How do Roads Spread AIDS in Africa?

Elodie Djemaï

Université Paris-Dauphine

Paris Seminar in Economic Demography
February 26, 2013

Introduction

- Large economic literature that emphasizes the role of transportation infrastructure on
  - poverty alleviation (Gibson and Rozelle 2003, Dercon *et al* 2008, Khandker *et al* 2009, Gachassin *et al* 2010, Dillon *et al* 2011)
  - access to markets (Jacoby 2000) and trade (Buys *et al* 2010)
  - economic performance (Straub 2008, Banerjee *et al* 2009, Donaldson 2010)
  - HIV prevalence (Tanser *et al* 2000)

- Mobility is known to be a risk factor of HIV-infection, among
  - truck drivers (e.g. Oruboloye *et al* 1993 ; Huygens 2001 ; Gouws and Ramjee 2002)
  - migrant workers (Meekers 2000 ; Adaji Nwokoji and Ajuwon 2004)
  - the general population (Oster 2011)

Introduction

- Trade along roads might appear as a transmitter of the epidemic from region to region
- Roads reduce the distance to markets and between people
- Roads might have two competing effects :
  - Lower the cost of protection
  - Increase the set of sexual partners
- Research Question : What is the net effect of road on the risk of infection ?

Introduction

Data :

- Study the general population at the individual level
- Combine survey data with geographical data on road infrastructure - for six African countries
- Exploit variations in the individuals' location, i.e. the distance to the road

Identification of the effect :

- Endogenous road placement
- Endogenous individual placement

Introduction

- Living far away from a paved road protects people from getting HIV
  - ↪ this holds when controlling for non-random placement of roads and individuals
  - ↪ this is robust to a number of checks
- Better access to condoms and greater level of HIV/AIDS-knowledge in proximity to road
- In proximity to a road, more casual sex even with a condom

Outline

Outline

Data description

Demographic and Health Surveys

DHS are standardized nationally representative household surveys in developing countries

We are using data from : Cameroon (2004), Ethiopia (2005), Ghana (2003), Kenya (2003), Malawi (2004) and Zimbabwe (2005/06)

- Homogeneous set of questions
- Blood sample collection to test for HIV
- GIS data on the sampled clusters

Data description

Demographic and Health Surveys : descriptive statistics (analytical sample)

	All	**CMR**	**ETH**	**GHA**	**KEN**	**MWI**	**ZWE**
HIV+	.078	.053	.019	.019	.065	.124	.179
HIV testing	.147	.177	.075	.091	.153	.152	.218
women	.542	.503	.539	.551	.527	.543	.573
age	28.89	29.08	29.16	30.11	28.71	28.83	27.72
urban	.326	.489	.249	.386	.295	.136	.317
no educ	.223	.158	.523	.287	.126	.176	.032
prim educ	.360	.403	.276	.188	.546	.634	.328
sec educ	.382	.405	.174	.490	.248	.181	.607
higher educ	.036	.034	.027	.035	.080	.009	.034
catholic	.172	.388	.011	.159	.245	.222	.105
protestant	.480	.356	.166	.538	.608	.623	.682
muslim	.157	.174	.320	.193	.108	.141	.007
Obs.	53,481	9,459	10,835	9,393	5969	5,121	12,704
Clusters	2,723	466	534	412	400	513	398

