

The Digital Revolution and Demography: Perspectives from Digital and Computational Demography

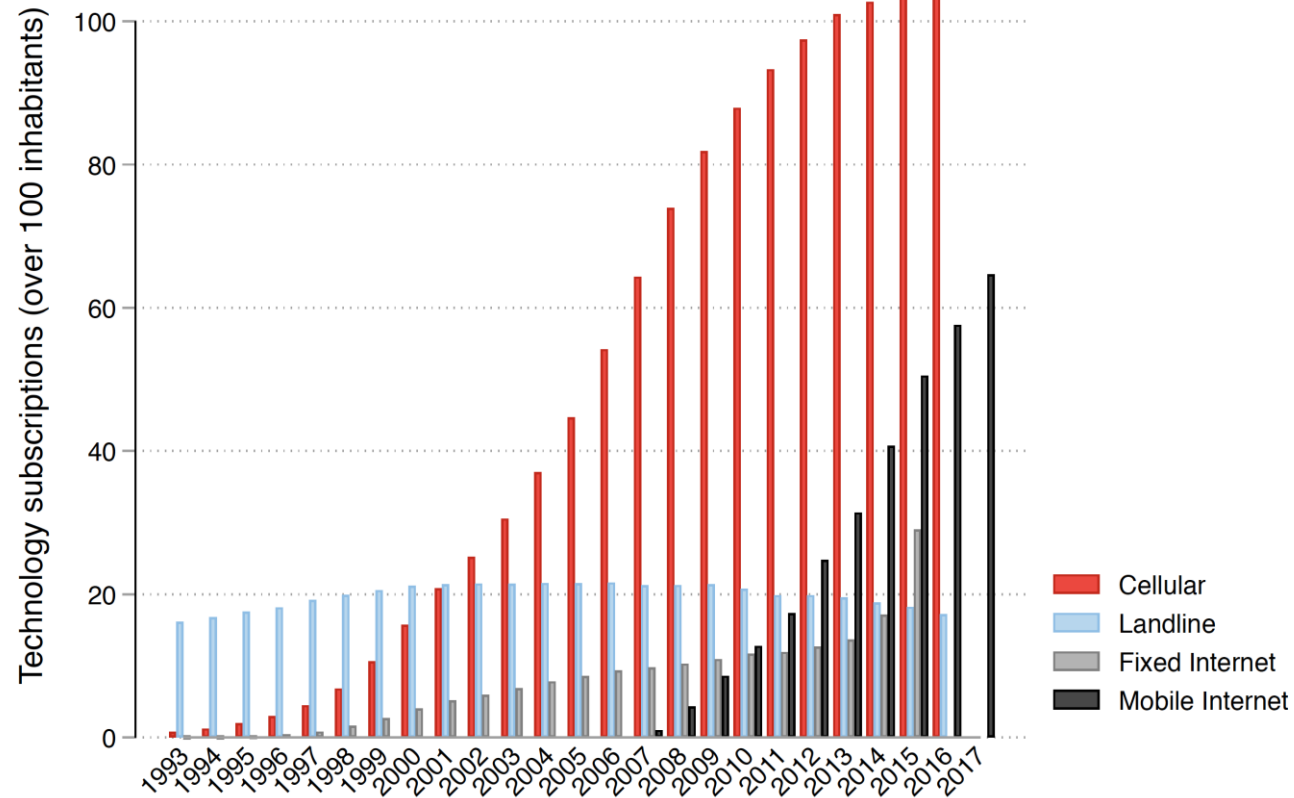
Ridhi Kashyap

Department of Sociology | Leverhulme Centre for Demographic Science | Nuffield College
University of Oxford

7th December, 2023

Wittgenstein Centre Conference

The Digital Revolution



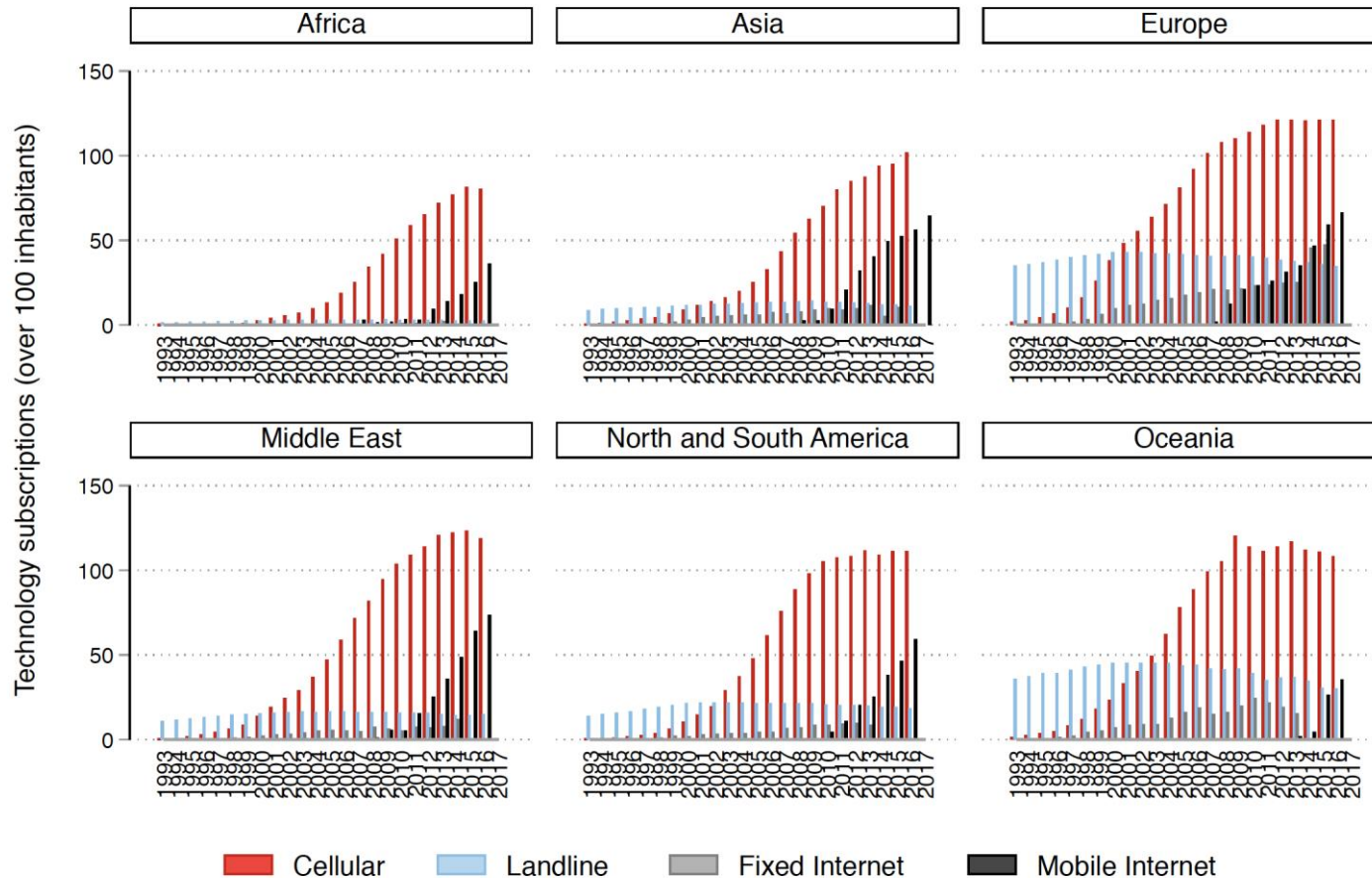
Technology subscriptions per capita, global, 1993-2017.

Data from International Telecommunications Union (ITU)

The Digital Revolution

Technology subscriptions per capita, by region, 1993-2017.

