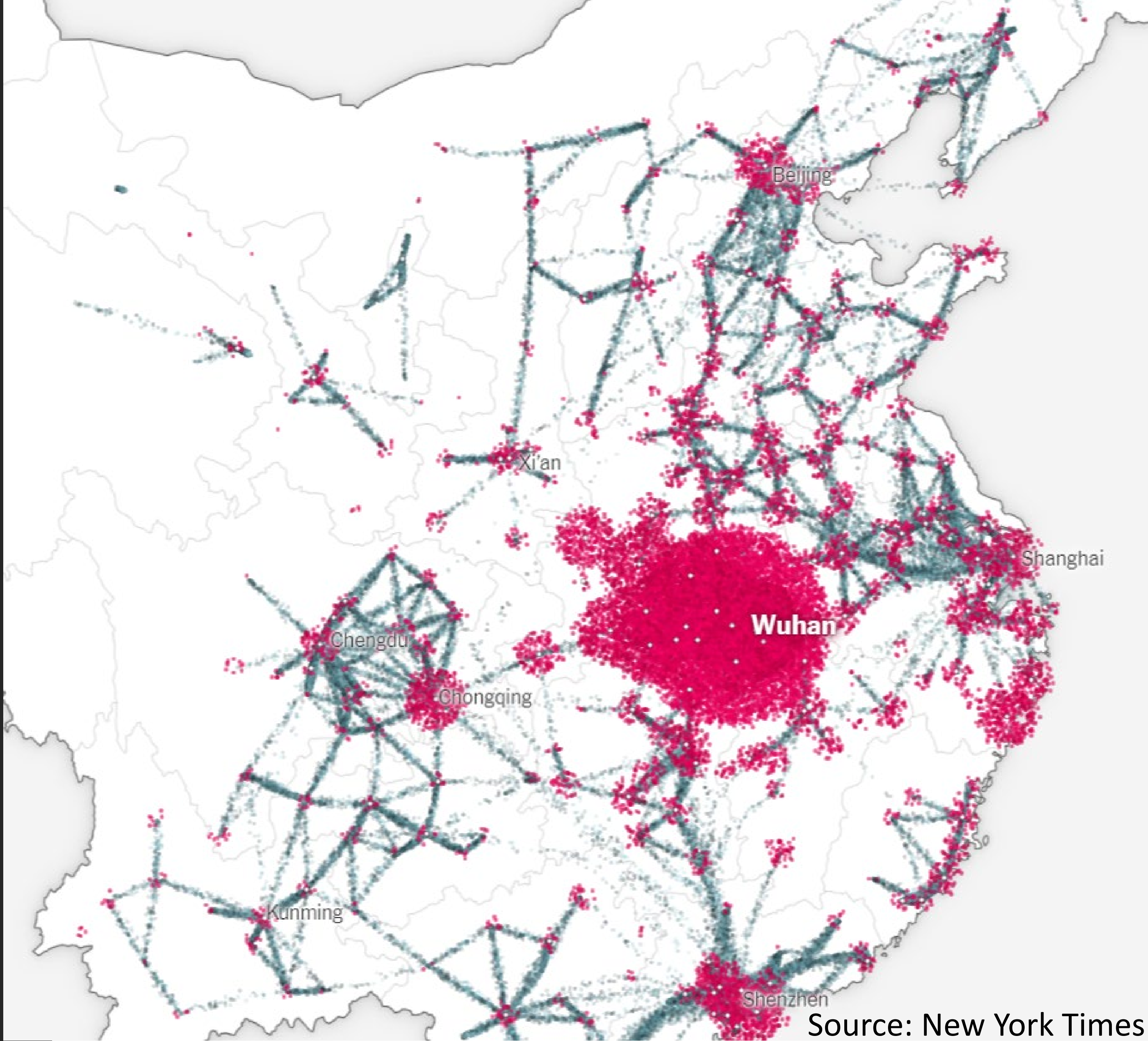
Combining mobile phone data with health surveillance data to understand the spread of disease

