The Great Revolt and its Legacy: Understanding Vaccine Hesitancy in Colonial India

Noel D. Johnson (GMU) & Malik Altaf Hussain (GMU & WVU)

Economic History Association, September 2025

This Version: September 4, 2025

Comments Welcome

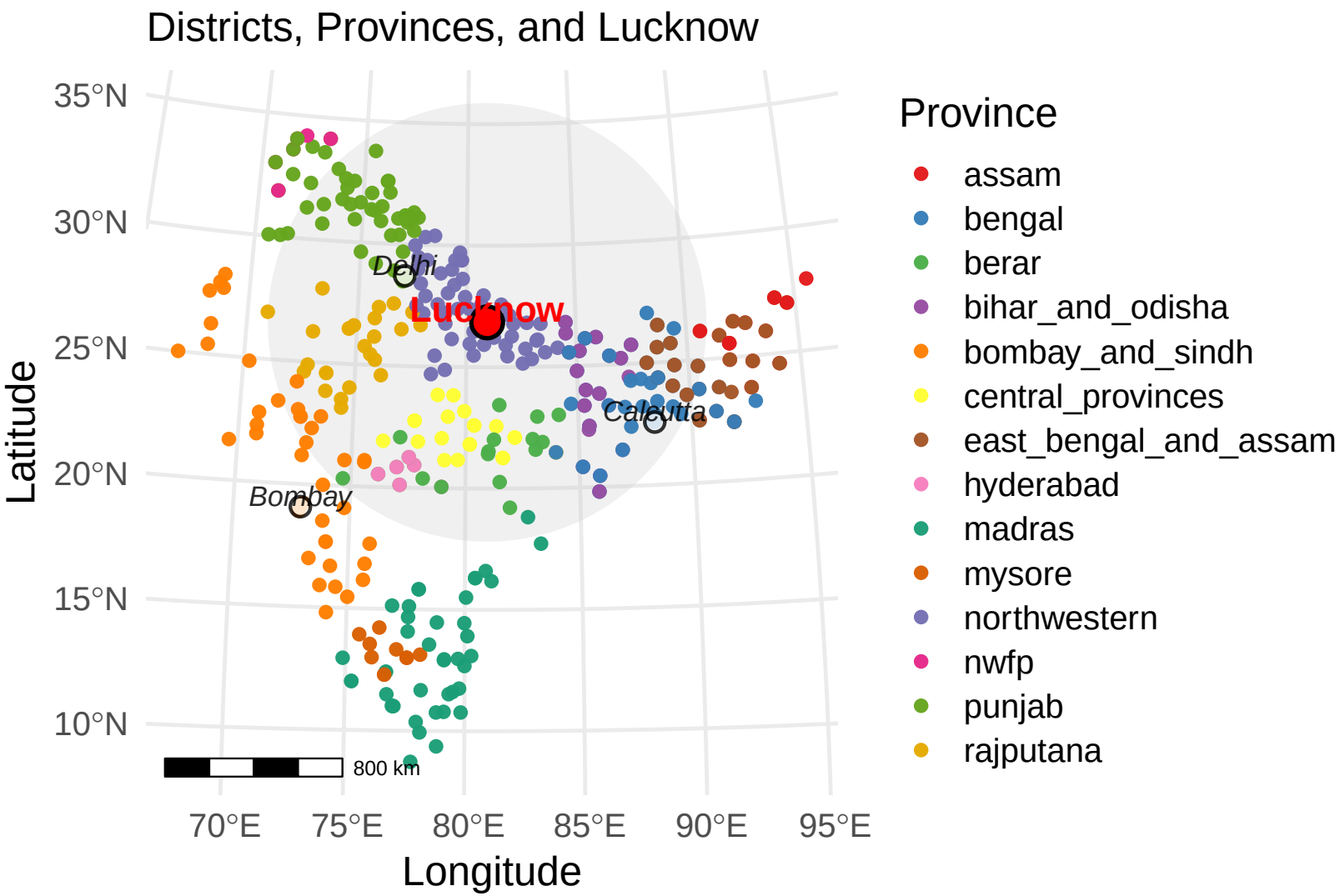
njohnsoL@gmu.edu

Trust in Government is related to Success of Public Health Policies

- From the Report on Vaccination in the Punjab (1868-1880)...

4. With reference to the complaint of the people of the Jullundhur district, that children already vaccinated are taken from one village to another to supply vaccine. His Honor desires that in future no children be so taken without the consent of their parents. The practice may appear to the Superintendent-General to involve no great hardship, but it is certainly unlawful, and if continued will do much to render vaccination unpopular.

Smallpox Vaccination in India: 1870-1919



Note: Map shows British Indian districts, and the location of Lucknow. Baseline sample restricted to districts within 1,000 km of Lucknow.

Smallpox Vaccination in 19th Century India

- Smallpox was a leading cause of death in 19th-century India, with fatality rates exceeding 25%.
- During 1870's up to 100,000 people died every year from smallpox (Banthia and Dyson, 1999, 660).
- The British colonial state expanded the vaccination campaign in the 1860s, supported by a growing health bureaucracy. ▸ Vaccination Coverage
- Yet uptake varied sharply across districts. ▸ District Vax Rates

Research Question: Did exposure to colonial repression during the 1857 Revolt reduce later engagement with state-led vaccination campaigns?

Mistrust and the Great Revolt of 1857

- The Great Revolt of 1857 (or Sepoy Rebellion) was a major disturbance in Colonial India.
- Roughly 6,000 British killed and, based on reconstructions from census data, up to 800,000 Indians were killed.
- Marked the transition of power from East India Company to British government.
- There were massacres on both sides, but the retribution by the British was particularly brutal and orders of magnitude worse.

Geographic Context of the Revolt

Meerut (initial mutiny) and Lucknow (key siege and repression site).

Historical Context: The Revolt of 1857 in Awadh

- **1856:** British annex Awadh, dismantle Nawabi authority, disrupt local power structures *(taluqdars)*.
- **1857–58:** Rebellion erupts, becoming one of the largest anti-colonial uprisings in Indian history.
- **Lucknow** emerges as a key center, suffering severe siege and brutal British reprisals.
- After suppression, widespread state violence and punitive governance enacted across Awadh:
  - Mass executions, public hangings, village burnings.
  - Confiscation of over 500,000 acres as collective punishment.
  - Direct military rule was imposed in the immediate aftermath.

Lucknow as the Epicenter of Revolt and Repression

- Lucknow was the **political and symbolic heart** of resistance.
- Siege lasted **over three months** with heavy British casualties, fueling brutal reprisals.
- Repression in Lucknow was **deliberately severe** to serve as a public example and deterrent.

British Residency under siege (1857).

Two Theoretical Questions

1. How does vaccine take-up diffuse through a population?

2. Who was most likely to be "treated" by the 1857 shock?

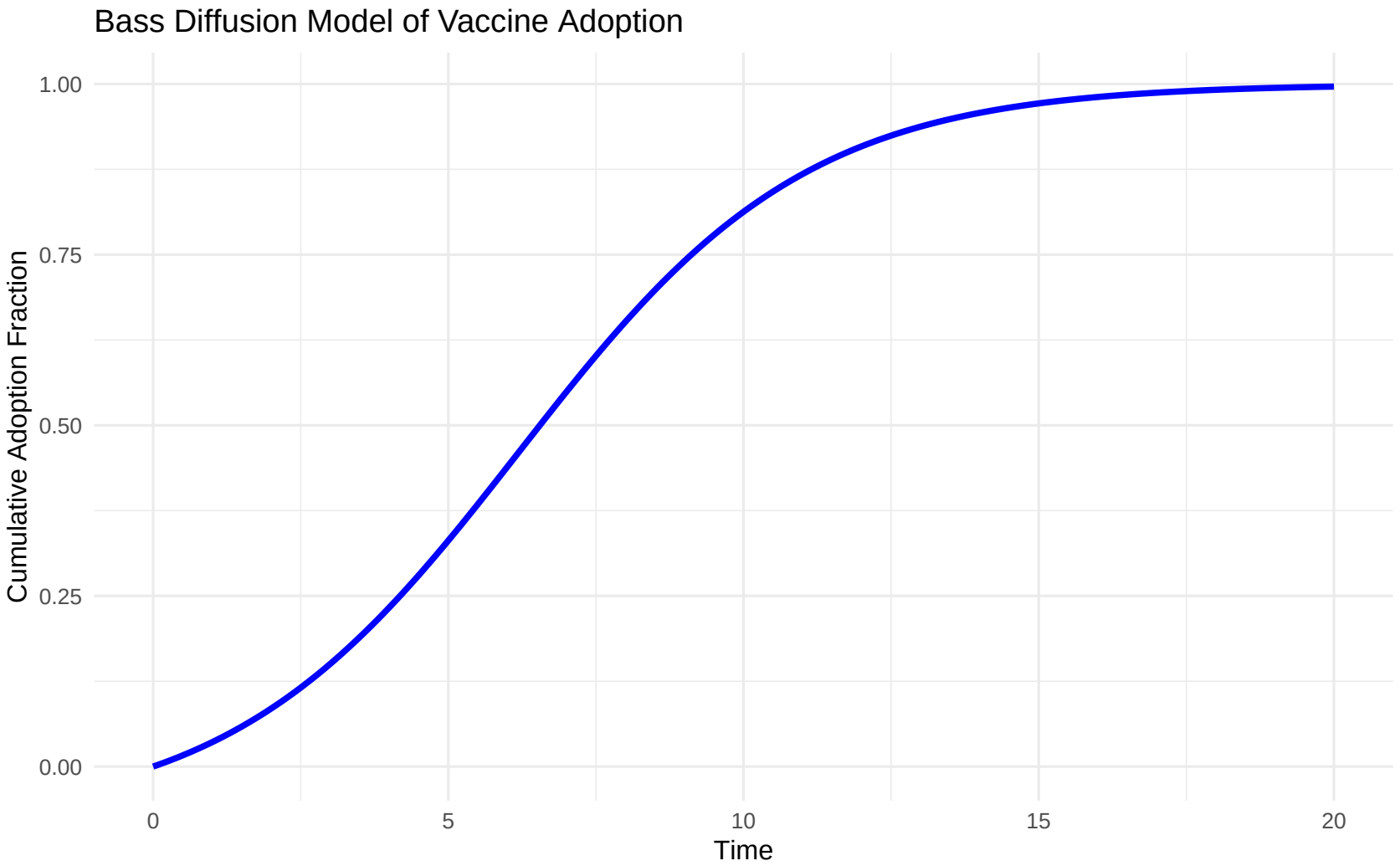
Modeling Vaccine Take-Up: The Bass Model

- The model divides the potential adopters into two categories:
  1. Innovators - those who adopt the vaccine without influence from others
  2. Imitators - those who adopt based on the influence of prior adopters
- The adoption rate over time can be modeled by the equation:

$$\frac{dF(t)}{dt} = (p + qF(t))(1 - F(t))$$

- Where:
  - p - coefficient of innovation
  - q - coefficient of imitation
  - $F(t)$ - cumulative fraction of adopters at time t

Graphical Representation



▸ Bombay Calibration

Solution: We have the vax rate for babies!

in the Punjab Province during the ye

6	7			
	PRIMARY VACCINA			
		Success		
erage ber of vaccinat- each nator.	Total.	Under one year.	One year and under six years.	
6	7	8	9	
..	46	29	5	
..	22	11	...	
..	38	22	1	
..	41	32	5	
..	5	5	...	
..	960	596	227	
..	101	61	5	
..	361	254	59	

Who was treated? I: Cohort

- Research in developmental psychology and economics suggests most influential window for forming attitudes toward institutions and societal trust is between ages roughly 10 and 25 years (Giuliano and Spilimbergo, 2025).

- The birth cohort 1832-47 would have been 10-25 years in 1857.

