Stockholm Doctoral Course Program in Economics

Topics in Applied Microeconometrics:
Using GIS

Lecture 4

Distance as Instruments

Masayuki Kudamatsu

IIES, Stockholm University

19 September, 2013

Outline

1. Distance as instruments in action
2. Issues for distance as instruments
3. Nunn (2008)
4. Nunn & Wantchekon (2011)
5. Replicate Nunn (2008)'s data with ArcGIS

1. Distance as instruments in action

A. Oster (2012)

- HIV prevalence $\uparrow \Rightarrow$ Risky sexual behavior $\downarrow$?
- Distance to DR Congo as instrument for HIV prevalence
- $\hat{\beta}_{IV} < 0$ while $\hat{\beta}_{OLS} > 0$
- Stronger effects for areas w/ higher life expectancy

Distance as instruments in action (cont.)

B. McKenzie et al. (2010)

- Migration to a richer country (Tonga to New Zealand) $\Rightarrow$ Income $\uparrow$?
- Distance to NZ Dept. of Labor office as instrument for migration from Tonga to NZ
- $\hat{\beta}_{OLS} > \hat{\beta}_{IV} > 0$

Distance as instruments in action (cont.)

C. Woodruff & Zenteno (2007)

- Migration rate in birth place $\uparrow \Rightarrow$
  Capital investment $\uparrow$
  for entrepreneurs in Mexico?
- Distance to early 20c railway stations from the capital of the birth state as instrument for migration rate
- $\hat{\beta}_{IV} > 0$

2. Issues on distance as instruments

A. Exclusion restriction

- Selection in space (those more affluent can live close to cities etc.)
- Distance to cities / roads: pick up access to markets, schools, clinics, etc.

2. Issues on distance as instruments (cont.)

B. LATE

- Treatment is induced not only by distance but by other factors

$\Rightarrow$ Is LATE of interest?

cf. McKenzie et al. (2010): Distance to NZ Dept. of Labor office the sole determinant of migration from Tonga to NZ

3. Nunn (2008)

Perhaps one of the best papers among those using distance as IV

3.1 Research Question

Do slave trades explain African underdevelopment?

- Interesting?
- Original?
- Feasible?

3.2 Measuring slave trades

1. # of slaves exported from each port/region of Africa

$\Rightarrow$ # of slaves exported from each coastal country

2. Ethnicity of a sample of 100,000+ slaves shipped from Africa
3. Murdock (1959)'s African ethnic group boundary map

$\Rightarrow$ Distribution of exported slaves from all countries

Taken from Figure II of Nunn (2008)