Sveta Milusheva
Development Impact Evaluation (DIME)
World Bank

Mobility is central to Spread of COVID-19

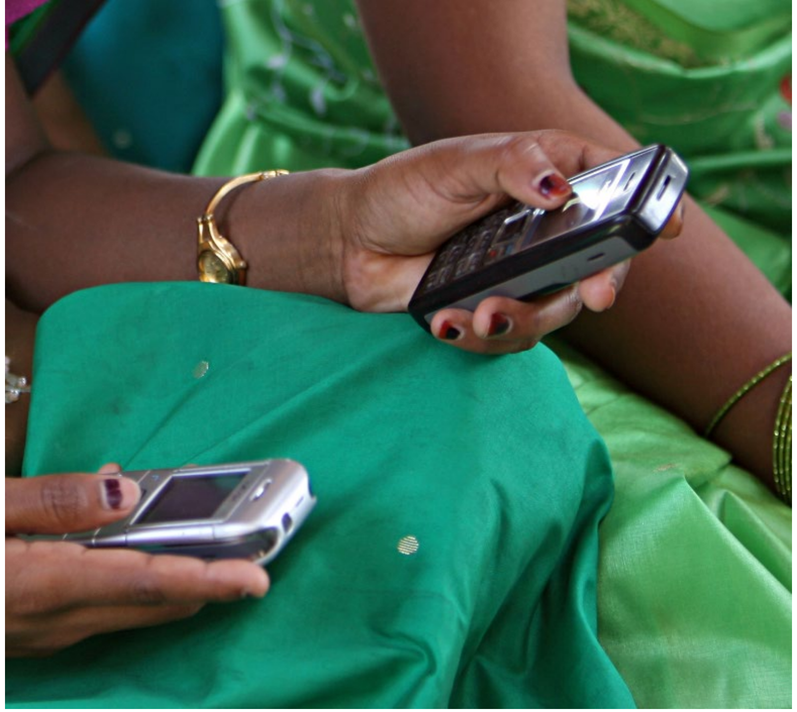


Source: New York Times

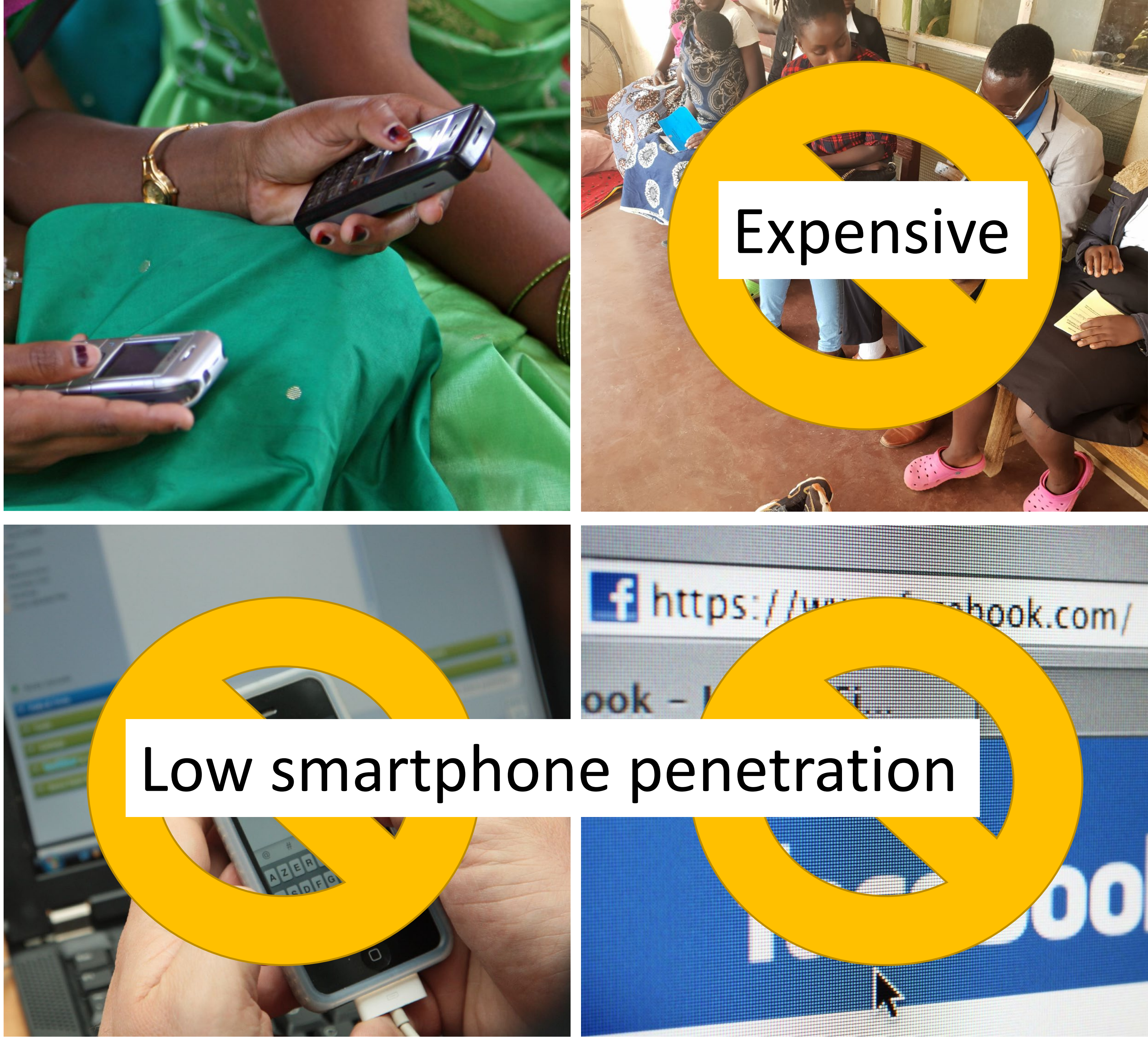
How do we measure the impact of mobility and target policies to mitigate the negative externality?

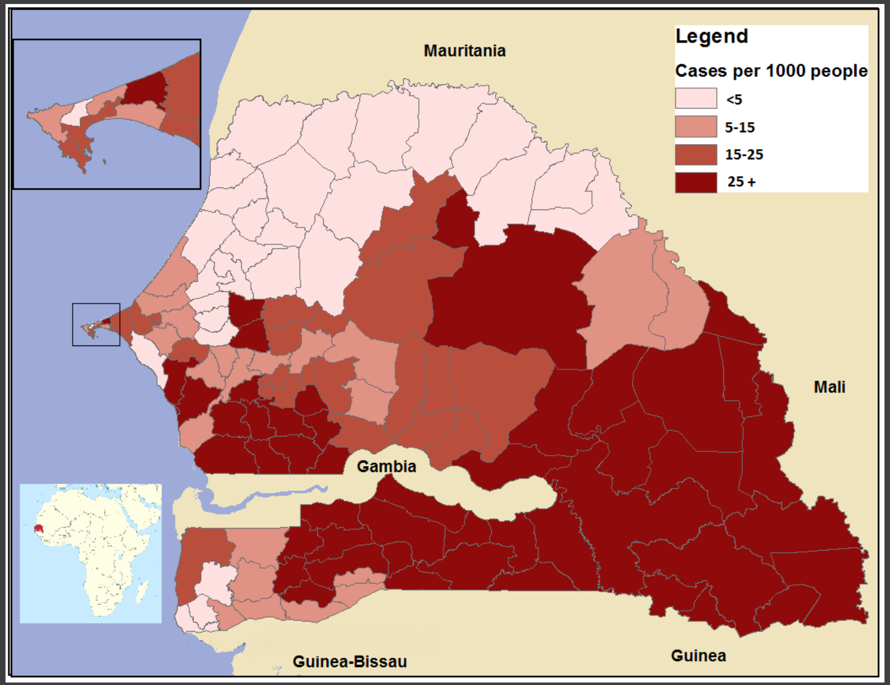
- Census data
- Detailed surveys
- Instruments for movement

