



Using Big Data to Solve Economic and Social Problems

Professor Raj Chetty

Head Section Leader: Gregory Bruich, Ph.D.

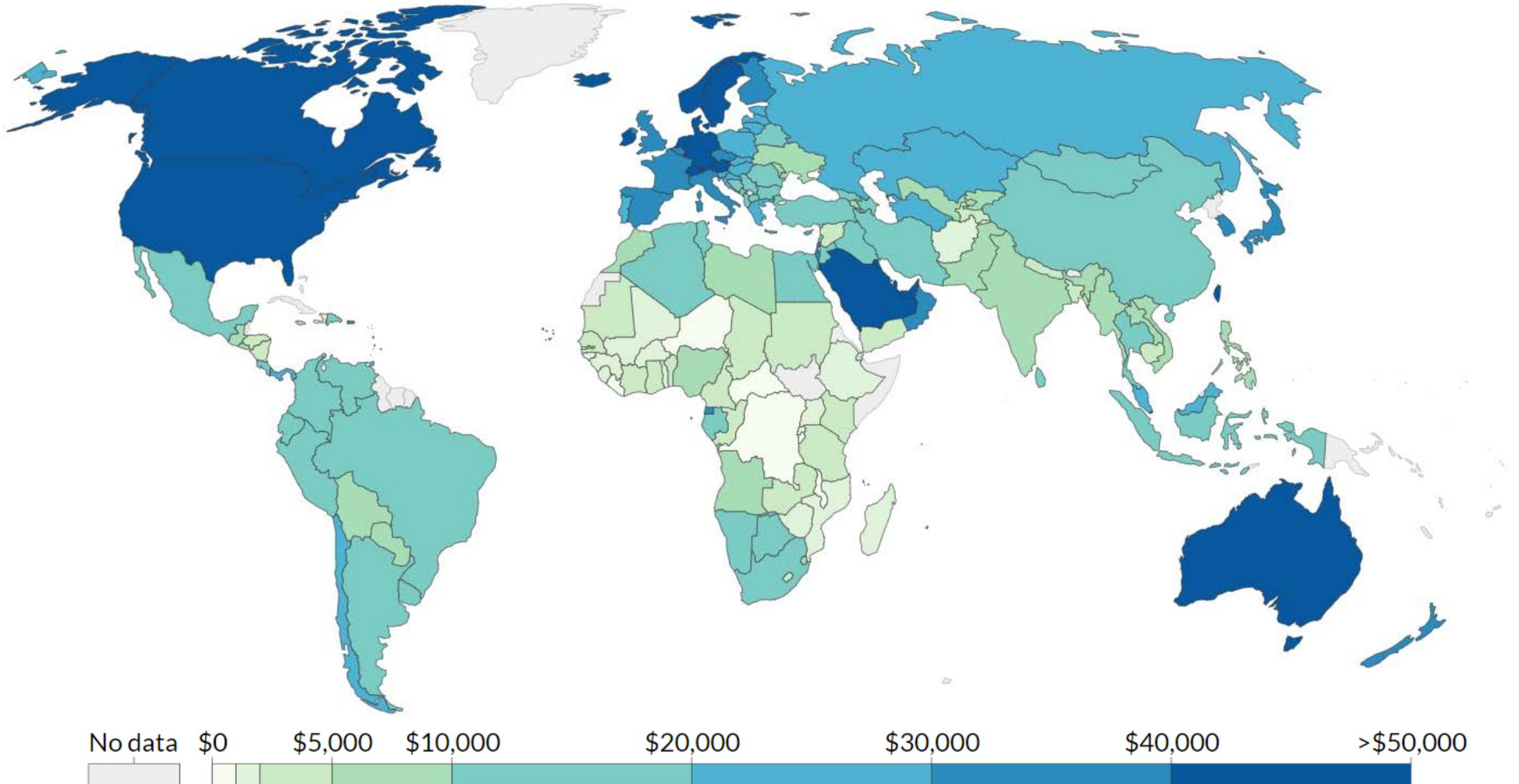
Spring 2019



HARVARD
UNIVERSITY



Gross Domestic Product Per Capita Across the World in 2016



Improving Outcomes in Developing Countries

- Development economics focuses on improving outcomes in developing countries
 - Not just doing the same work in lower income countries: questions and answers differ because of differences in technologies and institutions
- In this lecture, illustrate two types of approaches in modern development economics:
 1. Policy evaluation: improve design of specific policies
 2. Institutional design: improve basic institutions and government structures

Policy Evaluation Through Randomized Experiments

- Many recent studies evaluate the impacts of policies to improve outcomes in developing countries using randomized trials
- Duflo et al. research on immunization discussed in last lecture is one example of this approach
- Here, talk about a second example that has had substantial impact: policies to deworm children

The Impacts of Deworming Children

- Intestinal worms affect 25% of the world population
- Worm infections can reduce health and educational achievement
- Miguel and Kremer (2004) study the impacts of a deworming pill on health and educational outcomes in Busia, Kenya

Estimating The Effects of Deworming Children

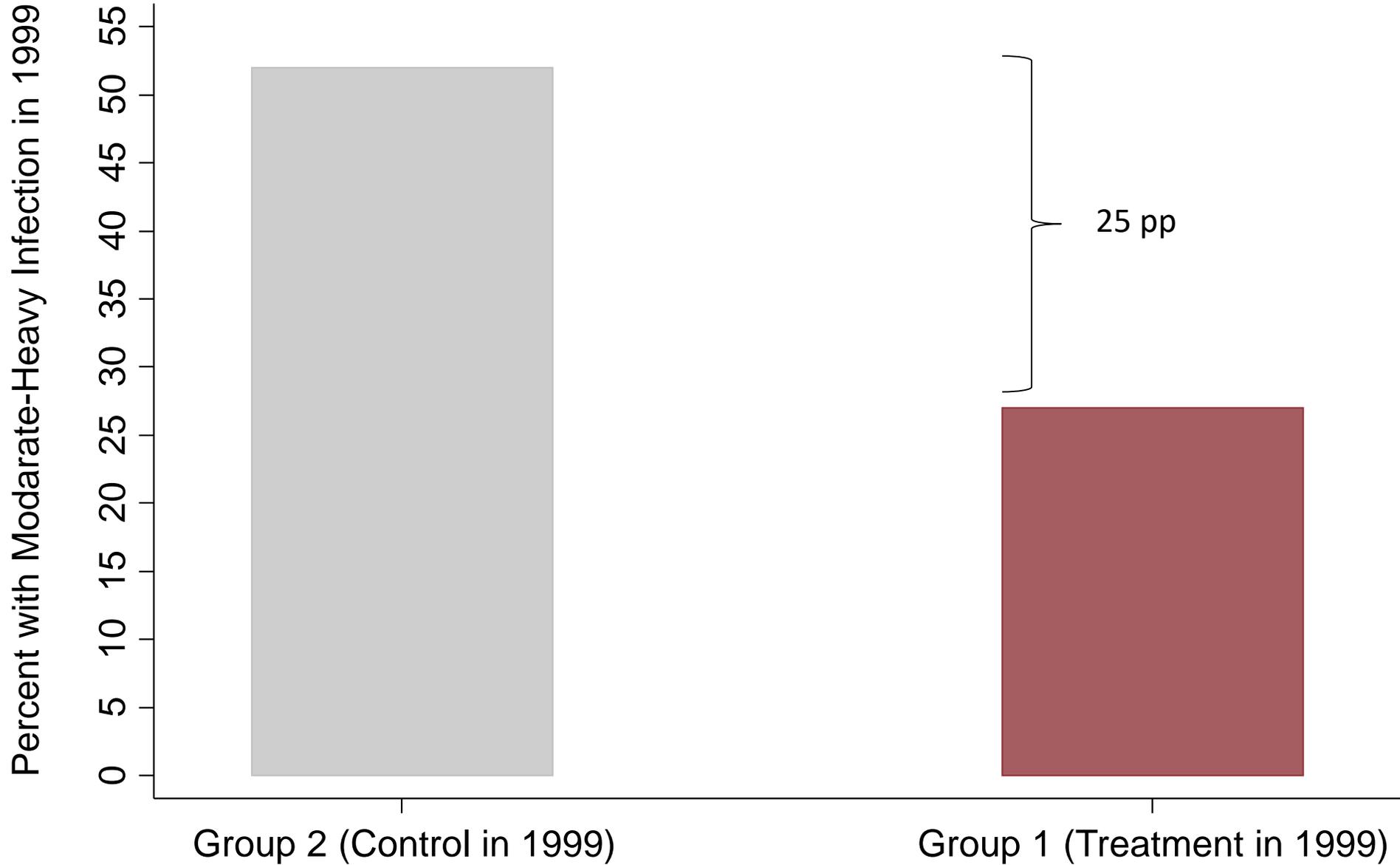
- One approach to evaluating impacts of deworming: randomly assign treatment to some children and compare outcomes to their peers
- But randomizing treatment at the individual level underestimates the benefits of treatment
 - Misses externality benefits to the control group from reduced disease transmission
 - And underestimates benefits for the treatment group because the control group also benefits (contamination of “control”)

Estimating The Effects of Deworming Children

- Miguel and Kremer (2004) solve this problem by evaluating an intervention that was randomized at the **school** level
 - This captures externalities across kids who attend the same school
- 75 schools in Busia were randomly divided into 3 groups:
 - Schools in group 1: deworming in 1998 and 1999
 - Schools in group 2: deworming in 1999
 - Schools in group 3: deworming in 2001