Data description

Geographical data on road infrastructure

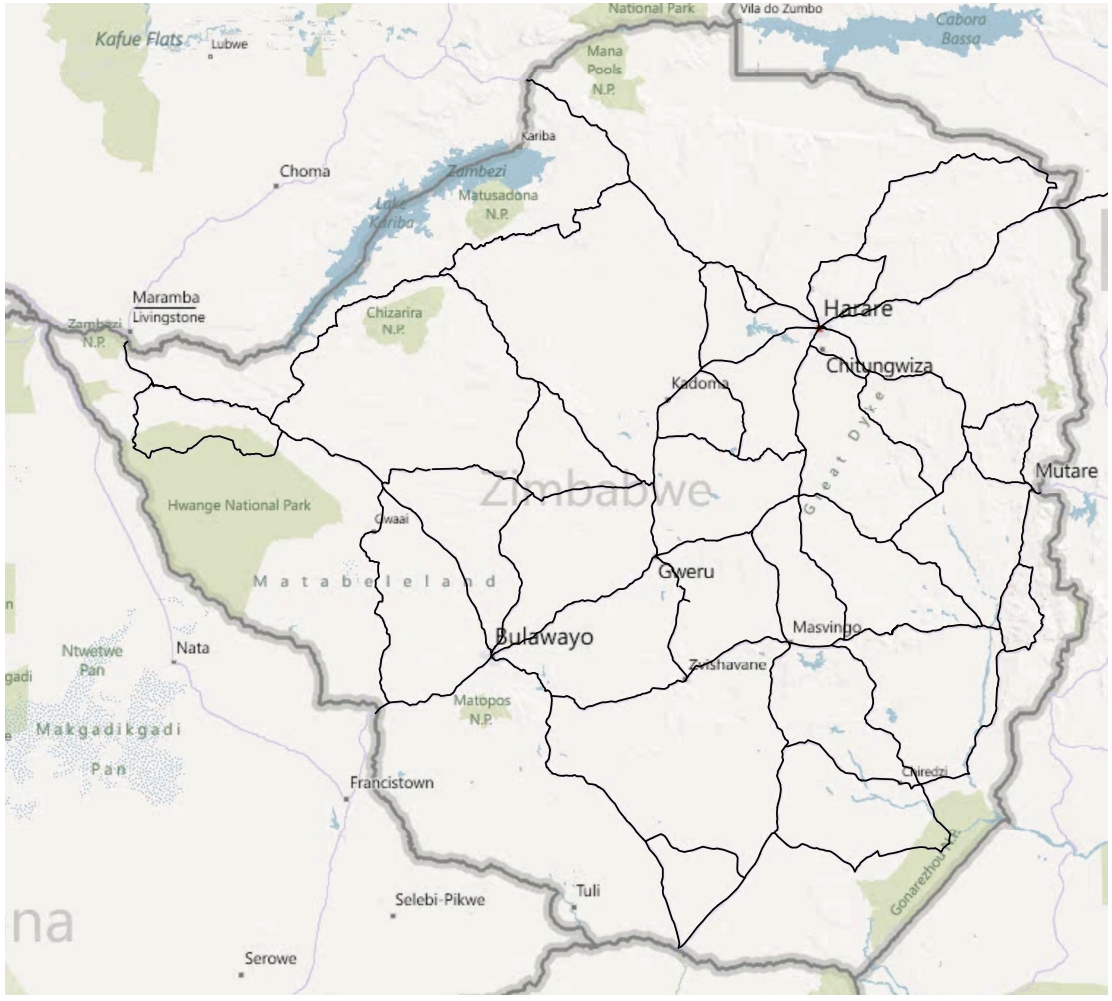
- Satellite image from *Bing map* as of July 2011
- Use ArcGIS to locate the sampled clusters and the road network on a country map and compute our variable of interest, i.e the *straight-line* distance
- Restrict to the network of paved or ”primary” roads

Data description

Zimbabwe : Road infrastructure, Bingmap (as of July 2011)

