

Multivariate and Geospatial Analysis of ICT Utilization in Africa

Avijit Sarkar, James Pick, Jeremy Johnson
University of Redlands, School of Business

Esri User Conference
San Diego CA
July 2015

Goals of this paper

- Examine the **digital divide** in Africa.
 - Enhance understanding of factors that impact ICT utilization and the digital divide in African countries
 - Digital divide: *“the gap between individuals, households, businesses and geographic areas at different socio-economic levels with regard both to their opportunities to access information and communication technologies and to their use of the Internet for a wide variety of activities”* (OECD, 2011).
- Overall research question
 - What factors determine ICT adoption and utilization for African nations and how does ICT adoption and utilization vary geographically?

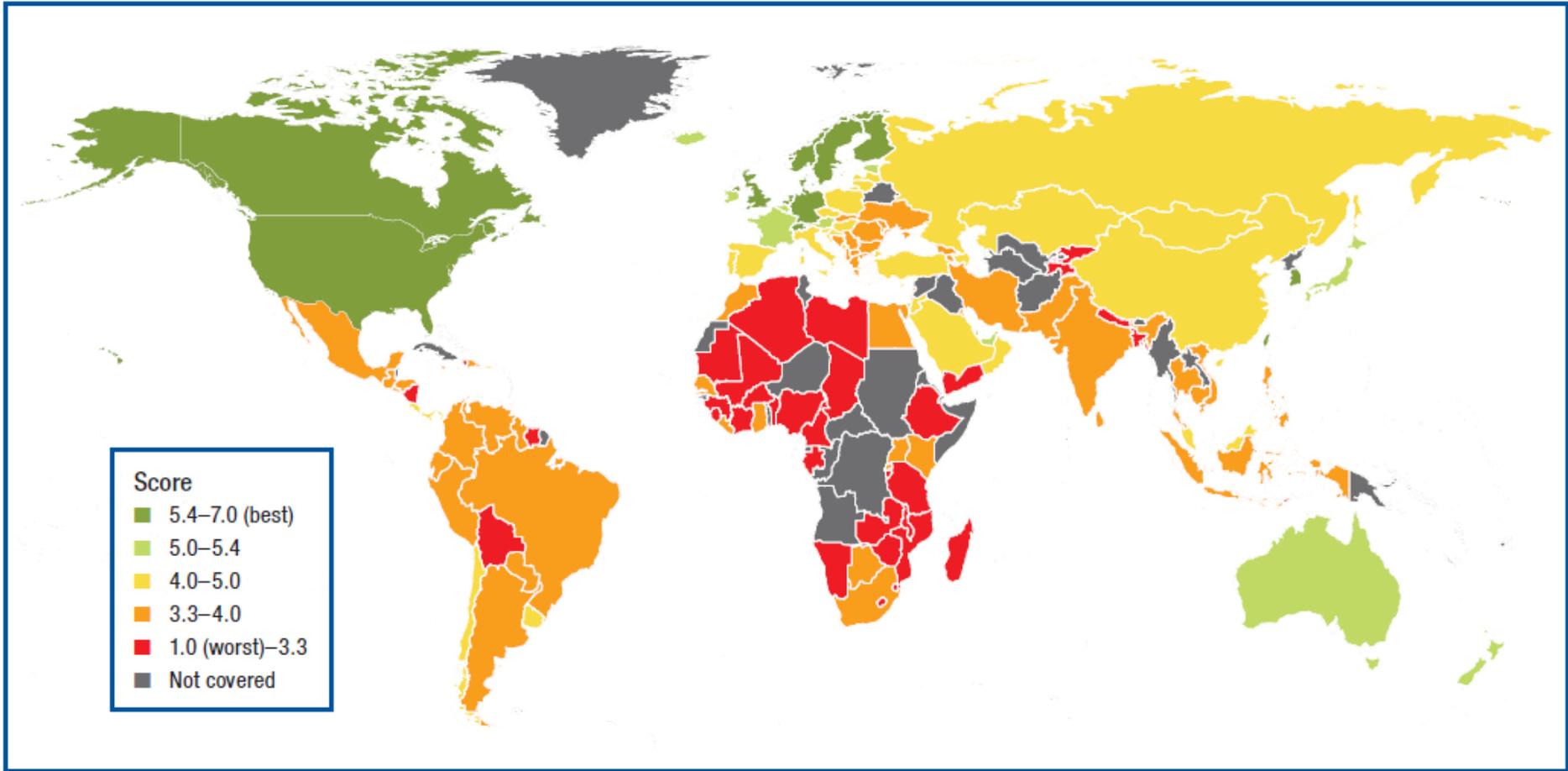
Outline of presentation

- Introduction
- Literature/prior research
- Research questions
- Conceptual model
- Methodology
- Findings
- Policy Implications
- Conclusion

Introduction: Africa & ICT in Africa

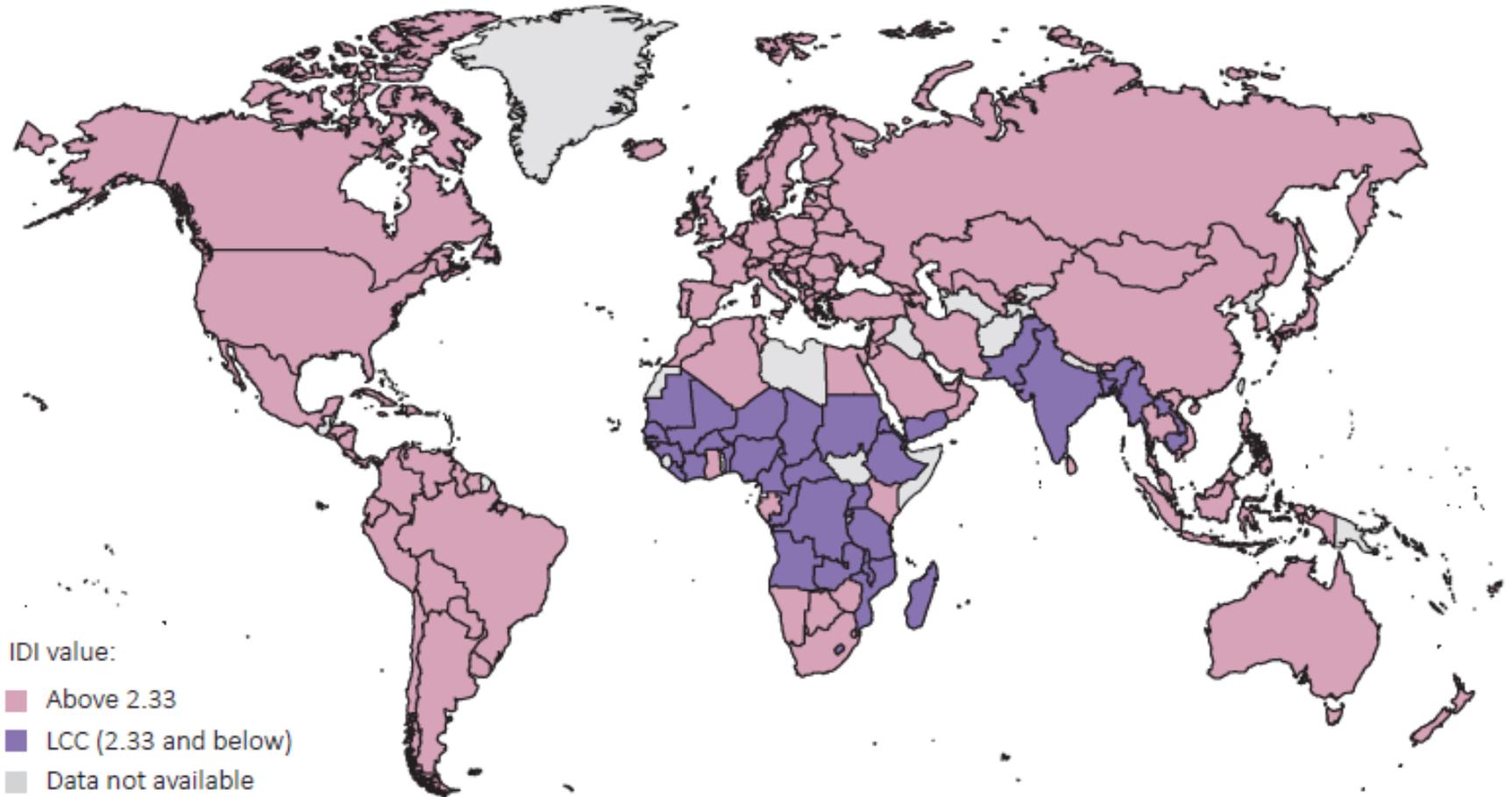
- Africa population: 1.03 billion in 2010.
- Several nations especially in sub-Saharan Africa classified as low income or lower middle income (The World Bank, 2012).
 - Lag behind developed and developing nations in social, health, economic frontiers.
- Most nations rank at the bottom end of the UN's Human Development Index ranking.
- Information and Communication Technology (ICT) in Africa:
 - **Fixed landline**: 1:100 (excluding South Africa 1:200, 1:50 worldwide).
 - **Internet usage**
 - **In 2005**: 14% of world population, but only 2% of world's internet users.
 - **In 2013**: 15.3% of world population, 7% of world's internet users.
 - **Internet penetration rate in 2012**: 15.6% of total population (world: 34.3%, North America: 78.6%, Europe: 63.2%, Latin America: 42.9%).