Impacts of Deworming Intervention on Infection Rates in 1999

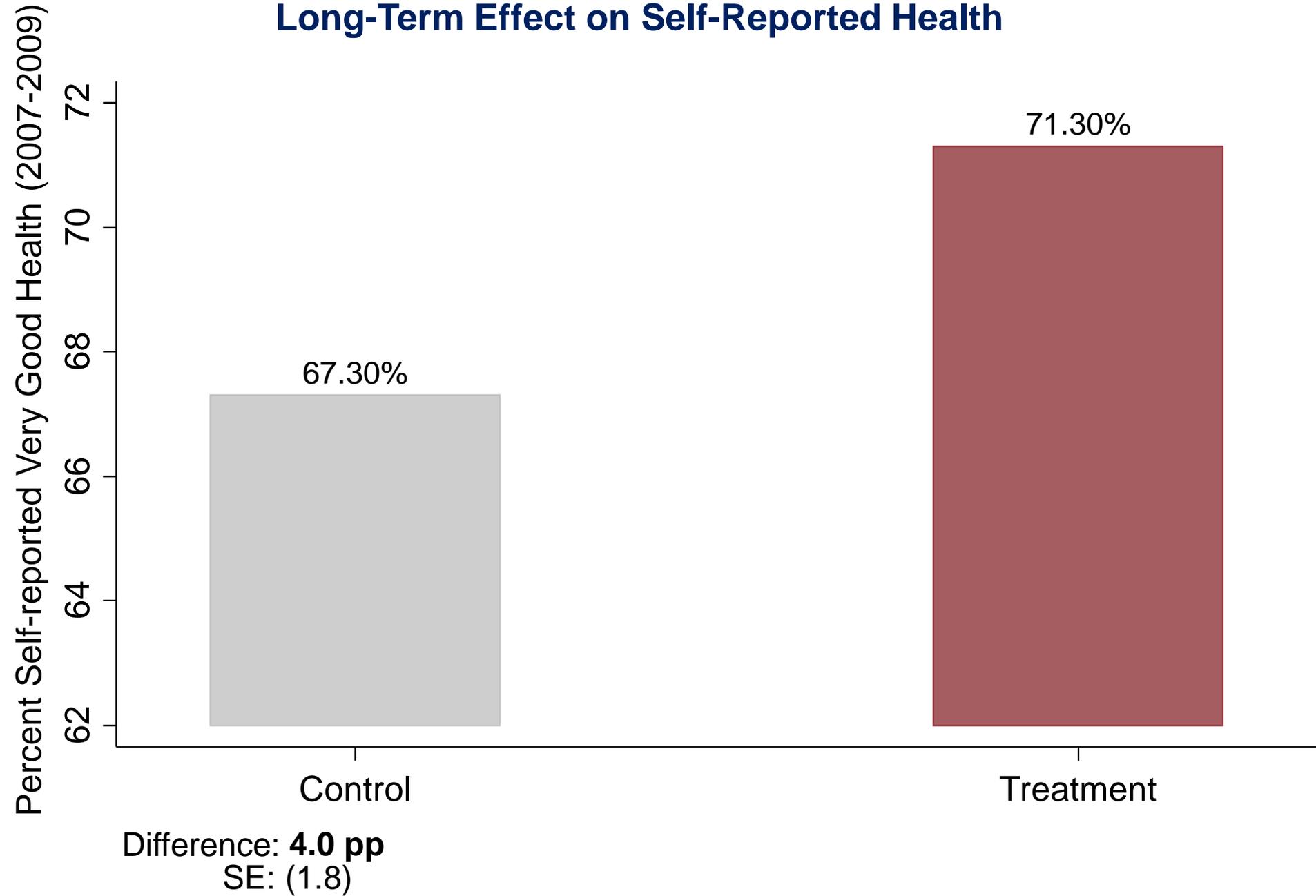
Group 1 vs. Group 2



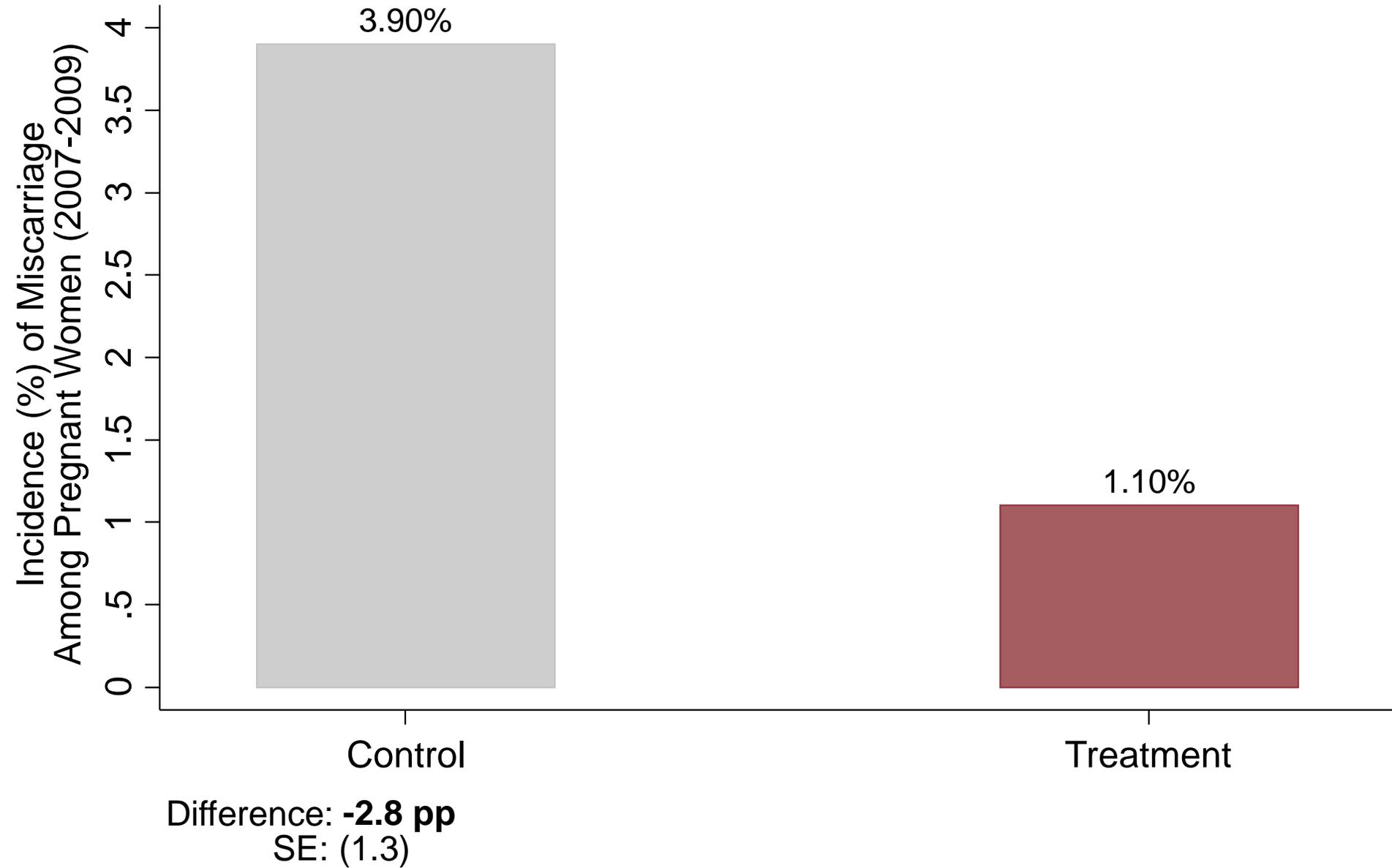
Long-Term Effects of Deworming Children

- Does deworming treatment have long-term benefits?
- Binder, Hicks, Kremer, Miguel (2016) measure impacts of program on kids' outcomes a decade after the intervention
 - Intervention cohort is between 19-26 years old at point of follow-up
 - 83% of children successfully located
- Since all groups were ultimately treated, use a **treatment intensity** design:
 - Compare groups with different duration of anti-worm medication
 - Groups 1 and 2 received about **2.5 years more** of exposure to the program than group 3

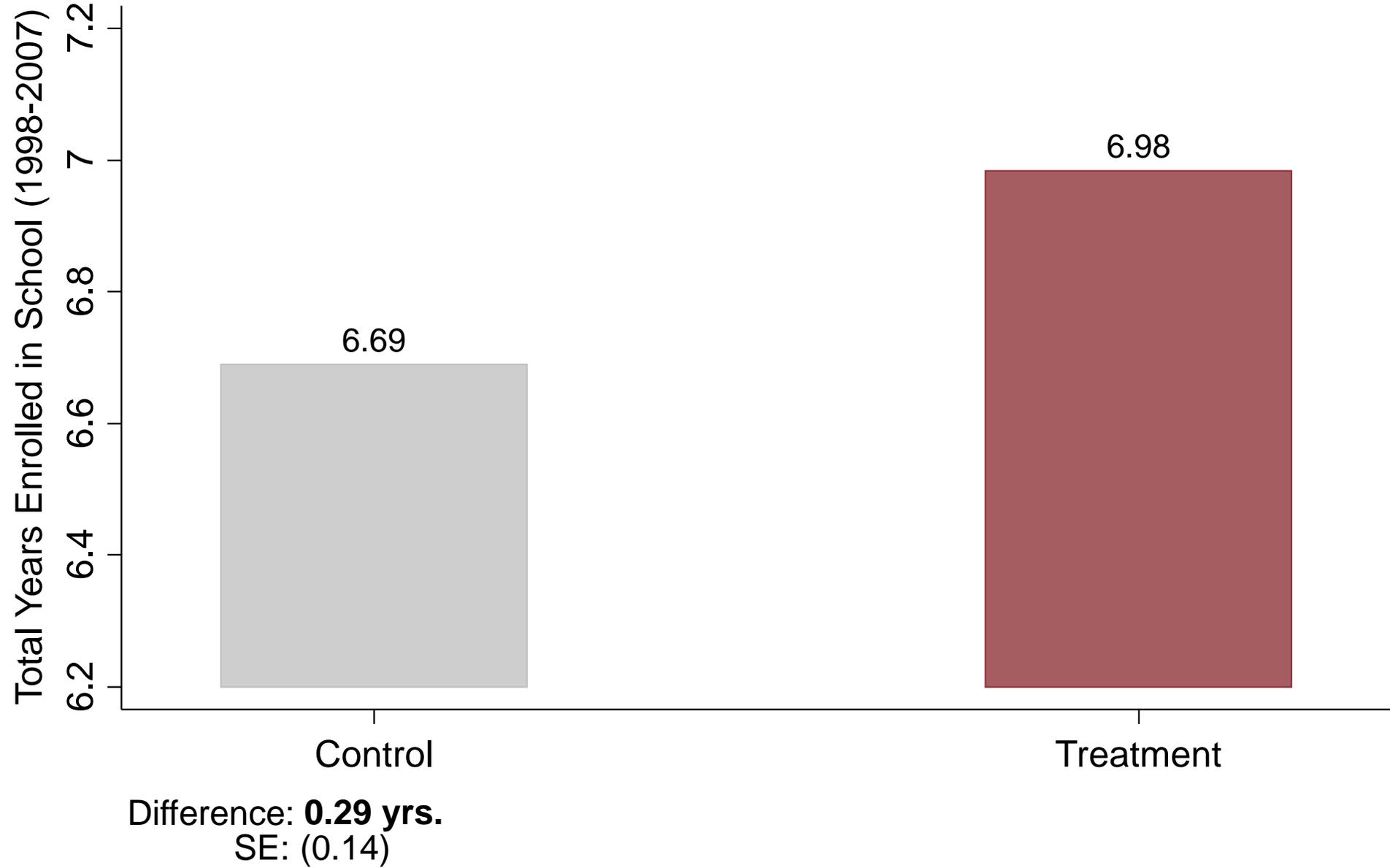
Long-Term Effect on Self-Reported Health



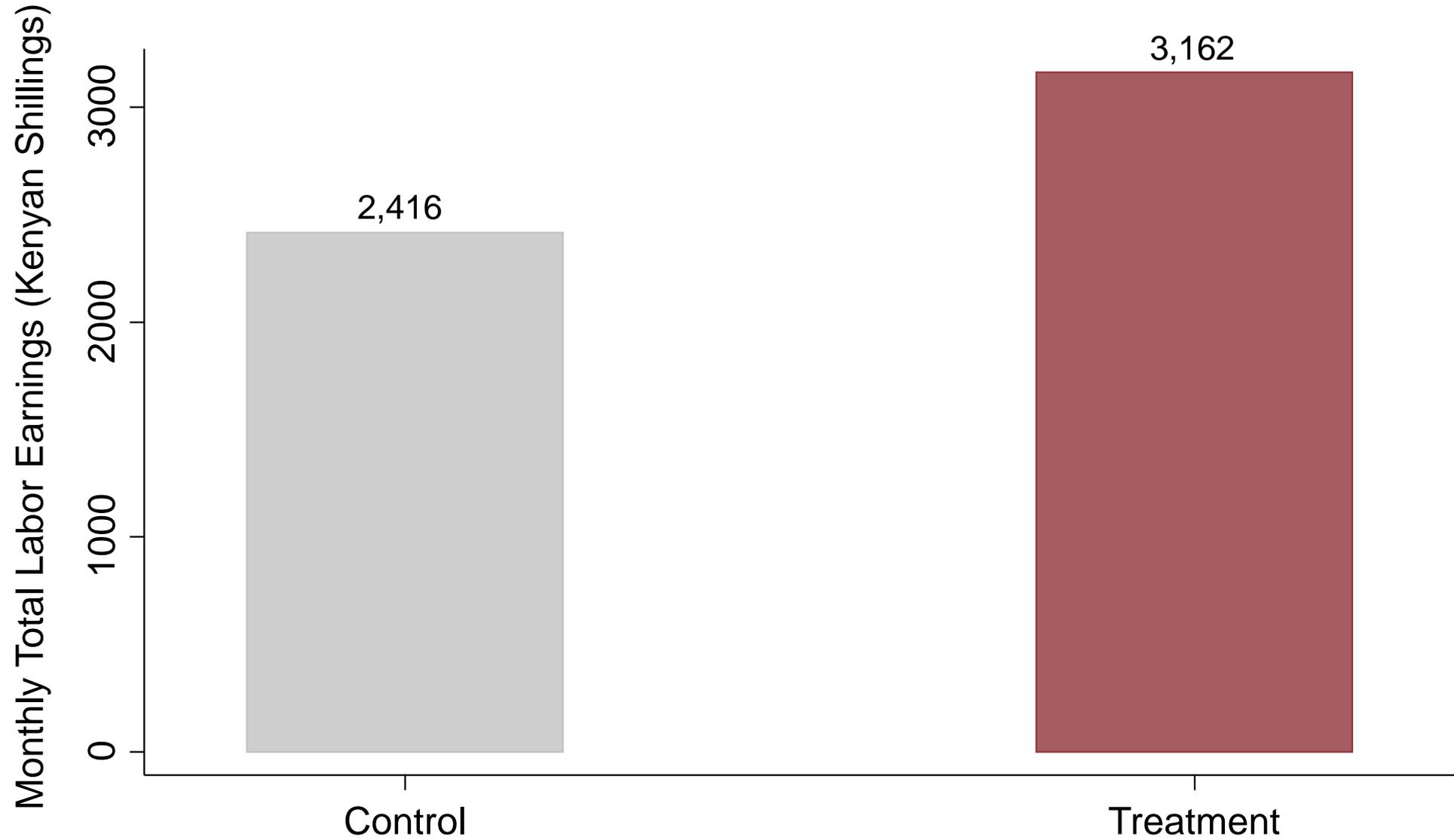
Effect on Miscarriage Among Pregnant Women



Effect on Total Years of School Enrollment (1998-2007)



Effect on Earnings

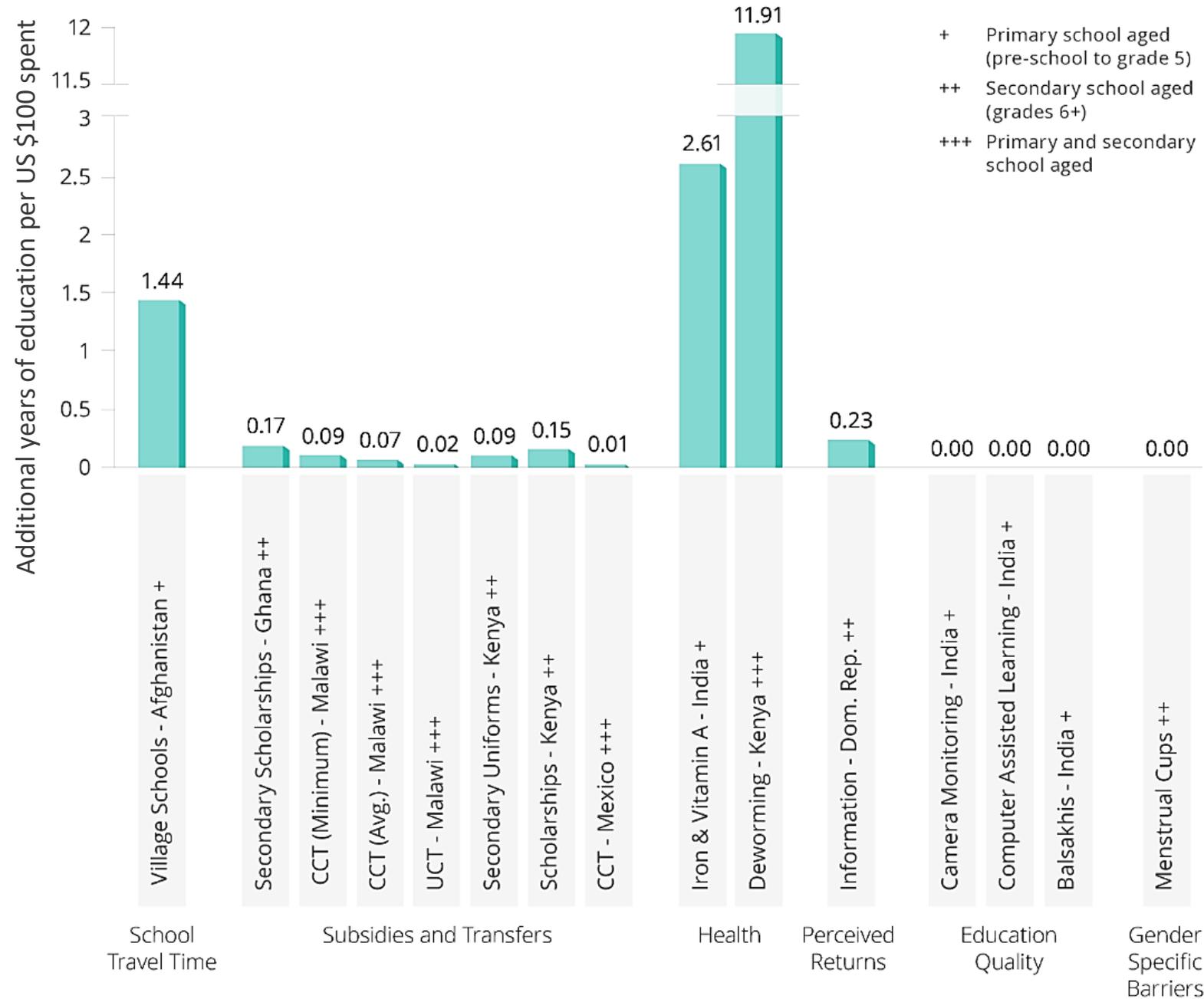


Difference: **745.8 shillings**
SE: (214.4)

Impacts of Deworming Interventions: Costs vs. Benefits

- Average per-person **cost** of de-worming treatment: \$1.07
- Average per-person benefits:
 - Years of education increase by 0.12 years
 - NPV increase in per-person earnings: \$142.43
- Is this a cost-effective program?
 - Program pays for itself in the long run: tax revenue to government from higher earnings more than offsets cost of program
 - And cost-effectiveness is very high compared to other interventions....

