Cell phone data

Martina Kirchberger

Trinity College Dublin and CEPR

STEG Virtual Course

24 May 2024

Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

Introduction: cell phone data in economics

Cell phone data are a typical example for ‘big data’ used in economics research.

Cell phone data: data on the movement patterns of mobile phones.

Allow studying phenomena previously difficult or impossible to measure.

Unique: spatio-temporal resolution and sample size.

Examples for recent papers using cell phone data

- Migration (Blumenstock, Chi, and Tan, 2023)
- Commuting (Kreindler and Miyauchi, 2021)
- High-frequency mobility (Blanchard, Gollin, and Kirchberger, 2023)
- Transport networks (Kreindler, Gaduh, Graff, Hanna, and Olken, 2023)
- Adoption of network goods (Björkegren, 2019)
- Targeting of humanitarian aid (Aiken, Bellue, Karlan, Udry, and Blumenstock, 2022)
- Firm behavior (Blumenstock, Ghani, Herskowitz, Kapstein, Scherer, and Toomet, 2020)
- Social segregation (Athey, Ferguson, Gentzkow, and Schmidt, 2021)
- Face-to-face interactions (Atkin, Chen, and Popov, 2022)

Example: measuring migration

Survey questions on permanent migration:

- How long have you been living continuously in ...?
- Just before you moved here, did you live in a city, in a town, or in a rural area?
- Before you moved here, which [PROVINCE/REGION/STATE] did you live in?

Survey questions on temporary migration:

- Have you spent between one and six months away from the village for work within the past year?

Cell phone data:

- Continuous information on location of devices.
- Can define migration at arbitrary scales.

Cell phone data

Open up entirely new possibilities: measurement and identification.

Distinct challenges compared to traditional data:

1. **Definition of new metrics**: not straightforward
2. **Selection concerns**: ownership of mobile phone, (selective) usage of mobile phone
3. **Distinct sources of measurement error**: coverage, data/call charges, sharing of phones, access to electricity to charge phone
4. **Lack of covariates**: makes understanding sample selection harder, identification, can observe location not intent
5. **Processing and visualization**: computational demands and different methods

1. Data structure
2. Selection and measurement error
3. Blumenstock, Chi and Tan (2023): Migration and networks
4. Blanchard, Gollin and Kirchberger (2023): High-frequency mobility
5. Resources and how to get started

What are cell phone data?

Part of a broader class of digital trace data.

Digital trace data: data that are "produced as the result of people's ordinary activities that leave behind a digital footprint, rather than data produced specifically for the purpose of scientific study" (Chi, Lin, Chi, and Blumenstock, 2020).

Discuss two broad types of cell-phone data:

1. Smartphone app location data: from one or more apps
2. Call detail records data: typically from one mobile phone operator

Illustrative smartphone app location data

deviceID	timestamp	longitude	latitude
id1	2018-03-24 06:00:00	3.34...	6.68...
id1	2018-03-24 06:10:00	3.34...	6.68...
id1	2018-03-24 06:12:00	3.34...	6.68...
id1	2018-03-24 06:13:00	3.34...	6.68...
id1	2018-03-24 12:40:00	3.19...	6.46...
id1	2018-03-24 12:41:00	3.19...	6.46...
id2	2018-03-24 18:21:00	3.41...	6.61...
id2	2018-03-24 18:45:00	3.64...	8.83...
id2	2018-03-24 23:21:00	3.34...	6.68...
...	...	...	...

Illustrative CDR data

Typically two components: 1. Call (and text) events

userID1	userID2	timestamp	tower1	tower2
id1	id2	2018-03-24 06:00:00	23	24
id1	id684	2018-03-24 06:10:00	23	67
id1	id45	2018-03-24 06:12:00	23	267
id1	id72	2018-03-24 06:13:00	23	45
id1	id72	2018-03-24 12:40:00	78	47
id1	id540	2018-03-24 12:41:00	78	36
id2	id67	2018-03-24 18:21:00	294	111
id2	id34	2018-03-24 18:45:00	287	35
id2	id1	2018-03-24 23:21:00	278	23
...	...	...	...	...

Illustrative CDR data (cont.)

Typically two components: 2. Tower location file

tower	longitude	latitude
1	3.34...	6.68...
2	3.19...	6.46...
3	3.64...	8.83...
4	3.67...	6.87...
5	3.19...	6.46...
...	...	...

One feature is that towers not equally spaced.

Construct map using Voronoi cells: each cell contains area closest to the point.

Note: These are also known as Voronoi tesselations or Thiessen polygons.