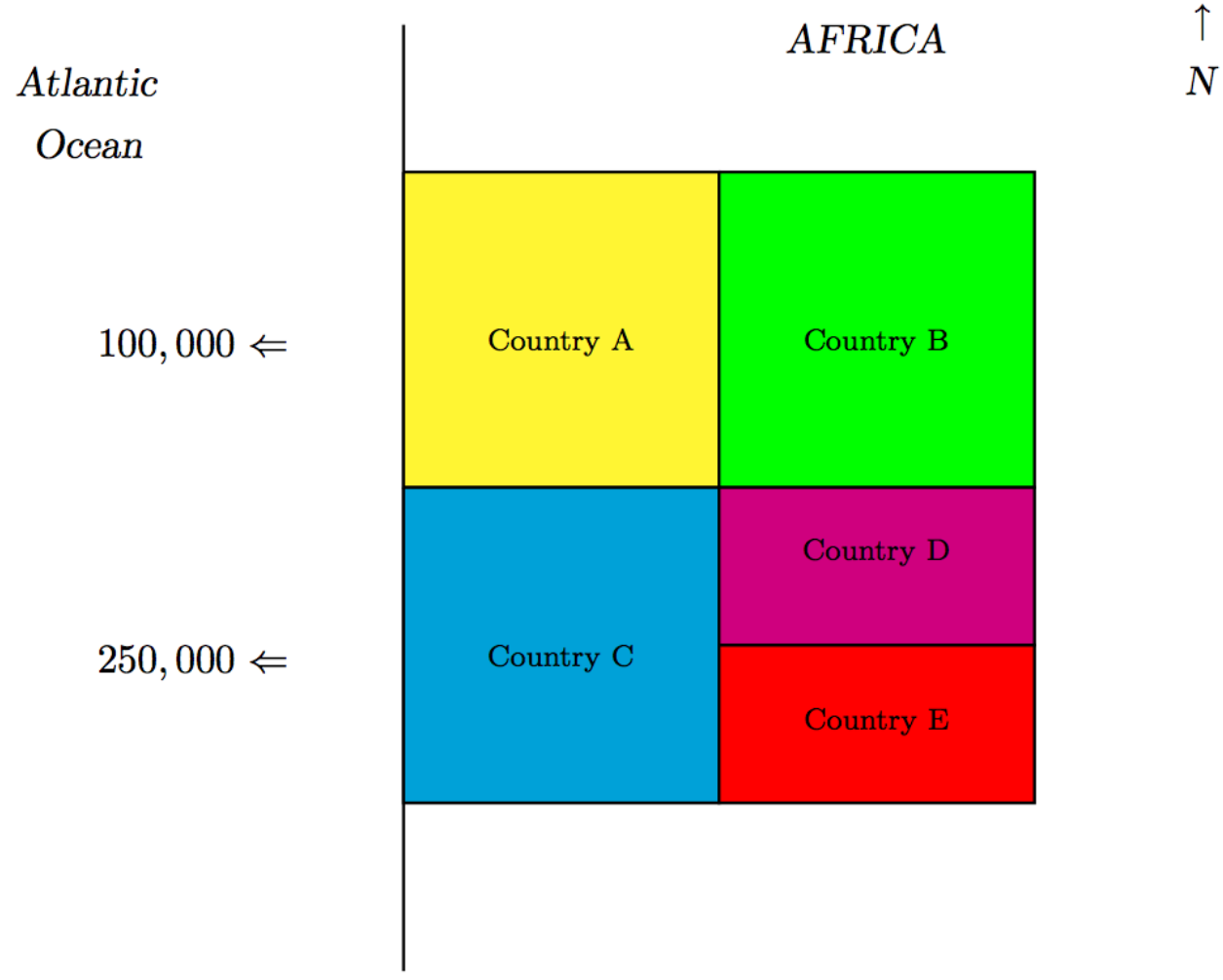


Figure 1: Artificial Map of the West Coast of Africa.

Taken from Nunn (2008)

Assigning pre-colonial ethnic groups into different countries can be done by ArcGIS's Intersect tool (Exercise 1 below)

3.3 Empirical method

$$\ln y_i = \beta_0 + \beta_1 \ln(exports_i / area_i) + \mathbf{C}_i' \delta + \mathbf{X}_i' \gamma + \varepsilon_i$$

- y_i: GDP per capita in country i in 2000 (Maddison 2003)

3.3 Empirical method

$$\begin{aligned} \ln y_i \;=\; & \beta_0 + \beta_1 \ln \left(exports_i / area_i \right) \\ & + \mathbf{C}_i' \delta + \mathbf{X}_i' \gamma + \varepsilon_i \end{aligned}$$

- $exports_i$: # of slaves exported 1400-1900 from country i
- $area_i$: Land area of country i

3.3 Empirical method

$$\begin{aligned} \ln y_i &= \beta_0 + \beta_1 \ln (exports_i / area_i) \\ &\quad + \mathbf{C}_i' \delta + \mathbf{X}_i' \gamma + \varepsilon_i \end{aligned}$$

- $\mathbf{C}_i$: Colonizer dummies
- $\mathbf{X}_i$: Geography & climate variables

- Some of the geography controls are created by ArcGIS (Feature to Point & Add XY Coordinates) (Exercise 2 below)
  - Distance from the equator (for tropical climate)
  - Longitude of country *i*'s centroid (for East-West differences)

3.3 Empirical method (cont.)

- Endogeneity concerns
  - Selection into slavery
- As an instrument, use distance to places with demand for slaves

- Specify the place with demand for slaves
  - Top 9 places in Atlantic slave import volume
  - Mauritius & Muscat for Indian Ocean trade
  - Algiers etc. for Saharan trade
  - Djibouti etc. for Red Sea trade
- Obtain the coordinate of
  - Slave import markets
  - Centroid of African countries
  - Point on the coast closest to each country's centroid