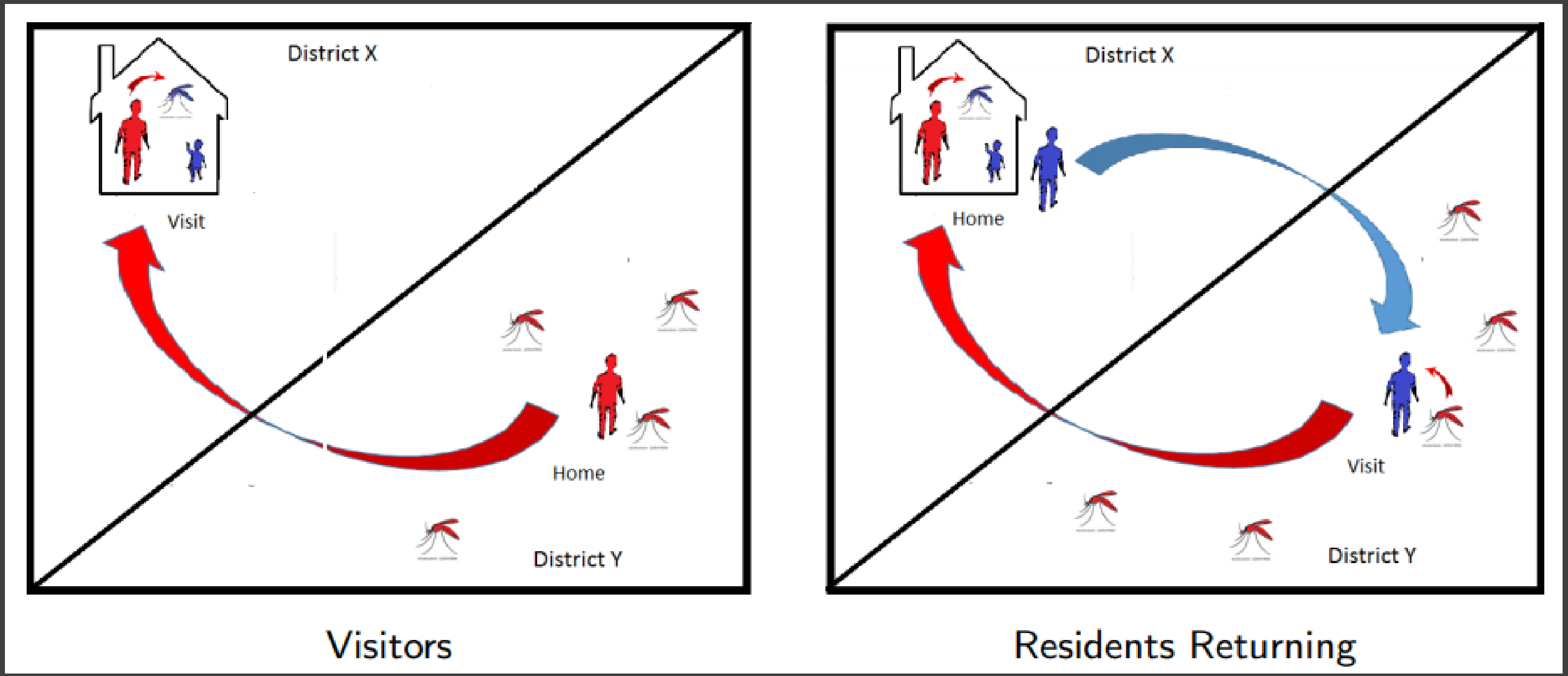
- Mobile phones
- Smartphone GPS
- Mobile applications



How to measure mobility across an entire low-resource country?







