Moving Opportunity

Local Connectivity and Spatial Inequality

Luke Heath Milsom

November 8, 2022

luke.milsom@economics.ox.ac.uk

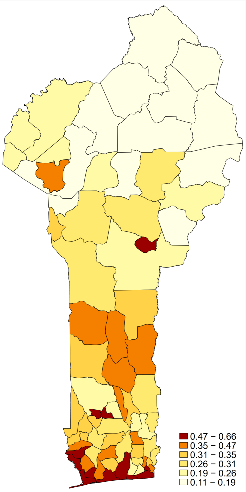
Within-country inequality is a pervasive problem globally

FINANCIAL TIMES myFT

UK north-south divide + Add to myFT

UK's regional inequality one of worst in developed world

Health, jobs, disposable income and productivity more polarised, says study



Observed primary education completion, Benin 2013

Within-country differences across space represent possibilities to reduce inequality

- Such inequalities are potentially costly and unfair, but could be leveraged to improve outcomes and reduce inequality.

- The literature has focused on the policy response of moving people to opportunity e.g. *Chetty, Hendren, and Katz (2016)*, and, *Bryan, Chowdhury, and Mobarak (2014)*.

Within-country differences across space represent possibilities to reduce inequality

- Such inequalities are potentially costly and unfair, but could be leveraged to improve outcomes and reduce inequality.

- The literature has focused on the policy response of moving people to opportunity e.g. *Chetty, Hendren, and Katz (2016)*, and, *Bryan, Chowdhury, and Mobarak (2014)*.

- Assumes the causal effect of place is immutable.

Within-country differences across space represent possibilities to reduce inequality

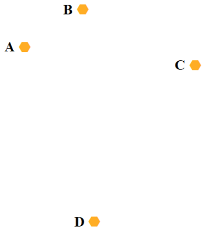
- Such inequalities are potentially costly and unfair, but could be leveraged to improve outcomes and reduce inequality.
- The literature has focused on the policy response of moving people to opportunity e.g. *Chetty, Hendren, and Katz (2016)*, and, *Bryan, Chowdhury, and Mobarak (2014)*.
- Assumes the causal effect of place is immutable.
- In this paper I exploit the alternative possibility of moving opportunity to people.

Within-country differences across space represent possibilities to reduce inequality

- Such inequalities are potentially costly and unfair, but could be leveraged to improve outcomes and reduce inequality.
- The literature has focused on the policy response of moving people to opportunity e.g. *Chetty, Hendren, and Katz (2016)*, and, *Bryan, Chowdhury, and Mobarak (2014)*.
- Assumes the causal effect of place is immutable.
- In this paper I exploit the alternative possibility of moving opportunity to people.
- Focus on a large (10% WB lending, 2021) and salient policy in a dev. context: Road building.

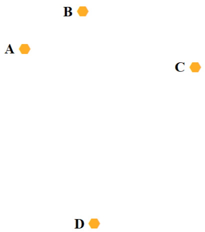
What determines the evolution of the distribution of place effects?

What determines the evolution of the distribution of place effects?



- Suppose outcomes are higher in **B** than **D**, why?

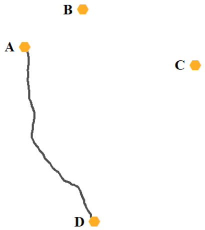
What determines the evolution of the distribution of place effects?



- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  $\rightarrow$ Opportunity is higher in **B**.

What determines the evolution of the distribution of place effects?

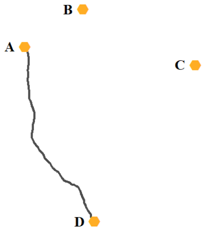


- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  $\rightarrow$ Opportunity is higher in **B**.

- What happens if **D** and **A** become connected?

What determines the evolution of the distribution of place effects?

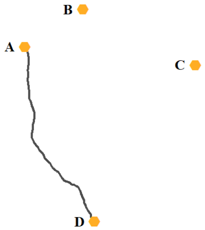


- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  → Opportunity is higher in **B**.

- What happens if **D** and **A** become connected?
  → **D** gains

What determines the evolution of the distribution of place effects?

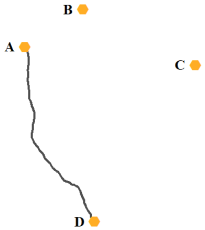


- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  $\rightarrow$ Opportunity is higher in **B**.

- What happens if **D** and **A** become connected?
  $\rightarrow$ **D** gains, **B** loses

What determines the evolution of the distribution of place effects?

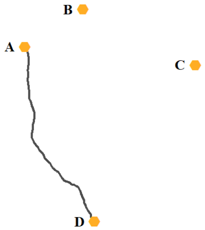


- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  $\rightarrow$ Opportunity is higher in **B**.

- What happens if **D** and **A** become connected?
  $\rightarrow$ **D** gains, **B** loses, **C**?

What determines the evolution of the distribution of place effects?

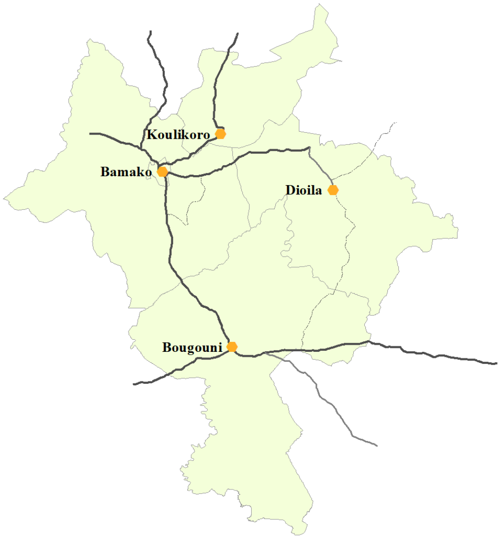


- Suppose outcomes are higher in **B** than **D**, why?
- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  $\rightarrow$ Opportunity is higher in **B**.
- What happens if **D** and **A** become connected?
  $\rightarrow$ **D** gains, **B** loses, **C**?

Research Question

How does road building alter spatial inequality of opportunity?

What determines the evolution of the distribution of place effects?



- Suppose outcomes are higher in **B** than **D**, why?

- Position within the spatial network matters.
  Suppose **A** has a high demand for non-agricultural goods.
  → Opportunity is higher in **B**.

- What happens if **D** and **A** become connected?
  → **D** gains, **B** loses, **C**?

Research Question

How does road building alter spatial inequality of opportunity?

Estimating the effect of changes in the road network on spatial inequality of opportunity

Estimating the effect of changes in the road network on spatial inequality of opportunity

$$\text{Road network} \longrightarrow \quad \text{MA}_i \quad \longrightarrow \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$

Estimating the effect of changes in the road network on spatial inequality of opportunity

$$\text{Road network} \longrightarrow \quad \text{MA}_i \quad \xrightarrow{\textbf{A}} \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$

A: Develop a sufficient statistic result.

- **Challenge 1**: Spillovers & general equilibrium effects.

- Within a broad class of models, market access terms capture all the effect of roads on opportunity.

Estimating the effect of changes in the road network on spatial inequality of opportunity

$$\text{Road network} \longrightarrow \begin{array}{c} \text{MA}_i \\ \uparrow_{\textbf{B}} \\ \text{MA}_i^{IV} \end{array} \xrightarrow{\textbf{A}} \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$

A: Develop a sufficient statistic result.

B: Novel identification strategy using not-on-least-cost-path variation.

- **Challenge 2**: Endogeneity of road placement.
- Use indirect variation in MA terms coming through induced changes in other locations MA.
- Overcome endogenous exposure (*Borusyak and Hull, 2021*) by permuting over road building.

Estimating the effect of changes in the road network on spatial inequality of opportunity

$$\begin{tikzcd} \text{Road network} \arrow[r] & \text{MA}_i \arrow[r, "A"] & \left\{\begin{array}{c}\text{Returns to education}\\ \text{Cost of education}\end{array}\right\} \arrow[r] & \mu_i \arrow[ll, bend right, "C"'] \\ & \text{MA}_i^{IV} \arrow[u, "B"'] \end{tikzcd}$$

A: Develop a sufficient statistic result.

B: Novel identification strategy using not-on-least-cost-path variation.

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

- **Challenge 3**: Network level characteristics matter.
- Parametize the model using the estimated sufficient statistic results.
- Impact of roads built since 1970, effect of future road building + how the network matters.

Estimating the effect of changes in the road network on spatial inequality of opportunity

$$
\begin{tikzcd}
\text{Road network} \arrow[r] & \mathrm{MA}_i \arrow[r, "A"] & \left\{\begin{array}{c}\text{Returns to education}\\ \text{Cost of education}\end{array}\right\} \arrow[r] & \mu_i \arrow[ll, bend right, "C"'] \\
 & \mathrm{MA}_i^{IV} \arrow[u, "B"'] & &
\end{tikzcd}
$$

A: Develop a sufficient statistic result.

B: Novel identification strategy using not-on-least-cost-path variation.

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

Summary of results & contributions

A: **Develop a sufficient statistic result.**

B: **Novel identification strategy using not-on-least-cost-path variation.**

C: **Estimate a structural general equilibrium spatial economics model to study counterfactuals.**

Summary of results & contributions

A: **Develop a sufficient statistic result.**

Literature: Quantitative spatial economic modeling.

B: **Novel identification strategy using not-on-least-cost-path variation.**

C: **Estimate a structural general equilibrium spatial economics model to study counterfactuals.**

Summary of results & contributions

A: **Develop a sufficient statistic result.**

Literature: Quantitative spatial economic modeling.

Contribution: Labor market access and goods market access are a suff. stat. for the effect of roads on opportunity.

Morten & Oliveira (2021); Donaldson & Hornbeck (2016); Redding & Rossi-Hansberg (2017); Allen et al. (2020); Franklin et al. (2022); Tsivanidis (2022); Ahlfeldt et al. (2015); Donaldson (2018).

B: **Novel identification strategy using not-on-least-cost-path variation.**

C: **Estimate a structural general equilibrium spatial economics model to study counterfactuals.**

Summary of results & contributions

A: **Develop a sufficient statistic result.**

Literature: Quantitative spatial economic modeling.

Contribution: Labor market access and goods market access are a suff. stat. for the effect of roads on opportunity.

Morten & Oliveira (2021); Donaldson & Hornbeck (2016); Redding & Rossi-Hansberg (2017); Allen et al. (2020); Franklin et al. (2022); Tsivanidis (2022); Ahlfeldt et al. (2015); Donaldson (2018).

B: **Novel identification strategy using not-on-least-cost-path variation.**

Literature: Identifying the impacts of transport networks.

C: **Estimate a structural general equilibrium spatial economics model to study counterfactuals.**

Summary of results & contributions

A: Develop a sufficient statistic result.

Literature: Quantitative spatial economic modeling.

Contribution: Labor market access and goods market access are a suff. stat. for the effect of roads on opportunity.

Morten & Oliveira (2021); Donaldson & Hornbeck (2016); Redding & Rossi-Hansberg (2017); Allen et al. (2020); Franklin et al. (2022); Tsivanidis (2022); Ahlfeldt et al. (2015); Donaldson (2018).

B: Novel identification strategy using not-on-least-cost-path variation.

Literature: Identifying the impacts of transport networks.

Contribution: Roads do indeed alter the spatial distribution of opportunity.

Jedwab & Storeygard (2021); Gonzalez-Navarro & Quintana-Domeque (2016); Moneke (2021); Faber (2014); Banerjee et al. (2020); Donaldson (2018); Adukia et al. (2020); Asher & Novosad (2020).

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

Summary of results & contributions

A: Develop a sufficient statistic result.

Literature: Quantitative spatial economic modeling.

Contribution: Labor market access and goods market access are a suff. stat. for the effect of roads on opportunity.

Morten & Oliveira (2021); Donaldson & Hornbeck (2016); Redding & Rossi-Hansberg (2017); Allen et al. (2020); Franklin et al. (2022); Tsivanidis (2022); Ahlfeldt et al. (2015); Donaldson (2018).

B: Novel identification strategy using not-on-least-cost-path variation.

Literature: Identifying the impacts of transport networks.

Contribution: Roads do indeed alter the spatial distribution of opportunity.

Jedwab & Storeygard (2021); Gonzalez-Navarro & Quintana-Domeque (2016); Moneke (2021); Faber (2014); Banerjee et al. (2020); Donaldson (2018); Adukia et al. (2020); Asher & Novosad (2020).

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

Literature: Place effects literature.

Summary of results & contributions

A: Develop a sufficient statistic result.

Literature: Quantitative spatial economic modeling.

Contribution: Labor market access and goods market access are a suff. stat. for the effect of roads on opportunity.

Morten & Oliveira (2021); Donaldson & Hornbeck (2016); Redding & Rossi-Hansberg (2017); Allen et al. (2020); Franklin et al. (2022); Tsivanidis (2022); Ahlfeldt et al. (2015); Donaldson (2018).

B: Novel identification strategy using not-on-least-cost-path variation.

Literature: Identifying the impacts of transport networks.

Contribution: Roads do indeed alter the spatial distribution of opportunity.

Jedwab & Storeygard (2021); Gonzalez-Navarro & Quintana-Domeque (2016); Moneke (2021); Faber (2014); Banerjee et al. (2020); Donaldson (2018); Adukia et al. (2020); Asher & Novosad (2020).

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

Literature: Place effects literature.

Contribution: Initial network conditions determine the impact of changes to the network since 1970.

Chetty & Hendren (2018); Deutscher (2020); Alesina et al. (2021); Laliberte (2021); Heath Milsom (2021); Chyn & Daruich (2022); Eckert & Kleineberg (2021); Atkin (2016); Hsiao (2022).

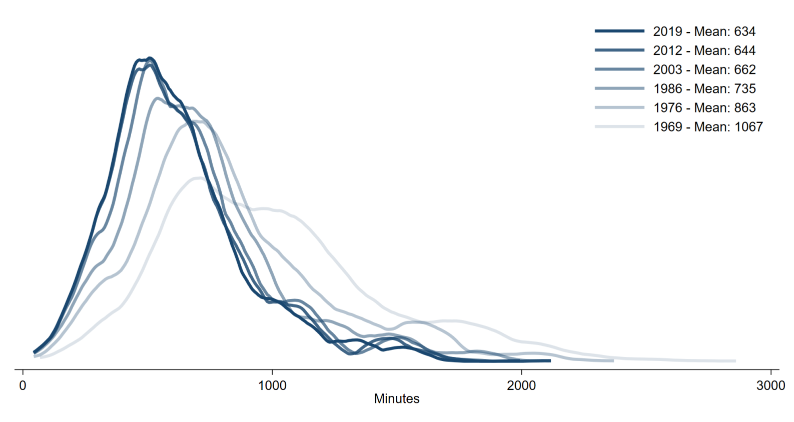
Data and descriptive evidence

Data and setting: Benin, Cameroon, and Mali 1970-2020

- Road data: I digitize historical Michelin road maps 1970-2020.

Data and setting: Benin, Cameroon, and Mali 1970-2020

- Road data: I digitize historical Michelin road maps 1970-2020.
    - Since 1970 travel times have fallen by on average 33%.



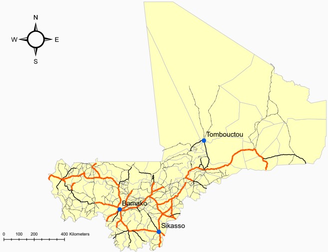
Data and setting: Benin, Cameroon, and Mali 1970-2020

- Road data: I digitize historical Michelin road maps 1970-2020.
  - Since 1970 travel times have fallen by on average 33%.

- Study three countries as...

  ... data prevents including more.

  ... can identify the importance of network level characteristics.



Data and setting: Benin, Cameroon, and Mali 1970-2020

- Road data: I digitize historical Michelin road maps 1970-2020.
  - Since 1970 travel times have fallen by on average 33%.

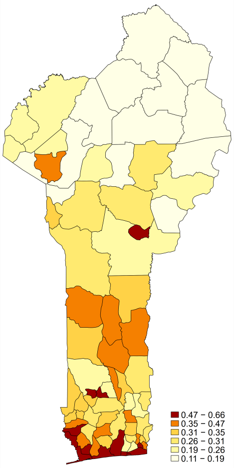
- Study three countries as...

  ... data prevents including more.

  ... can identify the importance of network level characteristics.

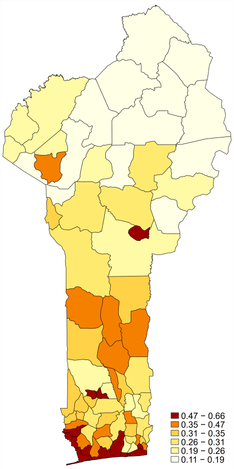
- Opportunity data: From *Heath Milsom (2022)*
  - Second administrative division level.
  - Estimates for 334 locality-year cells.