The Networked Readiness Index Map (WEF, 2013)



Only two African nations, Mauritius (55th) & South Africa (70th) rank in the top half (144 nations) of NRI Rankings (WEF 2013).

Least Connected Countries (ITU, 2012)



State of ICT in Africa (ITU, 2012)

Chart 2.9: IDI values compared with the global, regional and developing/developed-country averages, Africa, 2012

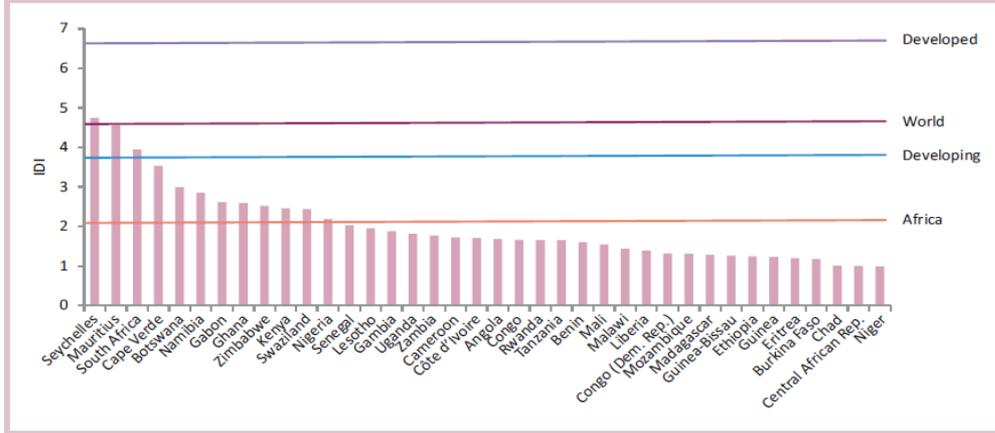


Chart 4.1: Digital natives as a percentage of total population, top countries (left) and bottom countries (right), 2012

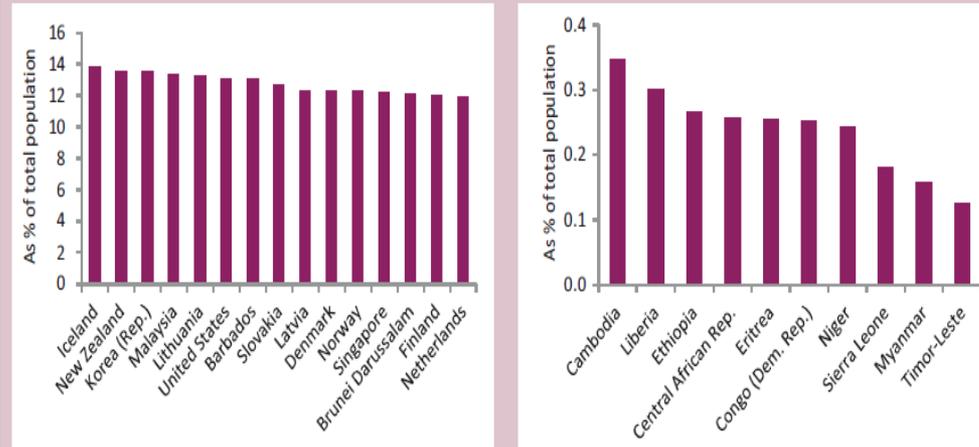
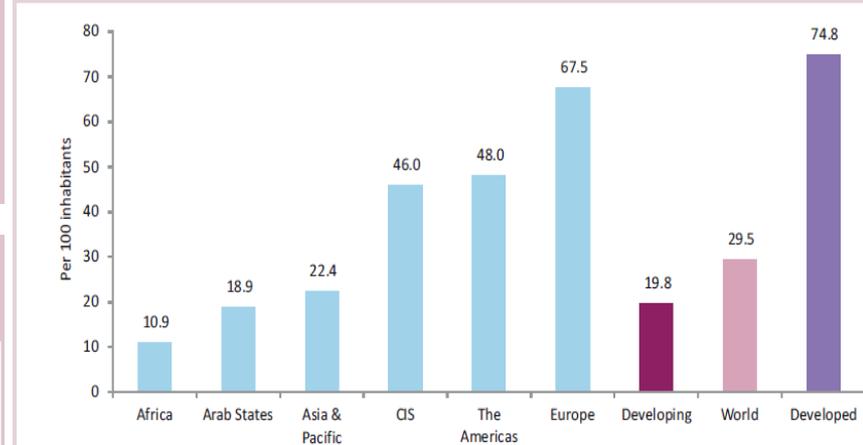
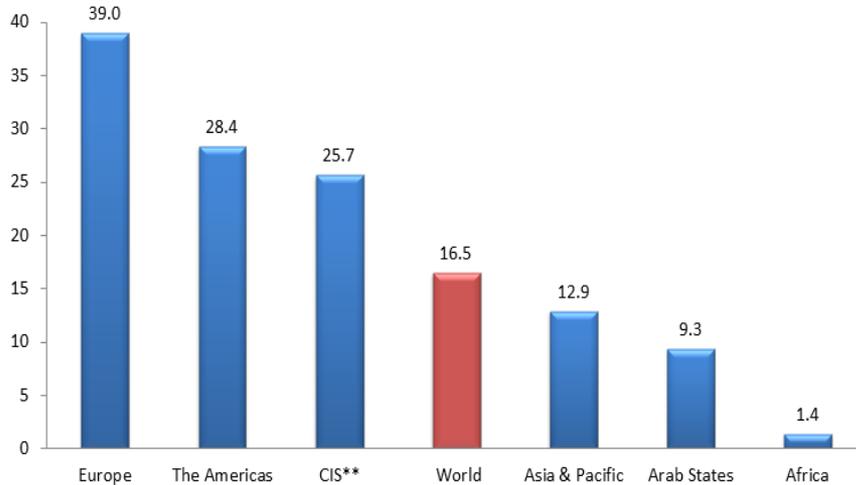


Chart 1.3: Active mobile-broadband subscriptions, by region and level of development, 2013*

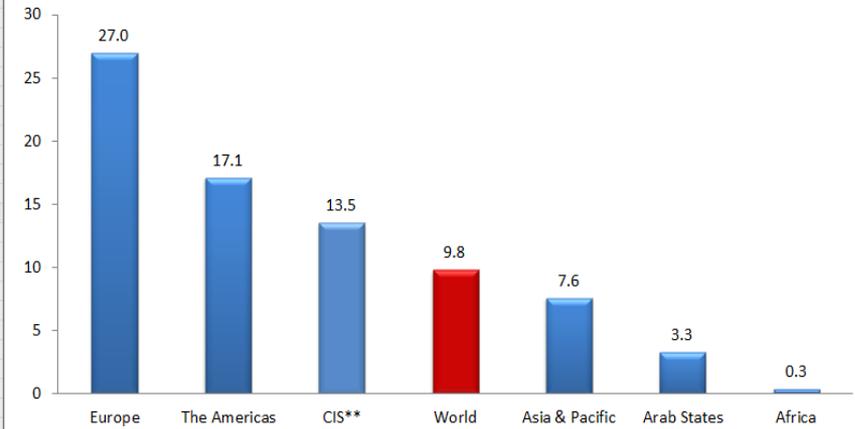


ICT in Africa vs. Other World Regions

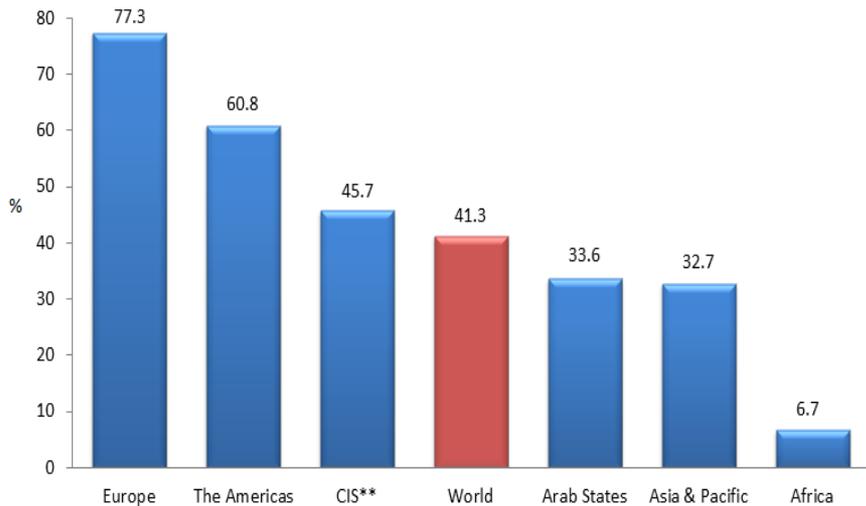
Fixed-telephone subscriptions per 100 inhabitants, 2013*



