Identifying Investment Opportunities In International Telecommunications Markets Using Regression Models

Joseph Callender
Bruce Golden
Shreevardhan Lele
Edward Wasil

ICS Conference
Chandler, Arizona
January 9, 2003

Introduction

- Background

Teledensity

Number of main telephone lines (copper lines) per 100 inhabitants

1995 Teledensity Data	
United States	62.57
Mexico	9.58

Source: *ITU World Telecommunications Development Report 1998*

Introduction

- Background

  Identify countries with teledensity levels that are *lower* than we might expect

  Underperforming countries are best candidates for investment

  Understand why a country has a low teledensity level

  National governments do not want to get left behind

Introduction

- Background

  1998 study by Herschel Shosteck Associates

  HSA built a simple linear regression model to estimate teledensity

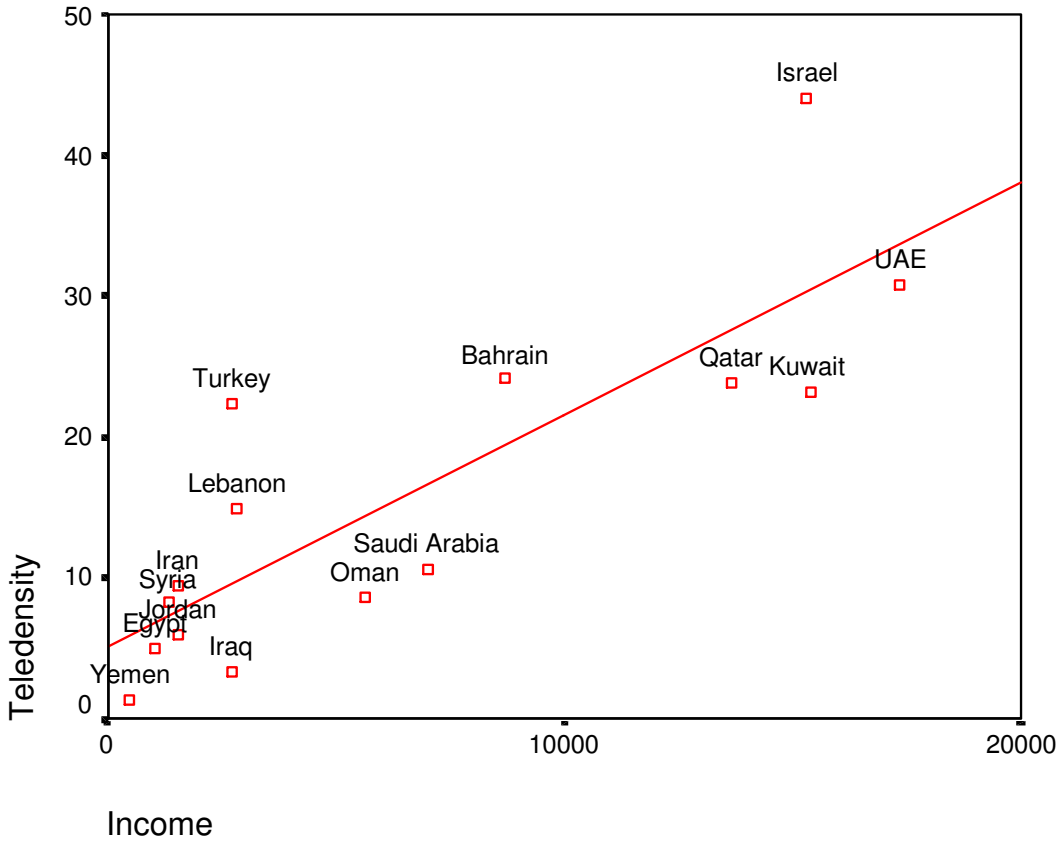
    Income (gross domestic product) was the independent variable

    15 countries from the Middle East

    GDP per capita explains 85% of a countries teledensity level

Introduction

- Background



Introduction

- Background

  Nine countries under the regression line have actual teledensity levels that are less than their predicted values

Yemen	Egypt
Jordan	Iraq
Oman	Saudi Arabia
Qatar	United Arab Emirates
Kuwait	

  Represent best opportunities for producing increased telephone utilization

Introduction

- Background

  HSA computed the residual value
  for each country

  Observed teledensity - Predicted teledensity

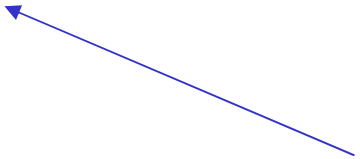
  Negative residual indicates underperformance
  with respect to telephone utilization