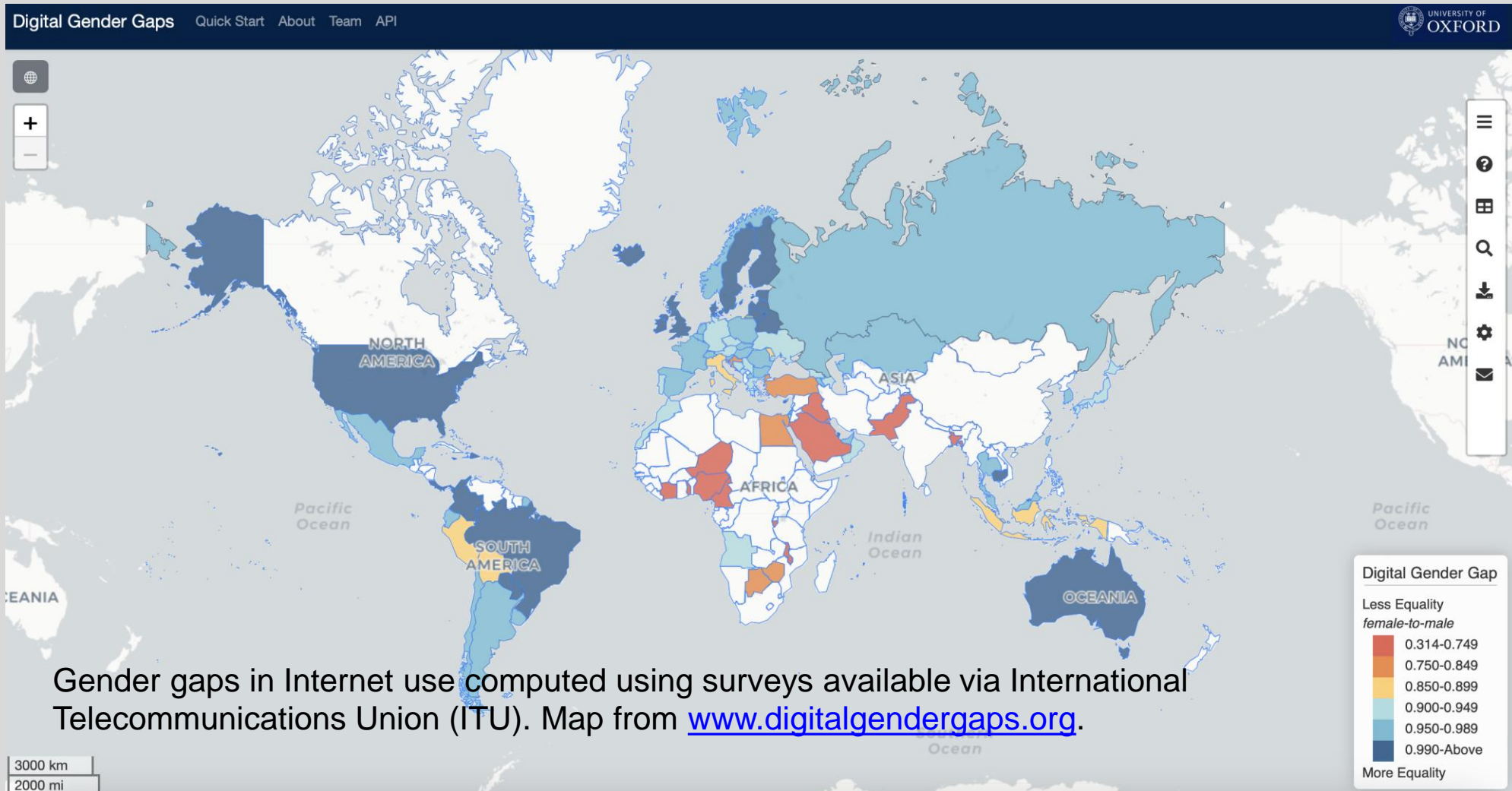
Data from International Telecommunications Union (ITU)



The Digital Revolution and Demography

- The digital revolution is:
 - **Social revolution:**
 - How – and for whom – are digital technologies affecting demographic behaviours and outcomes?
 - How do technologies (re)shape inequalities?
Technology has the *theoretical* potential to empower marginalized populations
 - SDG 5 on gender equality pledges to “enhance the use of...information and communication technologies to promote the empowerment of women.”
 - But is this potential being realized?

The Data Gap



The Digital Revolution and Demography

- The digital revolution is:
 - **Social revolution:**
How – and for whom – are digital technologies affecting demographic behaviours and outcomes (e.g. health and well-being, gender inequalities, family and fertility, migration)? How do technologies (re)shape inequalities?
 - **Data revolution:**
Our use of the web, social media, and mobile phones generates new data streams that have the potential to provide novel measures of digital adoption and behaviours

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.

The Digital Revolution and Demography

- The digital revolution is:
 - **Social revolution:**

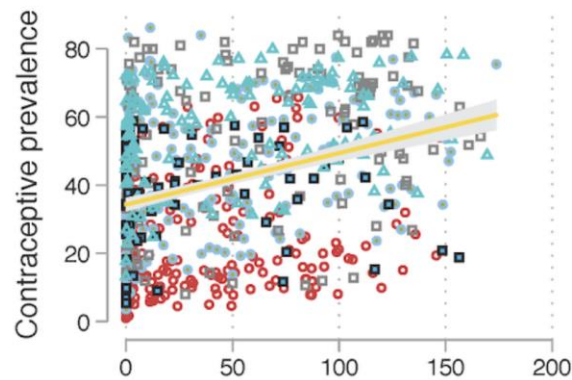
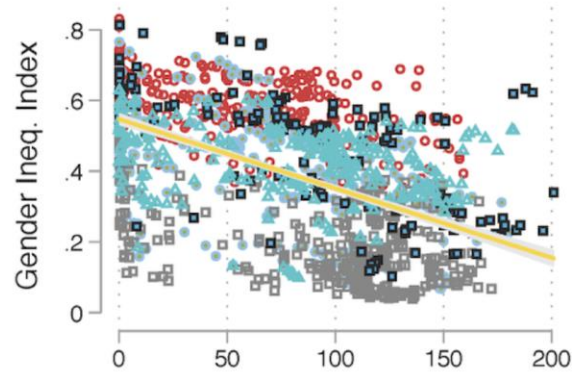
How – and for whom – are digital technologies affecting demographic behaviours and outcomes?

→ Digital expansion and sustainable development, especially SDG 5 on gender equality (The Digital Gender Gaps project, www.digitalgendergaps.org)
 - **Data revolution:**

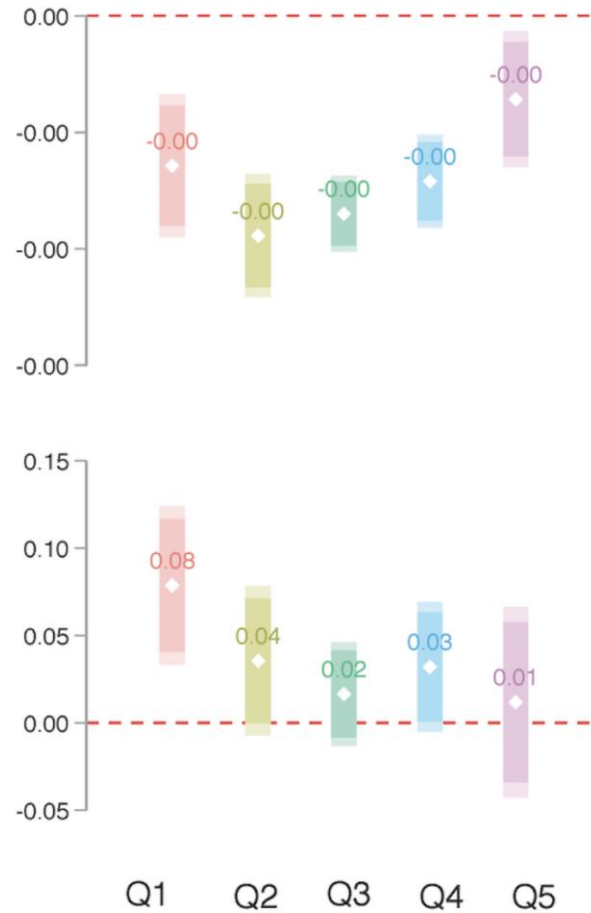
What are the demographic characteristics of online populations?
Can new data streams be used for population-generalizable measurement?

Technology impacts: prior pathways

- TV: exposure to new information and behaviours – shift in attitudes (e.g. Barber and Axinn 2004, Jensen and Oster 2009, La Ferrara et al 2012)
- Diffusion theories of fertility decline: technology can alter paths for social learning and social interaction
- **Mobile phones:**
 - Improved access to (private) information, as well as access to services (e.g. mHealth, mobile money)
 - Better connectivity and access to networks
 - Exposure to the ‘life of others’ including more globalized media content

