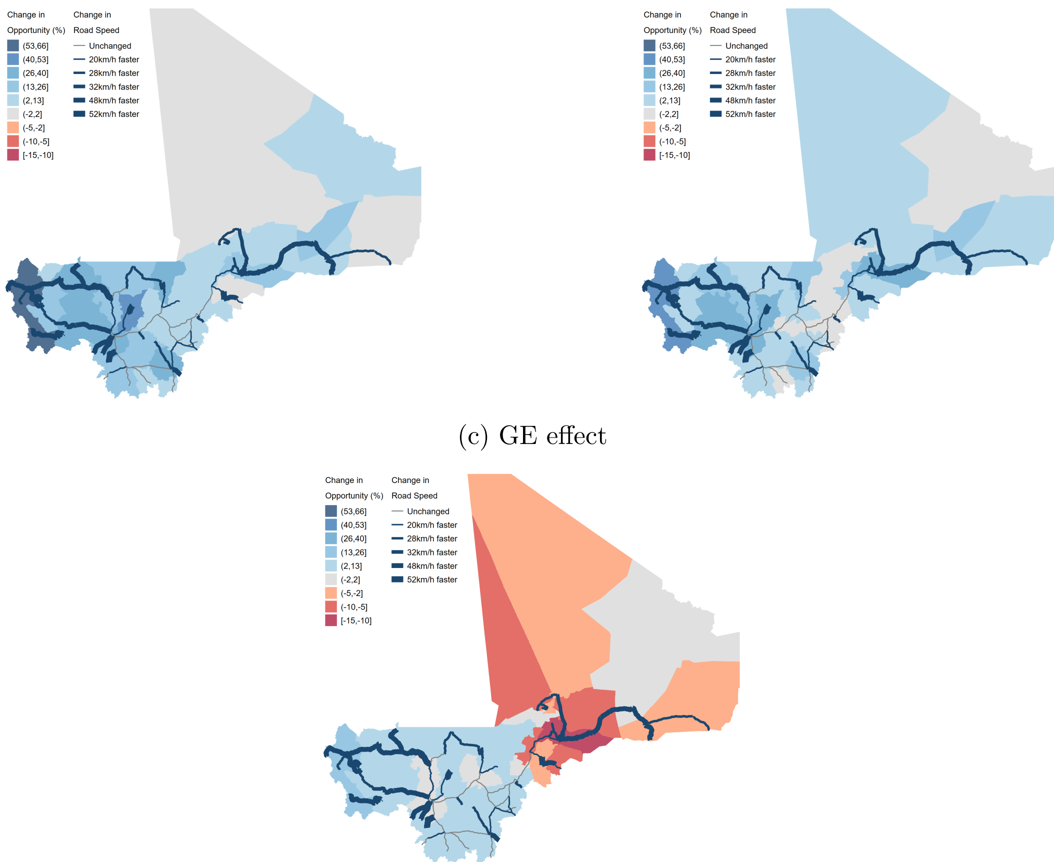
Figure B.3 Cameroon Counterfactual Decomposition

(a) Overall (b) Direct (c) GE effect



Notes: These maps show the spatial distribution across Cameroon's departments of the model-implied percentage change in opportunity due to road building since 1970, decomposed into its constituent components. Panel (a) reports the total full-equilibrium effect. Panel (b) reports the direct effect of lower migration costs, obtained by shutting down endogenous migration and trade responses. Panel (c) reports the net general-equilibrium adjustment, defined as the difference between the total effect and the direct effect. Shading indicates the percentage change in opportunity according to the legend; superimposed lines show changes in road speed since 1970, with thicker and darker lines indicating larger improvements.

Figure B.4 Mali Counterfactual Decomposition



(a) Overall (b) Direct (c) GE effect

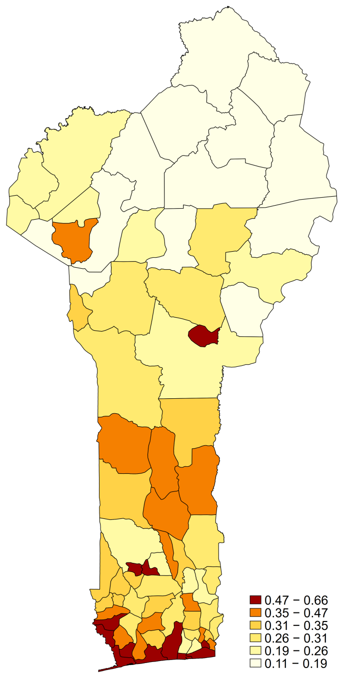
Notes: These maps show the spatial distribution across Mali's circles of the model-implied percentage change in opportunity due to road building since 1970, decomposed into its constituent components. Panel (a) reports the total full-equilibrium effect. Panel (b) reports the direct effect of lower migration costs, obtained by shutting down endogenous migration and trade responses. Panel (c) reports the net general-equilibrium adjustment, defined as the difference between the total effect and the direct effect. Shading indicates the percentage change in opportunity according to the legend; superimposed lines show changes in road speed since 1970, with thicker and darker lines indicating larger improvements.

C Descriptive statistics

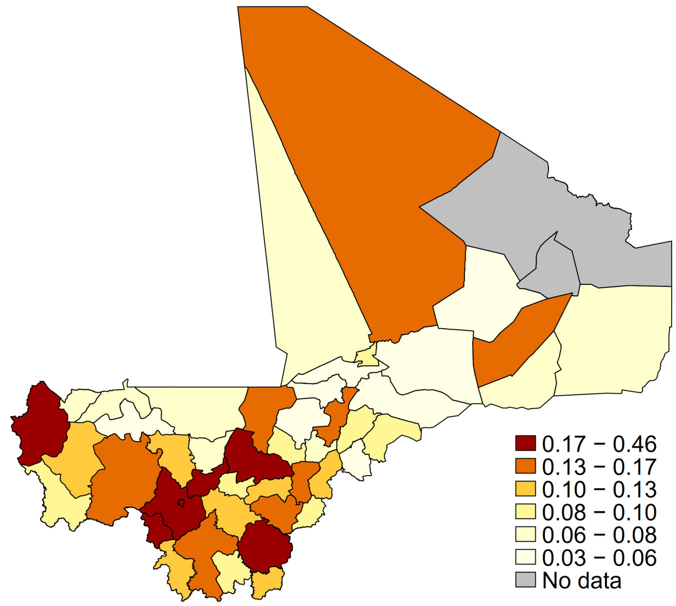
There is substantial variation in primary education completion within country across localities as can be seen by figures C.1a, C.1b and C.1c. In Benin in 2013 the proportion of individuals who had completed primary education in an area varied from 11% in the north to as high as 66% on the coast close to the capital. In addition, Parakou, a large city in the centre of the country had high completion rates. Mali, has consistently lower primary completion rates as compared to Benin or Cameroon. Figure C.1b also displays significant cross-locality variation with some areas' completion rates as low as 3% and some, especially round the capital, closer to 50%. Cameroon shows a similar pattern; the most educated areas are around the capital, or close to the large coastal city of Douala. Cameroon also has the highest completion rates over all, varying from 20% to almost 90%.

Figure C.1 Locality level primary completion rates

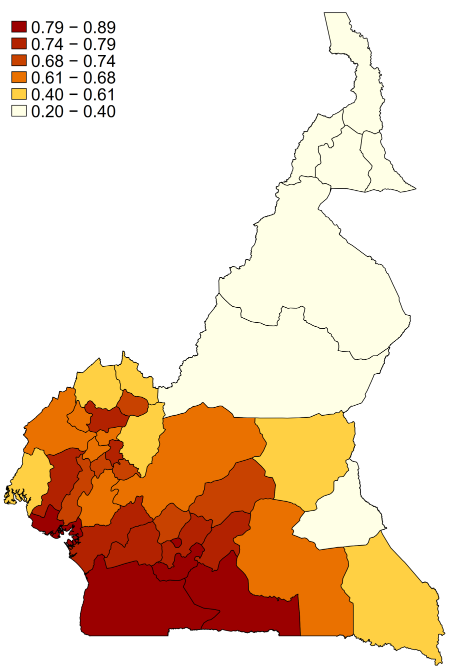
(a) Benin (2013)



(b) Mali (2009)



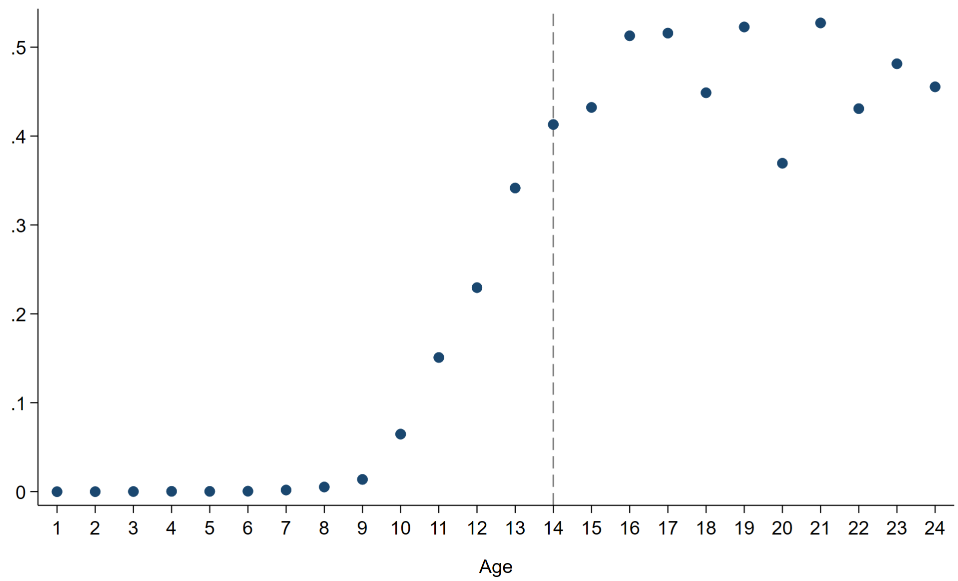
(c) Cameroon (2005)



Notes: this figure shows the spatial distribution of primary education completion rates for all those above the age of 12 in each of Benin, Cameroon, and Mali. Each figure has its own scale and corresponding legend, where darker orange/red indicates higher completion rates. The data for Benin comes from the 2013 census, for Mali the 2009 census, and for Cameroon the 2005 census.