Introduction

- Background

Country	Residual
Kuwait	-7.6
Iraq	-6.2
Saudi Arabia	-5.9
Oman	-5.7
Yemen	-4.6
Qatar	-3.4
UAE	-2.6
Egypt	-1.8
Jordan	-1.6

Country	Residual
Syria	0.9
Iran	1.9
Bahrain	4.8
Lebanon	5.2
Turkey	12.8
Israel	14.0



All nine countries under the line have negative residuals

Introduction

- Overview

Extend HSA study in three ways

Model 208 countries using regression

Expand the set of covariates

- Telecommunications
- Socioeconomic
- Political

Conduct a benchmarking exercise

- Identify peer group for each country
- Three partitioning methods
- Aggregate results to rank countries

Variables

- Database
  - HSA database
    - 29 variables for 208 countries
    - Telecommunications data
      - *World Telecommunications Development Report*, ITU, 1998
    - Socioeconomic data
      - *World Development Report* World Bank, 1998
    - Political data
      - *1998 Index of Economic Freedom* Heritage Foundation

Variables

- Independent variables

With HSA input, identified 12 independent variables that are candidate regressors

Variable	Definition
Income	Gross domestic product per capita
Connection Charge	Cost to install a telephone line
Income Growth	Compounded growth of GDP per capita from 1990 to 1995
Main Line Growth	Compounded growth of the number of main lines from 1990 to 1995
Main Lines per Employee	Number of main lines divided by the number of telecom employees

Variables

- Independent variables

Variable	Definition
Outgoing Telephone Traffic	Minutes of international telephone traffic
Traffic Growth	Compounded growth of outgoing traffic from 1990 to 1995
Percent Residential	Proportion of telephones in residential use
Political Risk	Measures the economic freedom in a country
Ratio	Ratio of teledensity in the largest city to that in the entire country
Telecom Revenue per Main	Revenue generated by each telephone line
Percent Waiting	Percent of people on a wait list for a telephone

Variables

- Dependent variable

Variable	Definition
Teledensity	Number of main lines per 100 inhabitants of a country Most commonly used measure of the prevalence of telephone lines in a country Reasonable proxy for the size of a telephone market

