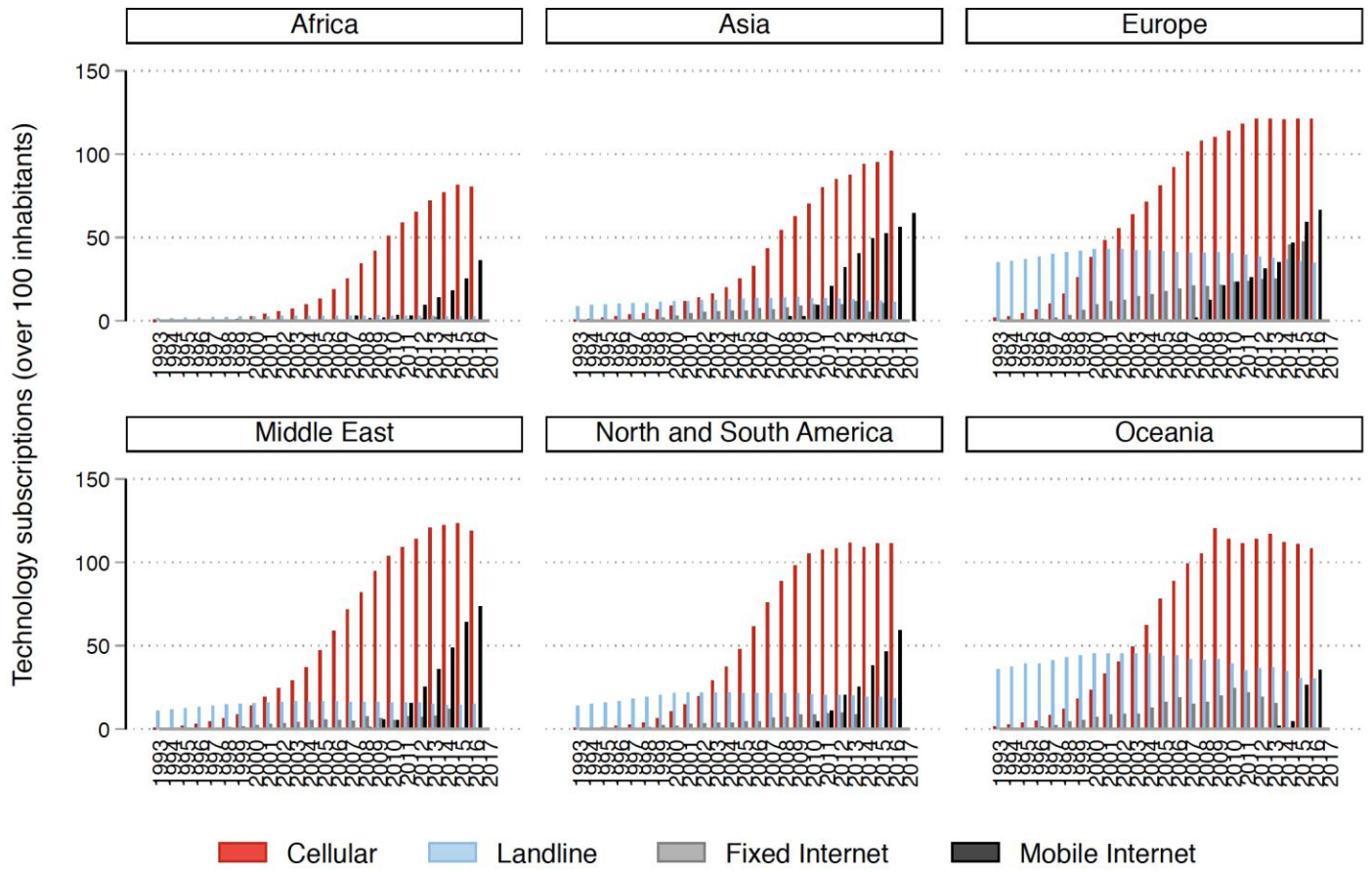
Gender Inequalities in a Digital World: Perspectives from Digital Demography

Ridhi Kashyap
Department of Sociology & Leverhulme Centre for Demographic Science
Nuffield College
University of Oxford
www.digitalgendergaps.org

12th July 2025
Pre-Conference Workshop
International Population Conference 2025

The Digital Revolution



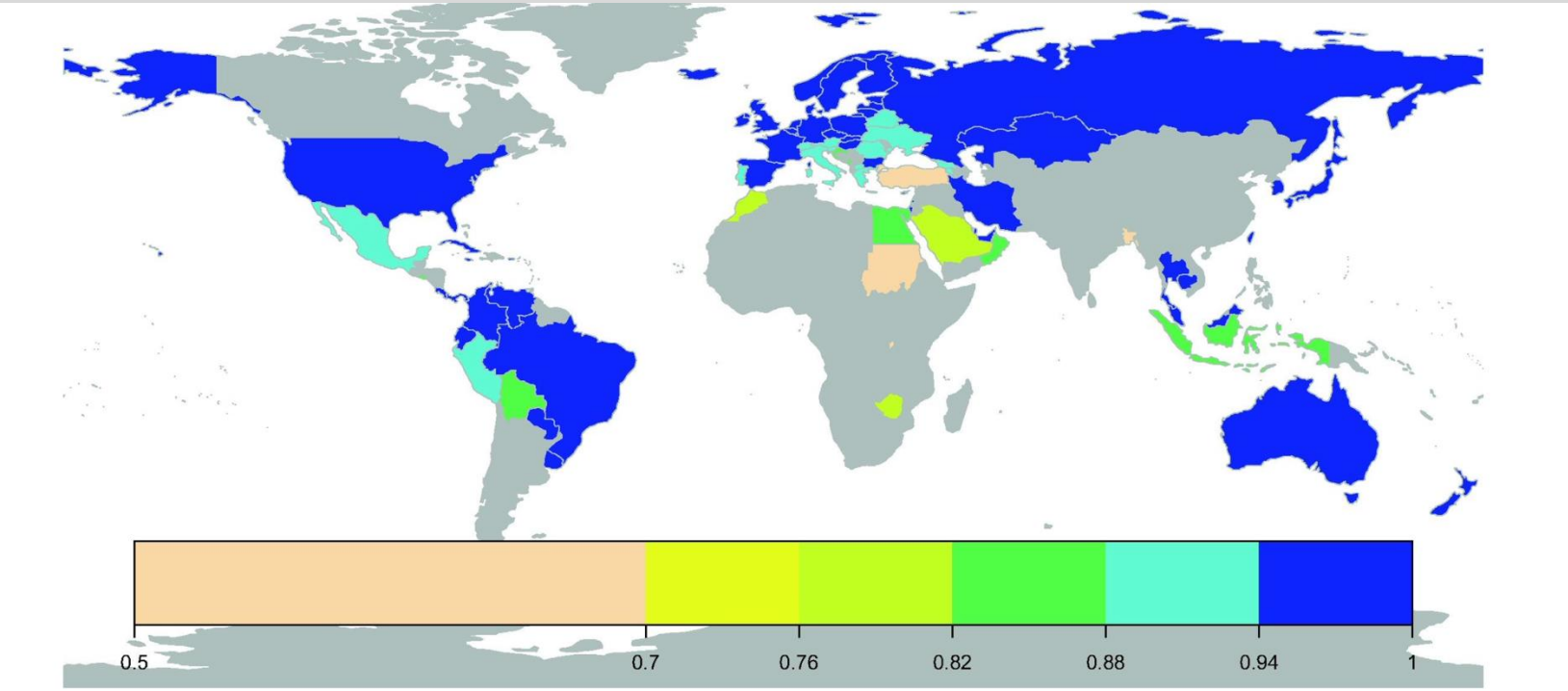
Technology subscriptions per capita, by region, 1993-2017.
Data from International Telecommunications Union (ITU)

The Digital Revolution

- The digital revolution is:
  - **Social revolution:**
    - → SDG 5 on gender equality pledges to "enhance the use of…information and communication technologies to promote the empowerment of women."
    - → What is the empirical evidence? How, for whom, under what conditions?
    - → Theoretically ambiguous effects – can empower the marginalised but also reinforce or exacerbate inequalities
  - **Data revolution:**

    Our use of new technologies such as mobile phones, web and social media generates new data streams that have the potential to provide real-time information on SDGs