Although not crucial for my analysis, it is helpful for the interpretation of results later that, indeed, most individuals who will at some point receive primary schooling do so by the age of 14. This is evident from figure C.2, which shows the proportion of the population who have completed primary education by age. This figure clearly shows, as anticipated, that most primary education is indeed completed by 14. However, it's worth noting that many children who haven't completed primary education by 12, which is the year officially primary schooling ends in the countries I study, go on to do so in the next few years.

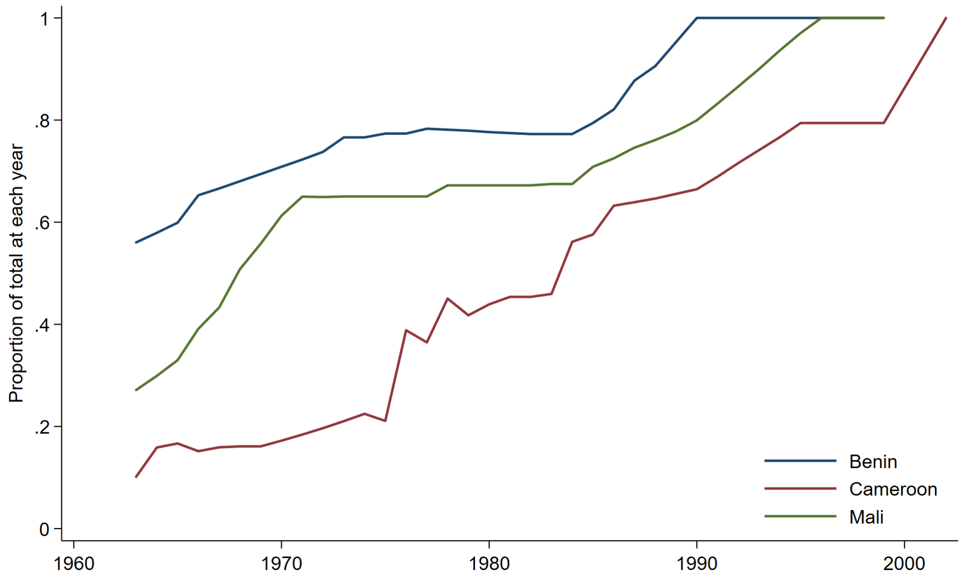
Figure C.2 Proportion of the population who have completed primary school by age



Notes: this figure shows the proportion of the sample who report having completed primary education at the time of each census against their age. This figure uses data from the full sample Benin (1992, 2002, 2013), Cameroon (1976, 1987, 2005), and Mali (1998, 2009).

Figure C.3 uses data from Canning and Pedroni (2008) to calculate the proportion of the completed paved road network existing in a given year over my study period. From this figure it's clear that, unlike railways in Benin, Cameroon, and Mali, roads were mainly a post-colonial technology displaying significant variation even in the recent past. In figure C.3, it's clear that Cameroon has seen the most intensive increase in road stock since 1960 when it had less than 20% of the length it does today. Mali and Benin, however, are not too far behind with less than 30% and less than 60% respectively of their modern road stock in place by 1960.

Figure C.3 Variation in paved roads



Notes: This figure shows the proportion of the 2000 total paved road stock in place in each given year for Benin, Cameroon, and Mali. It uses data from Canning and Pedroni (2008).

D Calculating expected travel times

A key object required to calculate market access terms is the iceberg-style movement and trade costs. Both of these are based on the fastest path from i to j in period t along the national transport network of the corresponding country, and so I turn first to estimating these travel times between i and j, which I denote by $\tilde{t}_{ijt}$. However, my data is available at the locality level, which means that $\tilde{t}_{ijt}$ is an aggregate measure of travel times across regions. In order to fully utilise the available variation and keep as close to the actual road network as possible, I don't just rely on centroid-to-centroid measures of distance across large localities. Instead, I take the interpretation that transport costs from i to j are measured as the expected cost of a randomly chosen individual in i travelling to a randomly chosen individual in j. That is, consider individuals $p \in i$ and $q \in j$ and denote their travel time as d_{pqt}. Then I estimate, $\tilde{t}_{ijt}$ as the following where $|i|$ and $|j|$ denote the population size of i and j respectively.

$$\tilde{t}_{ijt} = \frac{1}{|i|} \sum_{p \in i} \frac{1}{|j|} \sum_{q \in j} d_{pqt} \tag{D.1}$$

However, in order to estimate $\tilde{t}_{ijt}$ in this manner, I would need to observe the exact within locality distribution of the population. To focus on variation in road building rather than potentially endogenous changes in the population distribution, I estimate the within locality population distribution in the pre-sample year of 1970 and keep it fixed. To do this, I introduce a new data source, Africapolis, which maps all agglomerations in Sub-Saharan Africa that will achieve a population of at least 10,000 in 2015 and backdates each agglomerations' population to 1970. I take all such available agglomerations, their exact coordinates and 1970 populations. To this, I add the backdated approximate remaining population of each locality using census data and assign this to the locality centroid. This gives the best estimated within-locality and within-country population distribution in 1970 using available data.

Having completed the above steps, I have a set of locations p within each locality i that is $p \in i$. For each location, I associate a 1970 population $P_{p,1970}$ and time of travelling along the observed road network to each other point $q \in j$ for each j location in the same country, d_{pqt}. Then the expected travel time of a randomly chosen household in i travelling to a randomly chosen household in j in year t is given by the following.

$$\tilde{t}_{ijt} = \sum_{p \in i} \frac{P_{p,1970}}{P_{i,1970}} \sum_{q \in j} \frac{P_{q,1970}}{P_{j,1970}} d_{pqt} \tag{D.2}$$

This can be seen as a coarse discretisation of equation D.1, the best that can be done with the data available. Travel times d_{pqt} are calculated along the digitised actual road network, distinguishing between types of road as described in the main text.

To this object, I then add approximations for the internal cost of traveling through both the origin and destination locations. This ensures that the cost of travel and migration within a location scales with that across, otherwise the system would not be invariant to the size of geographic unit chosen for analysis.[12] I assume these costs scale with the size of each location, measured as the square root of the locality's area. I then assume that within a location, one can travel at 60km/h. This is intended to be faster than the average across-location travel time because within-location, much of the population is condensed into one or two main agglomerations, within which travel times are relatively fast. Denote within-location travel times calculated in this way by $\bar{t}_i$ and $\bar{t}_j$ for origin and destination locations, respectively. Then my final measure of $i - j$ travel times in period t is given by $t_{ijt} = \tilde{t}_{ijt} + 0.5 \times \bar{t}_i + 0.5 \times \bar{t}_j$.[13]

[12]For example, if we maintained that travel within-location was costless, then a country with larger geographic units would effectively have a higher cost of trade as compared to a country with smaller units.

[13]Note that the theory allows a generalisation of strictly symmetric travel times to incorporate costs of this form.

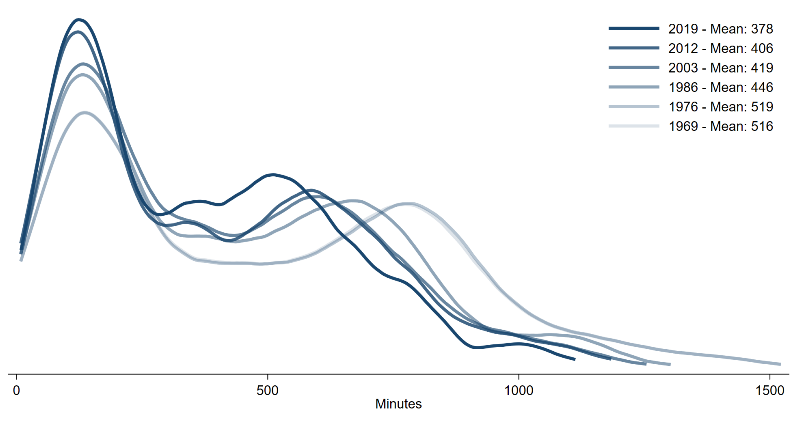
E Variation in connectivity over time and space

In figure E.1, I plot the distribution of expected travel times for each location, in each year maps are available over the digitised road network.[14] The expected travel time for a given location is defined as the length of time an individual should expect to travel for if they were to pick a person at random to travel to from the rest of the country. To calculate this over time, I fix the population distribution to 1970 levels and calculate each locality's expected travel time using the road network in each year. All three figures show considerable leftward shifts in the distribution of travel times over the study period, with mean travel times decreasing by 27%, 41%, and 44% in Benin, Cameroon, and Mali, respectively.