Methodology

- Value approach

  Identify countries that are underperformers

  Assess levels of the covariates linked to performance

  Use cross-sectional data

  Actual performance is below estimated potential performance

Methodology

- Naïve estimation approach

  Construct a single regression model

  Relate teledensity to 12 covariates for all 208 countries lumped together

  Model may not be robust for a group of similar countries

  Ignores particularities of a group that affect performance

Methodology

- Customized regression models

  Compare a country against appropriate peer groups

  Elaborate criteria may lead to very small peer groups

  Including too many variates for a peer group over fits the model

Methodology

- Constructing peer groups: Three methods

1. Assemble all countries in one group

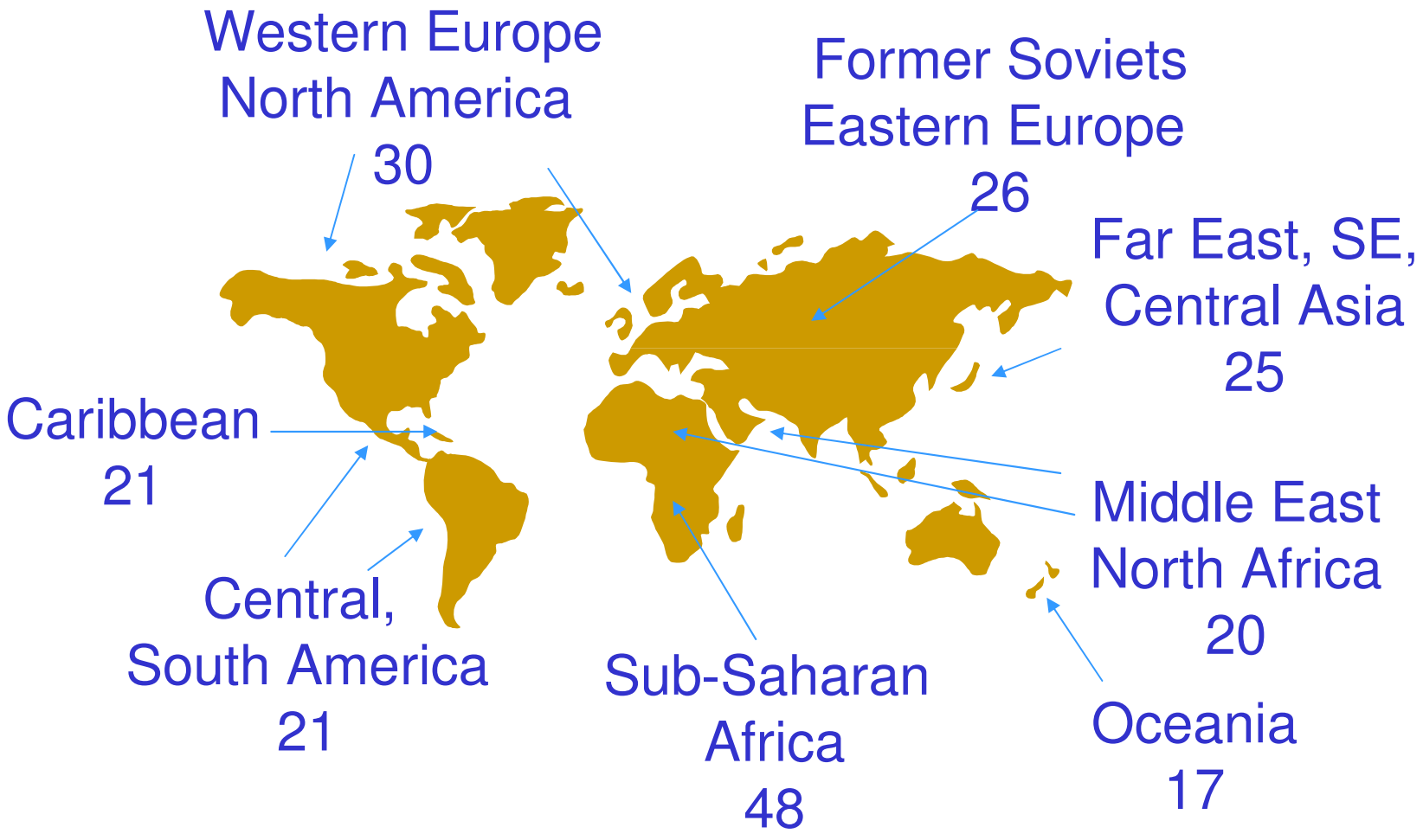
2. Partition on the basis of geography

   Countries in a geographic group likely to have similar economic, social, and political dimensions

   Eight geographic regions

Methodology

- Constructing peer groups: Geography



Methodology

- Constructing peer groups: Three methods

3. Form clusters using *k*-means algorithm

Reduced 12 correlated variables
to four orthogonal factors
- Income
- Industry characteristics
- National growth
- Telecommunication cost

Five clusters emerged

Methodology

- Constructing peer groups: Three methods

3. Form clusters using *k*-means algorithm

	Cluster				
	1	2	3	4	5
Number of Countries	16	30	58	47	57

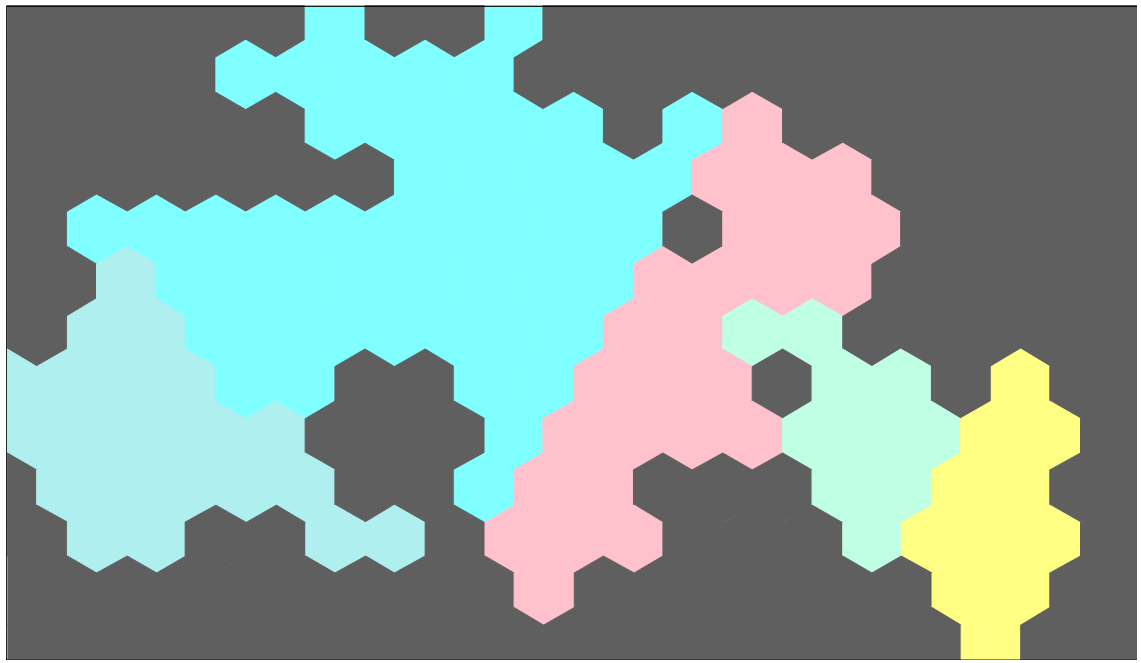
Countries in Cluster 1

Argentina	Croatia	Japan	Slovenia
Bangladesh	Finland	Mexico	Macedonia
Brazil	Greece	Paraguay	Uruguay
Columbia	Iran	Peru	Yemen