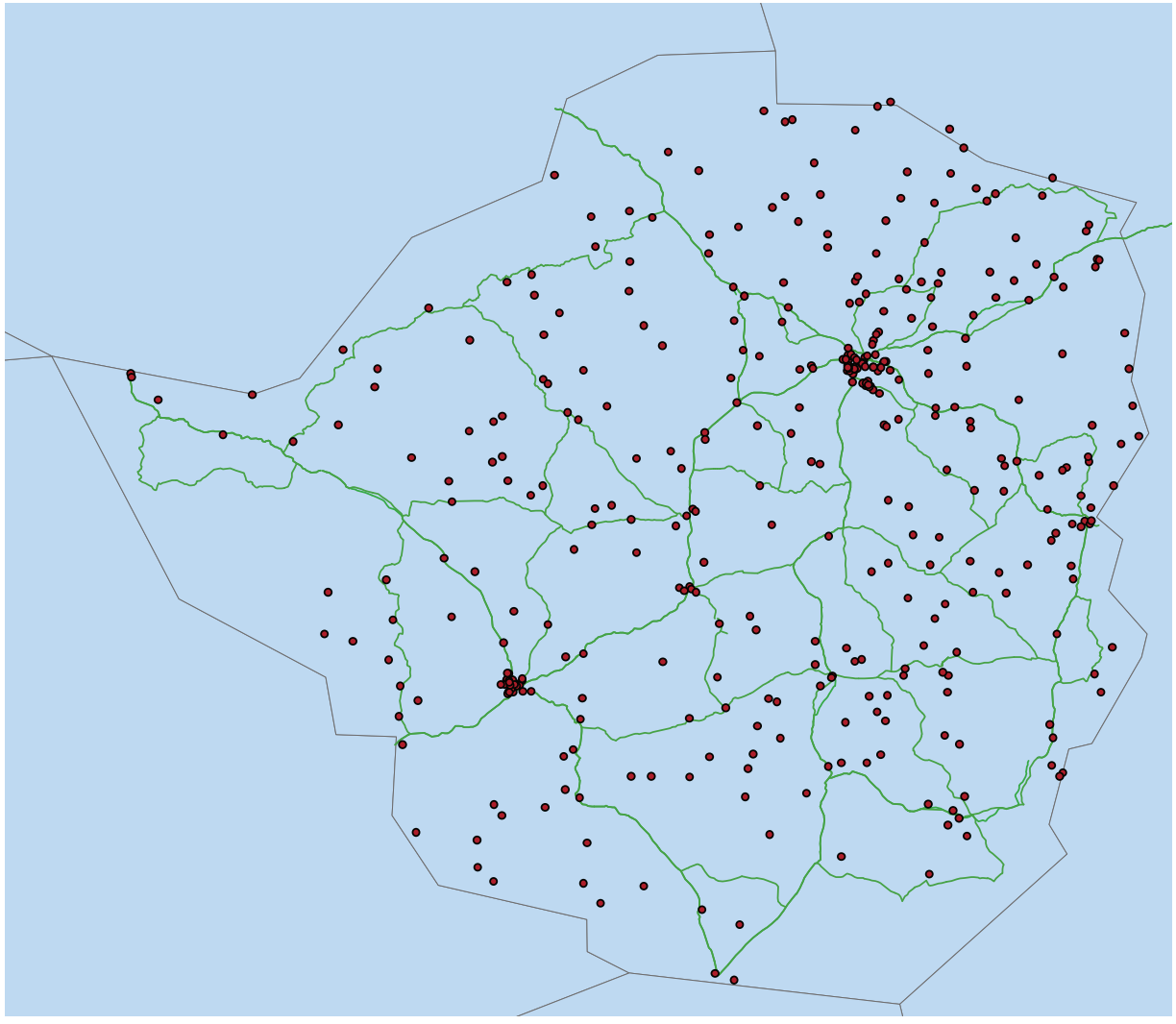
Data description

Zimbabwe : Road network (as of July 2011)