Example of map of cell phone towers

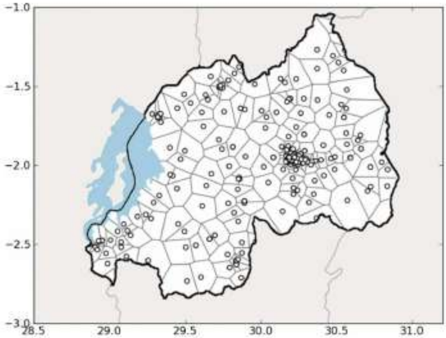


Figure 1. Map of Rwandan cell phone towers, January 2008. The median area covered by each tower is roughly 70 km^2.

Source: Blumenstock (2012).

Why do this? Important to understand variation in precision; convenient to link with other data; also indicates approximate area of coverage.

Call detail records data

Some key features of typical CDR data:

- Specific provider: consider market structure
- Data point typically linked to "event"
- No data on handovers
- Does not capture calling/messaging via apps
- Sample: any mobile phone
- Spatial detail: typically tower level (displacement?)

Note that datasets might also come at a tower-level, have regular pings, or include top-ups and call durations.

Smartphone app location data

Some key features of typical smartphone app location data:

- Linked to device rather than a phone number
- Captures calling/messaging via apps
- No information on specific apps
- Relies on data use (expensive)
- Sample: mobile phone with internet capacity
- Spatial detail: approximate GPS location

For differences between CDR and smartphone app data see also discussions in Blanchard et al. (2023) and Kirchberger (2021).

Sample selection

Different from traditional survey data.

Typically no covariates, with some exceptions.

Blumenstock, Gillick, and Eagle (2010) conduct a phone survey with 901 users in Rwanda: users are richer, older, more educated, more likely to be male.

Barwick, Liu, Patacchini, and Wu (2023) and Büchel, Ehrlich, Puga, and Viladecans-Marsal (2020) and have basic information on basic demographics from the operator.

Combine with other data: population density, census data, geo-coded survey data, surveys on mobile phone usage.

What are the cell phone penetration rates for the period of your data?

Distinct sources of measurement error

Some conceptual considerations:

- Lack of coverage: underestimate mobility
- Switched off devices: underestimate mobility
- Sharing of phone: overestimate network, overestimate mobility
- Multiple SIM cards: no issue with smartphone app location data (selection)
- Multiple phones: selection, selective usage
- Spammers/call centres in CDR data (remove numbers with large numbers of contacts)
- Missed calls: unobserved? (selection)
- Communication via mobile phones is subset of communication

Distinct sources of measurement error (cont.)

Some technical considerations:

- Load sharing: falsely attribute to mobility
- Missing tower locations (e.g., about 12% in Björkegren (2019))
- Missing data (data over certain periods might be missing)
- Discontinuities in tower availabilities
- Data sinks
- Mislabelling

Key: to what extent does this matter for your specific research question?

Quality of the judgements you make along the way is key for the quality of your research.

Blumenstock et al. (2023)

Migration is a key economic decision.

Networks at home and in the potential destination might play an important role in providing information and support.

Difficult question to answer because need detailed network and migration data.

This paper uses cell phone data to ask: how does an individual's network affect their decision to migrate?

Role of cell phone data: measurement and identification.

The potential role of networks

Three migrants: A, B and C.

Same number of contacts at home, but different network structure at destination.

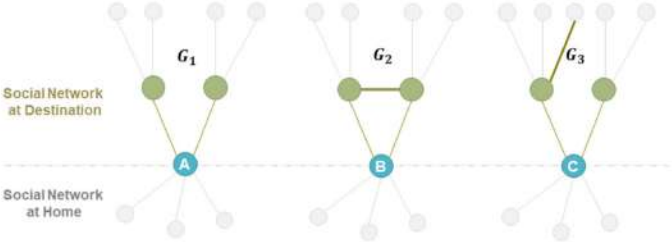


FIGURE 1

Schematic diagrams of the social networks of three migrants

Notes: Each of the blue circles (labeled A, B, C) represents a different individual considering migrating from their home to a new destination. Each individual has exactly three contacts in the home district (smaller grey circles below the dashed line) and two contacts in the destination district (larger green circles above the dashed line). The social network of these three individuals is denoted by G_1, G_2, and G_3.

Who is more likely to migrate?

The potential role of networks

Three migrants: A, B and C.

Same number of contacts at home, but different network structure at destination.