Methodology

- Constructing peer groups: Three methods

Number of clusters verified with Viscovery



Methodology

- Selecting covariates

  Three regression models

  1. Simple model of teledensity on income
  2. Quadratic model with income and income squared
  3. Stepwise model using all 12 covariates

Methodology

- Summary

14 peer groups

1 worldwide, 8 geographic, 5 clusters

3 regression models

Simple, quadratic, stepwise

Each country is in 9 different regression models

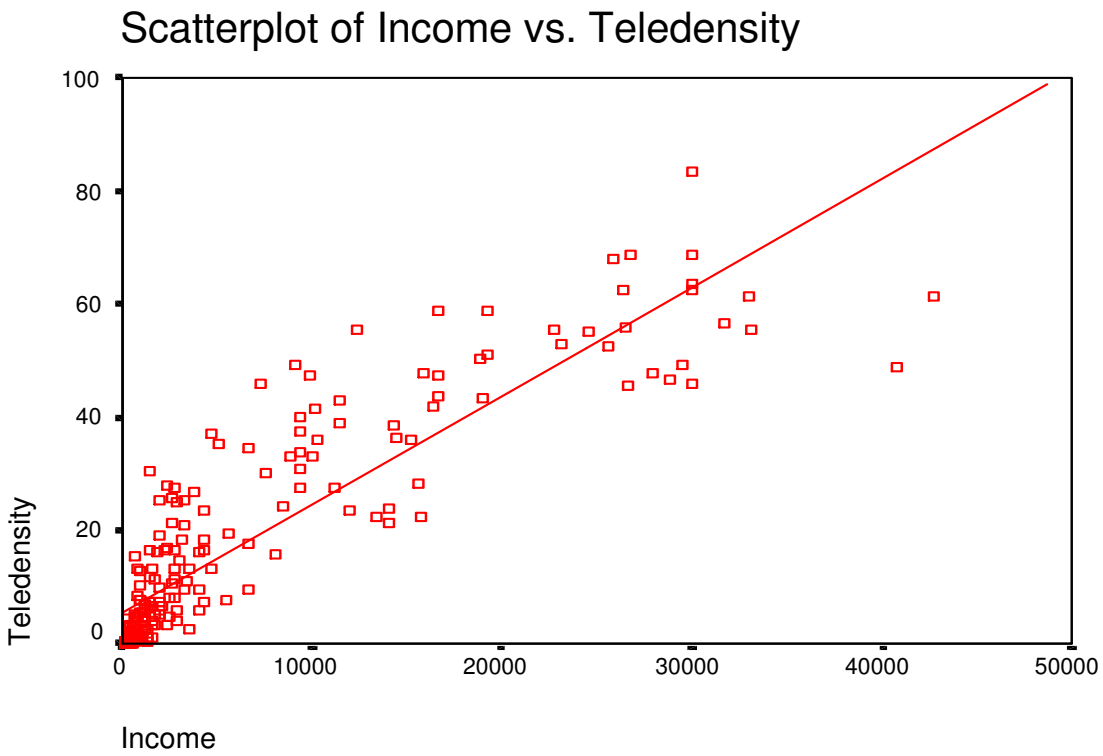
9 residual values

Modeling Results

- Worldwide models

Linear

5.59 + 0.00192 Income

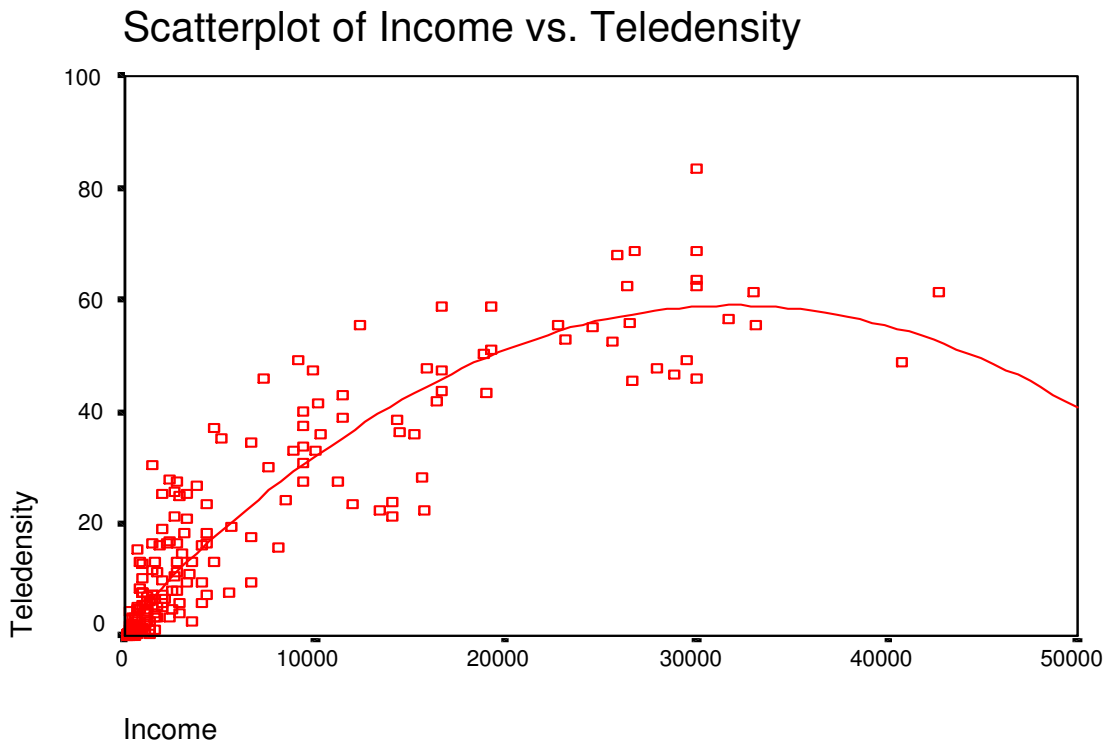


Modeling Results

- Worldwide models

Quadratic

$$1.99 + 0.00358\ \text{Income} - 5.61\text{E-}8\ \text{Income}^2$$



Modeling Results