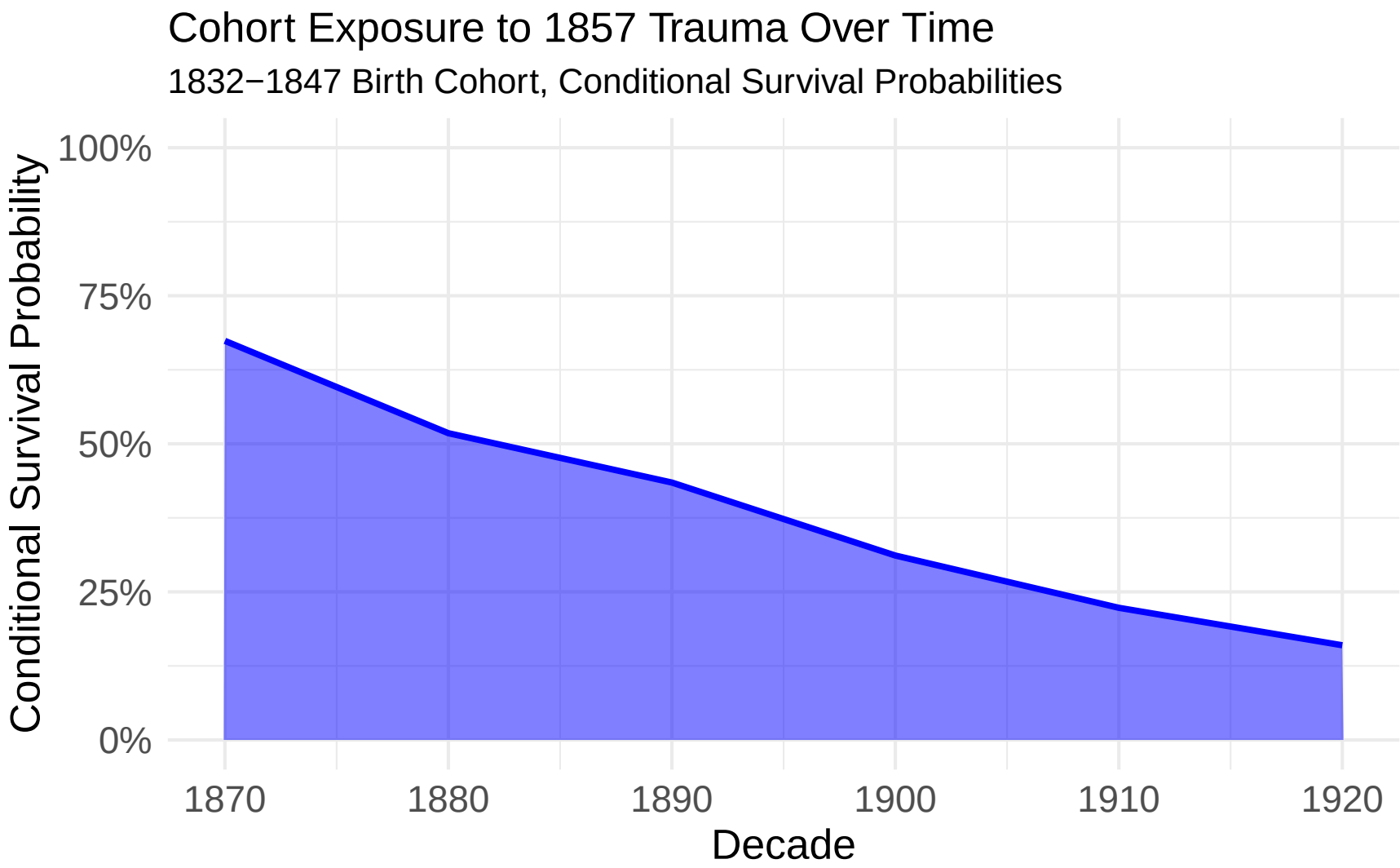
- So we construct a treatment exposure curve showing the probability an individual born in 1832-47 would be alive in a given year, *conditional on them being alive in 1857*.

- Standard exponential mortality function is $S(t) = e^{-\lambda t}$, where $S(t)$ is the probability of survival till period t and λ is the constant mortality over that time period.

- We can then calculate hazard rates for various ages as $\lambda_x \approx \frac{1}{e_x}$ using historical mortality data from Visaria and Visaria (1982).

- So, for example, conditional life expectancy at age 20 was $e_{20} = 38$, so $\lambda_{20} \approx 0.026$.

- We do this for multiple ages to construct our treatment exposure curve...


Cohort Exposure to 1857 Trauma Over Time
1832−1847 Birth Cohort, Conditional Survival Probabilities
Conditional Survival Probability
100%
75%
50%
25%
0%
1870
1880
1890
1900
1910
1920
Decade

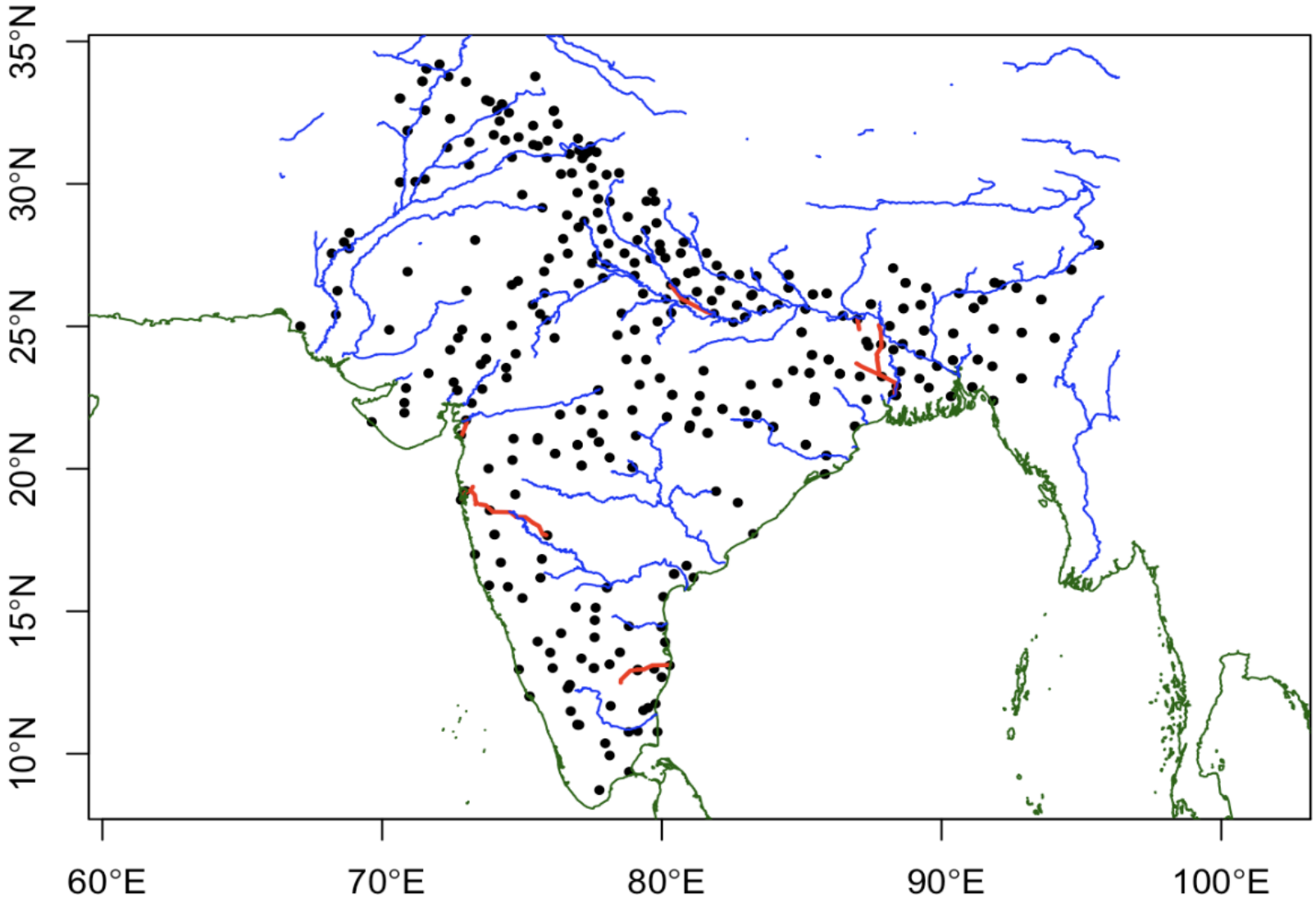
Who was treated? II: Distance

- Strategy is to look at whether people in districts "closer" to the massacres of 1857 are less likely to allow their children to be vaccinated by the British.

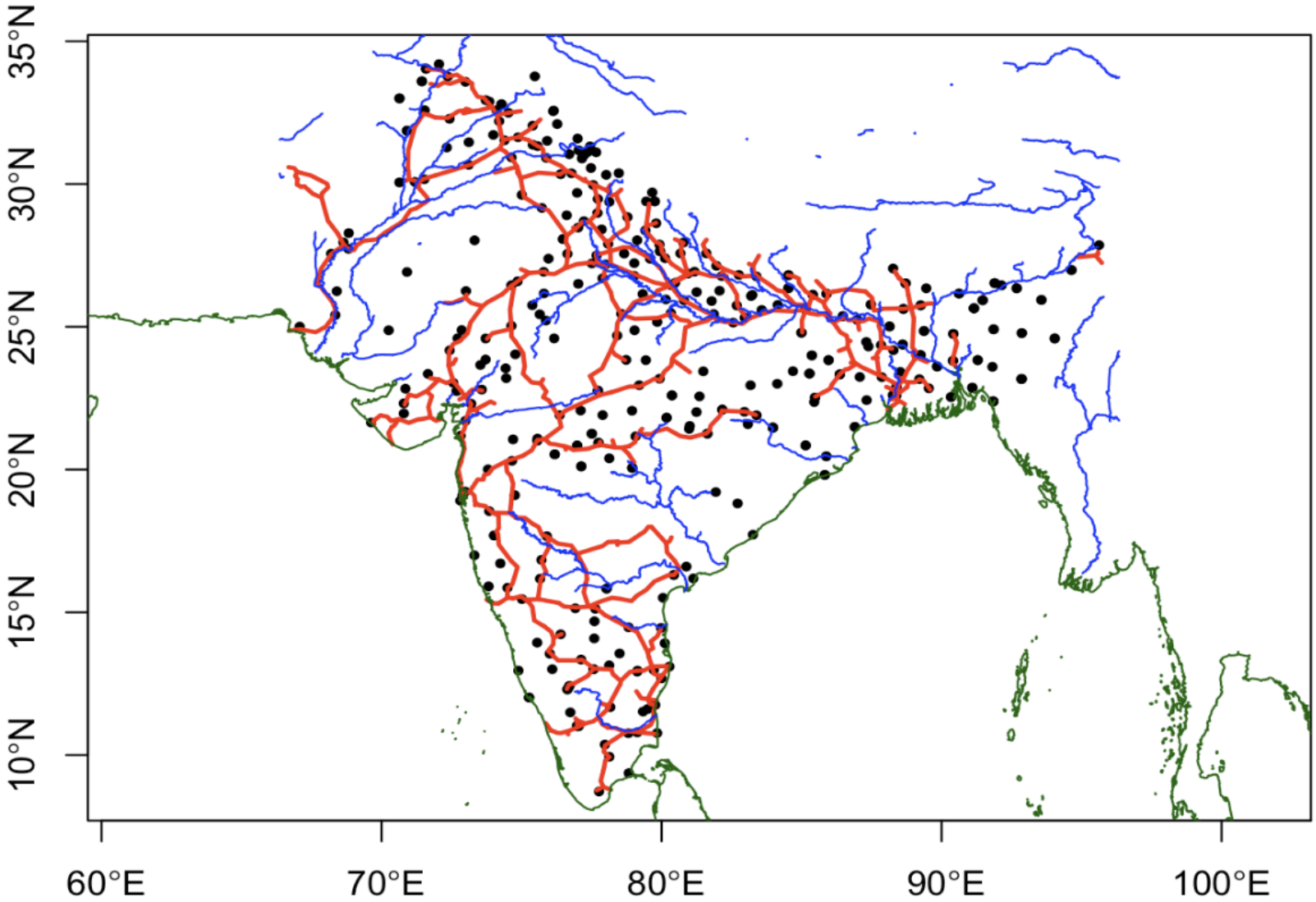
- We define "close" in two ways:
  - As the distance of the district from Lucknow measured along the least cost travel path
  - As the linguistic distance from Lucknow