Cost-Effectiveness of Programs to Improve Student Participation



Deworming to increase school attendance

After research found that school-based distribution of deworming pills in areas with high infection rates boosted health and school attendance; the approach has been scaled to reach over 285 million children in 2017.



Kenya's National School-Based Deworming Programme rolls out in Kwale province, Kenya.



The Deworm the World
Initiative

Institutions and Economic Development

The Role of Institutions in Economic Development and Growth

- Potential critique of policy evaluation approach: difficult to test and improve every single policy that countries implement
- Alternative strategy: try to improve the efficacy of governments and institutions so that countries are on a better path by themselves
 - Equip countries to figure out how to improve sanitation, implement public health efforts, expand schools, etc. by themselves
- Potentially a more scalable approach, but also more difficult to implement

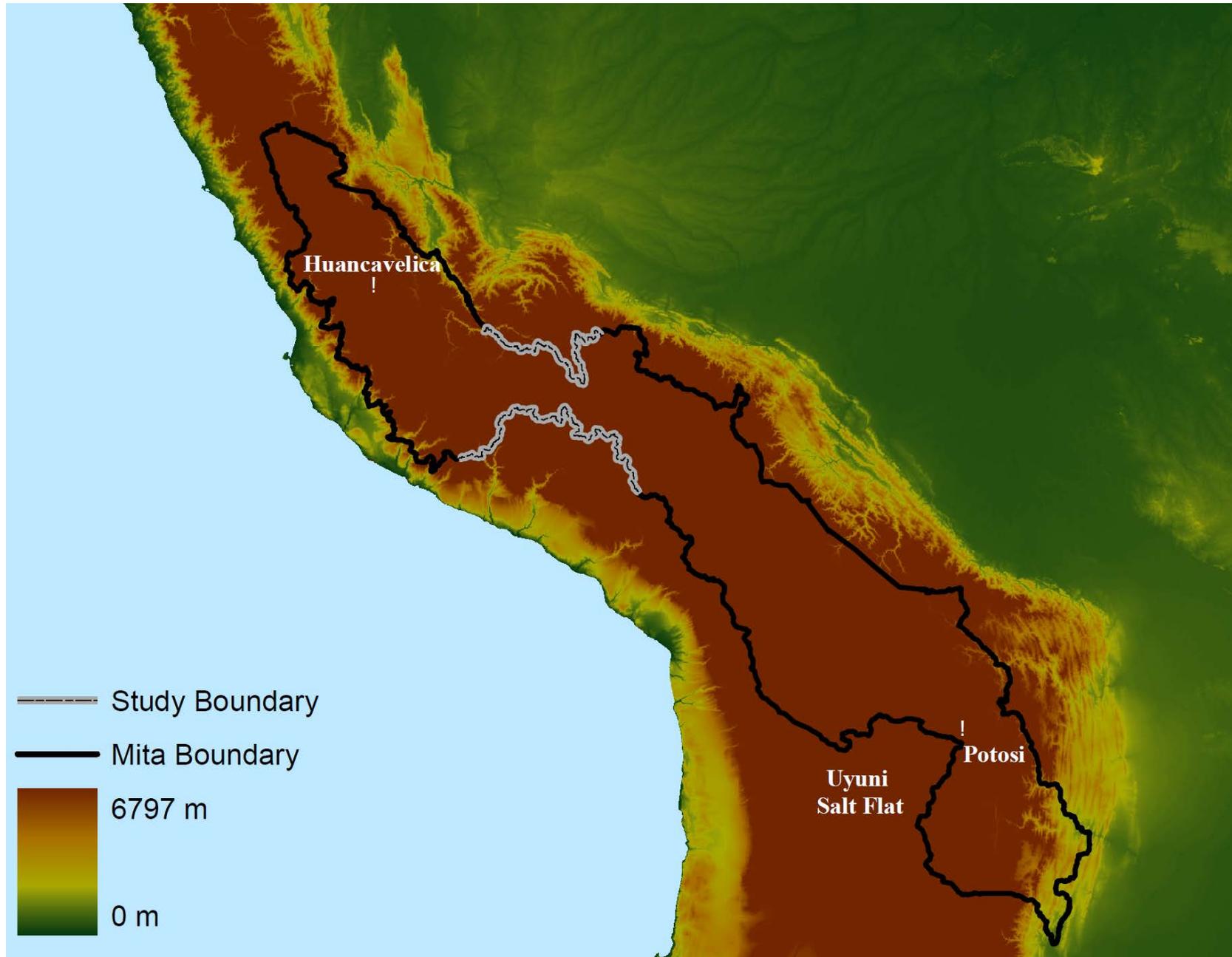
Institutions and Economic Development

- Theoretical work by Douglass North and others argued that strong institutions (property rights, political structures) are central for economic development
- Early empirical work testing this hypothesis relied on cross-country comparisons
 - Challenge: many differences across countries → difficult to be certain about mechanisms
- Most recent work focuses on **sub-national** differences in institutions and uses “big data” approaches to analyze long-term impacts of institutions
 - Method: historical data analysis, using information recovered from archives to apply modern empirical techniques

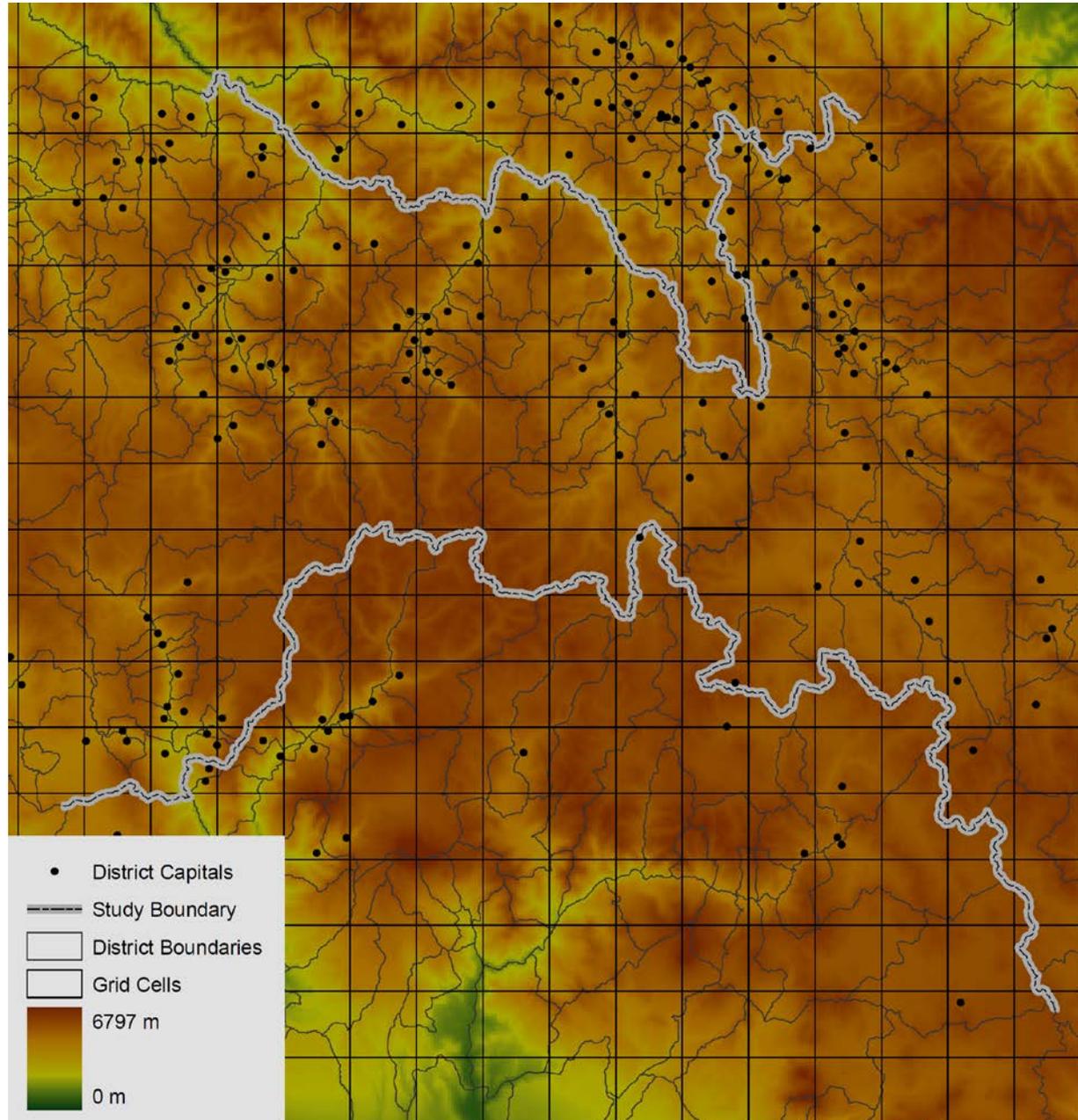
Evidence from Peru's Mining *Mita*

- Dell (2010) presents evidence on long-term impacts of institutions by studying the persistent impacts of Peru's mining *mita*
 - Mining *mita* was instituted by the Spanish government in 1573 and abolished in 1812
 - Required communities near Potosi and Huancavelica mines in Peru to send one-seventh of their adult male population to work in the mines

