Digital Gender Gaps Project: Measuring the Digital Divide and Explaining its Demographic Consequences

Max Planck Institute for Demographic Research

Casey F. Breen[1]

May 22, 2025

Project 1: Masoomali Fatehkia (Qatar Computing), Ingmar Weber (Saarland), Jiani Yan (Oxford), Xinyi Zhao (Oxford), Doug Leasure (Oxford), Ridhi Kashyap (Oxford)
Project 2: Till Koebe (Univ. of Saarland), Ridhi Kashyap (Oxford)

[1]University of Oxford

Roadmap for talk

1. **Mapping** trends in subnational digital gender inequality
2. **Impacts** of 3G internet expansion on fertility

Benefits of digital revolution

- The **digital revolution** has ushered in tremendous societal and economic benefits

Benefits of digital revolution

- The **digital revolution** has ushered in tremendous societal and economic benefits
  - Lower gender inequality, lower maternal/child mortality, higher contraception (Rotondi et al., 2020)

Benefits of digital revolution

- The **digital revolution** has ushered in tremendous societal and economic benefits
  - Lower gender inequality, lower maternal/child mortality, higher contraception (Rotondi et al., 2020)
  - Boost social connectivity, social learning, access to vital services (Unwin, 2009; DiMaggio and Hargittai, 2001; Suri and Jack, 2016)

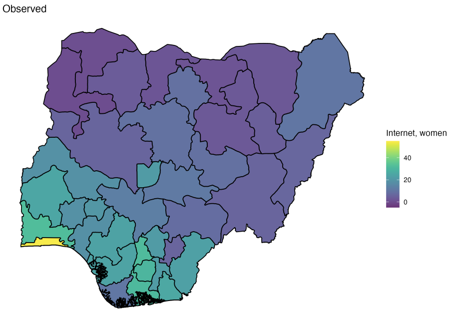
Benefits of digital revolution

- The **digital revolution** has ushered in tremendous societal and economic benefits
  - Lower gender inequality, lower maternal/child mortality, higher contraception (Rotondi et al., 2020)
  - Boost social connectivity, social learning, access to vital services (Unwin, 2009; DiMaggio and Hargittai, 2001; Suri and Jack, 2016)
  - Increases levels of education, economic benefits (Hjort and Poulsen, 2019; Kho, Lakdawala and Nakasone, 2018; Kharisma, 2022)

Benefits of digital revolution

- The **digital revolution** has ushered in tremendous societal and economic benefits
  - Lower gender inequality, lower maternal/child mortality, higher contraception (Rotondi et al., 2020)
  - Boost social connectivity, social learning, access to vital services (Unwin, 2009; DiMaggio and Hargittai, 2001; Suri and Jack, 2016)
  - Increases levels of education, economic benefits (Hjort and Poulsen, 2019; Kho, Lakdawala and Nakasone, 2018; Kharisma, 2022)
- Large **inequality** in who has access to digital technology

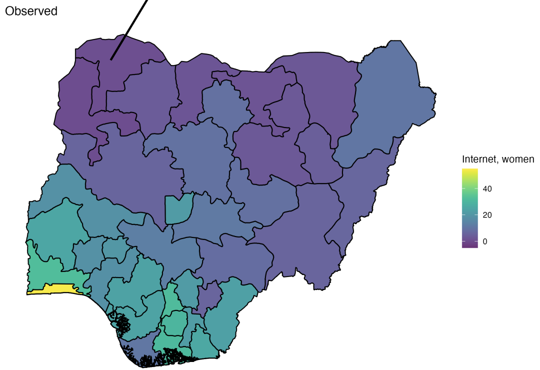
Adoption of digital technology varies geographically



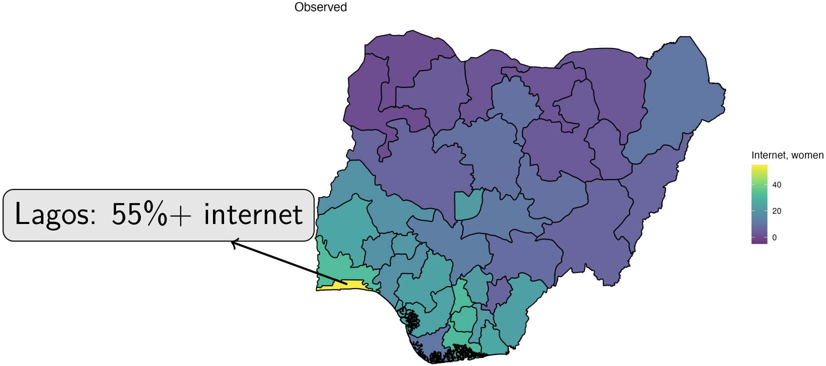
Source: Nigeria, Demographic and Health Survey

Women using internet, past 12 months

Sokoto: 2% internet



Women using internet, past 12 months

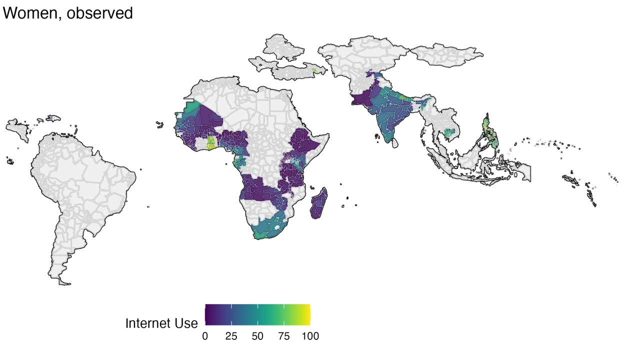


Develop subnational estimates of adoption

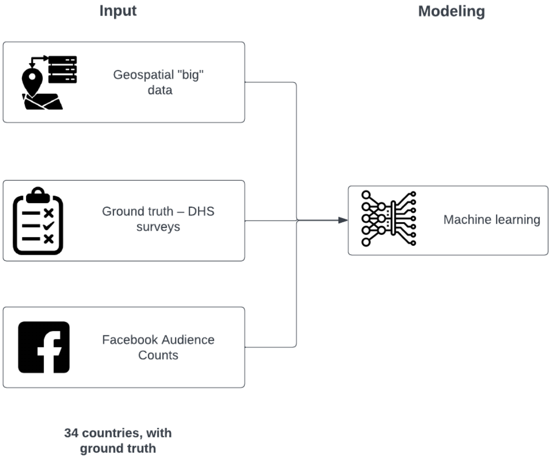
- **Goal**: Develop estimates of internet and mobile adoption by gender and digital gender gaps

