The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010

Remi Jedwab, George Washington University
Adam Storeygard, Tufts University

Working Paper, National Bureau of Economic Research, USA (July 2016)
Journal of the European Economic Association (2022)

Pankaj Kumar

Objective

- How has intercity road upgrading affected local economic development in Sub-Saharan Africa?
  - How do these effects differ by context?

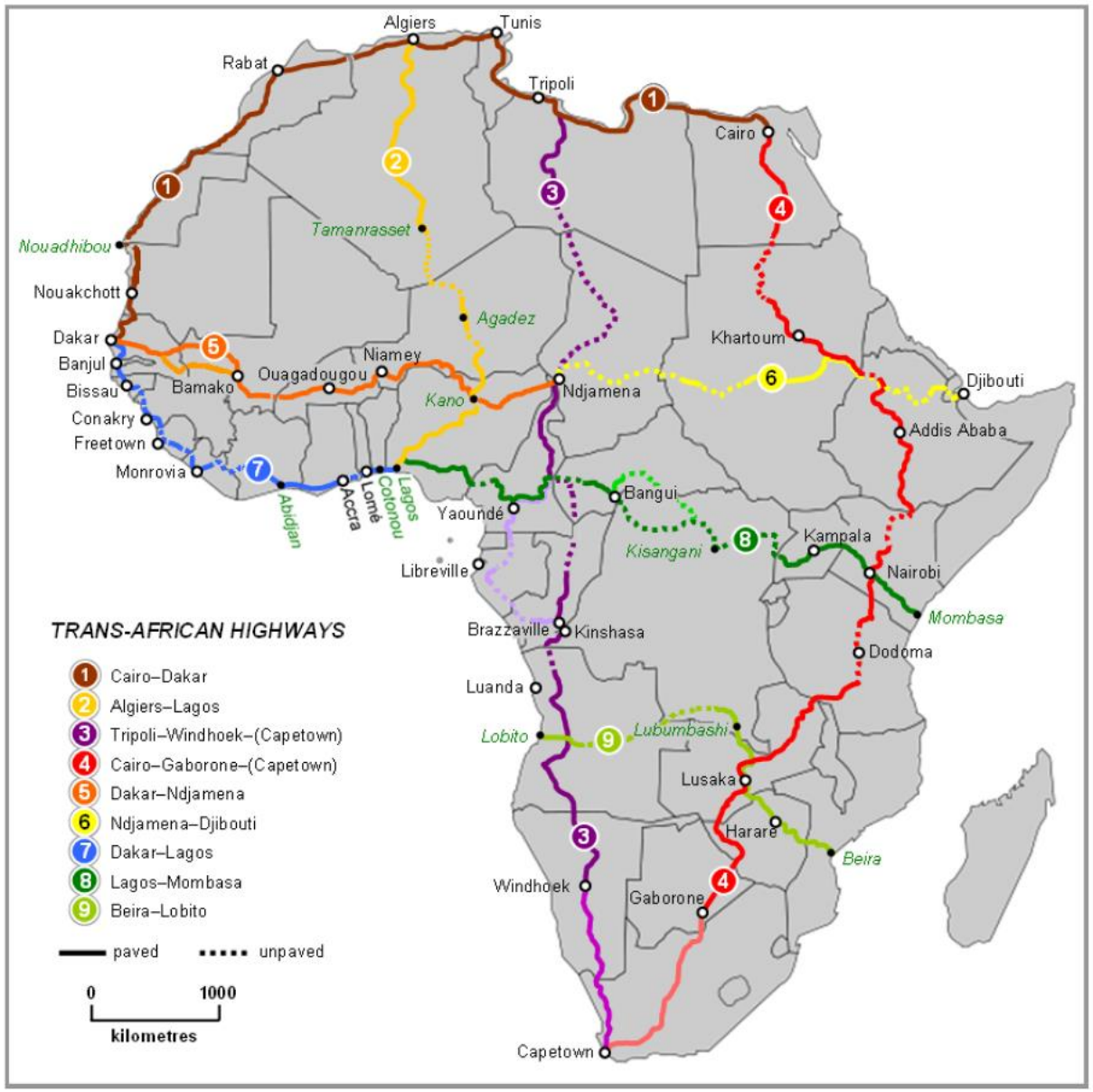
- Estimate the average effects of market access changes (as induced by road surface changes) on city population growth

- Investigate heterogeneous effects of road changes: remoteness, land suitability, ethnic homeland areas of heads of state

Background

- SSA: least urbanized world region
- 3.4km roads/1000 residents
- 0.7km paved roads/1000 residents

- 1960s-70s: Rapid road construction
- 1980s-90s: Slowed
- 2000 onwards: New international investments
  - World Bank, China
  - Large fraction of network still unpaved
  - Trans-African Highway network as coordinating mechanism: 55,000km of planned highways (vs. 1,000 km of highways c. 2012)



Related Literature

- Faber, Benjamin (2014): Impact of Trade cost reductions due to improved transportation infrastructure on distribution of economic activity

- Redding and Turner (2015): comprehensive literature on the idea of market access; explores the relationship between the spatial distribution of economic activity and transportation costs.

- Donaldson and Hornbeck (2016): Historical impacts of Railroads on US economy; impact of market access on agricultural land values

Related Literature

- Jaworski and Kitchens (2019): Effectiveness of policies aimed at integrating isolated regions by quantifying the relationship between market access and income from highways

- Casaburi et al. (2013): Impact of rural road infrastructure improvements on crop prices in rural markets in Sierra Leone.

- Micro road surface/quality impacts:
  - Gertler et al., 2015: estimate the impact of highway maintenance investments in Indonesia from 1990 to 2007
  - Asher and Novosad, 2020: estimated the impacts of India's national rural road construction program using a fuzzy regression discontinuity design and comprehensive household and firm census microdata

Research Novelty

- First systematic study of road-building across Africa
  - Scale: 140,000 km network, 39 countries, 1960-2010
  - Timing of the effects
  - Methodological: Instrumental Variable
  - Heterogeneous effects

- Not just building highways: paving and improving (gravelling)
  - Marginal changes more likely in most contexts

- Build new panel data set on road surface, city population and market access for 39 Sub-Saharan African countries 1960-2010