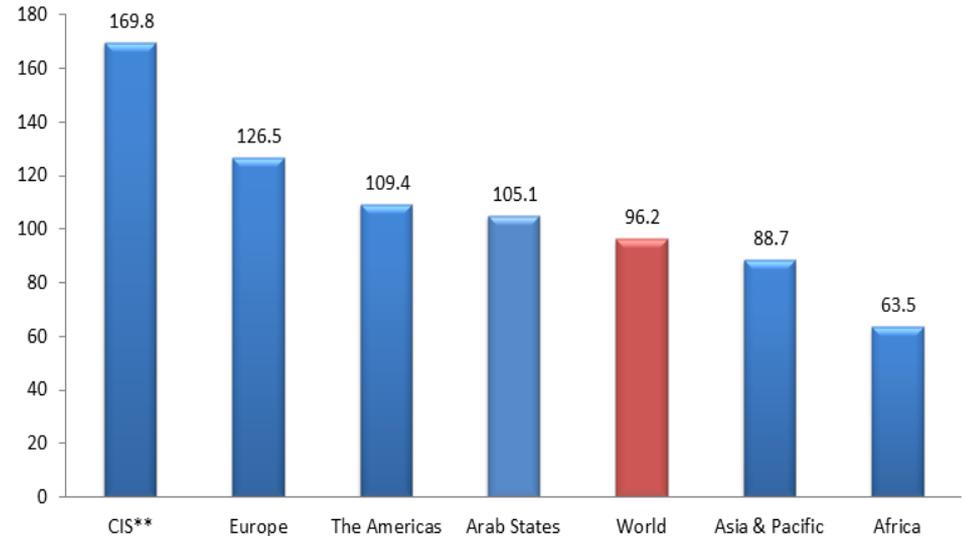
Fixed (wired)-broadband subscriptions per 100 inhabitants, 2013*



Percentage of households with Internet access, by region, 2013*

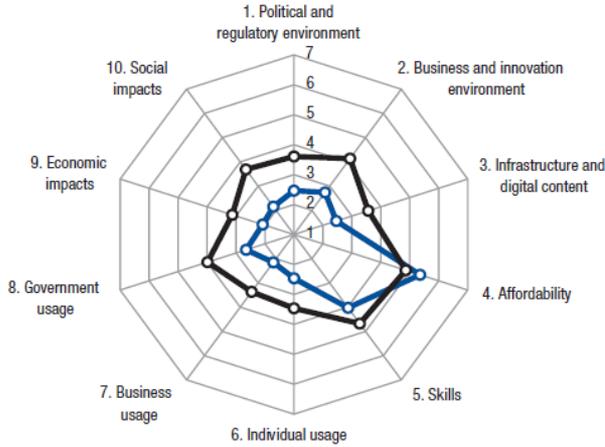


Mobile-cellular subscriptions per 100 inhabitants, 2013*

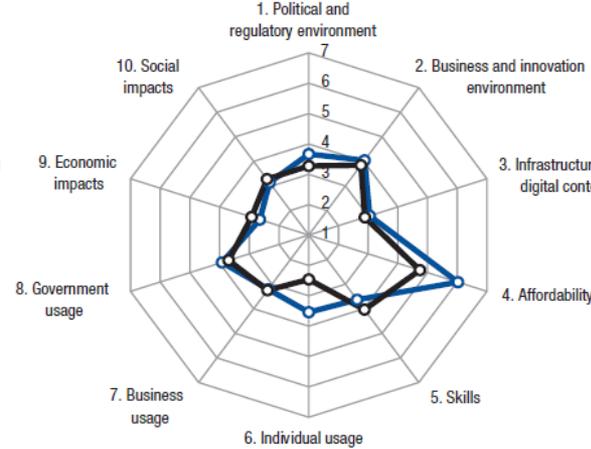


Dichotomy of Networked Readiness in Africa

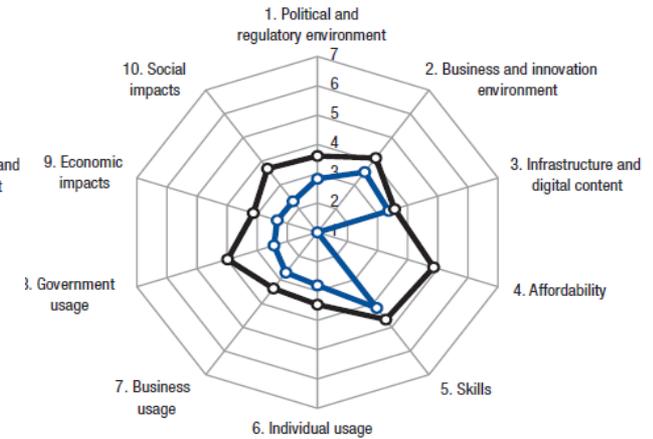
(WEF, 2013)



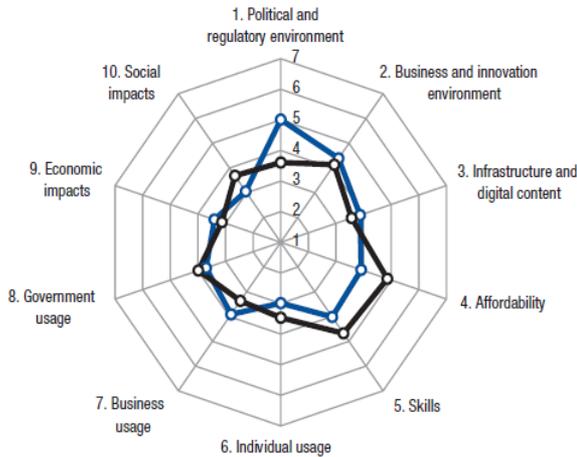
Algeria Upper-middle-income group average



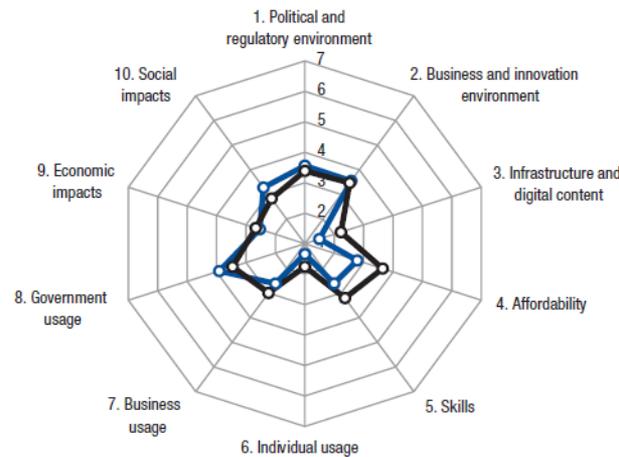
Morocco Lower-middle-income group average



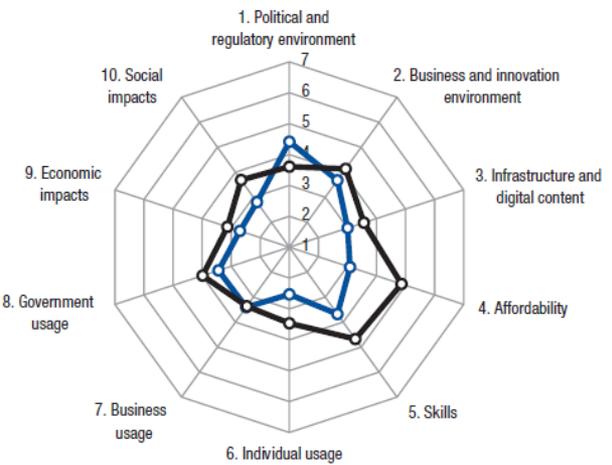
Libya Upper-middle-income group average



South Africa Upper-middle-income group average



Ethiopia Low-income group average



Namibia Upper-middle-income group average

Digital Divide Literature

- Digital divide differences have been studied extensively from a variety of perspectives.
- Significant empirical literature explains influences on access, adoption, and use of ICTs.
- Important determinants of ICT adoption and utilization from prior studies are:
 - Education
 - Income
 - Age
 - Ethnicity
 - Urban location
 - Higher education govt funding
 - ICT expenditure
 - Infrastructure
 - Innovation & R&D
 - Newspaper, magazine, & book publishers
 - Labor force
 - Societal openness
- For Africa
 - A handful of continent-wide studies at the country level.
 - Usually focused on a specific ICT.
 - A few country specific studies that focus on a specific context: **Egypt** (Warschaeur, 2003), **Ghana** (Alemna and Sam, 2006), **Nigeria** (Bankole, Bankole, and Brown, 2011), **South Africa** (Mbatha, Ocholla, and Le Roux, 2011), **Tanzania** (Furuholt and Kristiansen, 2007).