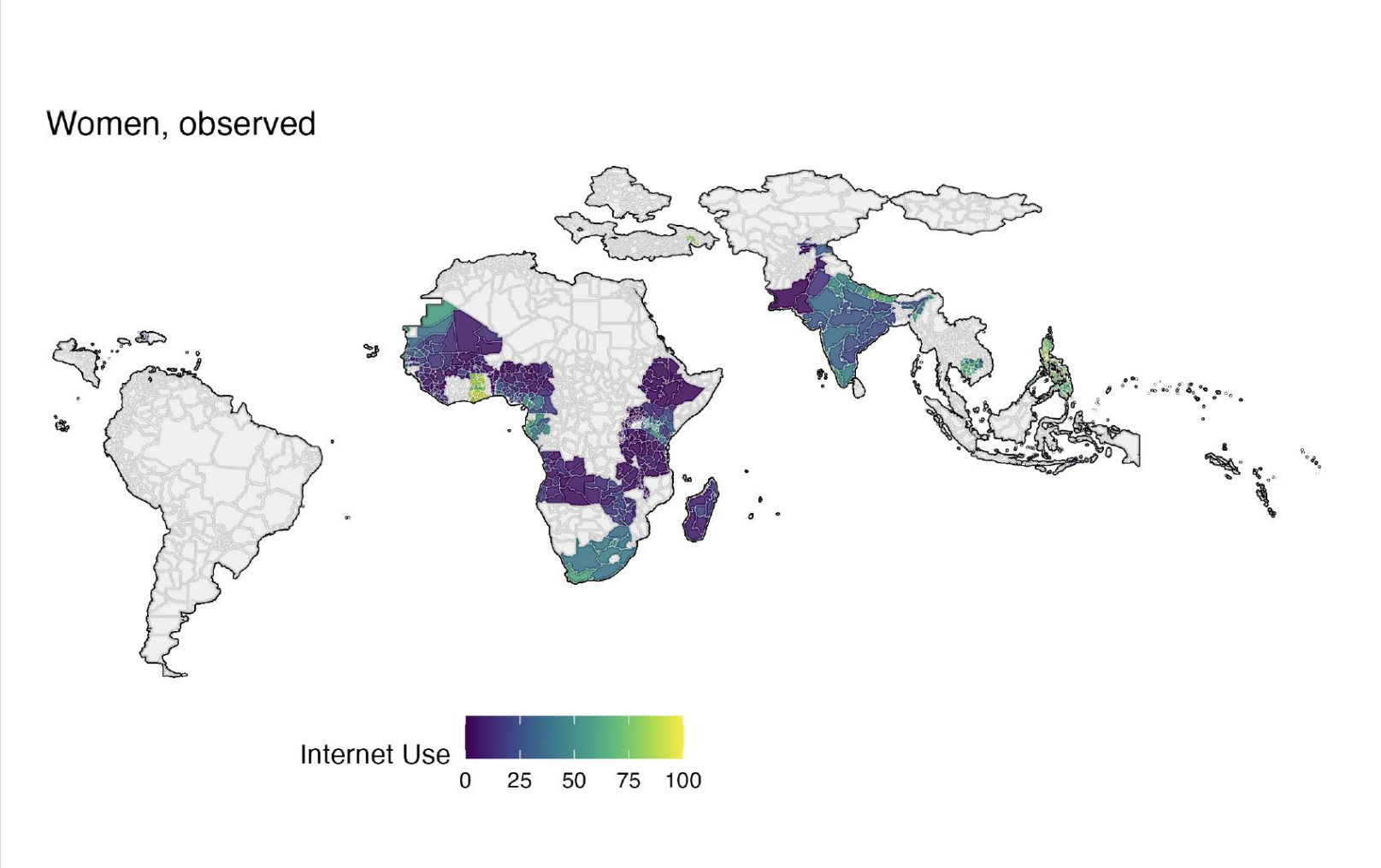
The Data Gap



(a) Internet Gender Gap Index computed using ITU ground truth data

Fatehkia, Masoomali, Ridhi Kashyap, and Ingmar Weber. "Using Facebook ad data to track the global digital gender gap." *World Development* 107 (2018): 189-209.

The Data Gap



Women who have used the internet at subnational resolution (admin-1) computed using Demographic and Health Surveys

The Digital Revolution

- The digital revolution is:
  - **Social revolution:**
    → SDG 5 on gender equality pledges to “enhance the use of…information and communication technologies to promote the empowerment of women.”
    → What is the empirical evidence? How, for whom, under what conditions?
    → Theoretically ambiguous effects – can empower the marginalised but also reinforce or exacerbate inequalities
  - **Data revolution:**
    Our use of the web, social media and mobile phones generates new data streams that have the potential to provide real-time information on digital connectivity and behaviours

Digital Gender Gaps Project

UNIVERSITY OF OXFORD

- The Digital Gender Gaps (DGG) project seeks to **deepen data and evidence** on:

  (1) **Gender gaps in digital connectivity and behaviours** at national- and subnational-levels, by integrating social media, geospatial data and population survey and census datasets;

Digital Gender Gaps Project

UNIVERSITY OF OXFORD

- The Digital Gender Gaps (DGG) project **seeks to deepen data and evidence** on:

  (1) **Gender gaps in digital connectivity and behaviours** at national- and subnational-levels by integrating social media, geospatial data and population survey and census datasets;

  (2) **Impacts of digital information and capabilities** on women's empowerment and gender inequalities, through quasi-experimental methods, and/or geospatial augmentation of surveys

The Digital Revolution and Demography

→ Advancing understanding of the **data** and **social** implications of the **digital revolution** requires the contributions of **digital and computational demography**.

1. Kashyap, Ridhi., 2021. “Has demography witnessed a data revolution? Promises and pitfalls of a changing data ecosystem” *Population Studies*, *75*(sup1), pp.47-75.

2. Kashyap, Ridhi, Gordon Rinderknecht et al. 2023. “Digital and Computational Demography”, Handbook of Digital Sociology, ed. Jan Skopek, Edward Elgar Publishing.

3. Kashyap, Ridhi, and Emilio Zagheni. 2023. "Leveraging Digital and Computational Demography for Policy Insights." Handbook of Computational Social Science for Policy. Springer International Publishing, 327-344.

Workshop

UNIVERSITY OF OXFORD

- Introduction to DGG research
  - Measuring digital inequalities using social media + population data
  - Impacts of digital technologies for gender inequalities

- 3-part lab session
  - Collecting social media data from FB marketing API
  - Application for modelling digital gender inequalities
  - Collecting digital inequality indicators from DGG web API

Jiaxuan Li
Research Assistant, DGG

Casey Breen
Postdoctoral Researcher, DGG

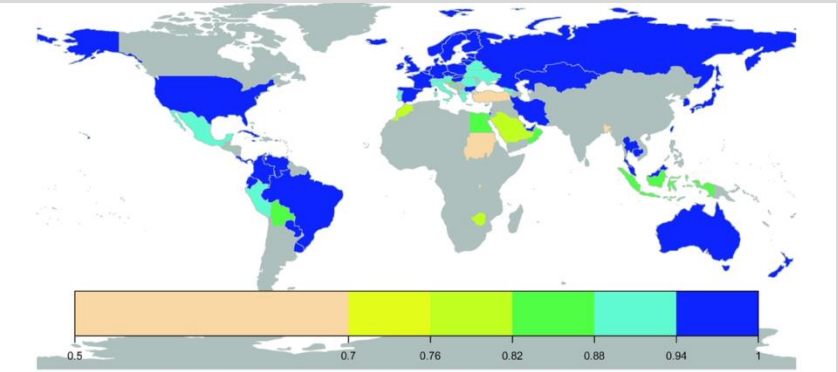
DGG Aim 1

- The Digital Gender Gaps (DGG) project **seeks to deepen data and evidence** on:

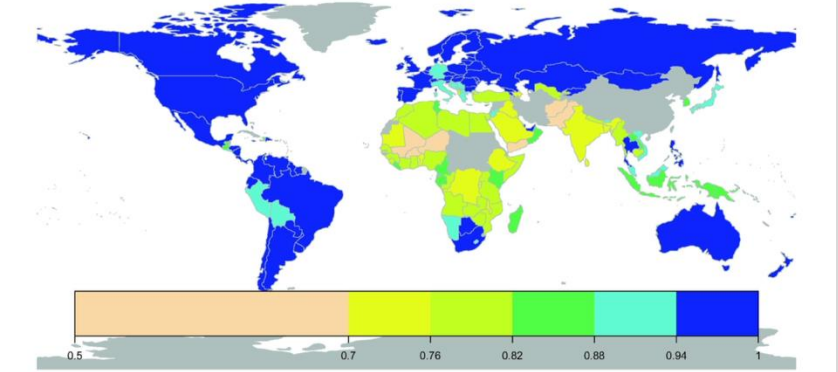
  (1) **Gender gaps in digital connectivity and behaviours** at national- and subnational-levels by integrating social media, geospatial data and population survey and census datasets;

  (2) **Impacts of digital information and capabilities** on women's empowerment and gender inequalities, through quasi-experimental methods, and/or geospatial augmentation of surveys

Internet GGI predicted using FB GGI



(a) Internet Gender Gap Index computed using ITU ground truth data



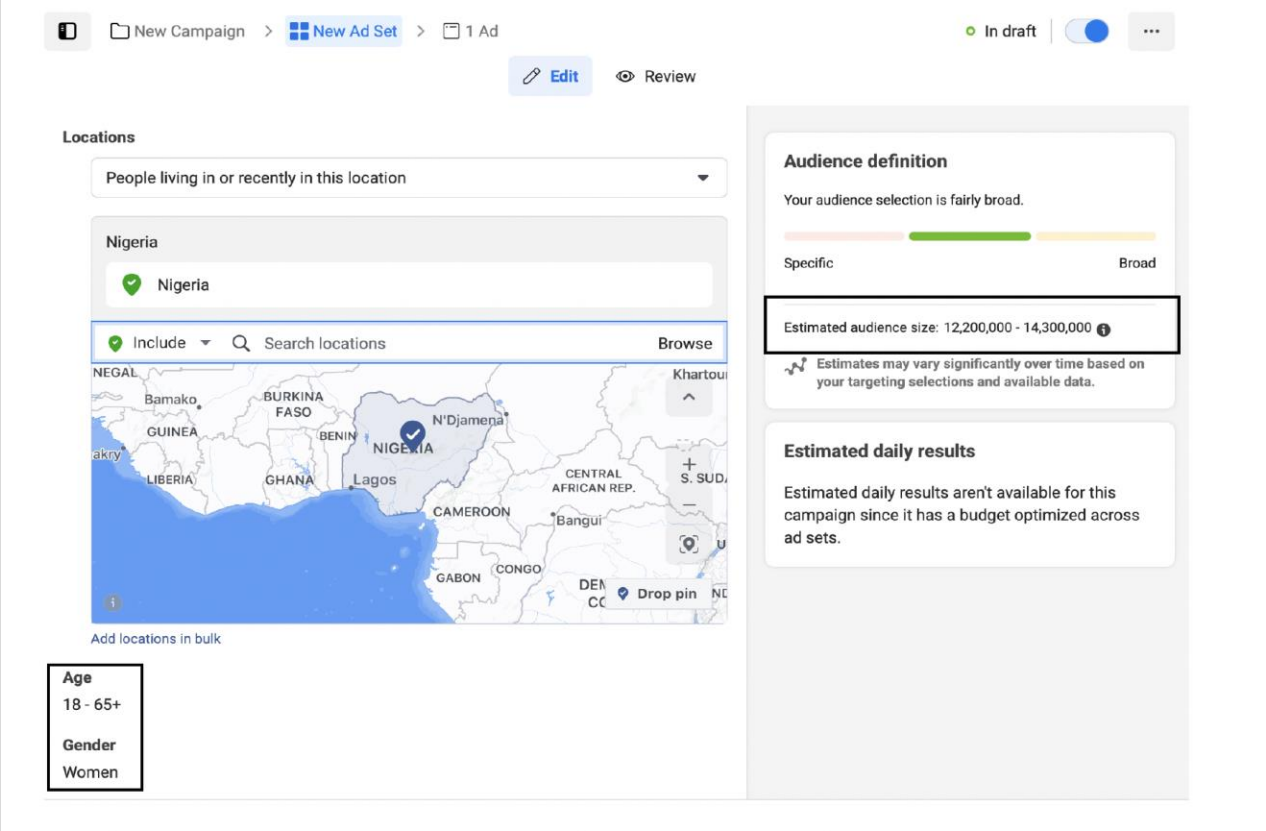
(b) Internet Gender Gap Index predicted from our Online Model using the Facebook 18+ user Gender Gap Index

Fatehkia, Masoomali, Ridhi Kashyap, and Ingmar Weber. "Using Facebook ad data to track the global digital gender gap." *World Development* 107 (2018): 189-209.

Facebook Marketing API

UNIVERSITY OF OXFORD

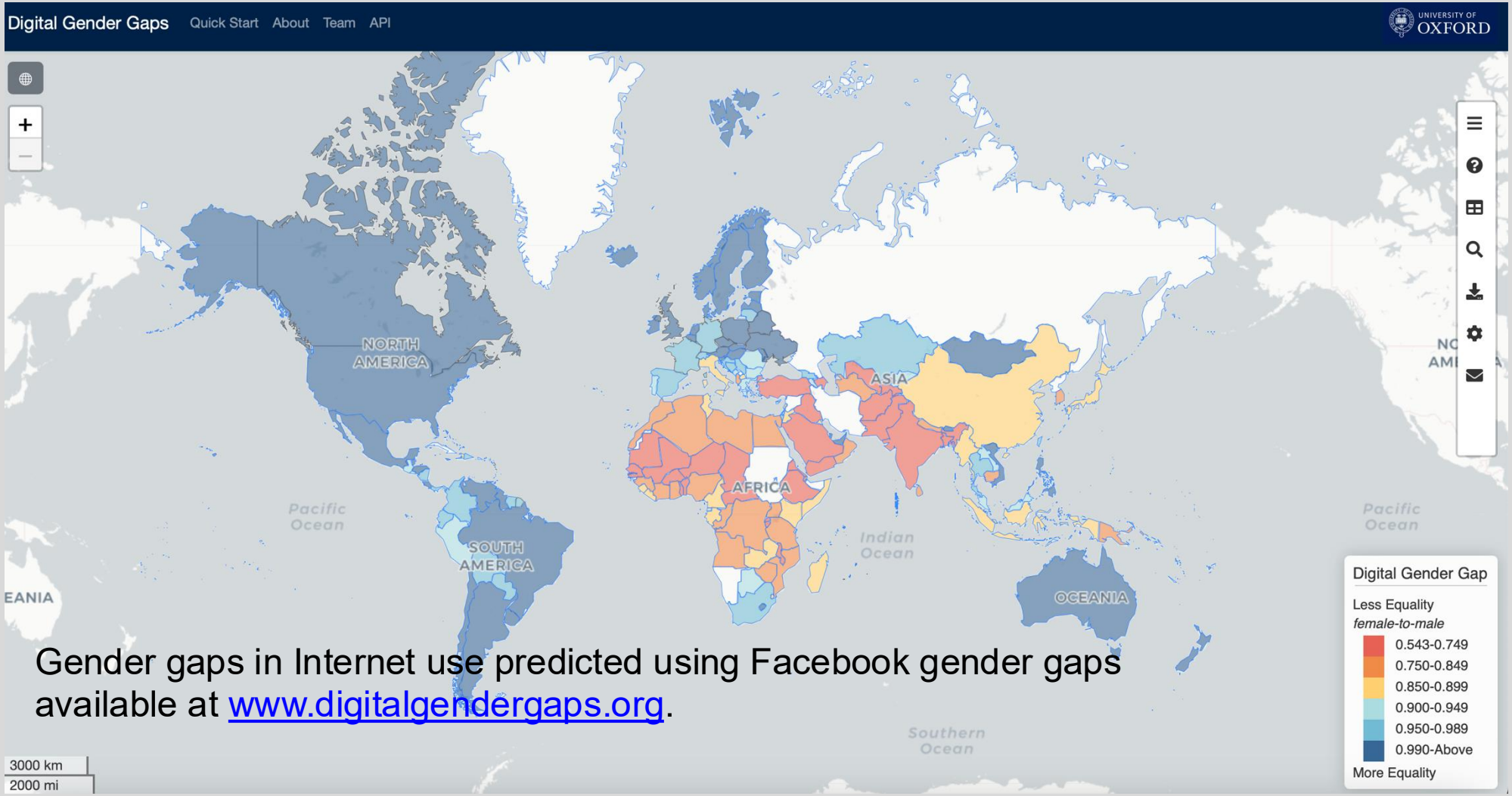
- Analysing global gender gaps in digital connectivity using **social media data advertising data validated against population surveys**



- How many users of ‘x’ characteristics (gender, location, age, device type, etc) are on a given platform?