Least Cost Travel Path

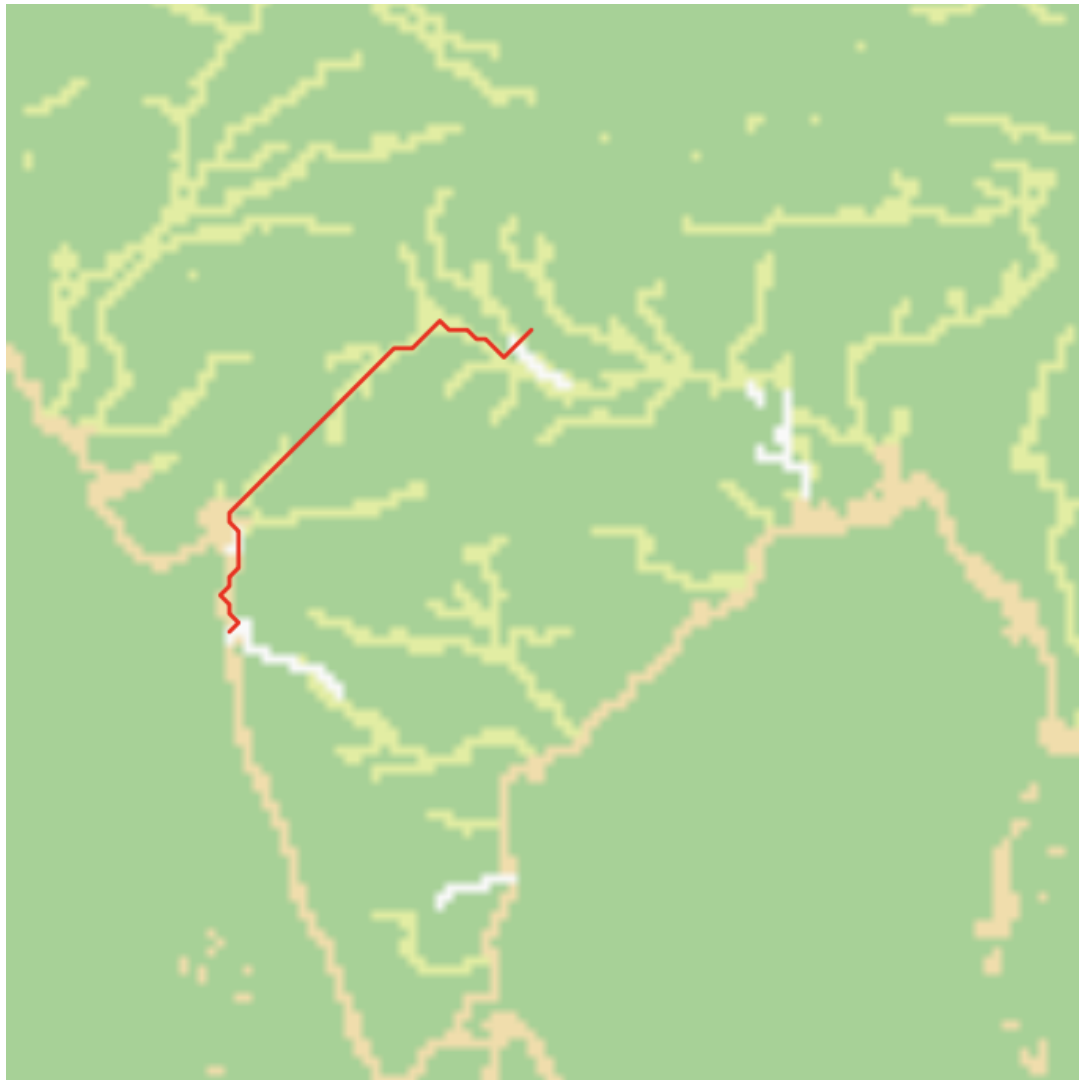
Transport Network 1860



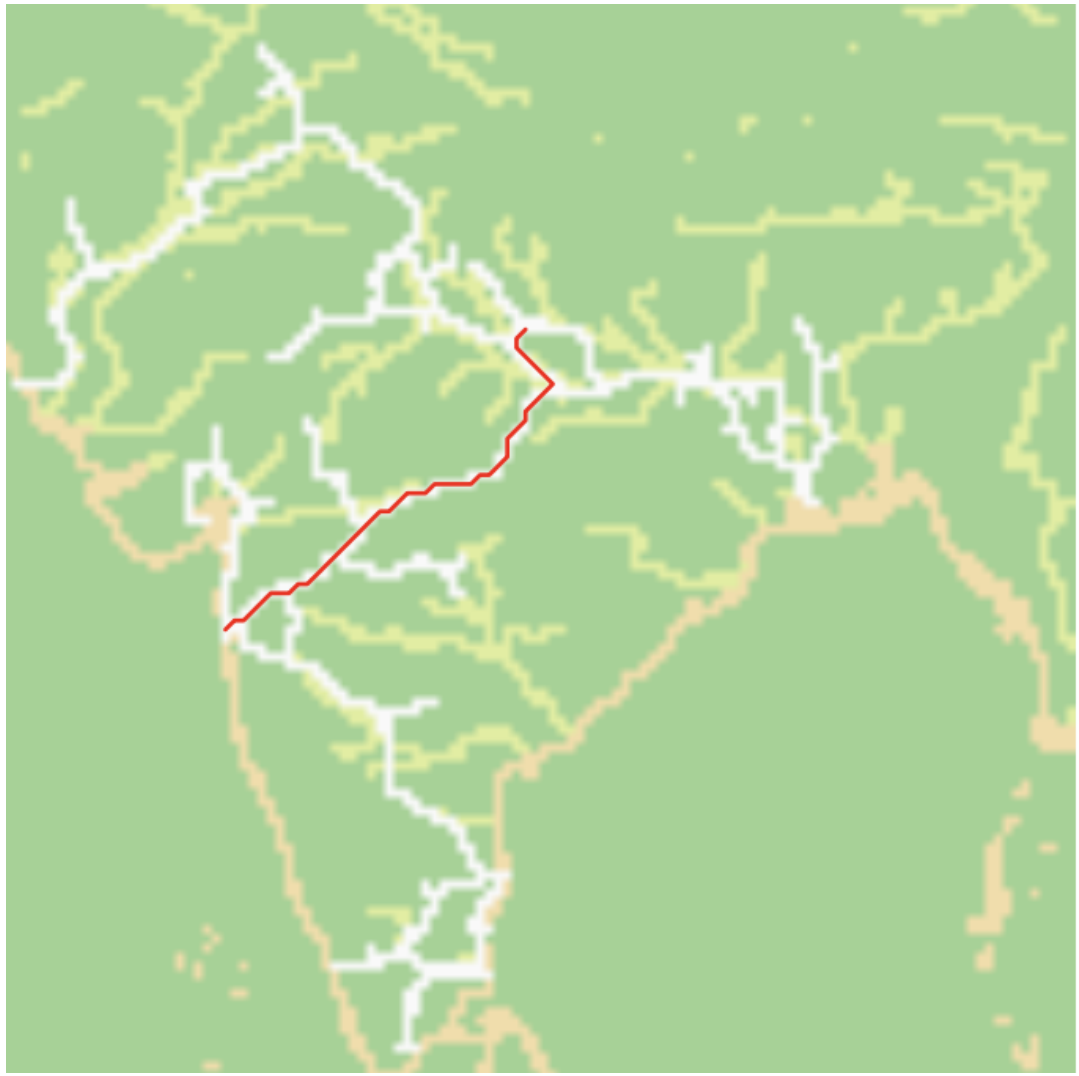
Transport Network 1890



Least Cost Travel Cost Lucknow to Bombay in 1860 = 126



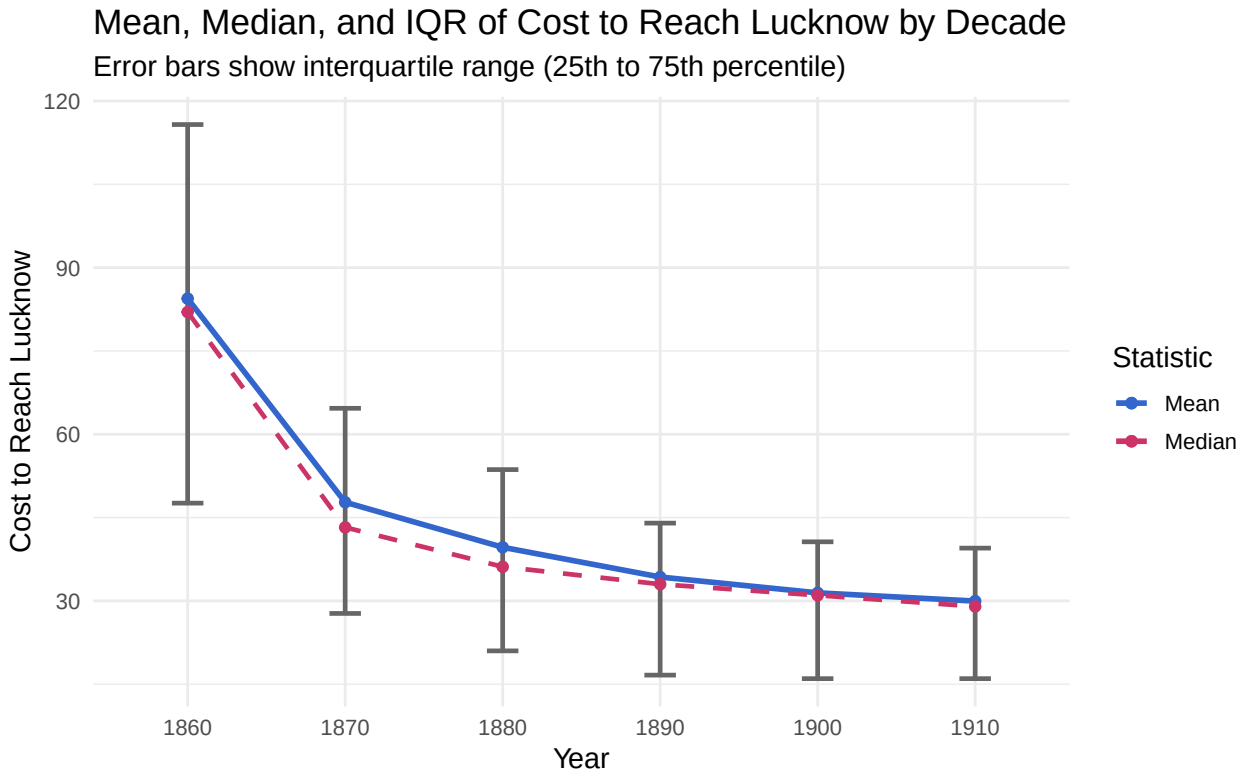
Least Cost Travel Cost Lucknow to Bombay in 1880 = 45



Capturing cumulative exposure over time

- Our primary measure of treatment exposure is:

$$\text{AccessLucknow}_{i,t} = \frac{1}{T} \sum_{s=1860}^{t} \text{PeriodCost}_{i,s} \tag{1}$$