My setting: Local educational opportunity in Benin, Cameroon, and Mali

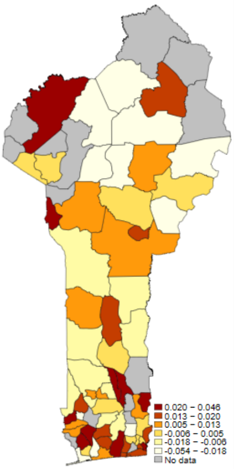


Observed primary education completion, Benin 2013

My setting: Local educational opportunity in Benin, Cameroon, and Mali



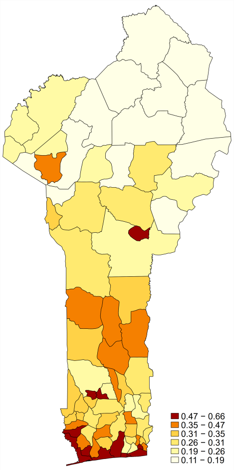
Observed primary education completion, Benin 2013



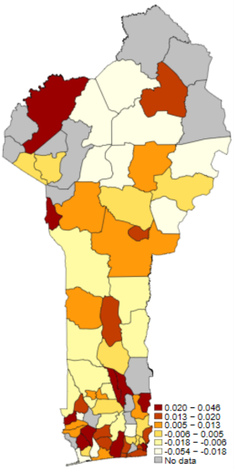
Causal place effects (Heath Milsom, 2022).

- This paper: Opportunity, μ_{it} = the causal effect of growing up in a location on the probability of completing primary school.
- Estimated in *Heath Milsom (2022)* using a movers design à-la *Chetty and Hendren (2018)*. Related to opportunity generally.

My setting: Local educational opportunity in Benin, Cameroon, and Mali



Observed primary education completion, Benin 2013



Causal place effects (Heath Milsom, 2022).

- This paper: Opportunity, μ_{it} = the causal effect of growing up in a location on the probability of completing primary school.
- Estimated in *Heath Milsom (2022)* using a movers design à-la *Chetty and Hendren (2018)*. Related to opportunity generally.
- Estimate: $y_\omega = \alpha_{odt} + \mu_{it} \cdot Exposure_{\omega i} + \varepsilon_\omega$
- Why should we care about the spatial distribution of local educational opportunity?

Data and setting: Benin, Cameroon, and Mali 1970-2020

- Road data: I digitize historical Michelin road maps 1970-2020.
  - Since 1970 travel times have fallen by on average 33%.

- Study three countries as...

  ... data prevents including more.

  ... can identify the importance of network level characteristics.

- Opportunity data: From *Heath Milsom (2022)*
  - Second administrative division level.
  - Estimates for 334 locality-year cells.

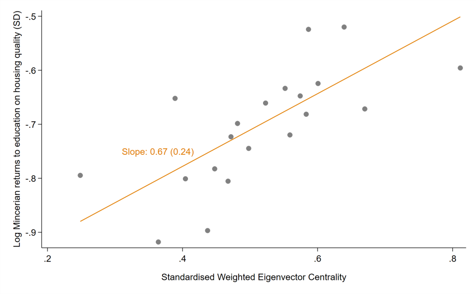
- Other data: Population, assets, migration histories from Censuses, wages from DHS, etc...

Increasing centrality is correlated with higher returns to educ which is correlated with higher μ

$$\text{Road network} \longrightarrow \quad \text{MA}_i \quad \longrightarrow \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$

Increasing centrality is correlated with higher returns to educ which is correlated with higher μ

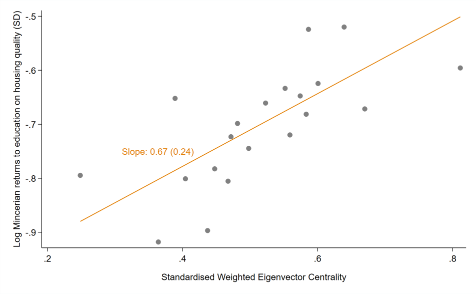
$$\text{Road network} \longrightarrow \text{MA}_i \longrightarrow \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$



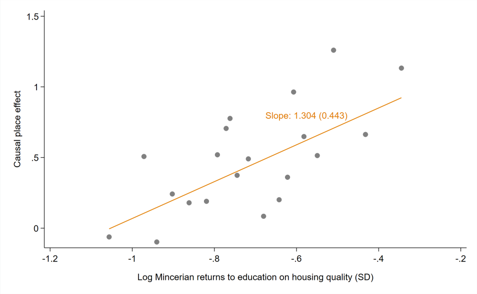
Centrality $\rightarrow$ Returns to Education

Increasing centrality is correlated with higher returns to educ which is correlated with higher μ

$$\text{Road network} \longrightarrow \quad MA_i \quad \longrightarrow \left\{ \begin{array}{c} \text{Returns to education} \\ \text{Cost of education} \end{array} \right\} \longrightarrow \mu_i$$



Centrality $\rightarrow$ Returns to Education



Returns to education $\rightarrow \mu$

A: The sufficient statistic result

Sufficient statistic result: Motivation

- **Challenge 1**: Spillovers & general equilibrium effects.

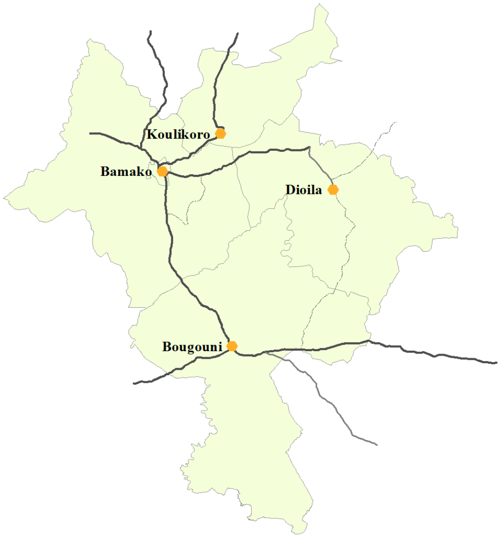
Sufficient statistic result: Motivation

- **Challenge 1**: Spillovers & general equilibrium effects.

- Market access terms act as a sufficient statistic (*Donaldson (2022)*) to describe the impact of changes in connectivity on the spatial distribution of local educational opportunity across a broad class of modeling specifications and assumptions.

- Modeling specifications: Armington *Allen and Arkolakis (2014)*, Eaton-Kortum *Bartelme (2015)*, Melitz *Di Giovanni and Levchenko (2013)*, Monopolistic competition *Krugman (1980)*, Quantitative spatial economic models *Tsivanidis (2019), Ahlfeldt et al. (2015)*.

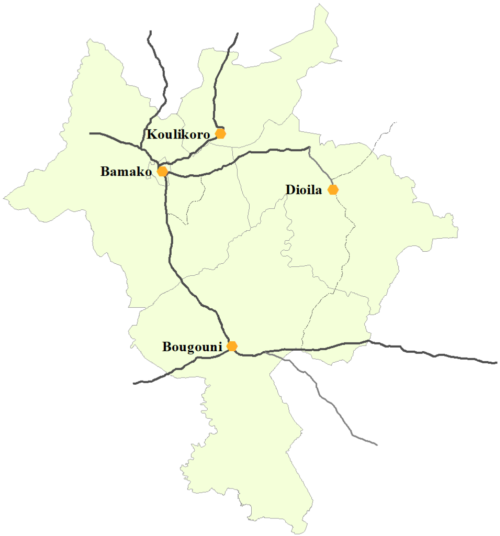
- Modeling assumptions: Can include land in production, land in consumption, endogenous land/housing, explicit agglomeration forces, endogenous amenities, generalize preferences, intermediate goods, and allow other factors to influence education choice.



Goods market

- Iceberg trade costs τ_{ij} so $p_{ij} = \tau_{ij} p_j$.
- CES aggregate demand $E_i = \left(\sum_j p_{ij}^{-\phi}\right)^{-1/\phi}$
  $\Rightarrow$ via Shephard's Lemma, gravity trade:
  $X_{ij} = \frac{(p_{ij})^{-\phi}}{\sum_k p_{jk}^{-\phi}} Y_j$.
- Goods market clears: $Y_i = E_i = \sum_k X_{ik}$.
- $Y_i = p_i^{-\phi} MA_i$, where $MA_i = P_i^{-\phi}$.

Sufficient statistic result: Labor Market



Labor market

- Iceberg movement costs κ_{ij} so $u_{ij} = \frac{1}{\kappa_{ij}} u_j$. Reflect more than just the pecuniary cost *Gollin, Kirchberger, Blanchard (WP); Bailey et al. (2018).*
- Gravity movement equation: $M_{ij} = \frac{u_{ij}^{\lambda}}{\sum_k u_{ik}^{\lambda}} L_j$.
- Labor markets clear $L_i = \sum_k M_{ik}$.
- $L_i = u_i^{\lambda} LMA_i$ where $LMA_i = \left(\sum_k u_{ik}^{\lambda}\right)^{1/\lambda}$.

Sufficient statistic result: Adding education

- Together: $w_i = \frac{Y_i}{L_i} = \frac{p_i^{-\phi} MA_i}{u_i^{\lambda} LMA_i} = \Omega_i \cdot MA_i^{\alpha_1} \cdot LMA_i^{\alpha_2}$

- Add two sectors/ types: educated and non-educated workers can only work in E or N sector.
    - Now wages depend on E and N market access and labor market access terms.

- Model education as part of a locations utility shifter $A_i = \mathcal{A}_i \mathcal{E}_i$, where $\mathcal{E}_i = (r_i)^{\beta} = (w_i^E / w_i^N)^{\beta}$.

- Then local opportunity, ($\mathbb{P}$[complete primary school$_i$]), is increasing in the local returns to education.

Sufficient statistic result: Adding education

- Together: $w_i = \frac{Y_i}{L_i} = \frac{p_i^{-\phi} MA_i}{u_i^{\lambda} LMA_i} = \Omega_i \cdot MA_i^{\alpha_1} \cdot LMA_i^{\alpha_2}$
- Add two sectors/ types: educated and non-educated workers can only work in E or N sector.
  - Now wages depend on E and N market access and labor market access terms.
- Model education as part of a locations utility shifter $A_i = \mathcal{A}_i \mathcal{E}_i$, where $\mathcal{E}_i = (r_i)^{\beta} = (w_i^E / w_i^N)^{\beta}$.
- Then local opportunity, ($\mathbb{P}[\text{complete primary school}_i]$), is increasing in the local returns to education.

Resulting sufficient statistic relationship

$$\mu_{it} = \gamma_1 \cdot \ln(MA_{it}^E) + \gamma_2 \cdot \ln(MA_{it}^N) + \gamma_3 \cdot \ln(LMA_{it}^E) + \gamma_4 \cdot \ln(LMA_{it}^N) + v_{it}$$

Sufficient statistic result: Intuition

Resulting sufficient statistic relationship

$$\mu_{it} = \gamma_1 \cdot \ln(MA^E_{it}) + \gamma_2 \cdot \ln(MA^N_{it}) + \gamma_3 \cdot \ln(LMA^E_{it}) + \gamma_4 \cdot \ln(LMA^N_{it}) + v_{it}$$

- Increases in MA^E_{it} increase demand for E-type goods, and therefore w^E_{it} and so RtE rises.

$\Rightarrow \gamma_1 > 0$

- Increases in MA^N_{it} increase demand for N-type goods, and therefore w^N_{it} and so RtE falls.

$\Rightarrow \gamma_2 < 0$

Coefficient interpretation details.

What is not consistent with the SS result.

Education supply.

Maps of market access variation.

B: Identifying the impact of changes in the transport network on local educational opportunity

Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.

Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.
- Placebo lines *Donaldson (2018)* or incidental connections *Faber (2014)* can't be used.

Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.
- Placebo lines *Donaldson (2018)* or incidental connections *Faber (2014)* can't be used.
- Instead I develop a novel IV approach extending the far-away variation strategy of *Donaldson and Hornbeck (2016), Jedwab and Storeygard (2022)*, by using not-on-least-cost-path variation.

Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.
- Placebo lines *Donaldson (2018)* or incidental connections *Faber (2014)* can't be used.
- Instead I develop a novel IV approach extending the far-away variation strategy of *Donaldson and Hornbeck (2016), Jedwab and Storeygard (2022)*, by using not-on-least-cost-path variation.
- Not-on-least-cost-path variation only uses variation in MA_i stemming from changes to all other locations MA_j freezing the least cost path from i to all j.

Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.
- Placebo lines *Donaldson (2018)* or incidental connections *Faber (2014)* can't be used.
- Instead I develop a novel IV approach extending the far-away variation strategy of *Donaldson and Hornbeck (2016), Jedwab and Storeygard (2022)*, by using not-on-least-cost-path variation.
- Not-on-least-cost-path variation only uses variation in MA_i stemming from changes to all other locations MA_j freezing the least cost path from i to all j.
    - Intuition 1: Far-away variation is defined over Euclidean distance, not-on-least-cost-path variation is defined over "network" distance.

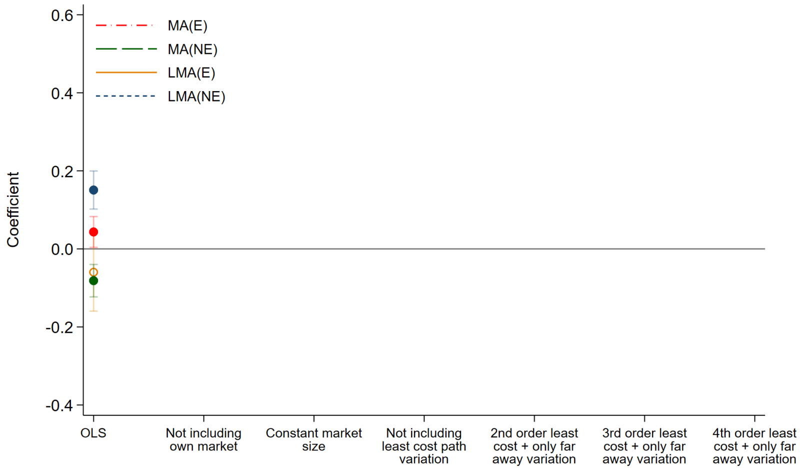
Not-on-least-cost-path identification strategy

- **Challenge 2**: Endogeneity of road placement.
- Placebo lines *Donaldson (2018)* or incidental connections *Faber (2014)* can't be used.
- Instead I develop a novel IV approach extending the far-away variation strategy of *Donaldson and Hornbeck (2016), Jedwab and Storeygard (2022)*, by using not-on-least-cost-path variation.
- Not-on-least-cost-path variation only uses variation in MA_i stemming from changes to all other locations MA_j freezing the least cost path from i to all j.
  - Intuition 1: Far-away variation is defined over Euclidean distance, not-on-least-cost-path variation is defined over "network" distance.
  - Intuition 2: Approximate the policy makers decision process.

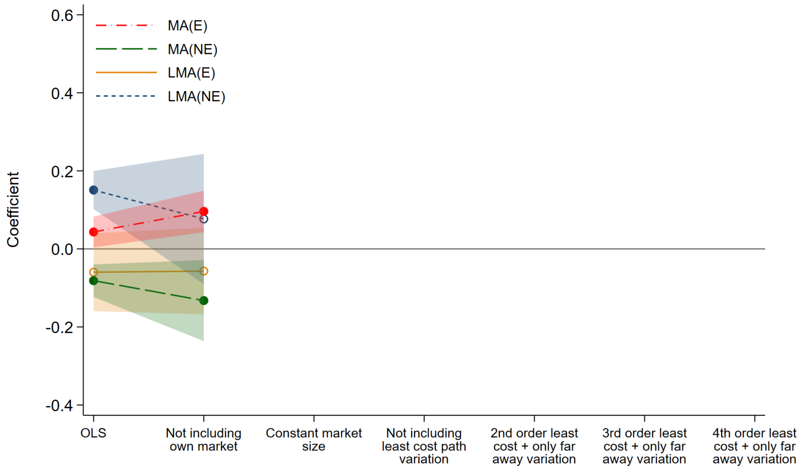
Graphical explanation.