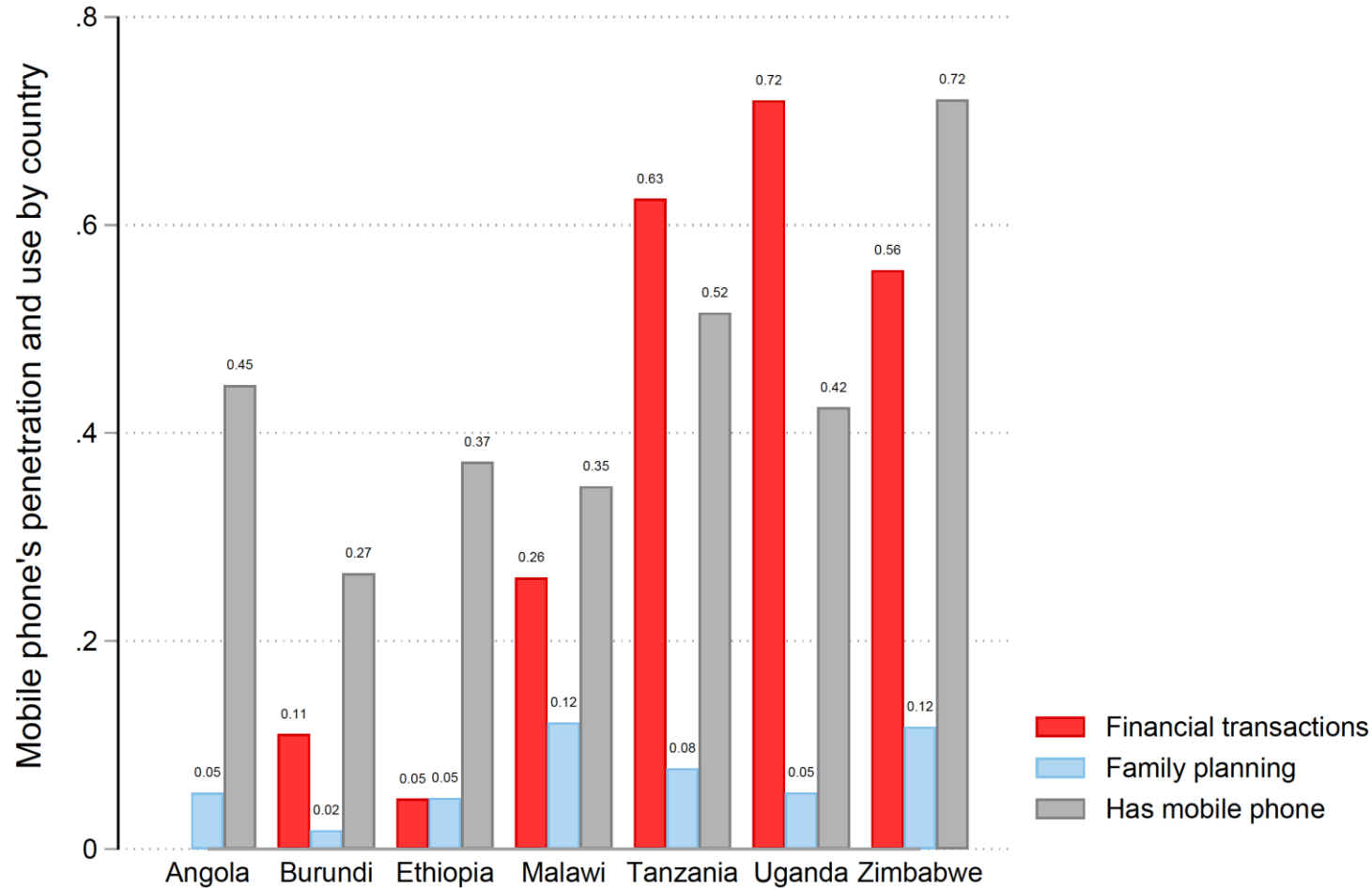


- Africa ● Asia ■ Europe
- Middle East ▲ North and South America ▲ Oceania

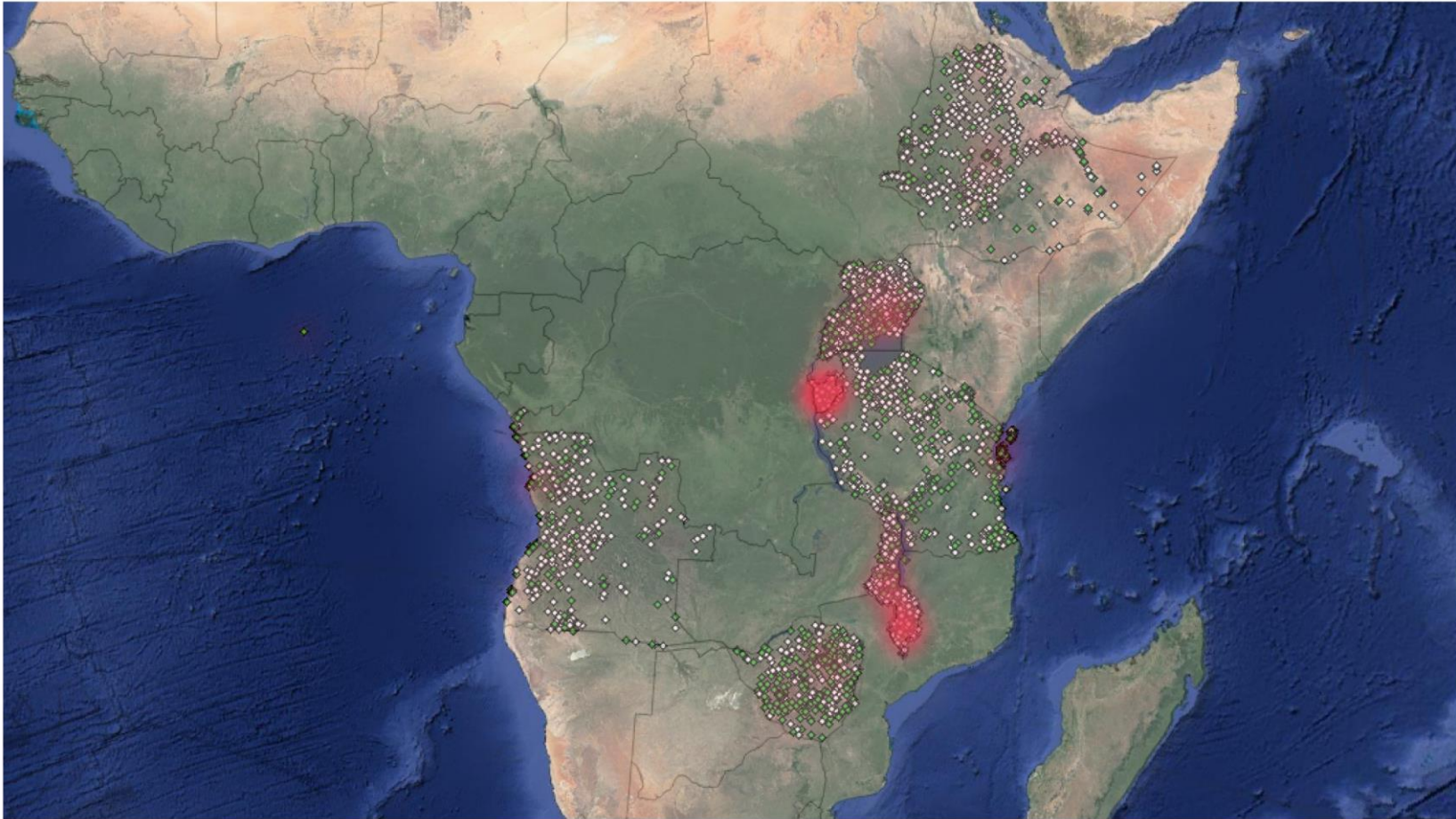


Global correlations between mobile phone diffusion and sustainable development outcomes (left-panel).

Standardized coefficients from models regressing mobile-phone diffusion on sustainable development outcomes by GDP-per-capita quintiles (right-panel).

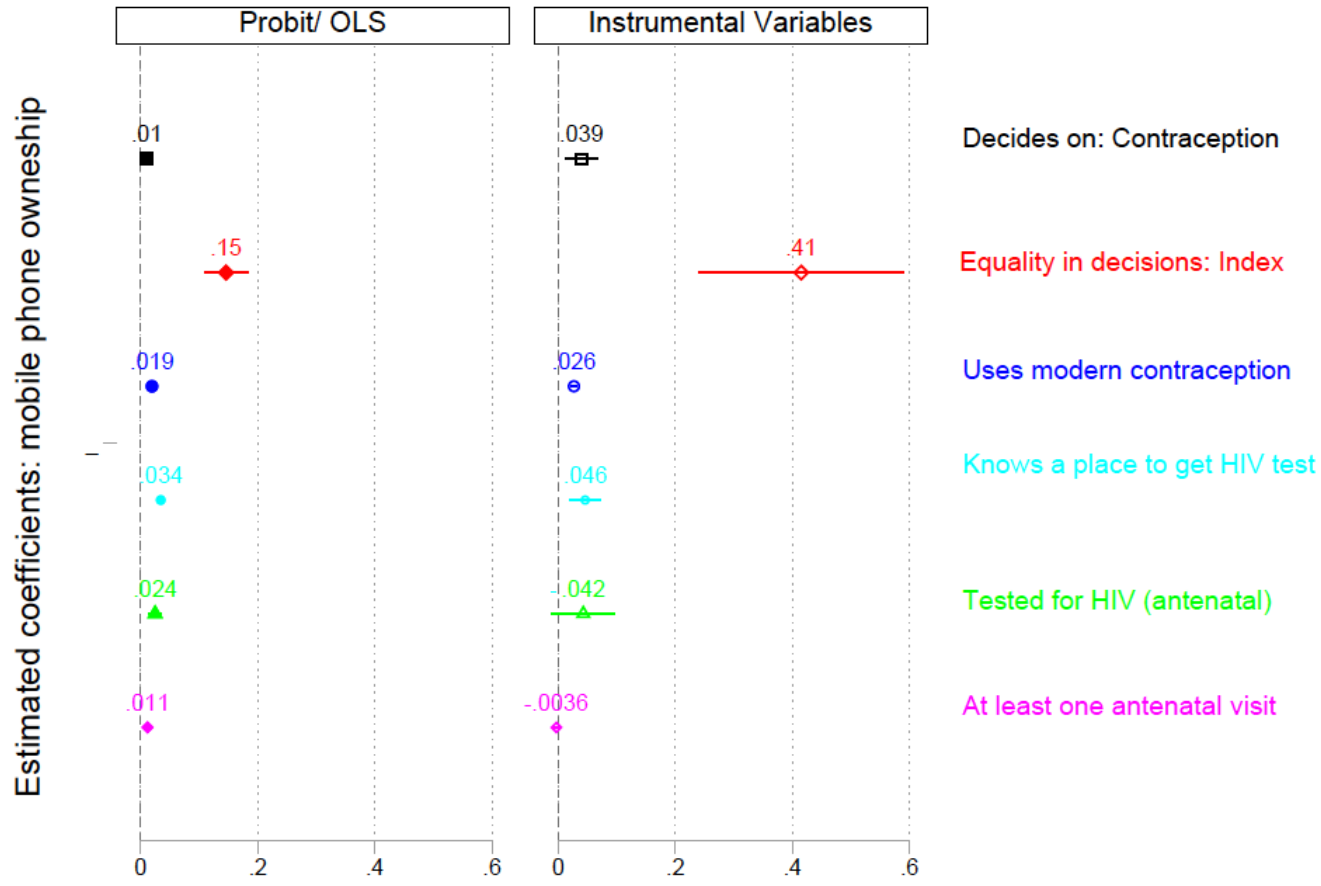


Mobile phone ownership and uses among women, 7 countries in Sub-Saharan Africa. Data from Demographic and Health Surveys.