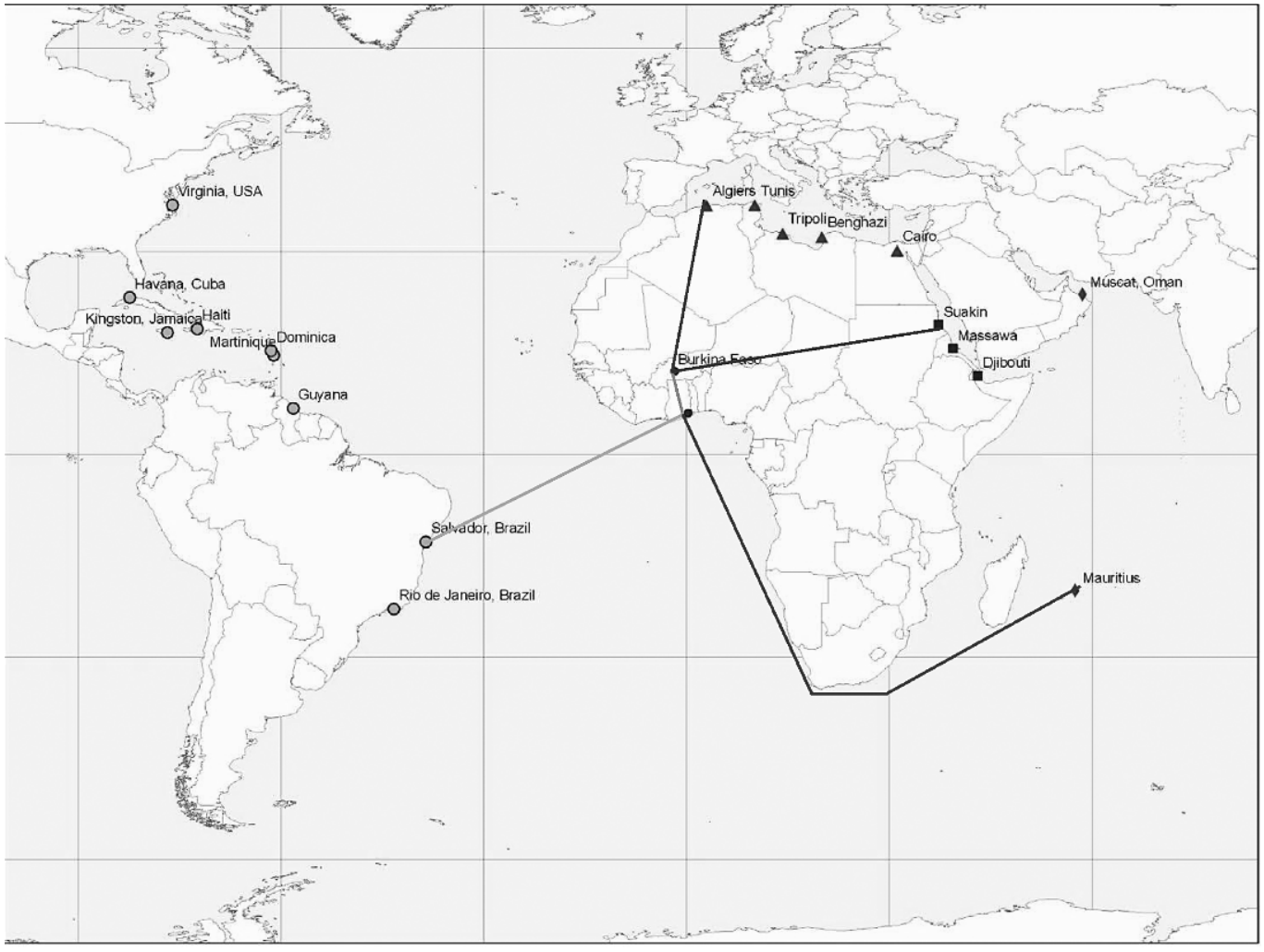


FIGURE V
Example Showing the Distance Instruments for Burkina Faso

Taken from Nunn (2008)

- Point on the coast closest to each country's centroid can be obtained by the Near tool of ArcGIS (Exercise 2 below)

$$
\begin{aligned}
d_{ij} \;=\; & \arccos\{\sin(La_i)\sin(La_j) \\
& + cos(La_i)cos(La_j)cos(Lo_i - Lo_j)\} \\
& \times 111.12
\end{aligned}
$$

- Use the great circle distance formula to obtain distance between i and j in kilometers
- Stata globdist ado does this

Remember...

2. Issues on distance as instruments

A. Exclusion restriction

- Selection in space (those more affluent can live close to cities etc.)
- Distance to cities / roads: pick up access to markets, schools, clinics, etc.

So in this case, what threatens IV validity?

Exclusion restriction?

- Location of slave supply does not affect location of slave demand
  - Location of demand: determined by climate (suitability for plantation crops) or by mines (p. 160)
- Distance to slave market $\neq$ distance to other economic opportunities
  - Reduced-form correlation is absent outside Africa (p. 163)

First stage results (what do you notice?)

First Stage. Dependent variable is slave exports, ln(exports/area)

Atlantic distance	−1.31***	−1.74***	−1.32*	−1.69**
	(0.357)	(0.425)	(0.761)	(0.680)
Indian distance	−1.10***	−1.43***	−1.08	−1.57*
	(0.380)	(0.531)	(0.697)	(0.801)
Saharan distance	−2.43***	−3.00***	−1.14	−4.08**
	(0.823)	(1.05)	(1.59)	(1.55)
Red Sea distance	−0.002	−0.152	−1.22	2.13
	(0.710)	(0.813)	(1.82)	(2.40)
F-stat	4.55	2.38	1.82	4.01
Colonizer fixed effects	No	Yes	Yes	Yes
Geography controls	No	No	Yes	Yes
Restricted sample	No	No	No	Yes
Hausman test (*p*-value)	.02	.01	.02	.04
Sargan test (*p*-value)	.18	.30	.65	.51

Taken from Table IV of Nunn (2008)

Weak instruments

- F-stat is very low
- With one endogenous regressor, one solution is to use Moreira's (2003) CLR confidence intervals (cf. Lecture 2)
- Table IV column (3), intervals are infinite as 1st stage F-stat < 2

Second Stage. Dependent variable is log income in 2000, $\ln y$				
ln(exports/area)	-0.208^{***}	-0.201^{***}	-0.286^{*}	-0.248^{***}
	(0.053)	(0.047)	(0.153)	(0.071)
	$[-0.51, -0.14]$	$[-0.42, -0.13]$	$[-\infty, +\infty]$	$[-0.62, -0.12]$
Colonizer fixed effects	No	Yes	Yes	Yes
Geography controls	No	No	Yes	Yes
Restricted sample	No	No	No	Yes
F-stat	15.4	4.32	1.73	2.17
Number of obs.	52	52	52	42

- As this is an over-identified case, one could also use LIML with Bekker (1994) standard error corrections (cf. Lecture 2)

OLS vs 2SLS

- $|\hat{\beta}_{2SLS}| > |\hat{\beta}_{OLS}|$
  - Hausman test rejects the equality at 5 percent
- Why?
  - Measurement error in $export_i$
  - Initially prosperous areas first export slaves
  - Another reason?