Results from estimating the sufficient statistic result

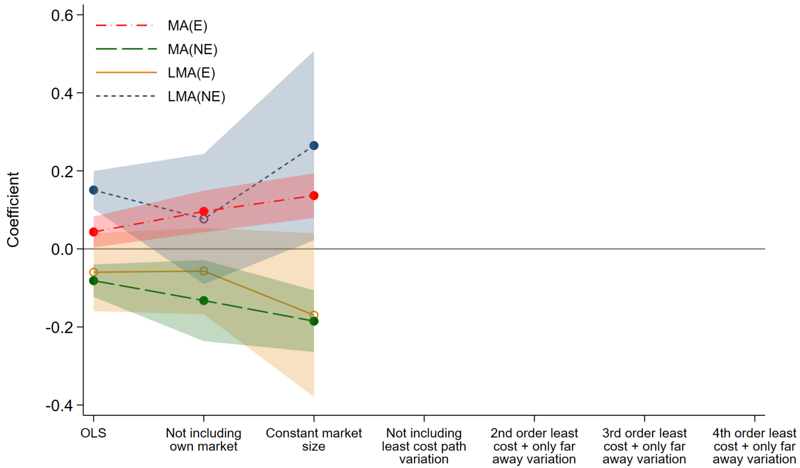
Changes in roads, through market access, affect local opportunity



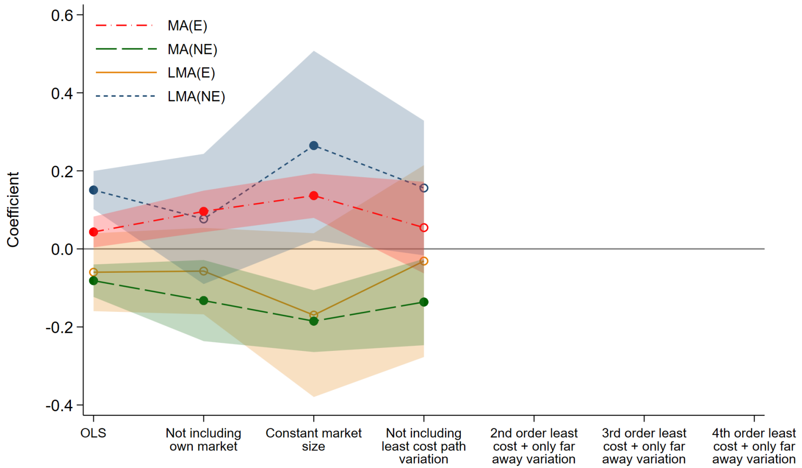
Changes in roads, through market access, affect local opportunity



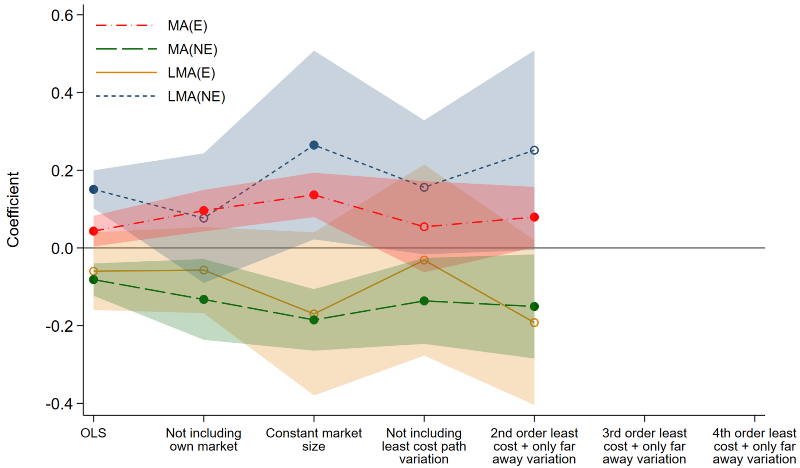
Changes in roads, through market access, affect local opportunity



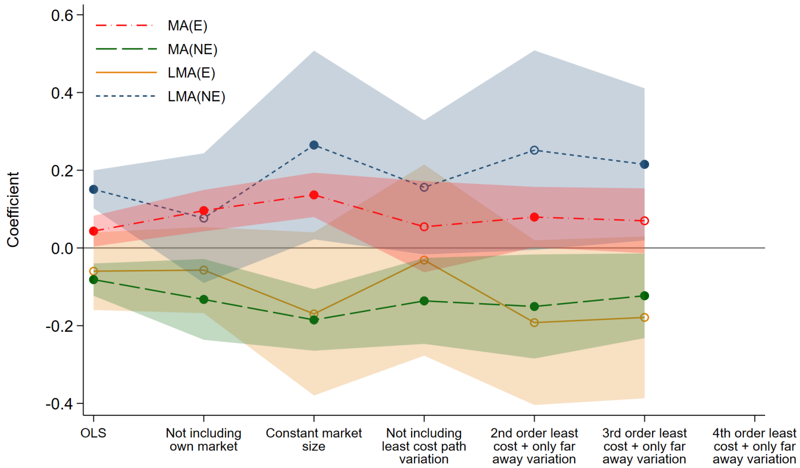
Changes in roads, through market access, affect local opportunity



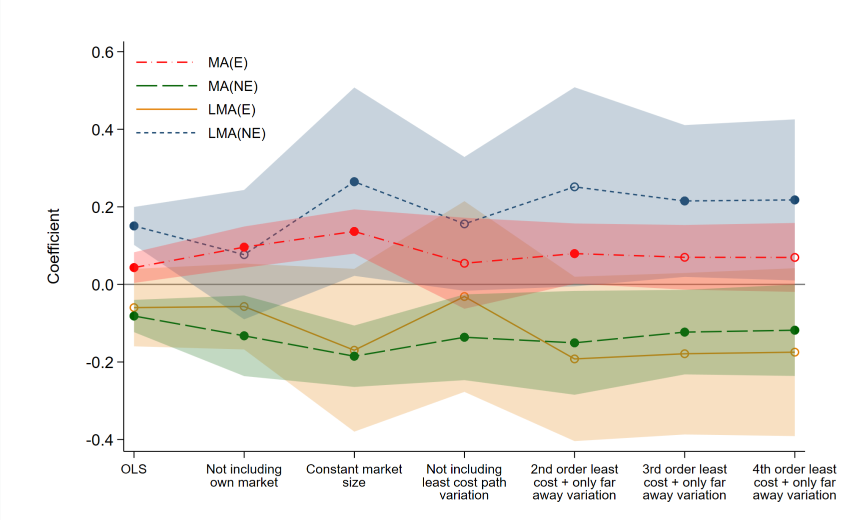
Changes in roads, through market access, affect local opportunity



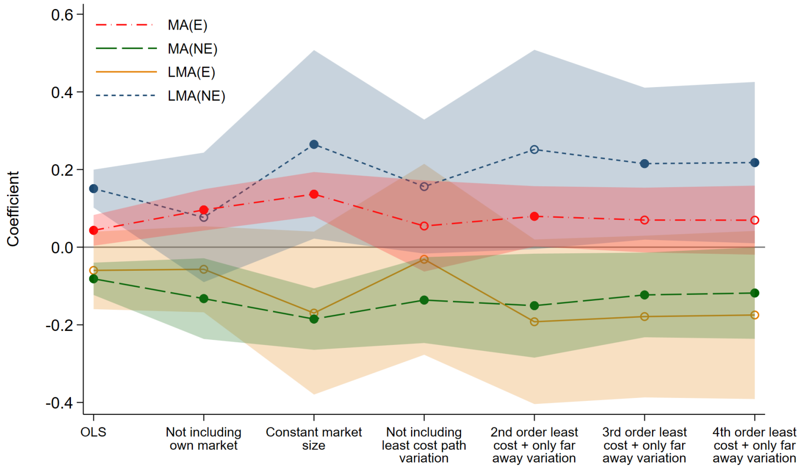
Changes in roads, through market access, affect local opportunity



Changes in roads, through market access, affect local opportunity



Changes in roads, through market access, affect local opportunity



	Combining Instruments
Log(LMA E)	-0.198**
	(0.0767)
Log(MA E)	0.118***
	(0.0281)
Log(LMA NE)	0.284***
	(0.0581)
Log(MA NE)	-0.171***
	(0.0329)
Locality and year FE	X
# Localities	127
N	334

Moving beyond the sufficient statistic approach

- **Research Question**: How does road building alter spatial inequality of opportunity?
- **Answer so far:** Road building does alter the spatial distribution of opportunity.

Moving beyond the sufficient statistic approach

- **Research Question**: How does road building alter spatial inequality of opportunity?
- **Answer so far:** Road building does alter the spatial distribution of opportunity.

Moving towards how: Counterfactuals.

- Remain agnostic about the modeling specification but take a stance on the modeling assumptions.
- Use the basic model described above.
  - Parsimonious whilst retaining key components.
  - Parameters can be directly estimated from the sufficient statistic result.
- One challenge remains: Network characteristics matter.

C: Counterfactual questions

Estimating the structural model

- Results in a succinct system of equations given below. Details.

$$MA_{it}^{n} = \sum_{j} K_{ijt}^{n} \prod_{h=1}^{4} \left(MA_{jt}^{h} \right)^{b_{nh}}$$

- *Allen, Arkolakis, and Li (WP)* show existence and give conditions for uniqueness.

Estimating the structural model

- Results in a succinct system of equations given below. Details.

$$MA_{it}^{n} = \sum_{j} K_{ijt}^{n} \prod_{h=1}^{4} \left(MA_{jt}^{h} \right)^{b_{nh}}$$

- *Allen, Arkolakis, and Li (WP)* show existence and give conditions for uniqueness.
- Kernel K_{ijt}^{n} is the product of iceberg movement costs and exogenous shifters, *unobserved*. Remove using exact-hat algebra *Dekle et al. (2008)*. Details.

$$\widehat{MA}_{i}^{n} = \sum_{j} \widehat{\rho}_{ij}^{n} \lambda_{ij}^{n} \prod_{h=1}^{4} \left(\widehat{MA}_{j}^{h} \right)^{b_{nh}}$$

- Coefficients b_{nh} are *unknown*. Recover from the estimated Suff. Stat. coefficients. Details.

Counterfactual questions

1. How did road building since 1970 effect spatial inequality of opportunity?

2. How might future road building effect inequality of opportunity over space?

3. How can we predict the effects of better connecting two given locations?

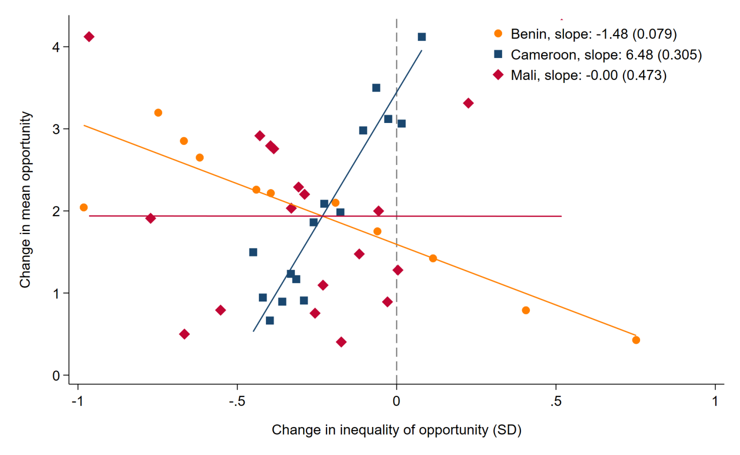
Equity-efficiency trade off

Equity-efficiency trade off

For each of 570 future road upgrades compare the change in average and variance of opportunity.

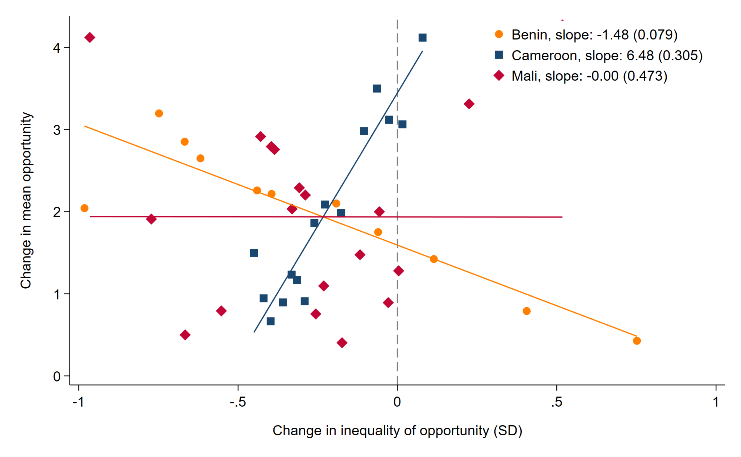
Equity-efficiency trade off

For each of 570 future road upgrades compare the change in average and variance of opportunity.



Equity-efficiency trade off

For each of 570 future road upgrades compare the change in average and variance of opportunity.



Bottom line: **If the social planner weights equity of opportunity this changes placement decisions.**

Summary

Research Question: How does road building alter spatial inequality of opportunity?

Summary

Research Question: How does road building alter spatial inequality of opportunity?

A: Develop novel sufficient statistic result.

- **Market access terms are a sufficient statistic for the effect of road building on opportunity.**

B: Novel identification strategy using not-on-least-cost-path variation.

- **Road building does alter the spatial distribution of opportunity.**

C: Estimate a structural general equilibrium spatial economics model to study counterfactuals.

- **Initial network conditions are important for determining the impact of changes since 1970.**

Appendix

Equity-efficiency trade off

Sufficient statistic results table. Return

	(1) OLS	(2) No including own market	(3) Constant market size	(4) Not including least cost path variation	(5) 1st order least cost + only far away variation	(6) 2nd order least cost + only far away variation	(7) 3rd order least cost + only far away variation	(8) 4th order least cost + only far away variation
Log(LMA Educ)	-0.0598	-0.0569	-0.170	-0.0312	-0.264	-0.192*	-0.179*	-0.175
	(0.0509)	(0.0564)	(0.107)	(0.125)	(0.356)	(0.108)	(0.106)	(0.110)
Log(MA Educ)	0.0434**	0.0961***	0.137***	0.0545	0.0904	0.0795**	0.0699	0.0696
	(0.0201)	(0.0271)	(0.0290)	(0.0598)	(0.237)	(0.0396)	(0.0425)	(0.0453)
Log(LMA No Educ)	0.151***	0.0767	0.265**	0.156*	0.619	0.252*	0.215**	0.218**
	(0.0249)	(0.0850)	(0.124)	(0.0879)	(0.475)	(0.131)	(0.0997)	(0.106)
Log(MA No Educ)	-0.0814***	-0.132**	-0.185***	-0.136**	-0.111	-0.151**	-0.123**	-0.118*
	(0.0212)	(0.0529)	(0.0403)	(0.0563)	(0.293)	(0.0682)	(0.0555)	(0.0601)
Locality by year FE	X	X	X	X	X	X	X	X
Kleibergen-Paap stat		7.39	8.01	5.65	0.52	2.93	4.23	4.74
Kleibergen-Paap p-value		0.007	0.005	0.017	0.470	0.087	0.040	0.029
SW under ID stat LMA(E)		108.53	34.45	7.84	2.61	15.97	26.00	23.65
SW under ID stat MA(E)		15.48	14.02	10.49	0.84	10.28	12.31	10.33
SW under ID stat LMA(NE)		13.40	13.21	12.59	1.16	4.57	9.81	11.49
SW under ID stat MA(NE)		12.93	27.01	23.74	0.69	5.00	12.45	11.25
SW weak ID stat LMA(E)		112.78	35.80	8.15	2.71	16.59	27.02	24.57
SW weak ID stat MA(E)		16.09	14.57	10.90	0.87	10.68	12.79	10.73
SW weak ID stat LMA(NE)		13.92	13.73	13.08	1.21	4.75	10.20	11.94
SW weak ID stat MA(NE)		13.43	28.07	24.67	0.71	5.19	12.94	11.69
# localities	127	127	127	127	127	127	127	127
N	334	334	334	334	334	334	334	334

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz_{lt}^{E}, Tz_{lt}^{N}, Tb_{lt}^{E}, Tb_{lt}^{N}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt} \left(\left(\frac{w^s_{lt}}{P^s_{lt}} \right)^{1-\beta} E^{\beta}_{lt} \right)^{\lambda_s}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz_{lt}^{E}, Tz_{lt}^{N}, Tb_{lt}^{E}, Tb_{lt}^{N}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u_{lt}^{s} = Tb_{lt}^{s} \left(\left(\frac{w_{lt}^{s}}{P_{lt}^{s}} \right)^{1-\beta} E_{lt}^{\beta} \right)^{\lambda_s}$
2. $E_{lt}^{\beta} = \left(\frac{w_{lt}^{E}}{w_{lt}^{N}} \right)^{\beta}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt}\left(\left(\frac{w^s_{lt}}{P^s_{lt}}\right)^{1-\beta} E^\beta_{lt}\right)^{\lambda_s}$
2. $E^\beta_{lt} = \left(\frac{w^E_{lt}}{w^N_{lt}}\right)^\beta$
3. $w^s_{lt} = \frac{Y^s_{lt}}{L^s_{lt}}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt}\left(\left(\frac{w^s_{lt}}{P^s_{lt}}\right)^{1-\beta} E^{\beta}_{lt}\right)^{\lambda_s}$
2. $E^{\beta}_{lt} = \left(\frac{w^E_{lt}}{w^N_{lt}}\right)^{\beta}$
3. $w^s_{lt} = \frac{Y^s_{lt}}{L^s_{lt}}$
4. $Y^s_{lt} = Tz^s_{lt}\left(w^s_{lt}\right)^{-\phi_k} MA^s_{lt}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz_{lt}^{E}, Tz_{lt}^{N}, Tb_{lt}^{E}, Tb_{lt}^{N}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u_{lt}^{s} = Tb_{lt}^{s} \left(\left(\frac{w_{lt}^{s}}{P_{lt}^{s}} \right)^{1-\beta} E_{lt}^{\beta} \right)^{\lambda_s}$
2. $E_{lt}^{\beta} = \left(\frac{w_{lt}^{E}}{w_{lt}^{N}} \right)^{\beta}$
3. $w_{lt}^{s} = \frac{Y_{lt}^{s}}{L_{lt}^{s}}$
4. $Y_{lt}^{s} = Tz_{lt}^{s} \left(w_{lt}^{s} \right)^{-\phi_k} MA_{lt}^{s}$
5. $L_{lt}^{s} = u_{lt}^{s} LMA_{lt}^{s}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt}\left(\left(\frac{w^s_{lt}}{P^s_{lt}}\right)^{1-\beta} E^{\beta}_{lt}\right)^{\lambda_s}$
2. $E^{\beta}_{lt} = \left(\frac{w^E_{lt}}{w^N_{lt}}\right)^{\beta}$
3. $w^s_{lt} = \frac{Y^s_{lt}}{L^s_{lt}}$
4. $Y^s_{lt} = Tz^s_{lt}\left(w^s_{lt}\right)^{-\phi_k} MA^s_{lt}$
5. $L^s_{lt} = u^s_{lt} LMA^s_{lt}$
6. $MA^s_{lt} = \left(P^s_{lt}\right)^{-\phi_k} = \sum_k \tau^{-\phi_k}_{lkt} \frac{Y^s_{kt}}{MA^s_{kt}}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt}\left(\left(\frac{w^s_{lt}}{P^s_{lt}}\right)^{1-\beta} E^{\beta}_{lt}\right)^{\lambda_s}$
2. $E^{\beta}_{lt} = \left(\frac{w^E_{lt}}{w^N_{lt}}\right)^{\beta}$
3. $w^s_{lt} = \frac{Y^s_{lt}}{L^s_{lt}}$
4. $Y^s_{lt} = Tz^s_{lt}\left(w^s_{lt}\right)^{-\phi_k} MA^s_{lt}$
5. $L^s_{lt} = u^s_{lt} LMA^s_{lt}$
6. $MA^s_{lt} = \left(P^s_{lt}\right)^{-\phi_k} = \sum_k \tau^{-\phi_k}_{lkt} \frac{Y^s_{kt}}{MA^s_{kt}}$
7. $LMA^s_{lt} = \sum_k \kappa^{-\lambda_s}_{lkt} \frac{L^s_{kt}}{LMA^s_{kt}}$