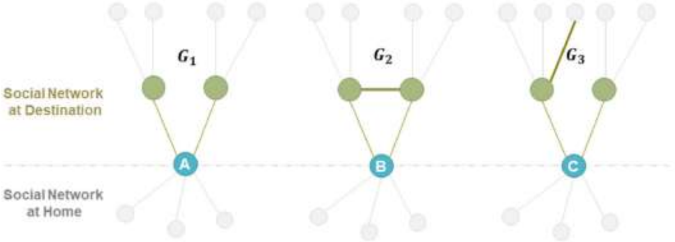


FIGURE 1

Schematic diagrams of the social networks of three migrants

Notes: Each of the blue circles (labeled A, B, C) represents a different individual considering migrating from their home to a new destination. Each individual has exactly three contacts in the home district (smaller grey circles below the dashed line) and two contacts in the destination district (larger green circles above the dashed line). The social network of these three individuals is denoted by G_1, G_2, and G_3.

Who is more likely to migrate? B, then A, then C.

Data

Call detail records for Rwanda from 2005 - 2009: anonymized phone numbers, timestamp, towers.

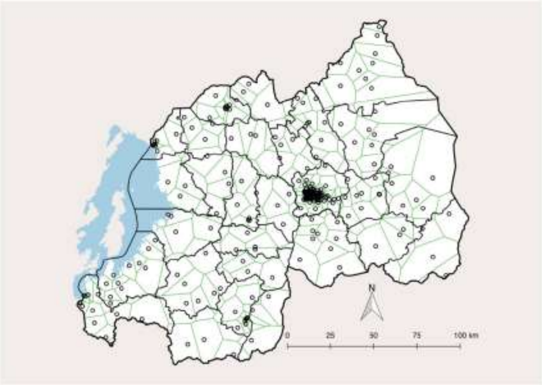


FIGURE 3
Location of all mobile phone towers in Rwanda, circa 2008

Notes: Circles indicate cell tower locations. Dark/black lines represent district borders. Light/green lines show the voronoi polygons roughly divide the country into the coverage region of each tower.

Blumenstock et al. (2023): Measuring migration

Building on Blumenstock (2012), Lai, Erbach-Schoenberg, Pezzulo, Ruktanonchai, Sorichetta, Steele, Li, Dooley, and Tatem (2019) and Chi et al. (2020).

Algorithm contains several steps:

1. Infer district of residence d every month t from evening location
2. Migration occurs if three conditions are met:
   - Observe an individual's home location in d at least k months prior to (and including) t.
   - Home location in d' is not equal to d
   - Observe an individual's new home location d' for at least k months after (and including) $t+1$.

Blumenstock et al. (2023): Measuring migration (cont.)

Migration rates

k	% Ever Migrate	% Repeat Migrant	% Long-distance Migrant	% Circular Migrants
1	34.6	11.2	23.2	18.5
2	21.6	1.9	13.8	5.9
3	14.0	0.4	9.2	2.0
6	5.3	0.0	3.5	0.1

Large fraction of migrants (in particular long-distance migrants and circular migrants).

Fraction decreases as k increases.

Cell phone data: allow for flexible definition of migration in terms of time and space; easily comparable with other data.

Blumenstock et al. (2023): Network of migrants

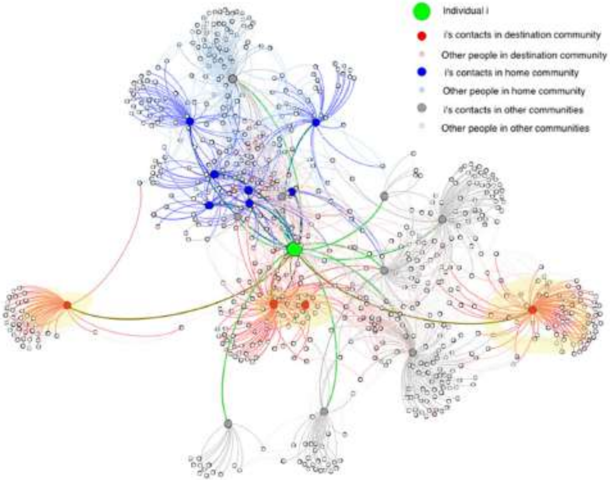


FIGURE 2
The social network of a single migrant

Blumenstock et al. (2023) Quantifying networks

Information capital: size of the second-degree neighborhood (friends-of-friends).
$\rightarrow$ flow of information about jobs and other opportunities.

Cooperation capital: probability that individual *i*'s friend *j* has one or more friends in common with *i*.
$\rightarrow$ proxy for insurance, support.

Also explore other measures such as degree centrality, strength of a social tie (e.g., more than 5 calls per month).

Blumenstock et al. (2023): Identification

Central identification concern: networks are not exogenous.

Simultaneity: migration shapes networks.

$\rightarrow$ Use lagged network.