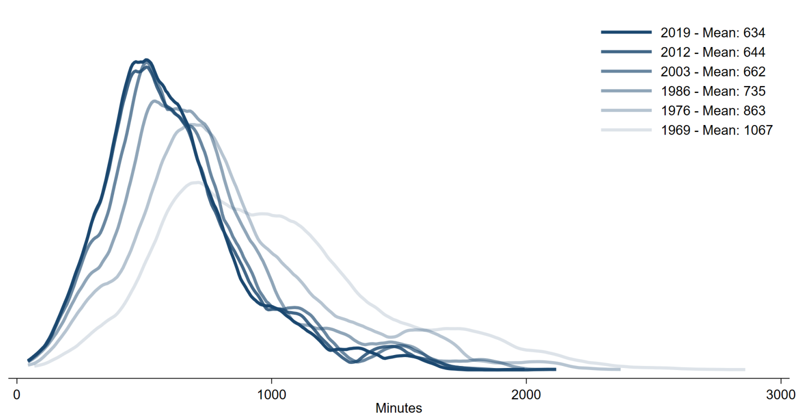
[14] Over my study period, in Benin, Cameroon, and Mali, other forms of transport such as railways or waterways exhibited little variation and are not modelled.

Figure E.1 Distribution of expected pairwise travel times between localities

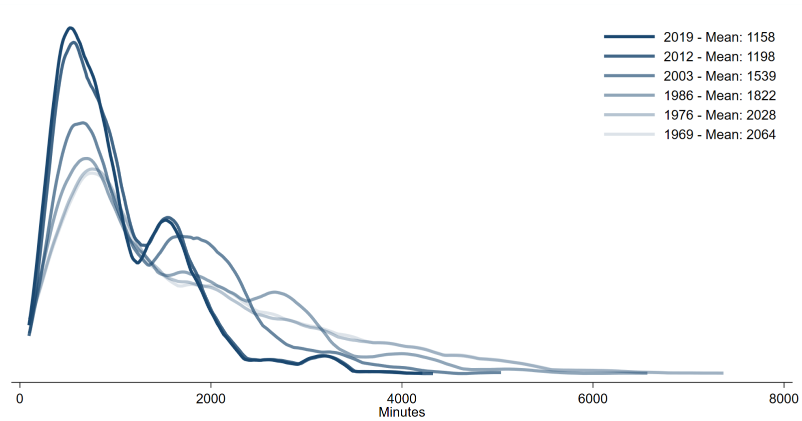
(a) Benin



(b) Cameroon



(c) Mali



Notes: These figures show the density of expected travel times from each locality in a given country-year. Expected travel time in a given location is defined as the time an individual in the location should expect to be traveling if one chooses an individual at random in the same country to travel to. The population distribution is kept fixed at 1970 levels, but the road network is allowed to vary. More recent years are denoted in a darker shade of blue. Population-weighted means across localities for each year are given in the top right.

F Calculating locality-year specific incomes

Census data does not provide information about wages or the total income/ output of localities. However, it does provide some limited information on the assets households own, such as flooring material, sanitation, and electricity. I can use this information, coupled with auxiliary regressions using income data from development health surveys (DHS), to impute approximate income at the locality-year level. Intuitively, this approach is similar to that of Young (2012) in that I use auxiliary Engle curve regressions to uncover parameters which are then used in a second stage with richer data to impute the outcome of interest at a broader and more granular geographic level. This approach requires some assumptions which are difficult to test; however, given the paucity of data available on wages/ incomes at sufficient geographic and temporal disaggregation, I believe that this is a good approximation. Additionally, due to high informality rates, it's unclear whether wages would be the most appropriate measure even if they were available.

Postulate that the (real) demand for an asset a by household h in locality i in year t is given by the following equation.

$$\ln(Q_{ahit}) = \alpha_a + \eta_a \ln(C^N_{hit}) + \xi_a \ln(P_{it}) + \beta X_{hit} + \varepsilon_{ahit} \tag{F.1}$$

Where α_a are product constants, η_a is the (quasi) income elasticity of demand, C^N_{hit} is nominal household consumption expenditure which is equal to household income in our setting, ξ_a is a vector of own and cross-price (quasi) elasticities of demand, $\ln(P_{it})$ is a vector of regional prices, X_{hit} and β are vectors of household characteristics and their coefficients. Finally ε_{ahit} is a white noise household-product preference shock. Elasticities are referred to as quasi above as for all assets considered, I use an indicator variable rather than a logarithm. To estimate this equation, I use data from the available DHS waves in Benin, Cameroon, and Mali that report income. Sadly, this is only two waves: Benin in 1995 and Mali in 1996. Additionally, these surveys don't include information on prices, and so I estimate equation F.1 using product-locality-year fixed effects (although year fixed effects are redundant given that localities are only observed once), which absorbs price variation. Results from running regressions are given in table F.1.

Focusing on column (4), these results suggest that a 1% increase in income is associated with a 0.07pp. increase in the probability of having a concrete floor, a 0.034pp increase in the probability of having access to electricity, and a 0.049pp. increase in having accessible sanitation.

In the second step, I use the inverted estimated coefficients from table F.1 to approximate income differences by assets households own as indicated in census data. Imputed average income in a locality-year cell is then given by $\tilde{Y}_{it} = 1/N_{it} \sum_{h \in \{i,t\}} \sum_a \frac{1}{\hat{\eta}_a} Q_{hait} + base$, where

Table F.1 Asset demand equations using DHS data

	(1)	(2)	(3)	(4)
Concrete floor	0.0744***	0.0744***	0.0528***	0.0733***
	(0.00606)	(0.00607)	(0.00538)	(0.00607)
Electricity	0.0351***	0.0351***	0.0228***	0.0341***
	(0.00472)	(0.00475)	(0.00378)	(0.00469)
Sanitation	0.0504***	0.0504***	0.0302***	0.0494***
	(0.00527)	(0.00526)	(0.00483)	(0.00531)
Asset × Region FE	X	X	X	X
Age polynomial		X	X	X
Asset × Region × Urban FE			X	
HH members control				X
R^2	0.394	0.394	0.481	0.395
N	13684	13684	13684	13684

Notes: This table shows the results from running regressions of the form given in equation F.1 using DHS data.

base is the average income calculated from DHS data. Intuitively, if we see households in an area with more assets than those in a different area, we infer that those in the first area have more income with which to purchase such assets. The Engle curve auxiliary regressions in the first stage allow me to approximate how much more income owning an asset may signal, and thus translate regional differences into a money-metric form.

G Digitizing Maps

Data on the changing connectivity of place comes from digitised historical Michelin road maps accessed from the Bodleian Library at the University of Oxford. In this section, I detail the digitising procedure taken. Figure G.1 gives examples of the original maps used for Mali. Throughout, I used the geographical mapping software ArcGIS. The procedure taken is detailed in the steps below.

1. Download the Open Street Maps shapefiles for Benin, Cameroon, and Mali, which representing the current road network in each country.

2. Remove minor roads, or other roads not represented on the most recent (2019) Michelin road maps.

3. Categorise all remaining roads as in the most recent Michelin road maps.

4. Add all settlements from Africapolis that includes all agglomerations that have a population of at least 10,000 in 2015.

5. Using settlement-level population estimates from Africapolis and back-dated locality level population estimates from Census data, calculate the remaining locality-level population not covered by the Africapolis settlements. Add this population to a location at the centre of each locality.

6. Make small adjustments to the 2019 road network so that roads hit the centroid of each settlement and form a connected network. To do this I used the topology tool in ArcGIS.

7. Iteratively delete or downgrade roads using maps increasingly in the past. In this way, for each year, a map is available, I create the complete road network. Note that I did not find any examples of Michelin roads that were not visible in the OSM map.

Figure G.1 Original maps

(a) Mali 1969 (b) Mali 2019

Notes: This figure shows pictures of the original Michelin road maps for Mali in 1969 on the left hand side and in 2019 on the right hand side.