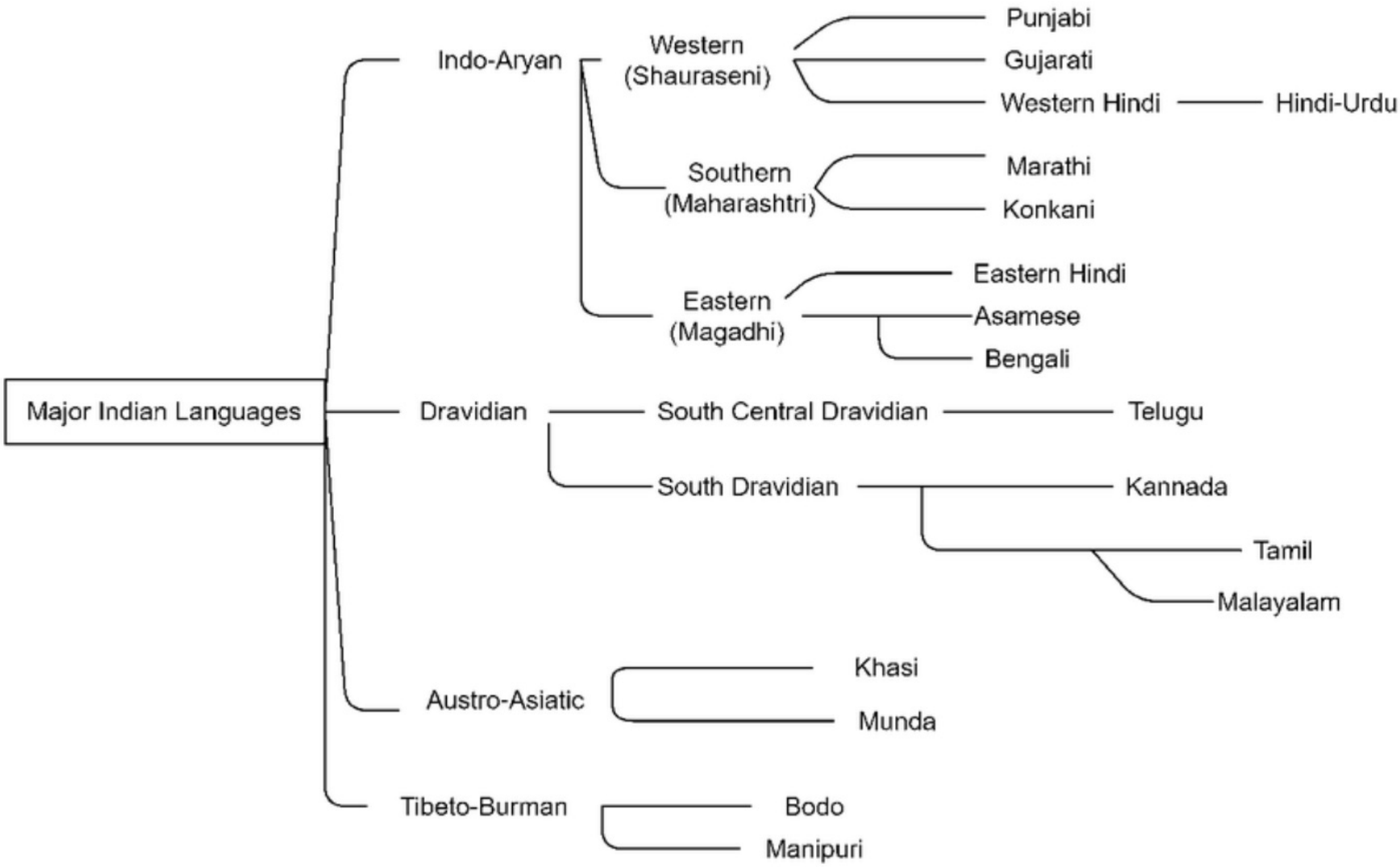
Linguistic Distance

Linguistic Distance

- Recent researc links linguistic distance to patterns of market development, trade, and cultural transmission (e.g. Fenske and Kala (2021), Spolaore and Wacziarg, (2018), Gupta (2014)).

- South Asia has remarkable linguistic diversity, containing numerous distinct languages from four major families: Indo-European, Dravidian, Austro-Asiatic, and Sino-Tibetan.

- We use data contained in the Ethnologue database to construct measures of the distance between Indian language groups (Eberhard et al., 2025).



Tree diagram to illustrate the language closeness of major Indian languages

Linguistic Distance

- A standard way to measure the similarity between two language groups is to count how many branches on the language tree they share (Fearon and Laitin, 2003; Esteban et al., 2012).

$$LinguisticDistance_{Lucknow} = 1 - \left(\frac{\text{SharedBranches}}{15} \right)^{\delta}$$

- $\delta$is a parameter indicating the sensitivity of the measure to differences in language groups. Larger $=$ more sensitive.

- We take $\delta = 0.5$

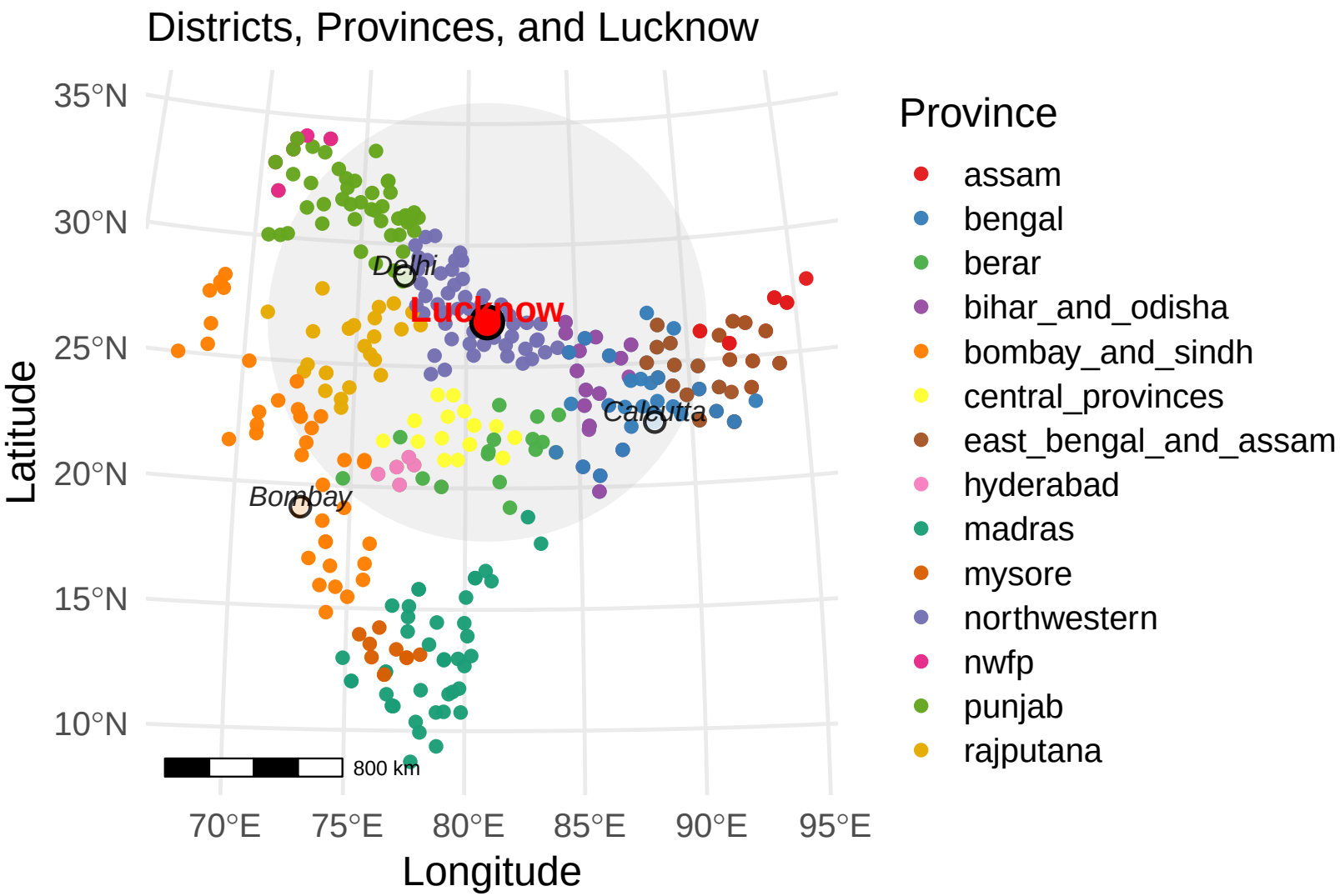
Data Sources and Construction

- **Vaccination Data (1870–1919)**: Annual district-level panel of infant vaccinations per capita (under age 1).
  - From provincial vaccination reports (e.g., Bengal, Bombay, Punjab, etc. . .).
  - Source: National Library of Scotland
- **Exposure to 1857 Repression**:
  - *Geographic Access*: Least-cost travel time to Lucknow via historical rail, river, coastal, and road networks.
  - *Cultural Distance*: Linguistic tree distance to Awadhi.

Main Sample, Controls, and Structure

- **Sample**: 7,746 district-year observations; unbalanced panel.
  - ≈240 districts observed annually by the 1910s.
  - Baseline sample: districts within 1,000 km of Lucknow.
- **Controls**:
  - Terrain ruggedness (Nunn and Puga (2012)), caloric suitability (Ozak and Galor (2016)), distance to sea/rivers (FAO).
  - Rail access (Donaldson (2018)), population, per capita vaccination spending from vaccination reports.
  - Province fixed effects capture administrative capacity and cultural variation.
- **Structure**: Data collapsed to decadal averages to reduce noise and reflect long-run patterns.

Smallpox Vaccination in India: 1870-1919



Note: Map shows British Indian districts, and the location of Lucknow. Baseline sample restricted to districts within 1,000 km of Lucknow.

Empirical Strategy

- We estimate the following using repeated cross-sections:

$$\text{Vaccination}_{it} = \beta \cdot \text{Proximity}_{it} + X'_{it}\gamma + \alpha_p + \varepsilon_{it}$$

- Outcome: Infant vaccinations per capita in district i and decade t.

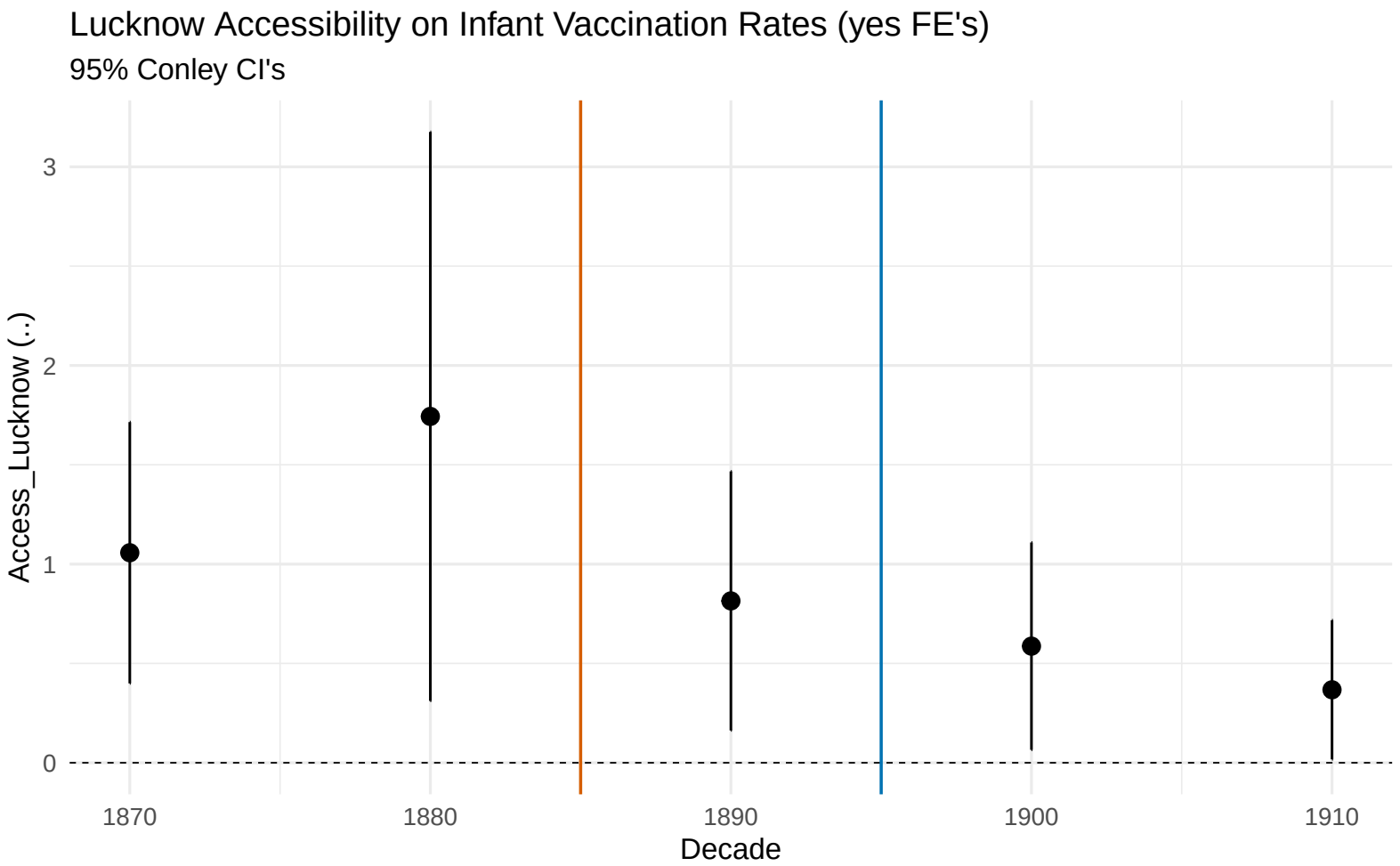
- Proximity is measured in two ways:
  - **Geographic access:** Average least-cost travel cost to Lucknow from 1860 to t.
  - **Linguistic distance:** Branch-weighted dissimilarity between district's language and Awadhi.