Review of prior Africa literature

- Onyeiwu (2002)
 - South Africa, Namibia, Morocco, Tunisia, Botswana, Cape Verde, and Seychelles have made significant progress in ICT adoption and access.
 - Sub-Saharan Africa lagged far behind (socio-economic foundations are shaky, e.g., Angola (148) and Cameroon (150) in UN HDI (2013)).
- A few studies have examined correlates of a specific form of ICT.
 - Roycroft and Anantho (2003)
 - Dependent variable: internet subscription in Africa
 - Independent variables: **economic development**, estimated by the UN HDI, **international internet bandwidth**, domestic internet hosts, **ISP market structure**, the cost of a local telephone call, and the use of **English as an official language**.
 - Oyelaran-Oyeynka and Lal (2005)
 - Dependent variable: internet diffusion (sub-Saharan Africa)
 - Independent variables: human capital, telecom infrastructure investment, **PC density**, telephone density, **presence of internet hosts**, and **economic wealth estimated by GDP per capita**.
- Multi-ICT studies
 - Andoh-Baidoo, Osatuyi, and Kunene (2013)
 - Dependent variable: ICT capacity – estimated by five dependent variables: ICT expenditure as a percent of GDP, ICT expenses per capita, investments in telecommunications with private sector participation, internet users, and mobile phone subscribers per capita.
 - Independent variables: economic, human development, political, and population conditions.
 - Human development factors such as life expectancy, enrollments in tertiary education and expenditures per student, and health expenses as percent of GDP are correlates of ICT capacity.
 - Onyeiwu (2002)
 - Dependent variable: IT index comprised of five components: the number of internet hosts, internet users, PCs, telephone lines, and cellular phones per capita.
 - Openness of the economy is the sole significant predictor of the extent of digitization as measured by the IT index.

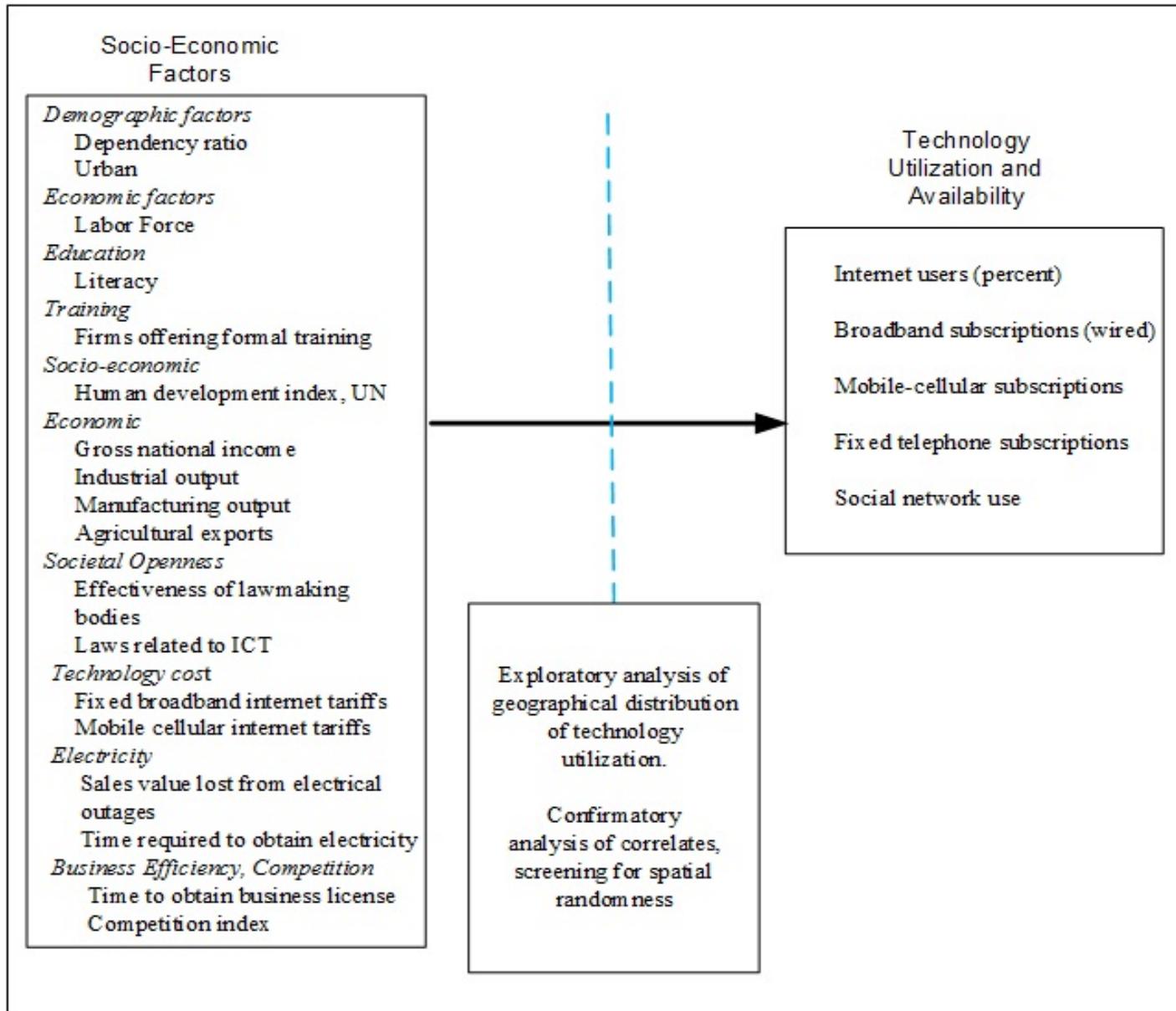
Unique features of this work

1. Examination of factors that influence the digital divide at the country level for Africa does not exist post 2010.
2. Inclusion of social media technologies of Facebook, Twitter, & LinkedIn among dependent variables.
3. Spatial analysis (clustering and detection of spatial autocorrelation that may bias regression results) are new to MIS research.

Research Questions

1. What is an **appropriate conceptual model** to understand the geography of, and influences on levels of technologies in African countries?
2. What is the **spatial clustering** of levels of technologies for African nations and how do the clusters differ?
3. How agglomerated are the technology variables in Africa, based on **spatial autocorrelation**?
4. What are the **social, economic, societal openness, infrastructure, and business factors** that influence levels of technologies for the entire sample of African nations, and for a lower-tech subsample?
5. Does the regression model account for spatial agglomeration so the **regression error terms** are spatially random?

Conceptual Model of ICT Utilization



DEPENDENT VARIABLES

	Variable	Definition	YEAR	Source	Mean (N)	SD
ICT use	Fixed broadband internet subscriptions per 100 people	Fixed (wired) broadband subscriptions per 100 persons	2012	ITU	13.83 (53)	14.15
ICT use	Fixed telephone subscriptions per 100 people	Fixed telephone subscriptions per 100 persons	2012	ITU	1.00 (52)	2.32
ICT use	Internet users per 100 people	Percent of individuals using the internet	2012	ITU	71.76 (54)	41.00
ICT use	Mobile subscriptions per 100 people	Mobile-cellular telephone subscriptions per 100 persons	2012	ITU	3.77 (54)	6.27
ICT use	Use of social networks per 100 people	How widely virtual social networks are used (e.g., Facebook, Twitter, LinkedIn) per 100 persons	2010-11	WEF	4.43 (54)	0.78