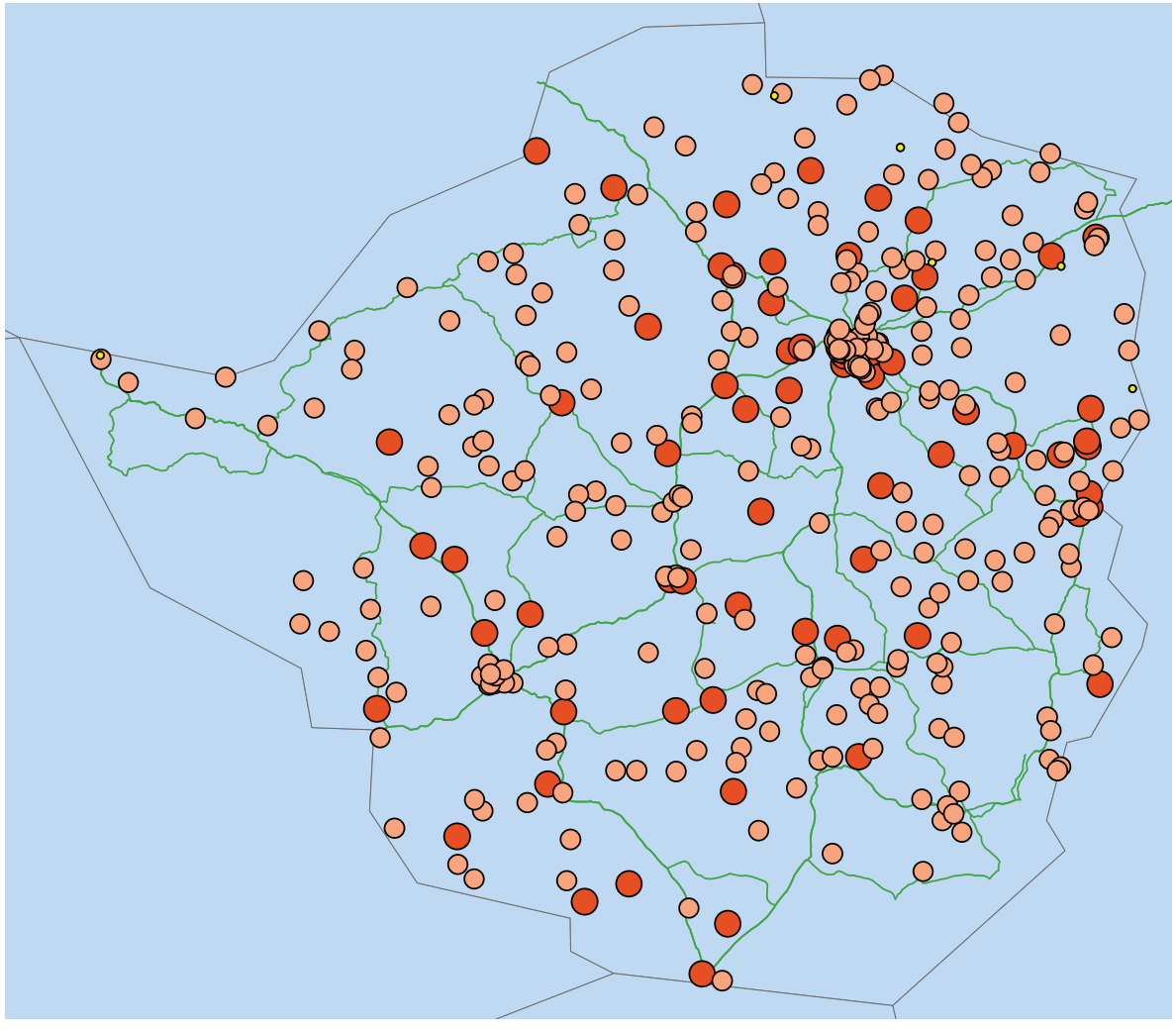
Data description

Zimbabwe : Road network and DHS sampled clusters



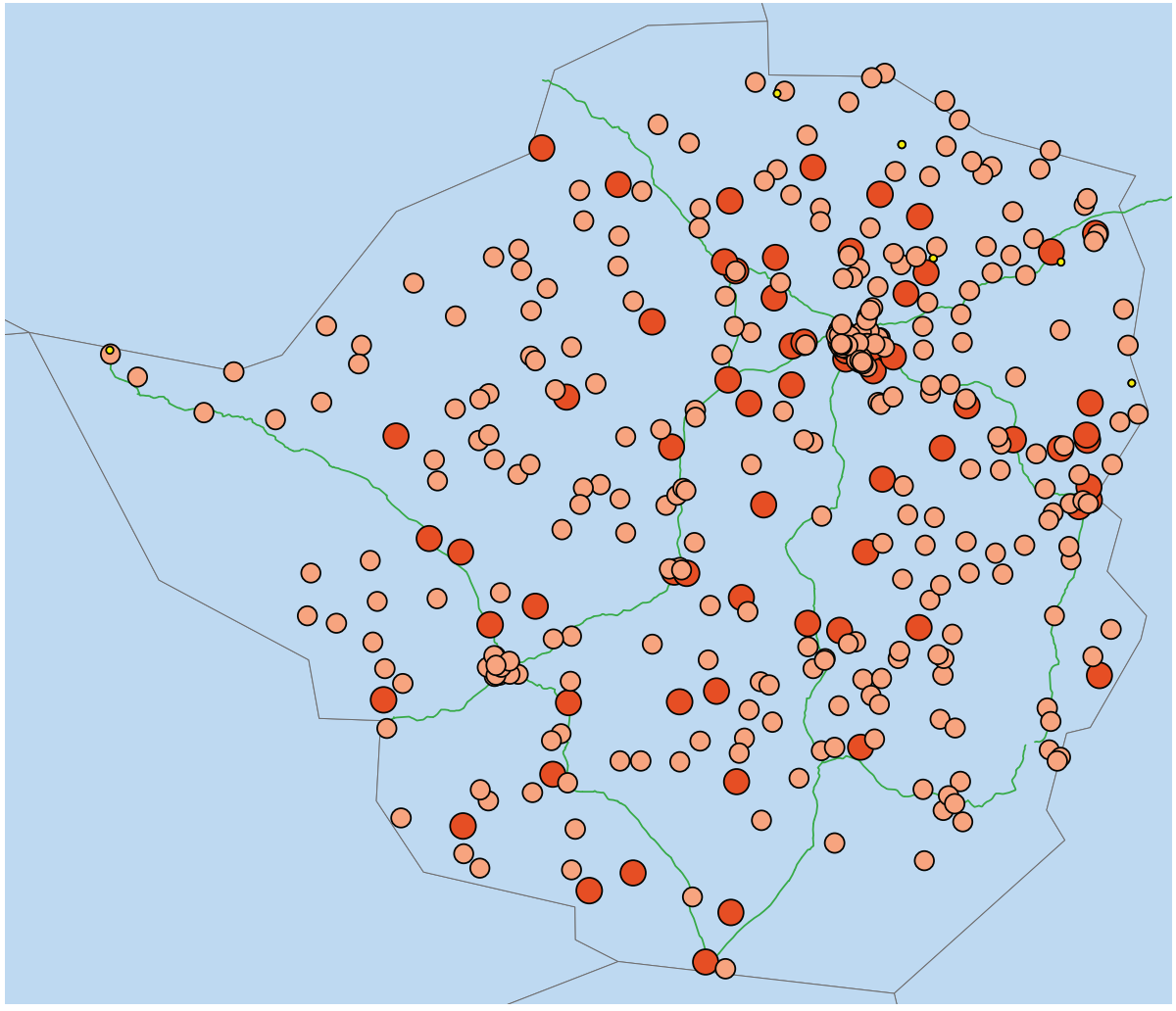
Data description

Zimbabwe : Road network and HIV prevalence rate in DHS sampled clusters



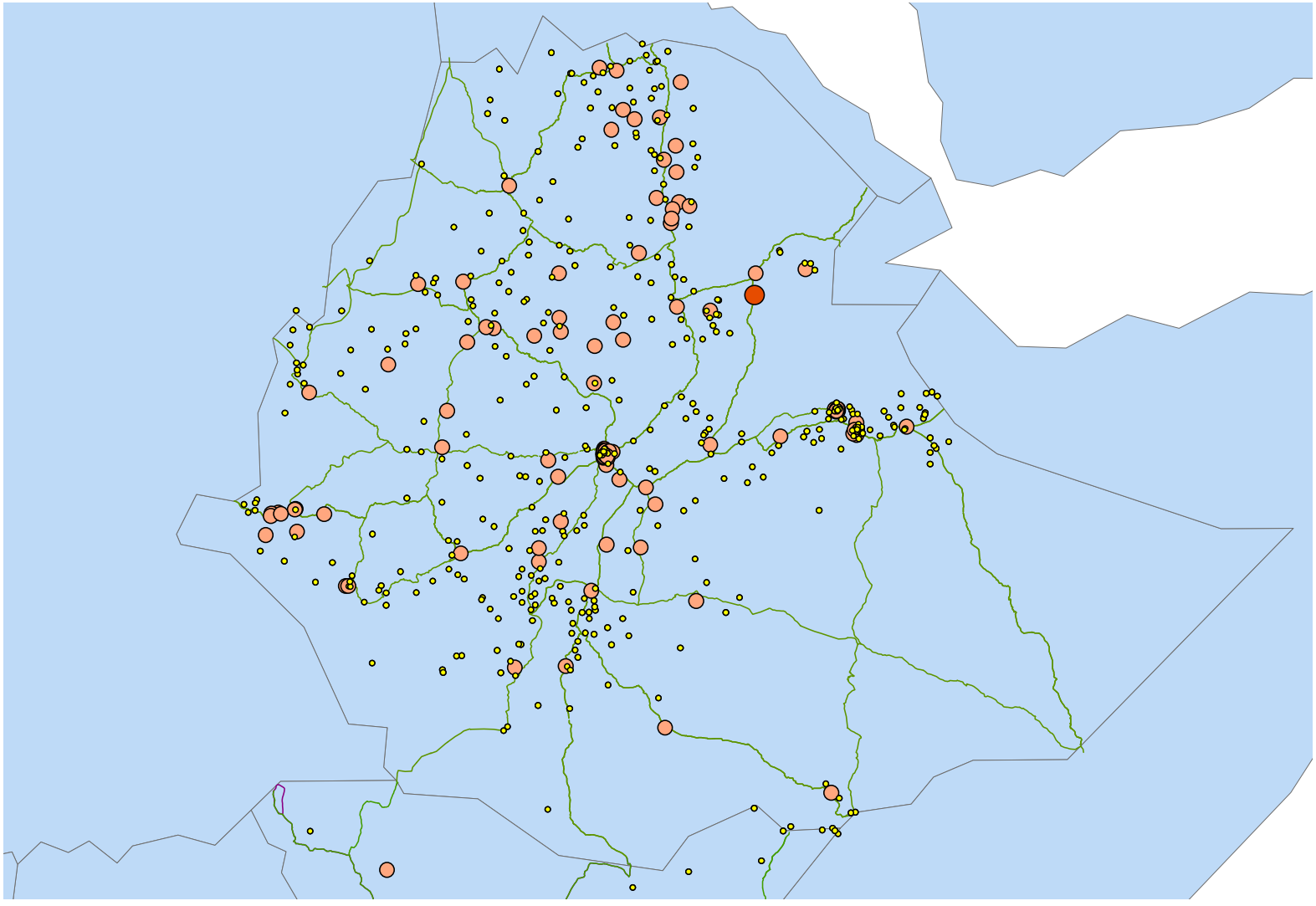
Data description

Zimbabwe : Road network (DCW) and HIV prevalence rate in DHS sampled clusters



Data description

Ethiopia : Road network and HIV prevalence rate in DHS sampled clusters



Data description

Geographical data on road infrastructure

- Respondents live on average 12.9 km away from the nearest paved road
- Median = 5.5 km
- Maximum value ranges from 63.3 km (Malawi) to 119.2 km (Kenya)

Outline

Distance to road and HIV-infection

Estimation strategy

If distance was randomly assigned across communities and people, we could estimate its effect through

$$Pr(HIV_{ijr} = 1) = \phi(\alpha + \beta log(1 + distroad_{jr}) + X_{ijr}^{I\prime}\delta_1 + X_{jr}^{J\prime}\delta_2 + \gamma_r + \varepsilon_{ijr})$$