- X_{it} Includes geography, rail access, caloric suitability, population density, and vaccination spending.

- Province (α_p) fixed effects

- Conley standard errors reported.

Regression: Distance to Lucknow and Infant Vaccination



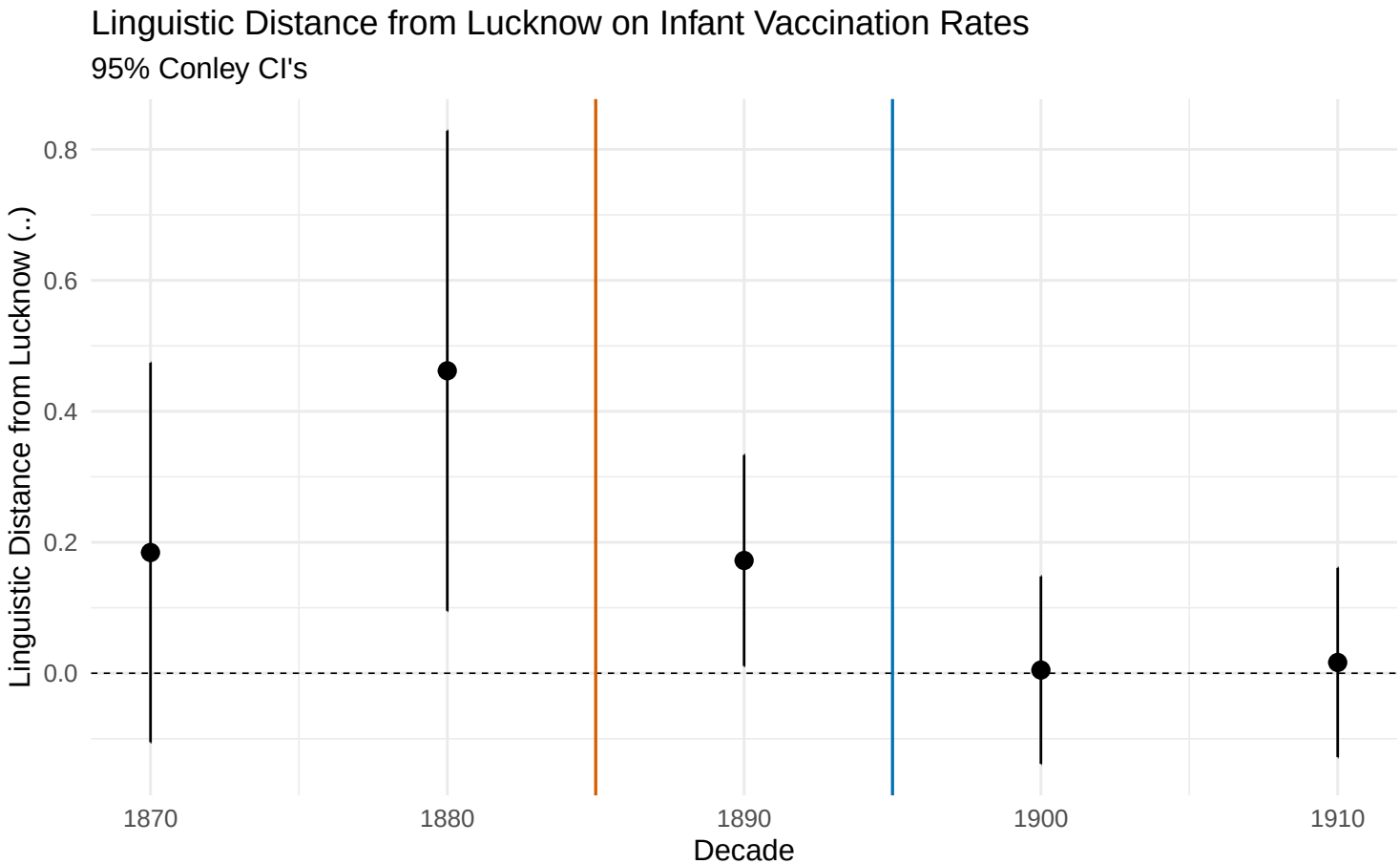
Notes: Estimated effect of least-cost travel access to Lucknow on infant vaccination rates (under 1), controlling for province fixed effects, geographic and infrastructure covariates. Conley SEs.

▸ Regs No FE's ▸ Regs Yes FE's

How big is the effect? Interpreting the 1880s Effect (Province FE Included)

- **Coefficient:** 1.74 in 1880 - positive and statistically significant
- All variables are standardized. A 1 SD *decrease* in travel cost to Lucknow $\Rightarrow$ **1.74 SD decrease** in infant vaccination rate.
- Mean infant vaccination rate (1880s): $\approx 2\%$ with a SD $\approx 1\%$.
- $\Rightarrow$ **Effect size:** -1.74 percentage points for every SD closer to Lucknow a district is.
- As expected, the coefficients shrink and become less significant as the treated cohort ages out.

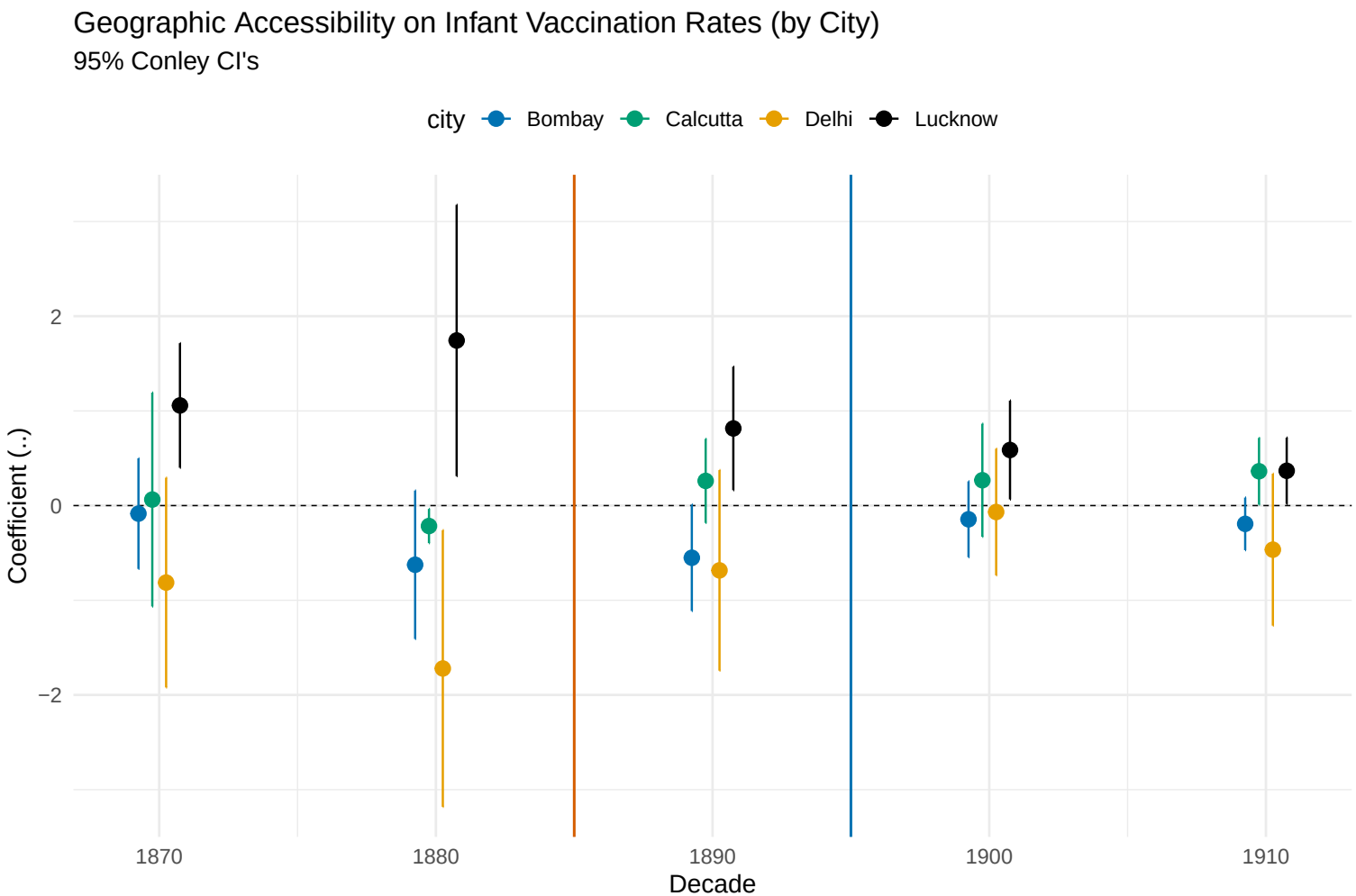
Linguistic Distance and Infant Vaccination



Note: All controls included. No province fixed effects. Conley SE's.

▸ Linguistic Regs

Placebo Tests: Other Major Districts



Note: Estimates from regressions of infant vaccination rates on distance to Delhi, Calcutta, and Bombay. Each specification includes province fixed effects and full set of controls. Conley SE's.

Placebo Tests

- Distance to other major colonial cities does not predict infant vaccination rates.
- No consistent pattern around Delhi, Calcutta, or Bombay.
- Helps rule out general urban proximity,transport access, or general massacres as explanations. Calcutta and Bombay important british centers of administration.
- Supports the argument that proximity to Lucknow captures exposure to the 1857 repression specifically.