Independent Variables

Demographic	Young dependency ratio	Ratio of population aged 0-14 to population aged 15-64	2010	WB	73.06 (45)	16.59
Demographic	Urban	Urban population as percent of total population	2010	WB	41.43 (46)	18.40
Education	Literacy	Adult literacy rate per 100 people	2005-10	WB	62.55 (42)	18.41
Training	Firms offering formal training	Percent of firms offering formal training	2005-10	WB	30.54 (42)	10.80
Socio-economic	Human development index of UN	Composite index for long and healthy life, knowledge and a decent standard of living.	2010	UN	0.80 (51)	0.401
Economic	Gross national income per capita	Gross national income per capita (\$)	2010	WB	3,194.08 (49)	6,673.95
Economic	Industry output	Industry output as a percent of GDP	2010	WB	32.92 (38)	17.34
Economic	Manufacturing output	Manufacturing output as a percent of GDP	2010	WB	10.50 (45)	7.65
Economic	Agricultural exports	Agricultural exports as percent of total exports	2010	WB	7.65 (37)	12.23
Economic	Foreign Direct Investment	Foreign Direct Investment per person	2010	WB	681.46 (45)	792.89
Societal openness	Effectiveness of lawmaking bodies	How effective is the national parliament/congress as a law-making institution (scale 1 lowest to 7 highest)	2010-11	WEF	3.49 (32)	0.784
Societal openness	Laws related to ICT	Laws relating to ICT (scale 1 lowest to 7 highest)	2010-11	WEF	3.35 (34)	0.72
Technology cost	Fixed broadband internet tariffs	Subscription charge for fixed (wired) broadband internet service (in PPP \$/month)	2010	WEF	225.44 (33)	385.10
Technology cost	Mobile cellular tariffs	Average per-minute cost of different types of mobile cellular calls in PPP (\$/min)	2010	WEF	0.41 (34)	0.18
Infrastructure	Sales value lost from electrical outages	Percent of sales value lost due to electrical outages	2005-10	WB	6.17 (42)	4.86
Infrastructure	Time required to get electricity in days	Time required (days)	2011	WB	133.40 (48)	109.72
Business efficiency/ competition	Time to obtain business license in days	Time required to obtain operating license (in days)	2005-10	WB	26.33 (40)	16.33
Business efficiency/ competition	Level of competition for internet, long distance fixed phone, & cell phone	Level of competition index for Internet services, international long distance services, & mobile telephone services on a 0 – 2(best) scale	2011	WEF	1.42 (34)	0.52

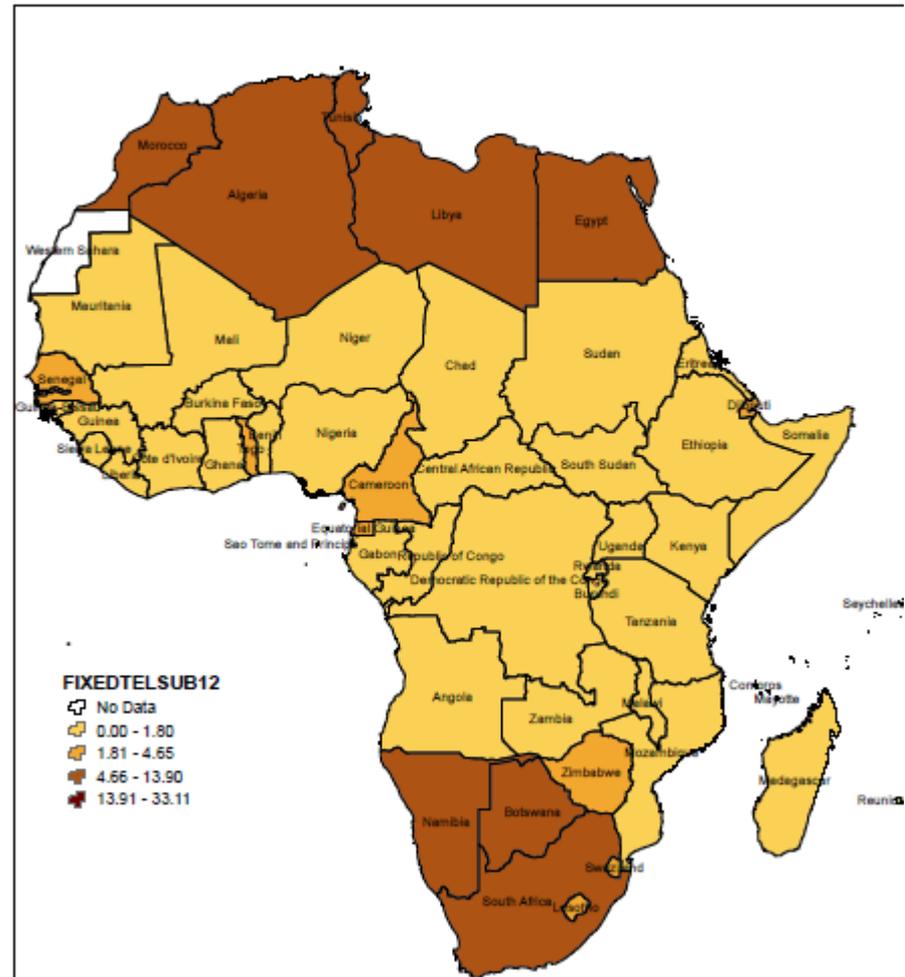
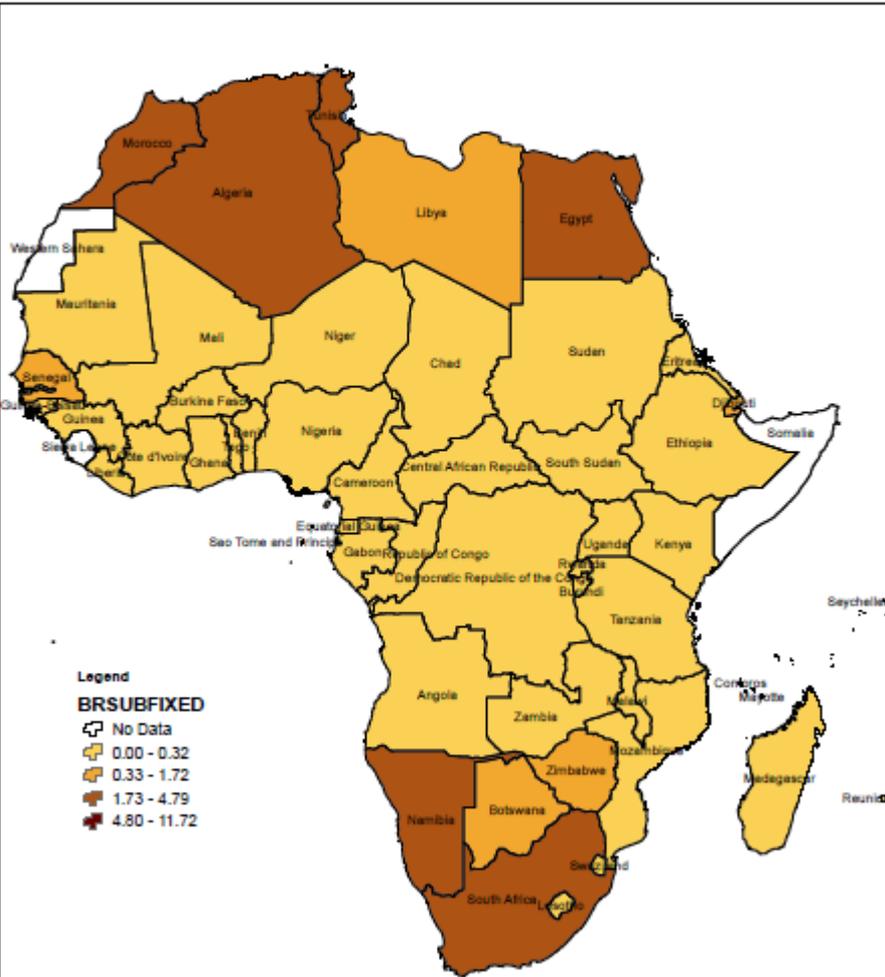
Methodology

1. Gather attributes for the factors in the conceptual model;
2. Test the factors for **multi-collinearity** and reject variables causing multi-collinearity;
3. Map the final set of variables to explore the rudimentary factors;
4. Apply local spatial autocorrelation and mapping of groupings from **k-means cluster analysis** to show clusters and outlier states based upon similarity of ICT dependent variables;
5. Apply **spatial autocorrelation methods** (Longley et al., 2011) on the dependent variables to test whether or not spatial agglomeration is present for high value states and low value states;
6. Apply **OLS stepwise regression** analysis for the full set of 50 nations to test for correlates as posited in the model;
7. Test the regressions for conformity to **OLS regression assumptions**;
8. Test for **spatial autocorrelation in the regression residuals**.