H Major Road Building Initiatives in Benin, Cameroon, and Mali, 1970–2020

H.1 Benin

Third and Fourth Highway Projects (1970s–1980s)

The World Bank's Third Highway Project (Credit 746) and Fourth Highway Project (Credit 1142) continued IDA involvement in Benin's roads sector. The Third Highway Project financed the rehabilitation of a 107 km section of the country's main north–south road axis to facilitate access to the port of Cotonou. *Sub-regions*: the Cotonou–Parakou corridor (communes along the Route Nationale Inter-État 1 (RNIE1): Cotonou, Abomey-Calavi, Allada, Bohicon, Dassa-Zoumé, Glazoué, Savè, Parakou). *Decades*: 1970s–1980s.

Transport Infrastructure Rehabilitation and Maintenance Project (late 1980s)

This World Bank project (Report No. 6000-BEN; IDA Credit of approximately US$19.5 million; total cost US$70.5 million) marked a strategic shift toward maintenance of existing road and port assets and protecting past investments (World Bank, 1987). It explicitly targeted the port–corridor system anchored by Cotonou, and continued and expanded earlier IDA assistance. *Sub-regions*: Cotonou port and main trunk roads linking Cotonou to the interior and to Burkina Faso. *Decade*: late 1980s–early 1990s.

Transport Sector Investment Program (1997–2001)

This US$40 million IDA-funded programme (1997–2001) aimed to safeguard the competitiveness of Benin's transport sector and its transit corridor through open modal competition, improve government planning and programming capacity, boost resource allocation to infrastructure maintenance, and expand private sector participation in public works (Mohan, 2003). The programme also included a substantial rural roads component. *Sub-regions*: Cotonou–Niger transit corridor and rural areas across northern and central Benin. *Decade*: 1990s–2000s.

Abidjan–Lagos Corridor Transport and Transit Facilitation Project (2010s)

This multi-country World Bank project (US$340 million across Benin, Côte d'Ivoire, Ghana, Nigeria, and Togo) targeted the 998.8 km Abidjan–Lagos coastal corridor, linking some of Africa's largest and most economically dynamic capitals. Benin received US$75 million (World Bank, 2010b). In Benin, the project specifically rehabilitated the Godomey–Pahou road section by widening it to two lanes in each direction and facilitated trade through a single window at the Port of Cotonou. *Sub-regions*: the coastal strip through Godomey and Pahou (Atlantique department, west of Cotonou), and the Port of Cotonou. *Decade*: 2010s.

AfDB: Lomé–Cotonou Road Rehabilitation (2011–)

The African Development Fund approved a multinational Benin/Togo project to rehabilitate the Lomé–Cotonou road as Phase I of the broader Abidjan–Lagos Corridor Organisation (ALCO) effort. Total financing was approximately UA 86.55 million. The project aimed to facilitate trade and transport along the east–west coastal corridor (African Development

Fund, 2016). *Sub-regions*: the coastal corridor from the Togo border through to Cotonou (Mono, Couffo, Atlantique, Littoral departments; communes including Grand-Popo, Comé, Ouidah, Cotonou). *Decade*: 2010s.

AfDB: Djougou–N'Dali Road Improvement Project (2000s)

The African Development Bank (ADF), together with BOAD, NTF, and the Government of Benin, co-financed the upgrade of the Djougou–N'Dali road in northern Benin. This was part of Benin's 1997–2006 Transport Sector Programme (TSP) (African Development Bank, 2009). Related segments include the Djougou–Ouaké–Togo border stretch (BOAD-financed) and the N'Dali–Chicandou–Nigerian border stretch. *Sub-regions*: Donga and Borgou departments (communes of Djougou, N'Dali, and connecting communes). *Decades*: 2000s–2010s.

H.2 Cameroon

World Bank Second and Third Highway Projects (1970s–1980s)

The World Bank's Second Highway Project (Loan 935-CM, Credit 429-CM) and Third Highway Project (Loan 1515-CM) targeted key trunk corridors, including the Transcameroon route and the Douala–Foumban road (World Bank, 1980). An associated Highway Engineering Project prepared improvements for the Ngaoundéré–Garoua road (part of the "Cameroon Route") and the Douala port. *Sub-regions*: the Transcameroon corridor (Douala–Yaoundé–Ngaoundéré; departments of Wouri, Moungo, Sanaga Maritime, Mfoundi, Lékié, Haute Sanaga, Mbam et Inoubou, Lom et Djerem, Vina), the Douala–Foumban road (Noun department), and the Ngaoundéré–Garoua axis. *Decades*: 1970s–1980s.

Bertoua Region Road Programme (1980s)

A road-building programme was completed in the Bertoua region in the southeast in 1986. This connected the East Region more firmly to the national network. The period also saw the completion of the first all-weather highway from Yaoundé to Douala and between Yaoundé and the western high plateau. *Sub-regions*: East Region (departments of Lom et Djerem, Kadey, Haut Nyong). *Decade*: 1980s.

CEMAC Transport and Transit Facilitation Project (2007–2020)

This landmark World Bank regional project (initially US$201 million IDA, expanded with US$217 million additional financing) was designed to facilitate trade among CEMAC member states and improve Cameroon's, Chad's, and CAR's access to world markets (World Bank, 2010a). The project covered: paving the Garoua Boulaï–Ngaoundéré road (254 km); rehabilitating rail track between Yaoundé and Belabo; port facilitation and customs reform at Douala; and road sections along the Douala–N'Djamena and Douala–Bangui corridors. *Sub-regions*: the Douala–Ngaoundéré–Garoua Boulaï corridor (Littoral, Centre, Adamawa, East regions), the Douala–N'Djamena corridor extending into the Far North, and the Douala–Bangui corridor through the East Region. *Decades*: 2000s–2010s.

Cameroon Multimodal Transport Project (2014–)

Approved in 2014 with US$71 million IDA credit (US$91 million total), this project aimed to establish good travel conditions along 90 per cent of the 1,842 km Douala–N'Djamena corridor. It rehabilitated the severely degraded Maroua–Mora road section (60 km) and supported performance-based road maintenance of a 270 km section between Maroua and Kousseri (World Bank, 2014). *Sub-regions*: Far North Region (departments of Diamaré, Mayo Tsanaga/Mayo Sava, Logone et Chari, Mayo Danay), plus maintenance along the full Yaoundé–Kousseri corridor. *Decade*: 2010s.

Yaoundé–Douala Tolled Highway (2014–ongoing)

Cameroon's first expressway, the 195 km Yaoundé–Douala tolled highway, is the most ambitious national road project in the country's history, forming part of Trans-African Highway 8 (Lagos–Mombasa). Phase 1 (60 km, Yaoundé to Bibodi) was launched in October 2014 and opened in December 2021, at a cost of approximately 350 billion FCFA. Phase 2 (141 km) was launched in October 2024 and budgeted at 1,072 billion FCFA, financed through Exim Bank of China and Standard Chartered Bank. *Sub-regions*: Centre Region (Mfoundi, Lékié) through to Littoral Region (Sanaga Maritime, Moungo, Wouri). *Decades*: 2010s–2020s.

H.3 Mali

World Bank First and Second Highway Projects (1970s)

IDA Credit 197-MLI (US$7.7 million, 1970): the First Highway Project included highway maintenance and betterment of feeder roads (1,450 km programmed, though only 470 km completed), plus feasibility studies for Mali's two oldest paved trunk roads linking Bamako with Bougouni to the south and with Ségou to the northeast (World Bank, 1983). IDA Credit 383-MLI (US$17.8 million, 1973/1975): the Second Highway Project provided for reconstruction of the Faladié–Ségou road (223 km), completed in 1981. *Sub-regions*: Bamako–Bougouni corridor (circles of Bamako, Kati, Bougouni), Bamako–Ségou corridor (circles of Bamako, Kati, Koulikoro, Ségou). *Decade*: 1970s.

Third Highway and Road Maintenance Projects (late 1970s–1980s)

IDA Credit 599-MLI (1976): The Third Highway Project continued feeder road programmes. IDA Credit 1104-MLI (US$17.0 million, 1981): the Road Maintenance Project explicitly targeted protection of a "priority" road network and combined rehabilitation with reforms to the construction and maintenance sectors (International Development Association, 1981b). *Sub-regions*: priority national network, with focus on the Bamako–Ségou and Bamako–southern corridors (circles of Bamako, Kati, Koulikoro, Ségou, Bougouni, Kayes). *Decades*: late 1970s–1980s.

West Africa Regional Transport and Transit Facilitation: Tema–Ouagadougou–Bamako Corridor (2000s–2010s)

This multi-country World Bank/UEMOA project (Burkina Faso, Ghana, and Mali) rehabilitated key sections of the Tema–Ouagadougou–Bamako road transport corridor (Independent Evaluation Group, 2018). The Malian component improved the Heremako–Bamako section, bringing roads in good condition from 40 per cent at baseline to 85.4 per cent at project closure. The project also addressed transit facilitation. *Sub-regions*: the corridor entering Mali from Burkina Faso toward Bamako (circles of Sikasso, Bougouni, and Bamako). *Decades*: 2000s–2010s.