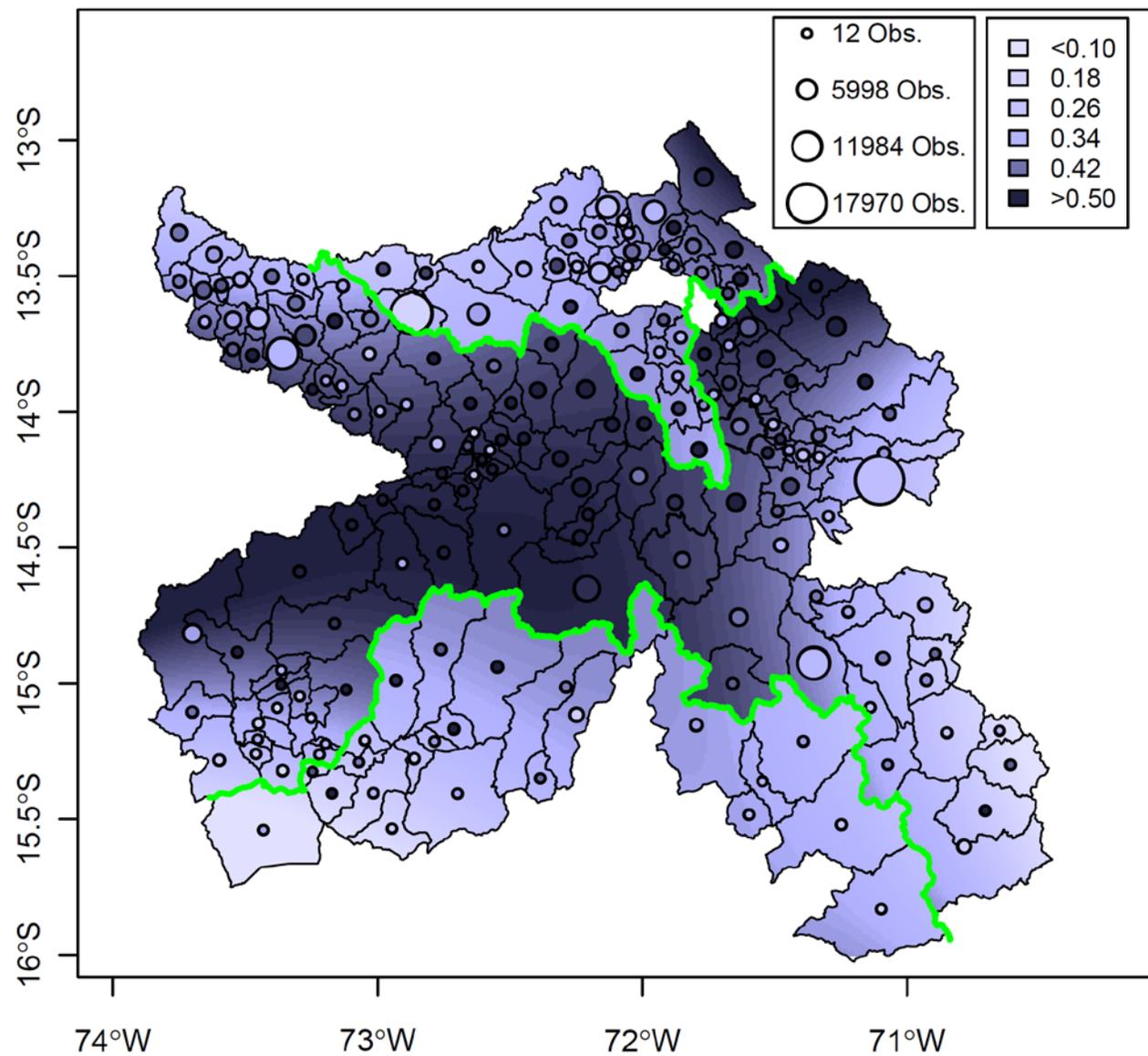
Mining *Mita* Boundaries



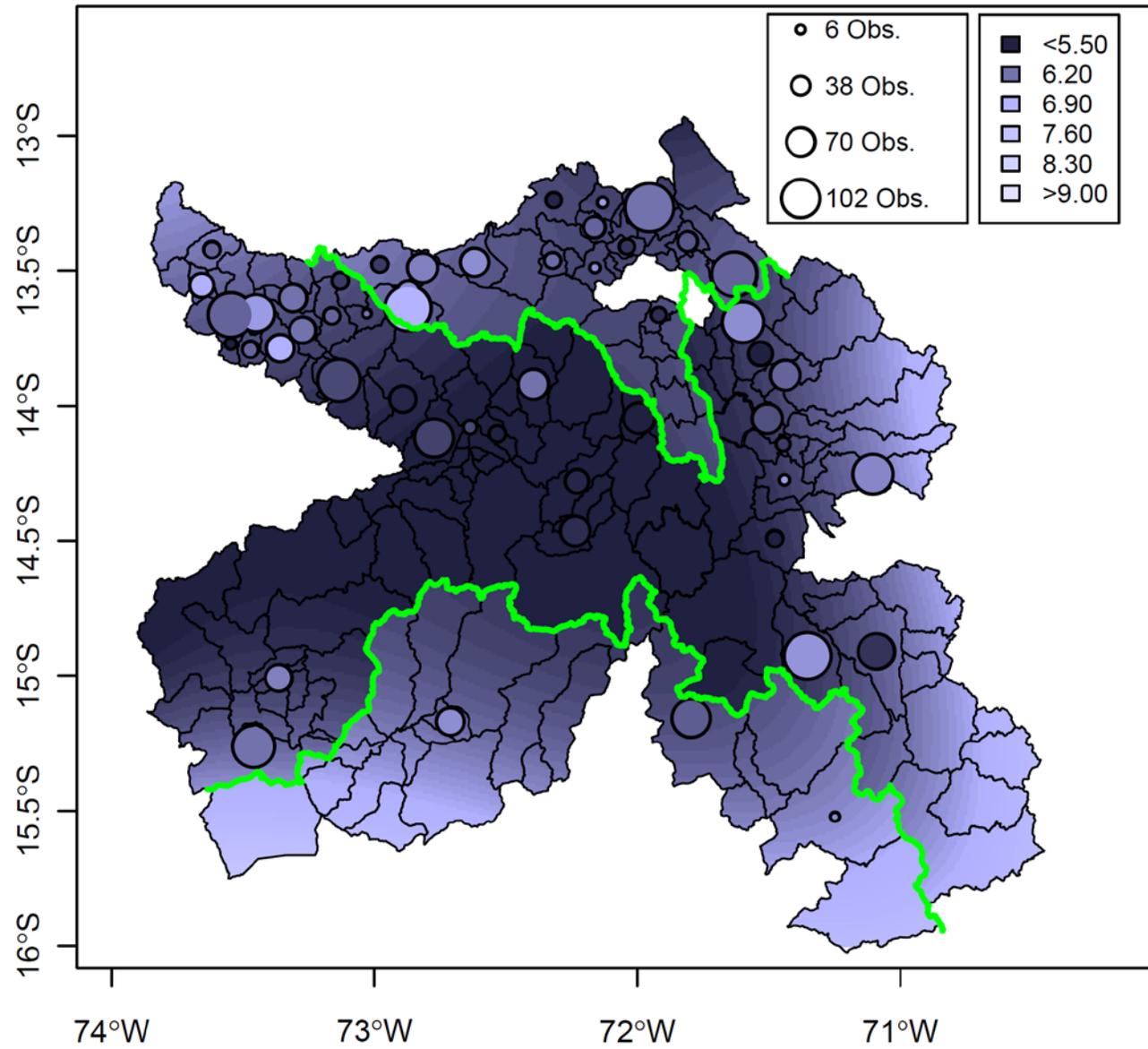
Boundaries Used in Study



Children's Heights in Present Day: Rates of Stunting



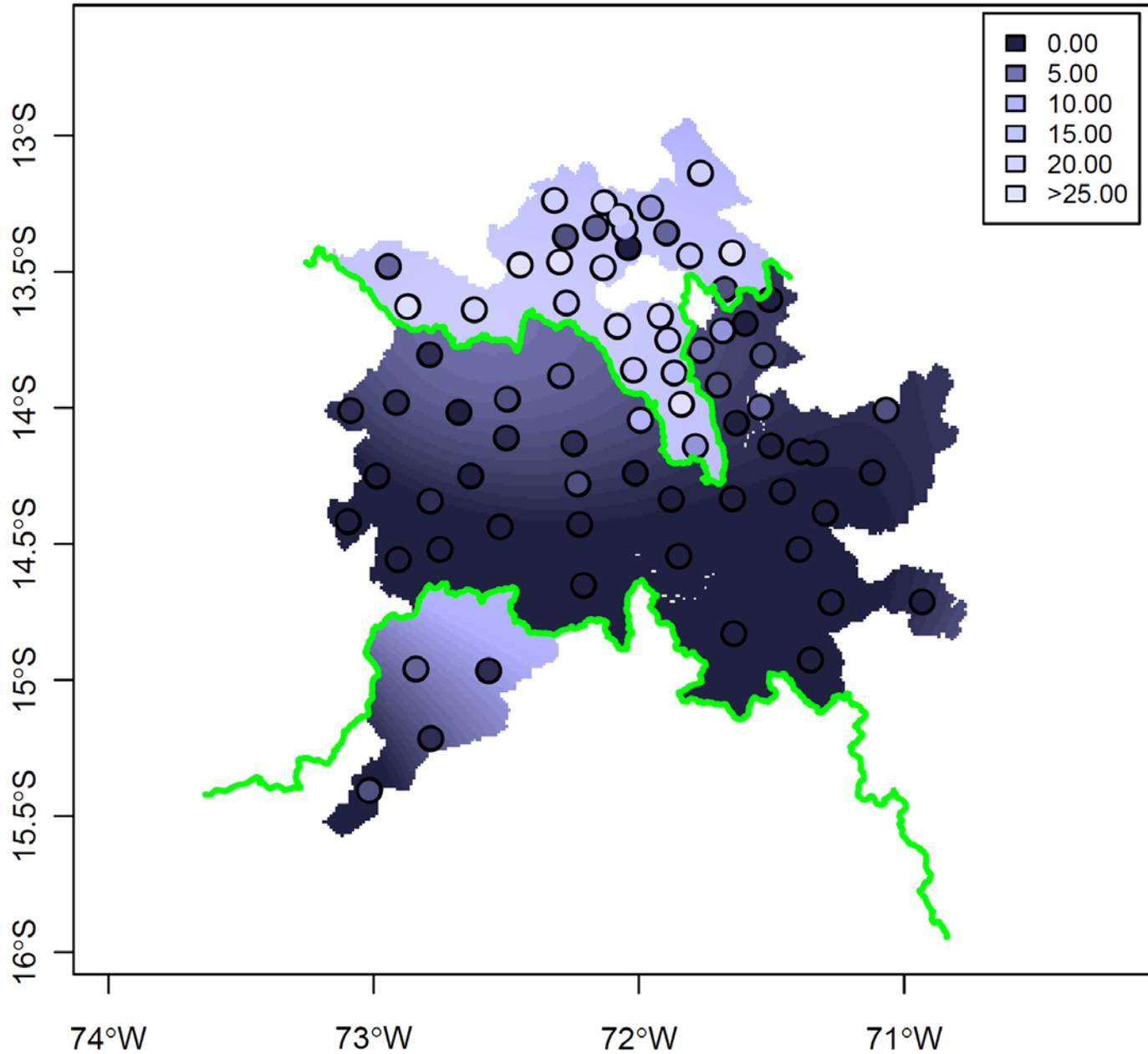
Household Consumption in 2001



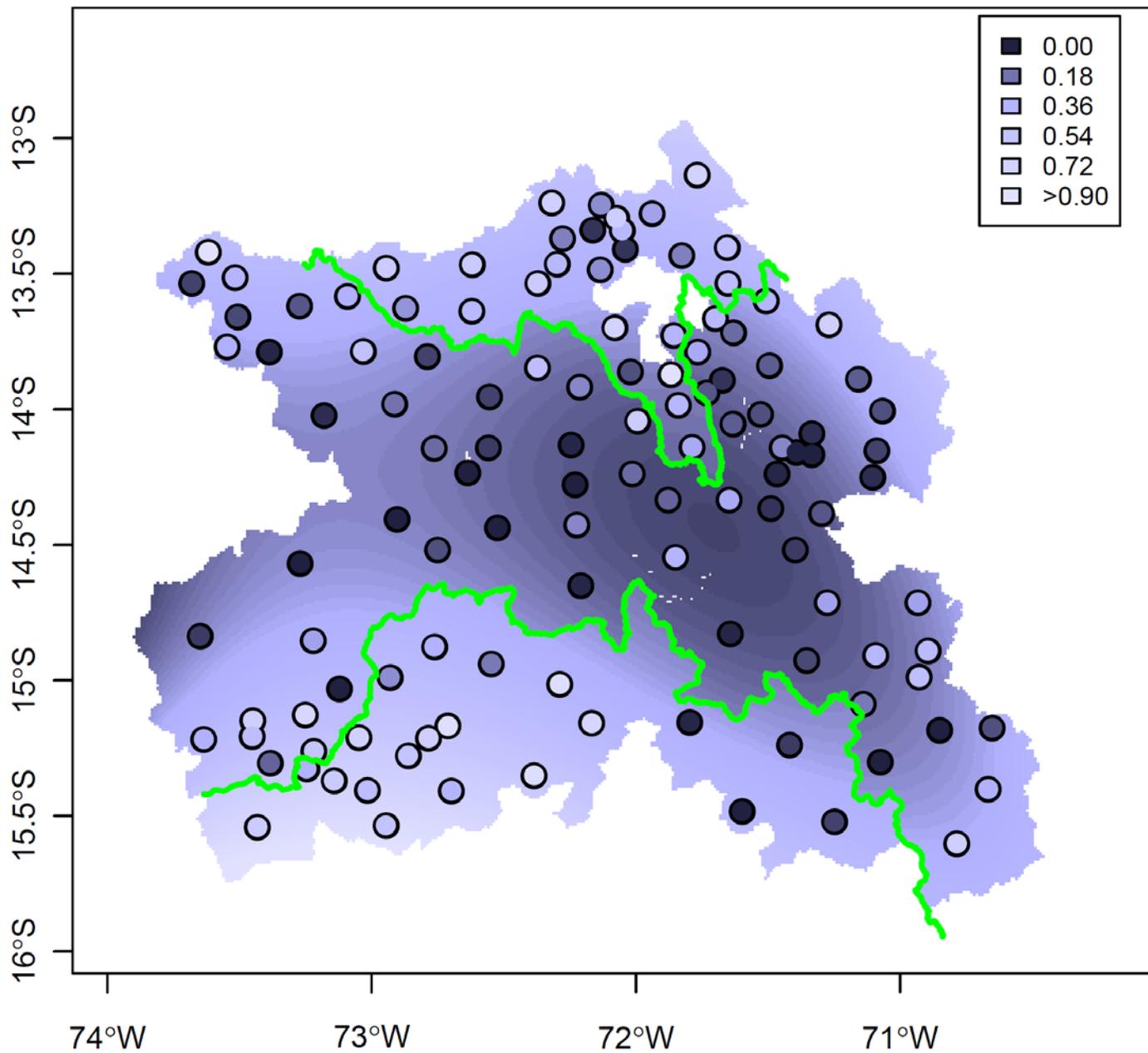
Channels: Impacts of *Mita* on Property Rights

- Why is there a persistent effect of *mita* on present-day outcomes?
- Non-*mita* areas had a well established system of land ownership and property rights
- But *mita* areas did not have clear property rights, and no system emerged even after *mita* system was ended
- Many more large rural estates in non-*mita* areas (*haciendas*), which provided employment opportunities and advocated for public goods