Maps of Dependent Variables

African Continent Countries
Fixed Wired-Broadband Subscriptions per 100 Persons

African Continent Countries
2012 Fixed Telephone Subscriptions per 100 Persons

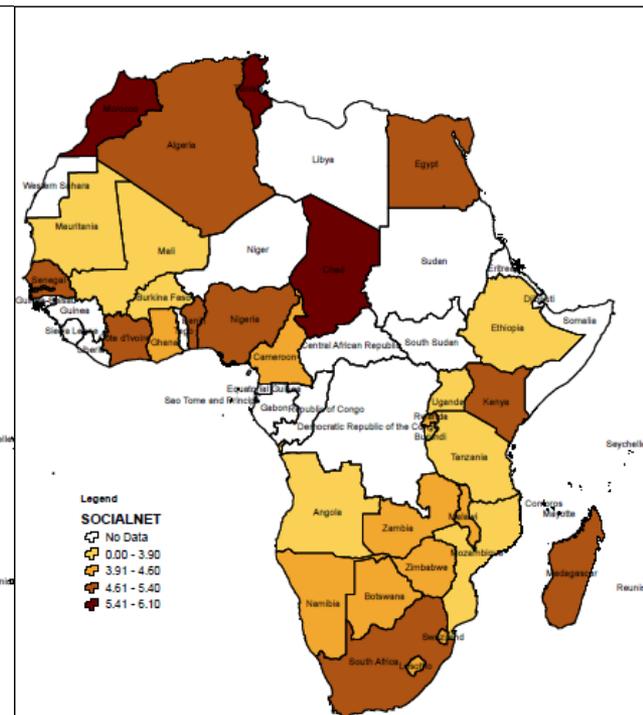
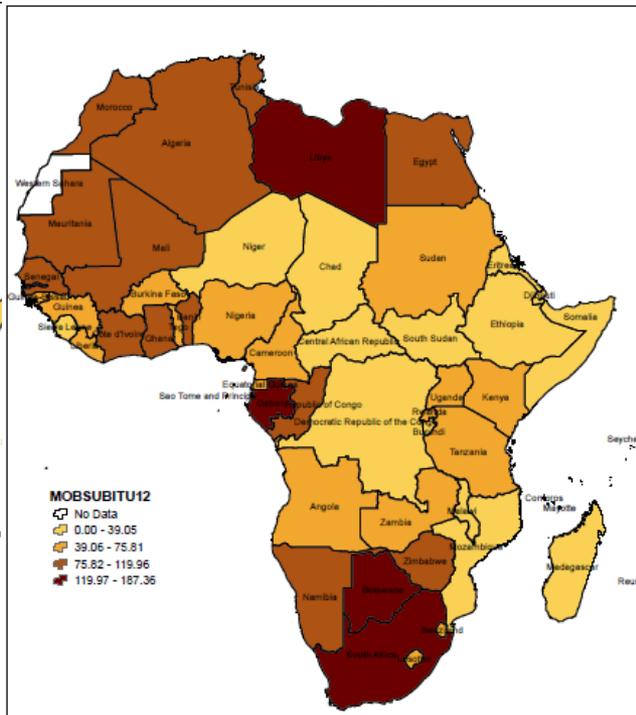
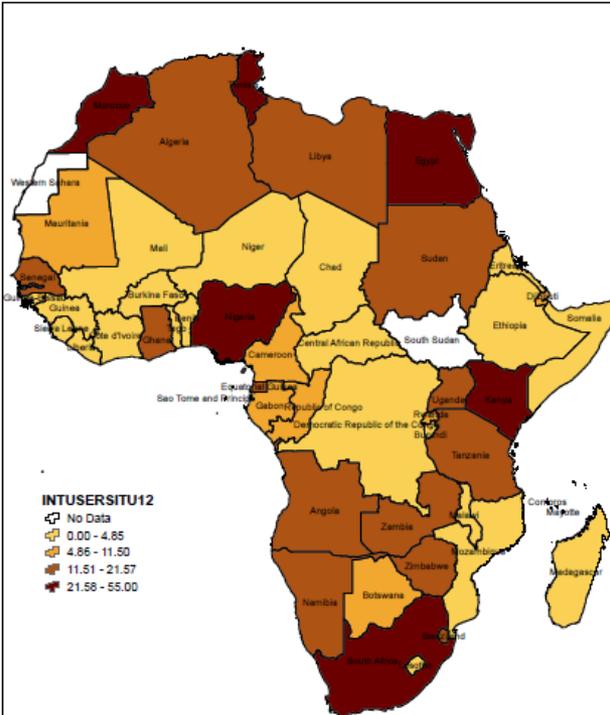


Maps of Dependent Variables

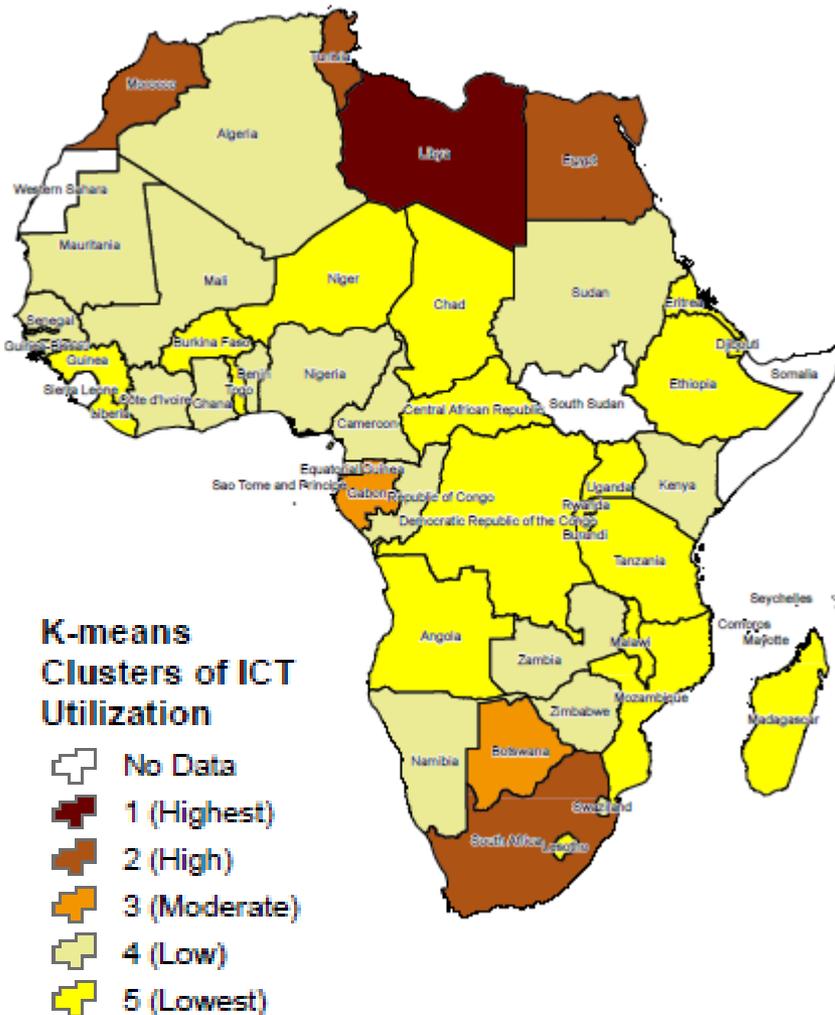
African Continent Countries
2012 Percent of Individuals Using Internet