- individual controls : gender, marital status, age, education, religion, HIV/AIDS-knowledge, wealth
- region-specific effects

Two concerns

- road placement might have been driven by characteristics that are also driving the spread of the epidemic
- individuals may sort non-randomly across accessible and remote areas, some might have moved to live close to a road

Distance to road and HIV-infection

Primary results

Table II : Road and HIV-risk- Probit coefficients

	(1)	(2)	(3)	(4)
Road distance	-0.0808***	-0.0858***	-0.0835***	-0.0623***
	(0.011)	(0.011)	(0.011)	(0.012)
Regional FE	yes	yes	yes	yes
X_{ijr}^{I}	no	yes	$+knowledge$	$+wealth$
N	53039	52993	50636	50636
Number of clusters	2703	2703	2703	2703

Note : Robust standard errors clustered at the community level in parentheses

Controls include gender, age, marital status, educational attainment, wealth, religion, HIV/AIDS-knowledge, regional dummies

Distance to road and HIV-infection

Primary results

Table II : Road and HIV-risk- Probit Estimates

	(1)	(2)	(4)	(3)
logdistroadkm	-0.0808***	-0.0858***	-0.0623***	-0.0835***
woman		0.2023***	0.2014***	0.2090***
married		0.4901***	0.5031***	0.4890***
prevmarried		1.0220***	1.0434***	1.0261***
age		0.0109***	0.0105***	0.0110***
primaryeduc		0.2234***	0.1710***	0.1968***
secondaryeduc		0.2965***	0.2026***	0.2657***
highereduc		0.0999	-0.0253	0.0719
catholic		0.0195	0.0313	0.0222
protestant		0.0014	0.0096	-0.0018
otherreligion		0.1480***	0.1740***	0.1565***
wpoorer			0.0460	
wmiddle			0.1396***	
wricher			0.2447***	
wrichest			0.2562***	
scoreclosed601			0.0061	0.0124

Note : Robust standard errors clustered at the community level in parentheses

Distance to road and HIV-infection

Primary results

Size of the effect : At sample means,

- a one-standard deviation increase in the distance to a road (2.24 km) reduces the risk of infection by between 0.5 and 0.9 percentage point

- the predicted probability is 4.25%

Distance to road and HIV-infection

Identification : Endogenous road placement

We control for potential confounders (community-level controls) :