Data

GIS database of roads

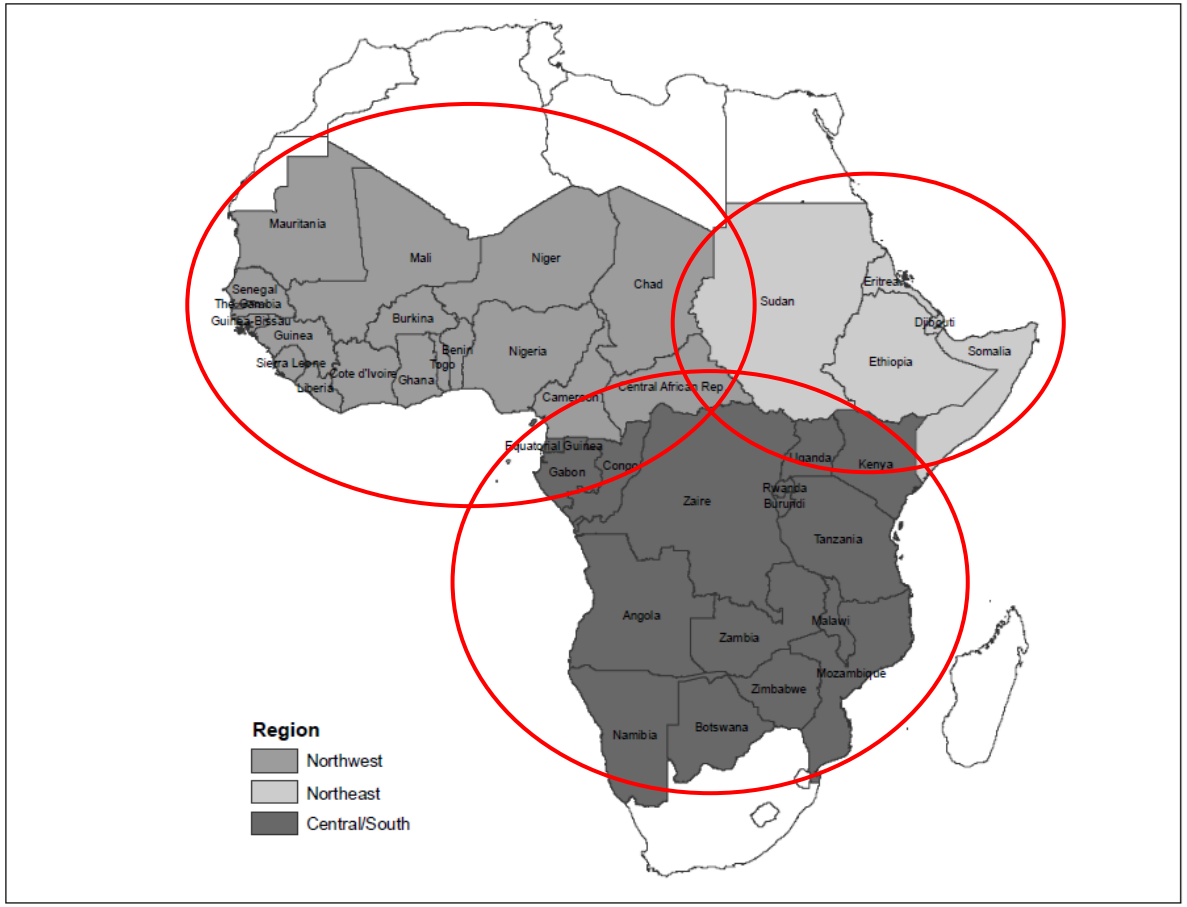
- Michelin paper road maps for 39 Sub-Saharan African countries from the early 1960s to date. Sources:
  - Government maps
  - Feedback from customers (large network of tire distributors and correspondents)

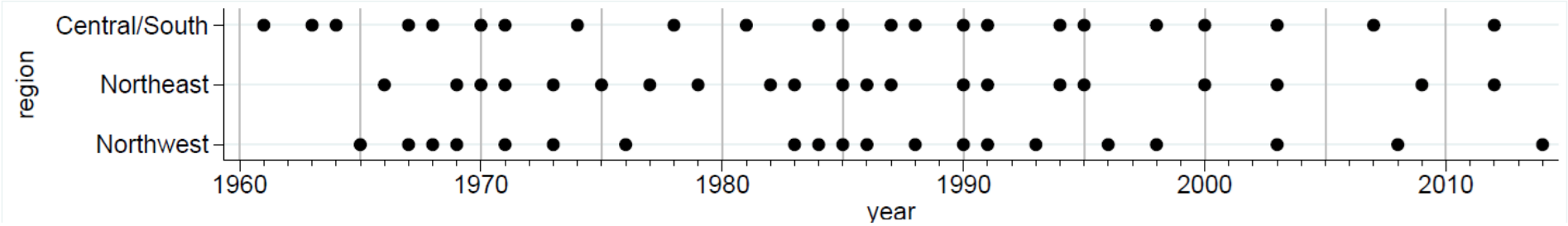
- Map ~ every 3 years, so 833 country-years

- Surface of each road: Highway, Paved, Improved and Dirt (vs. Primary, secondary, tertiary)