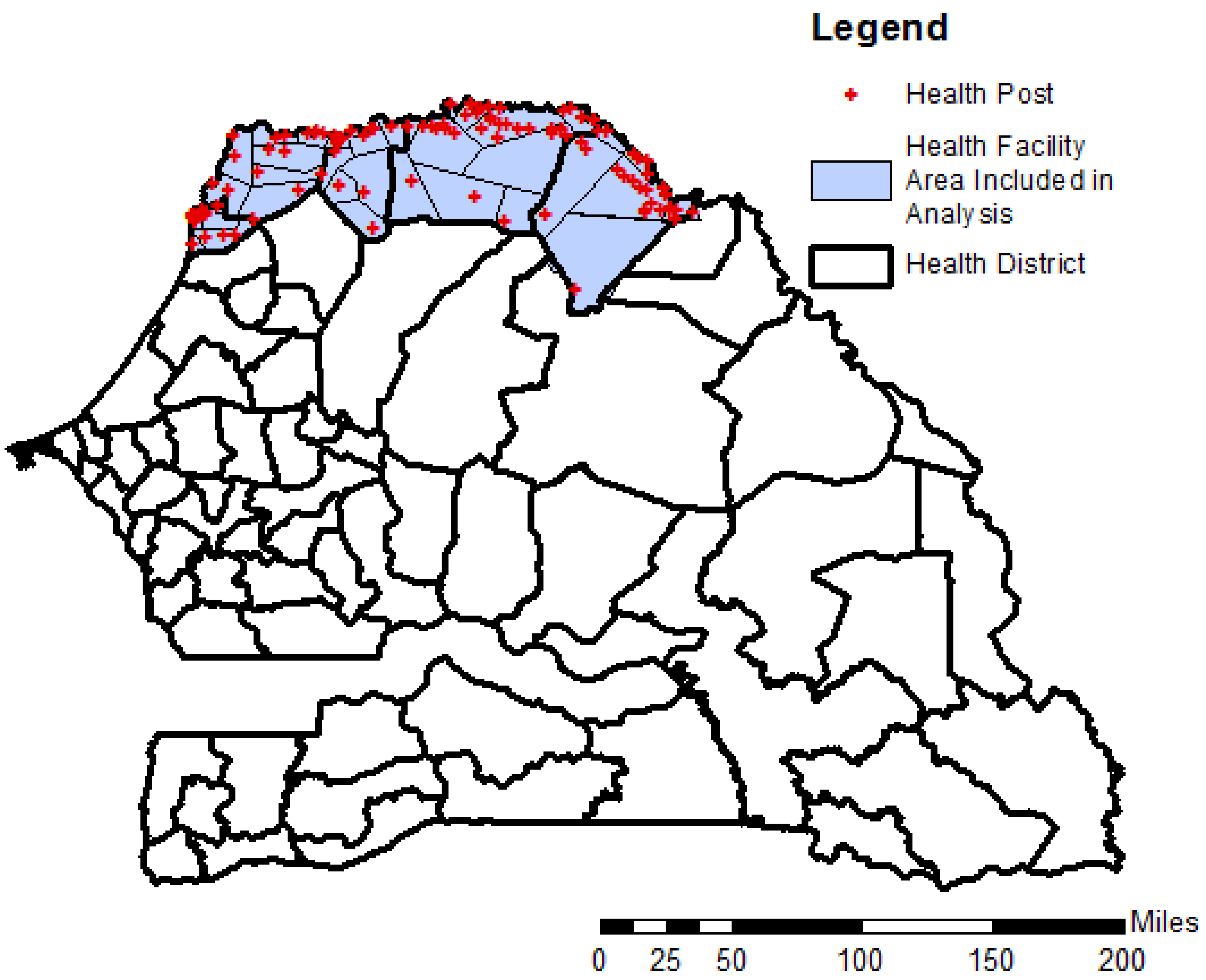
Example: Malaria in Senegal

Example: Malaria in Senegal

- Utilize high frequency temporal and spatial data in Senegal
  - Mobile phone data to measure population movement (15 billion records)
- Build and estimate an epidemiological model of malaria transmission
  - Combine mobility data with incidence data
  - Use a linear dynamic panel-data model and control for time fixed effects
- Conduct simulations to compare targeting strategies for travelers
  - Explore targeting of specific locations and travelers from specific locations

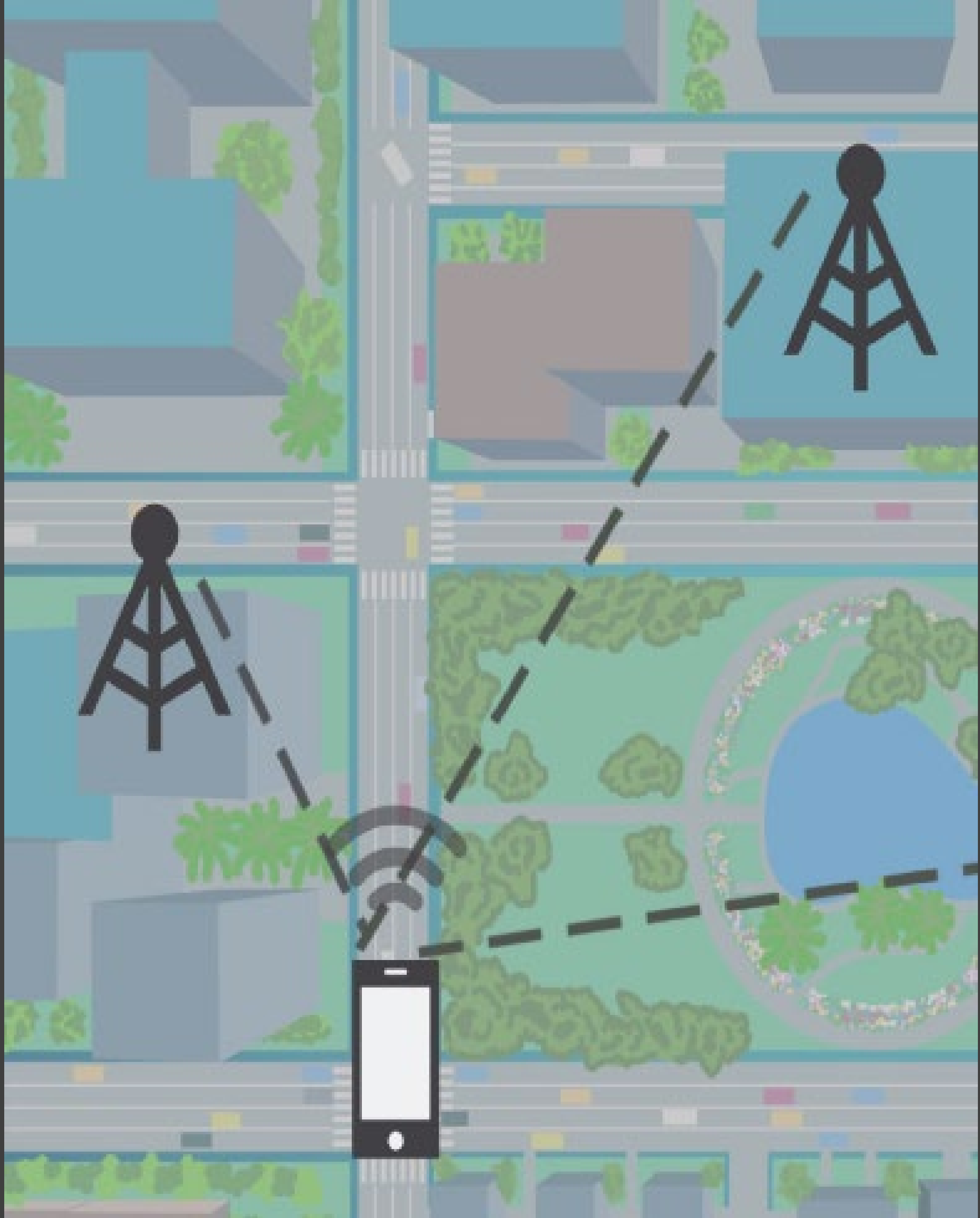
Malaria Data

- Monthly data on number of confirmed cases for all health districts
- Monthly data on number of confirmed cases at the health post level for 5 very low malaria health districts (117 health posts)
- 5 districts divided into health post catchment areas