The average number of lightning strikes (red) and mobile ownership (green dots) versus non-ownership (white dots). Source: authors' elaboration from augmented DHS data (2015-16)

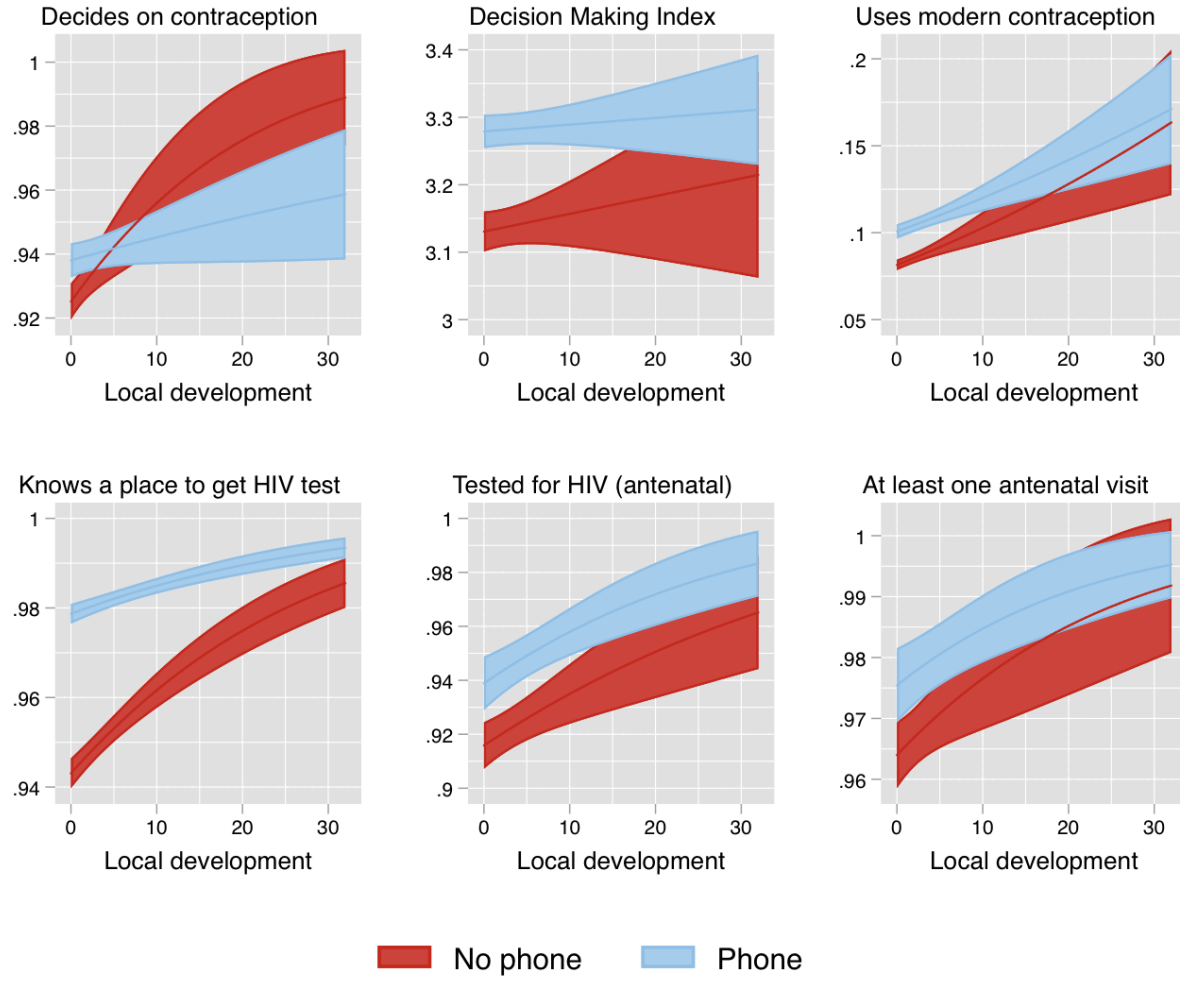
Lightning strikes as a quasi-experiment – poorer mobile connectivity in areas with more strikes due to poorer connectivity and broken down towers



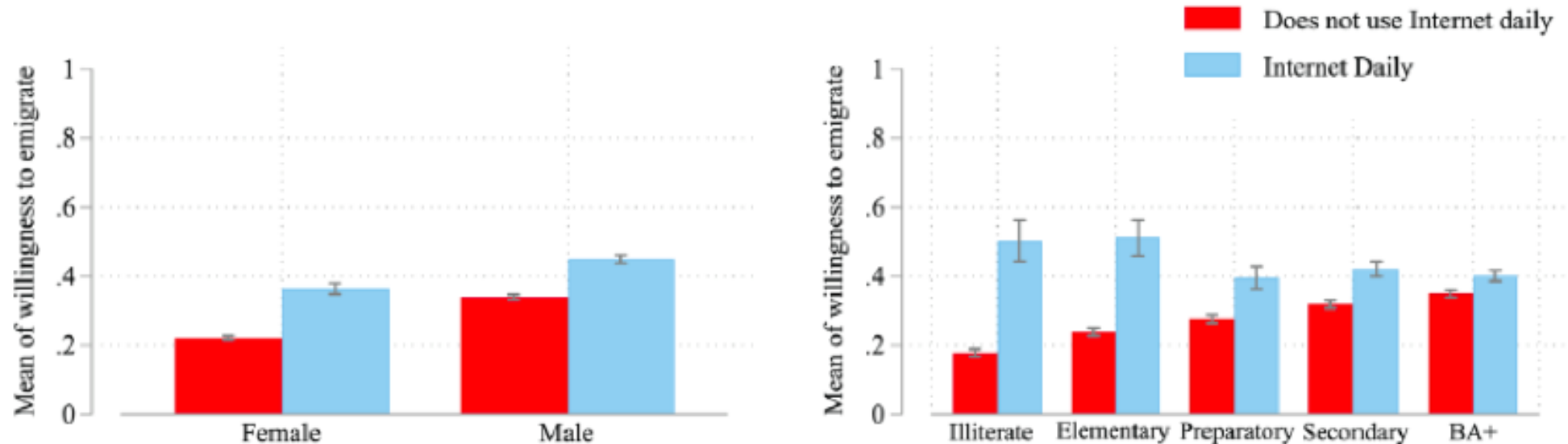
Individual-level effects of mobile phone ownership among women on sustainable development outcomes linked to women’s empowerment and sexual and reproductive health.

Covariates used in the models: education, age, household size, employment status, radio and tv ownership, urban, local development (nightlights). Country and year fixed effects. Standard errors clustered at the cluster level.

Leveraging mobile phones for sustainable development, *PNAS*
(with Valentina Rotondi, Luca Pesando, Simone Spinelli, Francesco Billari)



Marginal effects of the interaction between mobile-phone diffusion and local development (as proxied by nightlights).



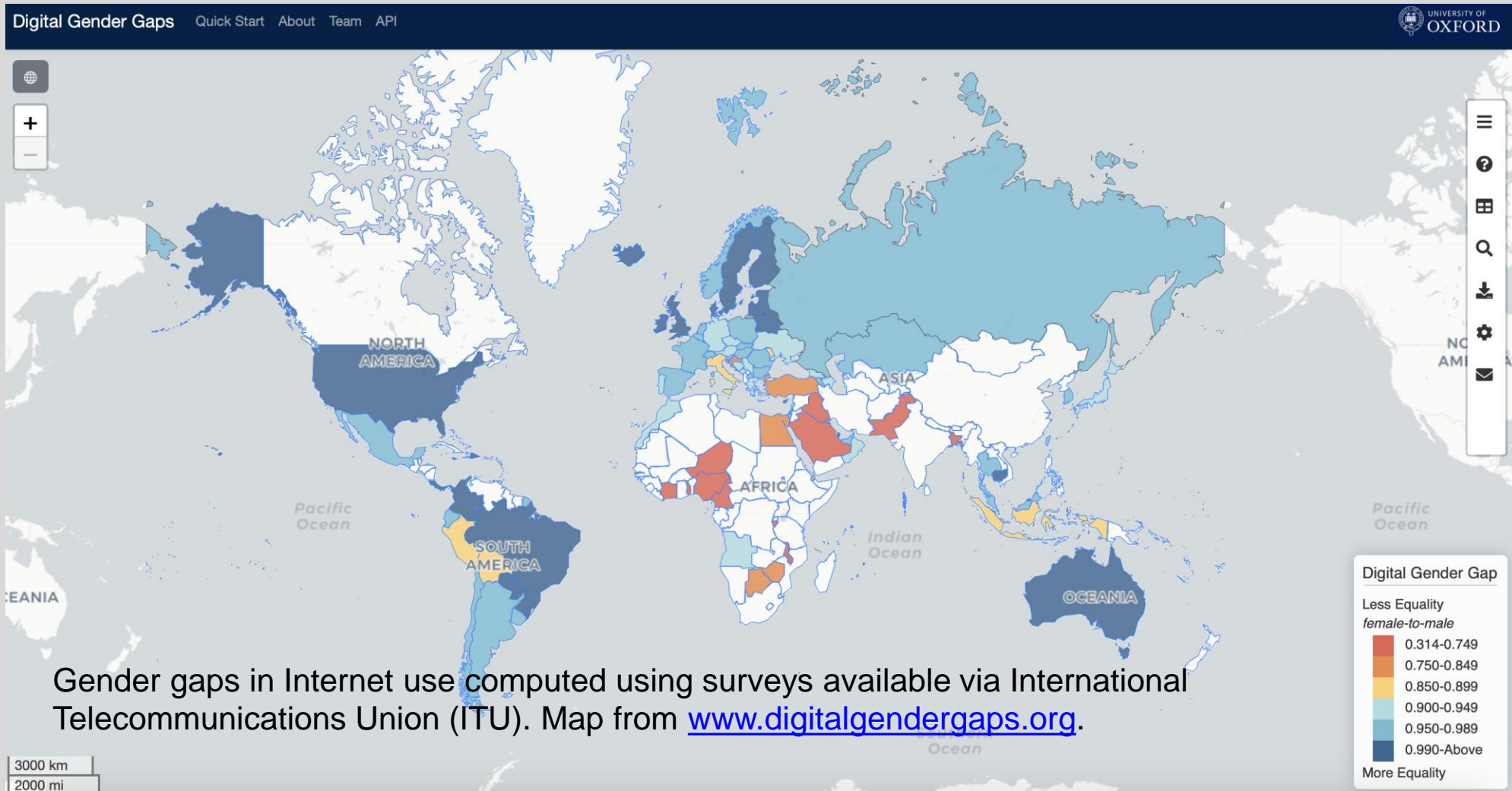
SOURCE: Arab Barometer.