African Continent Countries
2012 Mobile Telephone Subscriptions per 100 Persons

African Continent Countries
2010-11 Use of Virtual Social Networks



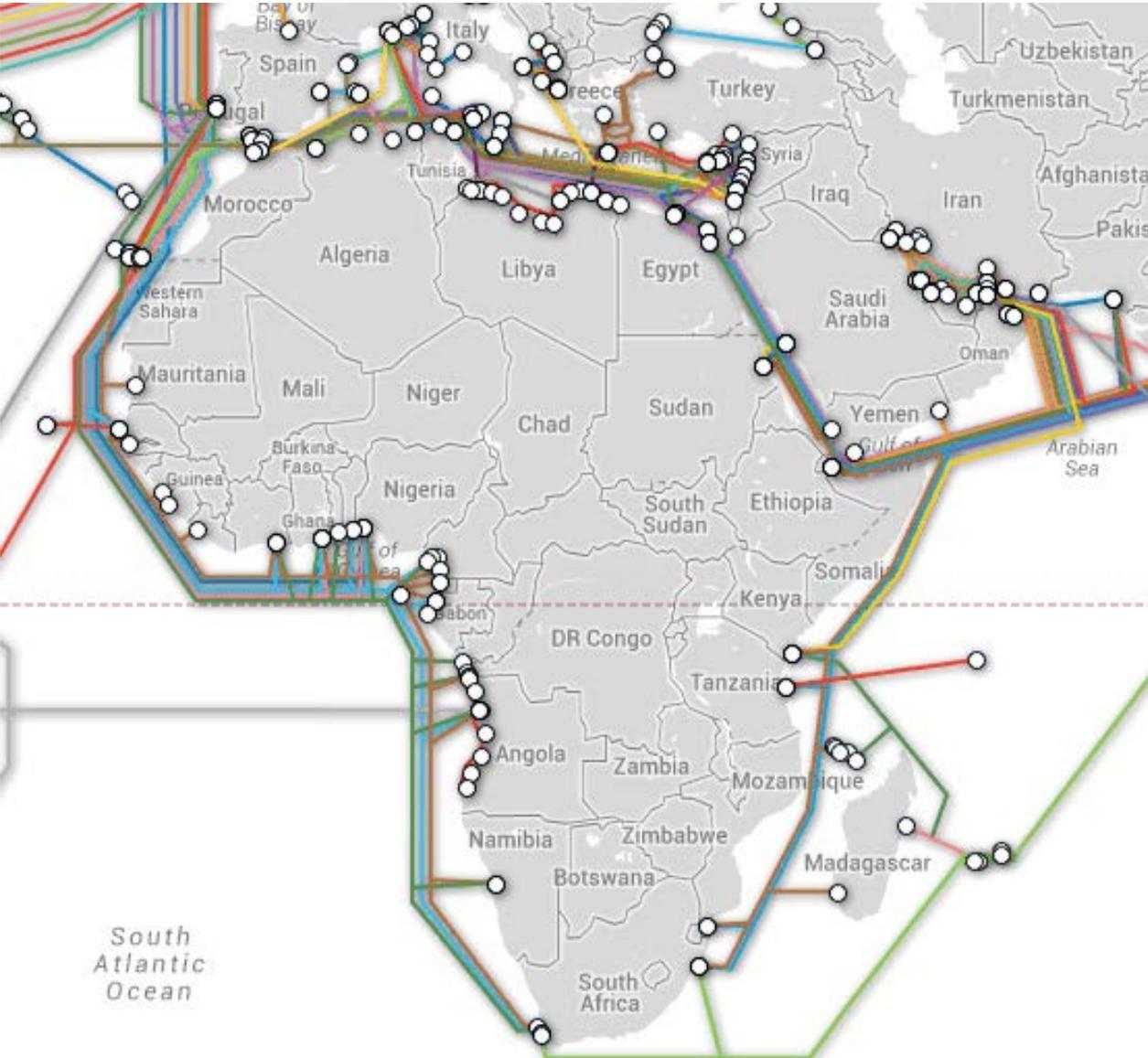
K-Means Cluster of ICT Utilization



- **Cluster 1:** Libya, Seychelles
- **Cluster 2:** Egypt, Mauritius, Morocco, South Africa, Tunisia
- **Cluster 3:** Botswana, Gabon
- **Cluster 4:** Algeria, Ghana, Kenya, Nigeria, Sudan, Zambia, Zimbabwe, etc.
- **Cluster 5:** Angola, Chad, Ethiopia, Niger, Rwanda, Uganda, etc.
- Strong geographic agglomeration, Tobler's Law (Longley et al, 2011)
- High-tech nations mostly coastal.

Submarine Cable Map

<http://www.submarinecablemap.com/#/>



Spatial Autocorrelation Patterns Measured by Moran's I

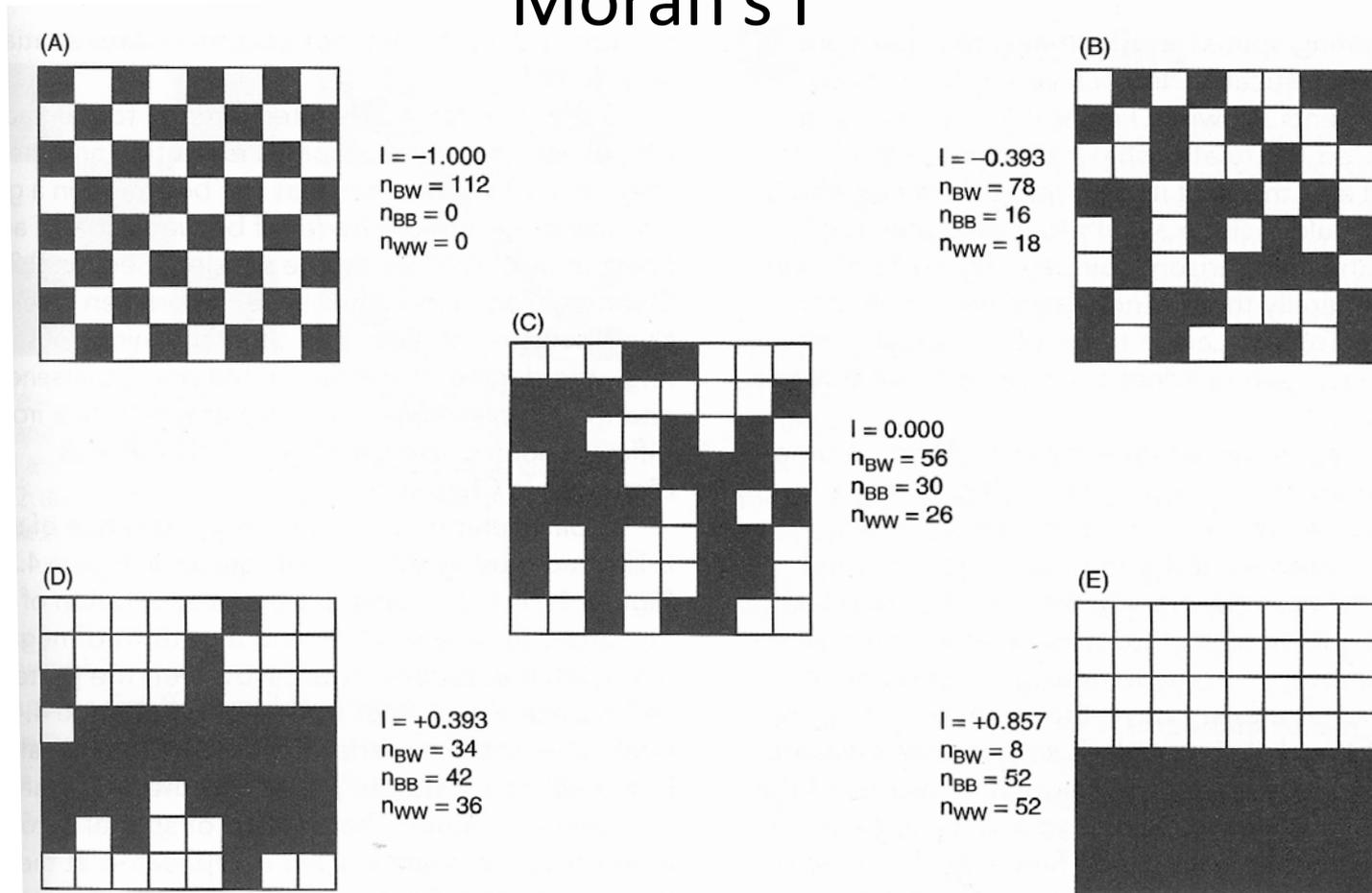


Figure 4.1 Field arrangements of blue and white cells exhibiting: (A) extreme negative spatial autocorrelation; (B) a dispersed arrangement; (C) spatial independence; (D) spatial clustering; and (E) extreme positive spatial autocorrelation. The values of the I statistic are calculated using the equation in Section 4.6. (Source: Goodchild 1986 CATMOG, GeoBooks, Norwich)

Source: Longley, P. et al. (2011). *Geographic Information Systems & Science*, Wiley, p. 103.

Spatial Autocorrelation Findings

<i>Moran's I values for African Nations, N=51, 2010</i>									
	Internet Users	Fixed-Broadband Internet Subscriptions		Mobile subscriptions	Fixed telephone subscriptions	Use of virtual social networks		AVERAGE	
	0.059	0.288***		0.277**	0.324***	0.001		0.223	
<i>Moran's I values for African Nations, N=44, 2010</i>									
	Internet Users	Fixed-Broadband Internet Subscriptions		Mobile subscriptions	Fixed telephone subscriptions	Use of virtual social networks		AVERAGE	
	0.034	-0.034		0.264**	0.133	-0.013		0.077	
<i>Moran's I values for United States, N=50, 2010</i>									
	Desktop/Laptop in Household	Log of Internet Access at Home	Broadband Adoption in Household	Persons in Cellphone-Only Household	Mobile Wireless High-Speed Devices	Persons in Fixed-phone-only Household	Facebook Users	Twitter Users	AVERAGE
	0.547***	0.471***	0.457***	0.621***	0.230***	0.648***	-0.004	0.069	0.381
<i>Moran's I values for China, N=31, 2009</i>									
PCs per 100 Urban Families	PCs per 100 Rural Families	Internet Users per 100 pop.	Broadband Subscribers per 100 pop.	Mobile Telephone Subscribers per 100 pop.	Fixed Phone Subscribers per Capita	No. of Domain Names per 100 pop.	No. of Web Pages per Capita	AVERAGE	
0.346***	0.206*	0.264**	0.272**	0.205*	0.252**	0.615***	0.086*	0.285	
<i>Moran's I values for China, N=29, Excluding Beijing and Shanghai, 2009</i>									
0.343***	0.239**	0.137	0.258**	0.061	0.143	0.236**	0.177*	0.199	

(Source for China, author, 2013a)

- Broadband, mobile, & fixed telephone are highly agglomerated.
 - Sharp contrast between high value agglomerations of the North African, Mediterranean nations plus southern tip ones, compared to the lower-valued agglomerated values in central sub-Saharan Africa.
- Internet users & virtual social network use are have random spatial patterns.