Robustness

- Sample selection. Holding sample constant to just 1890 districts.
  ▸ Sample Selection
- Using linear instead of least cost distance to Lucknow. ▸ Linear Distance
- Varying the perimeter defining the sample from 500km to 1,500 km.
  ▸ Varying Distance

Conclusions

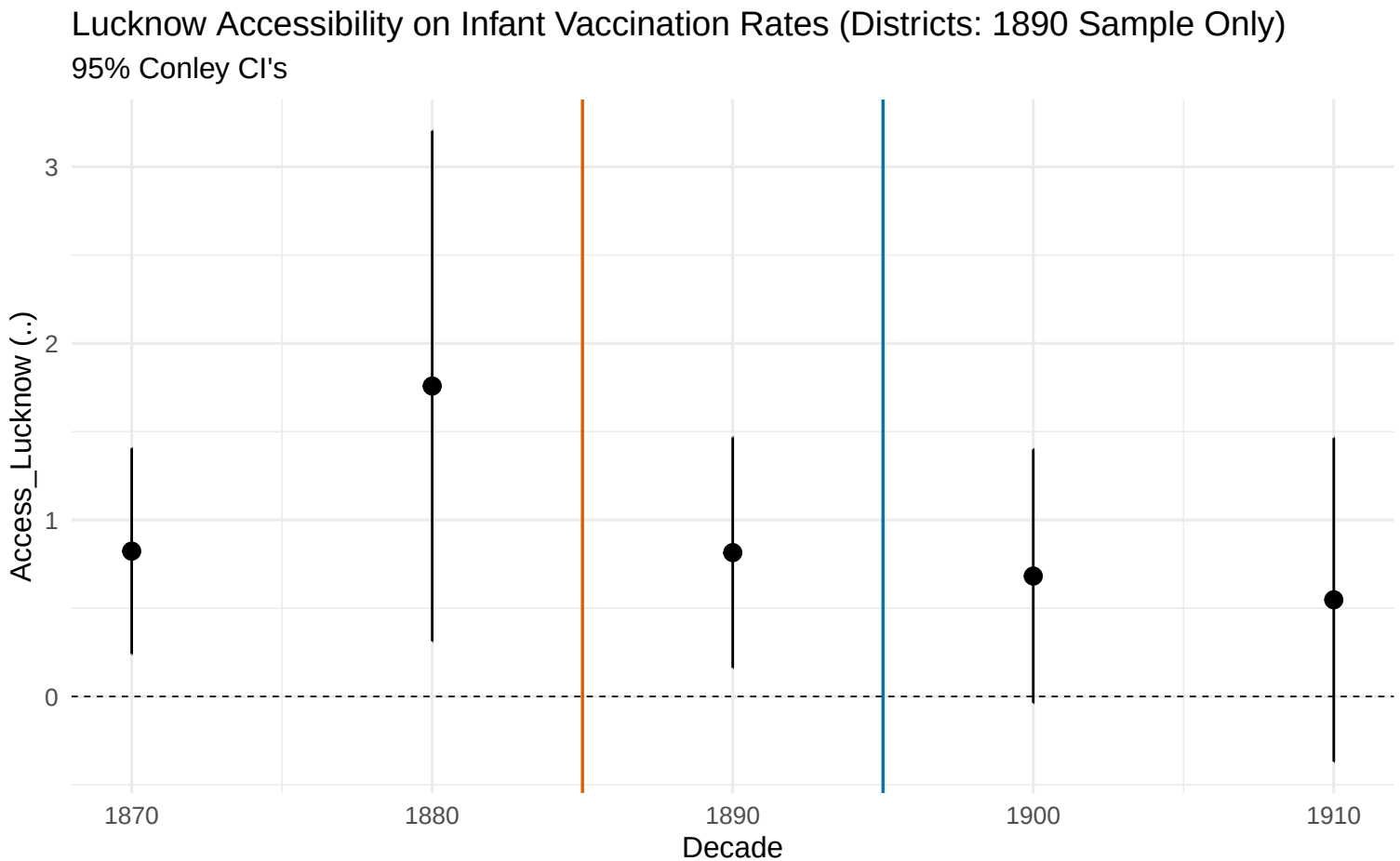
- Districts closer to Lucknow, the center of British reprisals in 1857, saw lower uptake of infant vaccination in the decades that followed.
- The effect was strongest when cohorts exposed to the revolt (ages 10–25 in 1857) reached decision-making age.
- Findings hold under alternative measures of exposure, placebo cities, and robustness checks.
- Historical experience with colonial violence reduced trust in the state and shaped later engagement with public health efforts.

Related Literature

- **Vaccine Hesitancy and Trust**: Alsan and Wanamaker (2018); Lowes and Montero (2021); Larson et al. (2018); Dubé et al. (2015).
- **The Impact of Historical Public Health Campaigns**: Cutler and Miller (2005); Alsan and Goldin (2019); Dupas (2011); Banthia and Dyson (1999); Arnold et al. (1993); Bhattacharya et al. (2005).
- **Economic and Political Impacts of Colonialism**: Acemoglu et al. (2001); Acemoglu et al. (2002); Nunn (2008); Dell (2010); Banerjee and Iyer (2005); Bose and Jalal (2022); Roy (2020); Dell and Olken (2020).

Extra Slides

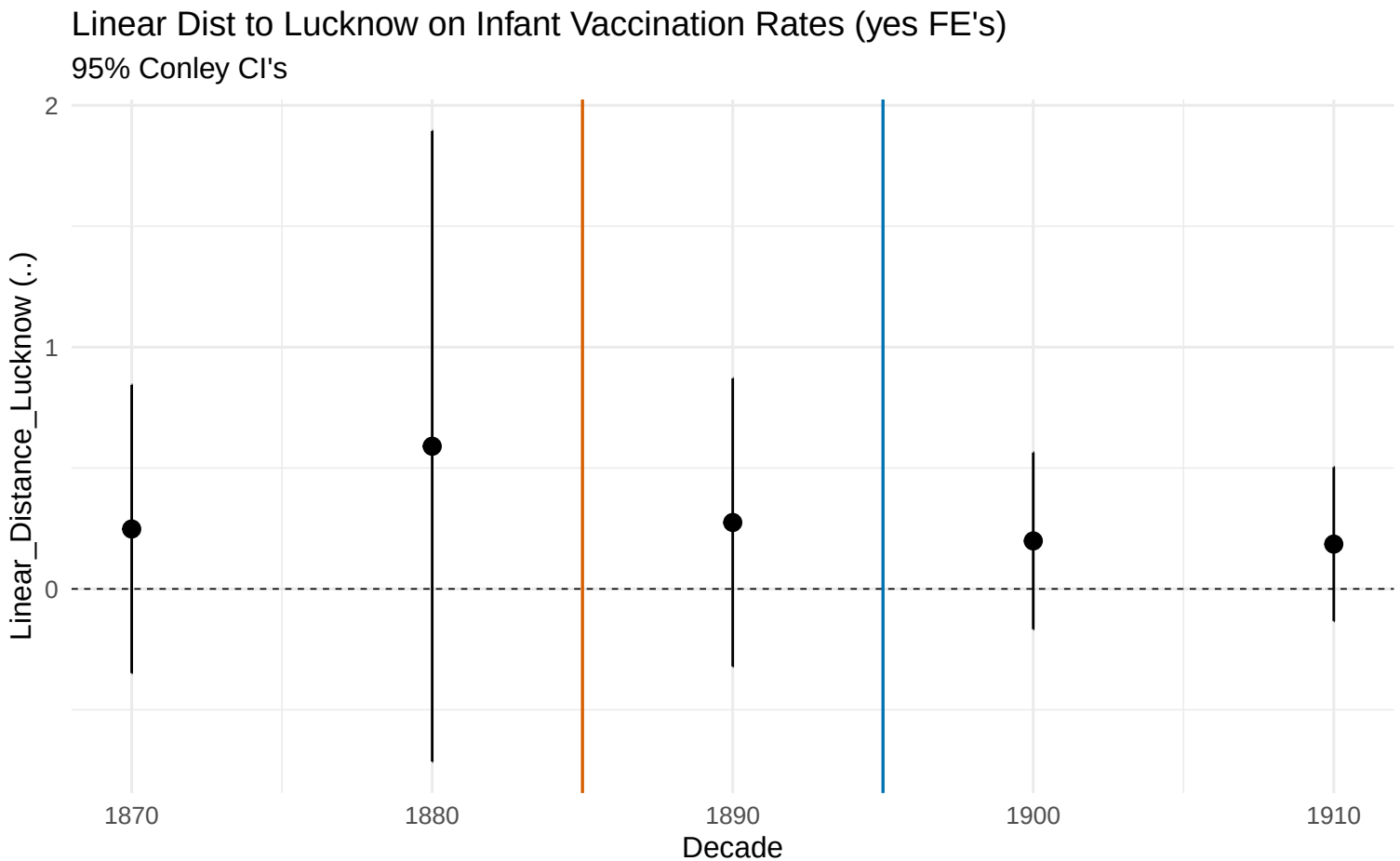
Robustness: Sample Restricted to 1890 Districts



Note: Results restricting sample to districts observed in the 1890s. Full controls and fixed effects included. Conley SE's.

Return

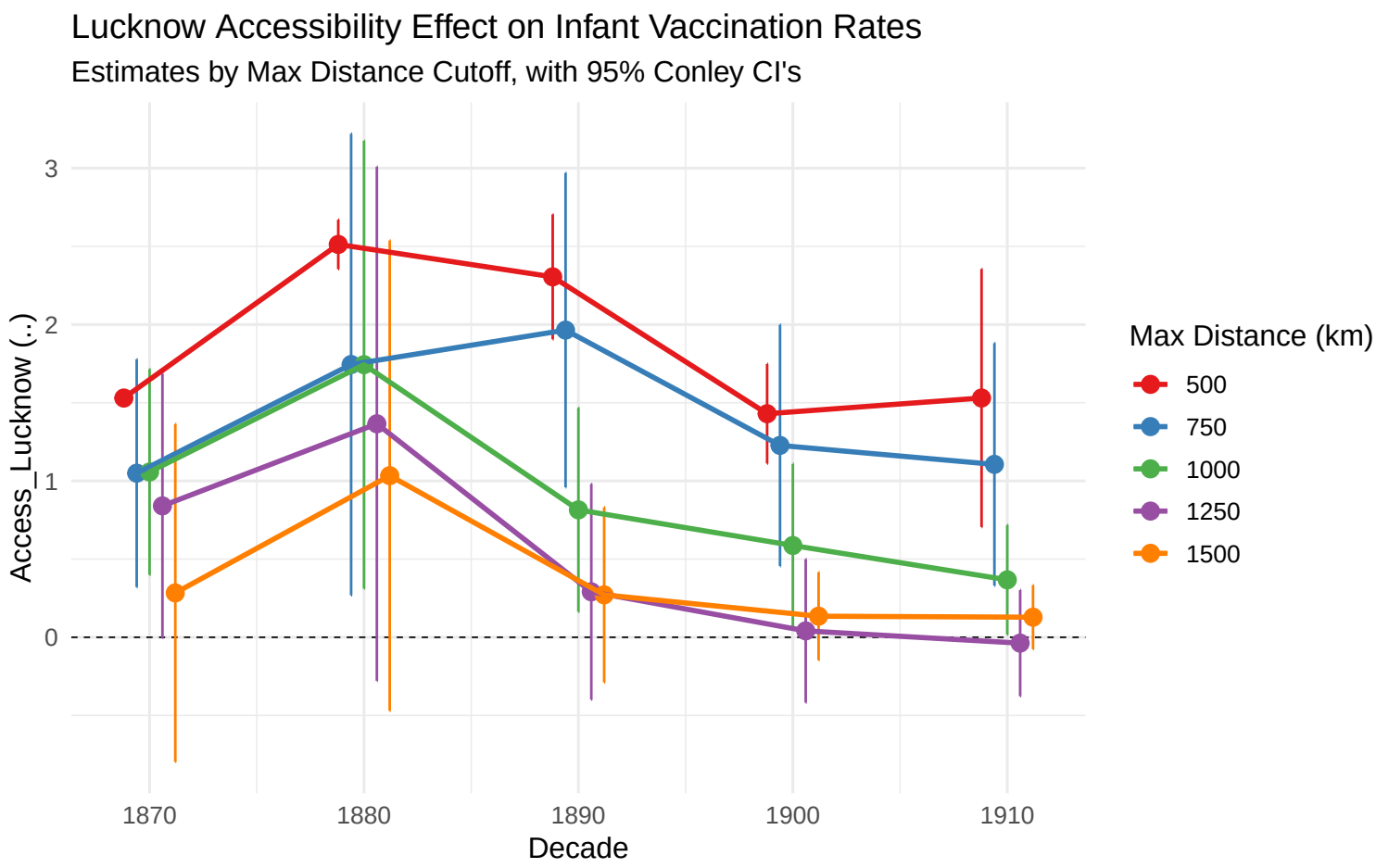
Robustness: Linear vs. Least-Cost Distance



Note: Effect of linear (Euclidean) distance to Lucknow on infant vaccination rates. Specifications mirror main model.