Data: Road categories

Michelin Road Map Countries and Years





Roads: 1960-2010

1960

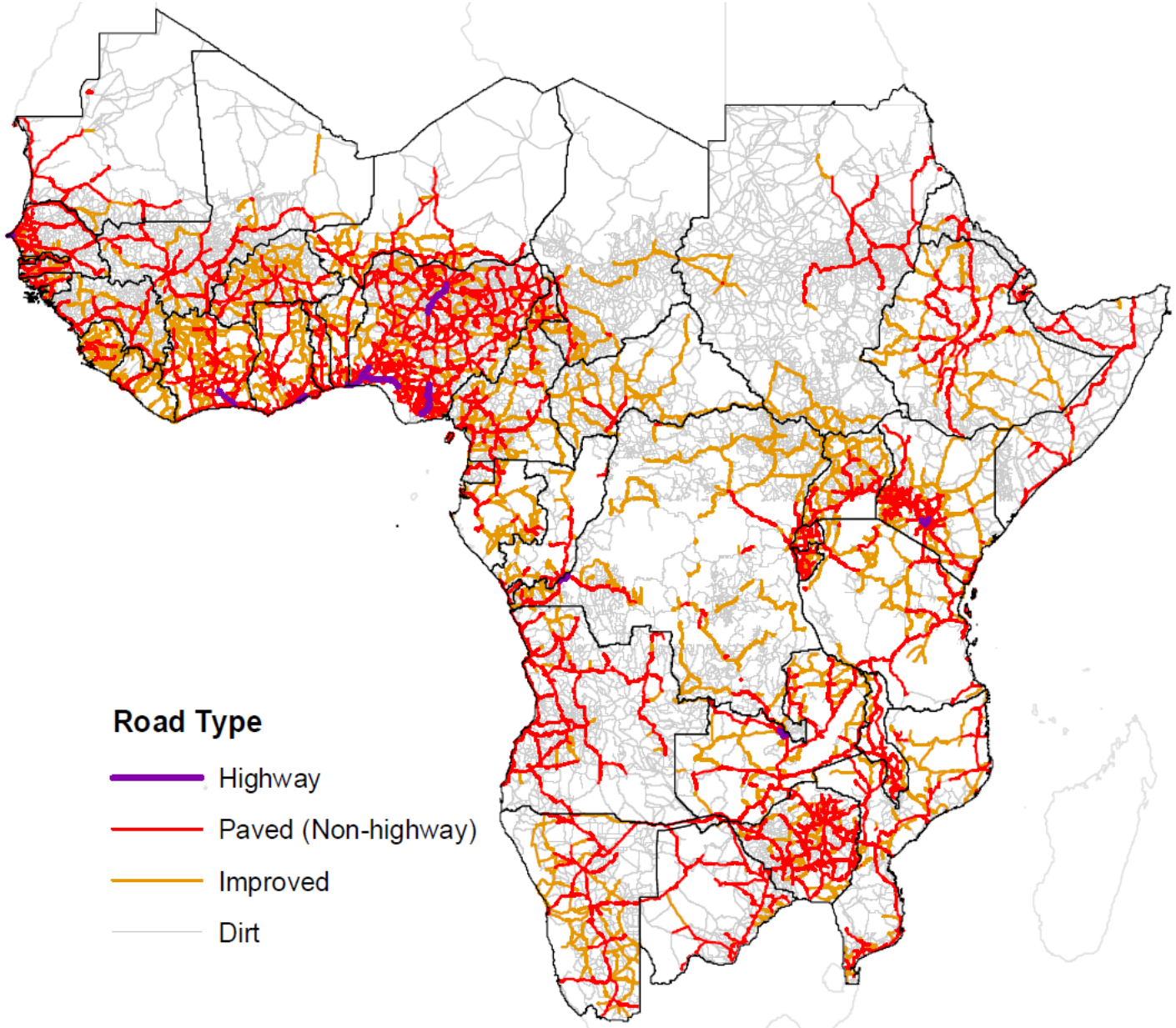
1970

1980

1990

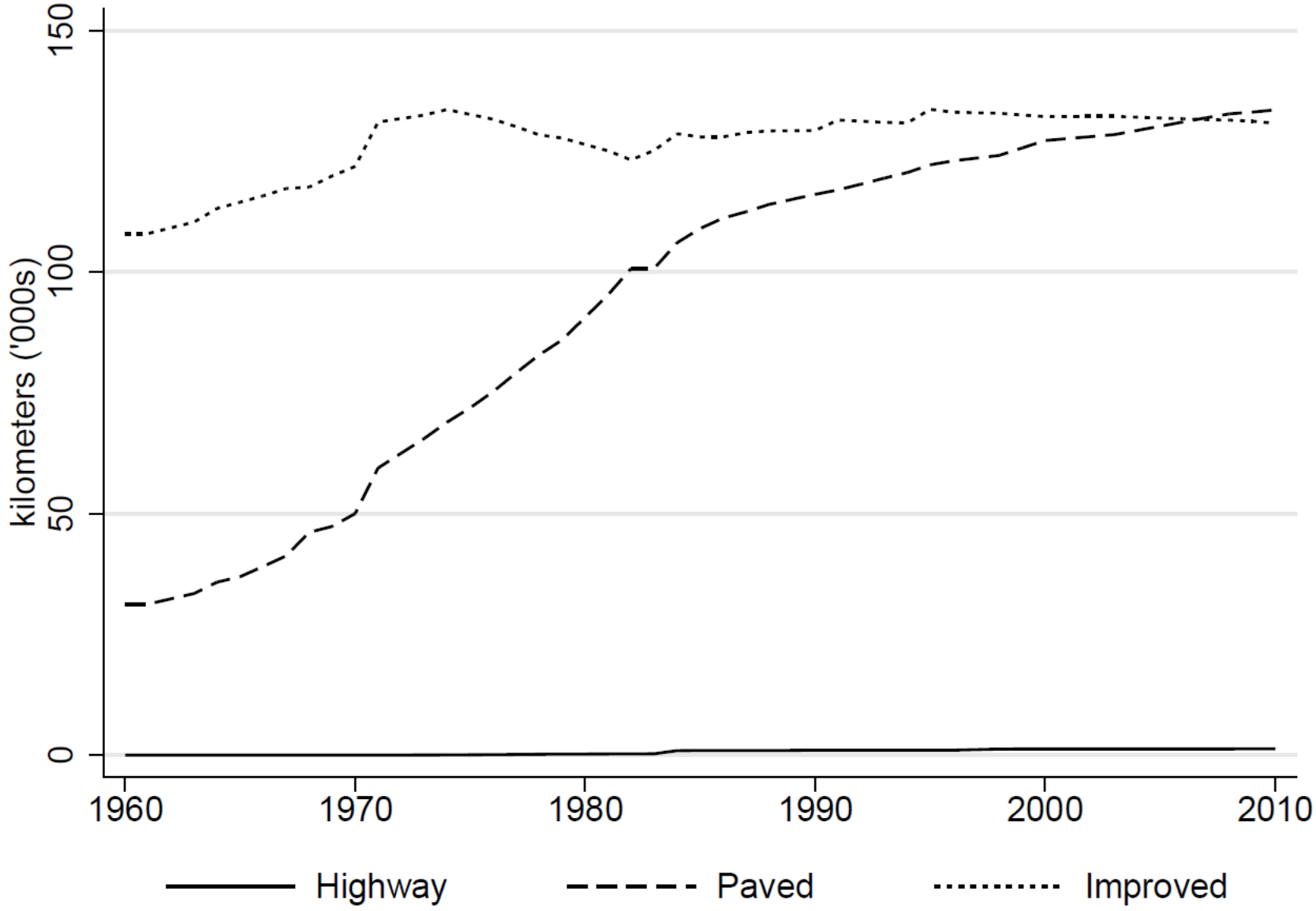
2000

2010



Data

Road Length in Sub-Saharan Africa (39 Countries)