Develop subnational estimates of adoption

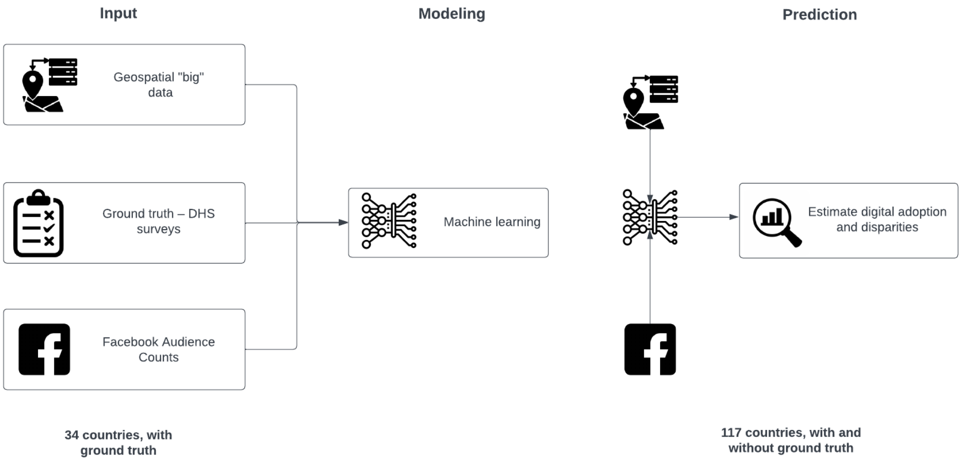
- **Goal**: Develop estimates of internet and mobile adoption by gender and digital gender gaps
- First subnational level
  - 117 countries, 2,075 subnational units



Overview of approach



Overview of approach

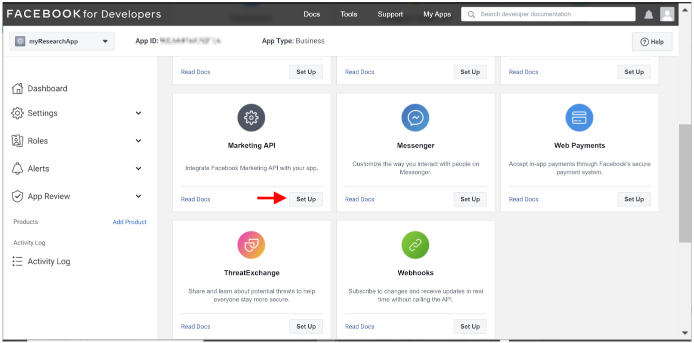


Ground truth – Demographic and Health Surveys (DHS)

- Household surveys representative at the first subnational level
  - Standardized sample design, questionnaire, implementation, etc.
  - Questions on **individual-level** internet use and mobile phone use (wave 7 onwards)
- Focus on 34 different DHS surveys, 2016-2023

Facebook monthly active users counts

- Collected through public marketing API
- Specify geographic region (FB template or custom region)
- Disaggregated counts by gender, age, device type, etc.



Big geospatial and population data

- Include ‘offline’ predictors that are uniformly available and consistent across subnational units
  - Satellite-derived nighlights data
  - Population density (World pop)
  - Subnational education index, income index, human development index (HDI), gender development index (GDI)

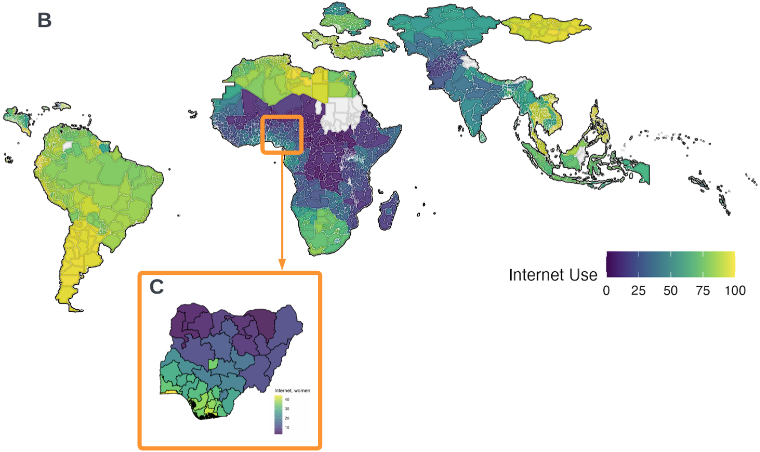
Outcomes of interest (from DHS)

Indicators	Women	Men	Gender Gap
Mobile Phone Ownership	✓	✓	✓
Internet Use, Past 12 Mo	✓	✓	✓