Willingness to migrate by internet use and gender (left panel) and education (right panel)

Summary (1)

- Mobile phone ownership and internet adoption linked to improved women's empowerment, reproductive and sexual rights, and health outcomes
- Bigger pay-offs at lower levels of economic development
- **Implication:** important to track digital connectivity by gender

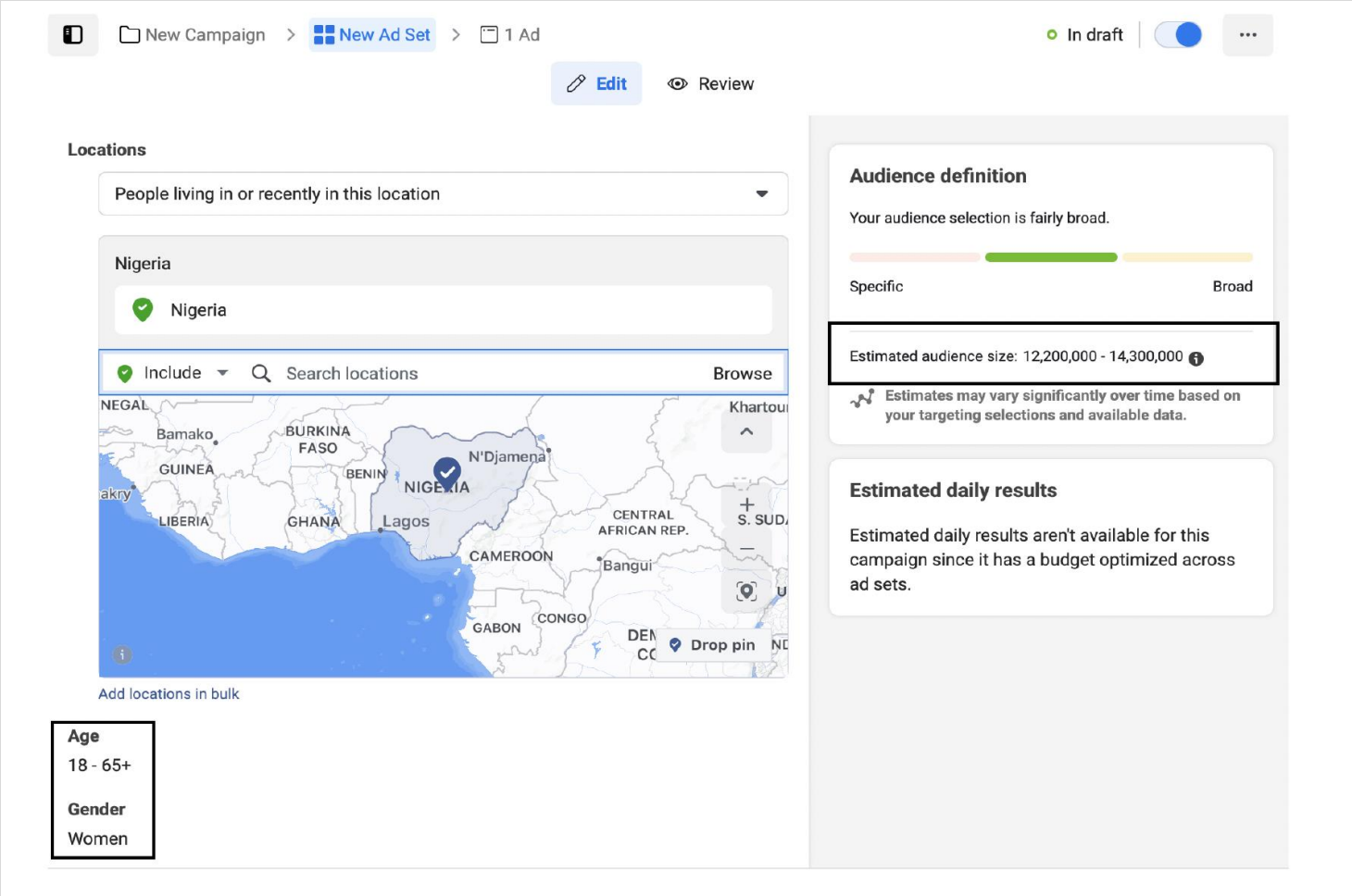
The Data Gap



The Digital Revolution

- The digital revolution is:
 - **Social revolution:**
How – and for whom – are digital technologies affecting demographic behaviours and outcomes?
 - **Data revolution:**
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

Social Media Ad Audience Estimates



The screenshot shows the Facebook Ads Manager interface for a new ad set. The location is set to Nigeria, and the audience definition is "People living in or recently in this location". The estimated audience size is 12,200,000 - 14,300,000. The audience definition is described as "fairly broad". The estimated daily results are not available because the campaign has a budget optimized across ad sets. The targeting criteria are set to Age: 18 - 65+ and Gender: Women.

Locations
People living in or recently in this location

Nigeria
Nigeria

Include Search locations Browse

Estimated audience size: 12,200,000 - 14,300,000

Estimated audience size: 12,200,000 - 14,300,000
Estimates may vary significantly over time based on your targeting selections and available data.

Estimated daily results
Estimated daily results aren't available for this campaign since it has a budget optimized across ad sets.

Age
18 - 65+

Gender
Women

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



Ad Audience Estimates

Use a matched audience (optional) Select
Custom targeting options to reach your website visitors, contacts, and target accounts.

Target by: the audience below

What location do you want to target? (required)

include Start typing a country, state, city, or town... See full list

include Netherlands

Target people who permanently live or work in the selected location(s).
Deliver ads to people who reside in the selected **location(s)** and are not recent visitors

What gender do you want to target?

All
 Female
 Male

Select specific targeting criteria to zero in on your ideal audience:

Company name	Company industry	Company size	Job title	Job function
Job seniority	Member schools	Fields of study	Degrees	Member skills

Your estimated target audience
4,500,000+ LinkedIn members

Netherlands

Male

Audience expansion: Enabled

LinkedIn tools may not be used to discriminate based on personal characteristics like gender, age, or actual or perceived race/ethnicity. [Learn more](#)

Applications

- Predicting SDG 5 indicators using gender gap indicators computed using social media ad data
 - Internet and mobile access gender gaps (www.digitalgendergaps.org, using Facebook)
 - Professional gender gaps (using LinkedIn)
- Expanding geographical coverage, more timely data on social indicators and for monitoring rapidly evolving situations

Limitations

- No past data; data access constraints
- Black box algorithms

Facebook Gender Gap Index

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

- Using ITU data, we compute a ground truth internet gender gap indicator derived from survey data:

- $$\textit{Internet GGI} = \frac{\% \textit{ of female population using the Internet}}{\% \textit{ of male population using the Internet}}$$