Quantitative spatial economics model. Return

Exogenous variables: $\{Tz^E_{lt}, Tz^N_{lt}, Tb^E_{lt}, Tb^N_{lt}, \tau_{lkt}, \kappa_{lkt}\}$ and parameters $\{\phi_E, \phi_N, \lambda_E, \lambda_N, \beta\}$ combine with the below series of equations (for each type s):

1. $u^s_{lt} = Tb^s_{lt}\left(\left(\frac{w^s_{lt}}{P^s_{lt}}\right)^{1-\beta} E^{\beta}_{lt}\right)^{\lambda_s}$
2. $E^{\beta}_{lt} = \left(\frac{w^E_{lt}}{w^N_{lt}}\right)^{\beta}$
3. $w^s_{lt} = \frac{Y^s_{lt}}{L^s_{lt}}$
4. $Y^s_{lt} = Tz^s_{lt}\left(w^s_{lt}\right)^{-\phi_k} MA^s_{lt}$
5. $L^s_{lt} = u^s_{lt} LMA^s_{lt}$
6. $MA^s_{lt} = \left(P^s_{lt}\right)^{-\phi_k} = \sum_k \tau_{lkt}^{-\phi_k} \frac{Y^s_{kt}}{MA^s_{kt}}$
7. $LMA^s_{lt} = \sum_k \kappa_{lkt}^{-\lambda_s} \frac{L^s_{kt}}{LMA^s_{kt}}$

to allow me to solve for the endogenous variables: $\{u^s_{lt}, E_{lt}, w^s_{lt}, Y^s_{lt}, L^s_{lt}, MA^s_{lt}, LMA^s_{lt}, \mu_{lt}\}$.

Far away and not-on-least-cost-path variation — the maths. Return

Consider some change in transport costs from $\{\tau_{lk}\}_{l,k}$ to $\{\tau'_{lt}\}_{l,k}$. Denote new variables with prime, so $MA^1_l = \sum_k \tau'_{lk} \frac{Y_k}{MA^1_k}$.

- Not-on-least-cost-path variation: $MA^2_l = \sum_k \tau_{lk} \frac{Y_k}{MA^1_k}$
- g-th order not-on-least-cost-path variation $MA^g_l = \sum_k \tau_{lk} \frac{Y_k}{MA^{g-1}_k}$

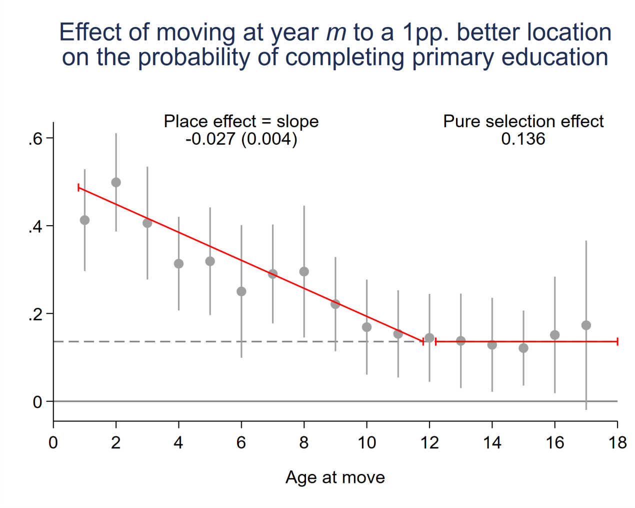
Estimating causal place effects. Return

Place effects are estimated in *Heath Milsom (2022)*, and follow *Chetty and Hendren (2018), Alesina et al (2021)* and others.

$$y_i = \alpha_{odt} + \mu_{lt} \cdot e_{il} + \varepsilon_i \qquad\qquad e_{il} = \begin{cases} 13 - m_i & \text{if } l = d(i) \\ m_i & \text{if } l = o(i) \\ 0 & \text{otherwise} \end{cases}$$

- μ_{lt} is the causal effect of spending an additional year of childhood in location l in period t on the probability of completing primary school.
- Can be used to decompose observed primary completion rates: $\bar{y}_{lt} - \bar{\bar{y}}_t = 13\mu_{lt} + \bar{\theta}_{lt}$.
- Relies on three key assumptions:
    1. (Functional form assumption) Place effects are additive and linear. Evidence from estimating age-at-move specific effects.
    2. (Interpretation assumption) Effect of place on movers is similar to that on stayers. Evidence from exogenous move events, and using to-be mover variation.
    3. (Identifying assumption) Selection effects do not systematically vary with the age at move. Evidence from HH fixed effects and over-identifying restrictions.

Functional form assumption. Return



- Assume place effects are additive.
- Study the linearity of place effects by estimating age-at-move specific effects across the whole sample (rather than for each individual location).
- $y_i = \sum_{m=1}^{18} \beta_m \cdot \mathbb{1}_{[m(i)=m]} \cdot (\bar{y}_{dt}^{p} - \bar{y}_{ot}^{p}) + \alpha_{ot} + \varepsilon_i$
- If $\Delta\beta_m$ is constant from 1-12 gives evidence of linearity.

Linear parametisation with varying FE specifications. | Robustness to LPM.

Interpretation assumption. Return

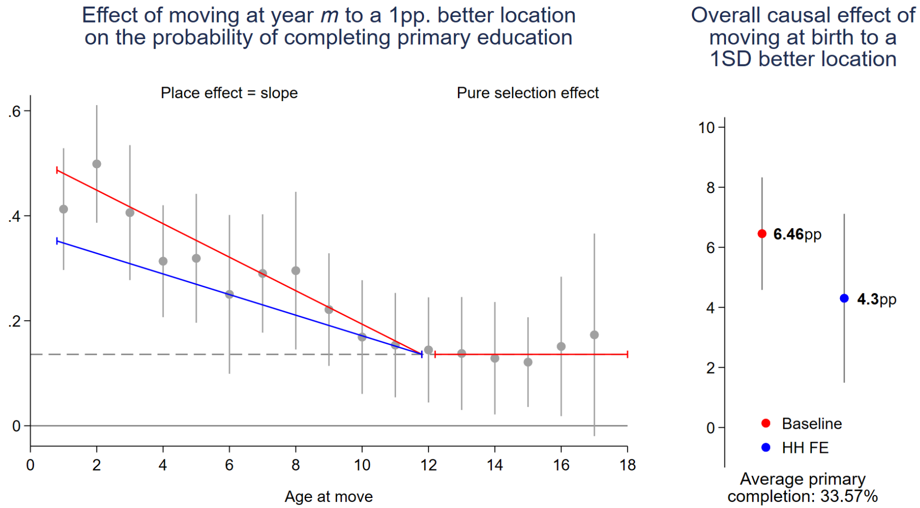
Assume

1. Movers and stayers do not systematically differ in a manner that interacts with place effects.
2. Movers and stayers interact similarly with their surroundings.

Evidence

- To help overcome (1) compare results using large movement events $\Rightarrow$ movers are less selected, find very similar results.
- To help overcome (2) use variation in origin quality enjoyed before a move occurs gives near identical results.

Include HH fixed effects. Return



Overidentification test. Return

- Individuals should react to group-specific location quality, not general location quality.
- Although gender-specific location quality is highly correlated this table shows in column (3) that what matters is group-specific quality in line with the CPE explanation over a selection story.

	(1)	(2)	(3)	(4)
Base	0.0209***			
	(0.0029)			
Same reported gender specific		0.0176***		0.0123***
		(0.0025)		(0.0041)
Other reported gender specific			0.0164***	0.00664
			(0.0025)	(0.0041)
N	65560	65560	65560	65560
R^2	0.358	0.356	0.358	0.358

Education supply. Return

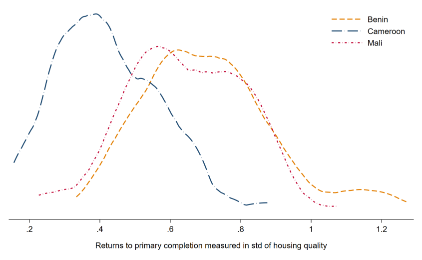
- Education sector is modeled as public, central planner maximises the number of students $\sum_l A_l$ with access to schooling by building schools S_l at fixed cost c with fixed budget I (Khanna 2015).
- The number of children with access to schools in a region is increasing in the number of schools, population density and easy of transportation: $A_l = \pi S_l^{\alpha_1} D_l^{\alpha_2} R_l^{\alpha_3}$.
- Solve the planners optimisation problem to find: $S_l = \gamma D_l^{\frac{\alpha_2}{1-\alpha_1}} R_l^{\frac{\alpha_3}{1-\alpha_1}}$.
- Endogenise the local cost of education as decreasing in the expected time taken to travel to a school (*DeStefano et al. (2007), Evans and Mendez Acosta (2021)*): $c_l^E = \kappa S_l^{\gamma_1} R_l^{\gamma_2}$.
- Then the causal effect of place on the probability of completing primary education is increasing in the returns to education and before, but decreasing in the cost of educating.
- Find also that changes in education quality don't jeopardize the sufficient statistic result.

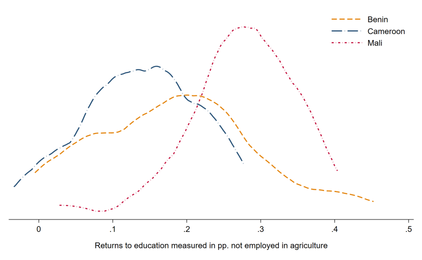
Evidence that local wages matter for education choices Return

- From the literature:
  - *Eckert and Kleineberg (WP)*, show that the college premium varies across CZ in the US and that where you grow up matters for future moving choices.
  - *Chetty and Hendren (2018)*, show that where you grow up matters for college attendance across CZ.
  - *Edmonds, Pavcnik, Topalova (2010), Adukia, Asher, and Novosad (2020), Atkin (2016)*, find that local actual cost, returns to, and opp cost, of education influence education choices.
  - *Hsiao (WP)*, effectiveness of school provision depends on local access to jobs which require schooling.

Evidence that local wages matter for education choices Return

- From the literature:
  - *Eckert and Kleineberg (WP)*, show that the college premium varies across CZ in the US and that where you grow up matters for future moving choices.
  - *Chetty and Hendren (2018)*, show that where you grow up matters for college attendance across CZ.
  - *Edmonds, Pavcnik, Topalova (2010), Adukia, Asher, and Novosad (2020), Atkin (2016)*, find that local actual cost, returns to, and opp cost, of education influence education choices.
  - *Hsiao (WP)*, effectiveness of school provision depends on local access to jobs which require schooling.
- Using my data:
  - Returns to education by birth location significantly disperate over space.





Linear parametization. Return

$$
\begin{aligned}
y_i = {} & FE \\
& + (\alpha_1 + \beta_1 \times \text{age move}_i) \cdot \mathbb{1}\left[\text{age move}_i \in [1, 12]\right] \cdot \Delta \bar{y}^p_{ods} \\
& + (\alpha_2 + \beta_2 \times \text{age move}_i) \cdot \mathbb{1}\left[\text{age move}_i \in [13, 18]\right] \cdot \Delta \bar{y}^p_{ods} + u_i
\end{aligned}
$$

	Year born - birth location and age at move FE	Year born, year born by $\bar{y}^p_{os}$ and age move by sample FE	Year born, year born by $\bar{y}_{os}$ and age move FE	Year born and age moved FE
$\hat{\beta}_1$	-0.0270***	-0.0272***	-0.0284***	-0.0301***
	(0.00530)	(0.00500)	(0.00554)	(0.00878)
$\hat{\beta}_2$	0.00302	0.00893	0.00553	0.0215**
	(0.00914)	(0.00631)	(0.0120)	(0.00717)
R^2	0.348	0.311	0.303	0.195

N = 79,970, SE clustered at origin and destination.

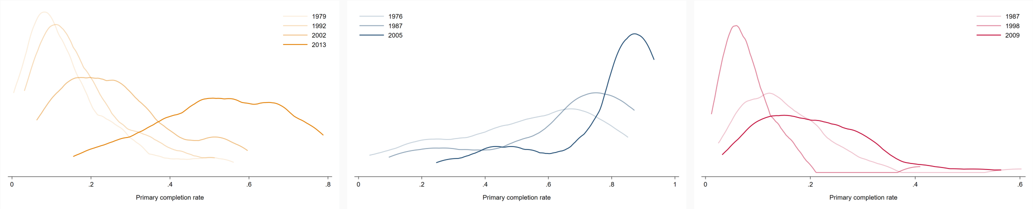
Robustness to LPM assumption. Return

	(1) LPM	(2) LPM	(3) PPML	(4) PPML	(5) Probit	(6) Logit
1 to 12 slope	-0.0270***	-0.0283***	-0.0452***	-0.0478***	-0.0871***	-0.145***
	(0.00530)	(0.00545)	(0.0113)	(0.0121)	(0.0113)	(0.0189)
13 to 18 slope	0.00302	0.00469	0.00210	0.00951	0.0220	0.0295
	(0.00914)	(0.0119)	(0.0185)	(0.0248)	(0.0246)	(0.0418)
1 to 12 constant	0.486***	0.514***	0.856***	0.950***	1.557***	2.574***
	(0.0496)	(0.0611)	(0.122)	(0.133)	(0.178)	(0.292)
13 to 18 constant	0.0932	0.0805	0.208	0.174	0.135	0.339
	(0.122)	(0.166)	(0.238)	(0.346)	(0.367)	(0.624)
Year born by birth location and age moved FE	X		X			
Year born, year born by birth location quality and age moved FE		X		X	X	X
R^2	0.348	0.304				
N	79778	79778	78950	79776	79776	79776

Top-coding. Return

Figure 6: Changes in the distribution over locations 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 re-sampling the same individuals.

Non-linearities 1/2. Return

Table 1: Relation between base level and changes (combined regressions)

	(1) Long Diff $Log(LMA_E)$	(2) Long Diff $Log(LMA_N)$	(3) Long Diff $Log(MA_E)$	(4) Long Diff $Log(MA_N)$	(5) Long Diff Prim Educ	(6) Log Diff μ
$Log(LMA_E)$	0.0546	0.0895	0.611**	0.652***	-0.186*	-0.0141
	(0.0935)	(0.117)	(0.291)	(0.231)	(0.102)	(0.0253)
$Log(LMA_N)$	0.139	0.0235	0.149	-0.322	0.570***	0.0502
	(0.141)	(0.173)	(0.402)	(0.312)	(0.143)	(0.0342)
$Log(MA_E)$	0.224**	0.230***	0.306	0.373**	-0.258***	-0.00717
	(0.0873)	(0.0806)	(0.228)	(0.180)	(0.0819)	(0.0165)
$Log(MA_N)$	-0.315*	-0.198*	-0.503	-0.318	0.0152	-0.0113
	(0.165)	(0.116)	(0.434)	(0.315)	(0.115)	(0.0237)
Prim Educ	-0.585***	-0.778***	-1.664***	-1.845***	1.027***	-0.0108
	(0.197)	(0.162)	(0.567)	(0.444)	(0.184)	(0.0262)
μ	0.0667	0.109	0.00185	0.145	1.207**	1.158***
	(0.522)	(0.488)	(1.661)	(1.260)	(0.588)	(0.167)
N	136	136	136	136	136	114

Notes: This table shows the results from running regressions of the long difference of each variable on the initial period level of all other variables. These regressions are weighted by 1970 locality population and include country fixed effects. Standard errors are robust.