- urban
- population density
- distance to the nearest city
- % of very rich people
- slope
- ruggedness (Nunn and Puga, 2011)
- latitude and longitude

Assumption : infrastructure placement is conditionally exogenous as in Koolwal and Van de Walle (2013) ; Nauges and Strand (2013),

Distance to road and HIV-infection

Identification : Endogenous road placement

Table III : Road and HIV-risk-Probit coefficients
Controlling for community-level characteristics

	(1)	(2)	(3)
Road distance	-0.0623***	-0.0520***	-0.0453***
urban		0.1405***	0.0719
percwrichest			0.2304***
popdensity			-0.0000
dist to city, km			-0.0102
longitude			0.0391*
latitude			-0.0469**
slope			-0.0000**
ruggedness			0.0000
Regional FE	yes	yes	yes
Individual covariates	yes	yes	yes
Community covariates	no	yes, urban	yes, all
N	50636	50636	50636
Number of clusters	2,703	2,703	2,703

Note : Robust standard errors clustered at the community level in parentheses
Controls include gender, age, marital status, educational attainment, wealth, religion, HIV/AIDS-knowledge, regional dummies

Distance to road and HIV-infection

Identification : Endogenous road placement

Table III : Road and HIV-risk-Probit coefficients
Controlling for community-level characteristics

	(1)	(2)	(3)
woman	0.2014***	0.1995***	0.1971***
married	0.5031***	0.5037***	0.5080***
prevmarried	1.0434***	1.0412***	1.0438***
age	0.0105***	0.0106***	0.0105***
primaryeduc	0.1710***	0.1693***	0.1660***
secondaryeduc	0.2026***	0.1974***	0.1920***
highereduc	-0.0253	-0.0264	-0.0422
wpoorer	0.0460	0.0441	0.0438
wmiddle	0.1396***	0.1300***	0.1300***
wricher	0.2447***	0.1996***	0.1823***
wrichest	0.2562***	0.1791***	0.0969*
catholic	0.0313	0.0356	0.0288
protestant	0.0096	0.0137	0.0025
otherreligion	0.1740***	0.1766***	0.1646***
scoreclosed601	0.0061	0.0055	0.0044
Regional FE	yes	yes	yes
Individual covariates	yes	yes	yes
Community covariates	no	yes, urban	yes, all
N	50636	50636	50636
Number of clusters	2,703	2,703	2,703

Note : Robust standard errors clustered at the community level in parentheses

Distance to road and HIV-infection

Identification : Endogenous individual placement

- Common source of bias as observed and unobserved factors at the individual level can affect both access to road infrastructure and HIV-infection and related behaviors
- Cluster-level analysis as in Dinkelman (2011), Koolwal and Van de Walle (2013) and Nauges and Strand (2013)
- $\hat{\beta}_{OLS,ind} = -0.0057^{***}$ and $\hat{\beta}_{OLS,cluster} = -0.0058^{***}$

Distance to road and HIV-infection

Identification : Endogenous individual placement

Two specific sources of estimation biases linked to migration as people might have moved to live close to a road

- Reverse causality : Infected people move to live close to a road to have access to ARV and/or to avoid stigma on their family
- Selection : More at risk individuals may have a higher likelihood to migrate and to move to live close to a road

Distance to road and HIV-infection

Identification : Endogenous road placement - reverse causality

Reverse causality : To rule out the possibility that HIV infection is a driver in the migration decision, that is driving our results

1. We look at whether the individual has ever been tested for HIV
   - only 15% of the sample has ever been tested (12.5% has done so and received the result)
   - do they know their current status ?
2. We remove the new movers, i.e. those who arrived less than 10 years ago
   - About 10 years is the median period between HIV infection and death in absence of treatment (Thirumurthy et al, 2005)

Distance to road and HIV-infection

Identification : Endogenous road placement - reverse causality

Table IV : Non-random Individual Placement
Probit Estimates

	Ever been tested (1)	Never been tested (2)	remove new movers (3)
road distance	-0.0273	-0.0482***	-0.0356**
	(0.024)	(0.013)	(0.015)
N	7470	42968	33877

Note : Robust standard errors clustered at the community level in parentheses

Distance to road and HIV-infection

Identification : Endogenous road placement- selection

Selection

1. We estimate the road distance effect for the subsample of people who have never migrated and for those who have ever migrated

    - Underlying assumption : the reasons why the respondent's parents were living in her birth place have no direct effect on her own risk of infection