- Similarly, we compute Mobile GGI

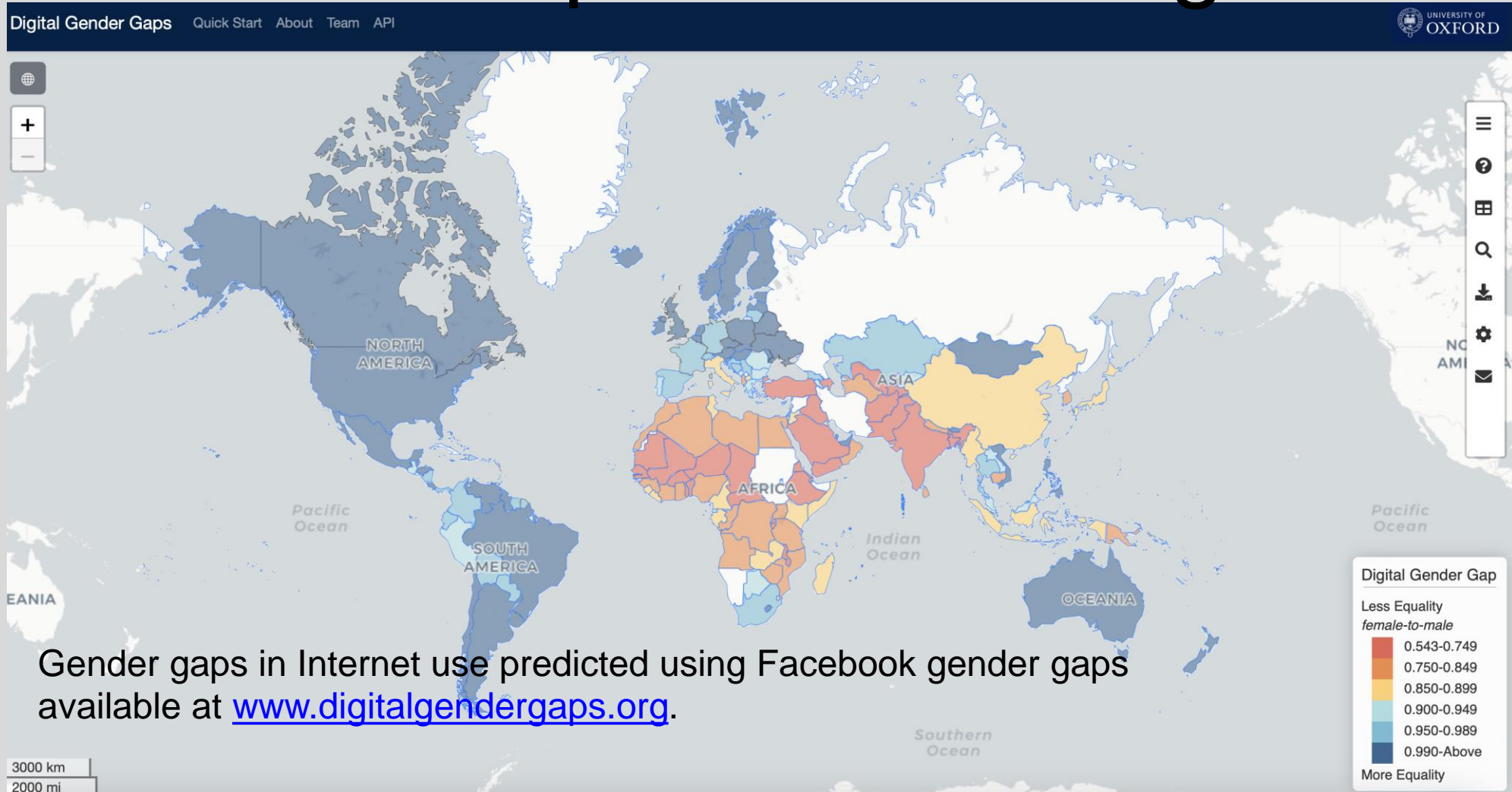
Facebook Gender Gap Index

FB GGI \approx Internet GGI

- Correlation of FB GGI 18+ with Internet GGI is 0.81
- Stronger than any other development indicator (including internet penetration)
- Correlation with Mobile GGI is 0.75
- We fit three types of predictive models:
 - online model
 - online-offline model
 - offline model

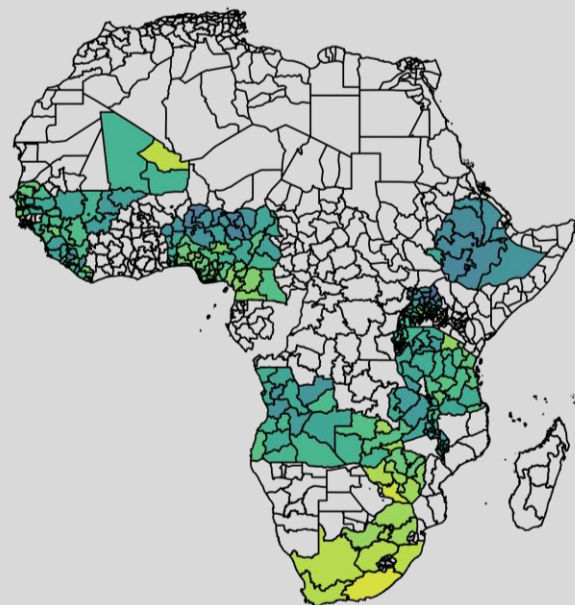
Using Facebook ad data to track the global digital gender gap.
World Development (joint work with Masoomali Fatehkia, Ingmar Weber)

Internet GGI predicted using FB GGI

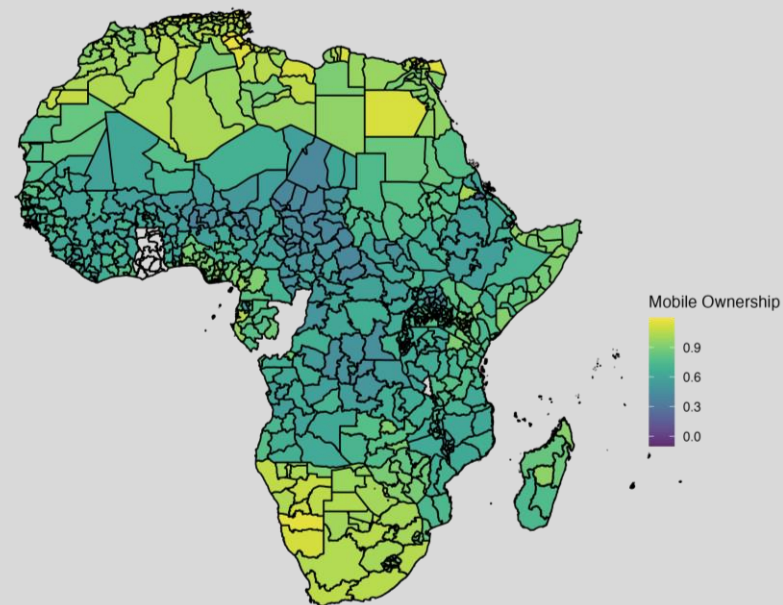


Next Output: Subnational Maps of Digital Gender Gaps (Africa)

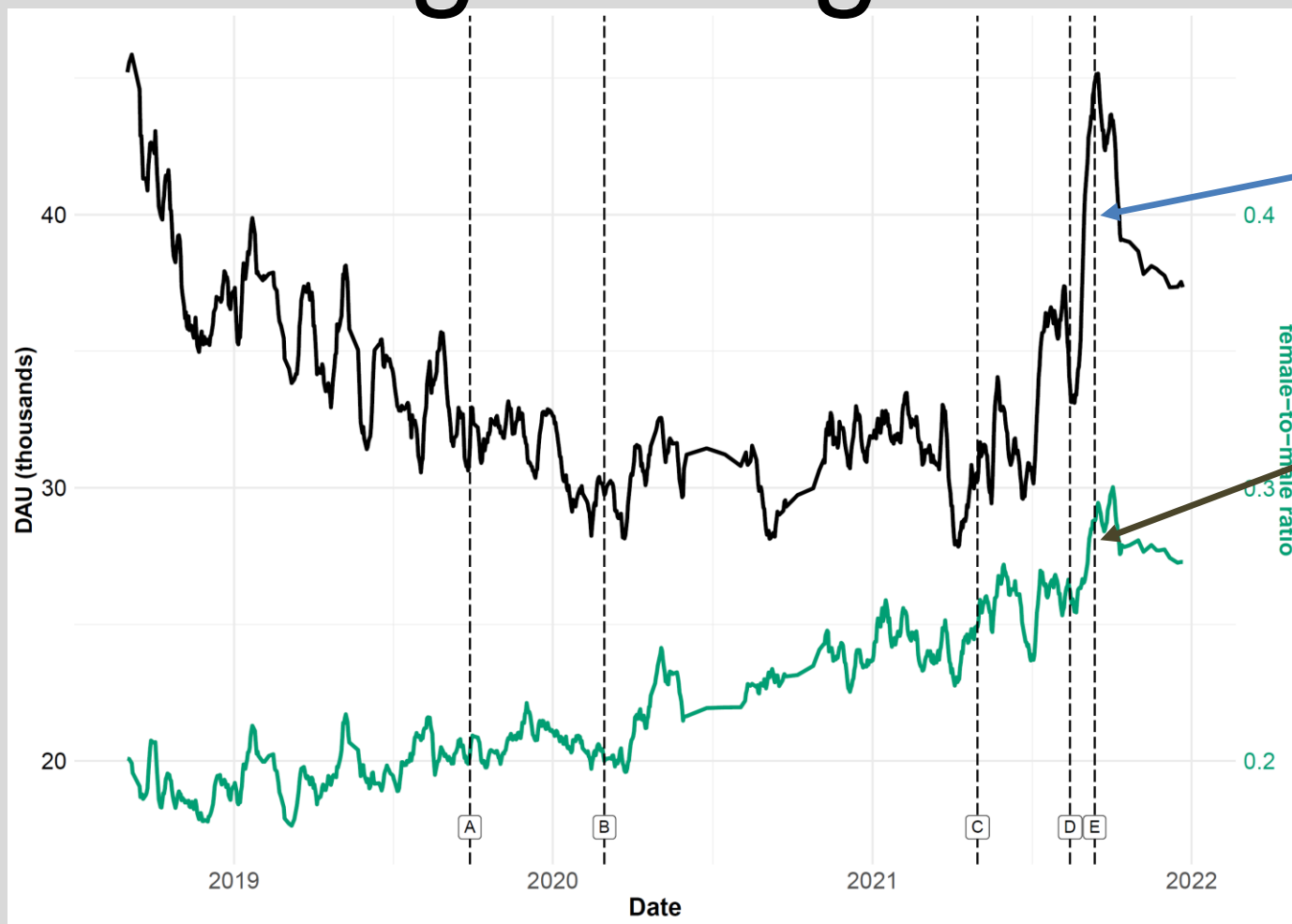
E Gender Gap (Female:Male Ratio)



F Gender Gap (Female:Male Ratio)



Tracking Changes



More teenage girls (15-19 year olds) are using Facebook in Afghanistan since Taliban offensive in Mid-2021