$\rightarrow$ Focus on connections of contacts, holding contacts fixed.

$\rightarrow$ Use shift-share specification.

Omitted variables: network is correlated with wealth, population density, wages, etc. that also affect migration.

$\rightarrow$ Include rich set of fixed effects:

- Individual (800,000 FE)
- Origin-destination-month (18,000 FE)
- Number of direct contacts in destination (100 FE)

Blumenstock et al. (2023): Results

Additional contact at origin (destination) decreases (increases) the probability of migration.

Migrants are more likely to migrate to places where they have interconnected networks.

Extensive networks are negatively affecting migrations.

This might be due to competition in information.

Higher-order network structure is important.

Blanchard et al. (2023): Motivation

Understanding human mobility in low-income countries has been limited by lack of data

- Migration: census or standard household surveys
- Commuting: commuting surveys

Newer sources of "Big data"

- More fine-grained measures for migration and commuting (Blumenstock et al., 2023; Kreindler and Miyauchi, 2021; Miyauchi et al., 2022)

Know little about human mobility in developing countries over other time scales.

Blanchard et al. (2023): Visits

Study a type of mobility previously difficult to capture.

Visits: movement of people from their home locations to other locations, not necessarily for daily work.

This type of mobility is common and frequent.

Visits may allow individuals to benefit from amenities without relocating.

Examples:
Enjoying consumption goods that are unavailable elsewhere, managing administrative and legal matters, market goods and services without paying traders.

Blanchard et al. (2023): What we do

Use fine-grained, anonymized data on smartphone app locations.

Each ping represents instance device connects to the internet.

More than one million devices over an entire year.

Three African countries: Kenya, Nigeria and Tanzania.

Provide new evidence on high-frequency mobility for large numbers of people at high spatial and temporal resolutions.

Blanchard et al. (2023): Contributions

1. Construct a novel set of metrics for characterizing mobility across space
   - Frequency, spatial extent, densities and places visited
   - Parsimonious and easily interpretable, characterize connectedness between locations

2. Use measures to provide a rich description of mobility in an African context
   - Do we observe much long-distance travel?
   - What patterns of mobility do we observe across cities?
   - Do we observe much connectivity between rural and urban areas?

3. Write down and test a model that characterizes the visits that we observe in the data (see paper).

Blanchard et al. (2023): Data

Each observation is a "ping": smartphone accesses internet via a set of apps (i.e. social, navigation, information).

For each ping we know: device identifier, timestamp and longitude/latitude.

Start by assigning home locations: modal 0.01 degree cell user is seen between 7pm and 7am; two additional restrictions:

1. User is observed for more than 10 nights
2. User is at inferred home location for at least 50 percent of total nights.

High-confidence users satisfy both restrictions.

Do the same for work locations.

Blanchard et al. (2023): Selection

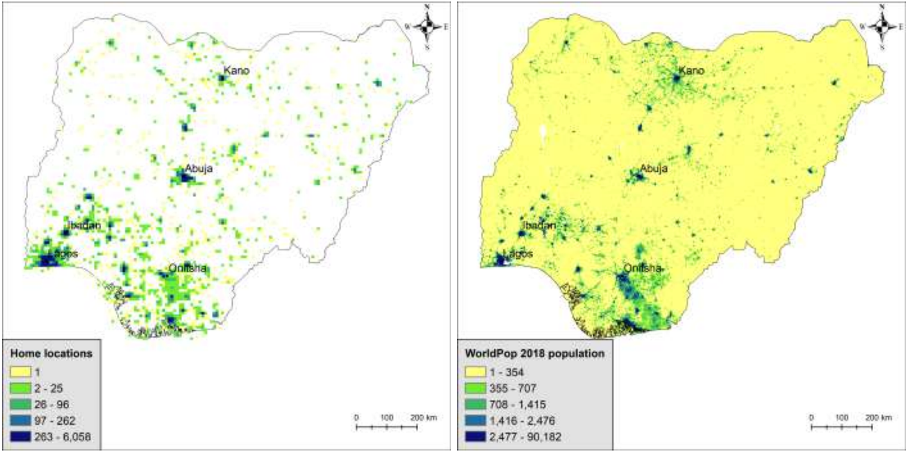
Smartphone app users in no way representative of the population in our setting.

Take three steps to examine representativeness:

1. Match home locations with population density.
2. Use ICT Access and Usage Survey to examine characteristics of users by device ownership.
3. Match home locations with DHS data to examine characteristics of places.