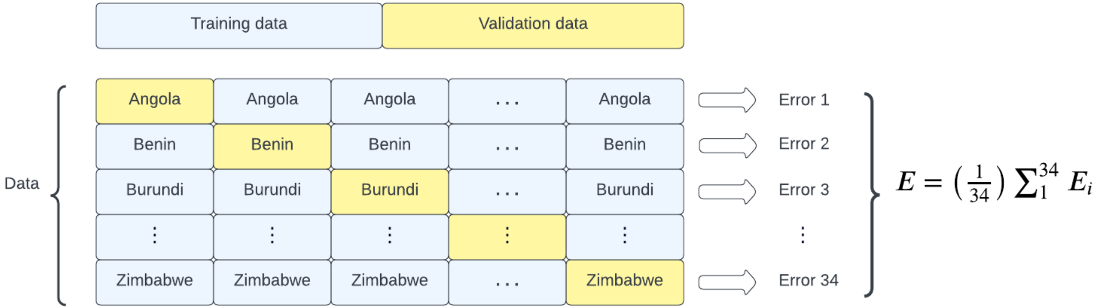
Modeling approach - ensemble machine learning

Algorithm	Description
glmnet (Lasso)	Lasso Regression
glmnet (Ridge)	Ridge Regression
glmnet (Elastic Net)	Elastic Net with 50% L1 Ratio
polspline	Polynomial Spline
ranger	Random Forest with 100 Trees
gbm	Gradient Boosted Machine
glm	Generalized Linear Model
xgboost	Extreme Gradient Boosting
SuperLearner	Ensemble method combining multiple learning algorithms

Greatly expanded geographic + temporal coverage



Leave-one-country-out cross validation

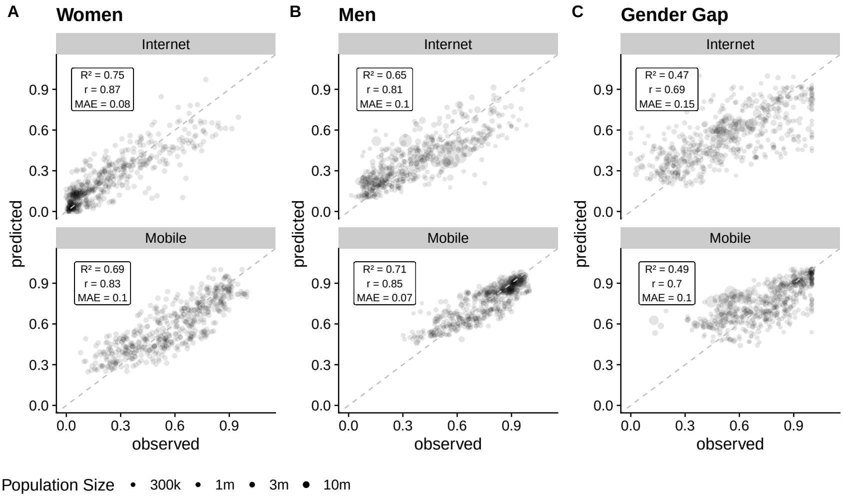


Validation Metric: R^2

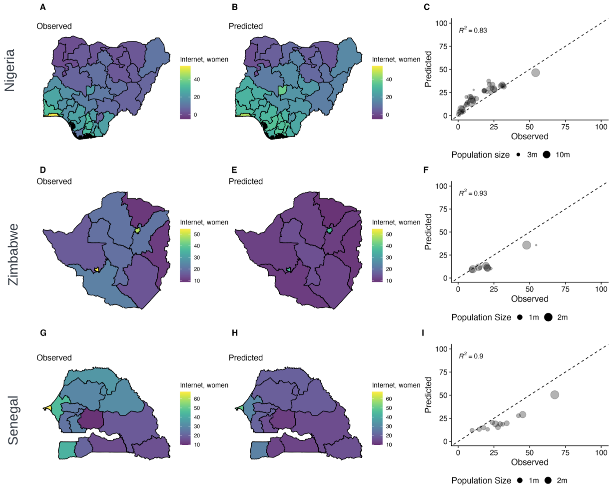
$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \tag{1}$$

