GEOGRAPHIC AND INSTITUTIONAL DETERMINANTS OF REAL INCOME: A SPATIO-TEMPORAL SIMULTANEOUS EQUATION APPROACH

Guyslain K. Ngeleza¹
Raymond J.G.M. Florax²,³
William A. Masters²

¹Mississippi State University, Starkville, USA
²Purdue University, West Lafayette, USA
³Vrije Universiteit, Amsterdam, The Netherlands

E-mail: ngeleza@agecon.msstate.edu
Introduction: Income per capita distribution

Data
- PWT 6.2

Facts
- In 1960, developed countries on average 5 times richer than Africa
- In 2000 developed are 8 times richer than Africa
- But also variation across regions within Africa

N B: The vertical axis represents the ratio of the average income per capita for developed country relative to developing country in a given region.
Africa Income per capita distribution

Data
- PWT 6.2

Facts
- income per capita significantly increased in North Africa and Southern Africa
- stagnant in West and East Africa
- falling in central Africa

N B: The vertical axis represents the ratio of the average income per capita for developed country relative to developing country in a given region.
Determinants of income per capita

Institutions

– North, Acemoglu et al.

- focus: property rights, market infrastructure, price incentives
- policy implication: adopt right institutions, convergence

Geography

– Sachs, Diamond

- focus: location, local technology
- Policy implication: overcoming geography comes at a cost
  - location-specific technology
    - transportation
    - public health
Determinants of Africa's poor performance

**External conditions (Bloom and Sachs)**
- Legacy of centuries of slave trading and colonial rule and manipulation of African politics during the cold war

**Social conditions (Easterly and Levine)**
- Ethnic divisions, religious diversity and low levels of social capital

**Trade specialization (Temple)**
- Heavy dependence on a small number of primary exports, with declines and volatility in terms of trade

**Low level of urbanization (Acemoglu et al)**
- Average 16% in 1960
- Average 35% in 2002
Central question
National policies and institutions interacting with trade, urbanization
- corruption, civil liberty, rule of law
- geography, transportation opportunities, and agriculture

Neighborhood effects
- characteristics of places
- spillovers among neighbors

Hypotheses
- spillovers among neighbors and shared characteristics of places could help better explain how institutions, trade and urbanization affect income
Challenges in testing these hypotheses:

- **Endogeneity**
  - correlation between institutions and income could be due to reverse causality, or to omitted variables that affect both of them

- **Spatial correlation**
  - interactions among neighbors such as trade or migration
  - geographic clustering could be due to spatially correlated attributes
  - misspecification
A new spatial analysis of economic growth

To account for both spatial effects and endogeneity

- apply new estimators (Kelejian and Prucha 2004)
- use panel data in simultaneous equations (Masters and Sachs 2001)

To distinguish between kinds of neighborhood effects

- spillovers among neighbors
- shared characteristics
<table>
<thead>
<tr>
<th>Eq.# and Variable</th>
<th>Description (all 5-yr. averages, centered on 1960, 1965...2000)</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Agric. output</td>
<td>net farm output/r capita (thousands of 2000 International dollars)</td>
<td>FAOStat</td>
</tr>
<tr>
<td>2. Infant mortality</td>
<td>deaths in first year, per thousand live births</td>
<td>WHO</td>
</tr>
<tr>
<td>3. Institutional qual.</td>
<td>average of six ICRG indexes, linked to two Freedom House indexes</td>
<td>ICRG+FH</td>
</tr>
<tr>
<td>4. Urbanization</td>
<td>percentage of the population living in urban areas</td>
<td>FAOStat</td>
</tr>
<tr>
<td>5. Trade</td>
<td>exports plus imports as a fraction of GDP</td>
<td>PWT 6.2</td>
</tr>
<tr>
<td>6. Income</td>
<td>real GDP per capita (thousands of 2000 PPP dollars)</td>
<td>PWT 6.2</td>
</tr>
<tr>
<td>1. Agricultural land</td>
<td>land used in agriculture (thousands of hectares)</td>
<td>FAOStat</td>
</tr>
<tr>
<td>1. Land quality</td>
<td>farmland in high soil fertility classifications (percent)</td>
<td>USDA</td>
</tr>
<tr>
<td>1. Frost</td>
<td>land receiving ≥5 frost-days/month in winter (share)</td>
<td>IPCC</td>
</tr>
<tr>
<td>1. Rainfall</td>
<td>average precipitation over the country’s land mass (mm/yr)</td>
<td>IPCC</td>
</tr>
<tr>
<td>2. Malaria</td>
<td>ecological index of malaria transmissibility</td>
<td>Kiszewski et al.</td>
</tr>
<tr>
<td>4. Catholic</td>
<td>estimated percentages of the population with the specified religion</td>
<td>Barro and Lee</td>
</tr>
<tr>
<td>4. Protestant</td>
<td>estimated percentages of the population with the specified religion</td>
<td>Barro and Lee</td>
</tr>
<tr>
<td>4. Muslim</td>
<td>estimated percentages of the population with the specified religion</td>
<td>Barro and Lee</td>
</tr>
<tr>
<td>5. Coastal</td>
<td>percent of a country’s land within 100 km of sea or navigable river</td>
<td>Sachs</td>
</tr>
<tr>
<td>6. Population</td>
<td>Millions</td>
<td>PWT 6.2</td>
</tr>
</tbody>
</table>
Data: Overview