Facebook Ads Manager:
https://www.facebook.com/adsmanager

Internet GGI predicted using FB GGI



Gender gaps in Internet use predicted using Facebook gender gaps available at www.digitalgendergaps.org.

Facebook Gender Gap Index

UNIVERSITY OF OXFORD

$$FB\ GGI = \frac{Female\ to\ male\ ratio\ of\ Facebook\ users}{Female\ to\ male\ ratio\ of\ the\ population}$$

- Using ITU data, we compute a ground truth internet gender gap indicator derived from survey data:
  - $Internet\ GGI = \frac{\%\ of\ female\ population\ using\ the\ Internet}{\%\ of\ male\ population\ using\ the\ Internet}$
- Similarly, we compute Mobile GGI

Facebook Gender Gap Index

UNIVERSITY OF OXFORD

$$\text{FB GGI} \approx \text{Internet GGI}$$

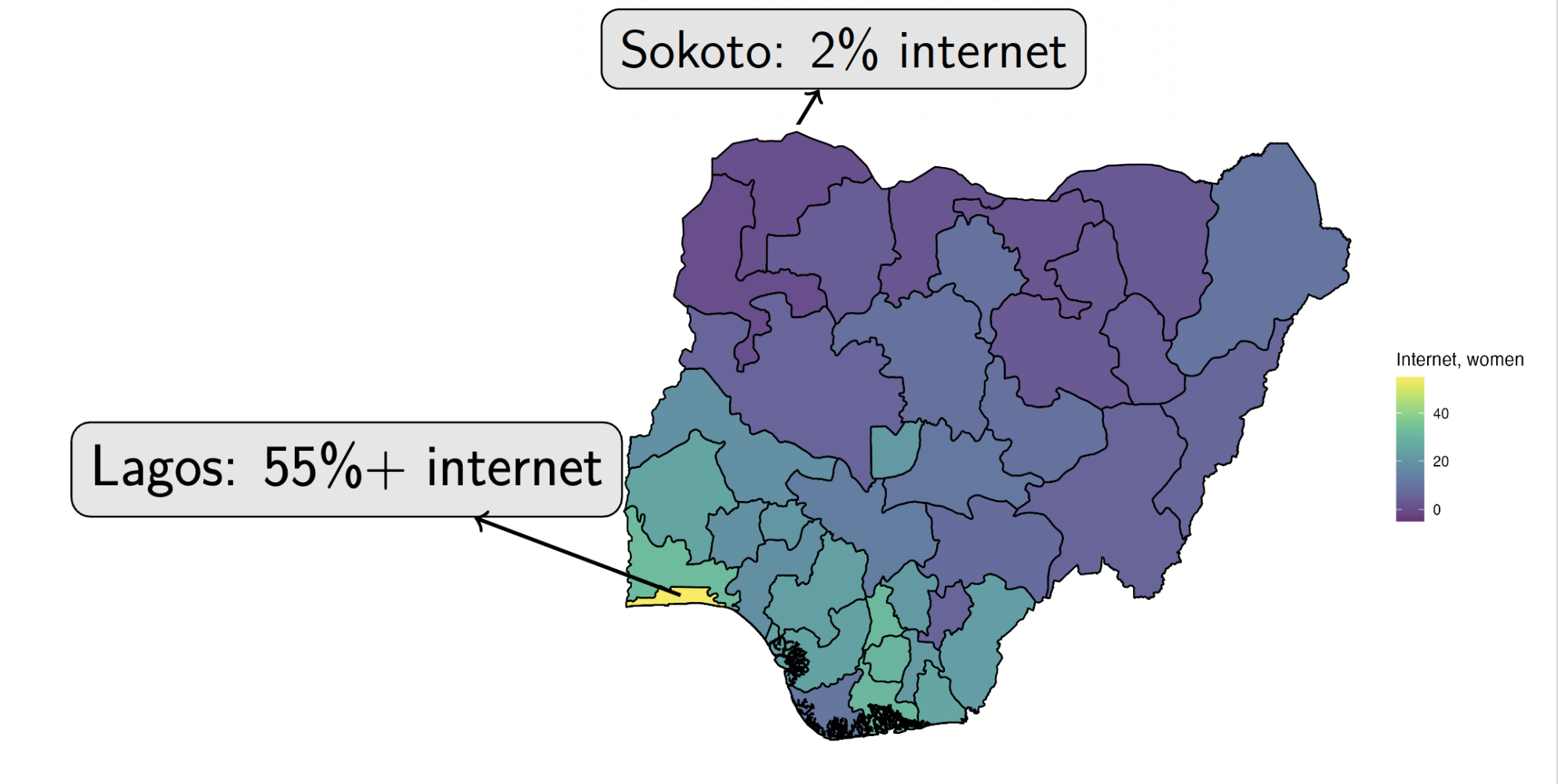
- Correlation of FB GGI 18+ with Internet GGI is ~0.8, mobile ~0.7
- Stronger than any other development indicator (including internet penetration)
- We fit three types of predictive models (linear models with cross-validation):
  - online model
  - online-offline model
  - offline model

Applications

UNIVERSITY OF OXFORD

- Using the FB Gender Gap Index to predict internet and mobile gender gap index
  - National-level gender gaps at www.digitalgendergaps.org
  - Ground truth data come from ITU-reported surveys (DHS, MICS, NSOs)
- Expanding geographical coverage, more timely data

Subnational Variation



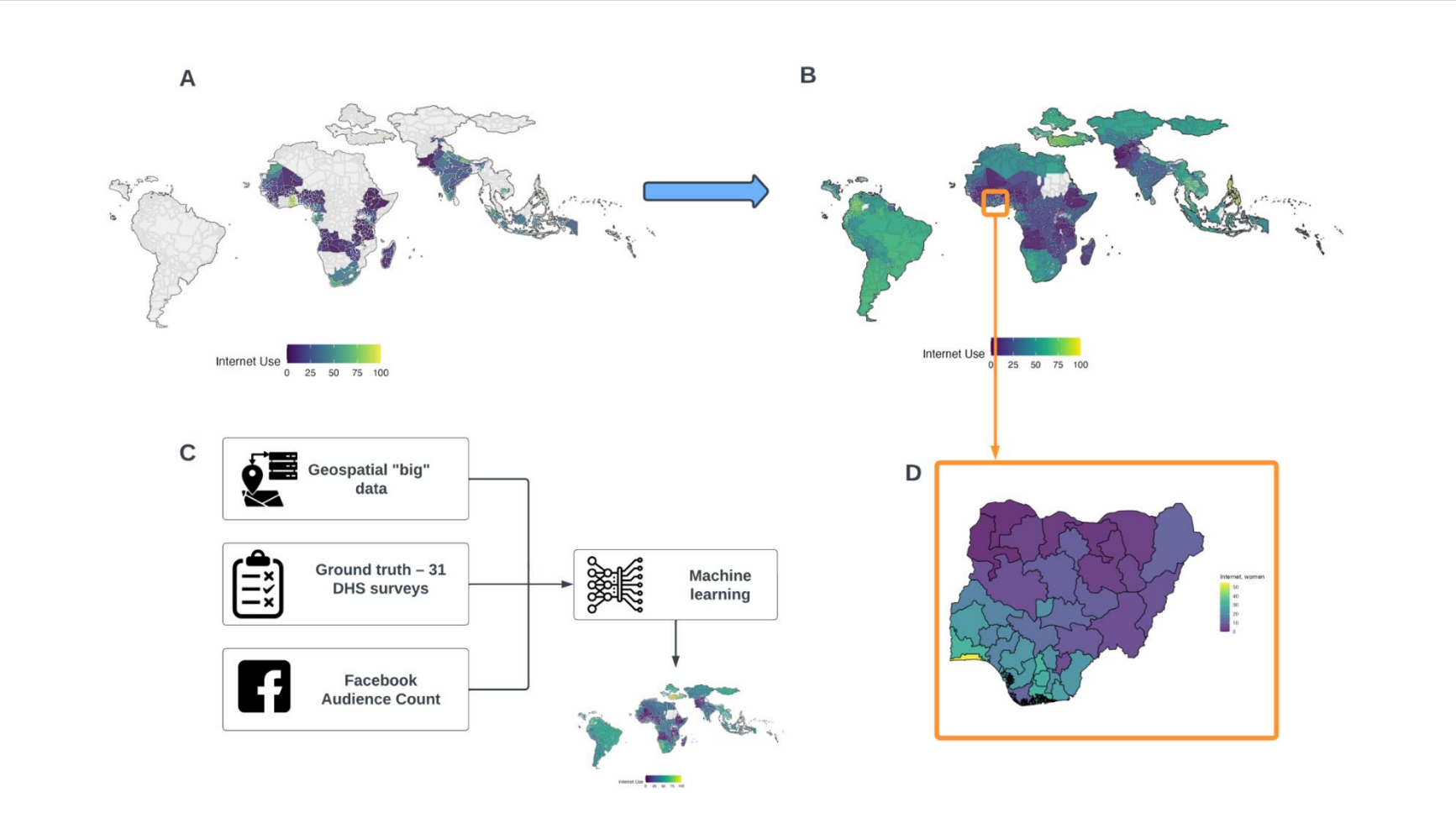
Internet use (last twelve months), women, Nigeria. From Demographic and Health Surveys (DHS).