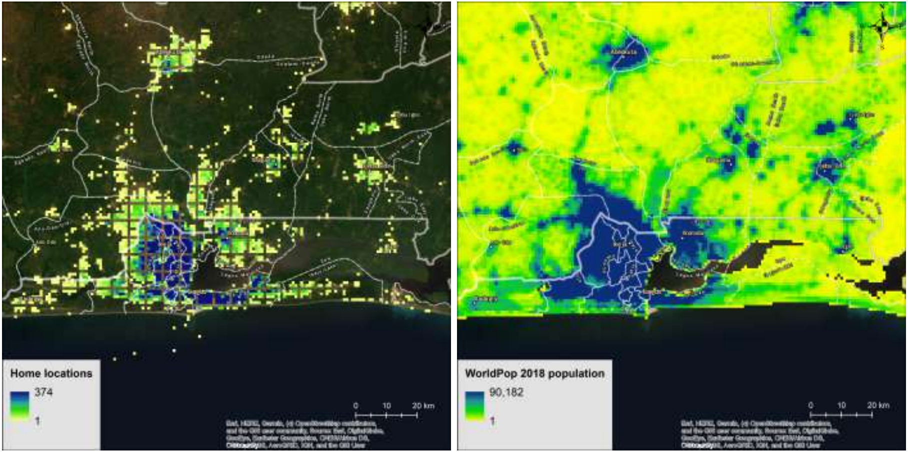
In paper: discuss device = users (device sharing, multiple devices/SIM cards), turned-off devices, coverage, smartphone app location data vs CDR data, transit pings, outliers.

Blanchard et al. (2023): Home locations in Nigeria and World Pop population



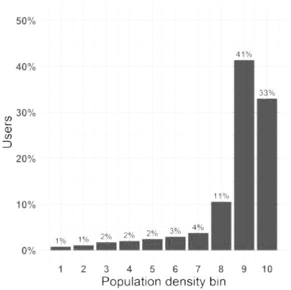
- Coverage of users is broadly national (users in 112/115 regional capitals).
- R-squared between population and users between 0.43 and 0.62 with coefficients of 0.95 (Kenya), 1.69 (Nigeria) and 0.96 (Tanzania).

Blanchard et al. (2023): Home locations in Lagos and World Pop population

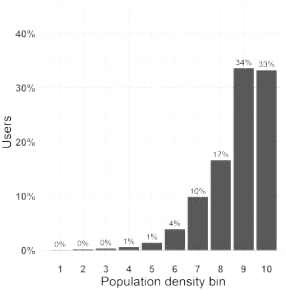


Broad coverage within cities.

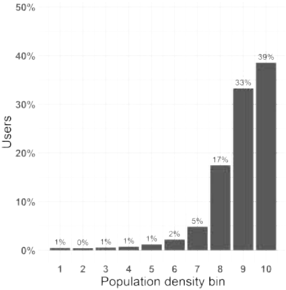
Blanchard et al. (2023): Population density



Kenya



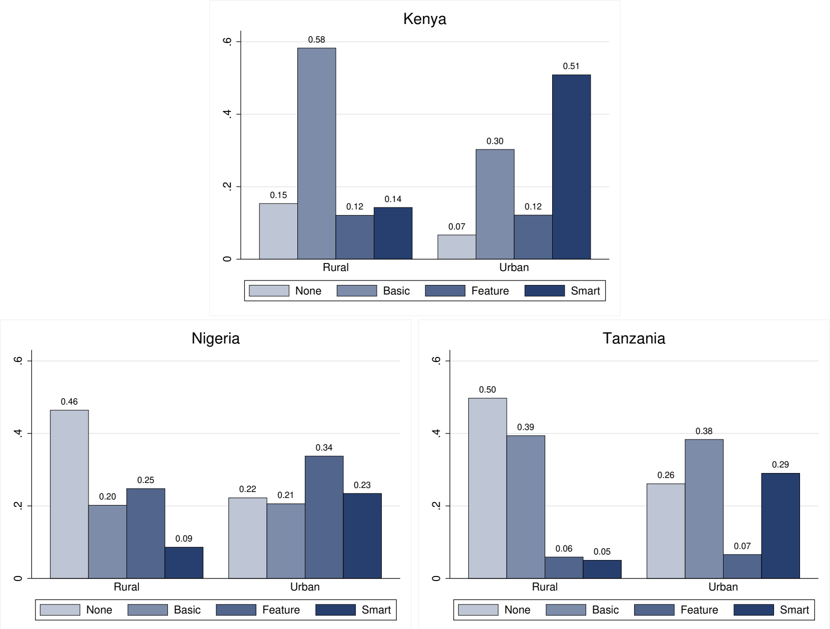
Nigeria



Tanzania

About 70% of users are falling into the highest density bins.