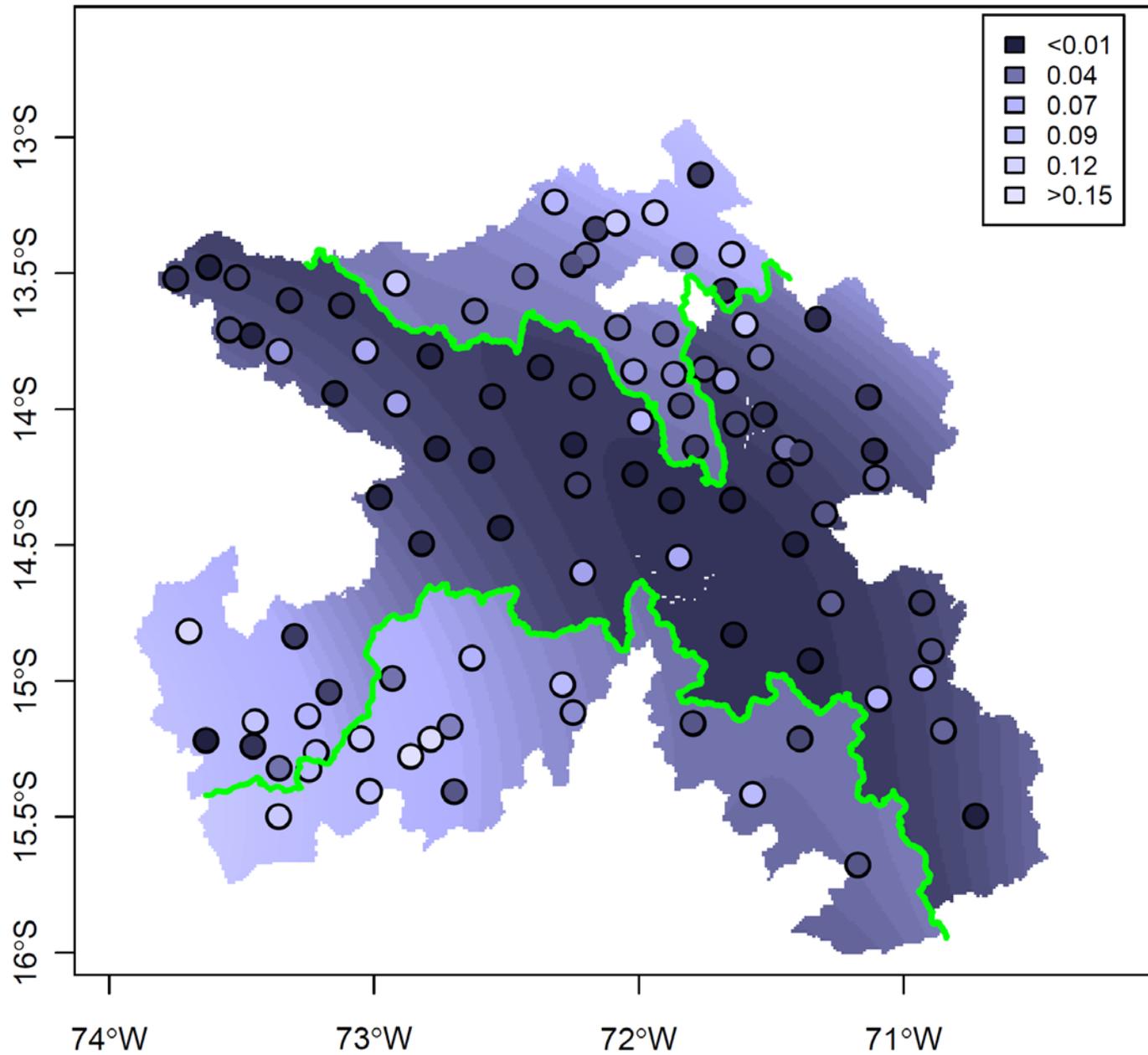
Number of Haciendas in 1689



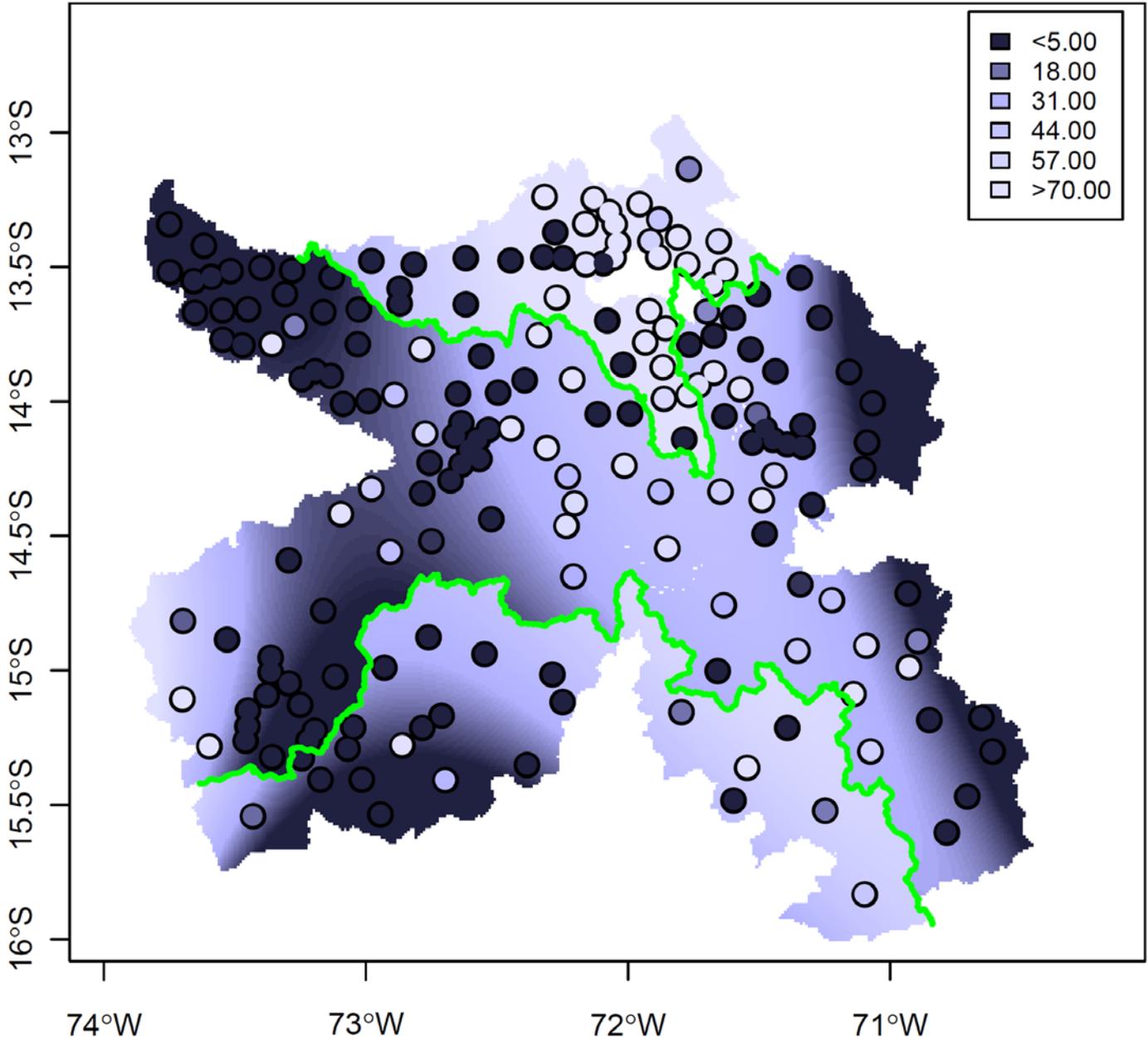
Number of Haciendas in 1940



Literacy Rates in 1876



Density of Regional Roads in 2006



Source: Dell (2010)

Summary: Tracing the Persistent Effects of Institutions

- The *mita* system had substantial adverse impacts on economic outcomes by changing institutions:
 1. Negative impact on property rights/land ownership
 - Large reduction in number of haciendas → few large landowners in *mita* districts
 2. Impact on public goods
 - Large property owners advocate for roads and public goods investment in their area
 3. Impact on markets and productivity
 - Residents of *mita* districts are substantially more likely to be subsistence farmers today because of lack of access to markets and human capital

Nation Building Through Foreign Intervention

- Given importance of institutions, one natural question is whether external (foreign) intervention could potentially help
- U.S. government has attempted to create more stable institutions in countries with weak institutions, particularly in the presence of conflict
 - Ex: military efforts to overthrow regimes or campaigns to sway public opinion
- These interventions are costly, but evidence on their impacts is limited
 - Experimentation is obviously unethical in the context of military interventions
 - But big data and careful historical analysis can provide a path forward even on such a challenging question

Causal Effects of U.S. Military Strategies: Evidence from the Vietnam War

- Dell and Querubin (2018) study how U.S. military strategies employed during the Vietnam War impacted state-building and institutions
- U.S. goal was to strength South Vietnamese government and weaken communist insurgency (Viet Cong)
- Used two different types of approaches:
 - Overwhelming Firepower: trying to establish control through military force (air strikes) by making it costly to oppose the state
 - Hearts and Minds: emphasizes providing public goods, creating economic opportunities, and promoting civic engagement

Estimating the Causal Effects of Bombing Vietnam

- U.S. Air Force determined which areas to bomb by evaluating where risk of Viet Cong insurgency was greatest
- Used 169 monthly and quarterly questions about hamlet security, politics, and economics (asked of local leaders) to assess security risk
- Then aggregated these questions using a statistical prediction model to create risk scores from 1-5
- Continuous scores were then rounded to the nearest integer value to determine which areas would be bombed
 - As a result, probability of being bombed jumped discontinuously for two similar areas that had scores of 4.49 vs. 4.5

Estimating the Causal Effects of Bombing Vietnam

- Dell and Querubin located algorithm used to create continuous scores in unclassified documents at Fort McNair and tapes at National Archives

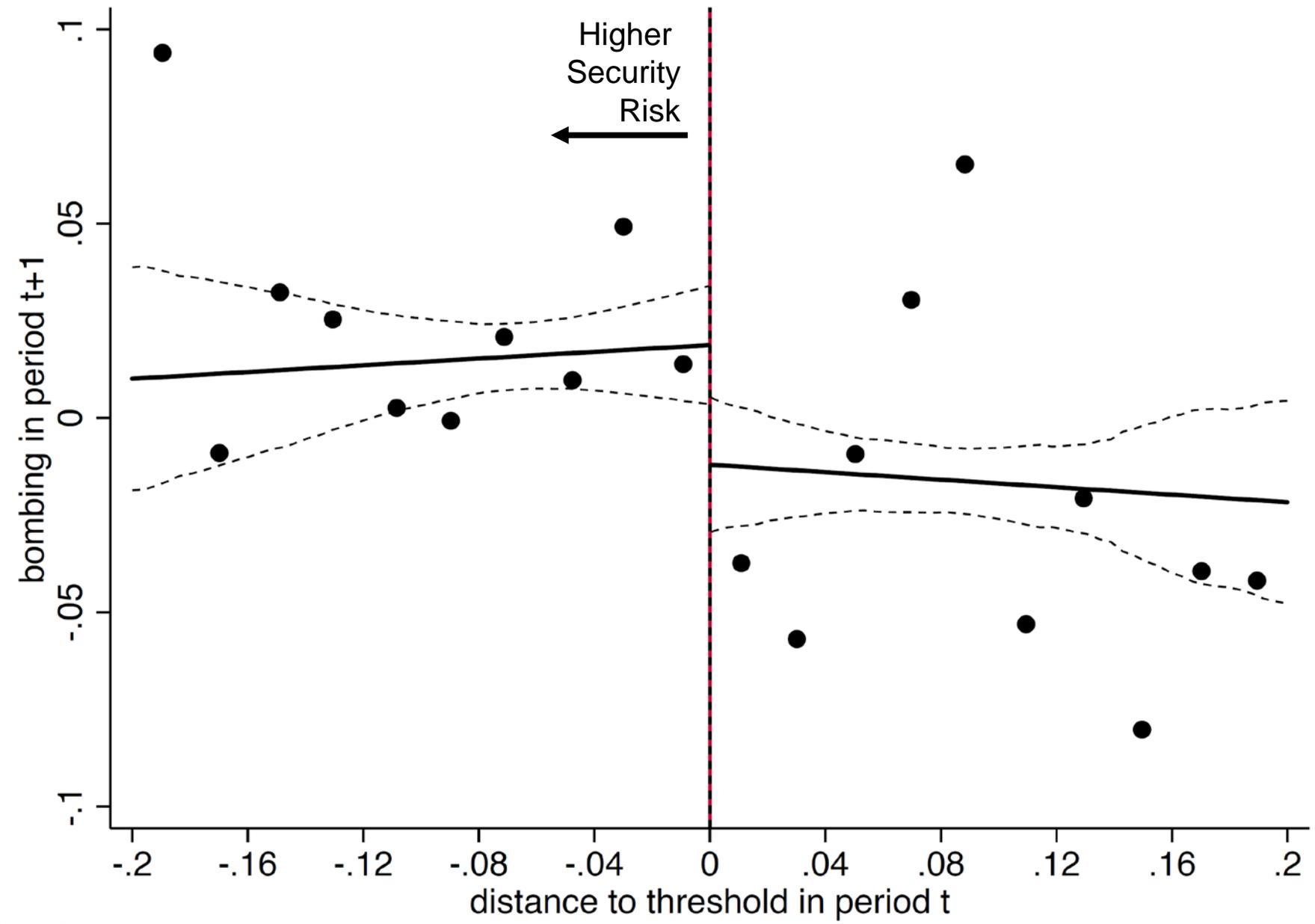
Constructing the Risk Scores

ENEMY POLITICAL ACTIVITY					
HMB05-0	.882	.788	.528	.571	.770
HMB05-1	.068	.140	.314	.286	.152
HMB05-2	.034	.070	.157	.143	.076
HMB06-0	.891	.797	.510	.557	.764
HMB06-1	.070	.134	.324	.294	.156
HMB06-2	.035	.067	.162	.147	.078
HMB07-0	.890	.785	.564	.500	.781
HMB07-1	.055	.105	.215	.250	.110
HMB07-2	.044	.084	.172	.200	.088
HMB07-3	.011	.021	.043	.050	.022
HMB08-0	.924	.858	.617	.278	.185
HMB08-1	.048	.100	.260	.342	.350
HMB08-2	.024	.038	.120	.378	.462
HQB01-0	.502	.273	.079	.011	.001
HQB01-1	.380	.469	.284	.074	.028
HQB01-2	.104	.213	.476	.315	.096
HQB01-3	.011	.026	.142	.440	.265
HQB01-4	.001	.017	.018	.160	.609
VQB01-0	.406	.244	.066	.010	.001
VQB01-1	.450	.533	.277	.097	.018
VQB01-2	.122	.183	.518	.272	.085
VQB01-3	.018	.032	.120	.416	.244
VQB01-4	.001	.003	.015	.203	.650