- $1 =$ **Perfect predictions**
- $0 =$ **Mean**

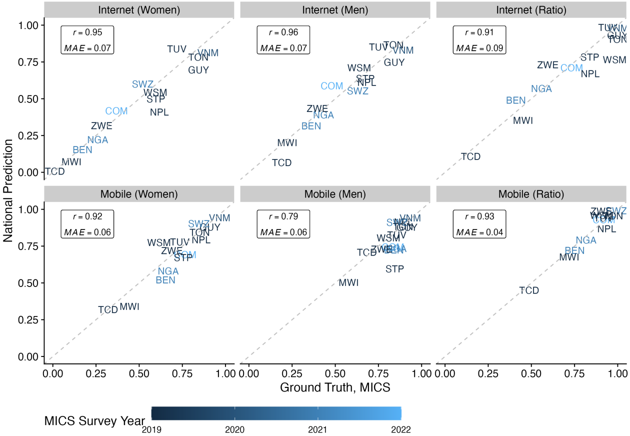
Overall predictive accuracy



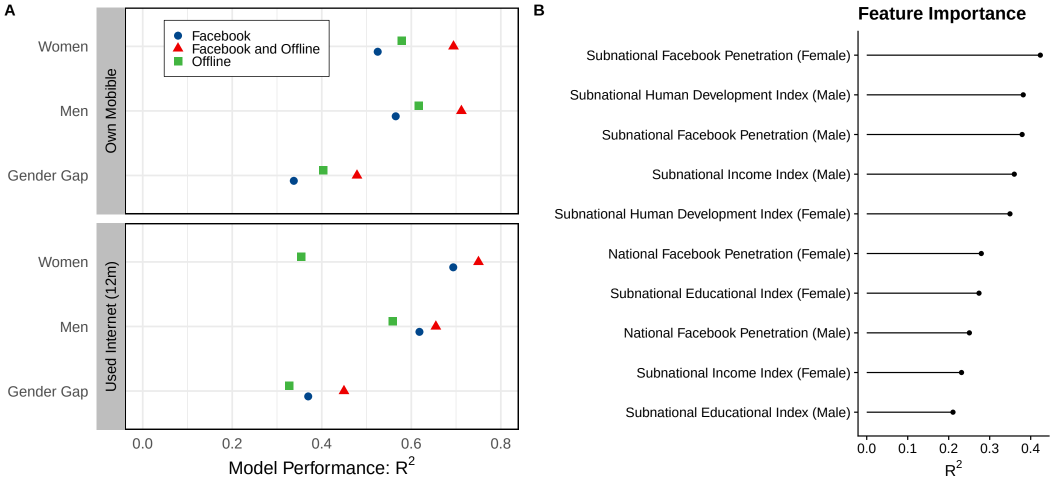
Error by country



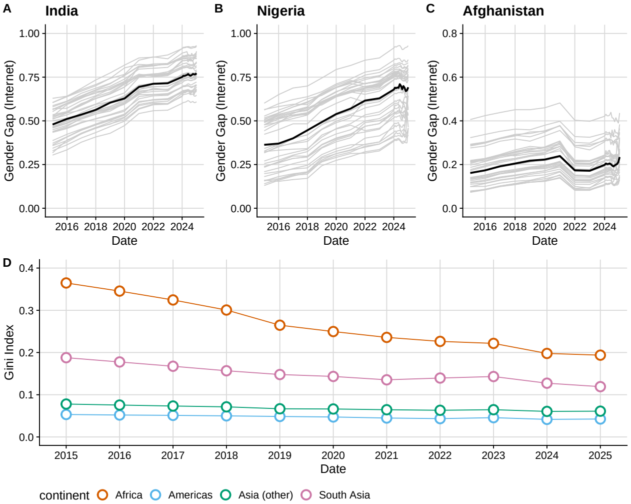
External benchmark



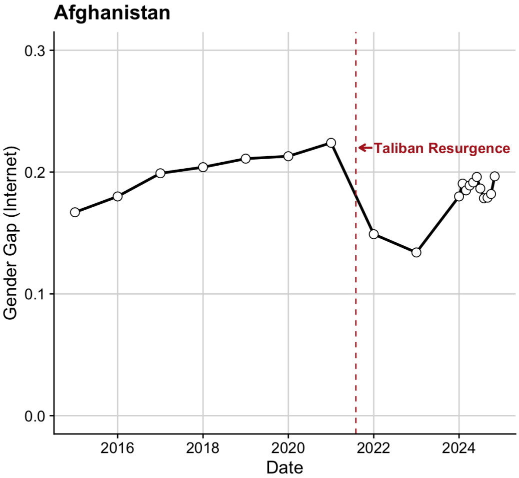
Most important predictors



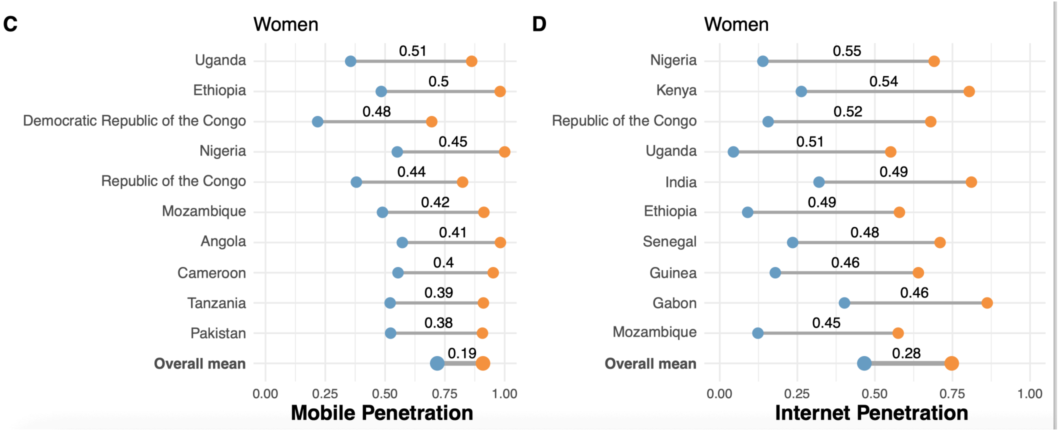
Trends from 2015 - present day



Extreme example



Within-country disparities



Summary

- Using Facebook audience counts **greatly expands** our ability to accurately predict internet adoption in countries with no ground truth
  - 2075 subnational units, 2015 - 2025

Summary

- Using Facebook audience counts **greatly expands** our ability to accurately predict internet adoption in countries with no ground truth
  - 2075 subnational units, 2015 - 2025
- Huge **disparities** in access to mobile and internet technologies between and within countries

Summary

- Using Facebook audience counts **greatly expands** our ability to accurately predict internet adoption in countries with no ground truth
  - 2075 subnational units, 2015 - 2025
- Huge **disparities** in access to mobile and internet technologies between and within countries
- New opportunities to study **population-level impacts** of digital technology using these subnational estimates

Project 2 - Demographic Impacts of Digitization

Digital revolution and fertility

- Diffusion theories of fertility decline have long emphasized the importance of mass media technologies in the spread of new ideas and norms (Montgomery and Casterline, 1996)

Digital revolution and fertility

- Diffusion theories of fertility decline have long emphasized the importance of mass media technologies in the spread of new ideas and norms (Montgomery and Casterline, 1996)
- Despite this theoretical potential, estimating the **causal impacts** of digital technology on fertility, especially in high-fertility contexts, has proven to be challenging

Digital revolution and fertility

- Diffusion theories of fertility decline have long emphasized the importance of mass media technologies in the spread of new ideas and norms (Montgomery and Casterline, 1996)
- Despite this theoretical potential, estimating the **causal impacts** of digital technology on fertility, especially in high-fertility contexts, has proven to be challenging
- Some evidence mobile phone ownership associated with lower parity / lower ideal family size (Billari, Rotondi and Trinitapoli, 2020); knowledge and access to contraception (Rotondi et al., 2020)