UNIVERSITY OF OXFORD

Subnational Estimates of Digital Gender Gaps

- **Goal:** develop subnational estimates of digital adoption and gender gaps for LMICs
- GADM-1 resolution, 117 countries, 2,075 regions
- **Indicators**: gender-disaggregated internet use and mobile ownership (% internet user, women/men, %mobile owner, women/men) and gender gaps (female-to-male ratio)

Our approach



Panel A shows the 33 countries for which ground truth data are available at subnational resolution. **Panel C** shows inputs to the ML model used to predict digital indicators for 117 LMICs. **Panel B** shows model-based estimates of internet adoption by women. **Panel D** shows enlarged estimates for Nigeria.

Data

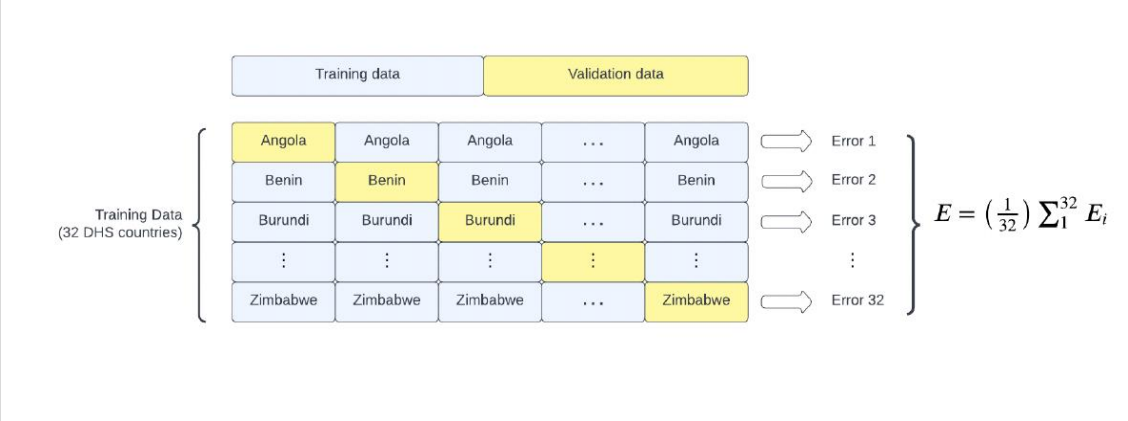
- **Ground truth:** Demographic and Health Surveys, representative at GADM-1 level, 2016-22 covering 33 countries

- **‘Online’ predictors:** Facebook penetration (proportion Facebook users in subnational area by gender, age, device type), or gender gaps, using Facebook marketing API

- **‘Offline’ predictors:**
  - Satellite-derived nightlight data
  - Population density (WorldPop)
  - Subnational education index, income index, human development index (HDI), gender development index (GDI)

Methods

- Ensemble Superlearner: best weighted combination of ML algorithms
- To assess predictive performance, three validation strategies:
  - Conventional 10-fold cross-validation
  - Leave-one-country-out cross-validation (LOCO-CV)
  - External benchmarking against Living Standards Measurement Studies (LSMS) and Multiple Indicator Cluster Surveys (MICS)



Methods

- We use different model performance metrics, incl. R2, MAE

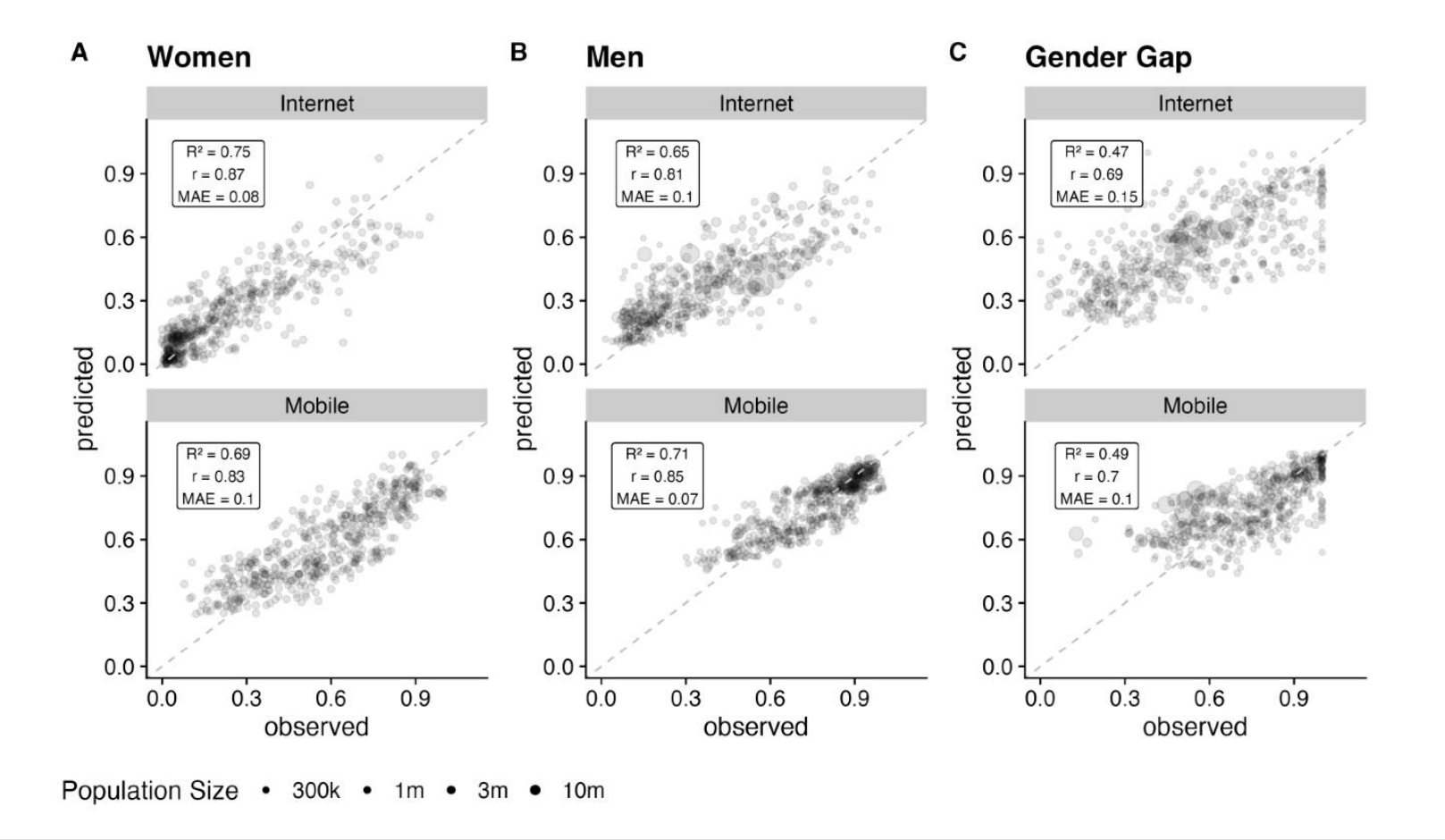
$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|.$$