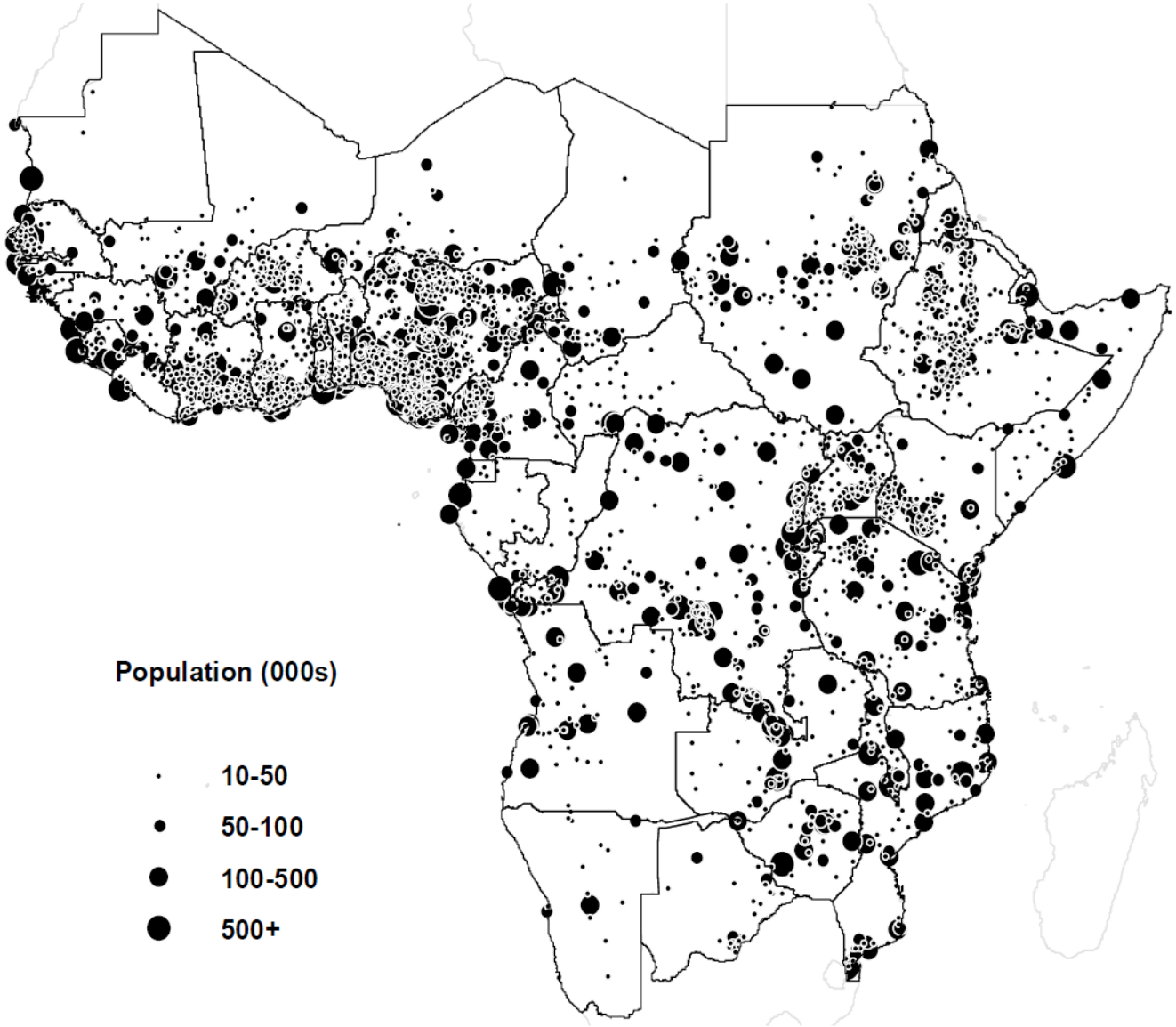


Data

GIS database of cities

- Population of localities ever above 10,000 for the same 39 countries in 1960, 1970, 1980, 1990, 2000 and 2010
- Proxy for local economic development in the absence of other data (no land prices, no systematic rural populations before 1990, no night lights before 1992).
- Sources: *Africapolis I* & *II* for 33 countries + Population Census data for 6 countries (similar methodology)
- Population estimates available for many city-year observations when below 10,000 (but not for all of them).

Cities (Pop>= 10000) (1960-2010)



1960

1970

1980

1990

2000

2010

Estimation

- Unit of Analysis: Grid squares
  - 0.1x0.1 degree (~11x11 km; due to computational constraints)

- Potentially multiple roads and cities per cell
  - Select best road in the cell:
    - Highway > paved > improved > dirt
  - Use the sum of city populations within cell

- Sample: 5906 city-years for 2127 cities (Pop > 10,000 in at least two consecutive years)
  - 4,725 city-years when including two lags

- Travel Speeds assumed to compute driving times: 80, 60, 40, 12, 6 kmph

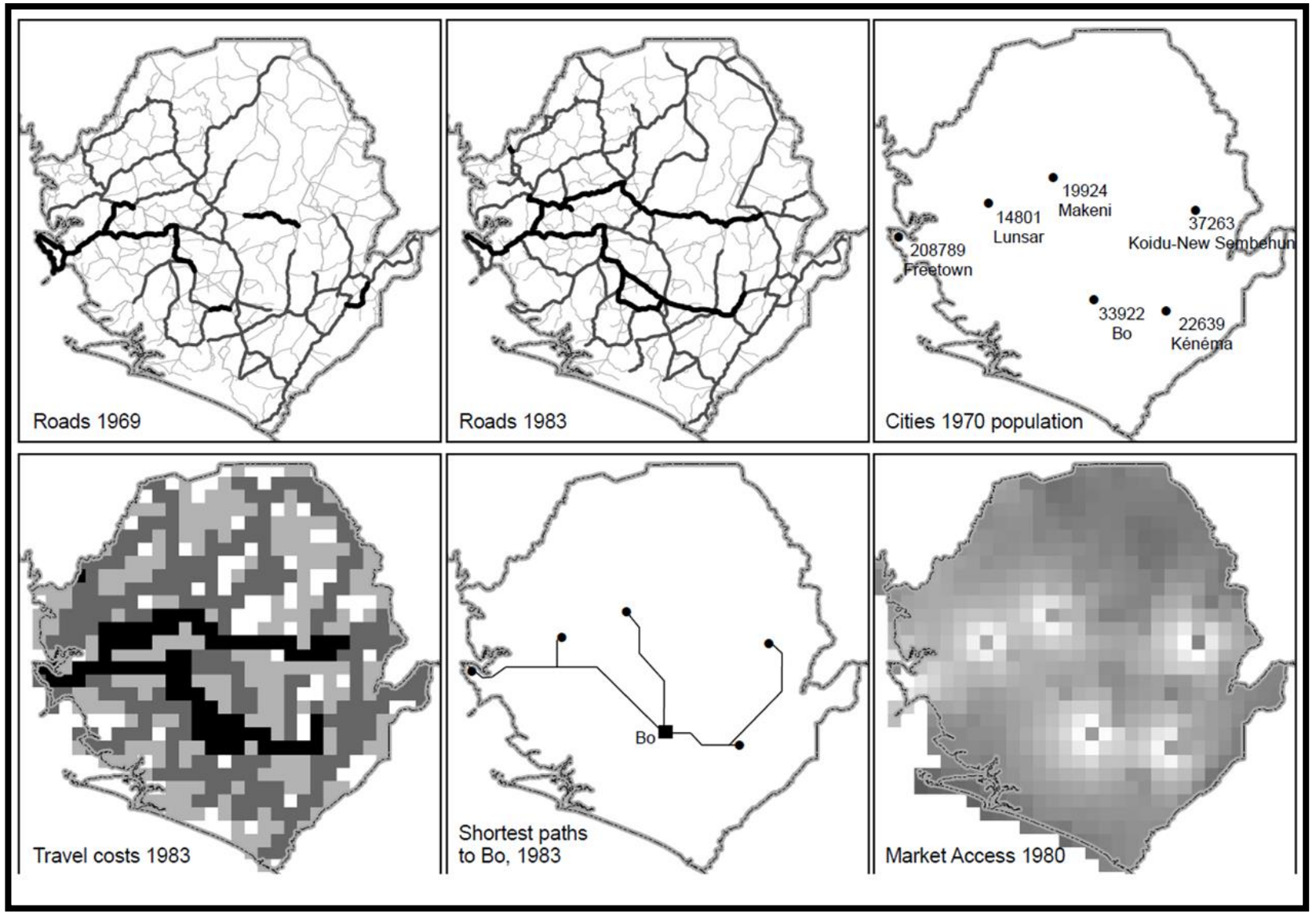
Estimation

Market Access

- Travel cost-discounted sum of the population of all other cities: $M_o = \Sigma_{d \neq o} P_d \tau_{od}^{-\theta}$ where
  - M_o is the market access of city *o*
  - P_d is the population of city *d*
  - $\tau_{od}^{-\theta}$ is the travel time from *o* to *d*, and
  - θ is the trade elasticity, baseline = 3.8

- Follows Donaldson and Hornbeck (2016) definition of market access
- Approximation to a recursive formulation that arises from Eaton-Kortum-type models.