- Worldwide models

  Stepwise

  3.00 + 0.00101 Income + 0.03 Traffic
  + 0.08 Main Lines per Employee
  + 0.60 Percent Waiting
  + 0.11 Percent Residential
  - 0.25 Main Line Growth
  - 2.34 Political Risk

  All three models are statistically useful

  p-values < 0.05

Modeling Results

- Worldwide models: Country performance

Standardized residuals

Model	Austria	Brazil	Gabon	Japan
Linear	-1.561	-0.687	-1.090	-3.826
Quadratic	-1.589	-1.160	-1.534	-0.770
Stepwise	-1.337	-1.253	-0.749	-2.952

All residuals < -1
Consistent *underperformer*

Modeling Results

- Geographic models

  Three models × Eight geographic groups

  Each country appears in three models that correspond to its group

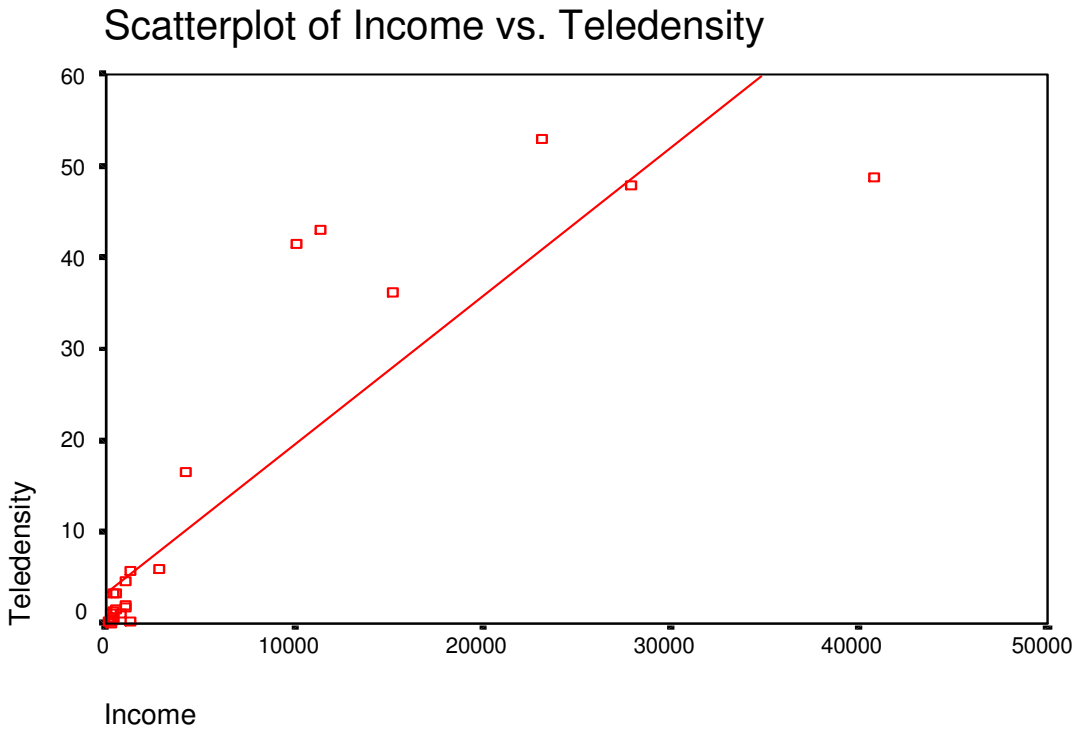
  Results for the Far East, Southeast, and Central Asia group