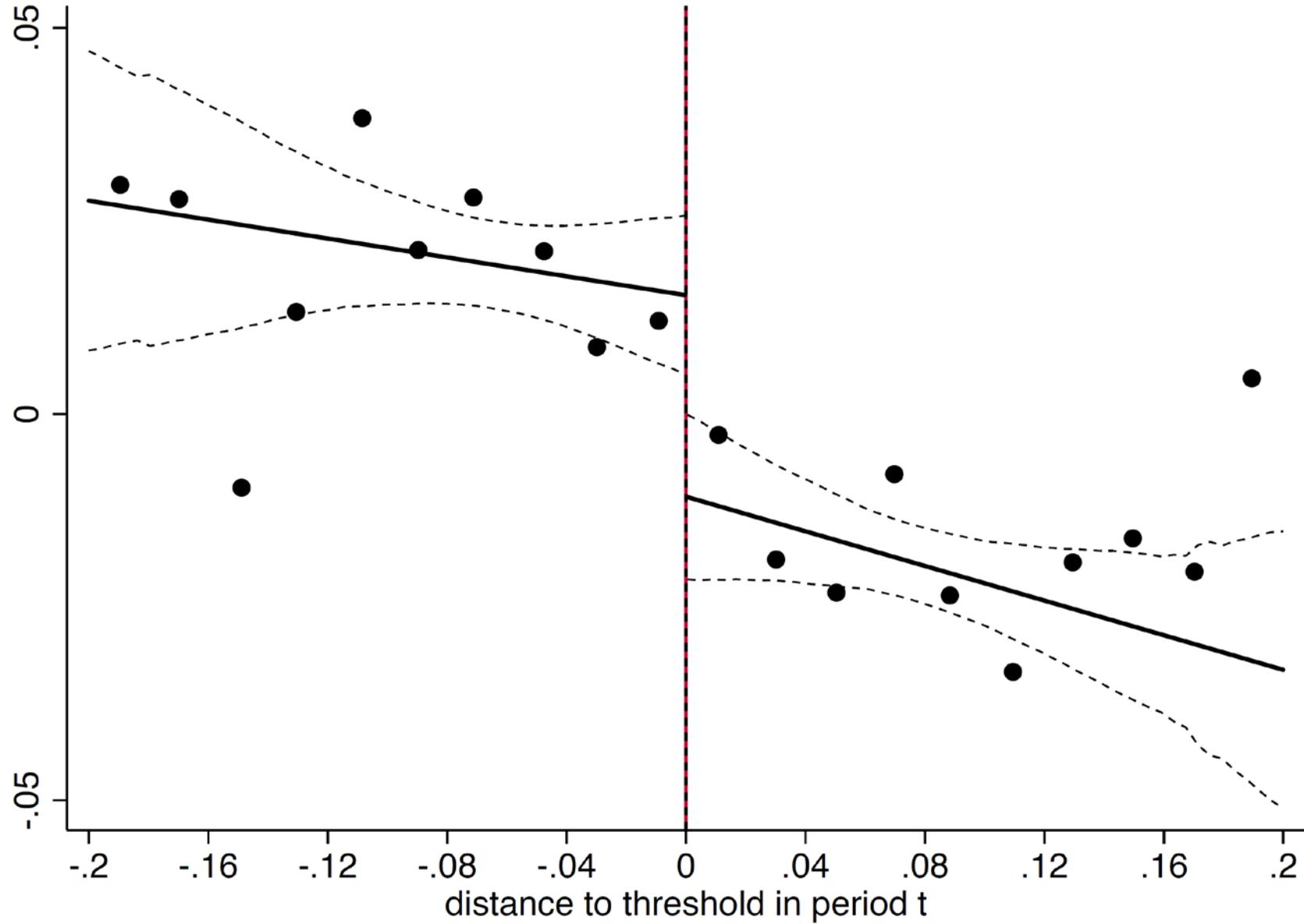
Estimating the Causal Effects of Bombing Vietnam

- Dell and Querubin located algorithm used to create continuous scores in unclassified documents at Fort McNair and tapes at National Archives
- Used these continuous scores to implement a regression discontinuity design

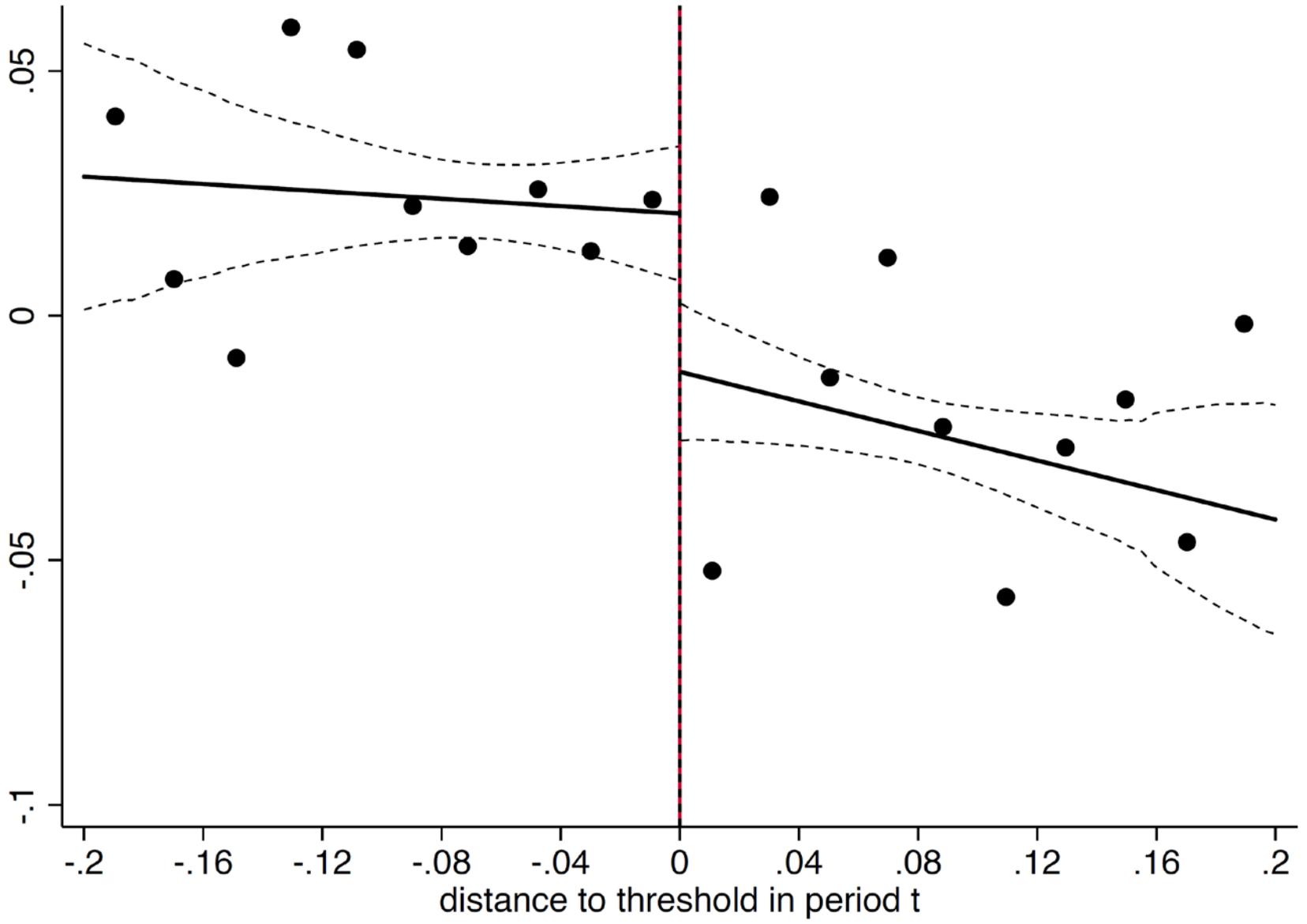
Effect of Reaching Threshold on Probability of Bombing



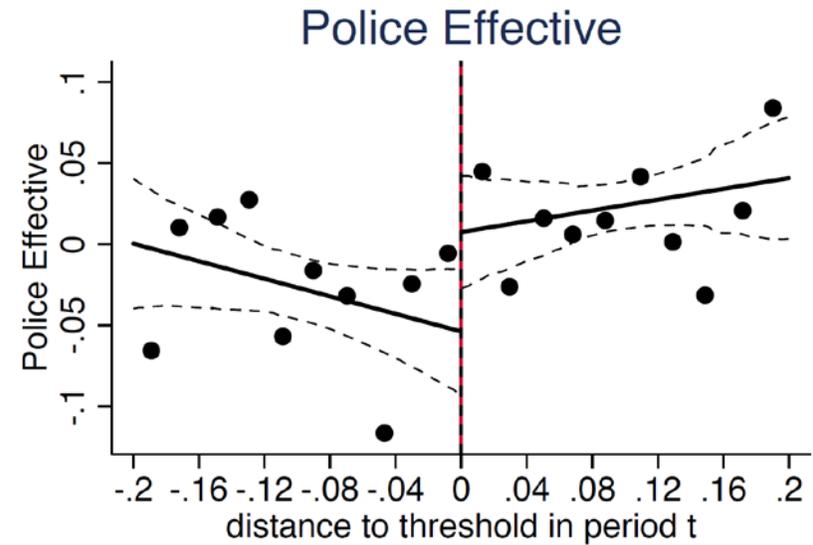
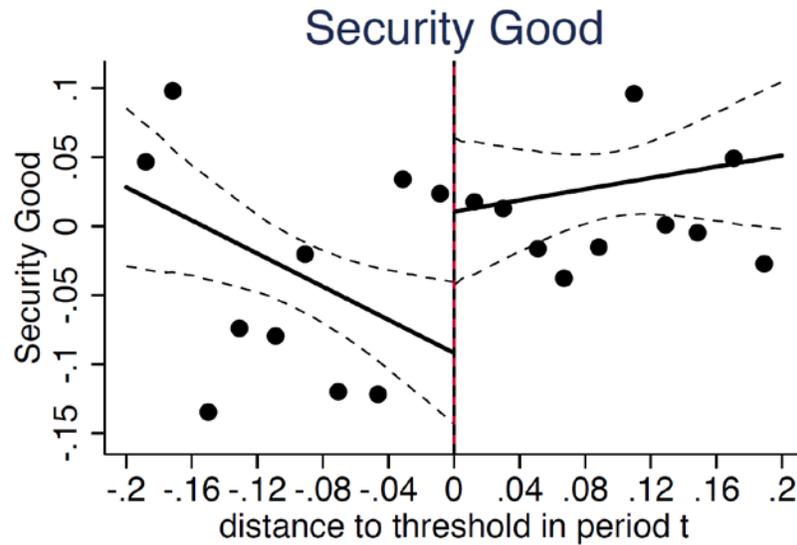
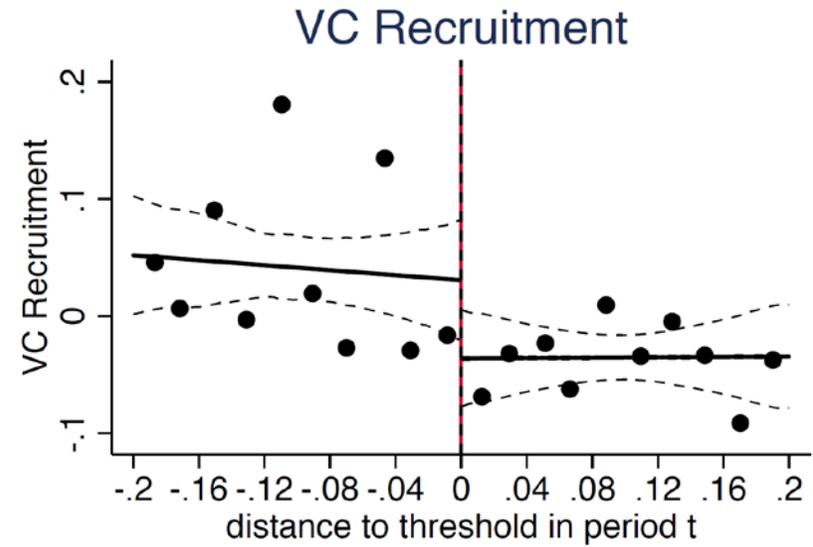
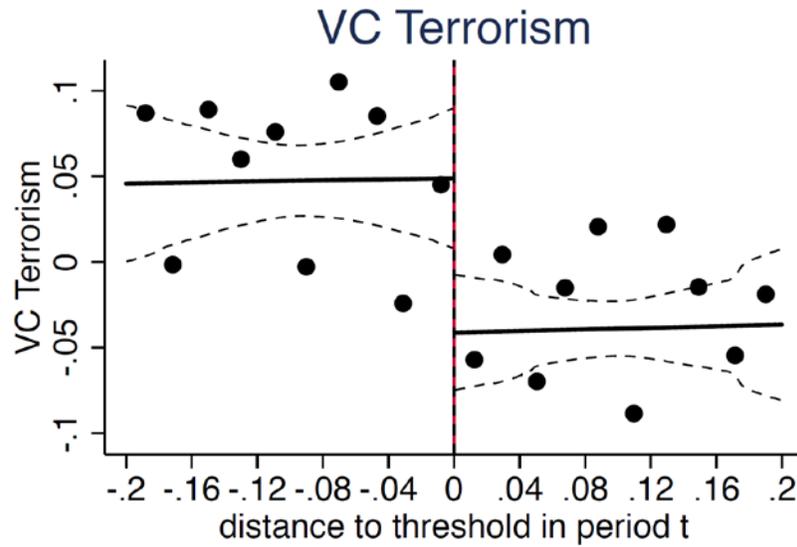
Armed Viet Cong Presence