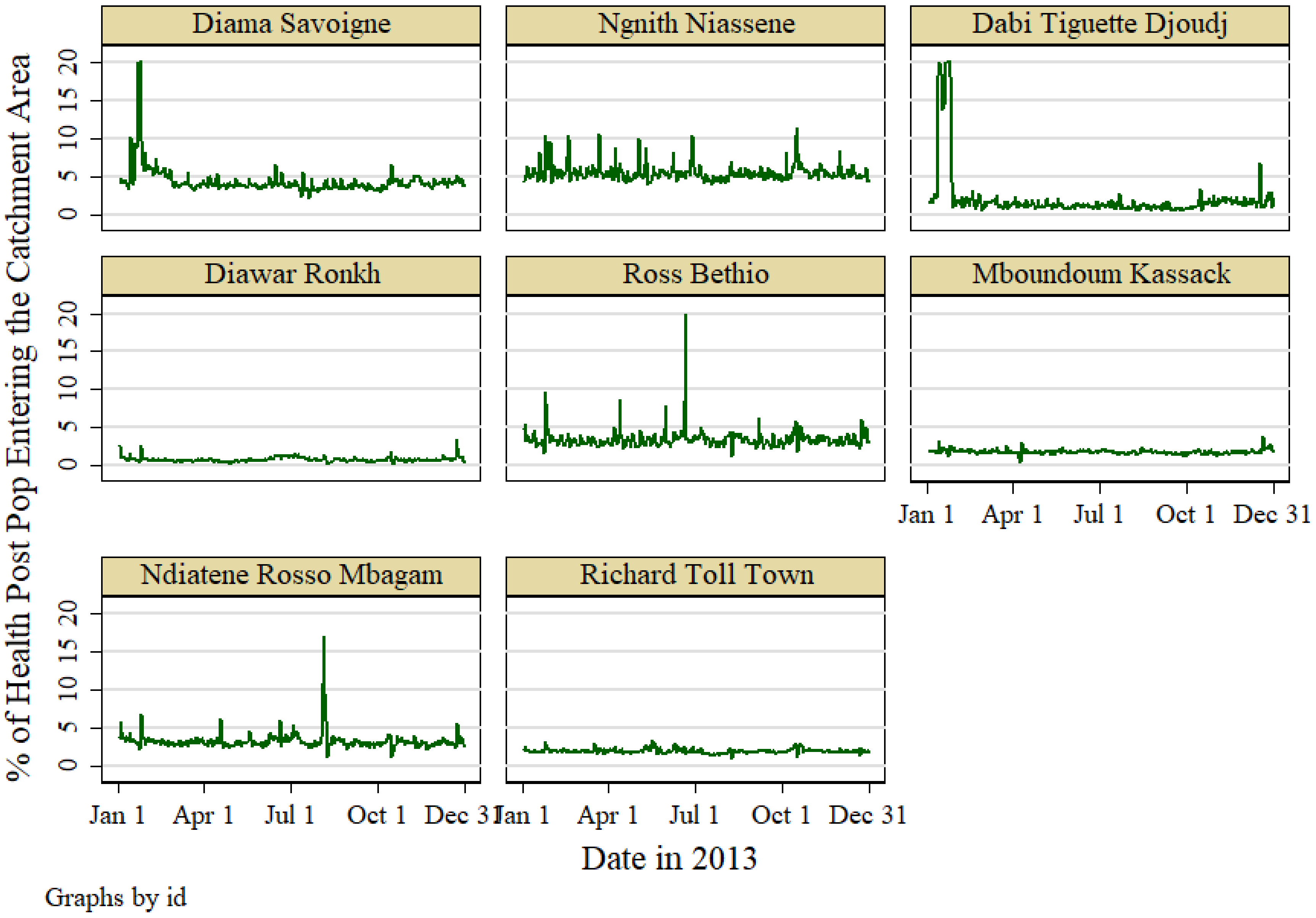


High Frequency Data

- Call Detail Record (CDR) data provided by Sonatel/Orange
- Data for 2013 on over 9 million SIM cards (15 billion observations)
- 83% of those with a phone use Sonatel (LSS 2014)
- Observation for every call/text made or received with timestamp
- GPS coordinates of tower from where call is made
- Data is anonymized but still sensitive



Movement Aggregated at Health Post Level



Empirical Model

$$
\begin{aligned}
x_{it} &= \beta_1 x_{it-1} + \beta_2 \mathbb{E}(\mathcal{I}_{it}) + \alpha Z_{it} + \gamma_i + \delta_t + \epsilon_{it} \\
\mathbb{E}(\mathcal{I}_{it}) &= \frac{1}{H_{it}} \sum_{j \neq i} \sum_{p_t \in j} T_{ip}(x_{jt} T_{jp})
\end{aligned}
$$