LATE?

- For countries that voluntarily engaged in slave exports, maybe a smaller effect on today's GDP per capita
  - Another possible reason for $|\hat{\beta}_{2SLS}| > |\hat{\beta}_{OLS}|$
- But LATE may be more of interest than ATE of slave exports

3.4 Mechanisms

- Discourage formation of broader ethnic identities
  - $EthnicFractionalization_i \propto export_i$ (Figure VI)
- Weaken the state
  - % pop from ethnic groups w/o pre-colonial state $\propto export_i$ (Figure VII)

3.4 Mechanism (cont.)

- Destroy interpersonal trust (Nunn & Wantchekon 2011)

4. Nunn & Wantchekon (2011)

4.1. Research question

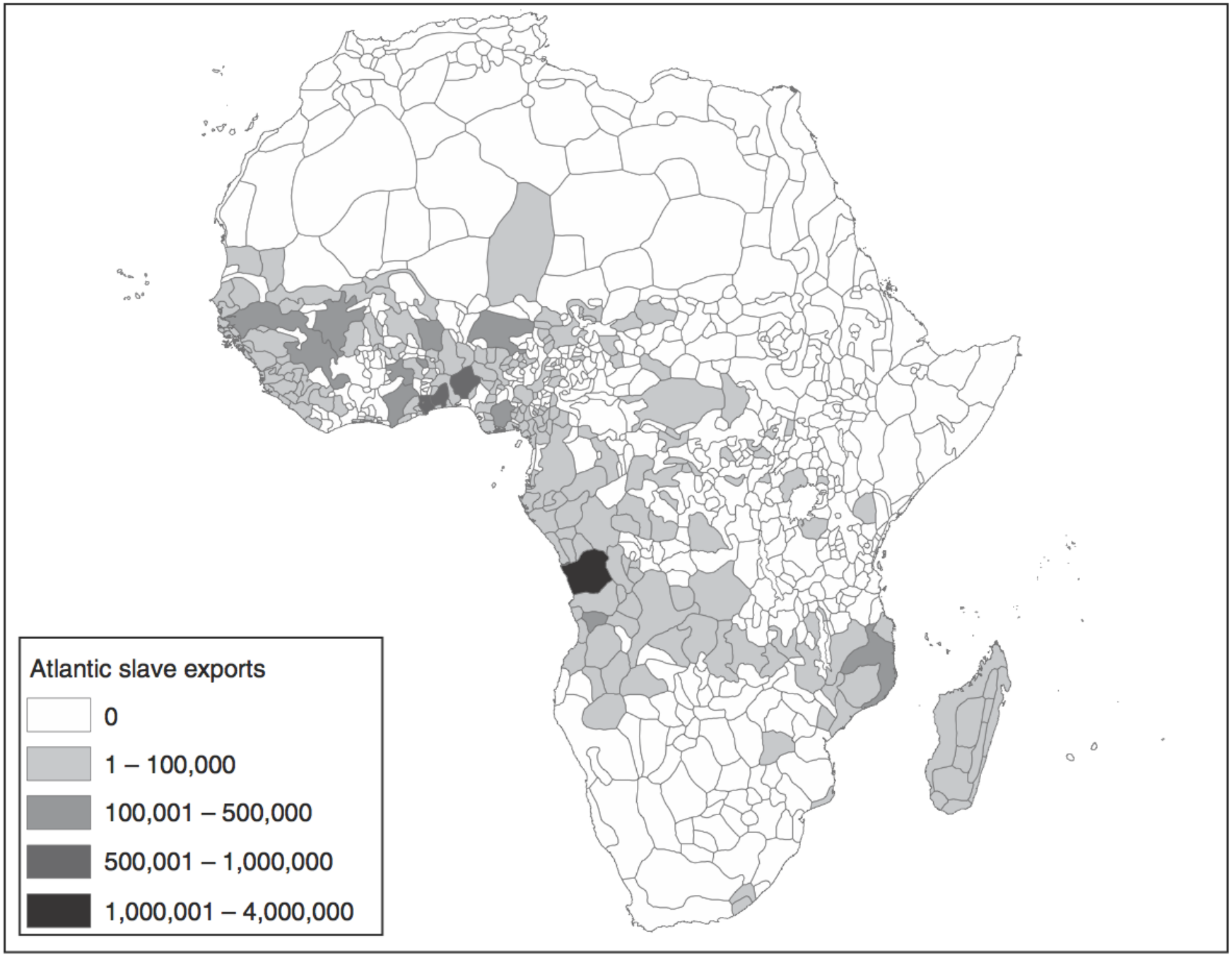
- Do slave trades have a long-run impact on the level of trust?
- Interesting?
  - Trust: correlated w/ economic performance (Knack & Keefer 1997; Algan & Cahuc 2010; Tabellini 2010)
- Original?
  - Yes indeed.
- Feasible?