Return

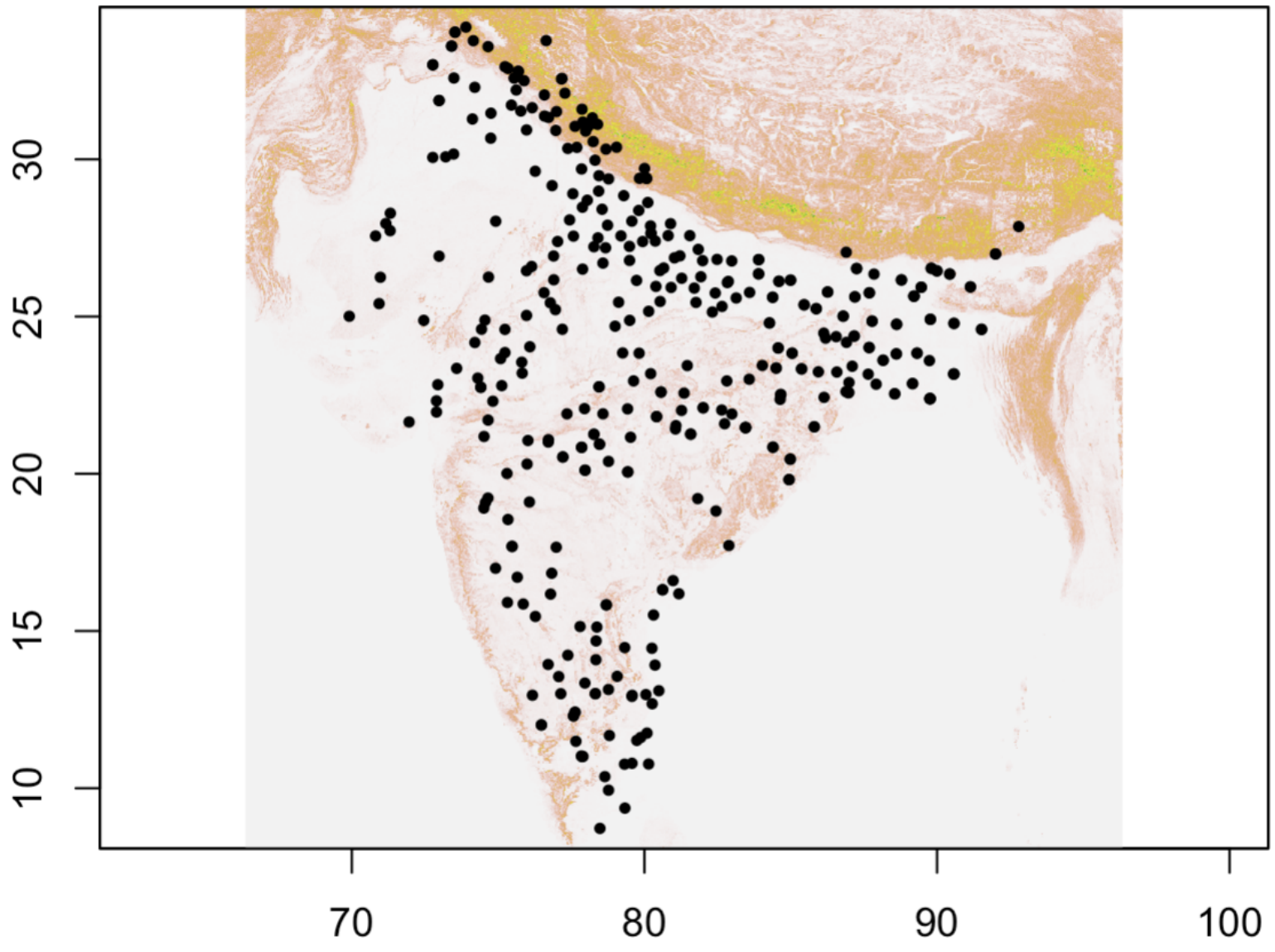
Robustness: Varying Distance Thresholds



Note: Estimated coefficients on proximity to Lucknow using alternative sample cutoffs (500–1,500 km). Full controls included. Conley SE's.

▶ Return

Terrain Ruggedness



Caloric Suitability

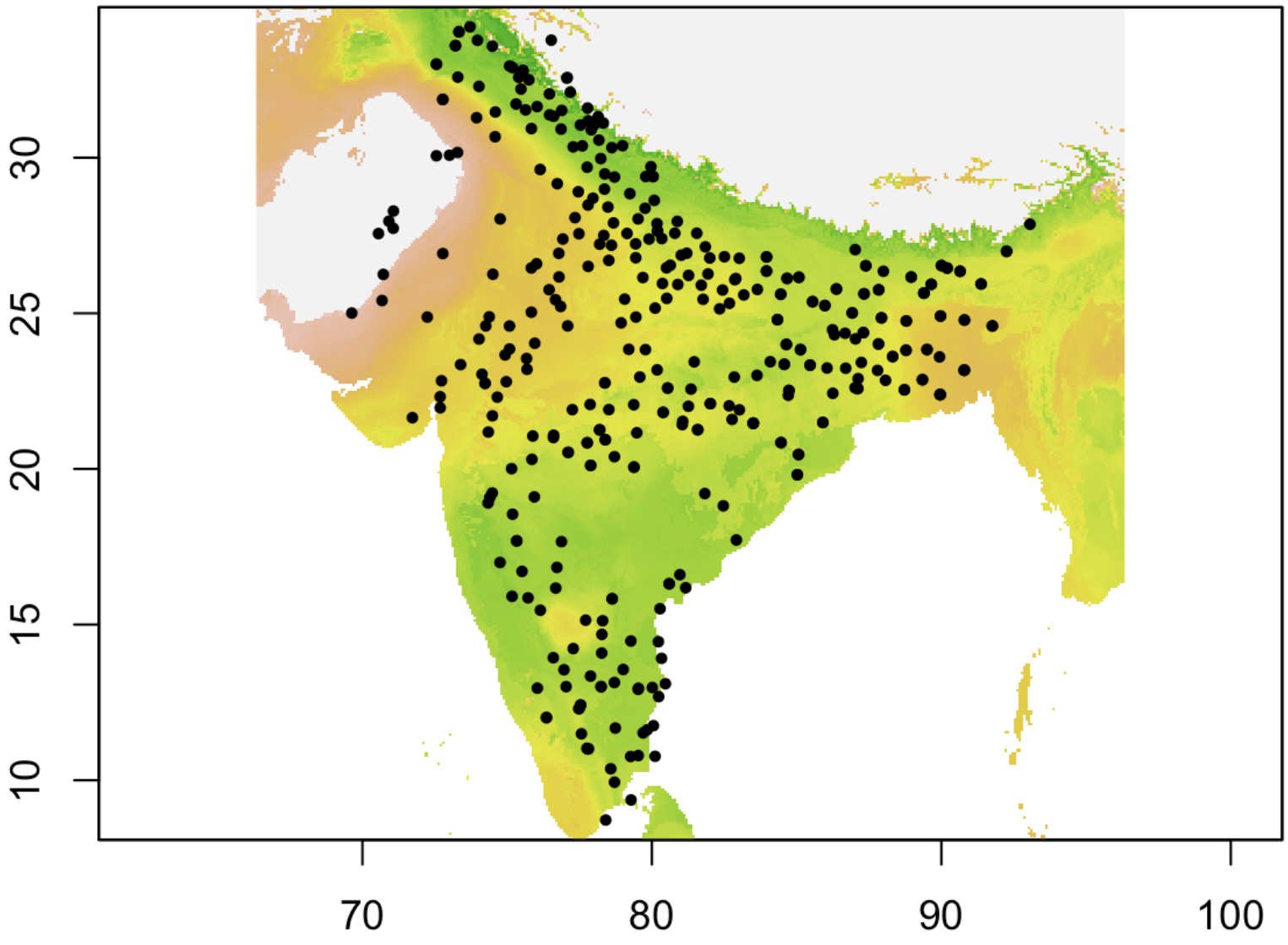
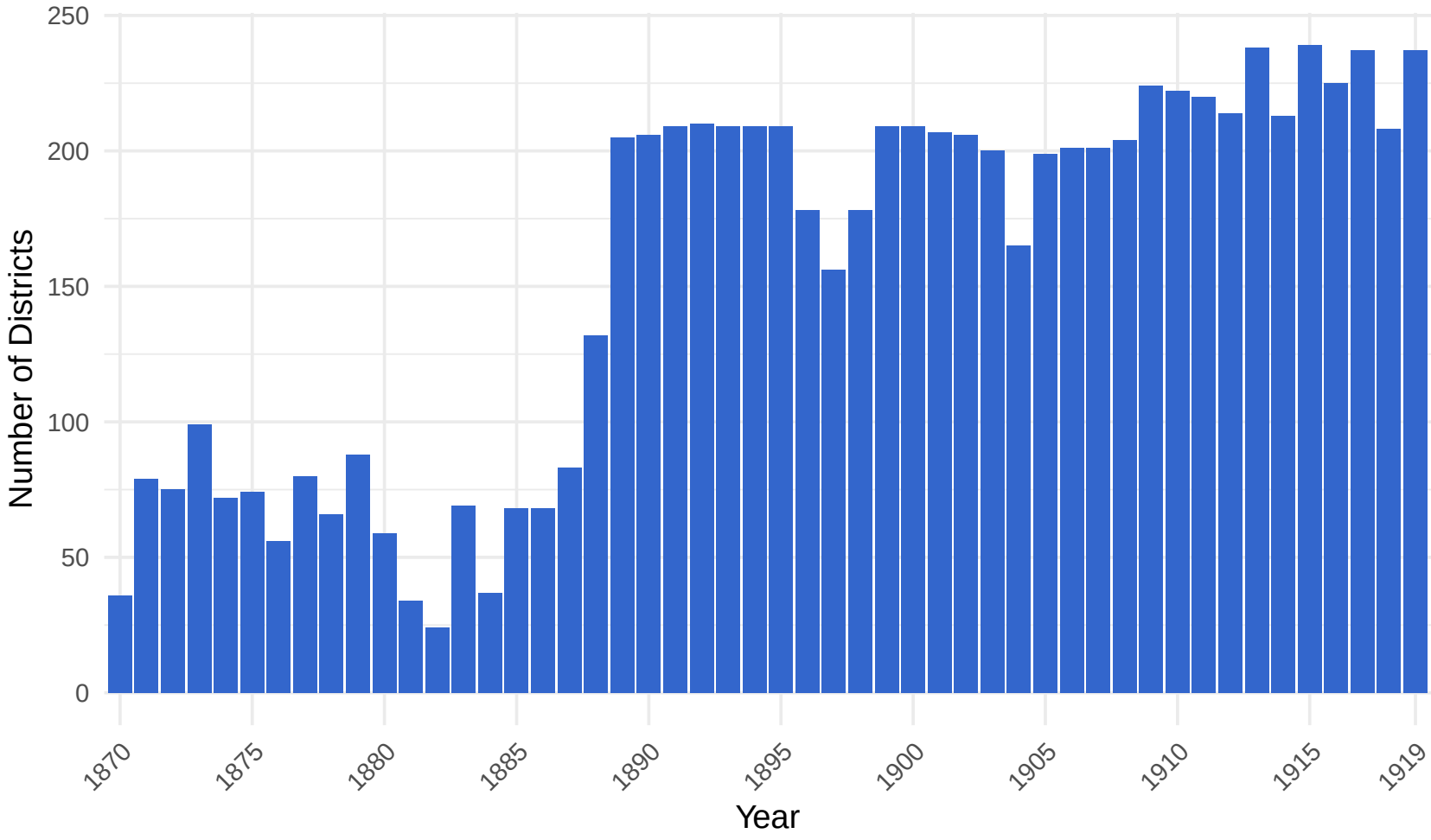