Coverage and time frame

- 30 countries
- 1960–2000

Weights matrix

- Inverse distance matrix (W)
- Cut-off point 2000 miles, standardized
- \( n = 30 \), 33% nonzero, min/max links 2/4, average 10

Spatial Variables

- Spatial lags (weighted average of the values of neighbors)
Creating the Weight Matrix

\[
W_{ij} = \begin{bmatrix}
0 & 1/d_{1,2} & \cdots & \cdots & 1/d_{1,30} \\
1/d_{2,1} & 0 & \cdots & \cdots & 1/d_{2,30} \\
\vdots & \vdots & \ddots & & \\
\vdots & \vdots & \ddots & \ddots & \\
1/d_{30,1} & 1/d_{30,2} & \cdots & \cdots & 0
\end{bmatrix}
\]

where \( w_{ij} = \begin{cases} 
1/d_{ij} & \text{if } 0 < d_{ij} \leq 2000 \text{miles} \\
0 & \text{otherwise}
\end{cases} \)

\( d_{ij} \) represents the distance between country \( i \) and country \( j \) measured from the center.
Standardized Weight Matrix (see Bell and Bockstael 2000)

\[
S_{ij} = \begin{bmatrix}
0 & S_{1,2} & \cdots & \cdots & S_{1,30} \\
S_{2,1} & 0 & \ddots & & \ddots \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
S_{30,1} & S_{30,2} & \cdots & \cdots & 0
\end{bmatrix}
\]

Where

\[
S_{ij} = \frac{d_{ij}}{\sum_{j=1}^{30} \frac{1}{d_{ij}}}
\]
Spatial lag variable for income per person

\[
\begin{bmatrix}
0 & S_{1,2} & \cdots & \cdots & S_{1,30} \\
S_{2,1} & 0 & \cdots & \cdots & S_{2,30} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
S_{30,1} & S_{30,2} & \cdots & \cdots & 0 \\
\end{bmatrix}
\times
\begin{bmatrix}
y_{1,2000} \\
y_{2,2000} \\
\vdots \\
y_{30,2000} \\
\end{bmatrix}
= 
\begin{bmatrix}
0 + y_{2,2000}S_{1,2} + \cdots y_{30,2000}S_{1,30} \\
y_{1,2000}S_{2,1} + 0 + \cdots y_{30,2000}S_{2,30} \\
y_{1,2000}S_{30,1} + y_{2,2000}S_{30,2} + \cdots 0 \\
\end{bmatrix}
\]

Weights Matrix  
Income per person  
Income spatial lags variable (Wpci)
Unequal world income distribution 2000

- year 2000 constant PPP US$
- strong geographic cluster
- dramatically low in Africa

- Moran scatterplot
- Gradient of the trend is the coefficient of spatial correlation
- strong cluster
Income variation across Africa

- year 2000 constant PPP US$
- moderate clustering
- high in North and South Africa
- moderate in coastal regions

GDP per capita
- 0-384
- 385 - 1113
- 1114 - 1822
- 1823 - 3072
- 3073 - 5664
- 5665 - 10593
- no data

moderate clustering
World Agricultural Output 2000

- dramatically low
  - Africa
- high
  - North America
  - Europe
  - Australia

Spatial autocorrelation low, but positive
Agricultural Output Across Africa 2000

- **high**
  - Egypt
  - South Africa
  - Tunisia
  - Uganda
- **lowest**
  - D. R. Congo

Spatial autocorrelation low, but positive
World Trade Share 2000

- patterns less obvious
- highest
  - Singapore
  - Guyana
  - Malaysia
  - Ireland