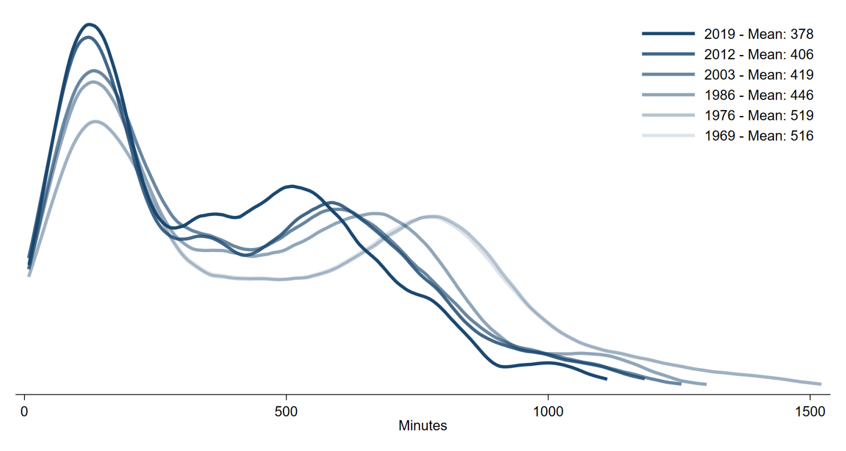
Non-linearities 2/2. Return

Table 2: Relation between base level and changes (individual regressions)

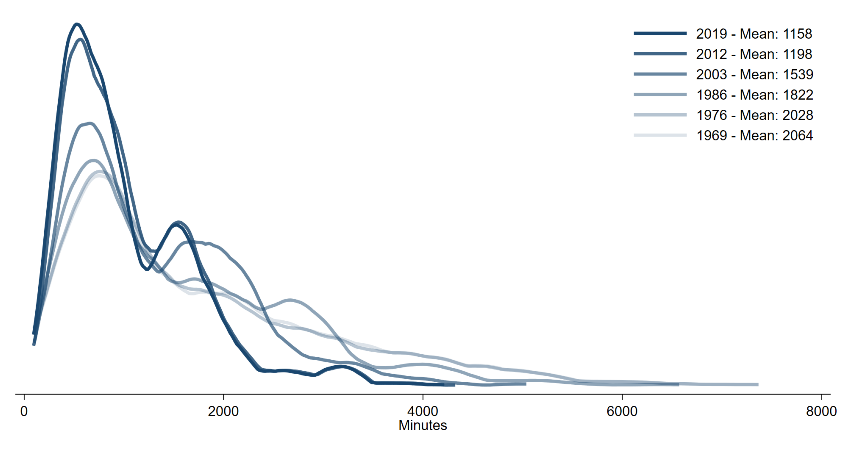
	(1) Long Diff $Log(LMA_E)$	(2) Log Diff $Log(LMA_N)$	(3) Long Diff $Log(MA_E)$	(4) Long Diff $Log(MA_N)$	(5) Long Diff Prim Educ	(6) Long Diff μ
$Log(LMA_E)$	-0.0581					
	(0.0471)					
$Log(LMA_N)$		0.199^{***}				
		(0.0728)				
$Log(MA_E)$			-0.0741			
			(0.0510)			
$Log(MA_N)$				0.0823		
				(0.0726)		
Prim Educ					0.124^{*}	
					(0.0735)	
μ						0.372^{*}
						(0.190)
N	162	162	162	162	162	114

Notes: This table shows the results from running regressions of the long difference of each variable on the initial period level of that variable only. These regressions are weighted by 1970 locality population and include country fixed effects. Standard errors are robust.

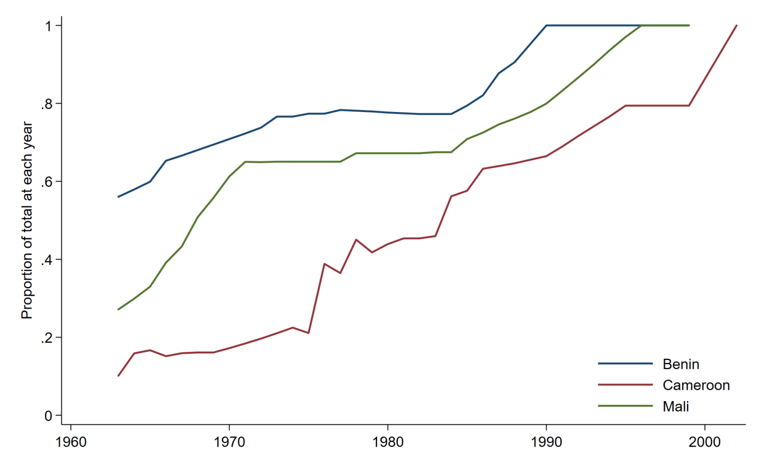
Since 1970 average travel times have fallen by roughly 33% in Benin, Cameroon, and Mali. Return



Since 1970 average travel times have fallen by roughly 33% in Benin, Cameroon, and Mali. Return



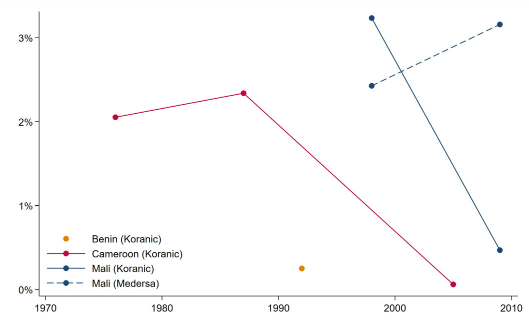
Alternative source showing aggregate changes in the transport network. Return



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)*.

Koranic schools and Medersas. Return

- Koranic schools are not counted as having completed education as they do not follow the curriculum.
- Medersas in Mali are religious schools that do follow the curriculum (*Boyle (2014)*) and so are counted.



Proportion of school age children enrolled in Koranic schools or Medersas.

Sample. Return

	1970	1975	1980	1985	1990	1995	2000	2005	2010	2015	2020
Michelin Maps	1969	1976		1986			2003		2012		2019
Benin Census					1992		2002			2013	
Cameroon Census	1976				1987			2005			
Mali Census							1998		2009		

Map Year	Census Year	Country
1969	1976	Cameroon
1976	1987	Cameroon
1976	1992	Benin
1986	1998	Mali
1986	2002	Benin
1986	2005	Cameroon
2003	2009	Mali
2003	2013	Benin

Specification details. Return

Estimate:

$$\mu_{lt} = \gamma_1 \cdot \ln(MA^{E}_{lt}) + \gamma_2 \cdot \ln(MA^{N}_{lt}) + \gamma_3 \cdot \ln(LMA^{E}_{lt}) + \gamma_4 \cdot \ln(LMA^{N}_{lt}) + \gamma_l + a_t + v_{lt}$$

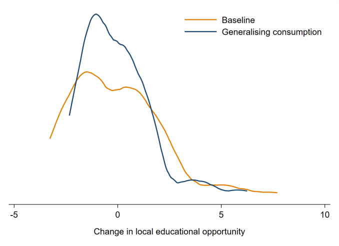
- $N = 334$
- Census years (market access measured ∼10 years before):
  - Benin: 1992, 2002, 2013. Cameroon: 1976, 198, 2005. Mali: 1998, 2009.
- Standard errors are clustered at the locality level.
- Kleibergen-Paap and Sanderson-Windmeijer weak-/ under-identification tests look okay.

Robustness to allowing more general consumption patterns. Return

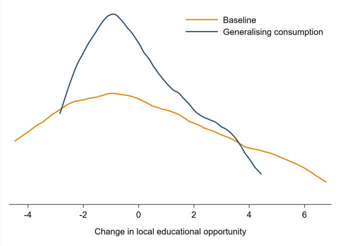
Allow both type of workers to consume goods from both sectors. Run-run counterfactual, correlation between results is >0.99, but variance is 25% smaller.

E-type spend 50% of their income on E-type goods, N-types spend 10%.

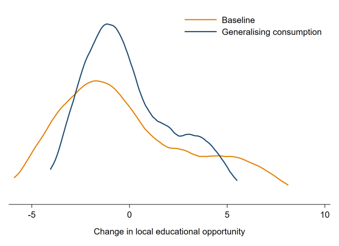
(a) Benin



(b) Cameroon



(c) Mali



Parameter estimates. Return

- Relative share of utility from education: $\beta = 0.2$.

- Sensitivity of goods and worker movement to transport costs: $\phi_E, \phi_N, \lambda_E, \lambda_N$

  - Sufficient statistic result gives us a system of non-linear equations mapping reduced form estimates to structural parameters: $\widehat{\gamma}_k = \widehat{\gamma}_k(\phi_E, \phi_N, \lambda_E, \lambda_N, \alpha, \beta)$, for $k = 1, 2, 3, 4$.

  - Solve the system to find: $\phi_E = 2.92, \phi_N = 1.77, \lambda_E = 3.10, \lambda_N = 1.88$.

- Alternatively calibrate from lit: Movement costs taken from Tsivanidis (2022) (commuting, Bogota, high/low educ), scaled by Morten and Oliveira (2022) (cross-country, Brazil): $\lambda_E = 1.74$, $\lambda_N = 2.11$. Trade costs taken from Zarate (2022) (Mexico city, formal/informal), scaled by Morten and Oliveira (2022): $\phi_E = 3.47$, $\phi_N = 4.52$.Quantitative results are very similar.

Exact hat results. Return

- Use exact-hat algebra, i.e. solve the model in changes and the T's drop out.
- Define λ^r_{ljt} as the proportion of l's market access (of type k) in t, which is due to j, then exact hat algebra gets us:

$$\widehat{MA}_i^E = \frac{(MA^E)_i'}{MA_i^E} = \frac{\sum_j (\tau_{ij}^c)^{-\phi_E} \frac{Y_j^{E,c}}{MA_k^{E,c}}}{\sum_k \tau_{ik}^{-\phi_E} \frac{Y_k}{MA_k^{E,c}}} = \sum_j \frac{(\tau_{ij}^c)^{-\phi_E} \frac{Y_j^{E,c}}{MA_j^{E,c}}}{\sum_k \tau_{ik}^{-\phi_E} \frac{Y_k}{MA_k^{E,c}}} = \sum_j \frac{\tau_{ij}^{-\phi_E} \frac{Y_j^E}{MA_j^E}}{\sum_k \tau_{ik}^{-\phi_E} \frac{Y_k}{MA_k^{E,c}}} \widehat{\tau}_{ij}^{-\phi_E} \widehat{Y}_j^E \left(\widehat{MA}_j^E\right)^{-1}$$

$$= \sum_j \lambda_{ij}^{MA^E} \widehat{\tau}_{ij}^{-\phi_E} \widehat{Y}_j^E \left(\widehat{MA}_j^E\right)^{-1}$$

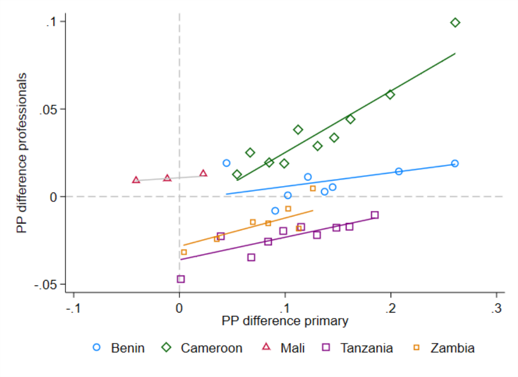
- $\hat{\rho}^r_{ijt}$ is either the change in τ or κ depending on market access type r.
- As we know λ^r_{ijt}, for any given counterfactual we can calculate $\widehat{\rho}^r_{ijt}$ and thus solve to find $\widehat{MA}^r_{it}$.

Education = Opportunity? Return

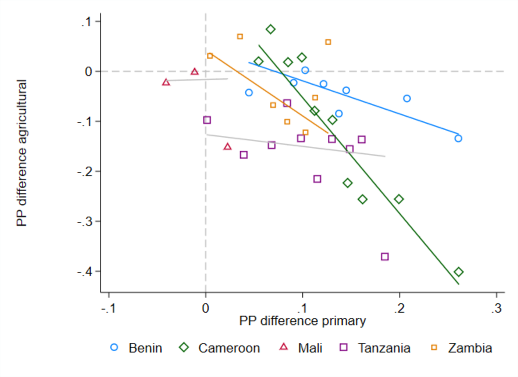
- Salient margin, 34% complete primary education, 6% complete secondary.
- High returns to education in Sub-Saharan Africa *(Psacharopoulos and Patrinos. (2018))*
- Primary completion is correlated with later life outcomes. E.g. occupation and housing quality.

Opportunity and primary education in Benin, Cameroon and Mali. Return

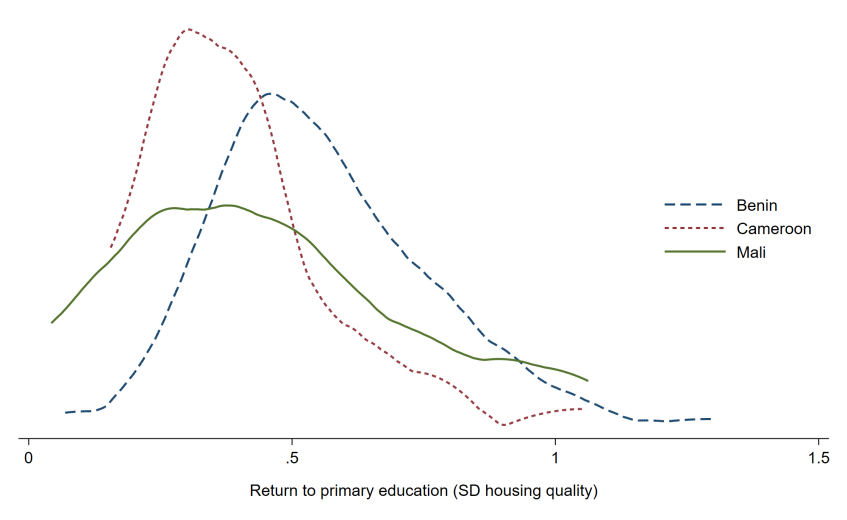
(a) Proportion Working Professionally



(b) Proportion Working in Agriculture



Opportunity and primary education in Benin, Cameroon and Mali. Return



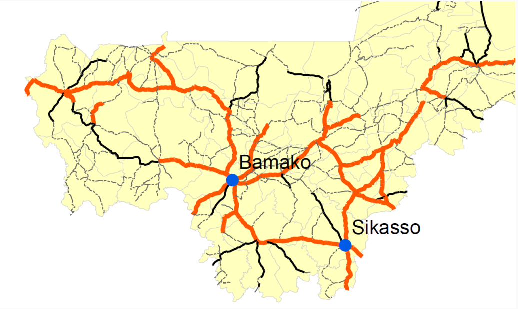
Measuring impedance terms. Return

Project bilateral iceberg transport costs onto the logarithm of the calculated least cost path travel times, t_{lk}, as follows $\tau_{lk}^{-\phi} = -\tilde{\phi}\ln(t_{lk})$, $\kappa_{lk}^{-\lambda_s} = -\tilde{\lambda}^s \ln(t_{lk})$.

1. Calculate t_{lkt} from digitized historical Michelin road maps.
2. Recover $\widehat{\kappa_{lk}^{-\lambda_s}}$ from internal migration gravity equations. Take $\tilde{\phi}$ from the literature.

Michelin map of Mali 2019

Digitised map of Mali 2019



Gravity equations: $M^s_{ltt} = \alpha^s_{lt} \cdot \lambda^s \ln(t_{lkt}) \cdot \rho^s_{kt} \cdot \varepsilon^s_{lkt}$. Return

	No primary education		Primary education	
	(1) Log-linear	(2) PPML	(3) Log-Linear	(4) PPML
Log(Travel time)	-1.270***	-1.477***	-0.933***	-0.985***
	(0.0193)	(0.0356)	(0.0171)	(0.0375)
Destination-time FE	X	X	X	X
Origin-time FE	X	X	X	X
N	11005	26234	10941	26234

Gravity equations: $M_{ltt}^s = \alpha_{lt}^s \cdot \lambda^s \ln(t_{lkt}) \cdot \rho_{kt}^s \cdot \varepsilon_{lkt}^s$. Return

	No primary education		Primary education	
	(1) Log-linear	(2) PPML	(3) Log-Linear	(4) PPML
Log(Travel time)	-1.270***	-1.477***	-0.933***	-0.985***
	(0.0193)	(0.0356)	(0.0171)	(0.0375)
Destination-time FE	X	X	X	X
Origin-time FE	X	X	X	X
N	11005	26234	10941	26234

$\hat{\tilde{\lambda}}^E = 0.985$

$\hat{\tilde{\lambda}}^N = 1.477$

Gravity equations: $M^s_{ltt} = \alpha^s_{lt} \cdot \lambda^s \ln(t_{lkt}) \cdot \rho^s_{kt} \cdot \varepsilon^s_{lkt}$. Return

	No primary education		Primary education	
	(1) Log-linear	(2) PPML	(3) Log-Linear	(4) PPML
Log(Travel time)	-1.270***	-1.477***	-0.933***	-0.985***
	(0.0193)	(0.0356)	(0.0171)	(0.0375)
Destination-time FE	X	X	X	X
Origin-time FE	X	X	X	X
N	11005	26234	10941	26234

$\hat{\tilde{\lambda}}^E = 0.985$

$\hat{\tilde{\lambda}}^N = 1.477$

$\hat{\tilde{\phi}} = 3.8$ from Donaldson and Hornbeck (2016)

Calculating wages. Return

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. Following Young (2012).