AfDB: Southern Bamako–Dakar Corridor (2010s)

The AfDB financed the Mali/Senegal Road Development and Transport Facilitation Project on the Southern Bamako–Dakar Corridor. This was part of the UEMOA PACITR programme, and JICA also contributed through bridge construction on the southern route. The southern corridor (via Tambacounda–Kédougou–Saraya–Moussala–Kita–Bamako) is approximately 200 km shorter than the northern route. *Sub-regions*: Kayes Region (circles of Kita, Kéniéba, and westward to the Senegalese border), circle of Kati. *Decade*: 2010s.

Northern Bamako–Dakar Corridor (2010s–2020s)

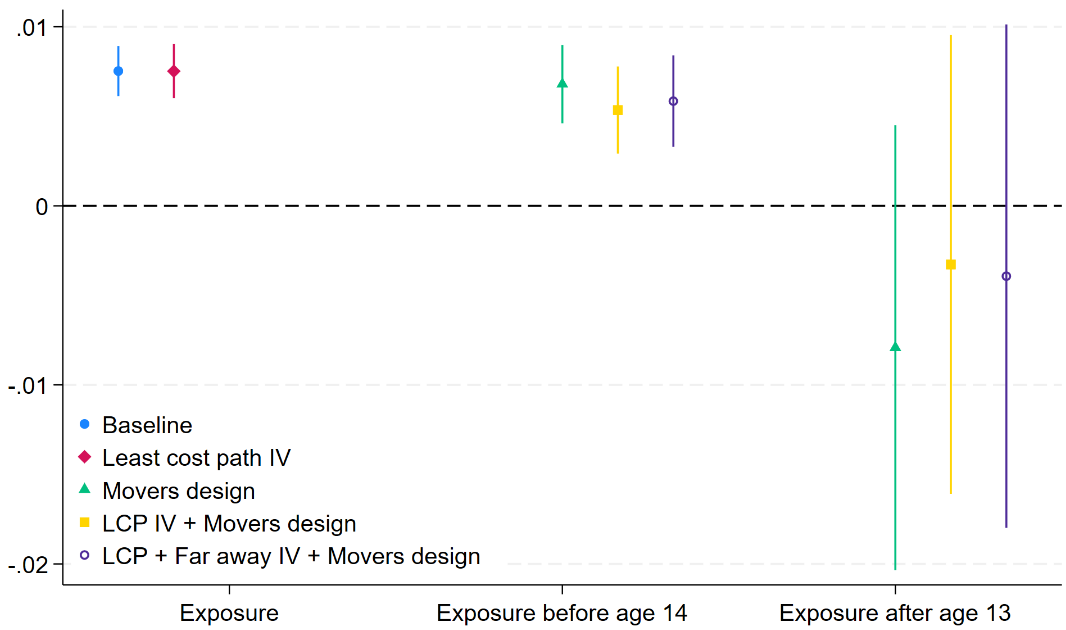
The Northern Bamako–Dakar Corridor (NBDC) is the major lifeline of Mali's economy, carrying the majority of the country's passenger and freight traffic. The route runs Bamako–Kati–Kolokani–Diéma–Kayes–Kidira/Diboli (approximately 1,470 km). Following the Ivorian crisis in 2002, this corridor absorbed the majority of Mali's maritime transit traffic. The AfDB has invested approximately US$400 million in the corridor, which now carries more than 50 per cent of Mali's imports and exports from Dakar (African Development Bank, n.d.). The most recent World Bank intervention is the Mali Connectivity and Road Resilience Programme (US$219.8 million, approved May 2025), rehabilitating the 137.7 km

Diéma–Sandaré section to climate resilience standards (World Bank, 2025). *Sub-regions*: Kayes Region (circles of Kayes, Diéma, Nioro, Yélimané, Bafoulabé), Koulikoro Region (circles of Kati, Kolokani), District of Bamako. *Decades*: 2010s–2020s.

H.4 Results controlling for road corridors and initiatives

Figure H.1 shows the results from Figure 4 in the main text, controlling for road corridor and initiative fixed effects at the destination-location level. The coefficients are qualitatively and quantitatively very similar to those in the main text. This is reassuring, as it suggests the results are not driven by spurious correlation at the region (or corridor) level. Although it is possible that I have not managed to capture all regional or corridor-level road strategies in the above analysis, it seems likely that the most prominent such examples are accounted for — and if these do not affect the estimates, it seems unlikely that less well-known or documented strategies would.

Figure H.1 Reduced-form results controlling for road corridors and initiatives

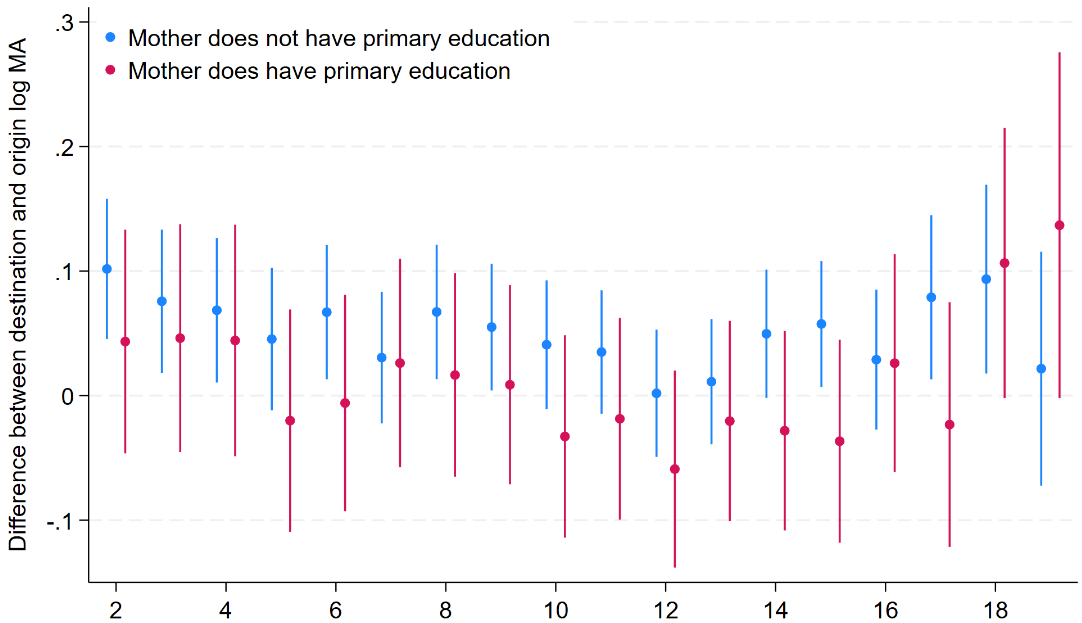


Notes: This figure replicates Figure 4 in the main text, controlling for road corridor and initiative fixed effects at the destination location level. All regressions are run on a sample of 14 to 18-year-olds and include origin location by year-born by census-year by age fixed effects, and age moved fixed effects. Standard errors are clustered at the origin by year of birth level, and indicated by horizontal lines. In blue with a circular marker, I show the baseline results of regressing average log market access over one's childhood on primary completion. In red with a diamond marker, I instrument market access using not-on-least-cost-path variation as discussed in the main text (the first stage Kleibergen-Paap rank LM statistic is 232.9). In green with triangular markers, I then report slope coefficients from estimating the movers design. These coefficients refer to movers before and after age fourteen as indicated on the x-axis. In yellow with square markers, I then combine the instrumental variables approach with the movers design (the first stage Kleibergen-Paap rank LM statistic is 417.2). Finally, in purple, I use a stricter instrument that combines the not-on-least-cost-path approach with that using far-away variation with a 50km cut off (the first stage Kleibergen-Paap rank LM statistic is 437.6).

I Reduced form results robustness and validity checks

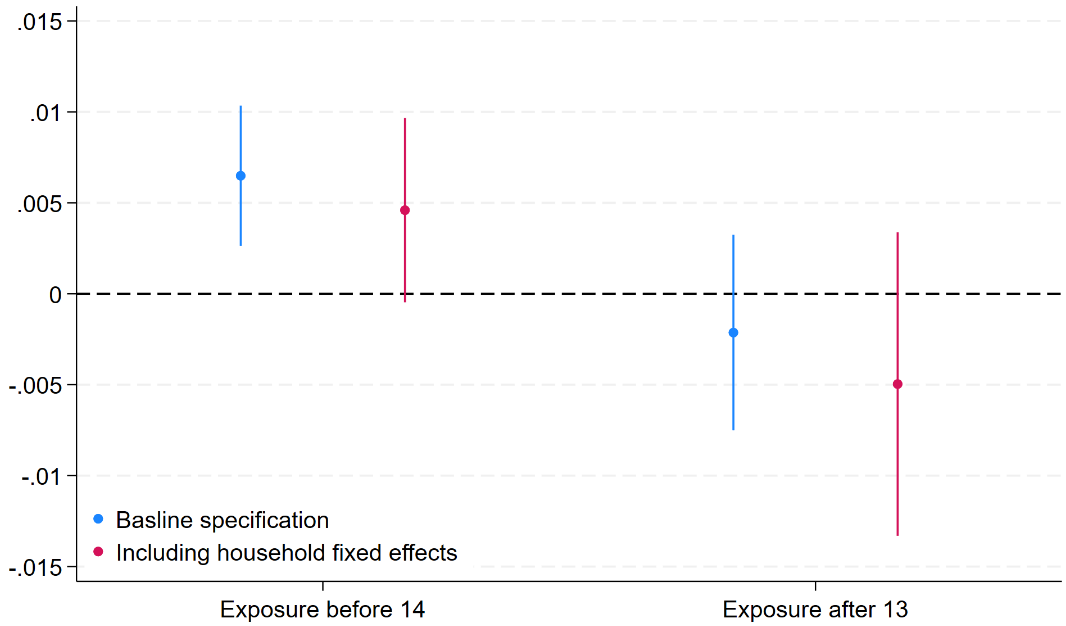
In this section, I report the results from the robustness and validity checks discussed in the main text.