4.2 Data

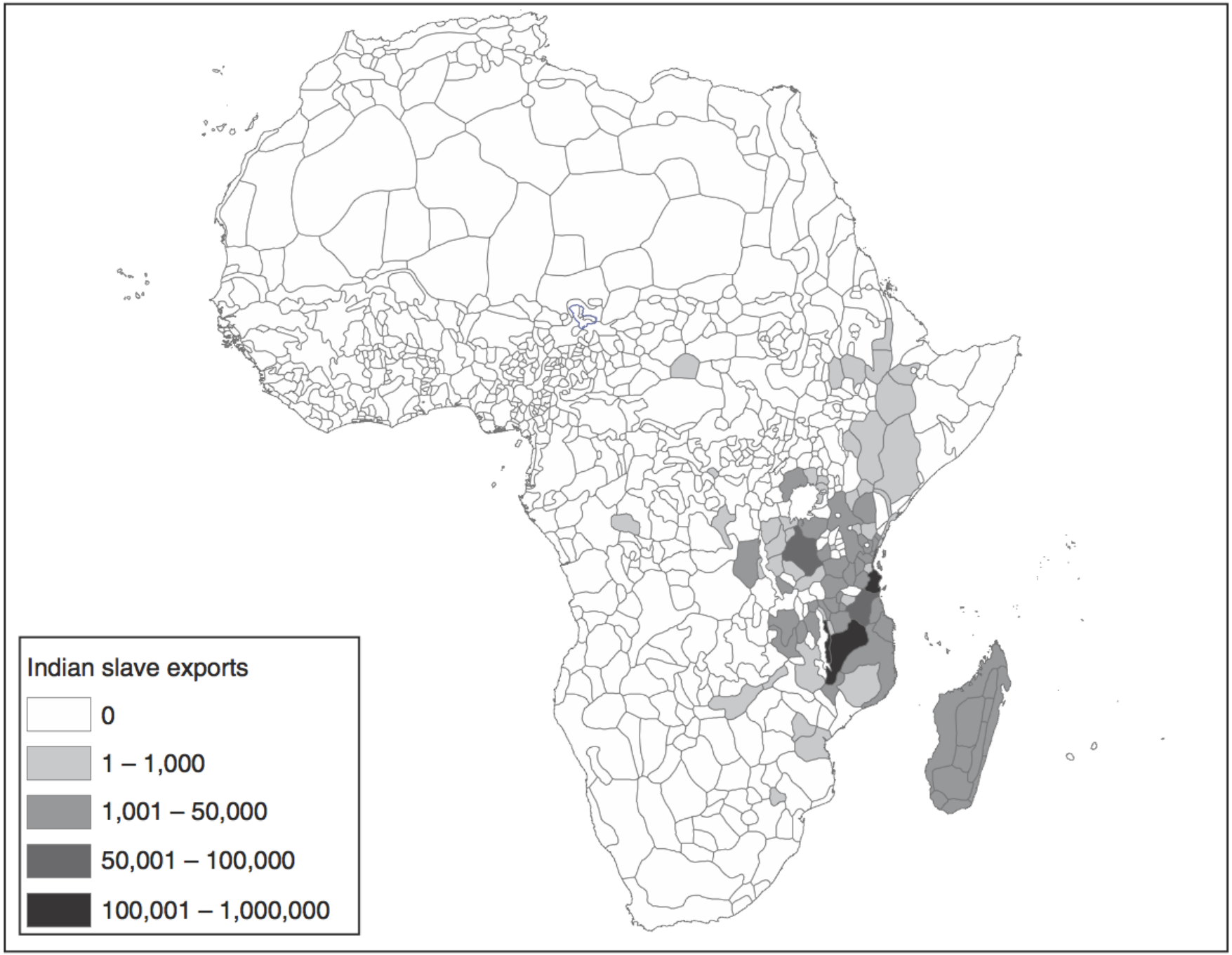
- Afrobarometor surveys in 17 African countries
  - How much they trust relatives, neighbors, local government, same/other ethnic group members
  - Ethnicity
- Ethnicity-level slave export data
- Murdock's precolonial ethnic boundary data

$\Rightarrow$ Figure 1A-1B

Panel A. Transatlantic slave trade



Panel B. Indian Ocean slave trade



4.3 Empirical specification

$$Trust_{iedc} = \alpha_c + \beta \ln(1 + export_e / area_e) + \mathbf{X}'_{iedc}\gamma + \mathbf{X}'_{dc}\omega + \mathbf{X}'_{e}\phi + \varepsilon_{iedc}$$

- $export_e$: instrumented by distance from ethnic group e's boundary centroid to the nearest point on the coast line

4.3 Empirical specification

$$Trust_{iedc} = \alpha_c + \beta \ln(1 + export_e/area_e) + \mathbf{X}'_{iedc}\gamma + \mathbf{X}'_{dc}\omega + \mathbf{X}'_{e}\phi + \varepsilon_{iedc}$$

- $\mathbf{X}'_e$: colonial variables, reliance on fishing, distance to the Saharan trade route (economic opportunities in early days)

4.3 Empirical specification

$$\begin{aligned} Trust_{iedc} &= \alpha_c + \beta \ln(1 + export_e/area_e) \\ &\quad + \mathbf{X}'_{iedc}\gamma + \mathbf{X}'_{dc}\omega + \mathbf{X}'_{e}\phi + \varepsilon_{iedc} \end{aligned}$$

- SE: two-way clustered by e & d

  (Cameron, Gelbach, and Miller 2011; Stata ado file available at Miller's website)

4.4 1st stage (Tab 4-5, col 2)