Standardized Regression Results: Full Sample (N=51)

- **ICT laws**: Most important correlate, consistent across 5 dependent variables.
 - Countries with well-developed laws for ICT were likely to have higher ICT levels.
- **Effectiveness of law-making bodies** reduced broadband subscriptions and fixed telephone subscriptions.
- **Economic factors** were influential selectively across all five dependent variables.
 - **Gross National Income per capita** – associated with broadband and fixed telephone subscriptions,
 - **Industrial and/or manufacturing output** – associated with internet users, mobile subscriptions, and social networks.
- Inverse association of **agricultural exports** on use of virtual networks.

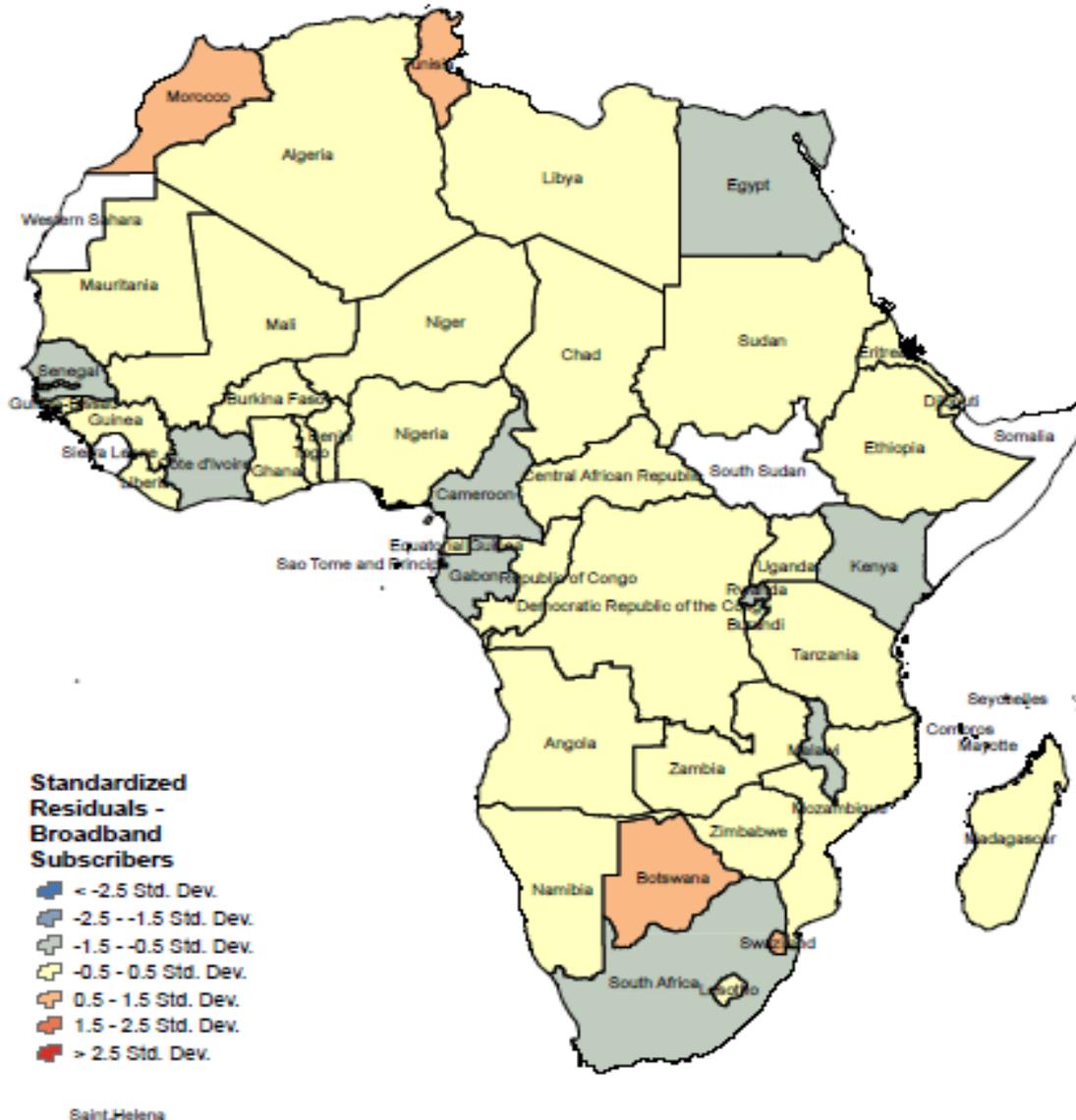
Standardized Regression Results: Lower Tech Sample (N=44)

- Economic categories dominate while societal openness is reduced.
- GNIPC, industrial output, manufacturing output are important correlates.
- Export of agricultural materials reduces use of virtual networks.
- Since ICT laws are much more developed in the seven “high tech” nations, eliminating them reduces ICT laws’ effect except for mobile subscriptions and social networks.
- With Urban included: urban, ICT laws, and manufacturing/industrial output account together for the effects.

Implications of Findings

- **Relatively high ICT countries** located in North Africa adjoining the Mediterranean and in the far south of the continent.
 - Greater industrial & urban development of these parts of the continent due to their proximity to ocean transport & earlier history of econ. dev.
- **Lowest ICT nations:** agglomerated in middle & middle south of Africa.
 - In 21st century, more of central Africa will industrialize, urbanize, especially since population is growing rapidly, a trend implying greater use of ICTs.
- **Internet & social network users:** spatial randomness relates to the greater evenness of their distribution throughout the continent.
 - Relatively inexpensive and hence not tied to the economic disparities of the continent, unlike *fixed broadband and fixed telephones*.
- **Regression results:**
 - At the low standard of living for most of Africa, education and training supplanted by more basic factors – income, urbanization, & industrialization.
 - Surprising that societal openness is highly influential, but less so if its recent influence at the national level of change in the Arab Spring and its often profound effects in advanced nations are considered.
 - Absence of Infrastructure (electricity) & business efficiency/competitiveness explained as less important in very basic and poor environments of most of Africa. Expected to grow in importance as Africa industrializes/urbanizes in this century.

Map of Regression Residuals



- The model’s weakness in incorporating spatial errors for certain dependent variables: stems from strong starting spatial autocorrelation of the variable for **broadband & fixed telephones**.
- Spatial errors diminish for the “lower technology” subsample.
- Measurement of spatial autocorrelation proves useful as a gauge of how much the regression independent factors account for spatial effects, and if they do not, then what external forces might account for the agglomeration of errors.
- Dangerous to trust solely regression for digital divide studies, since it ignores geographical proximity effects which often lead to having overconfidence in the findings.

Responses to Research Questions

1. *What is an appropriate conceptual model to understand the geography of, and influences on levels of technologies in African countries?*

From this single exploratory test, the conceptual model appears robust to account for highly significant regression findings.

2. *What is the spatial clustering of levels of technologies for African nations and how do the clusters differ?*

Cluster analysis demonstrates that there are distinctive clusters that are mostly spatially agglomerated, in concert with Tobler's Law.

3. *How agglomerated are the technology variables in Africa, based on spatial autocorrelation?*

About half of the dependent variables are spatially agglomerated. Reasons are given why internet users and virtual network users are spatially random.

4. *What are the social, economic, societal openness, infrastructure, and business factors that influence levels of technologies for the entire sample of African nations, and for a lower-tech subsample?*

The major categories of influences are economic, societal openness, and urban, followed by some technology cost effects. Surprisingly, literacy/training is not important, nor are infrastructure (electricity) and business efficiency/competitiveness.

5. *Does the regression model account for spatial agglomeration so the regression error terms are spatially random?*

The regression model does mostly account for spatial agglomeration, and explanations are offered for the instances where it does not.