Figure I.1 Age-at-move varying selection by mothers education



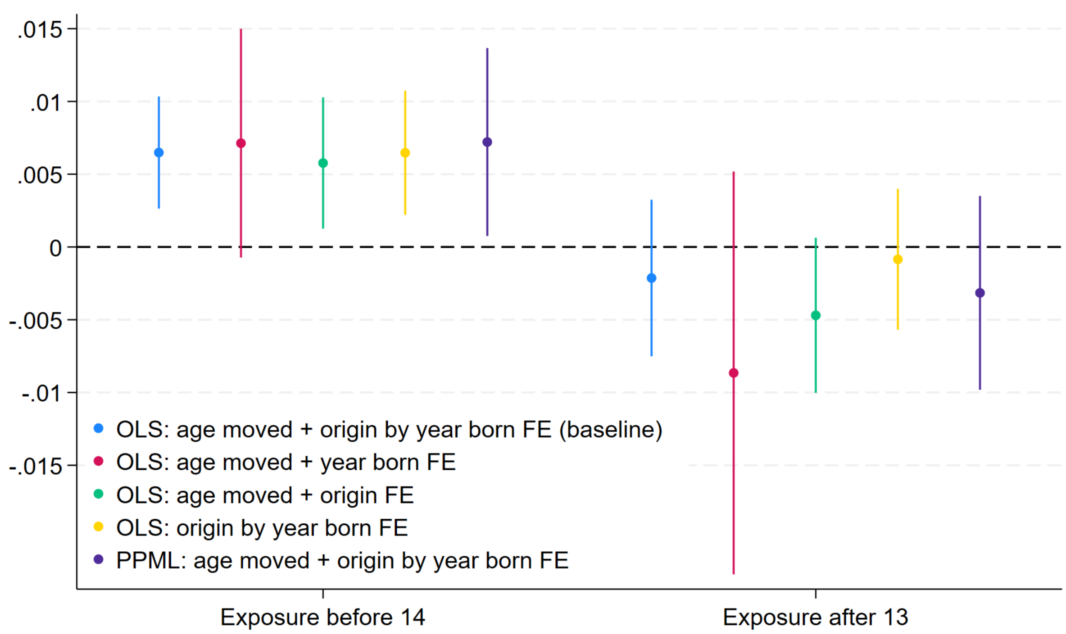
Notes: This figure shows the change in market access between destination and origin location by age at move for those whose mothers have and have not completed primary education.

Figure I.2 Including household fixed effects



Notes: This figure compares the estimated β_1 and β_2 from the main specification in the main text to the same coefficients estimated with the inclusion of household fixed effects.

Figure I.3 Specification robustness



Notes: This figure shows the robustness of the main specification in the main text to specification changes. Each colour refers to a different specification as indicated in the legend. As in the main results figure the coefficients β_1 and β_2 are plotted.

J Local clientelism and public good provision

One potential threat to the identification strategy used in this paper is that the provision of government services and public goods may vary over time and space. That is, if a government comes into power and builds roads and schools so as to benefit a given location in a potentially complex way that is not nullified by the not-on-least-cost-path identification strategy, this could bias the estimated coefficients. In my setting, this concern is most manifest when considering the interaction between local ethnic groups and that of the current political leader, as discussed in the Kenyan context by Burgess et al. (2015). However, the situation in Benin, Cameroon, and Mali is very different to that in Kenya. In Cameroon, Paul Biya has been in power since 1982, and thus, in Cameroon, there has been no temporal variation over my study period. In Mali, although there has been considerable variation in presidents since the 80s, ethnic favouritism or clientelism has been found to play only a minor, or perhaps even non-existent role (Dunning and Harrison, 2010; Basedau et al., 2011; Basedau and Stroh, 2012; Franck and Rainer, 2012).

In Benin, however, there is some evidence of politics having an ethnic component and clientelism (Battle and Seely, 2010; Fujiwara and Wantchekon, 2013; Wantchekon, 2003) and some correlational evidence that this may lead to less road building in politically marginalised locations Blimpo et al. (2013). To investigate whether these forces are driving my estimated effects, I construct a dummy variable equal to 1 if the ethnic majority in a location is equal to that of the leader of the time in Benin. Using the Geo-referencing of ethnic groups (GREG Weidmann et al. (2010)) database, I assign each locality in Benin to one of the four major Beninese ethnic groups.[15] Over my sample period, Benin has had three political leaders: in 1992, Nicéphore Soglo (Fon/ Ewe) was in power, in 2002, Mathiew Kérékou (Somba) was in power, and in 2013, Thomas Boni Yayi (Yoruba) was in power. In table J.1, I show the results from my baseline movers-design analysis in the first column and, in the second column, replicate this result, additionally controlling for the variable described above. The coefficients on the market access term are stable, and the coefficient on the same ethnicity variable is statistically indistinguishable from zero. I take this as evidence to suggest that threats to identification of this nature are minimal.

[15] Broadly defined as: Ewe, Yoruba, Somba or Barba.

Table J.1 Controlling for ethnicity by leader

	(1)	(2)
Exposure to MA before 14	0.00649***	0.00649***
	(0.00195)	(0.00195)
Exposure to MA over 13	-0.00213	-0.00215
	(0.00272)	(0.00273)
Current location has the same predominant ethnicity as the leader		0.0112
		(0.0566)
Observations	109207	109207
R^2	0.353	0.353

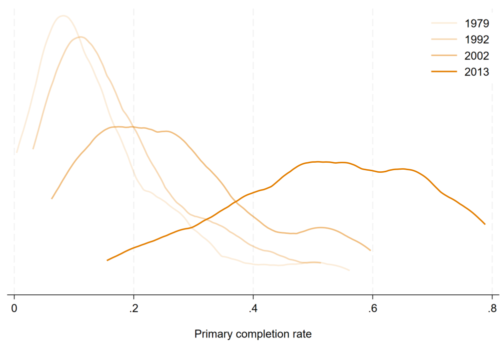
Notes: This table compares the baseline results to those including controls for clientelism. Column one replicates the baseline results using a movers design. Column two includes a dummy variable equal to one if the majority ethnicity of the locality is equal to that of the leader in Benin and zero otherwise.

K Top-coding

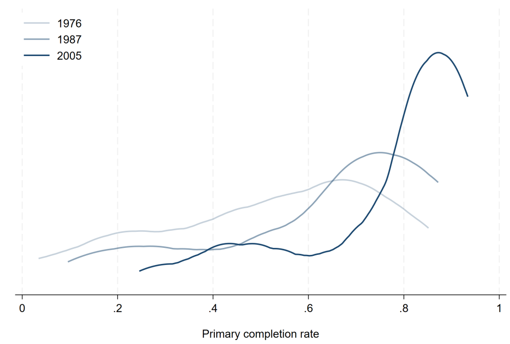
Figure K.1 shows how the distribution over localities of primary completion rates for those aged between 15 and 20 has changed in each country over the sample period. These figures show considerable rightward shifts in the distribution, but don't display bunching around 100% — that is they show evidence that top-coding at the upper limit of 100% primary completion is not present.