Postulate that the (real) demand for an asset p by household h in locality l in year t is give by the following equation.

$$\ln(Q_{phlt}) = \alpha_p + \eta_p \ln(C^N_{hlt}) + \xi_p \ln(P_{lt}) + \beta X_{hlt} + \varepsilon_{phlt} \quad (1)$$

Where α_p are product constants, η_p is the (quasi) income elasticity of demand, C^N_{hlt} is nominal household consumption expenditure which is equal to household income in our setting, ξ_p is a vector of own and cross-price (quasi) elasticities of demand, $\ln(P_{lt})$ is a vector of regional prices, X_{hlt} and β are vectors of household characteristics and their coefficients.

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.

Calculating wages. Return

	(1)	(2)	(3)	(4)
Conrete floor	0.00550***	0.00557***	0.00479***	0.00496***
	(0.000914)	(0.000923)	(0.000820)	(0.000915)
Electricity	0.00212***	0.00218***	0.00205***	0.00158**
	(0.000653)	(0.000668)	(0.000593)	(0.000665)
Sanitation	0.00312***	0.00319***	0.00313***	0.00258***
	(0.000890)	(0.000898)	(0.000814)	(0.000893)
Asset × Region FE	X	X	X	X
Age polynomial		X	X	X
Asset × Region × Urban FE			X	
HH members control				X
R^2	0.402	0.403	0.498	0.404
N	22586	22586	22586	22586

In the second step, I use the inverted estimated coefficients to approximate income differences by assets households own as indicated in census data. Imputed average income in a locality-year-education cell is then given by $\tilde{Y}^e_{lt} = 1/N_{lte} \sum_{h \in \{l,t,e\}} \sum_p \frac{1}{\hat{\eta}_p} Q_{hplt} + base$ where *base* is the average income calculated from DHS data.

Borusyak and Hull (2021) correction. Return

	(1) Baseline	(2) Including average MA variables
Log(LMA E)	-0.198**	-0.228***
	(0.0767)	(0.0844)
Log(MA E)	0.118***	0.136***
	(0.0281)	(0.0355)
Log(LMA NE)	0.284***	0.334***
	(0.0581)	(0.0953)
Log(MA NE)	-0.171***	-0.211***
	(0.0329)	(0.0536)
Locality and year FE	X	X
# localities	127	127
N	334	334

Clientelism. Return

	(1) Baseline	(2) Including ethnicity by leader
Log(LMA Educ)	-0.198**	-0.208***
	(0.0767)	(0.0780)
Log(MA Educ)	0.118***	0.114***
	(0.0281)	(0.0268)
Log(LMA No Educ)	0.284***	0.277***
	(0.0581)	(0.0549)
Log(MA No Educ)	-0.171***	-0.166***
	(0.0329)	(0.0324)
Same ethnicity as leader		0.00137
		(0.00584)
Locality by year FE	X	X
# localities	127	127
N	334	334

- Little variation in Cameroon as Paul Biya Beti has been in power since 1982.
- Clientelism thought not to be rife in Mali *Dunning and Harrison (2010), Basedau and Stroh (2012), Franck and Rainer (2012)*.
- Evidence clientelism may play a role in Benin *Fujiwara and Wantchekon (2013)*.
- Include dummy variable = 1 if in Benin the local ethnic group (Ewe, Yoruba, Somba, Barba) is the same as the leader.

Calculating market access terms. Return

$$MA^s_{lt} = \sum_k \tau_{lkt}^{-\phi} \frac{Y^s_{kt}}{MA^s_{kt}} \qquad LMA^s_{lt} = \sum_k \kappa_{lkt}^{-\lambda_s} \frac{L^s_{kt}}{LMA^s_{kt}}$$

I need data on: $\{\tau_{lkt}^{-\phi}\}$, $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{Y^s_{lt}\}$, $\{L^s_{lt}\}$.

Calculating market access terms. Return

$$MA^s_{lt} = \sum_k \tau_{lkt}^{-\phi} \frac{Y^s_{kt}}{MA^s_{kt}} \qquad LMA^s_{lt} = \sum_k \kappa_{lkt}^{-\lambda_s} \frac{L^s_{kt}}{LMA^s_{kt}}$$

I need data on: $\{\tau_{lkt}^{-\phi}\}$, $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{Y^s_{lt}\}$, $\{L^s_{lt}\}$.

- Digitize historical Michelin maps to find travel times and use gravity internal migration regressions to uncover $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{\tau_{lkt}^{-\phi}\}$.

Calculating market access terms. Return

$$MA^s_{lt} = \sum_k \tau_{lkt}^{-\phi} \frac{Y^s_{kt}}{MA^s_{kt}} \qquad LMA^s_{lt} = \sum_k \kappa_{lkt}^{-\lambda_s} \frac{L^s_{kt}}{LMA^s_{kt}}$$

I need data on: $\{\tau_{lkt}^{-\phi}\}$, $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{Y^s_{lt}\}$, $\{L^s_{lt}\}$.

- Digitize historical Michelin maps to find travel times and use gravity internal migration regressions to uncover $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{\tau_{lkt}^{-\phi}\}$.
- Population data, $\{L_{lt}\}$, from Censuses, gridded population of the world, and Africapolis.

Calculating market access terms. Return

$$MA^s_{lt} = \sum_k \tau_{lkt}^{-\phi} \frac{Y^s_{kt}}{MA^s_{kt}} \qquad\qquad LMA^s_{lt} = \sum_k \kappa_{lkt}^{-\lambda_s} \frac{L^s_{kt}}{LMA^s_{kt}}$$

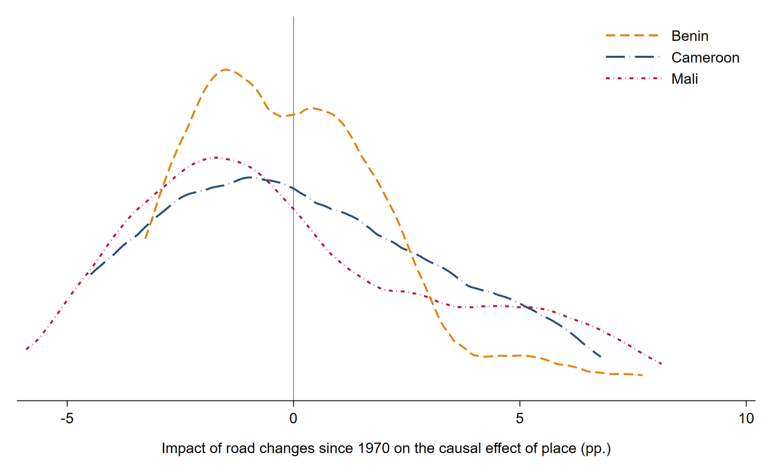
I need data on: $\{\tau_{lkt}^{-\phi}\}$, $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{Y^s_{lt}\}$, $\{L^s_{lt}\}$.

- Digitize historical Michelin maps to find travel times and use gravity internal migration regressions to uncover $\{\kappa_{lkt}^{-\lambda_s}\}$, $\{\tau_{lkt}^{-\phi}\}$.
- Population data, $\{L_{lt}\}$, from Censuses, gridded population of the world, and Africapolis.
- Use wage data from demographic health surveys and Engel curve regressions to estimate $\{Y^s_{lt}\}$ using an approach similar to Young (2012).

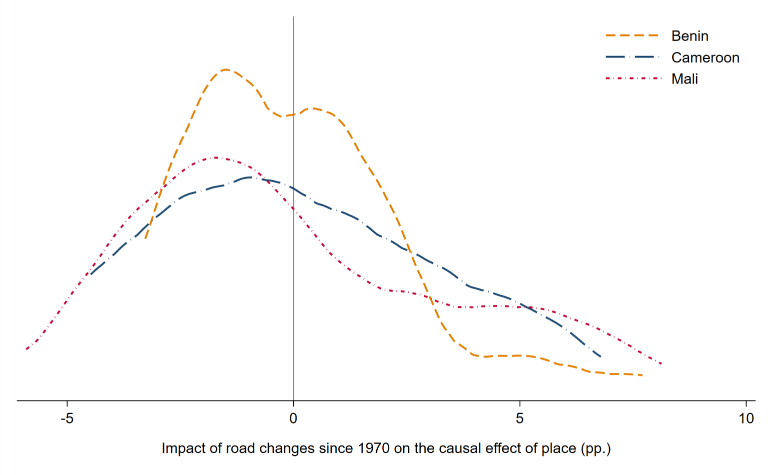
Details

How did road building since 1970 effect spatial inequality of opportunity? Return

How did road building since 1970 effect spatial inequality of opportunity? Return



How did road building since 1970 effect spatial inequality of opportunity? Return

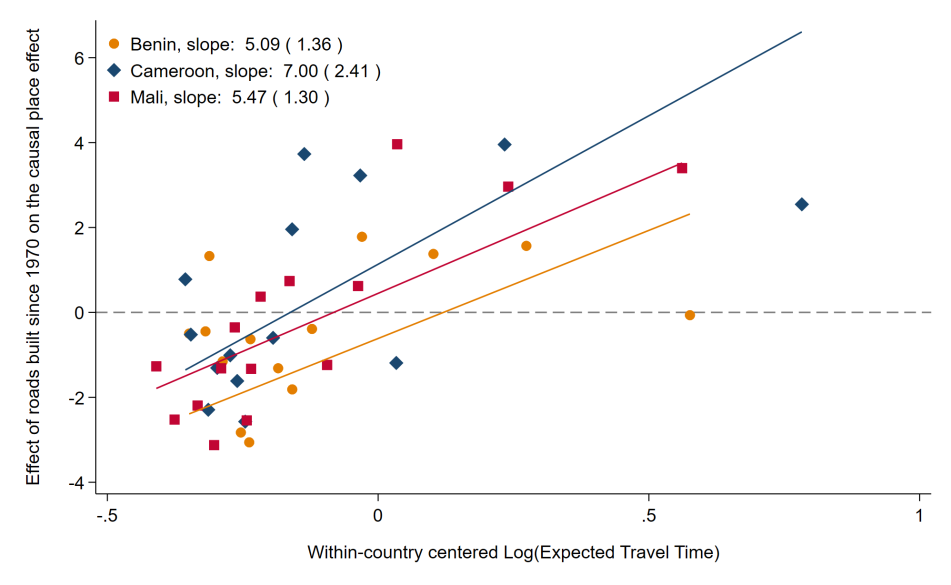


% change in variance of opportunity across space.

- Benin: $+0.04\%$
- Cameroon: $+5.81\%$
- Mali: -1.44%

Alternative calibration from the literature | Generalizing consumption

Within-country, larger effects are associated with remoteness in 1970. Return



Cross-country heterogeneity can be partly explained by initial conditions. Return

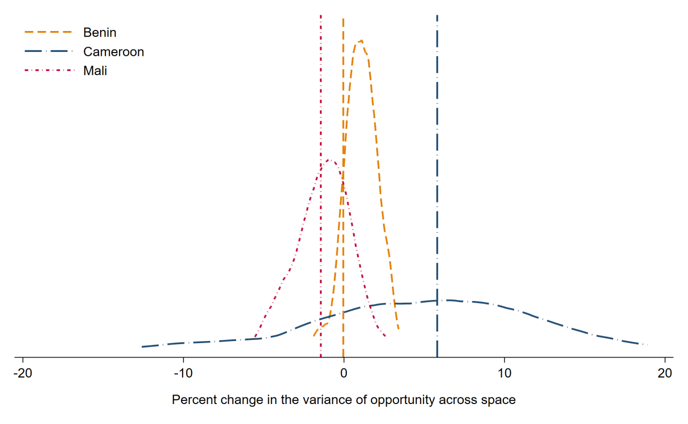
- Two explanations for cross-country heterogeneity: (1) Differing constraints faced by policy makers. (2) Differing efficacy of policy decisions.

Cross-country heterogeneity can be partly explained by initial conditions. Return

- Two explanations for cross-country heterogeneity: (1) Differing constraints faced by policy makers. (2) Differing efficacy of policy decisions.
- To investigate this compare the realised effects to those from 250 randomly generated counterfactual networks with the same overall change in travel time.

Cross-country heterogeneity can be partly explained by initial conditions. Return

- Two explanations for cross-country heterogeneity: (1) Differing constraints faced by policy makers. (2) Differing efficacy of policy decisions.
- To investigate this compare the realised effects to those from 250 randomly generated counterfactual networks with the same overall change in travel time.



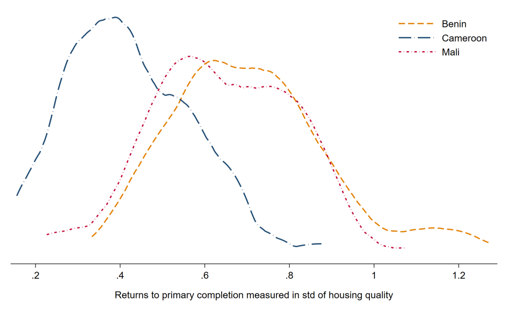
Returns to education vary over space. Return

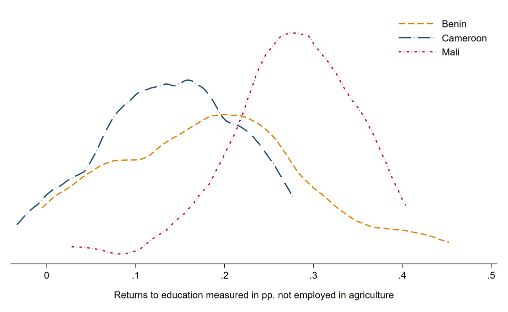
- From the literature: *Eckert and Kleineberg (WP)* and *Chetty and Hendren (2018)* in the US, *Edmonds, Pavcnik, Topalova (2010)*, and *Adukia, Asher, and Novosad (2020)* in India, *Atkin (2016)* in Mexico, and *Hsiao (WP)*, in Indonesia.

Returns to education vary over space. Return

- From the literature: *Eckert and Kleineberg (WP)* and *Chetty and Hendren (2018)* in the US, *Edmonds, Pavcnik, Topalova (2010)*, and *Adukia, Asher, and Novosad (2020)* in India, *Atkin (2016)* in Mexico, and *Hsiao (WP)*, in Indonesia.

- Using my data: Returns to education by birth location are significantly disperate over space.





Sufficient statistic result: Intuition. Return.

Resulting sufficient statistic relationship

$$\mu_{it} = \gamma_1 \cdot \ln(MA_{it}^E) + \gamma_2 \cdot \ln(MA_{it}^N) + \gamma_3 \cdot \ln(LMA_{it}^E) + \gamma_4 \cdot \ln(LMA_{it}^N) + v_{it}$$

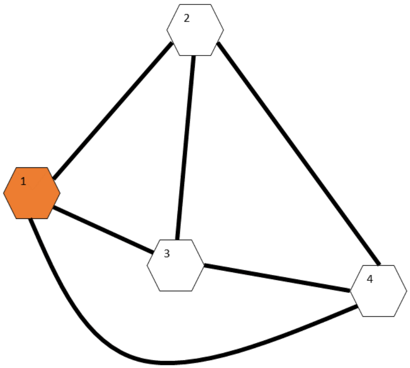
E type market access, MA_{it}^E, affects local opportunity through five channels.

1. Directly increases the demand for E-type goods and therefore w^E and so returns to education.
2. Decreases the local price level inducing E type migration and so reducing w^E and RtE.
3. Higher wages increase MC and so prices, reducing demand for local goods and dampening wages.
4. Higher wages induce more $E - type$ migration, dampening wages.
5. Higher E wages increase RtE inducing greater migration, may mitigate or multiply affects.

Theoretically and empirically (1) dominates (2), and as (3),(4),(5) attenuate or multiply existing effects in sum we have $\gamma_1 > 0$.

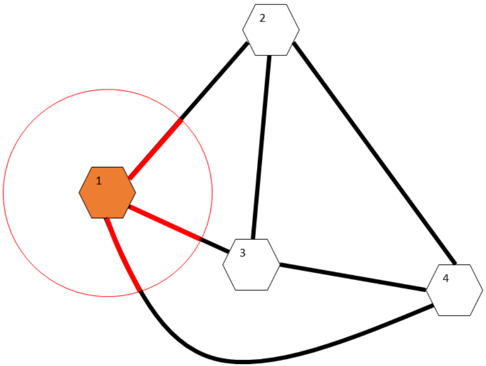
Can use similar arguments to find $\gamma_2 < 0$, $\gamma_3 < 0$, and $\gamma_4 > 0$.

Empirical strategy — Endogeneity, building MA instruments.



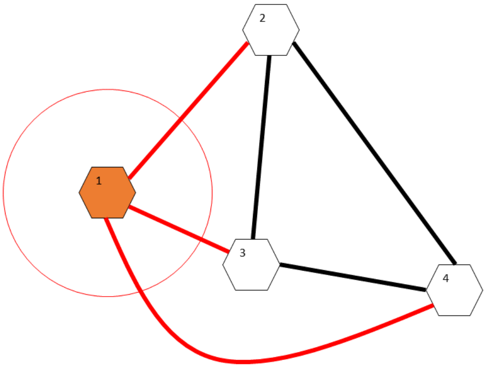
- **IV:** None.
- **Threats**: Any.
- [The maths](#).