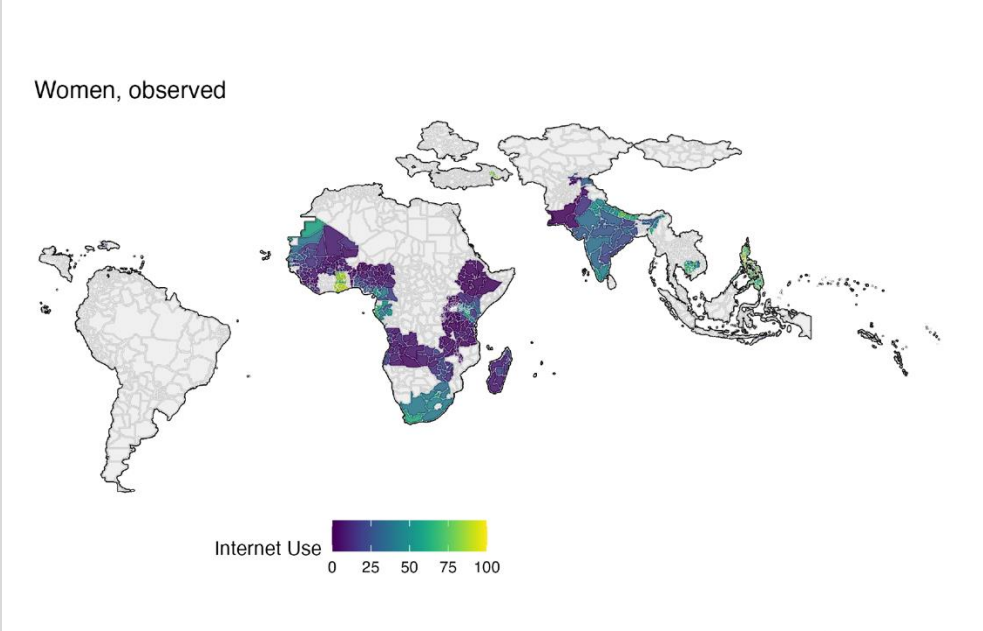
- 1 = perfect prediction
- 0 = mean

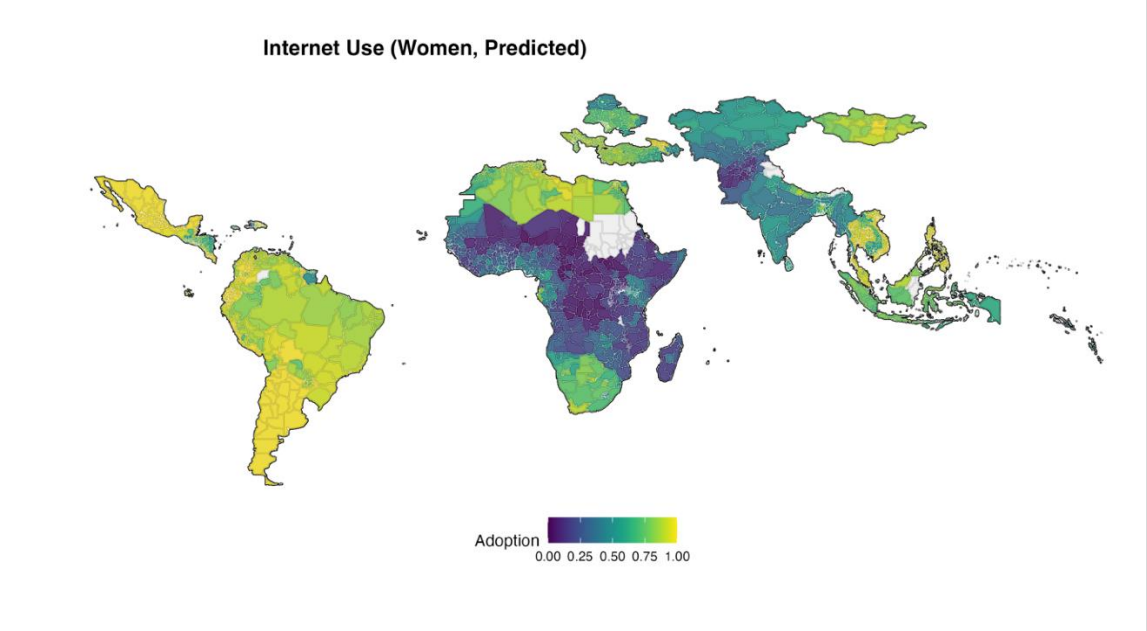
MAE provides an absolute measure of the average prediction error in dependent variable's units; lower values mean lower error

Overall predictive accuracy



Expanded geographical coverage





Improved temporal resolution

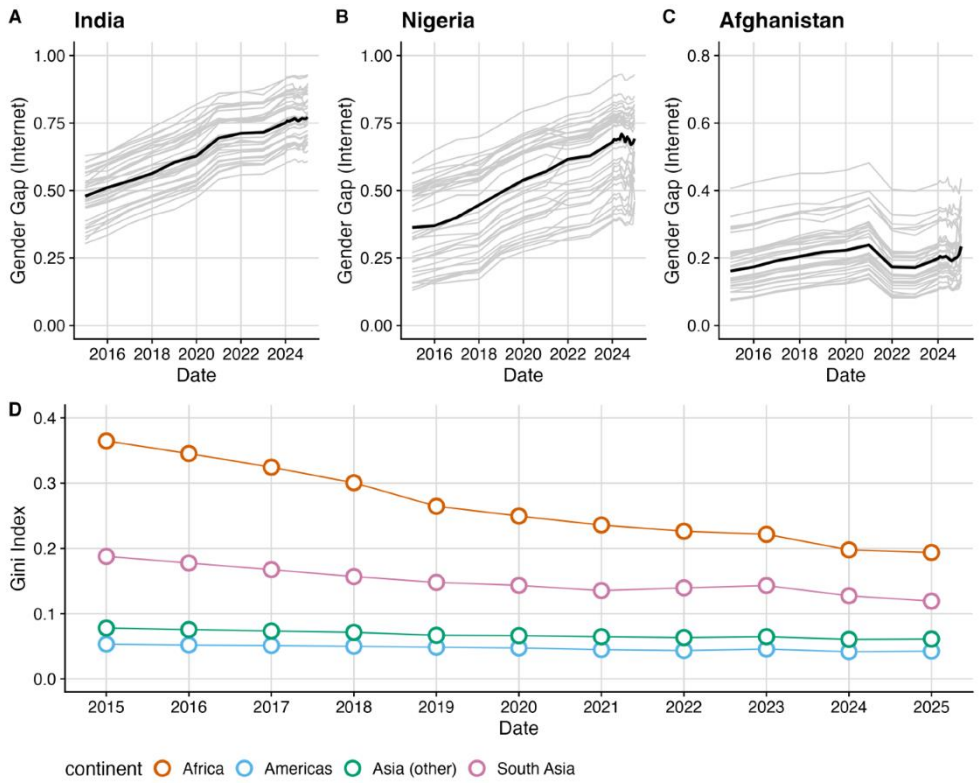


Figure 7: *(A–C)* Trends over time in India, Nigeria, and Afghanistan. The light grey lines show subnational change over time, and the dark shaded line shows the median trend for each country. *(D)* Change in within-country inequality for the internet gender gap, as measured by the subnational GINI index.