Table: Descriptive Statistics

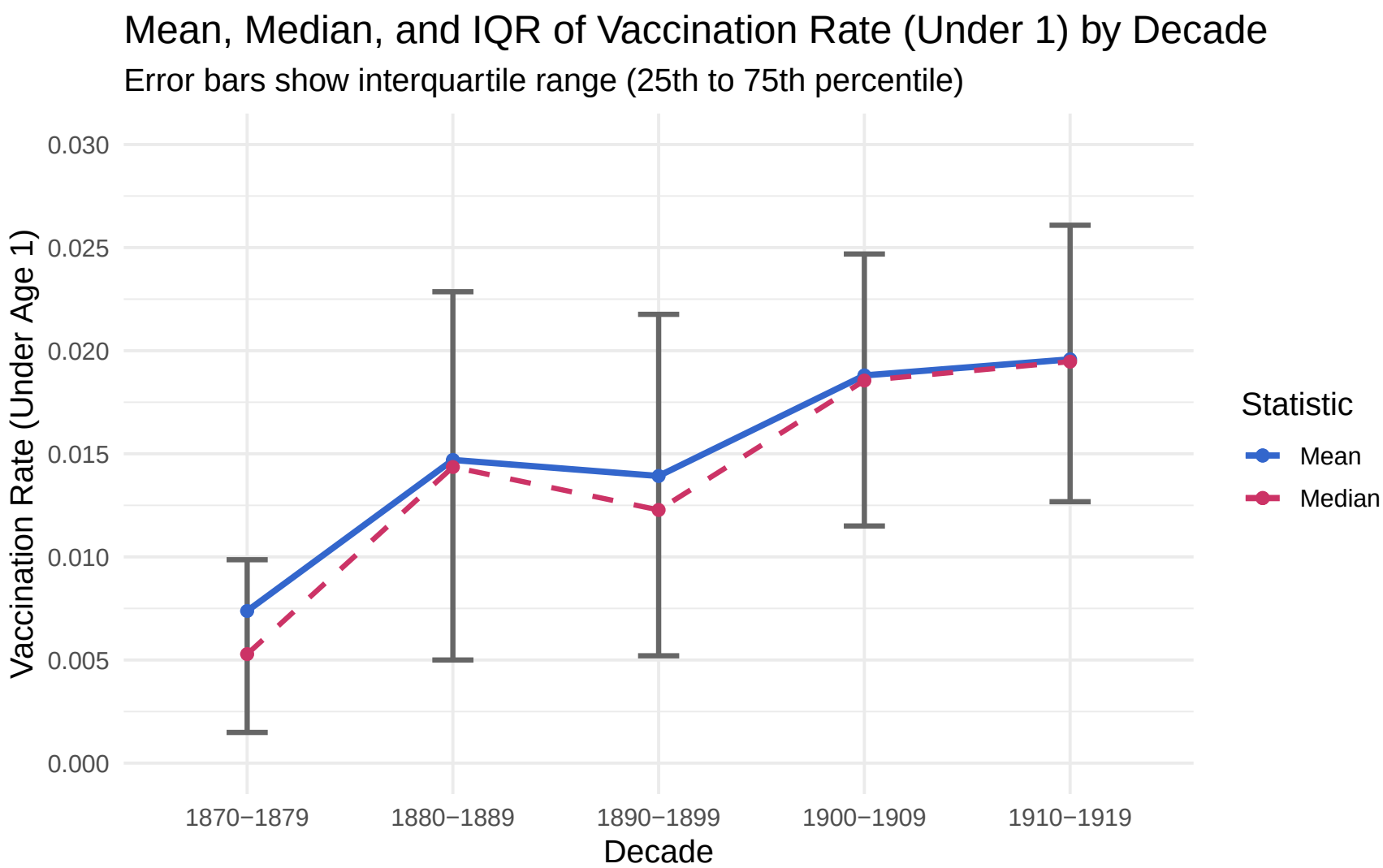
Variable	Mean	SD	Min	Max	N
Vaccination Rate Under Age 1	0.02	0.01	0.00	0.17	7622
Ruggedness (50km)	37561.17	63752.76	630.12	494843.44	7746
Caloric Avg (50km)	3232.31	829.88	0.00	4732.34	7746
Min Distance to Sea (m)	506135.70	347331.35	92.92	1259194.97	7746
Min Distance to River (m)	53670.59	51467.02	408.08	253496.63	7746
Total Population	925889.37	648094.64	2853.00	9138818.00	7694
Spending per Capita	0.01	0.01	0.00	0.15	6358
Distance to Rail (1870)	122202.51	128888.87	20.63	650085.13	7746
Distance to Rail (1880)	83098.38	109358.55	20.63	537276.94	7746
Distance to Rail (1890)	37774.42	60891.85	20.63	326556.30	7746
Distance to Rail (1900)	19303.14	28446.67	20.63	173759.48	7746
Distance to Rail (1910)	12832.09	21618.30	20.63	166421.10	7746
Access to Lucknow (1870)	66.09	37.09	0.00	170.04	7746
Access to Lucknow (1880)	57.28	32.62	0.00	145.33	7746
Access to Lucknow (1890)	51.53	29.64	0.00	131.66	7746
Access to Lucknow (1900)	47.52	27.45	0.00	123.05	7746
Access to Lucknow (1910)	44.59	25.86	0.00	117.21	7746
Vaccination Rate	0.04	0.02	0.00	0.39	7694
Distance to Lucknow	748871.71	421194.42	2191.69	2043062.67	7746
Linguistic Distance to Lucknow	0.43	0.29	0.01	1.00	5044
Smallpox Deaths per 1000	1.13	9.18	0.00	349.00	5831

Count of Districts Vaccinated by Year



Return

Vaccination Rate Under 1 Years Old



Return

Table 4: Access to Lucknow on Infant Vaccination Rates by Decade (no FE's)

	1870s	1880s	1890s	1900s	1910s
Access to Lucknow	0.711*	1.412***	0.669***	0.279*	0.192
	(0.295)	(0.262)	(0.123)	(0.118)	(0.119)
Ruggedness	0.288	0.068	−0.239***	−0.338***	−0.291***
	(0.285)	(0.163)	(0.070)	(0.072)	(0.076)
Caloric suitability	−0.199	−0.854**	−0.162+	0.106	0.107
	(0.226)	(0.263)	(0.095)	(0.081)	(0.077)
Distance to sea	0.099	0.834**	0.613***	0.358***	0.357***
	(0.288)	(0.262)	(0.086)	(0.082)	(0.079)
Distance to river	0.082	0.283**	0.372***	0.107	0.119
	(0.107)	(0.094)	(0.077)	(0.069)	(0.072)
Total population	−0.459***	−0.361**	−0.396***	−0.186*	−0.199**
	(0.130)	(0.108)	(0.070)	(0.071)	(0.065)
Spending per capita	0.063	−0.035	0.125	0.366***	0.213**
	(0.094)	(0.090)	(0.102)	(0.086)	(0.079)
Distance to railroad	−0.245*	−0.375**	−0.126	0.077	−0.139*
	(0.098)	(0.118)	(0.093)	(0.068)	(0.066)
Num.Obs.	79	78	146	203	206
R2	0.580	0.627	0.567	0.290	0.269

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All variables standardized. Standard errors are Conley (spatial HAC).

Table 5: Access to Lucknow on Infant Vaccination Rates by Decade (yes FE's)

	1870s	1880s	1890s	1900s	1910s
Access to Lucknow	1.057*	1.743+	0.815*	0.587+	0.367+
	(0.336)	(0.731)	(0.333)	(0.266)	(0.179)
Ruggedness	0.106	0.009	−0.236+	−0.322*	−0.216+
	(0.054)	(0.184)	(0.114)	(0.112)	(0.100)
Caloric suitability	−0.068+	−0.856*	−0.228	−0.128	−0.070
	(0.026)	(0.198)	(0.163)	(0.129)	(0.074)
Distance to sea	−0.138	0.840	0.889*	0.919**	0.621**
	(0.442)	(0.450)	(0.355)	(0.249)	(0.192)
Distance to river	0.185	0.237***	0.226***	−0.006	−0.021
	(0.095)	(0.023)	(0.034)	(0.069)	(0.086)
Total population	−0.444*	−0.344**	−0.301**	−0.148	−0.137*
	(0.116)	(0.044)	(0.066)	(0.089)	(0.050)
Spending per capita	0.093+	0.001	0.082	0.257*	0.092
	(0.040)	(0.022)	(0.122)	(0.100)	(0.105)
Distance to railroad	−0.369*	−0.489***	−0.069	0.041	−0.153
	(0.100)	(0.046)	(0.064)	(0.151)	(0.121)
Num.Obs.	79	78	146	203	206
R2	0.649	0.690	0.633	0.437	0.452

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All variables standardized. Standard errors are Conley (spatial HAC).

Table 6: Linguistic Distance from Lucknow on Infant Vaccination Rates by Decade (no FE's)

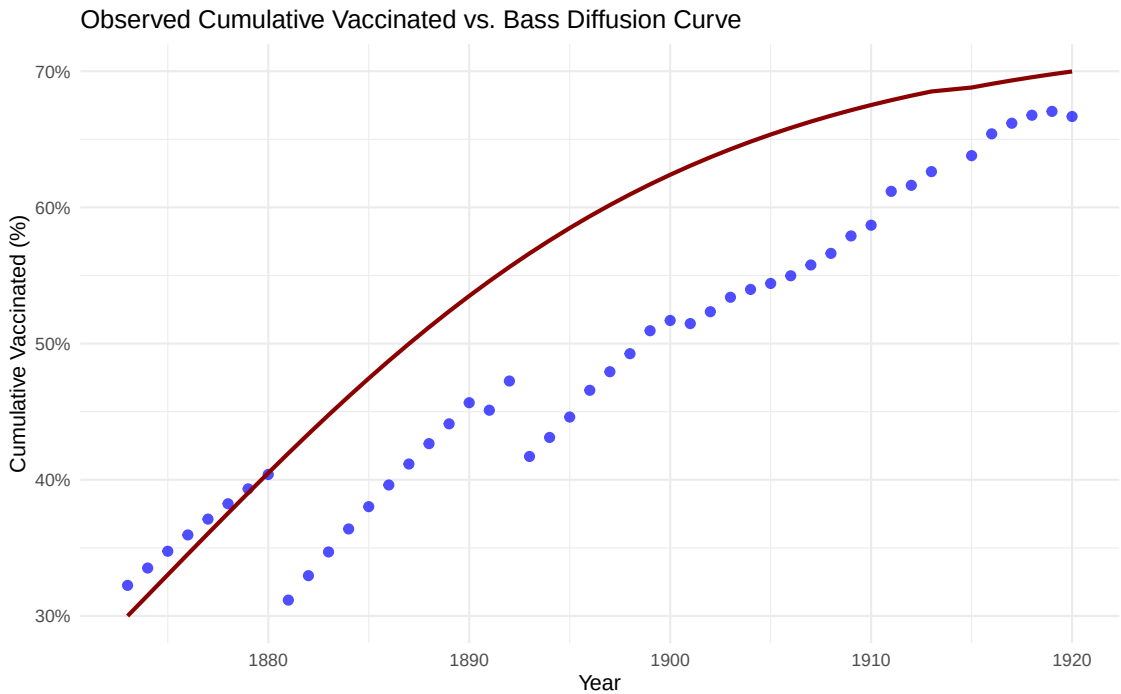
	1870s	1880s	1890s	1900s	1910s
Linguistic distance from Lucknow	0.184	0.462*	0.172*	0.005	0.016
	(0.148)	(0.187)	(0.082)	(0.073)	(0.074)
Ruggedness	−0.394	−0.027	−0.170	−0.220+	−0.309*
	(0.407)	(0.235)	(0.161)	(0.125)	(0.121)
Caloric suitability	−0.159	−0.253	−0.095	0.035	0.106+
	(0.161)	(0.208)	(0.082)	(0.058)	(0.059)
Distance to sea	0.005	0.598*	0.461***	0.293***	0.267***
	(0.193)	(0.264)	(0.095)	(0.071)	(0.071)
Distance to river	0.064	0.161+	0.203**	0.139*	0.170**
	(0.096)	(0.084)	(0.074)	(0.063)	(0.063)
Total population	−0.730***	−0.439***	−0.353***	−0.371***	−0.399***
	(0.142)	(0.095)	(0.074)	(0.072)	(0.074)
Spending per capita	0.203+	0.012	0.239*	0.105	−0.008
	(0.109)	(0.111)	(0.092)	(0.092)	(0.081)
Distance to railroad	−0.246**	−0.167+	−0.166*	−0.083	−0.264**
	(0.082)	(0.091)	(0.075)	(0.069)	(0.080)
Num.Obs.	80	78	136	164	156
R2	0.465	0.442	0.506	0.394	0.425

+ $p < 0.1$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

All variables standardized. Standard errors are Conley (spatial HAC).

Does our data match this pattern? Kolaba District

- 1872 Bombay census: Around 30% of the population had been vaccinated.
- We structurally estimate the diffusion function using the Bombay (Kolaba) data.



Return