Blanchard et al. (2023): Device ownership by location



Blanchard et al. (2023): T-tests for equality of means between DHS and matched DHS samples, Nigeria.

	Variable	DHS	Matched DHS	Difference	SE	p-value
Urban	Household size	4.44	3.83	-0.61	0.03	0.000***
	Age of HH head	45.21	45.18	-0.02	0.18	0.900
	Education of HH head	9.66	11.56	1.91	0.06	0.000***
	Access to piped water	0.14	0.14	-0.01	0.01	0.572
	Constructed floor	0.89	0.96	0.08	0.01	0.000***
Rural	Household size	4.85	3.92	-0.93	0.03	0.000***
	Age of HH head	45.34	44.77	-0.57	0.16	0.000***
	Education of HH head	6.03	10.23	4.20	0.06	0.000***
	Access to piped water	0.09	0.14	0.05	0.01	0.000***
	Constructed floor	0.64	0.96	0.32	0.01	0.000***

In almost two-thirds of 66 (11x2x3) comparisons, differences are less than 10 percent.

Blanchard et al. (2023): Selection summary

Users are more urban; likely more educated, richer.

Smartphone owners represent substantial fraction of the urban population (23% in Nigeria to 51% in Kenya).

Absence of evidence on this type of mobility on entire populations.

Provide insights into high-frequency mobility within a substantial fraction of overall population that is worthwhile to study.

Blanchard et al. (2023): Measuring mobility

- Characterizing mobility:
    1. Frequency.
    2. Spatial extent.
    3. Densities visited.
    4. Cities visited.
- Characterizing connectedness of locations (e.g. cities):
    1. Incoming flows: number of visitors, frequency of visits, distance travelled, density at origin.
    2. Outgoing flows: number of residents seen outside the area, frequency of movements, spatial extent, densities visited.

Blanchard et al. (2023): Measuring mobility: frequency of travel

Fraction of days with mobility beyond 10km:

- 13.8% Kenya
- 15.2% Nigeria
- 11.8% Tanzania

Average distance from home for non-home pings (km):

- 33.8 km Kenya
- 45.5 km Nigeria
- 31.8 km Tanzania

Blanchard et al. (2023): Measuring mobility: Share of users by home bin-visited bin pair

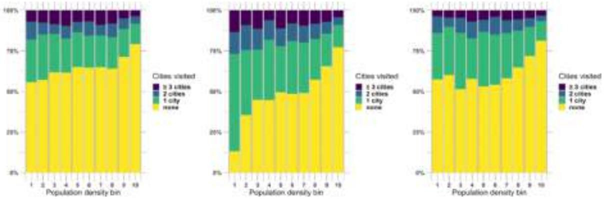
		Home density bin									
		1	2	3	4	5	6	7	8	9	10
	1	35.7%	19.6%	18.8%	6.8%	6.1%	3.8%	3.3%	3.1%	2.5%	1.5%
	2	23.8%	33.3%	35%	12.1%	12.6%	9.1%	6.8%	6.1%	5.3%	3.2%
	3	26.2%	29%	41.5%	32%	18.9%	13.1%	10.5%	8.8%	7.3%	4.7%
	4	31%	26.8%	45.3%	35.2%	32.6%	22.3%	15%	12%	11.2%	6.9%
Visited	5	23.8%	33.3%	43.6%	45.9%	51.3%	38.1%	27%	21%	20.1%	15.2%
density	6	33.3%	33.3%	37.6%	53.9%	60%	68.7%	45.8%	31.5%	26.8%	17.4%
	7	42.9%	55.8%	50.9%	52.7%	64%	69.9%	76.1%	56.1%	39.8%	25.5%
	8	71.4%	58.7%	54.7%	58.7%	61.5%	60%	72.8%	81.2%	63.7%	37.9%
	9	76.2%	61.6%	62.8%	62.6%	66.8%	64.1%	68.4%	81.2%	91.5%	64.7%
	10	42.9%	44.9%	43.2%	44.7%	50.2%	47.8%	46.9%	46.9%	61.9%	95.3%

Nigeria

Substantial flows of people across densities.

Blanchard et al. (2023): Distribution of users according to the number of non-home cities visited, by population density bin.

Figure 7: Distribution of users according to the number of cities visited, by population density bin.