DGG Aim 2

- The Digital Gender Gaps (DGG) project **seeks to deepen data and evidence** on:

  (1) **Gender gaps in digital connectivity and behaviours** at national- and subnational-levels by integrating social media, geospatial data and population survey and census datasets;

  (2) **Impacts of digital information and capabilities** on women's empowerment and gender inequalities, through quasi-experimental methods, and/or geospatial augmentation of surveys

Technology impacts: prior pathways

- Mass media technologies can spread of new ideas and norms (Barber and Axinn 2004)
- Diffusion theories of fertility decline: technology can alter paths for social learning and social interaction
- TV: exposure to new information and behaviours – impact on fertility ideals, son preference and women's status in the household (e.g. Jensen and Oster 2009, La Ferrara et al 2012)
- Mass media can spread 'liberal' ideas associated with global elites (Pierotti 2013), but also reinforce gender stereotypes (Woods 1994)

Mobile phones and internet technologies

- Improved access to (private) information, as well as access to services (e.g. mHealth, mobile money)
- Better connectivity and access to networks; information (but also dis-information) sharing is amplified
- Exposure to the ‘life of others’ including more globalized media content
- New outlets for the perpetuation of gender-based violence (tech-facilitated GBV)

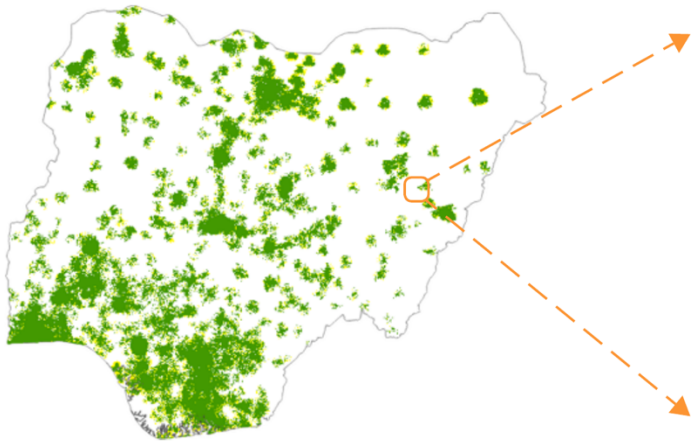
Ongoing Work

- Geographical linkages of Demographic and Health Surveys with
  - Mobile Coverage Maps to exploit geographical and temporal variations in the rollout of 3G internet technologies → impacts on fertility, SRH, IPV
  - Social connectedness (as measured via social media friendships) and sexual and reproductive health in Africa

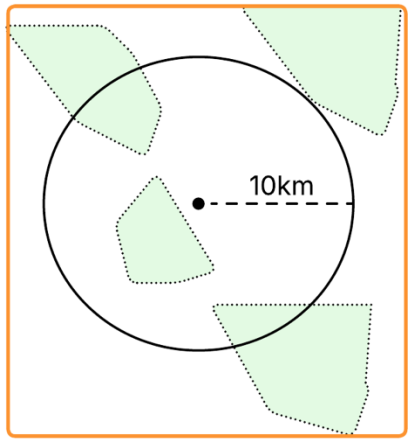
UNIVERSITY OF OXFORD

3G Expansion

Annual coverage maps (2010-2018)



Calculating 3G coverage in clusters



Combine with Nigeria 2018 DHS to create longitudinal panel

Woman ID	Year	Birth	3G Coverage	DHS Cluster	DHS Covariates
10 47 3	2010	0	0.000	001	X
10 47 3	2011	0	0.000	001	X
10 47 3	2012	1	0.000	001	X
10 47 3	2013	0	0.000	001	X
10 47 3	2014	1	0.716	001	X
10 47 3	2015	0	0.771	001	X
10 47 3	2016	0	0.780	001	X
10 47 3	2017	0	0.781	001	X
10 47 3	2018	0	0.916	001	X

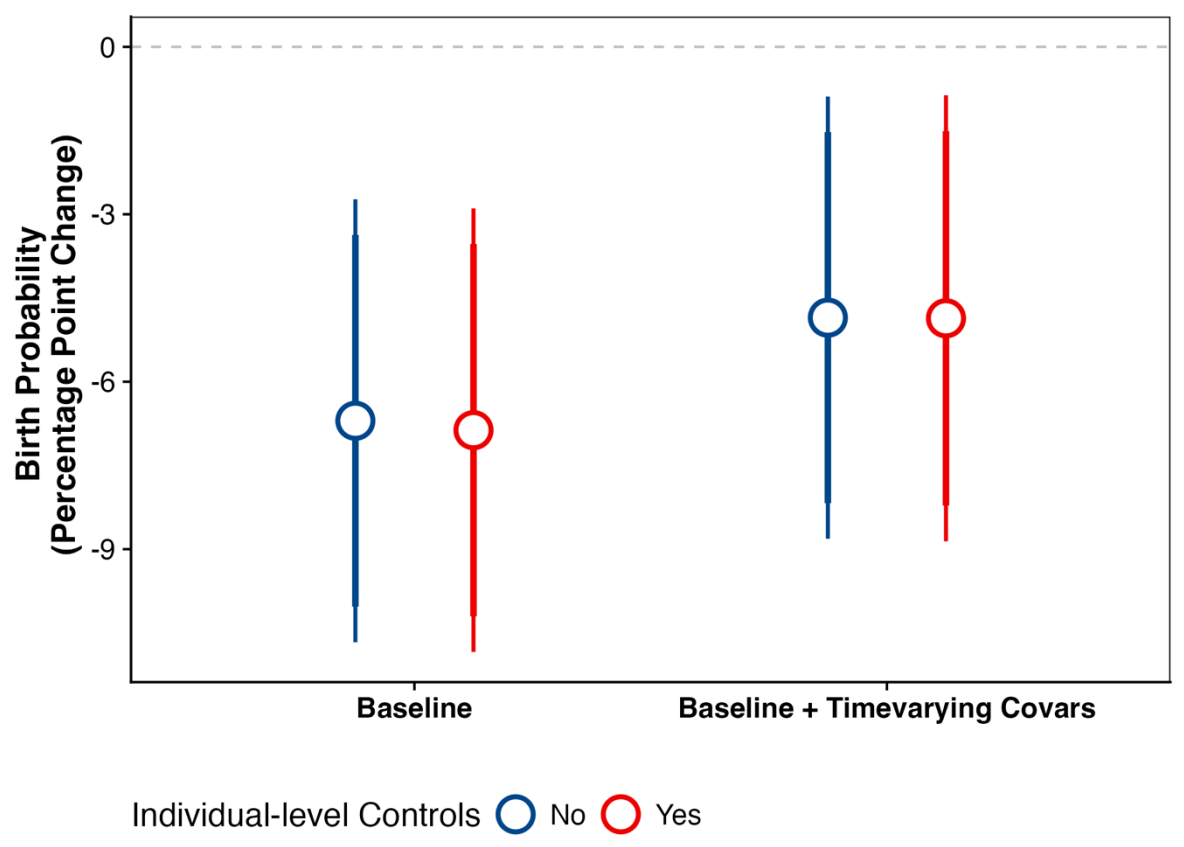
Note: only areas with DHS clusters are shown in the figure of 3G coverage

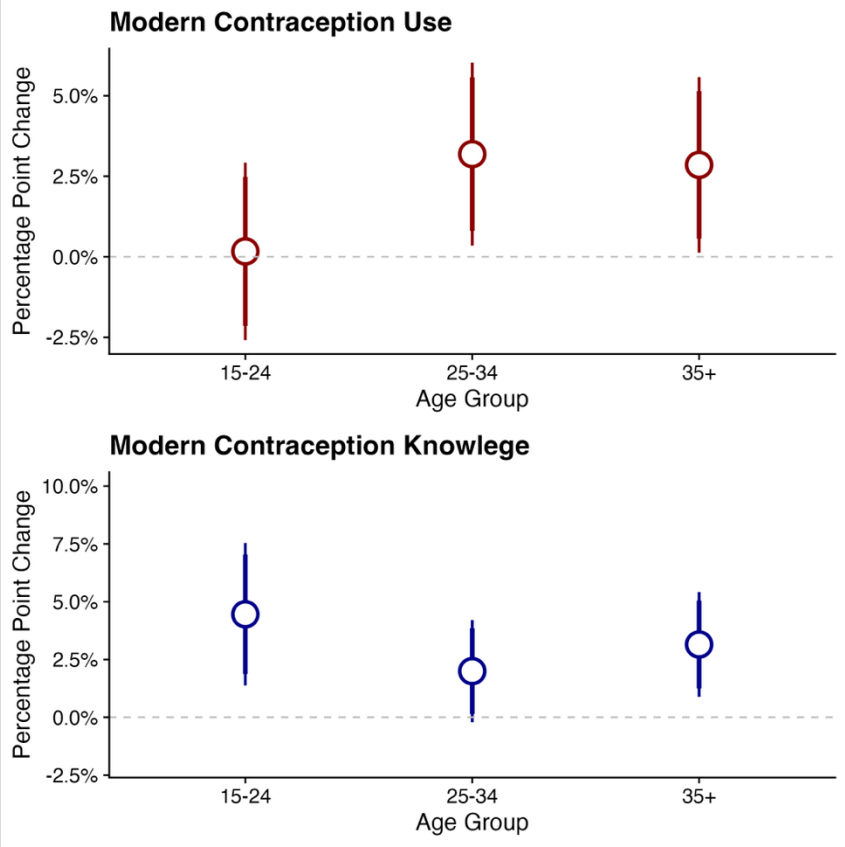
Data: longitudinal panel constructed by linking individual-level birth histories from DHS to annual mobile phone coverage maps