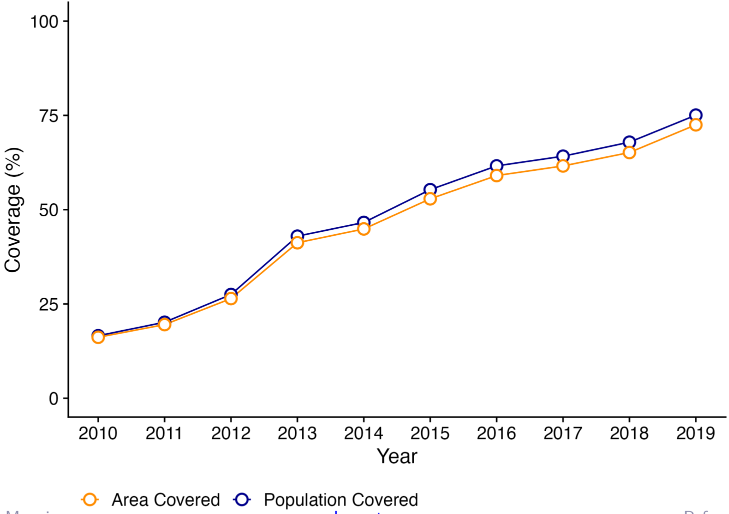
Research question

- Does expansion of 3G internet have a **causal effect** on fertility?
  - 2G coverage enables text/calling
  - 3G coverage enables mobile internet (social media, exposure to ideas from global elites, etc.)

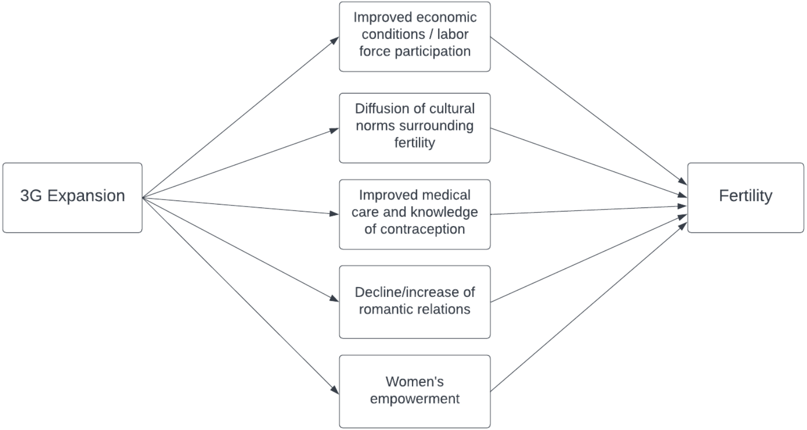
Research question

- Does expansion of 3G internet have a **causal effect** on fertility?
  - 2G coverage enables text/calling
  - 3G coverage enables mobile internet (social media, exposure to ideas from global elites, etc.)
- What are mechanisms linking 3G expansion with fertility behavior?

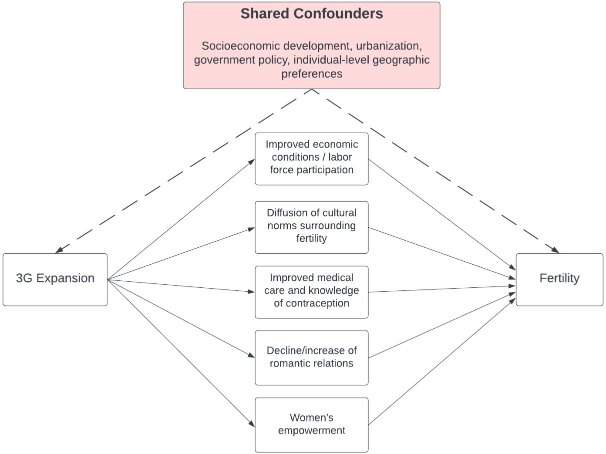
Focus on Nigeria: Rapidly expanding 3G infrastructure



Theoretical Framework

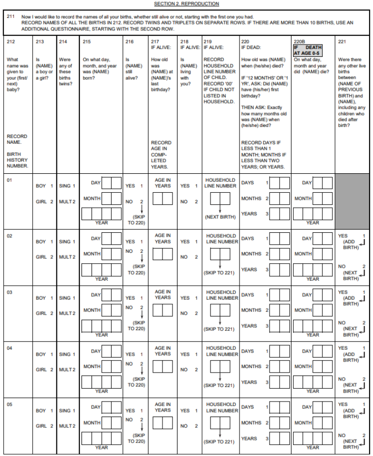


Theoretical Framework



Fertility + Mobile Coverage Data

- Generate longitudinal panel:
  - 2018 Nigeria DHS birth history (2010-2018), geo-referenced
  - Annual mobile coverage maps (2010-2018)



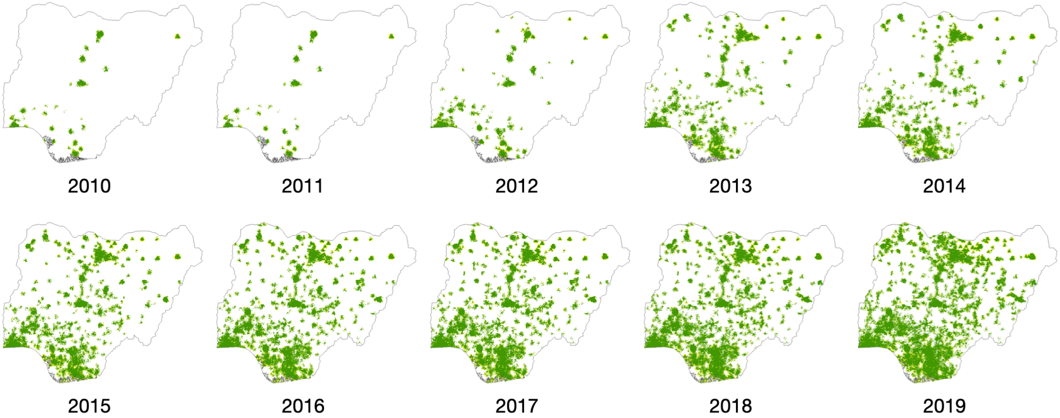
SECTION 2. REPRODUCTION

211 Now I would like to record the names of all your births, whether still alive or not, starting with the first one you had.
RECORD NAMES OF ALL THE BIRTHS IN 212. RECORD TWINS AND TRIPLETS ON SEPARATE ROWS. IF THERE ARE MORE THAN 10 BIRTHS, USE AN ADDITIONAL QUESTIONNAIRE, STARTING WITH THE SECOND ROW.