Figure K.1 Changes in the distribution of primary completion rates

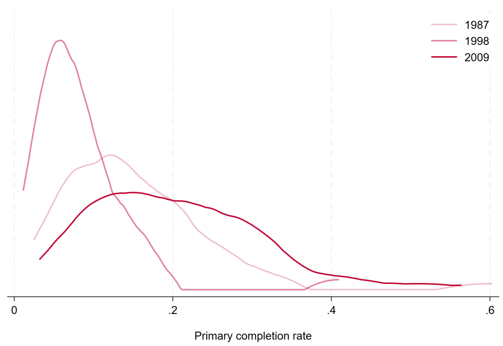
(a) Benin



(b) Cameroon



(c) Mali

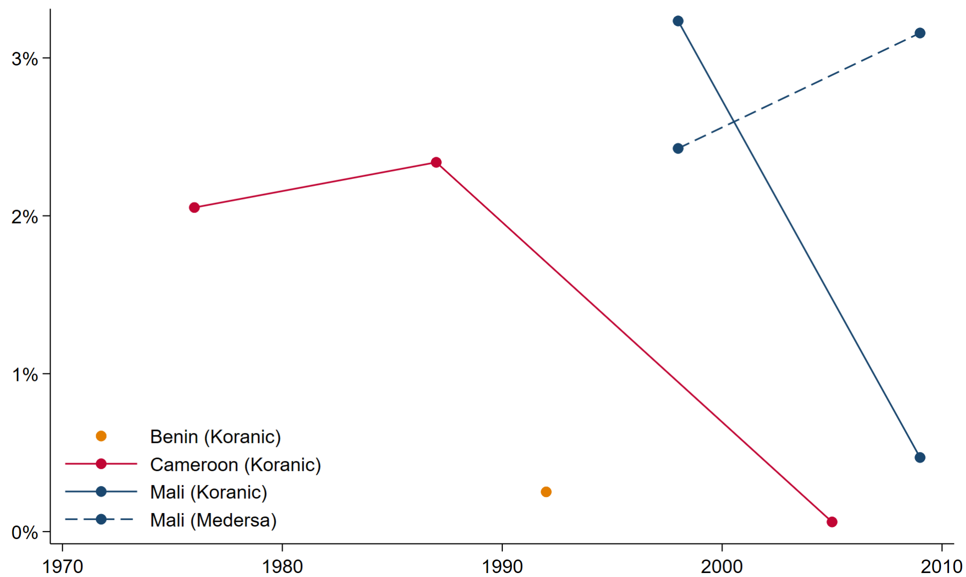


Notes: This figure shows the distribution of primary completion rates in each locality in each country in each census year of those between the ages of 15 and 20. Although primary schooling officially ends at 12 in each country over the time period I study, many children only complete in the years following, and thus I take 15 to be when most who will complete have done so. I cap at 20 in an attempt to capture more recent dynamics, and to remove mechanical correlation across censuses by resampling the same individuals.

L Koranic schools and Medersas

Koranic schools are a traditional method of education which involves memorising and reciting the Koran. They remain popular in many Muslim countries, and often offer a cheaper or more local method of schooling. In this paper, I don't count those who have solely had a Koranic education as having completed primary school, in line with the classification used by IPUMSi. Although these schools primarily concern themselves with memorizing and reciting the Koran, it maybe that they provide some opportunities to those who complete a course at them, and therefore this may be an important dimension I am missing from this analysis — or in the case where students switch from Koranic to state-sponsored schooling, I may be overstating the impact. Fortunately, in some of the censuses used, I can distinguish between those at a Koranic school and those at a secular school. Figure L.1 plots the proportion of 6 to 14 year old's in each census where data is available, who are at a Koranic school. It's clear from figure L.1 that Koranic education is in the vast minority (maximum 3% of children) and appears to be declining further. It's likely that many more students attend Koranic schools in the evening or on weekends in addition to attending state school, but this dimension is not covered in the data and is less consequential. Figure L.1 also shows the proportion of students in Mali at a Medersa (Boyle, 2014), which is a religious school in Mali that follows the national curriculum, toeing the line between Koranic schools and state schools. These schools are on the rise, but still constitute a very small proportion of the overall education system.

Figure L.1 Proportion of children enrolled in Koranic schools or Medersas



Notes: This figure shows the proportion of primary school aged children (6 to 14) who report attending a Koranic school or a Medersa in the Census.

M Theory Appendix

We can solve the model into one of three endogenous variables in three equations in each location in each period. These three endogenous variables are (1) the stock of human capital in a given location H_{it}, (2) market access MA_{it} and (3) goods market access TMA_{it}.

The stock of human capital is given by $H_i = \sum_j L_{jt-1}\pi_{ijt}\mathbb{E}[e_j^\eta]$. Note that as $e_j = \left(\gamma \text{MA}_j^{1/\theta}(\tau_j^{educ})^{-1}\right)^{1/(1-\eta)} \xi^{-1/(1-\eta)}$ we have that $\mathbb{E}[e_j^\eta] = \phi_\xi \left((\tau_j^{educ})^{-1}\text{MA}_{jt}^{1/\theta}\right)^{\eta/(1-\eta)}$ where $\phi_\xi = \gamma\mathbb{E}[\xi^{-\eta/(1-\eta)}]$. Into this expression, we can substitute our expression for migration probabilities π_{ijt} and, simplifying, find the following.

$$H_{it} = \phi_\xi a_{it}^\theta \left(\frac{r_{it}}{P_{it}}\right)^\theta \sum_j L_{jt-1}(\tau_{ijt}^m)^{-\theta}(\tau_{jt}^{educ})^{\frac{-\eta}{1-\eta}}\text{MA}_{jt}^{\frac{\eta-\theta(1-\eta)}{\theta(1-\eta)}} \tag{M.1}$$

Into this expression we can substitute our known expressions for real wages and prices $r_{it} = \left(\frac{\text{TMA}_{it}}{H_{it}}\right)^{1/\sigma} B_{it}^{\frac{\sigma-1}{\sigma}}$, and $P_{it}^{1-\sigma} = \text{TMA}_{it}$ to find the following.

$$H_{it}^{1+\frac{\theta}{\sigma}} = \phi_\xi a_{it}^{\theta} B_{it}^{\frac{\theta(\sigma-1)}{\sigma}} \text{TMA}_{it}^{\frac{\theta(1-2\sigma)}{\sigma(1-\sigma)}} \sum_j L_{jt-1}(\tau_{ijt}^{m})^{-\theta}(\tau_{jt}^{educ})^{\frac{-\eta}{1-\eta}} \text{MA}_{jt}^{\frac{\eta-\theta(1-\eta)}{\theta(1-\eta)}} \tag{M.2}$$

Finally, subbing in our expressions for real wages into market access and goods market access terms we can write.

$$\text{MA}_{it} = \sum_j \left(a_{jt} B_{jt}^{\frac{\sigma-1}{\sigma}} \text{TMA}_{jt}^{\frac{1-2\sigma}{\sigma(1-\sigma)}} H_{jt}^{-1/\sigma} (\tau_{ijt}^{m})^{-1} \right)^{\theta} \tag{M.3}$$

$$\text{TMA}_{it} = \sum_j \left(\tau_{ijt}^{t} \left(\frac{\text{TMA}_{jt}}{H_{jt} B_{jt}} \right)^{1/\sigma} \right)^{1-\sigma} \tag{M.4}$$

Equations M.2, M.3, and M.4 define three equations in three unknowns for each location and can jointly be solved to determine all endogenous variables. We can write the resulting system in terms of changes using the exact hat algebra approach.

$$\widehat{H}_{it}^{1+\frac{\theta}{\sigma}} = \widehat{\text{TMA}}_{it}^{\delta} \cdot \sum_j \omega_{ijt} \cdot \left(\widehat{\tau}_{ijt}^{m}\right)^{-\theta} \cdot \widehat{\text{MA}}_{jt}^{\lambda}$$

$$\widehat{\text{TMA}}_{it} = \sum_j \rho_{ijt} \cdot \left(\widehat{\tau}_{ijt}^{t}\right)^{1-\sigma} \cdot \widehat{\text{TMA}}_{jt}^{\frac{1-\sigma}{\sigma}} \cdot \widehat{H}_{jt}^{\frac{\sigma-1}{\sigma}}$$

$$\widehat{\text{MA}}_{it} = \sum_j \pi_{ijt} \cdot \left(\widehat{\tau}_{ijt}^{m}\right)^{-\theta} \cdot \widehat{\text{TMA}}_{jt}^{\delta} \cdot \widehat{H}_{jt}^{-\frac{\theta}{\sigma}}$$

Where $\lambda = \frac{\eta-\theta(1-\eta)}{\theta(1-\eta)}$ and $\delta = \frac{\theta(2\sigma-1)}{\sigma(\sigma-1)}$. Shares are defined as normal.

$$\text{Human capital share:} \quad \omega_{ijt} = \frac{H_{ijt}}{H_{it}} = \frac{L_{jt-1}(\tau_{jit}^{m})^{-\theta}\text{MA}_{jt}^{\lambda}}{\sum_q L_{qt-1}(\tau_{iqt}^{m})^{-\theta}\text{MA}_{qt}^{\lambda}}$$

$$\text{Trade share:} \quad \rho_{ijt} = \frac{X_{ijt}}{X_{it}} = \frac{(\tau_{jit}^{t})^{1-\sigma}\text{TMA}_{jt}^{(1-\sigma)/\sigma} H_{jt}^{(\sigma-1)/\sigma}}{\sum_q (\tau_{iqt}^{t})^{1-\sigma}\text{TMA}_{qt}^{(1-\sigma)/\sigma} H_{qt}^{(\sigma-1)/\sigma}}$$

$$\text{Migration share:} \quad \pi_{ijt} = \frac{L_{ijt}}{L_{it}} = \frac{(\tau_{jit}^{m})^{-\theta}\text{TMA}_{jt}^{\delta} H_{jt}^{-\theta/\sigma}}{\sum_q (\tau_{iqt}^{m})^{-\theta}\text{TMA}_{qt}^{\delta} H_{qt}^{-\theta/\sigma}}$$

We observed L_{ijt} and so π_{ijt} directly in the data. Using data on L_{it}, estimated market access terms, and estimated iceberg migration cost, I can calculate ω_{ijt}. Finally, given ω_{ijt}, I can back out H_{it}, combine with estimates of iceberg trade costs, and calculate TMA_{it} to find

ρ_{ijt}. To calculate TMA_{it}, I need data on output at the location-year level, Y_{it}. I find this by taking an Engel curve approach following Young (2012) as described in appendix F.