(a) Kenya (b) Nigeria (c) Tanzania

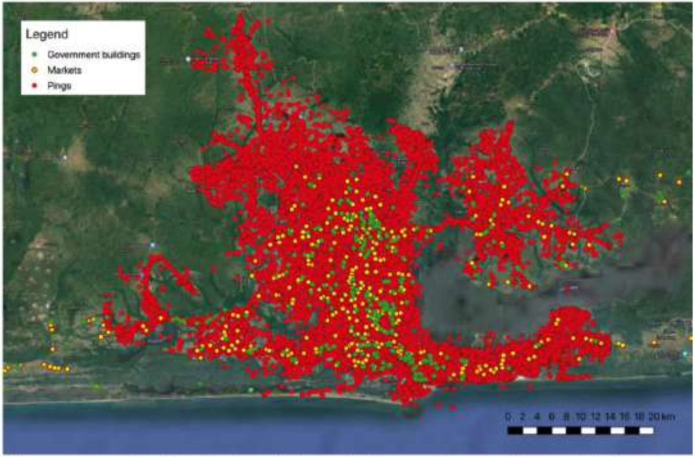
Note: This figure shows for each decile the distribution of users who are never seen in a city, those who visit exactly one city, those seen in two cities and those visiting three or more cities. These counts exclude the home city in the case of urban residents.

Across all densities a sizeable fraction of individuals make visits to one or more non-home cities.

Individuals make multiple visits to non-home cities on average.

Blanchard et al. (2023): Destinations visited within cities: Lagos case study.

Figure 9: Destinations visited within cities: Lagos case study.



Note: This figure shows the distribution of pings of visitors to Lagos. Data on the location of markets and government buildings come from Federal Republic of Nigeria (2021).

Blanchard et al. (2023): Destinations (cont)

Figure 10: Destinations visited within cities: Lagos case study examples.

(a) Shopping (b) Casino (c) Airport

(d) Stadium (e) Embassy (f) Hospital

Note: This figure shows the distribution of pings of visitors to Lagos at specific selected locations.

Blanchard et al. (2023): Distribution of users across places visited

Table 4: Distribution of users across places visited by density of origin.

	Visitors		Visitors from Below 300		Visitors from Above 300		Residents	
	Users	% of users	Users	% of users	Users	% of users	Users	% of users
Total	16,156	-	590	-	15,543	-	67,982	-
Total users matched with OSM	13,214	100.0%	438	100.0%	12,756	100.0%	60,432	100.0%
Residential	10,628	80.4%	288	65.8%	10,325	80.9%	54,633	90.4%
Roads and roadsides	6,815	51.6%	251	57.3%	6,560	51.4%	36,795	60.9%
Travel	4,825	36.5%	162	37.0%	4,652	36.5%	17,329	28.7%
Shops and markets	3,775	28.6%	159	36.3%	3,614	28.3%	30,076	49.8%
Commercial zone	2,835	21.5%	88	20.1%	2,745	21.5%	21,366	35.4%
Industrial zone	2,280	17.3%	78	17.8%	2,197	17.2%	21,445	35.5%
Public goods and services	1,540	11.7%	70	16.0%	1,469	11.5%	16,315	27.0%
Recreational	1,008	7.6%	45	10.3%	962	7.5%	10,516	17.4%
Other	733	5.5%	35	8.0%	696	5.5%	7,311	12.1%
Food and drinks	347	2.6%	16	3.7%	331	2.6%	3,893	6.4%
Worship	331	2.5%	16	3.7%	314	2.5%	4,733	7.8%

Note: This table links the locations of visitors to Lagos, Abuja, Nairobi, Mombasa, Dar es Salaam and Dodoma and residents of these cities with OSM data to show the type of locations visitors and residents are seen at.

Blanchard et al. (2023): Conclusion

Smartphone app location data allow to unveil a striking degree of mobility in an African context, that is neither commuting nor migration.
$\rightarrow$ Challenges idea that rural areas are isolated.

Theoretical framework where individuals optimally consume location-specific amenities via “visits”.

Mobility patterns consistent with theoretical framework (see paper):

1. Visits per person from smaller to larger settlement higher than vice versa.
2. Fraction of days users spend visiting a city follows gravity equation.
3. Between two equidistant locations, users visit larger one.

Blanchard et al. (2023): Conclusion, cont.

- Previous literature suggests (though contested!) that sectoral and spatial gaps are large.
- If true, this should lead to strong pressure for mobility across locations and sectors.
- An alternative / additional explanation:
  - 'Visits' offer an alternative technology to migration.
  - People in rural areas and small cities can get the benefits of larger urban centers without relocating.
  - Through visits, cities benefit people beyond their own residents and commuters.
  - Hard to square with any model in which gaps are sustained by direct costs of travel or information frictions...may be different for the poor.

Further resources

United Nations (2024) Methodological guide on the use of mobile phone data.

European Commission (2019) Measuring labour mobility and migration using big data.

Big data and development course by Josh Blumenstock.