212	213	214	215	216	217 IF ALIVE:	218 IF ALIVE:	219 IF ALIVE:	220 IF DEAD:	220B IF DEATH AT AGE 0-5	221
What name was given to your (first/ next) baby? RECORD NAME. BIRTH HISTORY NUMBER.	Is (NAME) a boy or a girl?	Were any of these births twins?	On what day, month, and year was (NAME) born?	Is (NAME) still alive?	How old was (NAME) at (NAME)'s last birthday? RECORD AGE IN COMP-LETED YEARS.	Is (NAME) living with you?	RECORD HOUSEHOLD LINE NUMBER OF CHILD. RECORD '00' IF CHILD NOT LISTED IN HOUSEHOLD.	How old was (NAME) when (he/she) died? IF '12 MONTHS' OR '1 YR', ASK: Did (NAME) have (his/her) first birthday? THEN ASK: Exactly how many months old was (NAME) when (he/she) died? RECORD DAYS IF LESS THAN 1 MONTH; MONTHS IF LESS THAN TWO YEARS; OR YEARS.	On what day, month and year did (NAME) die?	Were there any other live births between (NAME OF PREVIOUS BIRTH) and (NAME), including any children who died after birth?
01	BOY 1 GIRL 2	SING 1 MULT 2	DAY MONTH YEAR	YES 1 NO 2 (SKIP TO 220)	AGE IN YEARS	YES 1 NO 2	HOUSEHOLD LINE NUMBER (NEXT BIRTH)	DAYS 1 MONTHS 2 YEARS 3	DAY MONTH YEAR	
02	BOY 1 GIRL 2	SING 1 MULT 2	DAY MONTH YEAR	YES 1 NO 2 (SKIP TO 220)	AGE IN YEARS	YES 1 NO 2	HOUSEHOLD LINE NUMBER (SKIP TO 221)	DAYS 1 MONTHS 2 YEARS 3	DAY MONTH YEAR	YES 1 (ADD BIRTH) NO 2 (NEXT BIRTH)
03	BOY 1 GIRL 2	SING 1 MULT 2	DAY MONTH YEAR	YES 1 NO 2 (SKIP TO 220)	AGE IN YEARS	YES 1 NO 2	HOUSEHOLD LINE NUMBER (SKIP TO 221)	DAYS 1 MONTHS 2 YEARS 3	DAY MONTH YEAR	YES 1 (ADD BIRTH) NO 2 (NEXT BIRTH)
04	BOY 1 GIRL 2	SING 1 MULT 2	DAY MONTH YEAR	YES 1 NO 2 (SKIP TO 220)	AGE IN YEARS	YES 1 NO 2	HOUSEHOLD LINE NUMBER (SKIP TO 221)	DAYS 1 MONTHS 2 YEARS 3	DAY MONTH YEAR	YES 1 (ADD BIRTH) NO 2 (NEXT BIRTH)
05	BOY 1 GIRL 2	SING 1 MULT 2	DAY MONTH YEAR	YES 1 NO 2 (SKIP TO 220)	AGE IN YEARS	YES 1 NO 2	HOUSEHOLD LINE NUMBER (SKIP TO 221)	DAYS 1 MONTHS 2 YEARS 3	DAY MONTH YEAR	YES 1 (ADD BIRTH) NO 2 (NEXT BIRTH)

DHS Birth History

3G Coverage Rollout in Nigeria

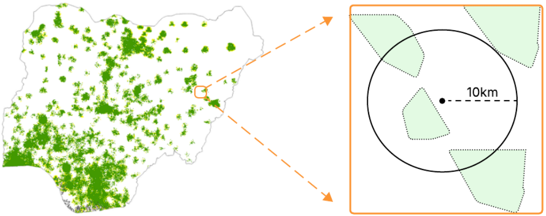


Source: GSMA Mobile Coverage Maps

Constructing longitudinal panel

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

Analytic Strategy (Two-Way Fixed Effects)

$$LB_{ict} = \beta_0 + \underbrace{\beta_1 \text{3G}_{ct}}_{\text{3G coverage intensity}} + \underbrace{\gamma_c}_{\text{Cluster FE}} + \underbrace{\delta_t}_{\text{Year FE}} + \underbrace{\boldsymbol{\beta}\mathbf{X}i}_{\text{Controls}} + \epsilon_{ict}$$

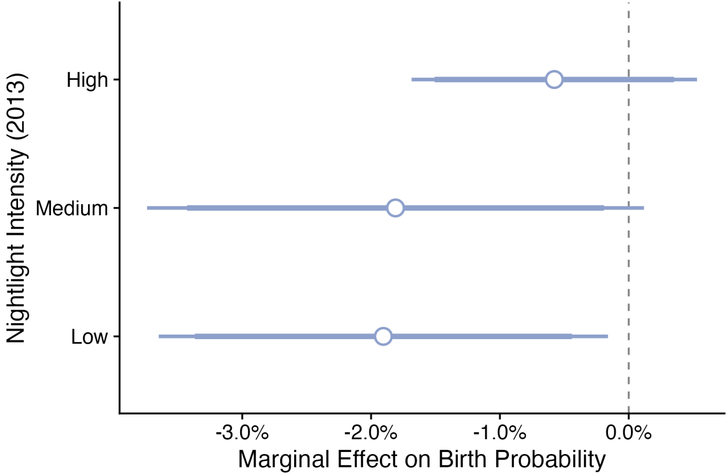
where

- LB_{ict} is an indicator for whether woman i in cluster c at time t had a live birth in the past year
- 3G_{ct} denotes the proportion of the population in cluster c covered by 3G service in year t