Modeling Results

- Geographic models: Asia

Linear

3.40 + 0.00162 Income

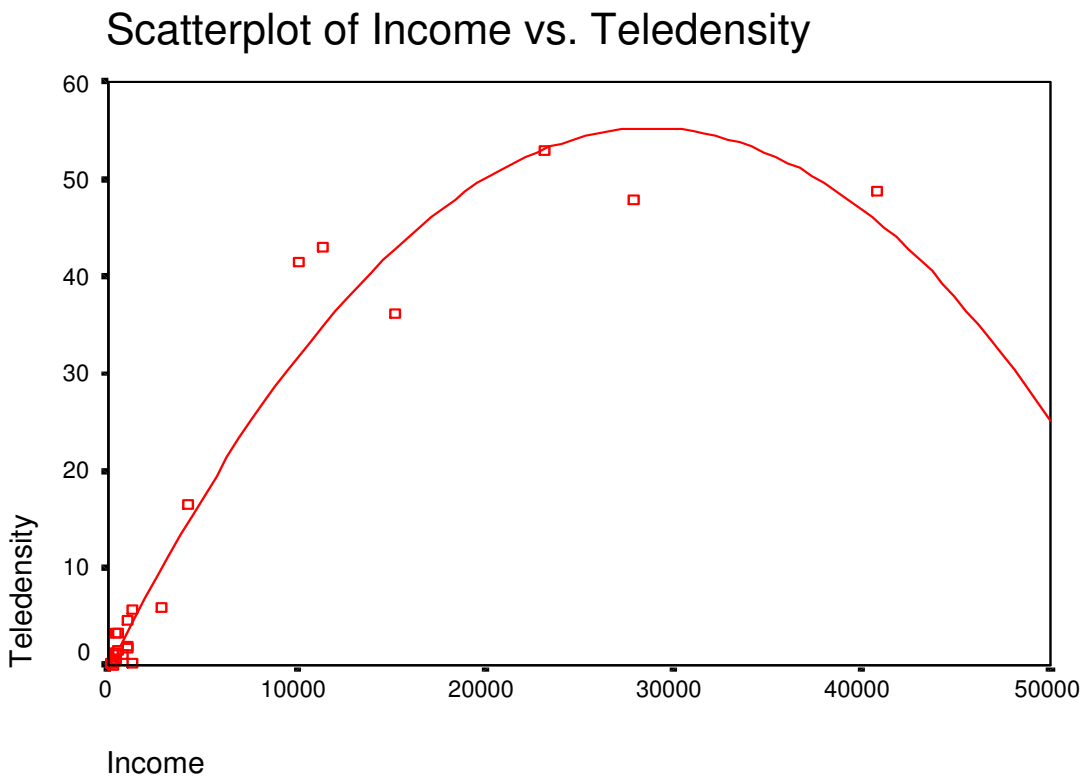


Modeling Results

- Geographic models: Asia

Quadratic

$$-0.32 + 0.00385\ \text{Income} - 6.69\text{E-}8\ \text{Income}^2$$



Modeling Results

- Geographic models: Asia

  Stepwise

  -2.00 + 0.000557 Income + 0.075 Traffic
  + 0.11 Main Lines per Employee

  All three models are statistically useful

  p-values < 0.05

Modeling Results

- Geographic models: Asia

Countries in 4 different groups

Standardized residuals

Model	Austria	Brazil	Gabon	Japan
Linear	-1.379	-1.374	-4.260	-2.359
Quadratic	-1.476	-1.291	-3.557	+0.826
Stepwise	-0.980	+0.534	-3.823	-0.509