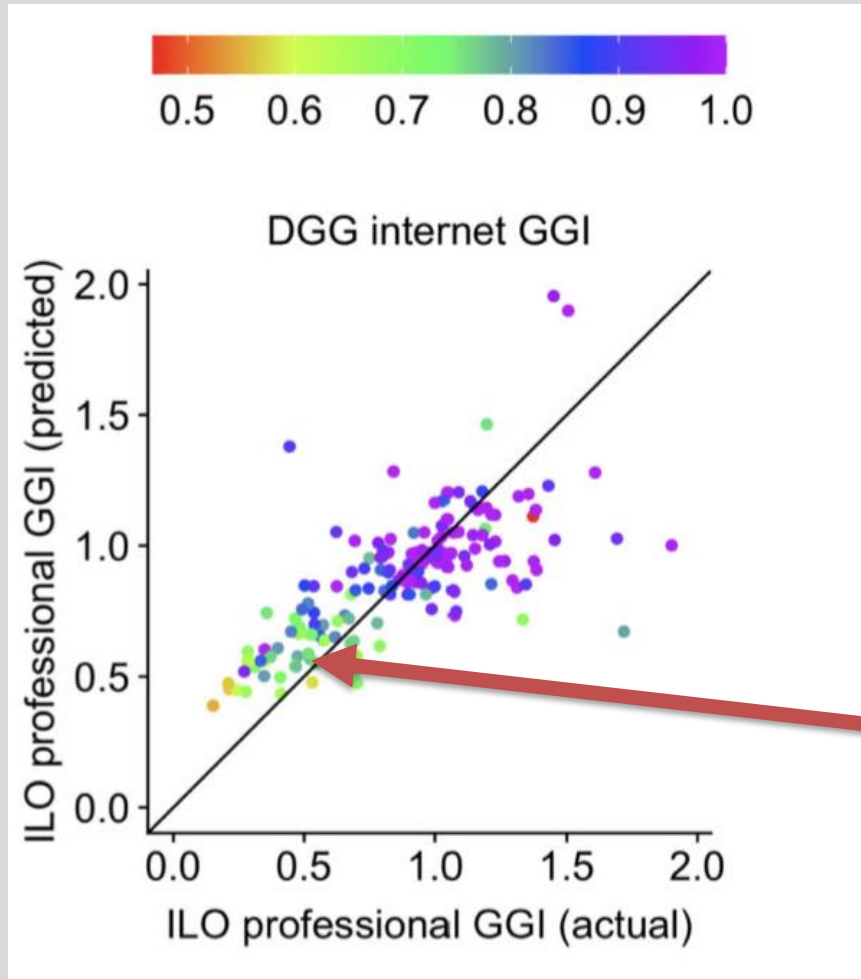
Girls are becoming a bigger share of teenage Facebook users

- [A] Afghanistan 2019 elections
- [B] Doha Agreement signed
- [C] Beginning of Taliban offensive
- [D] Fall of Kabul
- [E] #DoNotTouchMyClothes social media campaign

Selection and Bias

- Digital gender gaps are important to track themselves
- But they can also help illuminate biases in analyses using online samples/populations

Selection and Bias



In countries where women are less likely to be online, LinkedIn Gender Gaps over-predict professional gender equality

→ Women who are online are more selective

Summary

- The digital revolution has both **social** and **data** implications.
- Understanding differential **impacts of digitalization** is vital to understanding demographic and sustainable development dynamics.
- Harnessing the data revolution requires both ‘old’ and ‘new’ data sources, and their linkage, within a **diversifying data ecosystem**
- Demographic analyses of digital populations can help understand biases and generalizability – but important for illuminating population inequalities in new ways.
- **Digital and computational demography** has a vital role to play in advancing an **global vision of computational social science**.

Thank you!

References

- Rotondi, V., Kashyap, R., Pesando, L. M., Spinelli, S., & Billari, F. C. (2020). Leveraging mobile phones to attain sustainable development. *Proceedings of the National Academy of Sciences*, 117(24), 13413-13420.
- Fatehkia, M., Kashyap, R., & Weber, I. (2018). Using Facebook ad data to track the global digital gender gap. *World Development*, 107, 189-209.
- Kashyap, R., 2021. Has demography witnessed a data revolution? Promises and pitfalls of a changing data ecosystem. *Population Studies*, 75(sup1), pp.47-75.
- Kashyap, R., Fatehkia, M., Al Tamime, R., & Weber, I. (2020). Monitoring global digital gender inequality using the online populations of Facebook and Google. *Demographic Research*, 43, 779-816.
- Kashyap, R., and Verkroost, F.C.J. (2021) Analysing global professional gender gaps using LinkedIn advertising data. *EPJ Data Science* 10, no. 1, 39.
- Pesando, L. M., Rotondi, V., Stranges, M., Kashyap, R., & Billari, F. C. (2021). The internetization of international migration. *Population and development review*, 47(1), 79-111.

Thank you!

Learn more at:

www.digitalgendergaps.org

Contact:

ridhi.kashyap@nuffield.ox.ac.uk

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).

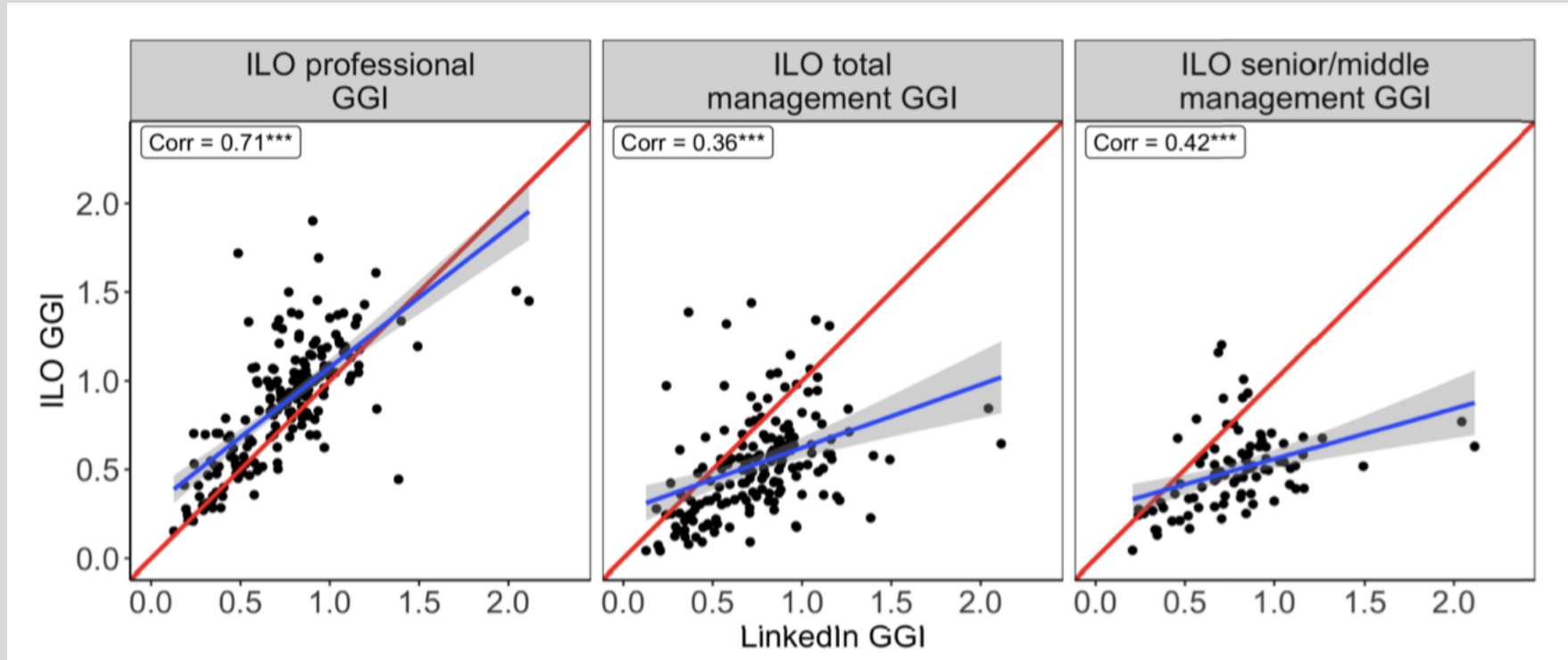
Social Media Ad Data Streams

- Social media ad data serve as **digital censuses** that provide aggregated demographic data on user populations
 - Expanding geographical coverage, filling data gaps
 - Finer temporal and spatial resolution
 - Valuable for rapidly evolving situations or for monitoring social indicators
 - Good complement to survey data, but not substitute