This system of equations falls within the category described in Allen et al. (2024) and therefore, existence and uniqueness can be proved by relying on their results and is given when $\rho(|A|) < 1$ where $A = B\Gamma^{-1}$ and

$$\Gamma = \begin{pmatrix} 1+\frac{\theta}{\sigma} & -\delta & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 0 & 0 & \lambda \\ \frac{\sigma-1}{\sigma} & \frac{1-\sigma}{\sigma} & 0 \\ -\frac{\theta}{\sigma} & \delta & 0 \end{pmatrix}$$

For the estimated parameter values, I find that $\rho(|A|) < 1$. Given parameter estimates σ, θ, η and estimates of the shares $\{\omega_{ijt}, \rho_{ijt}, \pi_{ijt}\}$ I can then solve for the endogenous variables for any given set of shocks to transport costs τ^{t}_{ijt} and τ^{m}_{ijt}.

Appendix References

African Development Bank (2009). *Benin—Djougou–N'Dali Road Upgrade Project—Project Completion Report*. Tech. rep. African Development Bank. URL: https://www.afdb.org/en/documents/document/benin-djougou-ndali-road-upgrade-project-project-completion-report-pcr-25873.

African Development Bank (n.d.). *The Quest to Integrate Africa: Cross-Border Road Corridors*. Tech. rep. African Development Bank. URL: https://www.afdb.org/fileadmin/uploads/afdb/Documents/Publications/Cross-border_road_corridors.pdf.

African Development Fund (2016). *STUDY FOR THE ABIDJAN – LAGOS CORRIDOR HIGHWAY DEVELOPMENT PROJECT*. Tech. rep. African Development Bank. URL: https://www.afdb.org/fileadmin/uploads/afdb/Documents/Project-and-Operations/Multinational_-_Approved_-_Study_for_the_Abidjan-_Lagos_Corridor_Highway_Development_Project.pdf.

Allen, Treb, Costas Arkolakis, and Xiangliang Li (2024). "On the equilibrium properties of spatial models". In: *American Economic Review: Insights* 6.4, pp. 472–489.

Alpern, Stanley B. (1999). "Dahomey's Royal Road". In: *History in Africa* 26, pp. 11–24. DOI: 10.2307/3172135.

Basedau, Matthias and Alexander Stroh (2012). "How ethnic are African parties really? Evidence from four Francophone countries". In: *International Political Science Review* 33.1, pp. 5–24.

Basedau, Matthias et al. (2011). "Ethnicity and party preference in sub-Saharan Africa". In: *Democratization* 18.2, pp. 462–489.

Battle, Martin and Jennifer C Seely (2010). "It's All Relative: Modeling Candidate Support in Benin". In: *Nationalism and Ethnic Politics* 16.1, pp. 42–66.

Blimpo, Moussa P, Robin Harding, and Leonard Wantchekon (2013). "Public investment in rural infrastructure: Some political economy considerations". In: *Journal of African Economies* 22.suppl_2, pp. ii57–ii83.

Boyle, Helen N (2014). "Between Secular Public Schools And Qur'anic Private Schools: The Growing Educational Presence Of Malian Medersas". In: *The Review of Faith & International Affairs* 12.2, pp. 16–26.

Burgess, Robin et al. (2015). "The value of democracy: evidence from road building in Kenya". In: *American Economic Review* 105.6, pp. 1817–51.

Canning, David and Peter Pedroni (2008). "Infrastructure, long-run economic growth and causality tests for cointegrated panels". In: *The Manchester School* 76.5, pp. 504–527.

Dunning, Thad and Lauren Harrison (2010). "Cross-cutting cleavages and ethnic voting: An experimental study of cousinage in Mali". In: *American Political Science Review* 104.1, pp. 21–39.

Franck, Raphael and Ilia Rainer (2012). "Does the leader's ethnicity matter? Ethnic favoritism, education, and health in sub-Saharan Africa". In: *American Political Science Review* 106.2, pp. 294–325.

Fujiwara, Thomas and Leonard Wantchekon (2013). "Can informed public deliberation overcome clientelism? Experimental evidence from Benin". In: *American Economic Journal: Applied Economics* 5.4, pp. 241–55.

Independent Evaluation Group (2018). *Burkina Faso, Ghana, and Mali: West Africa Transport and Transit Facilitation Project*. Tech. rep. World Bank. URL: https://ieg.worldbankgroup.org/reports/burkina-faso-ghana-and-mali-west-africa-transport-and-transit-facilitation-project.

International Development Association (1981a). "Mali – Road Maintenance Project: Credit 1104 – Credit Agreement – Conformed". In.

International Development Association (1981b). *Mali—Road Maintenance Project: Credit 1104—Credit Agreement—Conformed*. Tech. rep. Credit No. 1104 MLI. World Bank. URL: https://documents.worldbank.org/en/publication/documents-reports/documentdetail/450261468053091074/.

Mohan, P. C. (2003). *Benin—Transport Sector Investment Program*. Africa Region Findings & Good Practice Infobriefs 95. World Bank. URL: https://openknowledge.worldbank.org/handle/10986/9707.

Tchindjang, Mesmin et al. (2005). "Mapping of Main Artery Network and Colonial Porterage in Cameroon (1900–1940)". In: *ICC 2005: Proceedings of the 22nd International Carto-*

graphic Conference — Mapping Approaches into a Changing World. A Coruña, Spain: International Cartographic Association (ICA).

Wantchekon, Leonard (2003). "Clientelism and voting behavior: Evidence from a field experiment in Benin". In: *World politics* 55.3, pp. 399–422.

Weidmann, Nils B, Jan Ketil Rød, and Lars-Erik Cederman (2010). "Representing ethnic groups in space: A new dataset". In: *Journal of Peace Research* 47.4, pp. 491–499.

World Bank (1970). *Dahomey – Highway Maintenance and Engineering Project: Credit 0215 – Credit Agreement – Conformed*.

World Bank (1980). *Project Performance Audit Report: Cameroon Second and Third Highway Projects*. Tech. rep. Covers Loan 935-CM, Credit 429-CM (incl. supplementary financing), and Loan 1515-CM. World Bank. URL: http://documents1.worldbank.org/curated/en/971521507545682686/pdf/120305-Project-completion-report-ADD-Board-number-R1981-0015-ADD-BOTH-P000324-and-P000343-PUBLIC.pdf.

World Bank (1982). *Project Performance Audit Report: Cameroon Second and Third Highway Projects*.

World Bank (1983). *Mali—Second Highway Project: Project Performance Audit Report*. Tech. rep. World Bank. URL: https://documents1.worldbank.org/curated/en/743531468052445204/txt/multi0page.txt.

World Bank (1987). *Benin—Transport Infrastructure Rehabilitation and Maintenance Project*. Staff Appraisal Report 6000-BEN. World Bank. URL: https://documents1.worldbank.org/curated/en/903701468013151064/pdf/multi-page.pdf.

World Bank (Mar. 1994). *Republic of Benin: Transport Sector Strategy*. Report 12098-BEN. Confidential; prepared jointly by the Ministry of Public Works and Transport (MTPT) and the World Bank. World Bank. URL: https://documents1.worldbank.org/curated/en/101361680676614496/pdf/Benin-Transport-Sector-Strategy.pdf (visited on 01/28/2026).

World Bank (2010a). *New Road Paves the Way for Regional Integration in Central Africa*. CEMAC Transport-Transit Facilitation Project (P079736). URL: https://www.worldbank.org/en/news/feature/2010/07/16/new-road-paves-the-way-for-regional-integration-in-central-africa.

World Bank (2010b). *Regional Integration: Benin Receives US$75 Million to Facilitate Trade and Improve its Road Infrastructure*. Abidjan–Lagos Transport and Transit Facilitation Project. URL: https://www.worldbank.org/en/news/feature/2010/04/08/regional-integration-benin-receives-us75-million-to-facilitate-trade-and-improve-its-road-infrastructure.

World Bank (2014). *WB Helps Cameroon Rebuild Roads and Railways, Boosting Access to Clinics, Schools and Markets*. URL: https://www.worldbank.org/en/news/

press-release/2014/05/27/wb-helps-cameroon-rebuild-roads-and-railways-boosting-access-to-clinics-schools-and-markets.

World Bank (2025). *The World Bank Supports the Connectivity and Resilience of Road Infrastructure in Mali*. URL: https://www.worldbank.org/en/news/press-release/2025/05/28/world-bank-supports-the-connectivity-and-resilience-of-road-infrastructure-in-mali.

Young, Alwyn (2012). "The African growth miracle". In: *Journal of Political Economy* 120.4, pp. 696–739.