All residuals << -1
Highly *underperforming* when compared to the 47 countries in Sub-Saharan Africa group

Modeling Results

- Cluster models

  208 countries placed in five clusters using factor analysis and *k*-means analysis

  Three models × Five clusters

  Each country appears in three models that correspond to its clusters

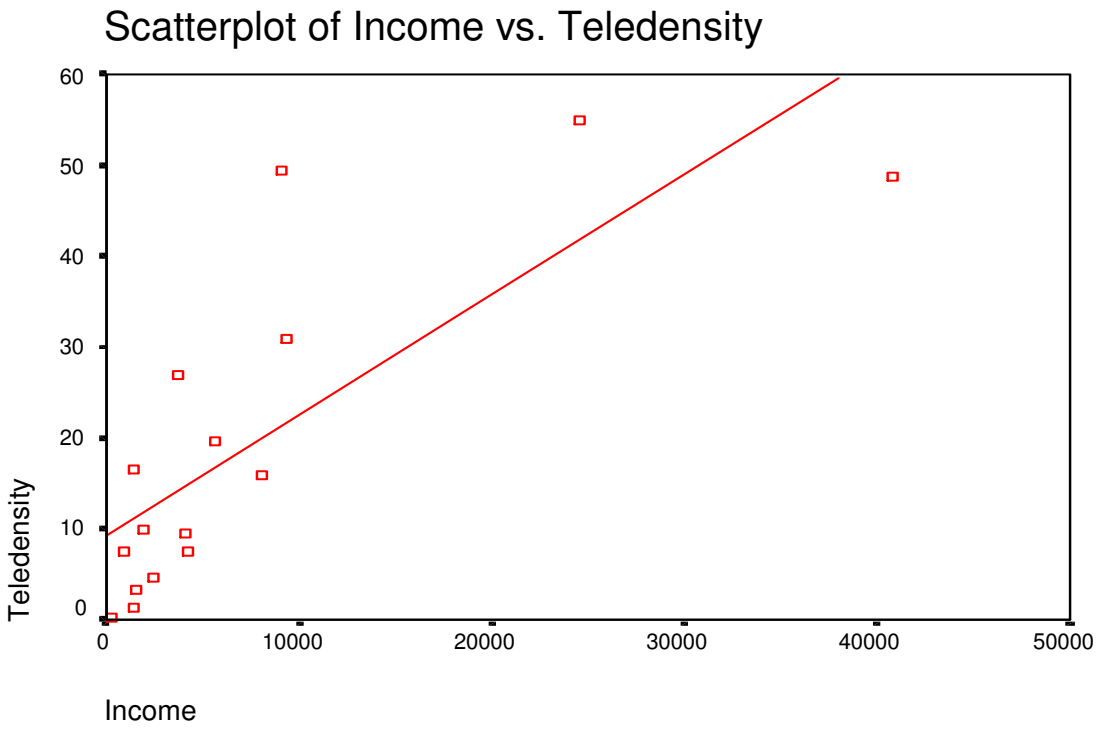
  Results for Cluster 1(Brazil, Japan)