- no significant spatial clustering in trade
Patterns less obvious

highest
- Angola
- Gambia
- Tunisia
- Senegal

no significant spatial clustering in trade
World Infant Mortality Rate 2000

- highest
  - India
  - Pakistan
  - Iraq
  - Sub-Saharan Africa

- low
  - North America
  - Europe
  - Australasia

- (+/+ cluster)
  - Africa, Asia
Infant Mortality Rate Across Africa 2000

- **Highest**
  - Central Africa
- **Low**
  - North Africa
  - South Africa

- **Strong cluster in central Africa**
World Institutions Quality 2000

- Lowest
  - Africa
  - Asia

- Highest
  - North America
  - Northern Europe
  - and Australasia

- Intermediate zone
  - Southern Europe

- Two clusters
  - (+/+ ) industrial
  - (−/− ) rest of the world
Institutional Quality Across Africa 2000

- highest
  - South Africa

Moderate clusters
World Urbanization 2000

- Uniform
- 50% for most countries
- But low in Africa and Southeast Asia

- strong clusters for ++ and --
High
- North Africa
- South Africa

strong clusters in east Africa
A simultaneous-equations model

(1) \[ agoutput_{it} = a_1 + \beta_{11}agland_{it} + \beta_{12}landqual_{i} + \beta_{13}frost_{i} + \beta_{14}rainfall_{it} + \beta_{15}malaria_{i} + \delta_{1t} + \epsilon_{1it} \]

(2) \[ imrate_{it} = a_2 + \beta_{21}income_{it} + \beta_{22}malaria_{i} + \delta_{2t} + \epsilon_{2it} \]

(3) \[ instqual_{it} = a_3 + \beta_{31}income_{it} + \beta_{32}imrate_{it} + \beta_{33}pctcath_{i} + \beta_{34}pctprot_{i} + \beta_{35}pctmus_{i} + \delta_{3t} + \epsilon_{3it} \]

(4) \[ urbanization_{it} = a_4 + \beta_{41}income_{it} + \beta_{42}agoutput_{it} + \beta_{43}coastal_{i} + \delta_{4t} + \epsilon_{4it} \]

(5) \[ trade_{it} = a_5 + \beta_{51}income_{it} + \beta_{52}agoutput_{it} + \beta_{53}coastal_{i} + \beta_{54}population_{it} + \delta_{5t} + \epsilon_{5it} \]

(6) \[ income_{it} = a_6 + \beta_{61}income_{it} + \beta_{62}coastal_{i} + \beta_{63}population_{it} + \delta_{6t} + \epsilon_{6it} \]

...
Econometric method

Issues
- endogeneity: system approach
- spatial dynamics: ARAR

ARAR single equation
- lag and error correlation
  \[ y = \rho Wy + X\beta + \varepsilon, \]
  \[ \varepsilon = \lambda W\varepsilon + \mu \]

System of \( m \) equations
\[ Y = \bar{Y}P + Y\Gamma + XB + U \]
\[ U = \bar{U}\Lambda + E \]

Estimators

Feasible GS3SLS estimator (Kelejian and Prucha 2004)
- estimate equations with 2SLS
  - choice of instruments
- obtain 2SLS residuals and use GM to determine \( \lambda \)
- perform Cochrane-Orcutt transformation, apply feasible GS2SLS
- use feasible GS3SLS to obtain full information results
Empirical results (i)

exploratory results (OLS)

<table>
<thead>
<tr>
<th>Variables</th>
<th>Agricultural output</th>
<th>Infant mortality rate</th>
<th>Institutional quality</th>
<th>Urbanization</th>
<th>Trade</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>LM-error</td>
<td>1.560</td>
<td>0.379</td>
<td>0.922</td>
<td>2.605**</td>
<td>3.284**</td>
<td>0.468</td>
</tr>
<tr>
<td>Robust LM-error</td>
<td>0.255</td>
<td>3.977**</td>
<td>2.564*</td>
<td>0.494</td>
<td>1.601</td>
<td>15.762**</td>
</tr>
<tr>
<td>LM-lag</td>
<td>1.654*</td>
<td>3.093**</td>
<td>2.217</td>
<td>2.392*</td>
<td>4.819**</td>
<td>18.572**</td>
</tr>
<tr>
<td>Robust LM-lag</td>
<td>0.349</td>
<td>6.690***</td>
<td>3.860**</td>
<td>0.281</td>
<td>3.137**</td>
<td>33.866**</td>
</tr>
</tbody>
</table>