Estimation: Sierra Leone, 1970-1980 (example)



Estimation

- Initial specification, how market access, MA affects urban population, P

$$\ln P_{ot} = \beta_0 \ln MA_{ot} + \lambda_o + \rho_{ct} + \varepsilon_{0ot}, \quad \text{.....(1)}$$

- In first differences (at 10-year intervals), cell fixed effects cancel and this becomes:

$$\Delta \ln P_{ot} = \beta_0 \Delta \ln MA_{ot} + \Delta\rho_{ct} + \Delta\varepsilon_{0ot}. \quad \text{.....(2)}$$

- Suppressing fixed effects and controls, stacking across all o, and defining the matrix Tt with off-diagonal elements in row o and column d equal to
- odt (and diagonal elements equal to zero), equation (1) becomes as follows:

$$\ln P_t = \beta_0 \ln(T_t P_t) + \varepsilon_{0t}, \quad \text{.....(3)}$$

A log-transformed spatial lag specification, where the log is applied elementwise

Estimation

Baseline specification

- Equation (3) in first differences can then be transformed as follows

$$\begin{aligned}
\Delta \ln P_t &= \beta_0 \Delta \ln(T_t P_t) + \Delta \varepsilon_{0t}, \\
&= \beta_0 \left[\ln(T_t P_t) - \ln(T_{t-10} P_{t-10})\right] + \Delta \varepsilon_{0t}, \\
&= \beta_0 \left[\ln(T_t P_t) - \ln(T_{t-10} P_{t-10}) + \ln(T_t P_{t-10}) - \ln(T_t P_{t-10})\right] + \Delta \varepsilon_{0t}, \\
&= \beta_0 \left[\ln(T_t P_t) - \ln(T_t P_{t-10})\right], \\
&\quad + \beta_0 \left[\ln(T_t P_{t-10}) - \ln(T_{t-10} P_{t-10})\right] + \Delta \varepsilon_{0t}.
\end{aligned} \qquad \ldots\ldots(4)$$

- Market Access is endogenous as city o's growth affects growth of other cities d

Estimation

Summary Statistics

- Baseline specification, 2 lags, N=4725

Main Variable:	Mean	Std. Dev.	Min	Max
Δ_{t-10}^{t} ln urban pop	0.318	0.209	-1.533	2.343
Δ_{t-10}^{t} ln *MA*	0.655	0.892	-8.236	10.618
Δ_{t-20}^{t-10} ln *MA*	0.901	1.099	-8.236	11.537
Δ_{t-30}^{t-20} ln *MA*	1.161	1.288	-8.236	13.291
ln urban pop$_{t-10}$	10.247	0.990	9.210	15.902

Estimation

Identification concerns

- Reverse causality
  - Roads built to cities expected to grow (or expected to lag).
- Omitted variables
  - Unobserved productivity shocks that drive road building and city growth.
- Measurement error
  - Speed assumptions are rough proxies, road quality, etc.

Estimation

- **Instrument fixing Population**

$$\Delta_R \ln MA_{ot} = \ln\left(\sum_{d\neq o} P_{d,\,\mathbf{t-10}}\,\tau_{o,d,\mathbf{t}}^{-\theta}\right) - \ln\left(\sum_{d\neq o} P_{d,\mathbf{t-10}}\,\tau_{o,d,\mathbf{t\text{-}10}}^{-\theta}\right). \qquad (5)$$

- **Instrument also Excluding Local Road Changes**

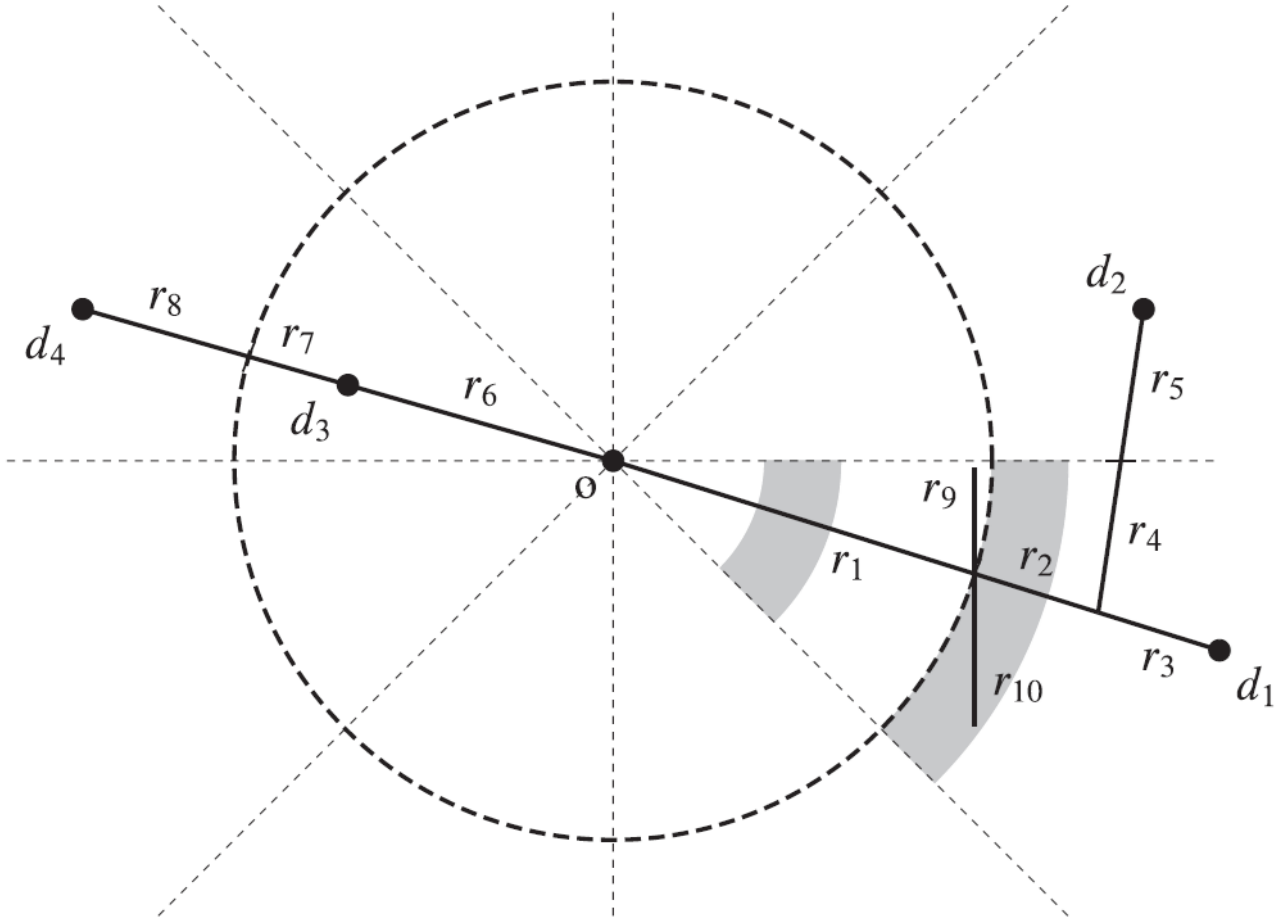
$$\Delta_R^{out,j} \ln MA_{ot} = \ln\left(\sum_{d:\delta(d,o)\geq j} P_{d,\mathbf{t-10}}\,\tau_{od,\mathbf{t}}^{-\theta} + \sum_{d:0<\delta(d,o)<j} P_{d,\mathbf{t-10}}\,\tau_{od,\mathbf{t-10}}^{-\theta}\right),$$

$$- \ln\left(\sum_{d\neq o} P_{d,\mathbf{t-10}}\,\tau_{od,\mathbf{t-10}}^{-\theta}\right), \qquad (6)$$