Dep. Var.	Slave exports	
Distance to coast	-0.0014***	-0.0015***
	[0.0003]	[0.0003]
F-stat on excl. IV	26.8	21.6
Fishing	NO	YES
Distance to Sahara	NO	YES
Other controls	YES	YES
# of ethnicity	147	147
# of districts	1187	1187
Observations	16,679	16,679

4.5 2nd stage (Tab 4-5, col 2)

Dep. Var.	Trust of neighbors	
Slave exports	-0.245***	-0.271***
	[0.070]	[0.088]
Hausman test p-value	0.88	0.42
Fishing	NO	YES
Distance to Sahara	NO	YES
Other controls	YES	YES
# of ethnicity	147	147
# of districts	1187	1187
Observations	16,679	16,679

- Download the data folder
  Computer > teacher (ink-edu)>
  economics > Lecture4
  to C:/TEMP now
- Launch ArcMap 10 now

Now unzip the following files:

- 10m-admin-0-countries.zip
- 10m-coastline.zip
- Murdock_shapefile.zip

The first two are obtained from Natural Earth (www.naturalearthdata.com).

5. Replicate Nunn (2008)'s data with ArcGIS 10

- Exercise 1: Obtain country-level slave exports
- Exercise 2: Measure each country's distance to slave trade centers

5.1 Country-level slave exports

- You have two datasets
  - (a) # of slaves shipped from each coastal country ($\tilde{n}_c$)
  - (b) Each ethnic group's share in the total # of slaves exported from Africa (s_e)
- Let N_c be the set of country c and the inland countries bordering with country c.
- We want to estimate # of slaves exported from country $d \in N_c$ (denote this by n_d)

- Let E_{N_c} be the set of ethnic groups in all the countries in N_c

$\Rightarrow$ # of slaves of ethnicity e exported from country c is given by

$$\tilde{n}_c \frac{s_e}{\sum_{f \in E_{N_c}} s_f}$$

- Let E_d be the set of ethnic groups in country d
- Then n_d is obtained by

$$n_d = \tilde{n}_c \frac{\sum_{g \in E_d} s_g}{\sum_{f \in E_{Nc}} s_f}$$

- If an ethnic territory spans across more than one country, its share of slaves will be assigned to each country in proportion to areas
  - (footnote 4 of Nunn 2008)

$\Rightarrow$ Need to calculate the area of each ethnicity-country intersection

- Each row in the output table is country-ethnicity intersection w/ 3 attributes:
  - Country ID
  - Ethnicity ID
  - Area

- Then in Stata...
  1. merge by ethnicity w/ slave share data (s_e)
  2. egen sum of areas by ethnicity
  3. generate fraction of area over [2] for each country-ethnicity intersection
  4. generate product of s_e & [3]
  5. egen sum of [4] by country (ie. $\sum_{g \in E_d} s_g$)

- Ethnicity boundary: a digitized map of Murdock (1959), available from Nathan Nunn's website
- Country boundary: we use Natural Earth 1:10m Cultural Vector "Admin 0 - Countries"
  - Downloaded at www.naturalearthdata.com
  - Better than GAUL (cf. Lec 1 Ex 1), which contains separate tiny polygons for islands belonging to each country

To calculate the area of country-ethnicity intersection polygons, we use the following geoprocessing tools:

- Intersect
- Project
- Add Field
- Calculate Field

Intersect

To create country-ethnicity intersection polygons

- Input Features: borders_tribes.shp & ne_10m_admin_0_countries.shp
- JoinAttributes: ALL or NO_FID
- Output Type: INPUT

Difference from Spatial Join

- Output features are intersections (if Spatial Join, target features)
- Useful if you need to obtain statistics at the intersection level (e.g. area)
- If you only need to know which ethnic groups live in which country, the Spatial Join (with “one to many” option) suffices.

Project

- Output Coordinate System:
  Projected Coordinate System >
  Continental > Africa >
  Africa Sinusoidal

Sinusoidal: equal area projection often used for the whole globe or a continent (Africa, Latin America)

Sinusoidal projection

- Longitude: $X \times (\lambda - \lambda_0) \cos(\phi)$
- Latitude: $X \times \phi$