Replication files for published papers (see for example here for Blumenstock et al. or here for Kreindler and Miyauchi).

City-scale and longitudinal dataset of anonymized human mobility trajectories (2024), *Scientific Data*

Research published outside economics journals (e.g., Nature, Science, PNAS, PLOS ONE).

Upcoming NetMob Data Challenge

NetMob 2024 Data Challenge is launched on 3 June 2024.

Will release a dataset on mobility to work with.

This year the challenge "targets countries in the Global South, aligning with the Sustainable Development Goals (SDGs)."

Past editions: mobile data traffic, mobile phone data.

Note that you will need to ask for permission to keep working with the data after the conference.

Supplemental spatial data

Some basic datasets to get started:

- Global Administrative Boundaries
- Population Density
  - WorldPop
  - Gridded Population of the World (GPW) v4
  - Global Human Settlement Layer
- Global Rural-Urban Mapping Project (GRUMP), v1

Many sources for further data (e.g., see here for various spatial datasets) and other parts of this course.

How to get started

1. Understand data structure
   - CDR, digital trace data
   - Events, smartphone app pings, regular pings
   - Spatial resolution (tower, precise coordinates, displacement)
   - Number of devices, period, frequency of observations
2. Carefully explore data or portions of the data (!)
3. Understand your sample
   - Infer home location
   - Link with population density, administrative boundaries
   - Link with other micro-data: surveys, census
   - Link with other spatial data: city outlines, transport infrastructure, firm location, natural disasters
4. What is the right research question to answer with this dataset?

Good luck!

Q&A

References

Aiken, E., S. Bellue, D. Karlan, C. Udry, and J. E. Blumenstock (2022): "Machine learning and phone data can improve targeting of humanitarian aid," *Nature*, 603, 864–870.

Athey, S., B. Ferguson, M. Gentzkow, and T. Schmidt (2021): "Estimating experienced racial segregation in US cities using large-scale GPS data," *Proceedings of the National Academy of Sciences*, 118, e2026160118.

Atkin, D., K. Chen, and A. Popov (2022): "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley," Unpublished working paper.

Barwick, P. J., Y. Liu, E. Patacchini, and Q. Wu (2023): "Information, mobile communication, and referral effects," *American Economic Review*, 113, 1170–1207.

References (cont.)

Björkegren, D. (2019): “The adoption of network goods: Evidence from the spread of mobile phones in Rwanda,” *The Review of Economic Studies*, 86, 1033–1060.

Blanchard, P., D. Gollin, and M. Kirchberger (2023): “Perpetual Motion: High-Frequency Human Mobility in Three African Countries,” Tech. rep., Trinity College Dublin, Department of Economics.

Blumenstock, J., T. Ghani, S. Herskowitz, E. B. Kapstein, T. L. Scherer, and O. Toomet (2020): “How do firms respond to insecurity? Evidence from Afghan corporate phone records,” *Mimeo*.

Blumenstock, J. E. (2012): “Inferring patterns of internal migration from mobile phone call records: evidence from Rwanda,” *Information Technology for Development*, 18, 107–125.

References (cont.)

Blumenstock, J. E., G. Chi, and X. Tan (2023): "Migration and the value of social networks," *Review of Economic Studies*, rdad113.

Blumenstock, J. E., D. Gillick, and N. Eagle (2010): "Who's calling? Demographics of mobile phone use in Rwanda," in *2010 AAAI Spring Symposium Series*.

Büchel, K., M. V. Ehrlich, D. Puga, and E. Viladecans-Marsal (2020): "Calling from the outside: The role of networks in residential mobility," *Journal of urban economics*, 119, 103277.

Chi, G., F. Lin, G. Chi, and J. Blumenstock (2020): "A general approach to detecting migration events in digital trace data," *PloS one*, 15, e0239408.

Kirchberger, M. (2021): "Measuring internal migration," *Regional Science and Urban Economics*, 91, 103714.

References (cont.)

Kreindler, G., A. Gaduh, T. Graff, R. Hanna, and B. A. Olken (2023): "Optimal Public Transportation Networks: Evidence from the World's Largest Bus Rapid Transit System in Jakarta," Tech. rep., National Bureau of Economic Research.

Kreindler, G. E. and Y. Miyauchi (2021): "Measuring Commuting and Economic Activity inside Cities with Cell Phone Records," Unpublished Manuscript.

Lai, S., E. z. Erbach-Schoenberg, C. Pezzulo, N. W. Ruktanonchai, A. Sorichetta, J. Steele, T. Li, C. A. Dooley, and A. J. Tatem (2019): "Exploring the use of mobile phone data for national migration statistics," *Palgrave communications*, 5, 1–10.