Estimation

Instruments:

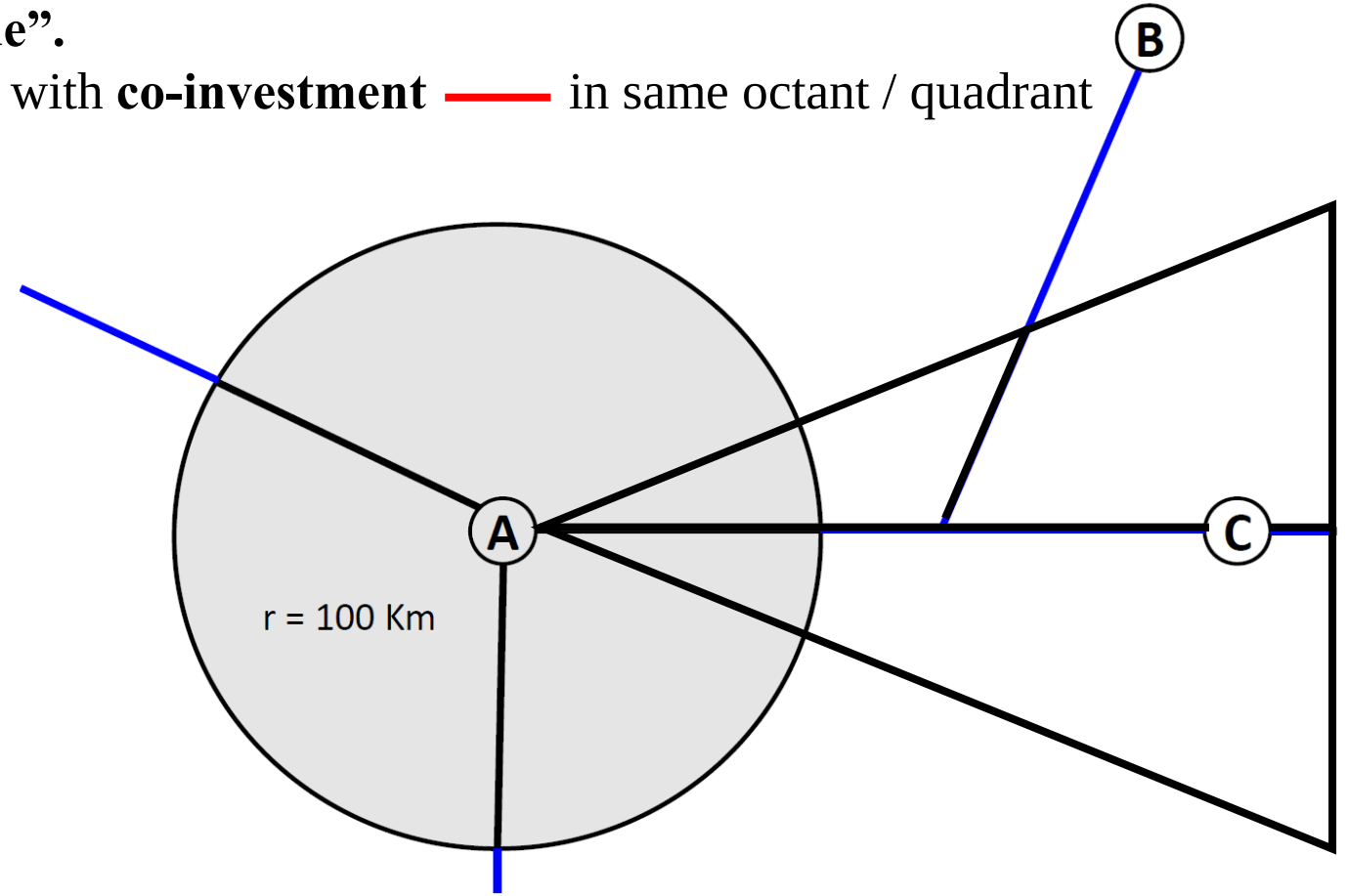
1. Population of all cities fixed at their initial levels.
2. Road changes outside the circle are exogenous



Estimation

Instrument Excluding Selected Non-local Road Changes

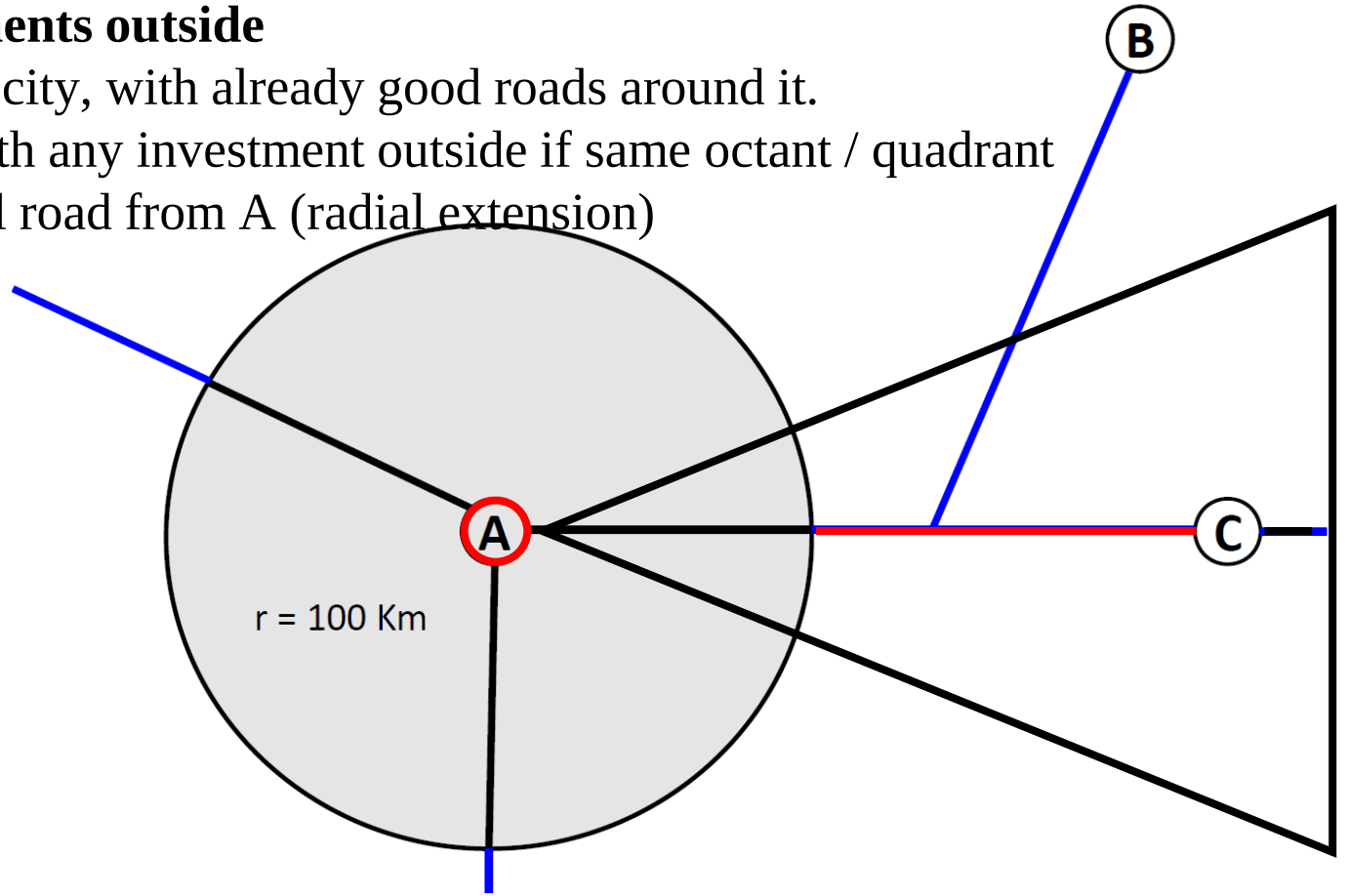
- **Non-local** road changes ——
- Non-local changes on A-C are still endogenous if there are **spatially correlated road investments "inside" and "outside".**
- Exclude observations with **co-investment** —— in same octant / quadrant