2. We remove the "potential selection drivers" : who are defined as those who have migrated after reaching 15 years old and before getting married

    - more likely to have initiated the decision to relocate
    - and to benefit from extended set of sexual partners

Distance to road and HIV-infection

Identification : Endogenous road placement- selection

Table IV : Non-random Individual Placement
Probit Estimates

	Born here	Migrant	remove selection drivers
	(4)	(5)	(6)
road distance	-0.0454***	-0.0429***	-0.0418***
	(0.017)	(0.015)	(0.014)
N	23929	26563	43208

Note : Robust standard errors clustered at the community level in parentheses

Distance to road and HIV-infection

Sensitivity analysis

- No gender effect
- rural sample more protected by road distance
- Mobility scenario : Close to road people are in touch with more people

↪ We found that people who own a bike, a motorcycle or a car are not protected against HIV by living in remote areas

Threats to validity

Robustness Checks

1. Refusal to be tested for HIV
   - 15% of eligible individuals were not tested (refusal, absence, error)
   - Does distance to the road affect the likelihood of showing up ?
     - *Yes*, Probability of showing up increases with the distance

2. Random reallocation of communities : up to 2 km of random error is added to cluster locations in urban areas and up to 5 in rural areas
   - Results are robust when removing the potential "overlapping" communities

Outline

Access to and demand for self-protection

Goal : Study whether the increase in the individual risk of infection is due to deficiencies in the supply or in the demand for self-protection

Approach : Estimate the effect of road distance on

- Knowledge of HIV-transmission
- Access to condoms
- Choice of condom use and sexual partner

Access to and demand for self-protection

Knowledge is improved through increased access to media

Table V : Road Distance and Knowledge (score of 6 items)
OLS coefficients (analytical sample)

	(1)	(2)	(3)	(4)
road distance	-0.1252***	-0.0163*	-0.0154*	-0.0121
knows someone HIV+			0.1759***	0.1584***
ever been tested			0.0407***	0.0347***
magazines				
less than once a week				0.2365***
at least once a week				0.1556***
almost every day				0.1068***
radio				
less than once a week				0.1758***
at least once a week				0.2381***
almost every day				0.2866***
tv				
less than once a week				0.0996***
at least once a week				0.0253
almost every day				0.0909***
Regional FE	yes	yes	yes	yes
Individual covariates	no	yes	yes	yes
Community covariates	no	yes	yes	yes +*media*
mean y	4.60	4.61	4.61	4.61

Access to and demand for self-protection

Access to condoms

Proximity to road increases the likelihood of

- knowing at least one place where one could find a condom
- citing a place from the non health private sector
- declaring being able to buy a condom

No effect on the likelihood of citing a place from the public health sector or the private health sector

Access to and demand for self-protection

Demand : more mitigated results

Table VI : Last sexual intercourse with spouse and condom
Probit Model

	Condom use		Sex with spouse	
	All	Rural	All	Rural
	(1)	(2)	(3)	(4)
Panel A : analytical sample				
Road distance	-0.0019	-0.0034*	0.0012	0.0024*
	(0.002)	(0.002)	(0.001)	(0.001)
Panel B : all surveyed respondents				
Road distance	-0.0027**	-0.0036**	0.0012	0.0023**
	(0.001)	(0.001)	(0.001)	(0.001)

note : Analytical sample (similar qualitative results on cluster-level observations)

Access to and demand for self-protection

Discussion and implications

- Access to protection seems not sufficient to prevent people from being infected
- Proximity to road increases the likelihood of engaging in casual sex
- As condoms become available, people use them but increase or maintain their willingness to have casual partners
  - ↪ Related to the literature on risk compensation (road safety)

Conclusion

Summary

This empirical analysis of the relationship between proximity to road and HIV-infection reaches the following conclusions :

- Living close to a paved road increases the risk of HIV-infection
- despite the increased access to condom and knowledge
- as the likelihood of having casual sex increases there and offsets the increase in condom use (at least for rural people)

Conclusion

Policy recommendations

- Persistent spatial disparities in access to information and protective devices (from any source)
  - ↪ in favor of drawing specific programs for accessible and remote areas
- Increased general knowledge and condom availability are somehow necessary but not sufficient to prevent from being infected
  - ↪ Need to provide people more incentives to self-protect
- Road are found to have additional costs and benefits that were not explored beforehand