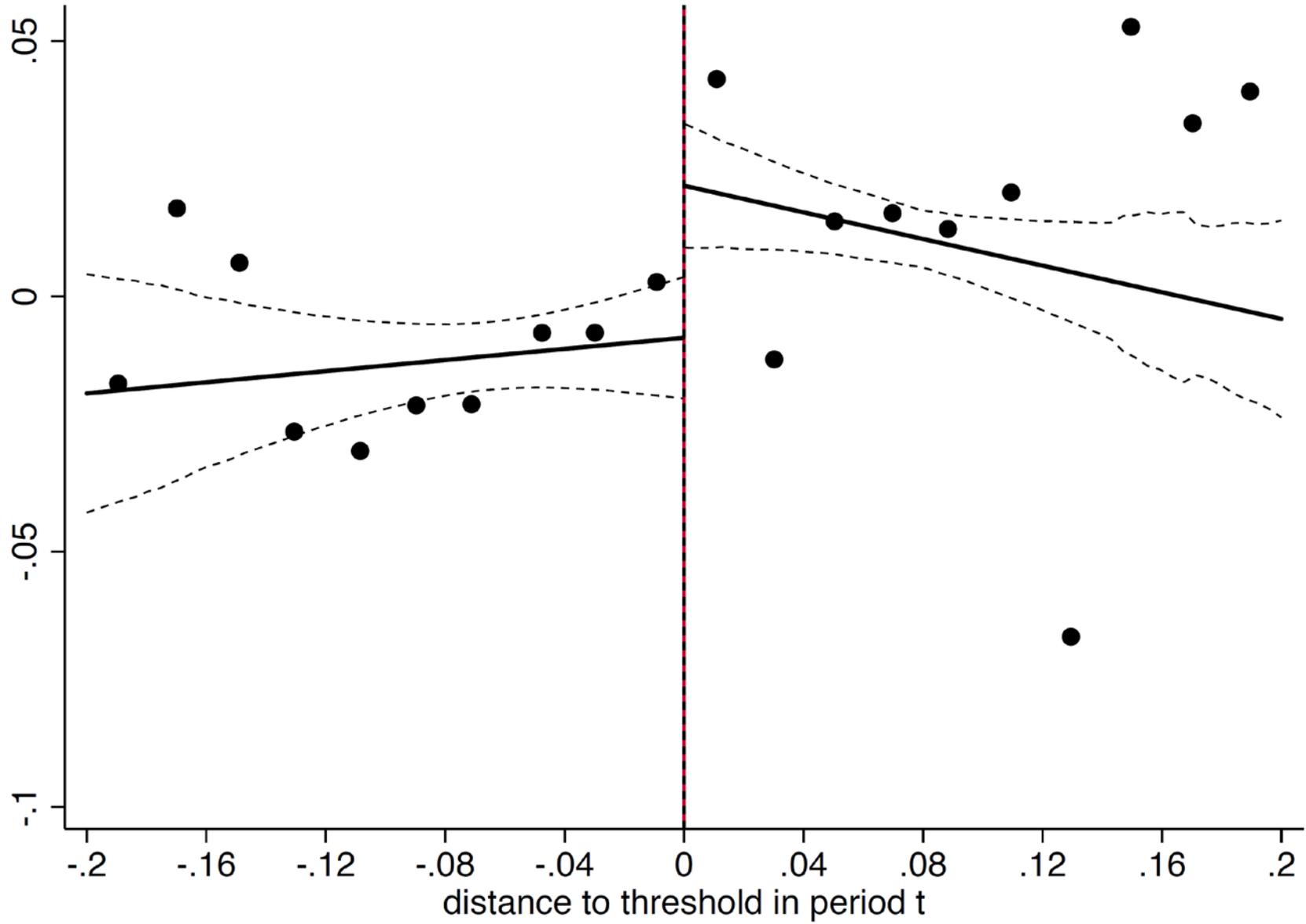
Viet Cong Recruitment Infrastructure



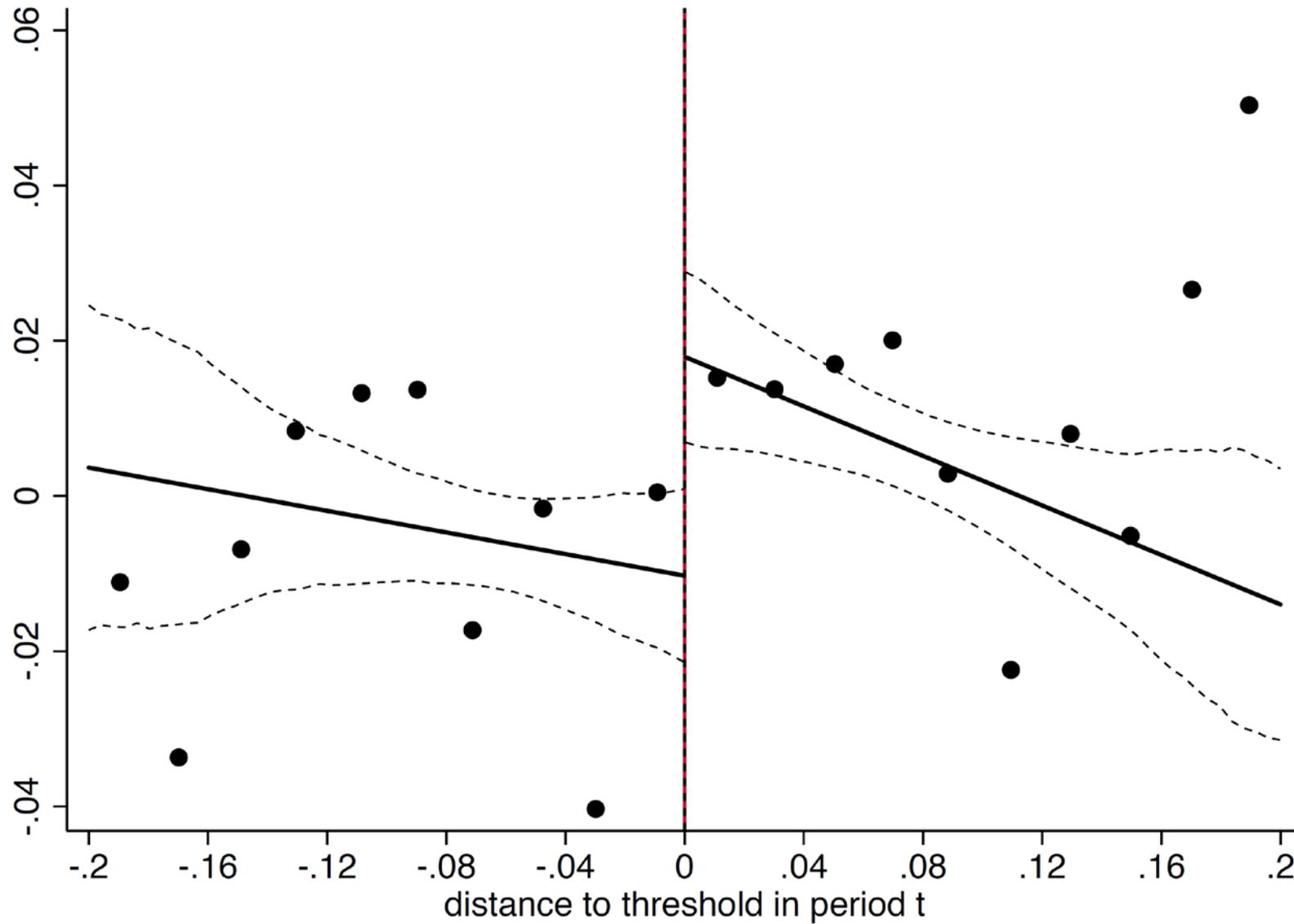
Public Opinion About Security



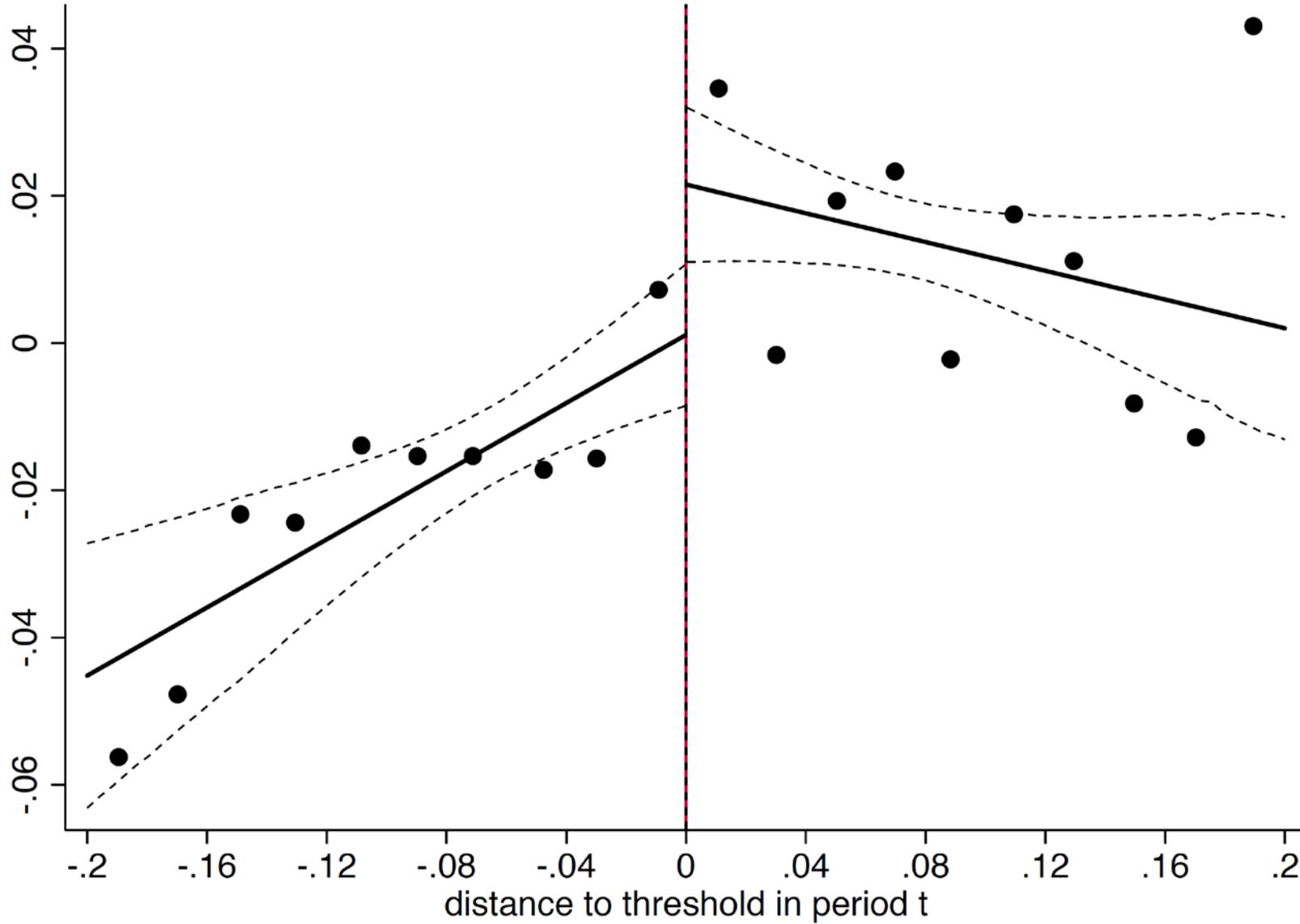
Village Committee Positions Filled



Access to a Primary School



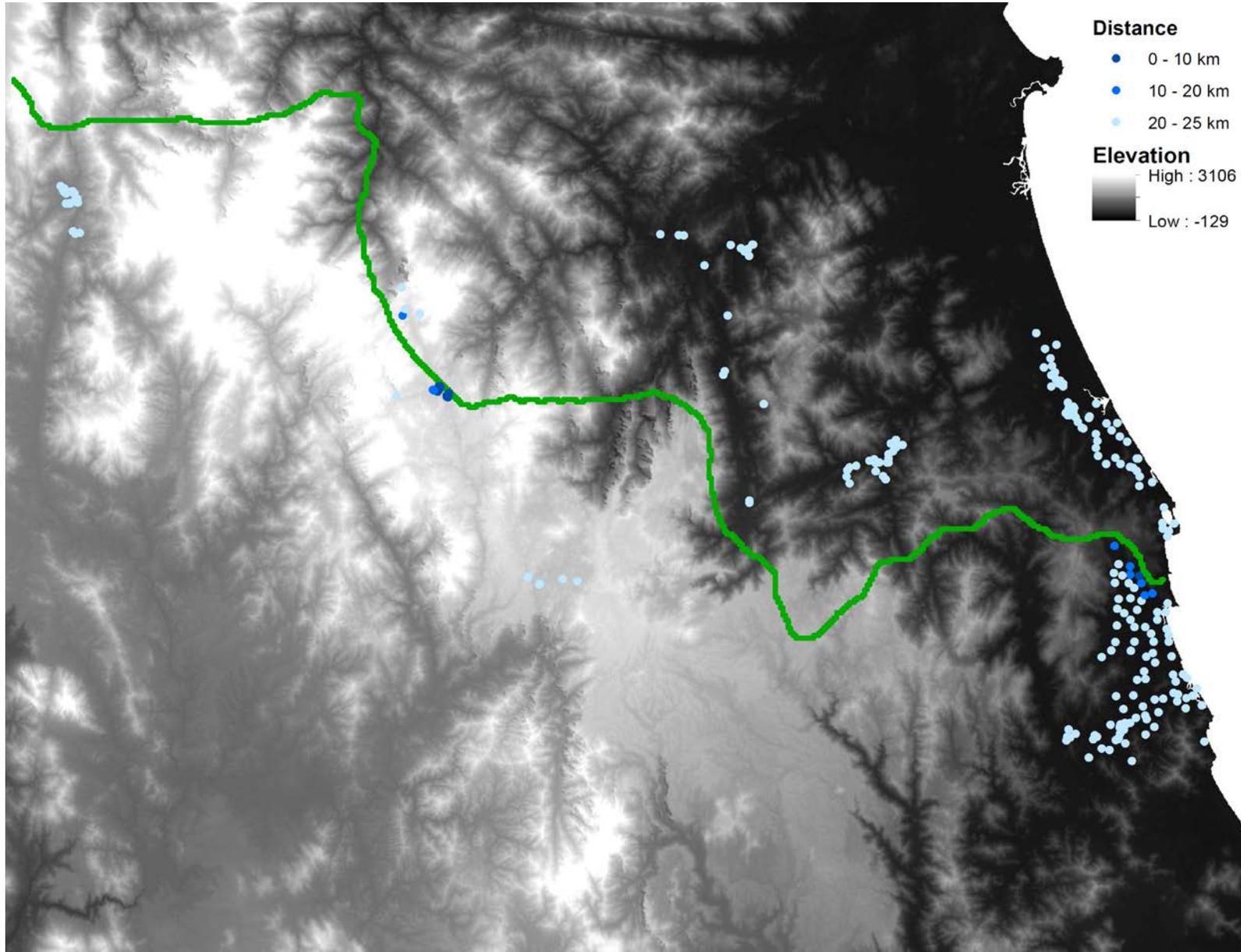
Participation in Civic Organizations



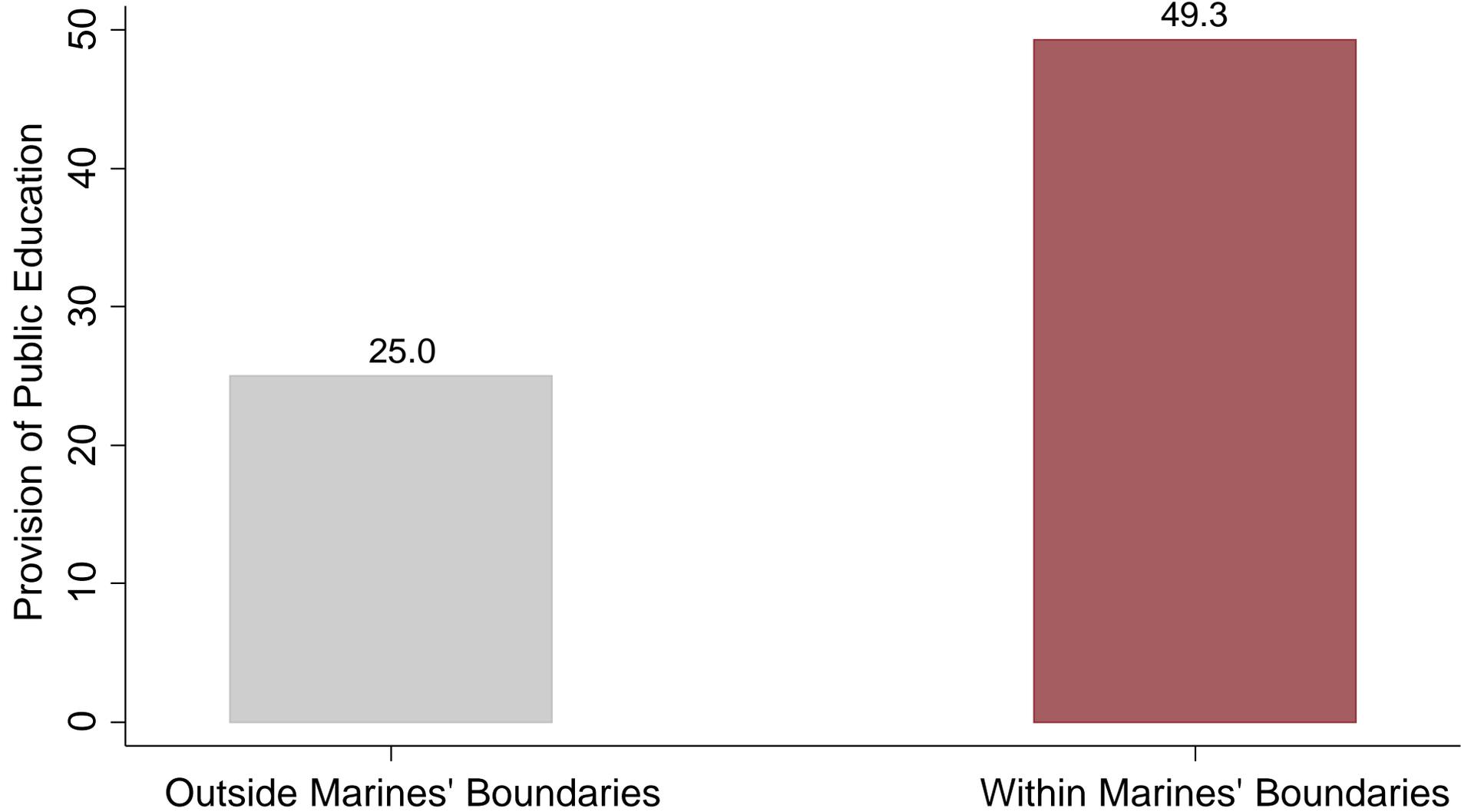
Air Strikes vs. Hearts and Minds Approach

- Bombing increased the military and political activity of the Viet Cong and **weakened** governance and non-communist civic engagement