Empirical strategy — Endogeneity, building MA instruments. Return.



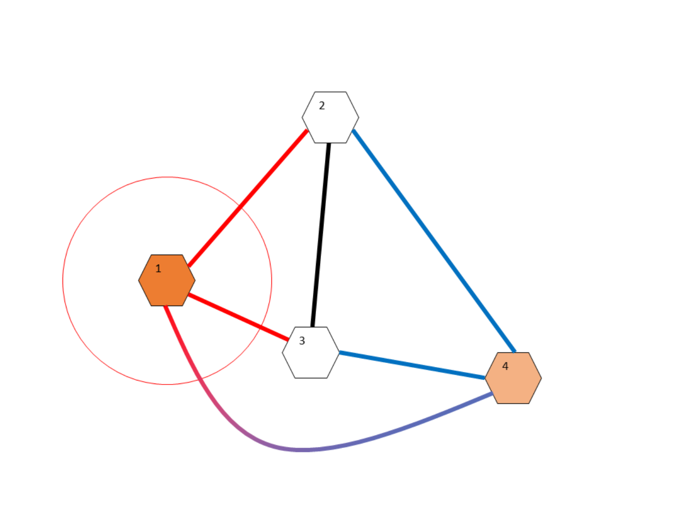
- **IV:** Only consider far away variation.
- **Threats**: (i) long connections (ii) far away connections (iii) indirect improvements.
- The maths.

Empirical strategy — Endogeneity, building MA instruments. Return.



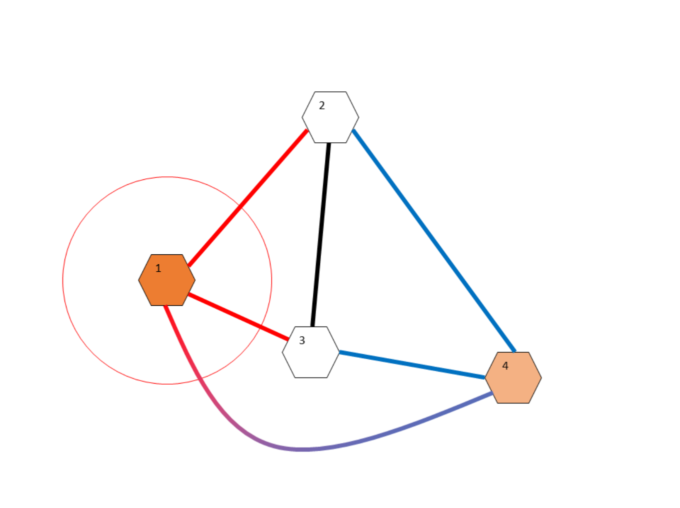
- **IV:** Only consider far away variation and not-on-least-cost-path variation.
- **Threats**: (i) ~~long connections~~ (ii) ~~far away connections~~ (iii) indirect improvements.
- The maths.

Empirical strategy — Endogeneity, building MA instruments. Return.



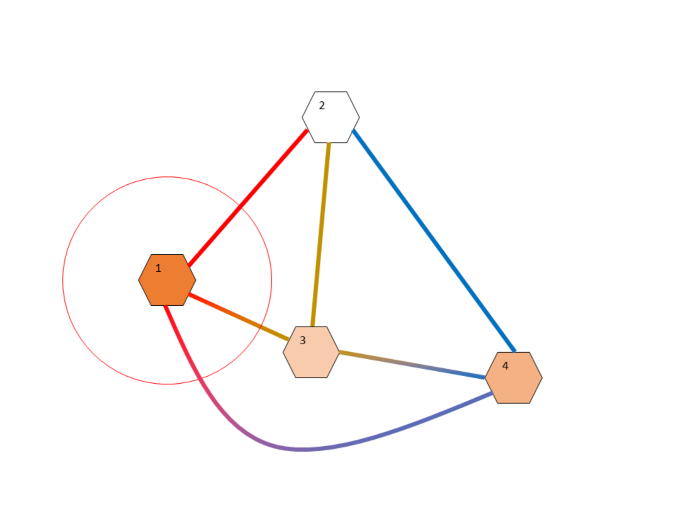
- **IV:** Only consider far away variation and 2nd order not-on-least-cost-path variation.
- **Threats**: (i) ~~long connections~~ (ii) ~~far away connections~~ (iii) ~~indirect improvements.~~
- The maths.

Empirical strategy — Endogeneity, building MA instruments. Return.



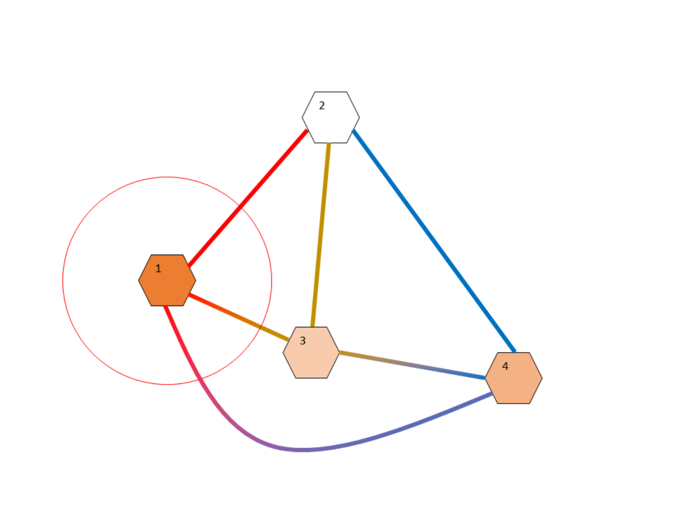
- **IV:** Only consider far away variation and 2nd order not-on-least-cost-path variation.
- **Threats**: (i) ~~long connections~~ (ii) ~~far away connections~~ (iii) ~~indirect improvements~~ (iv) second order indirect improvements.
- The maths.

Empirical strategy — Endogeneity, building MA instruments. Return.



- **IV:** Only consider far away variation and 3rd order not-on-least-cost-path variation.
- **Threats**: (i) ~~long connections~~ (ii) ~~far away connections~~ (iii) ~~indirect improvements~~ (iv) ~~second order indirect improvements.~~
- The maths.

Empirical strategy — Endogeneity, building MA instruments. Return.



- **IV:** Only consider far away variation and 3rd order not-on-least-cost-path variation.
- **Threats**: (i) ~~long connections~~ (ii) ~~far away connections~~ (iii) ~~indirect improvements~~ (iv) ~~second order indirect improvements~~ (v) third order indirect improvements.
- The maths.

Far away and not-on-least-cost-path variation — the maths

Consider some change in transport costs from $\{\tau_{ij}\}_{Ii,j}$ to $\{\tau'_{ij}\}_{i,j}$. Denote new variables with prime, so $MA_i^1 = \sum_j \tau'_{ij} \frac{Y_j}{MA_j^1}$.

- Not-on-least-cost-path variation: $MA_i^2 = \sum_j \tau_{ij} \frac{Y_j}{MA_j^1}$
- g-th order not-on-least-cost-path variation $MA_i^g = \sum_j \tau_{ij} \frac{Y_k}{MA_j^{g-1}}$

Return

Education supply. Return.

Thus far I have focused on education demand, but I can also incorporate education supply *(Hsiao (2022))*:

Education supply. Return.

Thus far I have focused on education demand, but I can also incorporate education supply *(Hsiao (2022))*:

- Endogenise education cost as depending on distance to the closest school *(DeStefano, et al. (2007), Prakash, et al. WP)* which decreases with population: $c_{it} = \tilde{c} L_{it}^{-\chi}$.

- Alternatively endogenise education supply following *Khanna (2022)*. Central planner builds schools to maximise coverage $\rightarrow$ builds more schools in denser areas with better transport networks. Details.

- Education *quality*, is also an important dimension.
  - Areas that attract E-type individuals also attract teachers.
  - Control directly for education quality (proportion completing primary school but remaining illiterate).

Education supply. Return.

Thus far I have focused on education demand, but I can also incorporate education supply *(Hsiao (2022))*:

- Endogenise education cost as depending on distance to the closest school *(DeStefano, et al. (2007), Prakash, et al. WP)* which decreases with population: $c_{it} = \tilde{c}L_{it}^{-\chi}$.
  - $\rightarrow$ Included in sufficient statistic result.
- Alternatively endogenise education supply following *Khanna (2022)*. Central planner builds schools to maximise coverage $\rightarrow$ builds more schools in denser areas with better transport networks. Details.
  - $\rightarrow$ Sufficient statistic result should also include local transport connections. Find not significant.
- Education *quality*, is also an important dimension.
  - Areas that attract E-type individuals also attract teachers.
    - $\rightarrow$ Included in the sufficient statistic result.
  - Control directly for education quality (proportion completing primary school but remaining illiterate).
    - $\rightarrow$ Not significant and doesn't impact estimates.

How can we predict the effects of better connecting two given locations?

- The first order effect of connecting any two locations is to strengthen pre-existing trade and migration flows.

How can we predict the effects of better connecting two given locations?

- The first order effect of connecting any two locations is to strengthen pre-existing trade and migration flows.

- If only a small proportion of i's exports of E-type goods come from j then connecting i and j will not have much effect on demand for E-type goods in i.

How can we predict the effects of better connecting two given locations?

- The first order effect of connecting any two locations is to strengthen pre-existing trade and migration flows.

- If only a small proportion of i's exports of E-type goods come from j then connecting i and j will not have much effect on demand for E-type goods in i.

- Helps explain the previous result: main-periphery connections decrease inequality of opportunity more than periphery-periphery connections.

Conclusion.

What is not consistent with the sufficient statistic result. Return.

- Models where locations don't admit a single aggregate good representation in each sector.
- Models that don't exhibit constant demand and supply elasticities.
- Forward looking behavior.
- More general preferences & non-homotheticities result in a non-linear version of the SS result.

How can we predict the effects of better connecting two given locations? Return.

- Consider the counterfactual of building a road between i and k.

- Recall: $\widehat{MA}_i^n = \sum_j \widehat{\rho}_{ij}^n \lambda_{ij}^n \prod_{h=1}^4 \left(\widehat{MA}_j^h\right)^{b_{nh}}$.

How can we predict the effects of better connecting two given locations? Return.

- Consider the counterfactual of building a road between i and k.

- Recall: $\widehat{MA}_i^n = \sum_j \widehat{\rho}_{ij}^n \lambda_{ij}^n \prod_{h=1}^4 \left(\widehat{MA}_j^h\right)^{b_{nh}}$. Suppose $\hat{\rho}_{pq}^n = 0$ for all $p \neq i$ and $q \neq k$ i.e. only capture first order effects of improving the connection between i and k.

How can we predict the effects of better connecting two given locations? Return.

- Consider the counterfactual of building a road between i and k.

- Recall: $\widehat{MA}_i^n = \sum_j \widehat{\rho}_{ij}^n \lambda_{ij}^n \prod_{h=1}^4 \left(\widehat{MA}_j^h\right)^{b_{nh}}$. Suppose $\hat{\rho}_{pq}^n = 0$ for all $p \neq i$ and $q \neq k$ i.e. only capture first order effects of improving the connection between i and k.

- Then we have: $\widehat{MA}_i^n = \hat{\rho}_{ik}^n \lambda_{ik}^n \left(\prod_{h=1}^4 \left(\widehat{MA}_k^h\right)^{b_{nh}}\right)$ which plugging into the sufficient statistic result gives:

$$\hat{\mu}_i = \sum_{n=1}^{4} \gamma_n \left(\ln\left(\hat{\rho}_{ik}^n\right) + \ln\left(\lambda_{ik}^n\right) + \ln\left(\prod_{h=1}^{4} \left(\widehat{MA}_k^q\right)^{b_{nh}} \right)\right)$$

How can we predict the effects of better connecting two given locations? Return.

- Consider the counterfactual of building a road between i and k.

- Recall: $\widehat{MA}_i^n = \sum_j \widehat{\rho}_{ij}^n \lambda_{ij}^n \prod_{h=1}^4 \left(\widehat{MA}_j^h\right)^{b_{nh}}$. Suppose $\hat{\rho}_{pq}^n = 0$ for all $p \neq i$ and $q \neq k$ i.e. only capture first order effects of improving the connection between i and k.

- Then we have: $\widehat{MA}_i^n = \hat{\rho}_{ik}^n \lambda_{ik}^n \left(\prod_{h=1}^4 \left(\widehat{MA}_k^h\right)^{b_{nh}}\right)$ which plugging into the sufficient statistic result gives:
$$\hat{\mu}_i = \sum_{n=1}^4 \gamma_n \left(\ln\left(\hat{\rho}_{ik}^n\right) + \ln\left(\lambda_{ik}^n\right) + \ln\left(\prod_{h=1}^4 \left(\widehat{MA}_k^q\right)^{b_{nh}}\right)\right)$$

- $\lambda_{ik}^n =$ proportion of i's trade/migration of type n which is due to $k \rightarrow$ current trade and migration.

Building a road between i and k is better for i if it gets more "goods" than "bads" from k. R.

Estimate: $\hat{\mu}_{ir} = \beta_1 \cdot \ln\left(\lambda^{MA(E)}_{i,k(i,r)}\right) + \beta_2 \cdot \ln\left(\lambda^{MA(N)}_{i,k(i,r)}\right) + \beta_3 \cdot \ln\left(\lambda^{LMA(E)}_{i,k(i,r)}\right) + \beta_4 \cdot \ln\left(\lambda^{LMA(N)}_{i,k(i,r)}\right) + \varepsilon_{ir}$

Building a road between i and k is better for i if it gets more "goods" than "bads" from k. R.

Estimate: $\hat{\mu}_{ir} = \beta_1 \cdot \ln\left(\lambda^{MA(E)}_{i,k(i,r)}\right) + \beta_2 \cdot \ln\left(\lambda^{MA(N)}_{i,k(i,r)}\right) + \beta_3 \cdot \ln\left(\lambda^{LMA(E)}_{i,k(i,r)}\right) + \beta_4 \cdot \ln\left(\lambda^{LMA(N)}_{i,k(i,r)}\right) + \varepsilon_{ir}$

	(1)	(2)	(3)	(4)
$\ln(\lambda^{MA(E)}_{ik})$	3.175***	1.837*	2.858***	**2.475****
	(0.914)	(1.070)	(0.928)	**(1.149)**
$\ln(\lambda^{MA(N)}_{ik})$	-2.694**	-0.858	-2.321	**-2.718**
	(1.298)	(1.657)	(1.432)	**(1.775)**
$\ln(\lambda^{LMA(E)}_{ik})$	-10.52***	-8.784***	-5.927***	**-5.595****
	(2.306)	(2.936)	(2.190)	**(2.662)**
$\ln(\lambda^{LMA(N)}_{ik})$	4.189*	1.947	1.368	**3.012**
	(2.403)	(2.865)	(2.959)	**(2.979)**
Country FE	X	X	X	X
Population weighted		X		X
Region FE			X	X
R^2	0.0881	0.117	0.248	0.232
N	396	368	368	368

How might future road building effect inequality of opportunity over space?. Return.

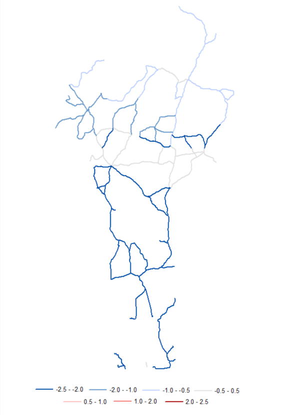
How might future road building effect inequality of opportunity over space?. Return.

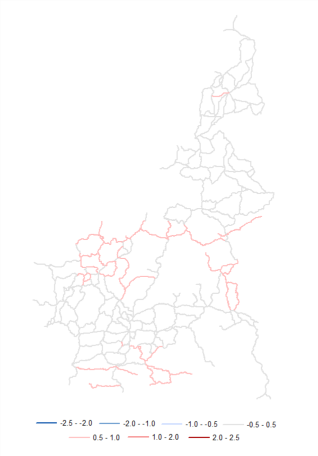
- Most changes to the network in the last 20 years have been upgrades $\rightarrow$ restrict attention to upgrading existing roads (*Balboni 2021*).

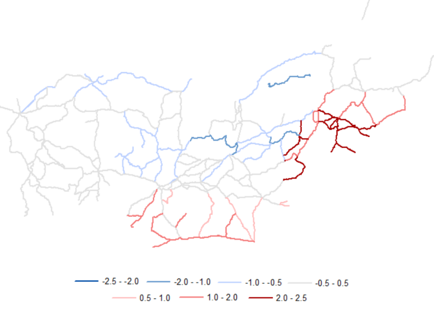
- Upgrade each road r, and calculate the resulting change in spatial inequality of opportunity (variance) $\Delta\mathbb{V}(\mu_i)_r$, keeping the rest of the network constant at 2019 level.

- Upgrade each road to have a speed of 80km/h.

Future road upgrading can increase or decrease inequality of opportunity over space. Return.







Plot the change in the variance of inequality over space due to upgrading a given road.

Within-network (country) heterogeneity. Return.

Differences by road type:

1. main = connects main city to the periphery.
2. periphery = connects two periphery areas.
3. Other

Regression: $\Delta\mathbb{V}(\mu_i)_r = \beta_r \text{RoadType}_r + \gamma_{c(r)} + \varepsilon_r$

Within-network (country) heterogeneity. Return.

