Purchase patterns, socioeconomic status, and political inclination

Xiaowen Dong *1  Eaman Jahani *1  Alfredo J. Morales 1,2
Burçin Bozkaya 3  Bruno Lepri 1,4  Alex ‘Sandy’ Pentland 1

* (equal contribution)

1 MIT Media Lab  2 New England Complex Systems Institute
3 Sabanci University  4 Fondazione Bruno Kessler

Annual Bank Conference on Development Economics 2016
Data and Development Economics
Motivation

• Large-scale behavioral data collected via pervasive technologies

- 5B mobile phone users by 2019 [Statista]
- 400M users’ WiFi locations resolved every day by Skyhook Wireless (2013)
- 4B credit cards in use in 2013 with $2.2 trillion spent [Nilson reports]
Motivation

• Such quantitative data provide novel insight into understanding ourselves and human/social behavior
Motivation

• Such quantitative data provide novel insight into understanding ourselves and human/social behavior

[deMontjoye13]
Motivation

• Such quantitative data provide novel insight into understanding ourselves and human/social behavior

[deMontjoye13]

[Krumme13]
Motivation

• Such quantitative data provide novel insight into understanding ourselves and human/social behavior

[deMontjoye13]

[Krumme13]

[Pentland14]
Motivation

- Understanding human/social behavior helps understand economic development
Motivation

- Understanding human/social behavior helps understand economic development

- Previous studies have shown link between diversity of communication/mobility and economic prosperity using phone records [Eagle10, Smith13]
Motivation

• Understanding human/social behavior helps understand economic development

• Previous studies have shown link between diversity of communication/mobility and economic prosperity using phone records [Eagle10, Smith13]

• **Objective of this study**: Use large-scale bank transaction data to
  - strengthen such link from economic perspective
  - understand individual economic behavior with sociopolitical implications
  - provide policy recommendation
Data and method

- Two major financial institutions in Europe and Latin America
- Credit card transaction records
  - European: >10M records about >100K customers at >100K stores in a few months
  - Latin American: >60M records about >3M customers at >400K stores in a few months
Data and method

- Two diversity patterns in purchase behavior
Data and method

- Two diversity patterns in purchase behavior

**Outgoing purchase diversity:** Level of exploration

Diversity as Shannon entropy of customer $i$’s merchant visits:

$$H(i) = - \sum_{j=1}^{N} p_{ij} \log(p_{ij})$$
Data and method

• Two diversity patterns in purchase behavior

**Outgoing purchase diversity:** Level of exploration

**Incoming purchase diversity:** Level of attractiveness
Data and method

• Two diversity patterns in purchase behavior

**Outgoing purchase diversity:** Level of exploration

**Incoming purchase diversity:** Level of attractiveness

Physical mixing
Access to information
Idea exchange
Data and method

• We would like to investigate link between purchase diversity patterns and important social indicators
Data and method

• We would like to investigate link between purchase diversity patterns and important social indicators

• Socioeconomic status
  - European: Index (0-100) computed using distribution of population of different classes
  - Latin American: Categorical variable (1-5) provided by national statistics office
Data and method

• We would like to investigate link between purchase diversity patterns and important social indicators

• Socioeconomic status
  - European: Index (0-100) computed using distribution of population of different classes
  - Latin American: Categorical variable (1-5) provided by national statistics office

• Political inclination
  - Computed based on percentage of votes for parties labeled as “liberal” or “conservative” in recent national elections
  - An approximation of the political inclination of each district
  - Interpretation of liberal/conservative labels are not consistent across countries
Visualization: Latin American country

Purchase diversity (outgoing)

Socioeconomic status

$r = 0.53$
Visualization: Latin American country

Purchase diversity (outgoing) $r = 0.53$

Socioeconomic status $r = 0.26$

Political conservatism
Quantitative results

<table>
<thead>
<tr>
<th></th>
<th>Socioeconomic Status</th>
<th>Political Conservatism</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>European:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase Diversity (outgoing)</td>
<td>0.77</td>
<td>-0.55</td>
</tr>
<tr>
<td>Purchase Diversity (incoming)</td>
<td>0.35</td>
<td>-0.25</td>
</tr>
<tr>
<td><strong>Latin Amer.:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Purchase Diversity (outgoing)</td>
<td>0.53</td>
<td>0.26</td>
</tr>
<tr>
<td>Purchase Diversity (incoming)</td>
<td>0.49</td>
<td>0.33</td>
</tr>
</tbody>
</table>

All correlations statistically significant with $p$-value < 0.001

- Diversity measures are positively correlated with socioeconomic status
- No universal relationship between diversity measures and political inclination
Significance and application

• Economic link underlying mobility diversity and economic development
Significance and application

• Economic link underlying mobility diversity and economic development

• Political inclination might have different economic characteristics in different context
Significance and application

• Economic link underlying mobility diversity and economic development

• Political inclination might have different economic characteristics in different context

• Real-time monitoring and prediction of such indexes from purchase behavior
  - Higher spatial and temporal resolutions
  - Less costly than census/survey
Significance and application

• Economic link underlying mobility diversity and economic development

• Political inclination might have different economic characteristics in different context

• Real-time monitoring and prediction of such indexes from purchase behavior
  - Higher spatial and temporal resolutions
  - Less costly than census/survey

• Application:
  - Prediction of socioeconomic status
  - Approximation of segregation patterns in socioeconomic status
Prediction of socioeconomic status
Approximation of segregation patterns

Consider a travel-time network: Level of social exposure
Approximation of segregation patterns

Consider a travel-time network: Level of social exposure

Segregation in networks measured by assortativity [Newman02]
Approximation of segregation patterns

Consider a travel-time network: Level of social exposure

Segregation in networks measured by assortativity [Newman02]

Assortativity in socioeconomic status can be approximated by assortativity in purchase diversity
Approximation of segregation patterns

Consider a travel-time network: Level of social exposure

Segregation in networks measured by assortativity [Newman02]

Assortativity in socioeconomic status can be approximated by assortativity in purchase diversity

Socioeconomic status  
Purchase diversity

\[ r = 0.52 \]
\[ p < 0.001 \]

\[ r = 0.57 \]
\[ p < 0.001 \]
Discussion

• Bank transaction records constitute a new and complementary data source to understand human economic behavior
Discussion

• Bank transaction records constitute a new and complementary data source to understand human economic behavior

• Investigate causal relation as next step
Discussion

• Bank transaction records constitute a new and complementary data source to understand human economic behavior

• Investigate causal relation as next step

• Diversity (mixing level) is generally associated with economic prosperity

• Increasing mixing level may promote economic development. Some empirical evidences:
  - La Línea Verde in Mexico
  - Riyadh Metro
  - Trans-Sahara Highway
La línea Verde in Mexico

In Mexico, a city’s scar becomes its most prized park, La Línea Verde

ANA ARANA  JANUARY 20, 2014

Aguascalientes turned a scrubby strip of pipeline land into a 12-km linear park that cuts through 90 neighborhoods. Photo Ags Prensa

Riyadh Metro
Trans-Sahara Highway

Economic integration is helping boost trade and investment in Africa

While Europe is on the verge of breaking up, Africa is reaping the benefits of integrating, growing and developing its trading blocks