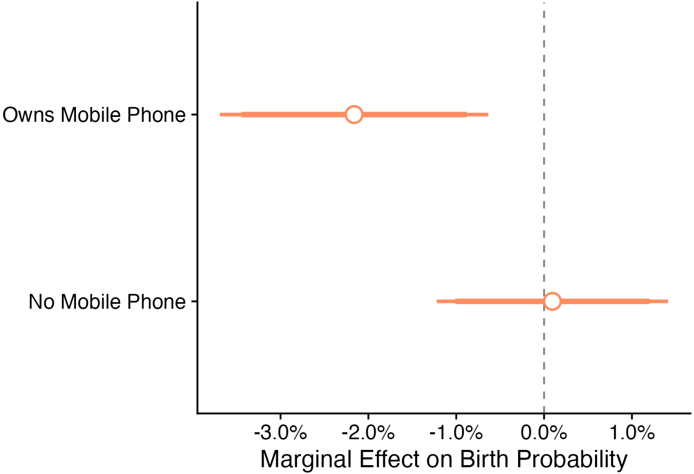
Effect of 3G Coverage on Recent Birth

	Model 1	Model 2	Model 3
Intercept	0.184***		
	(0.002)		
3G Coverage (Population Share)	-0.049***	-0.011*	-0.011*
	(0.005)	(0.005)	(0.005)
2G Coverage (Population Share)		0.005	0.005
		(0.017)	(0.017)
Individual-level controls			X
Fixed Effects: DHS Cluster		X	X
Fixed Effects: Year		X	X
Fixed Effects: Age		X	X
Fixed Effects: Parity (Lagged)		X	X
Observations	116178	116178	116178
R^2	0.002	0.087	0.089

Heterogeneity by development level (nightlights proxy)



Heterogeneity by mobile phone ownership



Alternative specifications - mother fixed effects

	Model 1	Model 2 (0 Parity)
3G Coverage (Population Share)	-0.043***	-0.054***
	(0.006)	(0.009)
2G Coverage (Population Share)	-0.003	0.026
	(0.024)	(0.031)
Fixed Effects: Mother	X	X
Fixed Effects: Year	X	X
Fixed Effects: DHS Cluster	X	X
Fixed Effects: Parity (Lagged)	X	X
Observations	194,067	60,732
R^2	0.292	0.277

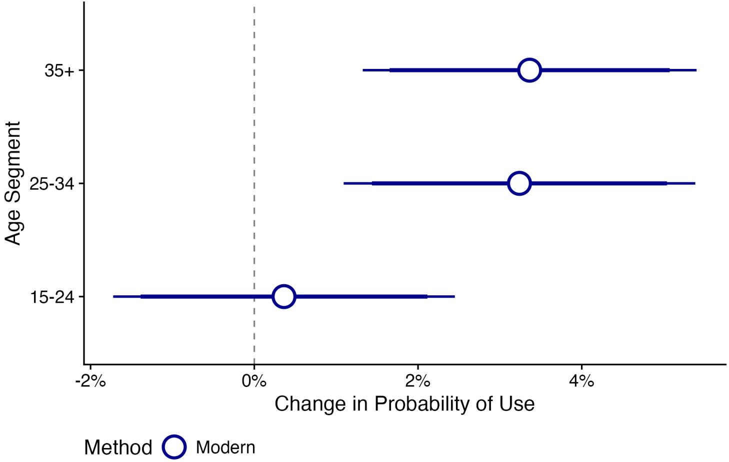
Testing Mechanisms

- Not longitudinal, measured in 2018 Nigeria DHS...
- Suggestive descriptive evidence of association between 3G expansion (2010-2018) and outcomes (not causal...)

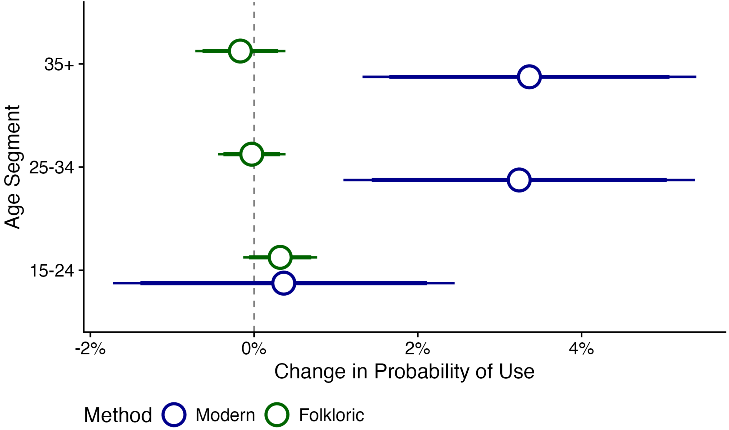
Delayed cohabitation, decreased ideal family size

	Age at First Cohabitation	Ideal family size
3G coverage expansion (cluster-level)	0.160	-0.248**
	(0.100)	(0.092)
Wealth quintile	0.088*	-0.085***
Currently working	-0.343***	0.207***
Education level	0.711***	-0.482***
Religion (Islam)	0.003	-0.002
Access to radio	0.015	0.083**
Access to television	0.090	-0.164***
FE - Birth Cohort	X	X
FE - State	X	X
Cluster covariates (rainfall, nighlights, IMR)	X	X
Observations	7202	23566