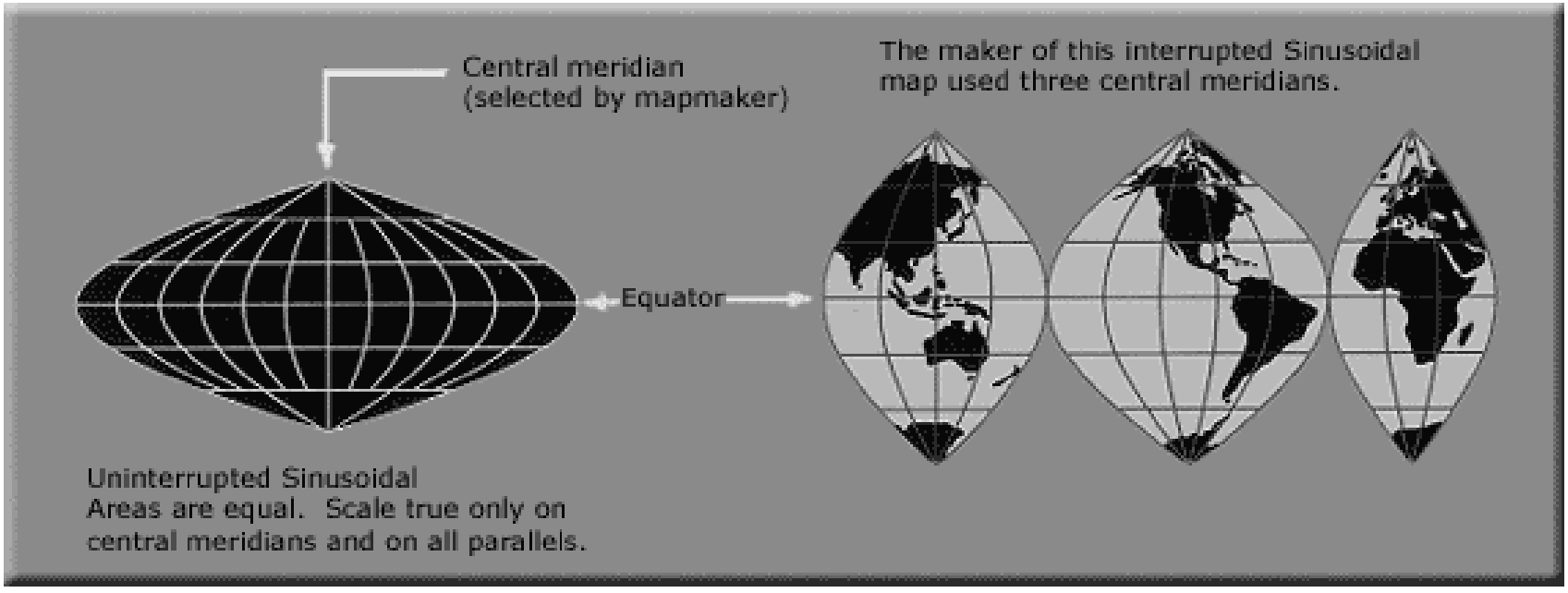
X : length in meters of 1 degree on the equator

λ : Longitude (in degrees)

ϕ : Latitude (in degrees)

λ_0 : Central meridian

  - In ArcGIS, 0° for world, 15° for Africa



Taken from

www.absoluteastronomy.com/topics/Map_projection

Add Field

Before using the Calculate Field tool, you always need to create a blank field by the Add Field tool.

- Field Name: area (or whatever)
- Field Type: FLOAT

Note: This tool overwrites the input file

Calculate Field for area calculation

- Field Name: area (or the one created by Add Field)
- Expression:

  float(!SHAPE.AREA!)
- Expression Type: PYTHON_9.3

Now export the script, copy & paste local variables and geoprocessing commands to the template script, edit it, and run.

Scripting tip for Intersection tool

- Intersect tool's first argument (input file names) should be a string
- If you use local variable names, the separator needs to be + ";" +
  - ; : separator in ArcGIS for file names
  - + : concatenate two strings (cf. L3, Sec 5)
  - A string should be quoted with double (or single) quotation marks in Python scripting
- Same applies to other tools with multiple inputs

Now run the script.

5.2 Distance to slave trade centers

- Need to obtain geographic coordinates of
  - Slave trade centers
    - $\Leftarrow$ Online gazetteer (e.g. www.fallingrain.com)
  - Country centroids
    - Used as controls as well
  - Closest point on the coast from country centroid
- Then use globdist (ado) in Stata