x_{it} = Malaria incidence per 1000 in area i in month t

Z_{it} = 0,1,2 Lags of rainfall

γ_i = Health post catchment area fixed effects

δ_t = Month fixed effects

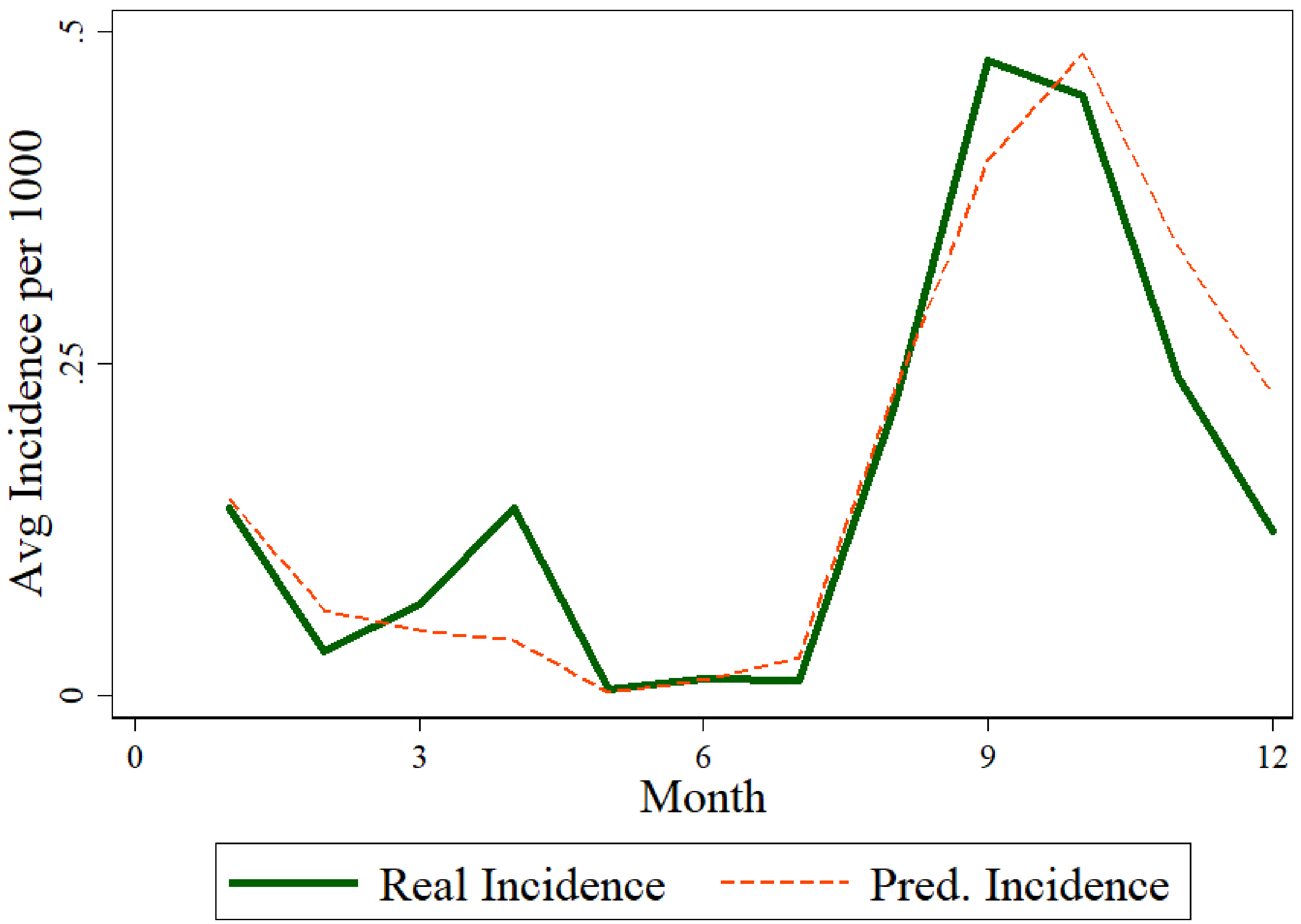
ϵ_{it} = Error term

H_{it} = Population in i in month t

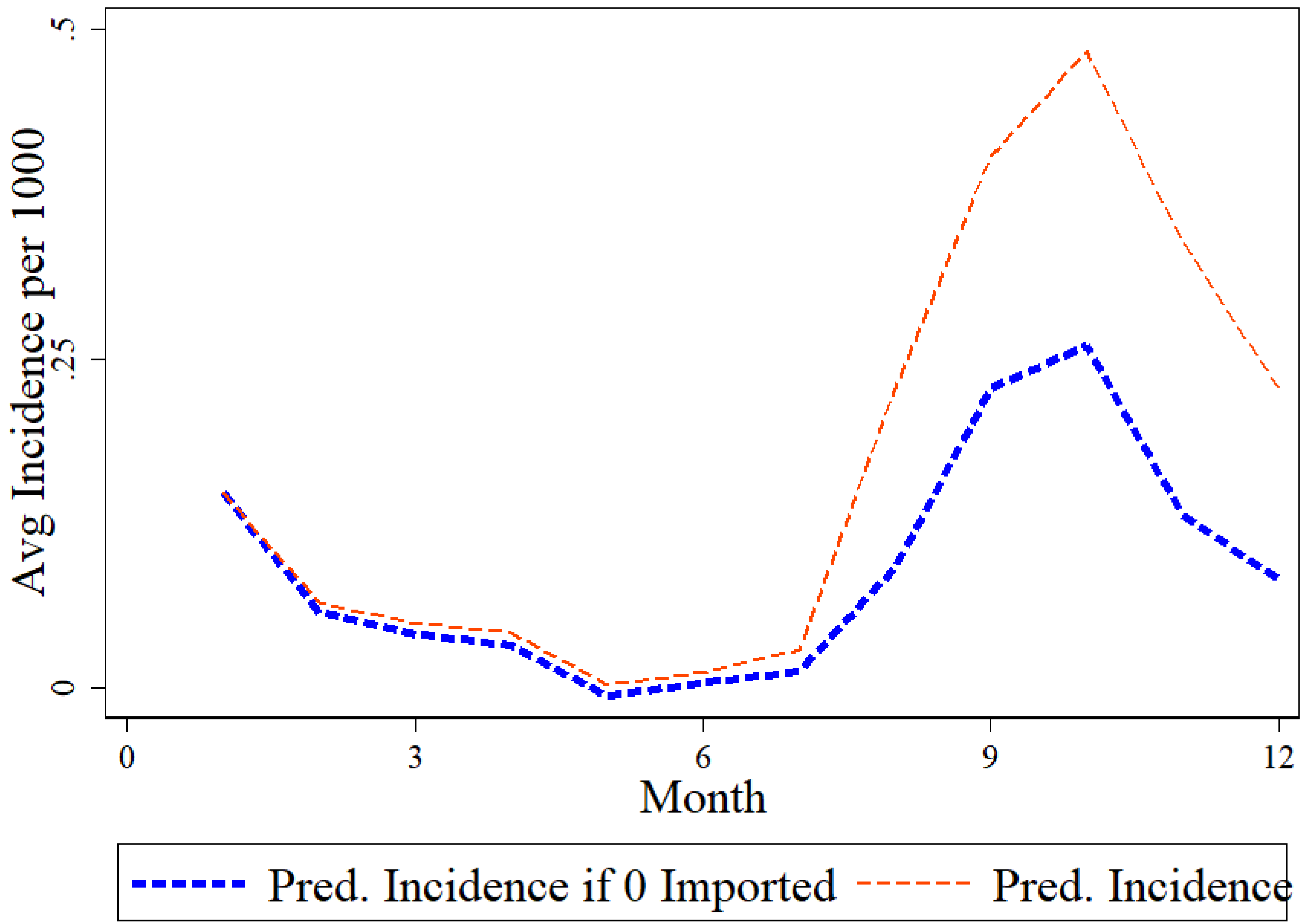
T_{ip} = Proportion of 15 days person p spent in i in month t

T_{jp} = Proportion of month up to 15 days p spent in j before entering i

Goodness of Fit for Example District



Effect of Travelers for Example District



	Arellano Bond
Imported Incidence	1.094***
	(0.357)
Lag Incidence	0.563***
	(0.129)
Rain in cm	-0.000654
	(0.00790)
Lag Rain in cm	0.0269
	(0.0181)
Lag 2 Rain in cm	0.0319**
	(0.0156)
Constant	-0.0688**
	(0.0290)
Month FE	Yes
Health Post Area Controls	No
Health Post x Month Obs	432

How do we Use this for Policy?