Modeling Results

- Cluster models: Cluster 1

Linear

9.30 + 0.00133 Income

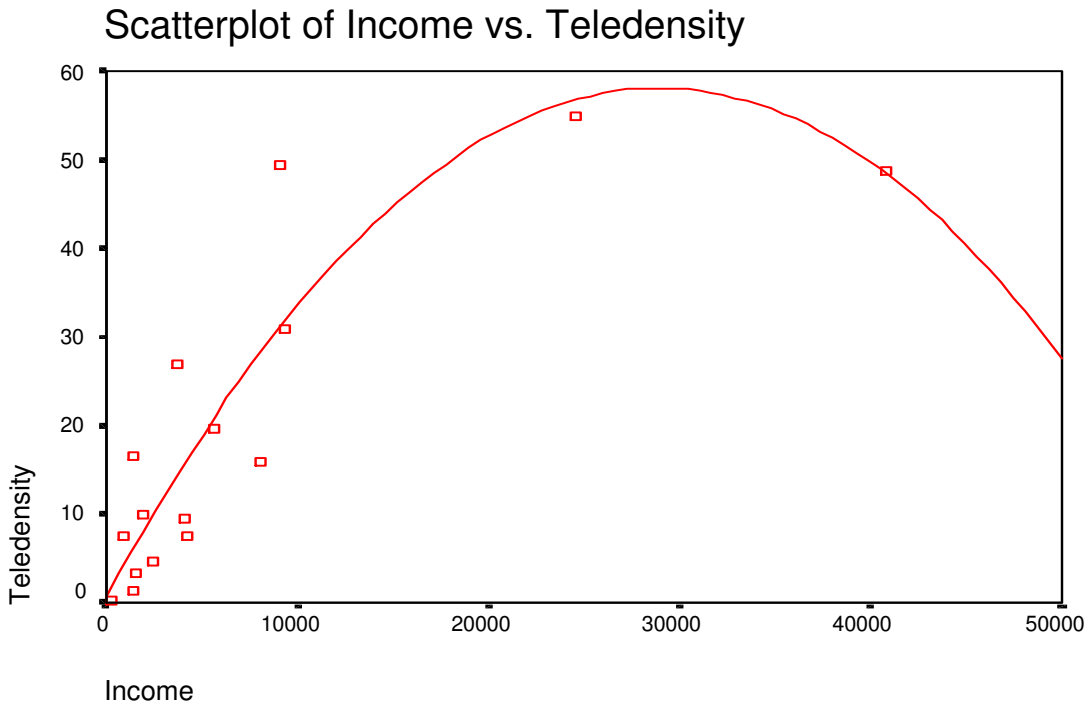


Modeling Results

- Cluster models: Cluster 1

Quadratic

$$0.74 + 0.00397\ \text{Income} - 6.86\text{E-}8\ \text{Income}^2$$



Modeling Results

- Cluster models: Cluster 1

  Stepwise

  2.83 + 0.000955 Income + 0.52 Traffic

  All three models are statistically useful

  p-values < 0.05

Modeling Results

- Cluster models: Cluster 1

Countries in 3 different clusters

Standardized residuals

Model	Austria	Brazil	Gabon	Japan
Linear	-1.099	-0.647	-4.483	-1.264
Quadratic	-1.113	-1.049	-4.181	+0.043
Stepwise	-1.095	-0.065	-4.097	+0.032

Consistently *underperforms* in worldwide, geographic, and cluster models

All residuals << -1
Highly *underperforming* when compared to the 57 countries in Cluster 3

Analyzing Results

- Combine results from different models: First cut

  Each country appears in nine models

  Focus on 63 countries that have at least one standardized residual less than -1

  Average standardized residual

  Weighted average standardized residual

  Number of standardized residuals less than -1

Analyzing Results

- Combine results from different models: Gabon

Model		(1) Adjusted R^2	(2) Stand. Residual	(1)×(2)
Worldwide	Linear	0.791	-1.090	-0.862
	Quadratic	0.858	-1.534	-1.316
	Stepwise	0.882	-0.749	-0.661
Sub-Saharan Africa	Linear	0.884	-4.260	-3.766
	Quadratic	0.921	-3.557	-3.276
	Stepwise	0.948	-3.823	-3.624
Cluster 3	Linear	0.907	-4.483	-4.066
	Quadratic	0.907	-4.181	-3.792
	Stepwise	0.934	-4.079	-3.827
Total		8.032	-27.774	-25.190

Analyzing Results

- Combine results from different models: Gabon

Average standardized residual

$U_{Gabon} = -27.774/9 = -3.086$

Weighted average standardized residual

$W_{Gabon} = -25.190/8.032 = -3.136$

Number of residuals less than -1

$N_{Gabon} = 8$

Analyzing Results

- Combine results from different models: Second cut

  Focus on countries that satisfy at least one of the following three criteria

  - Average standardized residual < -1
  - Weighted average standardized residual < -1
  - At least five standardized residuals less than -1

Analyzing Results

- Combine results from different models: Second cut

Country	W_i	U_i	N_i
Gabon	-3.14	-3.09	8
Kuwait	-1.71	-1.77	7
Botswana	-1.55	-1.53	7
French Polynesia	-1.47	-1.50	7
Saudi Arabia	-1.32	-1.42	6
Austria	-1.31	-1.29	8
Qatar	-1.20	-1.26	5
Singapore	-1.18	-1.12	5
Guam	-1.16	-1.15	6
Iraq	-1.15	-1.18	5
Japan	-1.14	-1.20	4
Belgium	-1.11	-1.12	5
Luxembourg	-1.01	-0.87	3
Marshall Islands	-0.99	-1.02	4
Argentina	-0.97	-0.99	6
Oman	-0.86	-1.00	3
Brazil	-0.76	-0.78	5

- First seven countries maintain ordering based on W_i and U_i

- Eleven countries satisfy all three criteria – good candidates for investment

Conclusions

- Summary

  Partitioned 208 countries into three groups for benchmarking purposes

  Large negative standardized residual indicated underperformance relative to peers

  Aggregated results from nine regression models

  Identified a candidate list of potential investment opportunities

  Combine results with domain-specific knowledge (missing covariates?) to evaluate countries