Differences by road type:

1. main = connects main city to the periphery.
2. periphery = connects two periphery areas.
3. Other

Regression: $\Delta\mathbb{V}(\mu_i)_r = \beta_r \text{RoadType}_r + \gamma_{c(r)} + \varepsilon_r$

	(1) Overall	(2) Benin	(3) Cameroon	(4) Mali
Main	0.145*	-0.0194	-0.0492	0.247
	(0.0831)	(0.0816)	(0.0598)	(0.166)
Periphery	0.310***	0.954***	0.00485	0.421***
	(0.0643)	(0.101)	(0.0435)	(0.126)
Observations	534	94	260	180
R^2	0.500	0.147	0.003	0.041

Cross-network (country) heterogeneity. Return.

- First-order difference between Benin, Cameroon, and Mali is their scale. Average travel time between any two locations is: 369, 698, and 1176 minuets in Benin, Cameroon, and Mali.

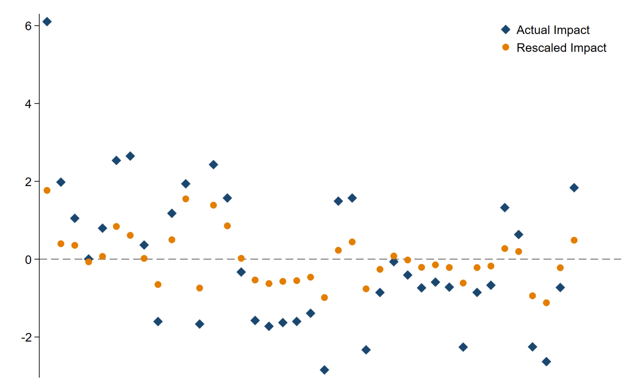
Cross-network (country) heterogeneity. Return.

- First-order difference between Benin, Cameroon, and Mali is their scale. Average travel time between any two locations is: 369, 698, and 1176 minuets in Benin, Cameroon, and Mali.
- Re-run counterfactuals re-scaling Benin & Cameroon so their averages are now equal to Mali.

Cross-network (country) heterogeneity. Return.

- First-order difference between Benin, Cameroon, and Mali is their scale. Average travel time between any two locations is: 369, 698, and 1176 minuets in Benin, Cameroon, and Mali.
- Re-run counterfactuals re-scaling Benin & Cameroon so their averages are now equal to Mali.

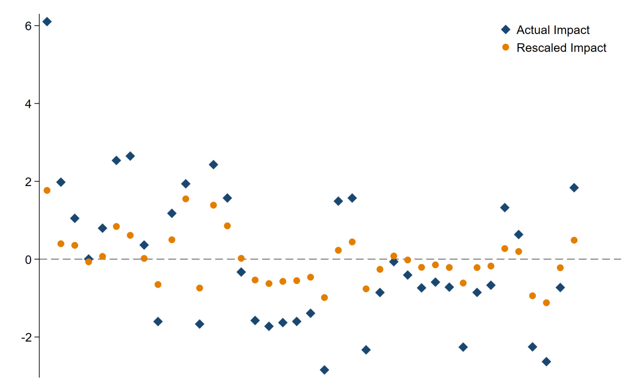
Locality level effects of a random road



Cross-network (country) heterogeneity. Return.

- First-order difference between Benin, Cameroon, and Mali is their scale. Average travel time between any two locations is: 369, 698, and 1176 minuets in Benin, Cameroon, and Mali.
- Re-run counterfactuals re-scaling Benin & Cameroon so their averages are now equal to Mali.

Locality level effects of a random road



Average effect over all roads on the variance of opportunity

	Actual	Same scale as Mali
Benin	-1.49	-0.38
Cameroon	0.18	0.00
Mali	0.26	

Inference. Return.

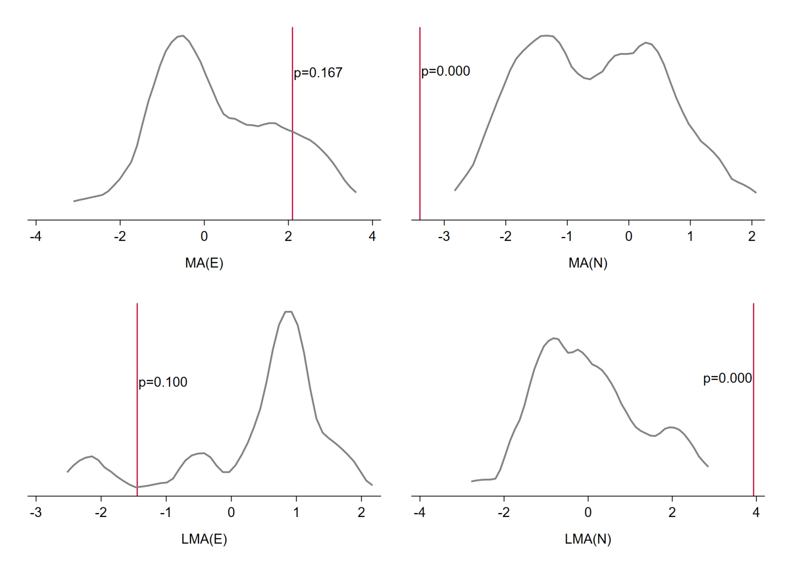
1. Problem: Serial correlation. Solution: Clustering.
2. Problem: Spatial correlation. Solution: Conley SE.
3. Problem: μ is estimated. Solution: classical measurement error $\rightarrow$ coefficients attenuated.
4. Problem: Dependencies introduced by MA terms. Solution: Permutation inference as suggested by *Borusyak and Hull (2021)*.

Inference

	Coefficient	Unadjusted	Clustered	Conley d=100	Conley d=100,l=30	Conley d=1000,l=infinity
Log(LMA Educ)	-0.0819	0.0442	0.0566	0.0513	0.0584	0.0588
Log(MA Educ)	0.0458	0.0163	0.0218	0.0186	0.0205	0.0254
Log(LMA No Educ)	0.1662	0.0450	0.0423	0.0331	0.0383	0.0561
Log(MA No Educ)	-0.0886	0.0214	0.0261	0.0183	0.0217	0.0280

Notes: This table shows the results from performing various inference procedures. In column one I report the coefficients from estimating equation **??** by OLS. Column two reports the unadjusted standard errors associated with these coefficients. Column three adjusts standard errors by clustering at the locality level. Column four reports Conley standard errors with a distance cut off of 100km. Column five reports Conley standard errors with a distance cut off of 100km and allowing auto-correlation up to 30 years. Finally, column six reports Conley standard errors with a distance band of 1000km and allowing general autocorrelation.

Inference

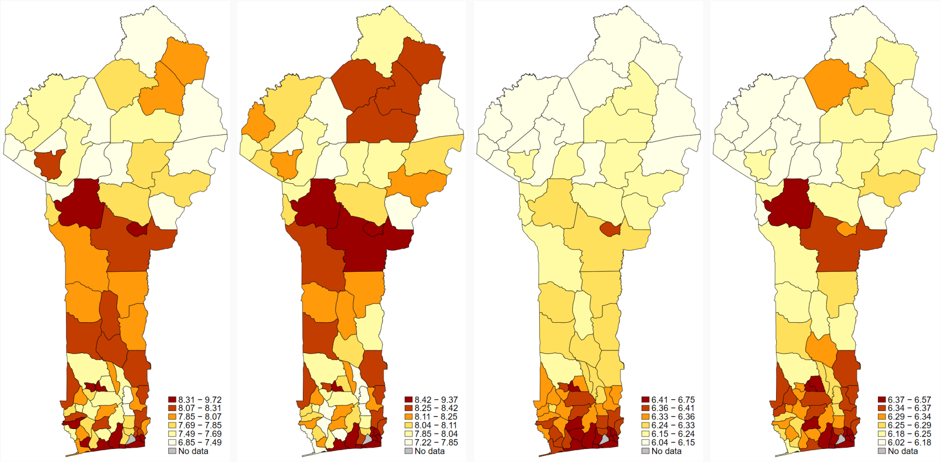

p=0.167
-4
-2
0
2
4
MA(E)
p=0.000
-3
-2
-1
0
1
2
MA(N)
p=0.100
-3
-2
-1
0
1
2
LMA(E)
p=0.000
-4
-2
0
2
4
LMA(N)

Sufficient statistic result by country regressions. Return.

	Overall	Benin		Cameroon		Mali	
	Coef (O)	Coef (B)	Pval (O-B)	Coef (C)	Pval (O-C)	Coef (M)	Pval (O-M)
LMA E	-0.198	-0.175	0.883	-0.298	0.535	-0.228	0.800
	(0.077)	(0.153)		(0.160)		(0.117)	
MA E	0.118	0.017	0.222	0.123	0.815	0.043	0.142
	(0.028)	(0.082)		(0.023)		(0.050)	
LMA N	0.284	0.527	0.203	0.325	0.688	0.388	0.413
	(0.058)	(0.188)		(0.102)		(0.126)	
MA N	-0.171	-0.069	0.238	-0.186	0.712	-0.213	0.479
	(0.033)	(0.086)		(0.040)		(0.058)	

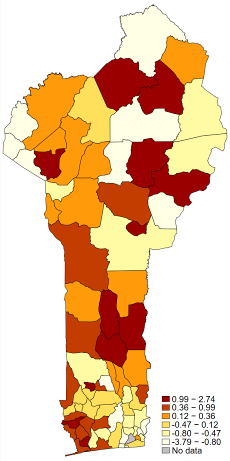
Maps of MA variation — raw. Return.

(a) Raw MA(E) **(b)** Raw MA(NE) **(c)** Raw LMA(E) **(d)** Raw LMA(NE)

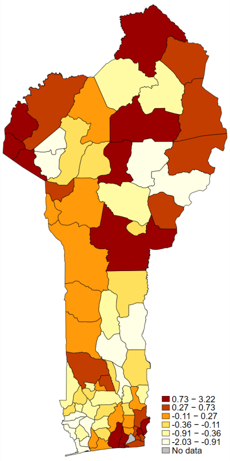


Maps of MA variation — residualised. Return.

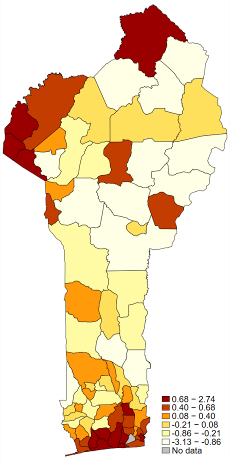
(a) Resid MA(E)



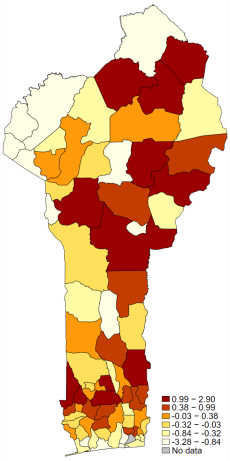
(b) Resid MA(NE)



(c) Resid LMA(E)

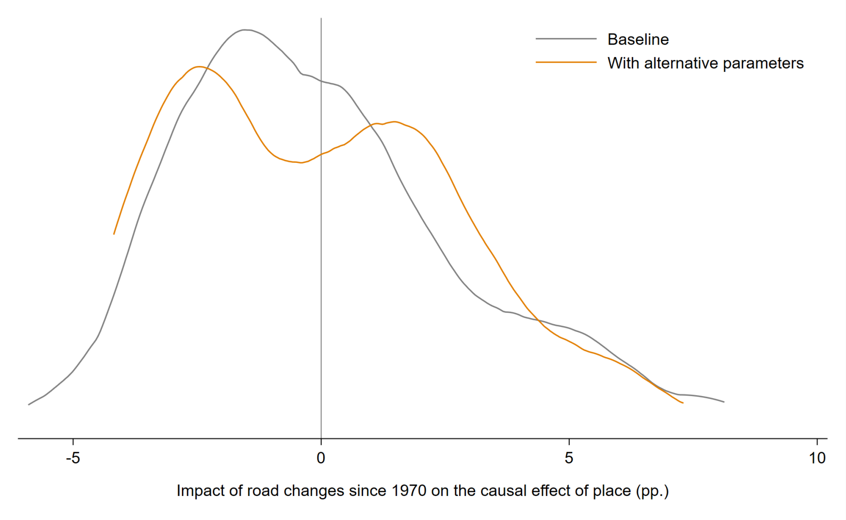


(d) Resid LMA(NE)

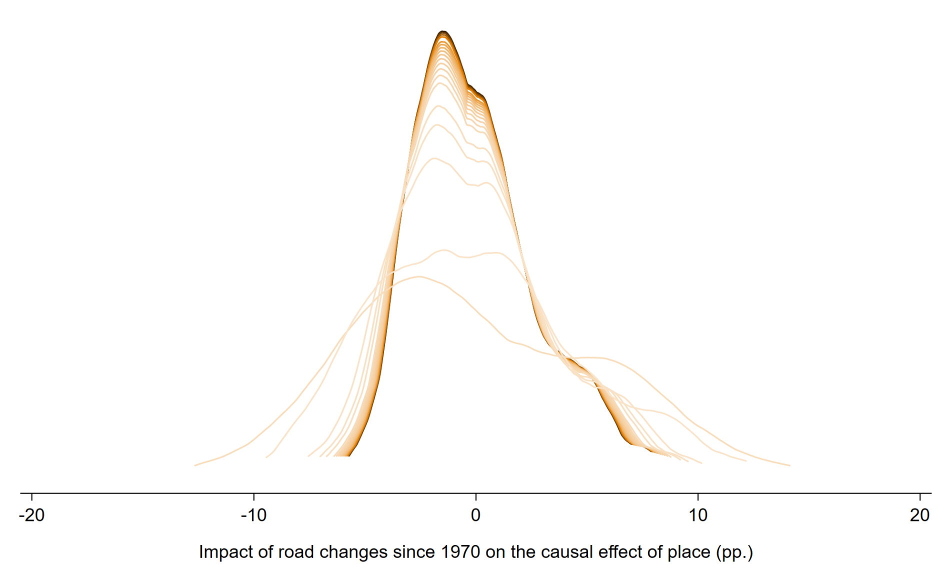


Alternative Calibration. Return

Following *Zarate (2020), Tsivanidis (2019), Morten and Oliveira (2021)* find: $\lambda_E = 1.74, \lambda_N = 2.11, \phi_E = 3.47,\ \phi_N = 4.52$. Correlation 0.87.

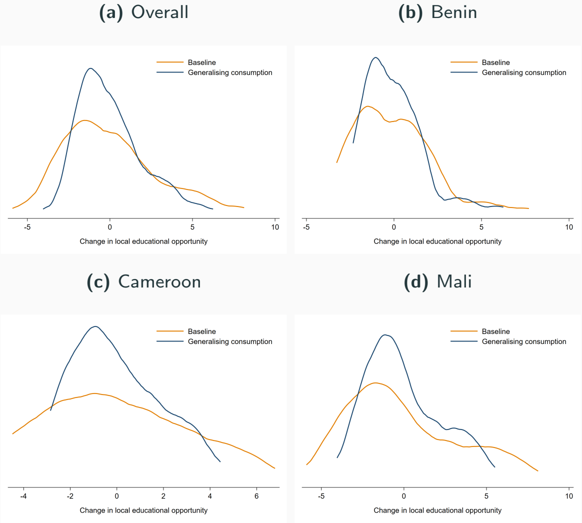


Alternative values of β. Return



Generalising consumption. Return

E-type consume 50-50, N-types only spend 10% on E-type goods.



Why should we care about the spatial distribution of local educational opportunity? Return

- If you happen to grow up in an area with high μ_i you're more likely to complete primary school than if you grew up in an area with low μ_i. Not the case if we consider observable variables.

Why should we care about the spatial distribution of local educational opportunity? Return

- If you happen to grow up in an area with high μ_i you're more likely to complete primary school than if you grew up in an area with low μ_i. Not the case if we consider observable variables.

- One measurable dimension proxying the general effect of place.
  - Most important margin of education — 30% complete primary school, 7% complete secondary school.
  - Correlated with later life opportunity more broadly defined — 67% less likely to work in agriculture.

Why should we care about the spatial distribution of local educational opportunity? Return

- If you happen to grow up in an area with high μ_i you're more likely to complete primary school than if you grew up in an area with low μ_i. Not the case if we consider observable variables.

- One measurable dimension proxying the general effect of place.
    - Most important margin of education — 30% complete primary school, 7% complete secondary school.
    - Correlated with later life opportunity more broadly defined — 67% less likely to work in agriculture.

- Normative statement: Care about inequality of opportunity from an equity/ fairness perspective.

Why should we care about the spatial distribution of local educational opportunity? Return

- If you happen to grow up in an area with high μ_i you're more likely to complete primary school than if you grew up in an area with low μ_i. Not the case if we consider observable variables.

- One measurable dimension proxying the general effect of place.
  - Most important margin of education — 30% complete primary school, 7% complete secondary school.
  - Correlated with later life opportunity more broadly defined — 67% less likely to work in agriculture.

- Normative statement: Care about inequality of opportunity from an equity/ fairness perspective.

- **Bottom line:** Giving weight to equality of opportunity changes road placement decisions.