Significance is indicated by ***, ** and * for the 1, 5, and 10 per cent level, respectively

LM – lag test:

\[ H_0 : \rho = 0 \]

\[ H_1 : y = \rho Wy + X\beta + \varepsilon \]

LM – Error test:

\[ H_0 : \lambda = 0 \]

\[ H_1 : \varepsilon = \lambda W \varepsilon + u \]

All variables display spatial dependence

no clear spatial-error for agricultural output
Empirical results (ii)

Regression output, system estimation, 3SLS not allowing for spatial spillovers

<table>
<thead>
<tr>
<th>Variables</th>
<th>Agricultural output</th>
<th>Infant mortality rate</th>
<th>Institutional quality</th>
<th>Urbanization</th>
<th>Trade</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural output</td>
<td></td>
<td>-0.410***</td>
<td></td>
<td></td>
<td></td>
<td>0.430***</td>
</tr>
<tr>
<td>Agricultural land</td>
<td>0.102**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land quality</td>
<td>0.021*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frost</td>
<td>0.016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainfall</td>
<td>-0.164***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaria</td>
<td>0.040** 0.038***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant mortality rate</td>
<td>-0.637***</td>
<td></td>
<td></td>
<td>-1.037***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-0.337*** -0.146*** 0.551*** 0.624***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institutional quality</td>
<td>-0.166</td>
<td></td>
<td></td>
<td></td>
<td>0.312***</td>
<td></td>
</tr>
<tr>
<td>trade</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>-0.006***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>-0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>0.002</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urbanization</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.507***</td>
</tr>
<tr>
<td>Coastal</td>
<td>0.035*** -0.013*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td></td>
<td></td>
<td>-0.079***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significance is indicated by ***, ** and * for the 1, 5, and 10 per cent level, respectively

- Comparison purposes
- Account for endogeneity
- No spatial dependency
- System context
Empirical results (iii)

Regression output, system estimation, full information estimator for the ARAR

<table>
<thead>
<tr>
<th>Variables</th>
<th>Agricultural output</th>
<th>Infant mortality rate</th>
<th>Institutional quality</th>
<th>Urbanization</th>
<th>Trade</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>W· agricultural output</td>
<td>0.837***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W· infant mortality</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W· institutional quality</td>
<td>0.821***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W· urbanization</td>
<td>0.501***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W· trade</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W· income</td>
<td>0.196***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significance is indicated by ***, ** and * for the 1, 5, and 10 per cent level, respectively.
Empirical results (iv)

Regression output, system estimation, full information estimator for the ARAR

<table>
<thead>
<tr>
<th>Variables</th>
<th>Agricultural output</th>
<th>Infant mortality rate</th>
<th>Institutional quality</th>
<th>Urbanization</th>
<th>Trade</th>
<th>Income</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agricultural output</td>
<td>−0.186***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.112**</td>
</tr>
<tr>
<td>Agricultural land</td>
<td>0.122**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Land quality</td>
<td>0.004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frost</td>
<td>0.026</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rainfall</td>
<td>−0.055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaria</td>
<td>0.054** 0.068***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Infant mortality rate</td>
<td>−0.188**</td>
<td></td>
<td></td>
<td>−0.753***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>−0.257*** 0.026</td>
<td>0.485*** 0.211**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Institutional quality</td>
<td></td>
<td></td>
<td>0.613***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trade</td>
<td></td>
<td></td>
<td></td>
<td>−0.120*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>−0.003**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Muslim</td>
<td>−0.001</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urbanization</td>
<td></td>
<td></td>
<td></td>
<td>0.536***</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coastal</td>
<td></td>
<td></td>
<td></td>
<td>0.035*** 0.016</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td></td>
<td></td>
<td></td>
<td>−0.128***</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Significance is indicated by ***, ** and * for the 1, 5, and 10 percent level, respectively.
Summary

determinants of country income within Africa.

- “institutions first”
  - Institutions shaped by external conditions and social conditions

- “geography first”
  - Geography influences both trade and urbanization

Novelty of the paper:

- Correct omitted variable due to spillovers among neighboring countries
- Correct misspecification of error structure due shared characteristics of places

In doing so, we found

- Significant spatial lags among most endogenous variables
- Support for both “geographic” and “institutional” claims
Implications

- Good institutions needed in poorest regions in Africa
- Investment where technologies are location-specific
  - transportation sector
  - public health

Implications of spatial analysis for growth theory
- Significant lags suggest that determinants of economic growth should not be examined in isolation
- Spatial method rigorously applied produces better results
Questions