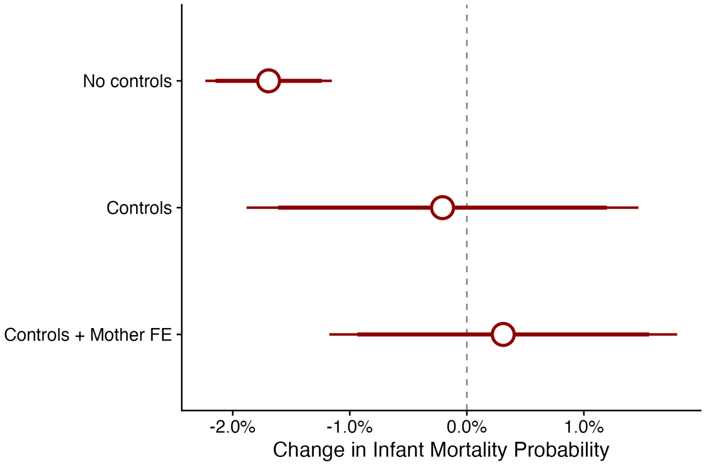
3G associated with higher modern contraception uptake



But not folkloric methods...



No evidence of reduced infant mortality...



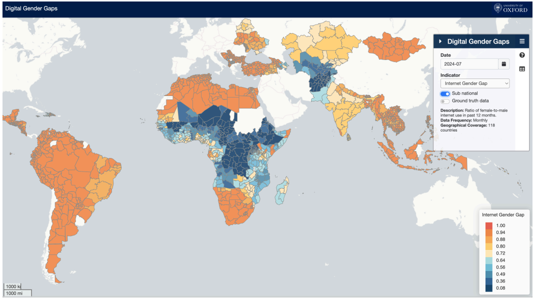
Conclusion

- Used mobile coverage maps and retrospective fertility histories to create a **longitudinal panel**, exploiting plausibly exogenous rollout of 3G coverage
- Full 3G coverage expansion has **causal effect** of approximately 7% reduction in probability of birth over baseline
- Plausible mechanisms:
  - Evidence for lower ideal family size and increase in modern contraception usage
  - Not driven delayed cohabitation or improved child survival
- **Next steps**: Robustness check with instrument variable

Thank You

Funders:

- Bill and Melinda Gates Foundation (INV-045370)
- Leverhulme Trust (Grant RC-2018-003) for the Leverhulme Centre for Demographic Science



digitalgendergaps.org

Contact:

caseyfbreen

casey.breen@demography.ox.ac.uk

References

Billari, Francesco C., Valentina Rotondi and Jenny Trinitapoli. 2020. "Mobile Phones, Digital Inequality, and Fertility: Longitudinal Evidence from Malawi." *Demographic Research* 42:1057–1096.

DiMaggio, Paul and Eszter Hargittai. 2001. "From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use as Penetration Increases." p. 25.

Hjort, Jonas and Jonas Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa." *American Economic Review* 109(3):1032–1079.

Kharisma, Bayu. 2022. "Surfing Alone? The Internet and Social Capital: Evidence from Indonesia." *Journal of Economic Structures* 11(1):8.

Kho, Kevin, Leah K Lakdawala and Eduardo Nakasone. 2018. "Impact of Internet Access on Student Learning in Peruvian Schools.".

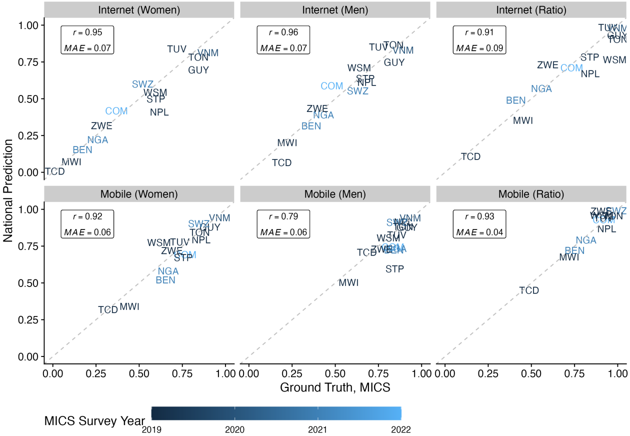
Montgomery, Mark R. and John B. Casterline. 1996. "Social Learning, Social Influence, and New Models of Fertility." *Population and Development Review* 22:151.

Rotondi, Valentina, Ridhi Kashyap, Luca Maria Pesando, Simone Spinelli and Francesco C. Billari. 2020. "Leveraging Mobile Phones to Attain Sustainable Development." *Proceedings of the National Academy of Sciences* 117(24):13413–13420.

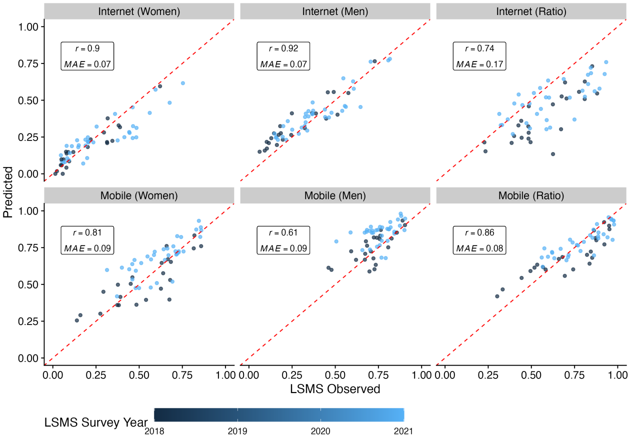
Suri, Tavneet and William Jack. 2016. "The Long-Run Poverty and Gender Impacts of Mobile Money." *Science* 354(6317):1288–1292.

Unwin, P. T. H. 2009. *ICT4D: Information and Communication Technology for Development*. Cambridge University Press.

External benchmark



External Benchmark



Trends validation