Fertility impacts of 3G Expansion: Evidence from Nigeria
(with Casey Breen, Till Koebe)

UNIVERSITY OF OXFORD

3G Expansion





Meta's Social Connectedness Index

- Builds on the number of friendship links on Facebook (Bailey et al 2018)
- Answers the question: "How likely is it that two Facebook users from separate locations are friends on Facebook?"

$$\text{SCI}_{a,b} = \frac{\text{FB\_Connections}_{a,b}}{\text{FB\_Users}_a \times \text{FB\_Users}_b},$$

$$\text{scaled\_SCI}_{a,b} = \frac{\text{SCI}_{a,b}}{\max_{i,j}(\text{SCI}_{i,j})},$$

- Looks at the ***edges***: social connectedness *between* locations
- Geographical granularity varies
  - US: Zip codes; Europe: NUTS3; Rest of the world: Admin-level 1
- Applications in high-income countries, e.g. Chetty et al (2022), Bailey et al (2021)
- We link SCI with Demographic and Health Surveys at subnational levels for Africa

Social capital mediates knowledge gaps in informing sexual and reproductive health behaviours across Africa, *Social Science and Medicine* (with Till Koebe, Theophilius Aidoo, Valentina Rotondi, Douglas Leasure and Ingmar Weber)

UNIVERSITY OF OXFORD

- Social connectedness between regions (measured using social media friendship links) and sexual and reproductive health, and women's empowerment using DHS

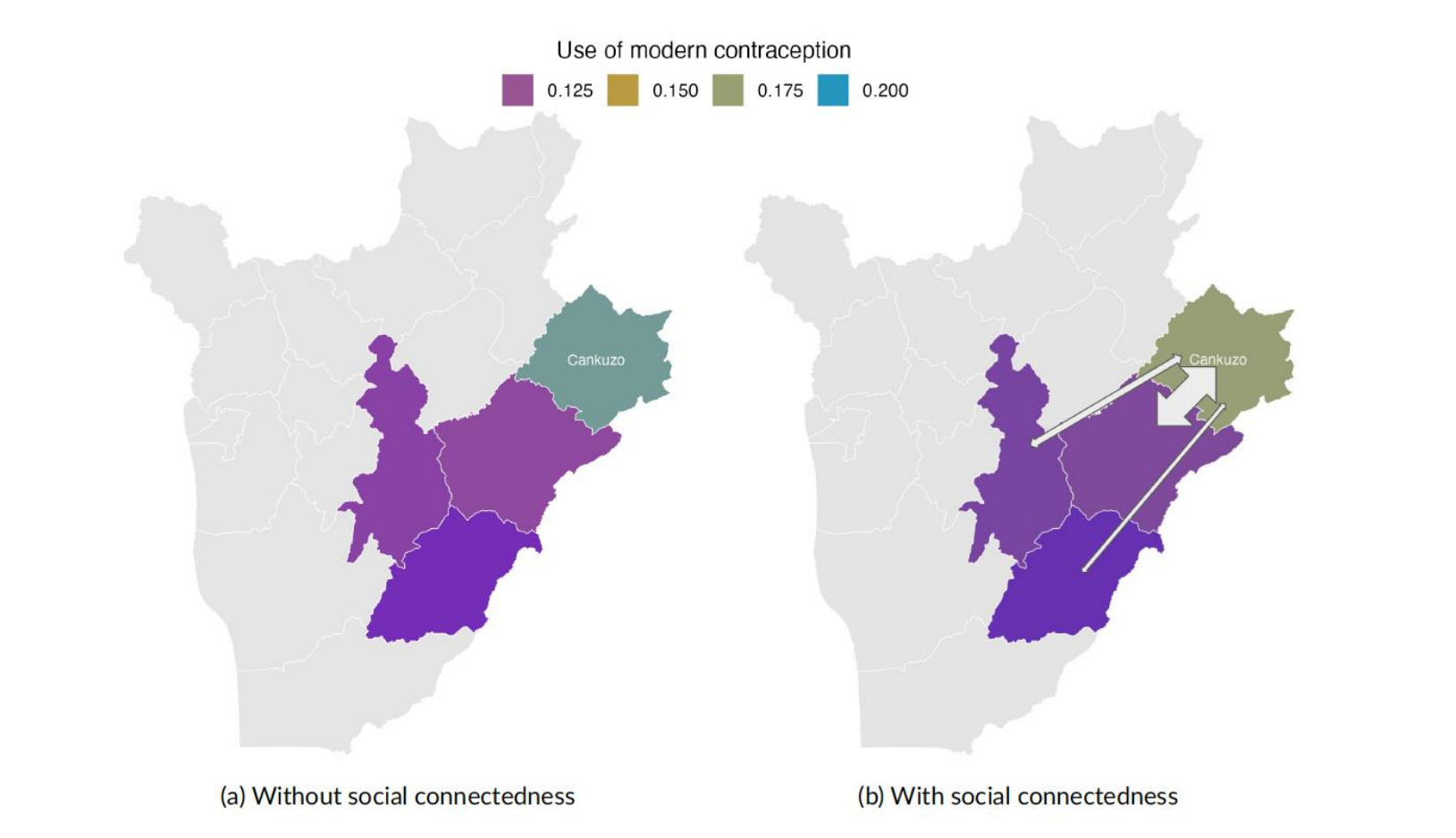
Regions that are more socially connected have smaller differences in health behaviours and knowledge

The mediating role of social connectedness becomes stronger, the larger the knowledge gaps between regions (communities *learn* from each other when they are connected)

	Dependent variable: Positive regional differences in...			
	...use of...		...knowledge about...	
	modern contraception	HIV tests	modern contraception	HIV
	(1)	(2)	(3)	(4)
	b/se	b/se	b/se	b/se
Constant	0.035***	0.037*	0.140***	0.062**
	(0.011)	(0.019)	(0.009)	(0.027)
SCI	−0.260***	−0.131***	−0.360***	−0.165***
	(0.040)	(0.027)	(0.050)	(0.032)
ΔContraceptive knowledge	0.217***			
	(0.011)			
SCI x ΔContraceptive knowledge	−14.395***			
	(5.257)			
ΔKnowledge about HIV		0.548***		
		(0.027)		
SCI x ΔKnowledge about HIV		−25.121***		
		(8.307)		
ΔControl variables (20)	Yes	Yes	Yes	Yes
Observations	122,760	122,760	122,760	122,760
R^2	0.726	0.932	0.697	0.932
Adjusted R^2	0.726	0.932	0.697	0.932

Note: OLS. SE clustered at the regional level. Country FE included. *p<0.1; **p<0.05; ***p<0.01

Social Connectedness and SRH



The effect of being socially connected: the predicted contraception adoption rates in Cankuzo, Burundi and the three regions Cankuzo is most socially connected to, respectively, with and without social connectedness

Summary

- The digital revolution has both **social** and **data** implications.
- Empirical evidence for impacts on gender inequalities (SDG 5), with the potential for largest payoffs at lower levels of economic development.
- Digital gender gaps are an important dimension of population inequality
- Social media marketing data streams can help fill data gaps on these inequalities

UNIVERSITY OF OXFORD

Thank you!

www.digitalgendergaps.org

We are grateful to funding from the Gates Foundation (INV-045370) and Leverhulme Trust for the Leverhulme Centre for Demographic Science (Grant RC-2018-003).