Estimation

Instrument Excluding Selected Non-local Road Changes

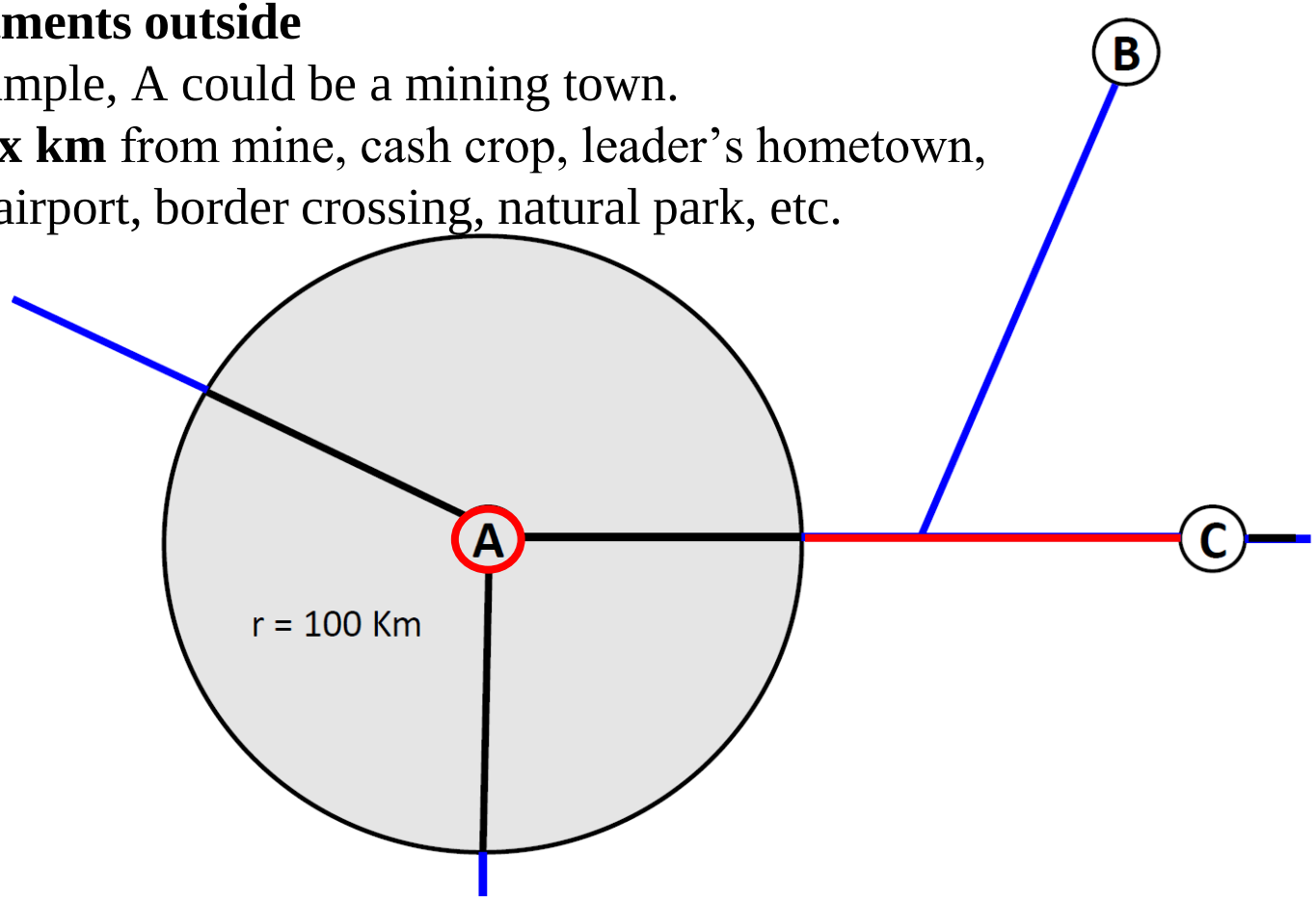
- **Non-local** road changes
- Non-local changes on A-C are also endogenous if there are correlated **non-road investments inside** and **road investments outside**
- A could be an important city, with already good roads around it.
- Exclude observations with any investment outside if same octant / quadrant as paved/improved radial road from A (radial extension)



Estimation

Instrument Excluding Selected Non-local Road Changes

- Only use **non-local** road changes
- Non-local changes on A-C are still endogenous if there are correlated **non-road investments inside** and **road investments outside**
- **Growth hubs**: For example, A could be a mining town.
- **Exclude cities within x km** from mine, cash crop, leader's hometown, regional capital , port, airport, border crossing, natural park, etc.



Results

Table 1: Average Effect of Market Access on Urban Population

Panel A: OLS	(1)	(2)	(3)	(4)	(5)
Δ^{t}_{t-10} ln Market Access	1.33***	1.33***	1.56***	1.57***	1.52***
	[0.38]	[0.42]	[0.37]	[0.46]	[0.41]
Δ^{t-10}_{t-20} ln Market Access		0.98***	1.18***	1.49***	1.10***
		[0.27]	[0.31]	[0.36]	[0.34]
Δ^{t-20}_{t-30} ln Market Access			0.73***	0.80**	0.76**
			[0.24]	[0.32]	[0.30]
Δ^{t-30}_{t-40} ln Market Access				0.31	
				[0.25]	
Δ^{t+10}_{t} ln Market Access					0.72
					[0.56]
Overall Effect	1.33***	2.31***	3.47***	4.18***	3.39***
($t-40$ to t)	[0.38]	[0.58]	[0.63]	[0.83]	[0.71]
Observations	5,906	5,472	4,725	3,630	2,607
Adj. R-squared	0.23	0.19	0.17	0.16	0.19

Results

Panel B: IV	IV: Exclude 5 (1)	IV: Exclude 10 (2)	IV: Exclude 15 (3)
Δ_{t-10}^{t} ln Market Access	3.09***	4.45**	5.55*
	[1.10]	[1.82]	[2.99]
Δ_{t-20}^{t-10} ln Market Access	3.04***	5.56***	6.68**
	[0.87]	[1.50]	[2.62]
Δ_{t-30}^{t-20} ln Market Access	2.23**	2.88**	4.27**
	[0.88]	[1.38]	[1.94]
Overall Effect	8.35***	12.89***	16.49***
($t-30$ to t)	[2.16]	[3.23]	[4.60]
Observations	4,725	4,725	4,725
1st stage Kleibergen-Paap F	98.56	42.94	11.90

Results

TABLE 2. Main robustness checks.