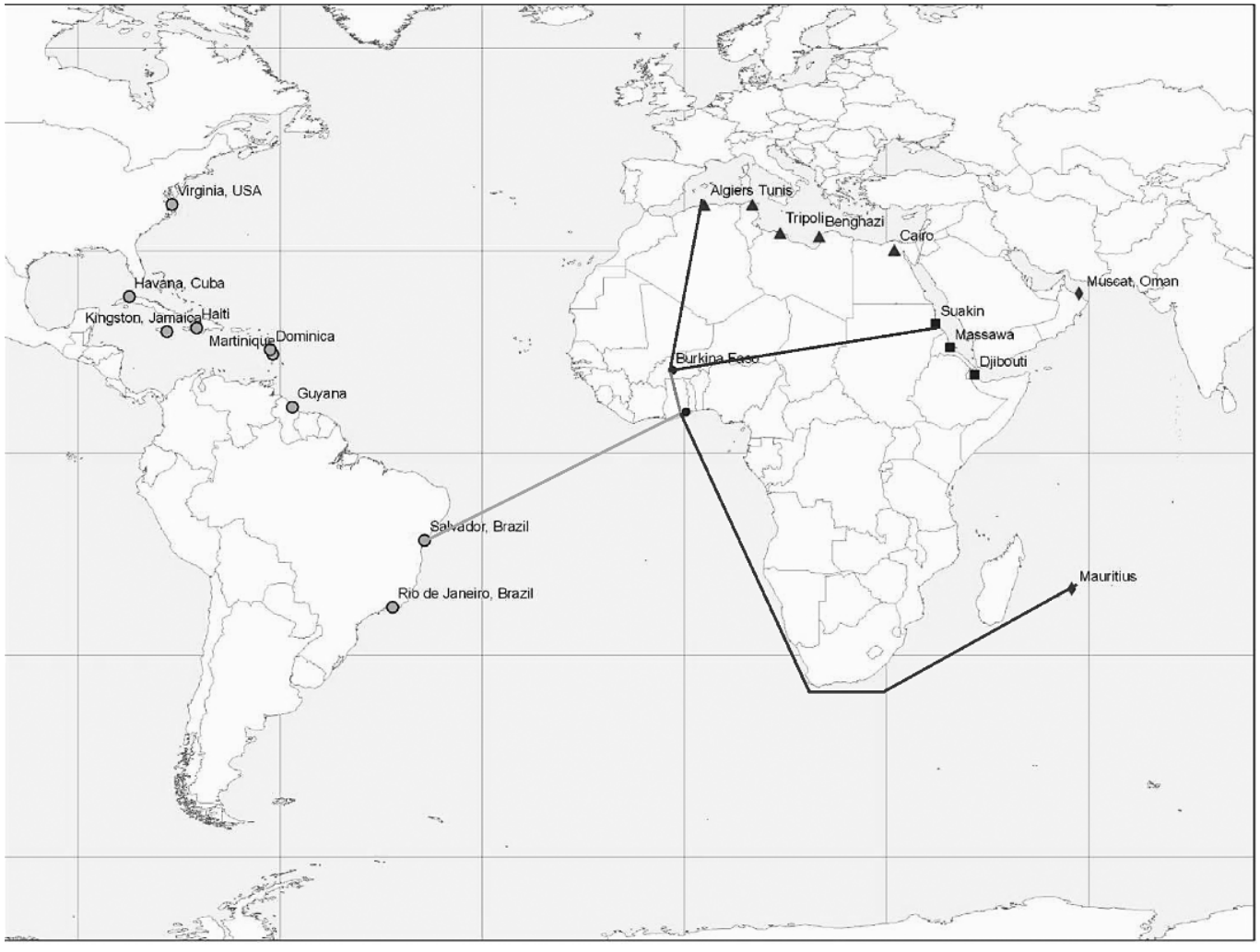


FIGURE V
Example Showing the Distance Instruments for Burkina Faso

Country centroids

- Feature To Point
- Add XY Coordinates

(cf. Lec 2 Ex 1)

Closest point on the coast

- Use the Near tool.

Near

- Input Features: country centroid points (the one created by Add XY Coordinates)
- Near Features: 10m-coastline.shp
- Check “Location”
  - Otherwise the coordinate of the nearest point on the coast won’t be attached as attributes.

Note: this tool overwrites the input file

Now export the script and edit it.

- For the Near tool command, 2nd argument does not use the local variable
- Because near features can be more than one
- Replace this with your own local variable

Now browse the attribute table of the output

- The Near tool creates a field “NEAR_DIST”
- Which is rubbish as the distance is calculated in degrees
- If the input features use the UTM projection (cf. L3), this field can be useful.

Assignment 4

- Obtain the distance from each ethnic boundary's centroid to its nearest point on the coast (the instrument in Nunn & Wantchekon 2011)

6. What we've learned on ArcGIS

- Calculate the area of intersections of countries and ethnic boundaries in Africa
- Obtain the coordinate of a polygon centroid
- Obtain the distance between a polygon centroid and its nearest location on the polyline features.

Do you remember which geoprocessing tools you used for each of these tasks?

References for Lecture 4

Algan, Yann, and Pierre Cahuc. 2010. "Inherited Trust and Growth." *American Economic Review* 100(5): 2060-2092.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference With Multiway Clustering." Journal of Business & Economic Statistics 29(2): 238–249.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2008. "Bootstrap-based Improvements for Inference with Clustered Errors." Review of Economics & Statistics 90(3): 414-427.

Knack, Stephen, and Philip Keefer. 1997. "Does Social Capital Have an Economic Payoff? A Cross-country Investigation." *Quarterly Journal of Economics* 112: 1251-1288.

McKenzie, David, John Gibson, and Steven Stillman. 2010. "How Important Is Selection? Experimental vs. Non-Experimental Measures of the Income Gains from Migration." *Journal of the European Economic Association* 8(4): 913-945.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71(4): 1027-1048.

Murdock, George Peter. 1959. *Africa: Its People and Their Cultural History*. New York: McGraw-Hill Book Company.

Nunn, Nathan. 2008. "The Long-Term Effects of Africa's Slave Trades." *Quarterly Journal of Economics* 123(1): 139-176.

Nunn, Nathan, and Leonard Wantchekon. 2011. "The Slave Trade and the Origins of Mistrust in Africa." American Economic Review 101(7): 3221–3252.

Oster, Emily. 2012. "HIV and sexual behavior change: Why not Africa?" Journal of Health Economics 31(1): 35–49.

Tabellini, Guido. 2010. "Culture and Institutions: Economic Development in the Regions of Europe." Journal of the European Economic Association 8(4): 677–716.

Woodruff, Christopher, and Rene Zenteno. 2007. "Migration networks and microenterprises in Mexico." *Journal of Development Economics* 82(2): 509-528.

Wooldridge, Jeffrey M. 2002. Econometric Analysis of Cross Section and Panel Data. Cambridge, Massachusetts: MIT Press.