Limitations

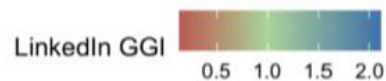
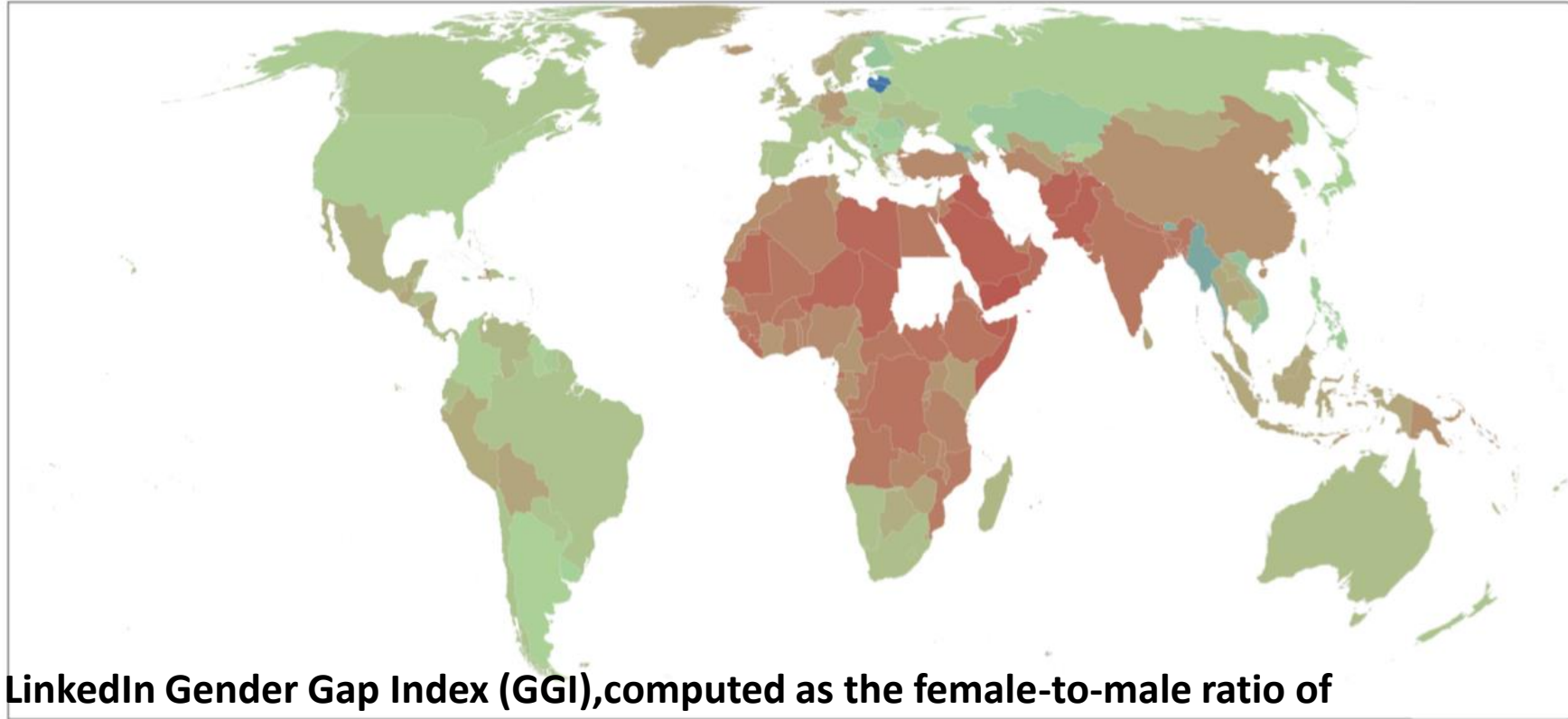
- No past data
- Black box algorithms

Professional Gender Gaps



LinkedIn Gender Gap Index (GGI) and its correlations with different International Labour Organization (ILO) Gender Gaps

LinkedIn Gender Gap Index



Predictive Fit

Table 3

Summary of results for three regression models predicting ITU internet Gender Gap Index using (i) a single online Facebook variable; (ii) online and offline variables; (iii) offline variables. Bootstrap estimates of the coefficient standard errors are reported in parentheses.

	Online Model	Onl.-Offl. Model	Offline Model
Intercept	0.933*** (0.006)	0.932*** (0.005)	0.933*** (0.007)
FB GG (age 18+)	0.071*** (0.011)	0.093*** (0.017)	
log(GDP per capita)		0.018* (0.008)	
GGGR – Literacy		–0.018 (0.016)	
GGGR – Education		–0.019 (0.019)	
Internet Penetration			0.040*** (0.009)
GGGR – Tertiary Educ.			0.032 (0.021)
GGGR – Economy			0.043** (0.014)
GGGR Score			–0.024 (0.012)
Adjusted R-squared	0.691	0.791	0.615
Mean Abs. Error	0.0325	0.0288	0.037
SMAPE	3.92%	3.90%	4.97%
F-statistics	169	67.38	29.79
df	74	66	68
N	76	71	73
# predicted countries ^a	152	127	132

*** p < 0.001, ** p < 0.01, * p < 0.05.

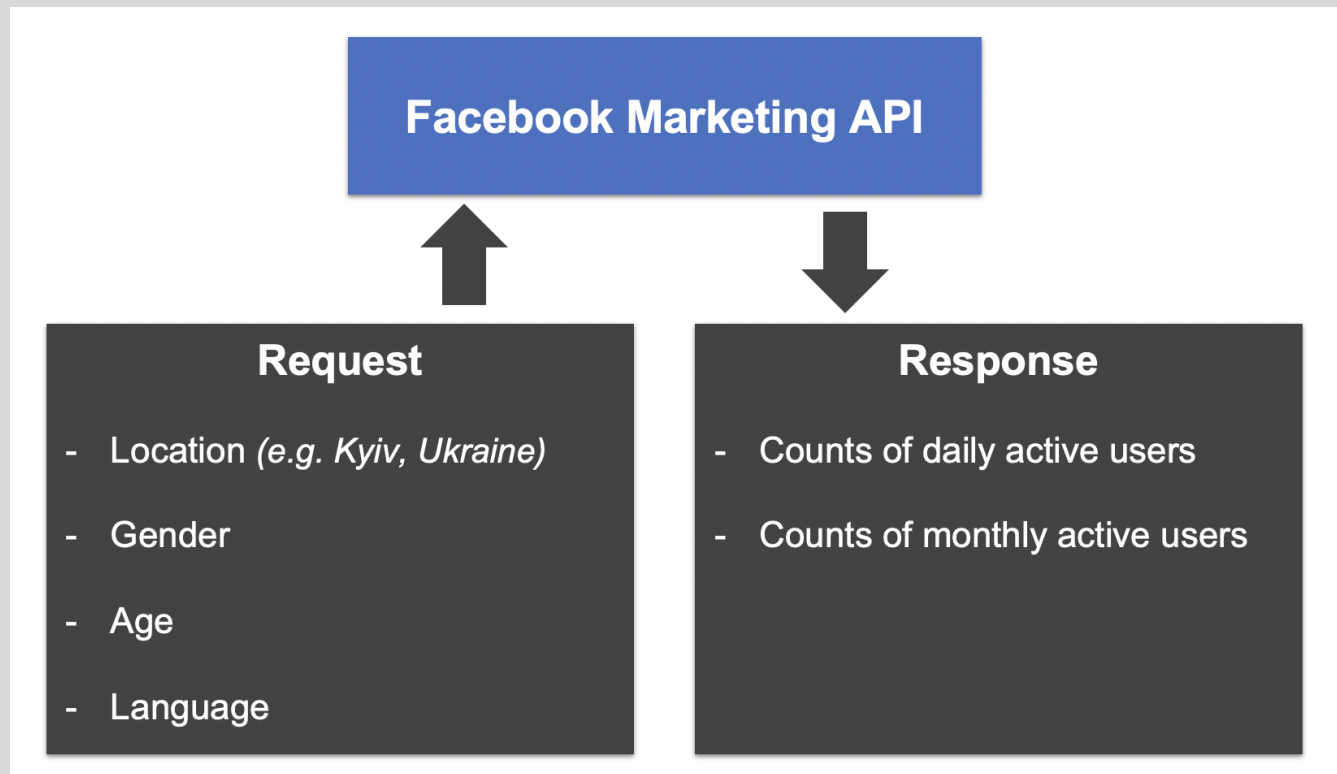
First stage regression

	(1) Flash rates	(2) Coverage	(3) Flash+Coverage
Gray Flash rates (log)	-0.013*** (0.002)		-0.013*** (0.002)
Cell phone coverage		0.017*** (0.004)	0.014*** (0.004)
<i>F-stat</i>	19.96***	7.99***	18.15***
<i>Hansen J statistic</i>			0.484

Note: First stage. Instruments: Flash rates (columns 1, 3), cellphone network (column 2,3). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Further applications

- Monitoring population displacement during Ukraine war through the same social media marketing database
- Help fill critical data gaps to inform humanitarian response



From FB estimates to Population Counts

Population Estimation

*30-34 year old women
in Kyiv before the conflict*

Baseline population = 100

Baseline Facebook users = 25

Baseline Facebook penetration = 25%



From FB estimates to Population Counts

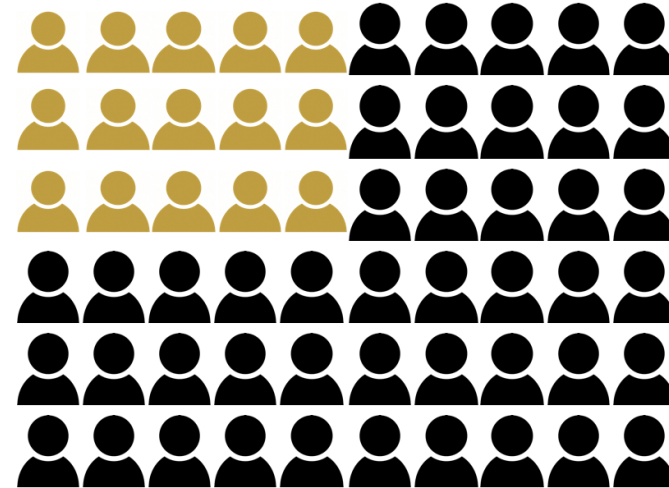
Population Estimation

*How many 30-34 year old women
are in Kyiv **today**?*

Facebook users = 15

Baseline Facebook penetration = 25%

Current population = $15 / 0.25 = 60$



POPULATION = USERS / PENETRATION

Daily net population change per Oblast