	OLS (1)	IV: Exclude 5 (2)	IV: Exclude 10 (3)	IV: Exclude 15 (4)
(1) Co-investment: inner: 2, outer: 15	3.65***	10.21***	14.39**	19.22*
(N = 2,260; F: _; 60.6; 12.2; 5.1)	[0.98]	[3.34]	[6.24]	[9.97]
(2) Radial extension outward	3.86***	9.60***	13.75***	17.93***
(N = 1,603; F: _; 76.6; 21.4; 5.1)	[0.92]	[2.44]	[3.89]	[6.20]
(3) Radial extension inward	3.11***	8.76***	10.94***	13.17***
(N = 2,867; F: _; 107.2; 52.9; 9.9)	[0.85]	[2.34]	[3.43]	[4.86]
(4) Excl. changes convex hull 100k+		8.00***	11.83***	14.01***
(N = 4,725; F: _; 43.7; 35.6; 12.0)		[2.19]	[3.07]	[4.59]
(5) Excl. Δtranscontinental road		8.22***	12.78***	14.43***
(N = 4,725; F: _; 44.6; 27.8; 6.7)		[2.32]	[3.16]	[5.12]
(6) Excl. nat'l, regional and top 5 cities	3.61***	7.73***	12.96***	15.30**
(N = 3,799; F: _; 87.8; 8.3; 7.3)	[0.75]	[2.56]	[4.56]	[6.63]
(7) Fix population to 1960 in IVs		7.54***	11.81***	16.20***
(N = 4,723; F: _; 49.4; 18.8; 4.6)		[1.90]	[2.73]	[3.97]
(8) Fix population to 1960 in *MA*	3.18***	9.83***	16.45***	24.71***
(N = 4,723; F: _; 134.3; 24.0; 9.7)	[1.10]	[2.54]	[3.95]	[6.04]
(9) Control for regional growth	2.34***	6.06***	7.97**	8.69*
(N = 3,498; F: _; 31.2; 8.2; 4.2)	[0.65]	[2.23]	[3.29]	[4.61]
(10) Province (1960)-year *FE*	1.34*	5.54	13.68*	13.13
(N = 4,725; F: _; 44.1; 5.5; 2.2)	[0.76]	[3.38]	[7.34]	[11.24]
(11) Quadrant-year *FE*	2.38***	6.28***	8.81**	7.87
(N = 4,725; F: _; 97.4; 16.8; 6.6)	[0.63]	[2.33]	[4.11]	[6.91]

Notes. This table is structured like Table 1 but only reports overall effects. Robust standard errors, clustered by 1960 province, are in brackets. *, **, *** = 10%, 5%, and 1% significance. FE = fixed effects.

Results

Table 3: Effect of Market Access on Night Lights

	(1) OLS	(2) IV: Excl. 5	(3) IV: Excl. 10	(4) IV: Excl. 15
Δ_{t-10}^{t} ln MA	0.39	22.29**	43.98***	69.37***
	[3.06]	[9.97]	[11.98]	[18.56]
Δ_{t-20}^{t-10} ln MA	1.70	12.52	8.82	5.84
	[2.90]	[8.12]	[12.50]	[16.91]
Δ_{t-30}^{t-20} ln MA	0.84	3.90	0.71	-3.01
	[2.11]	[4.32]	[7.33]	[10.21]
Overall Effect	2.93	38.70***	53.51***	72.19***
	[5.22]	[11.39]	[18.31]	[26.49]
1st stage Kleibergen-Paap F	53.24		29.74	9.837

Notes: See Table 1. Outcome variable is $100\Delta_{t-10}^{t}$ ln (Light Intensity). N = 3,591. Robust SEs, clustered by 1960 province, are in brackets. *, **, *** = 10, 5, 1% significance.

Results

Induced Urbanization vs. Reallocation, Natural Increase

- Cannot directly test whether increases are due to
  - *Induced rural-urban migration*
  - *Reallocation across cities*
  - *Natural increase (births and deaths)*

- Evidence consistent with primary role for induced urbanization
  - Restrict sample to initially less urbanized country-periods (< 10% or 7%; less scope for urban reallocation): similar results
  - Use mega-cells of 3-9 cells (33-99 km) (and drop those with national/ regional/ largest cities, since reallocation less likely between non-neighboring local cities). Noisy but still sizable effect
  - Test for direct effect on natural increase (Demographic and Health Survey data): suggests no more than half of overall effect

Results

Summary of Average Effects

- Naive effect of a 10% change in market access: ~0.1%-0.15% per decade for three decades (total 30-year effect: 0.3-0.4%).
- IV: ~0.8-1.3% over 30 years.
- Concentrated in first two decades (i.e. decade of construction and following decade)
- No measurable effect in fourth decade.
- Source of growth: rural areas, other cities, or natural increase?
  - Natural increase and urban reallocation unlikely to be large share of effect. Rural-urban migration must be large share.

Results

Table 5: Heterogeneous Effects of Market Access on Urban Population

	OLS	Col. (2)–(4): IV5			IV10	IV15
	Diff.	0	1	Diff.	Diff.	Diff.
	(1)	(2)	(3)	(4)	(5)	(6)
(1) > Med. Dist. Top Cities	7.54***	2.16	9.21***	7.05***	13.41***	21.13***
(F: _; 41.3; 10.5; 2.1)	[1.30]	[2.65]	[2.20]	[2.66]	[3.74]	[5.76]
(2) Crop Suitability < 25%	-0.92	6.95***	11.75***	4.80	11.28*	20.54***
(F: _; 17.0; 8.6; 3.9. Sh: 0.16)	[1.34]	[2.12]	[4.32]	[4.42]	[6.03]	[7.87]
(3) Leader's Origin 150km	-2.20*	9.68***	1.68	-8.00**	-7.35	-9.88
(F:_; 12.6; 5.7; 5.2, Sh: 0.24)	[1.26]	[2.20]	[3.94]	[4.00]	[5.21]	[6.29]

- Larger effects for smaller and more remote places (decentralization).
- Bigger effects for areas with worse agricultural land (trade specialization?)
- Smaller in ethnic homeland areas of head of state ("roads to nowhere")

Results

Heterogeneous Effects

- Classify the cities into two groups depending on:
  - High vs. low initial market access.
  - High vs. low land suitability for crops.
  - Ethnic homeland areas of head of state vs. rest (newly collected data on place of origin and ethnicity of 189 heads of state 1960-2010).
- See if the overall effect of a same change in road market access varies across the two groups.
- Possibly important for policy.

Conclusion

- Study the effects of road construction and market access on city population growth in Sub-Saharan Africa in 1960-2010.
- New panel data set on road surface and city population for 39 African countries every ten years in 1960-2010.
- Average effect of a 100% change in market access ~8-13%.
- Heterogeneity in the effects. Need to understand local context when evaluating the impact of transport investment.

Thank you for listening!!