More Effective Targeting

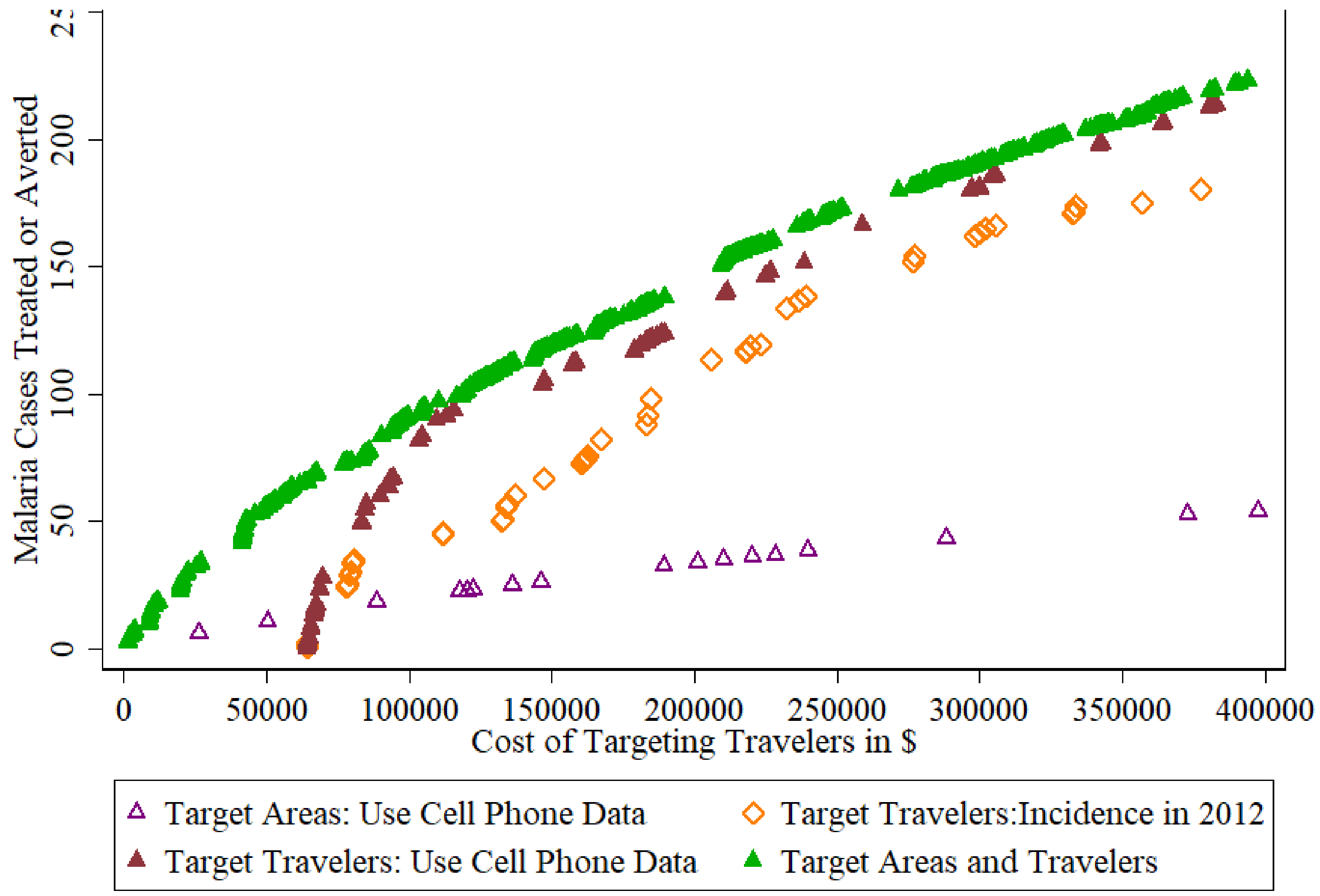
- Number of travelers entering 5 low malaria districts in 2013: 6,956,197 travelers
- Cases generated by these travelers: 607 cases (∼40% of case load)

Cost of any testing policy:

- Fixed costs such as training health workers
- Variable costs such as cost of rapid diagnostic test

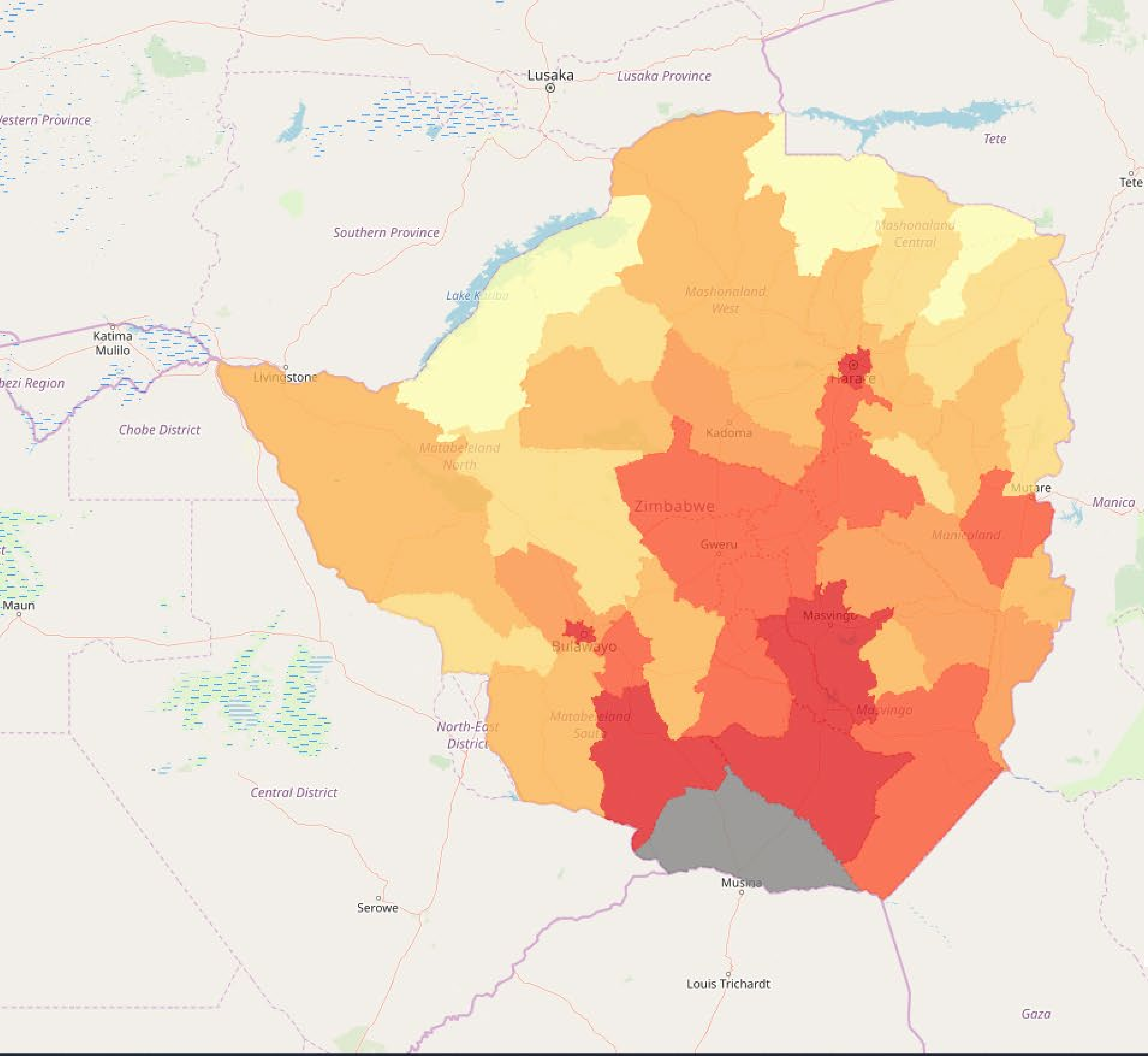
Targeting

- Target particular locations and test all travelers in these locations
- Target only high risk travelers to all locations



Cost/Benefit of Different Targeting Strategies

How does this relate to COVID-19?

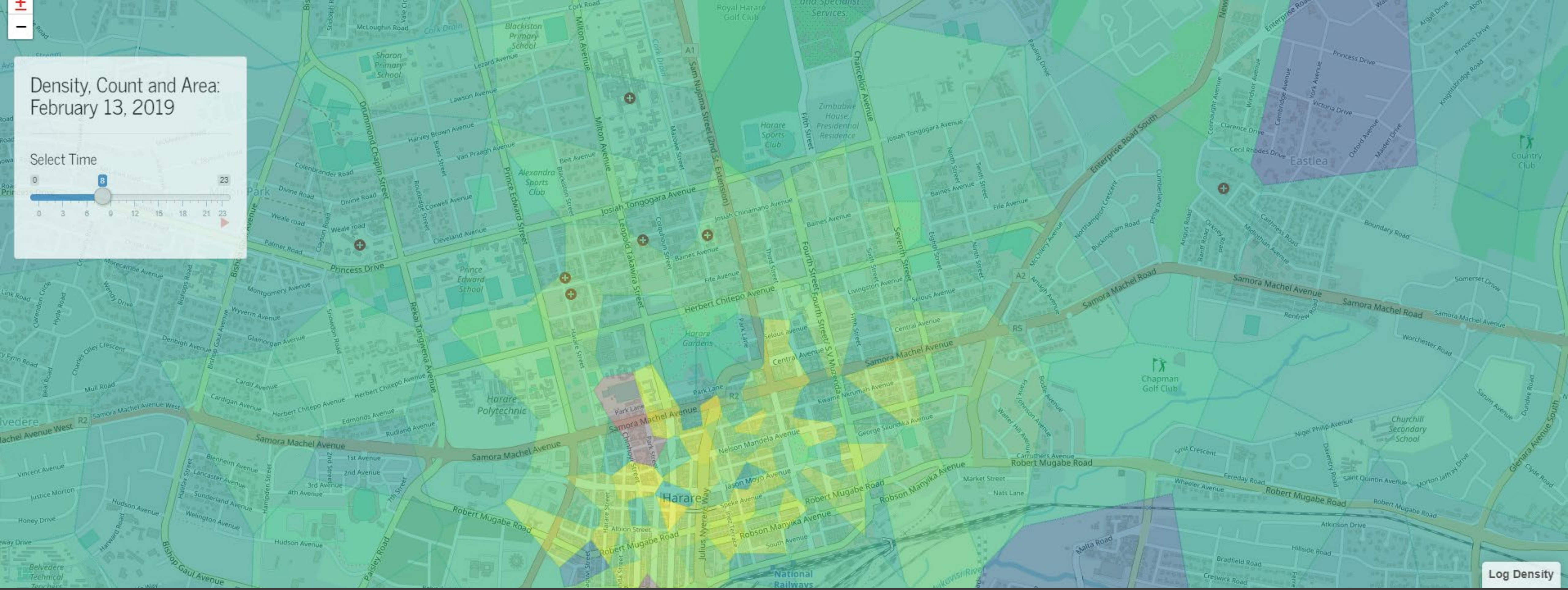


Use mobile phone data to study mobility and identify high-risk areas for spread

Example: Zimbabwe

COVID-19 Mobile Analytics Task Force

Data Access	Establish access to telecom data and case/incidence data
Country Cases	Build out additional country cases
Analytics	Develop systematic analytics to directly support policymakers
Dashboards	Communicate analytics through user-friendly dashboards to directly support policy targeting



Thank You

Contact: Sveta Milusheva

smilusheva@worldbank.org