- Now contrast this with alternative “hearts and minds” approach
- U.S. Marines pursued an approach of civilian engagement rather than military force
- Use a spatial regression discontinuity based on sharp border that determined where Army vs. Marines engaged

US Marines vs. Army Engagement Boundary

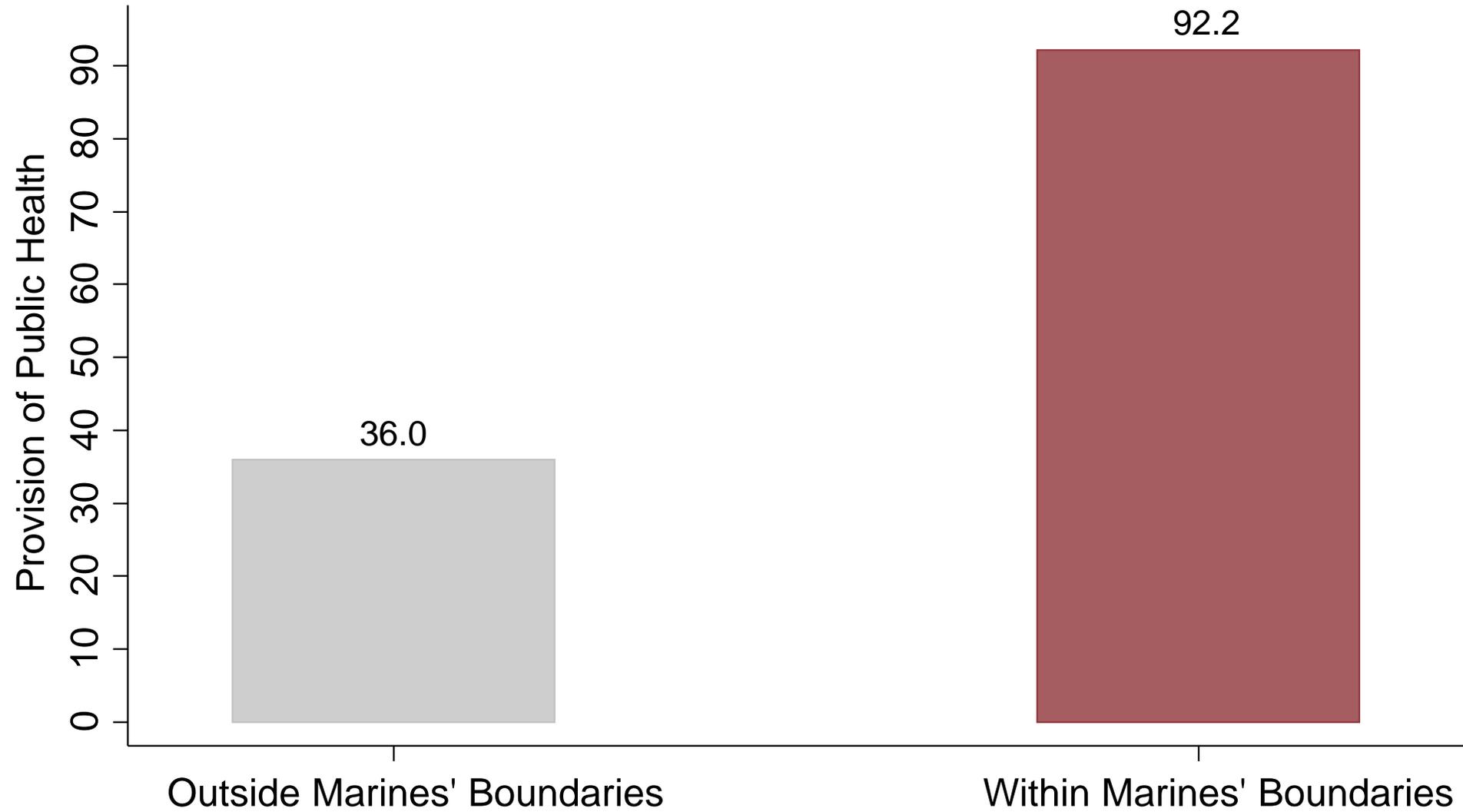


Effect on Provision of Public Education



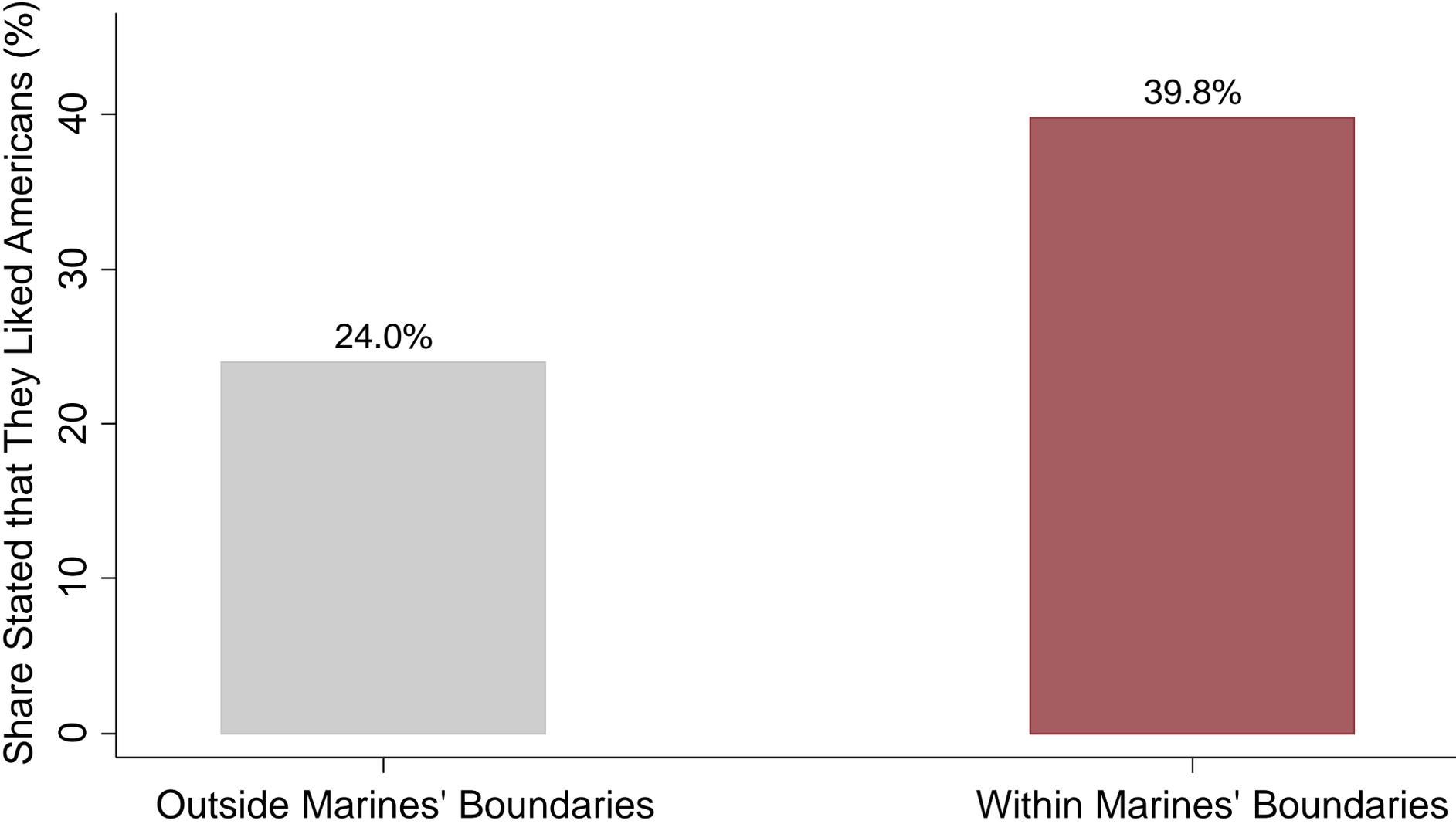
Difference: **24.3**
SE: (9.6)

Effect on Provision of Public Health



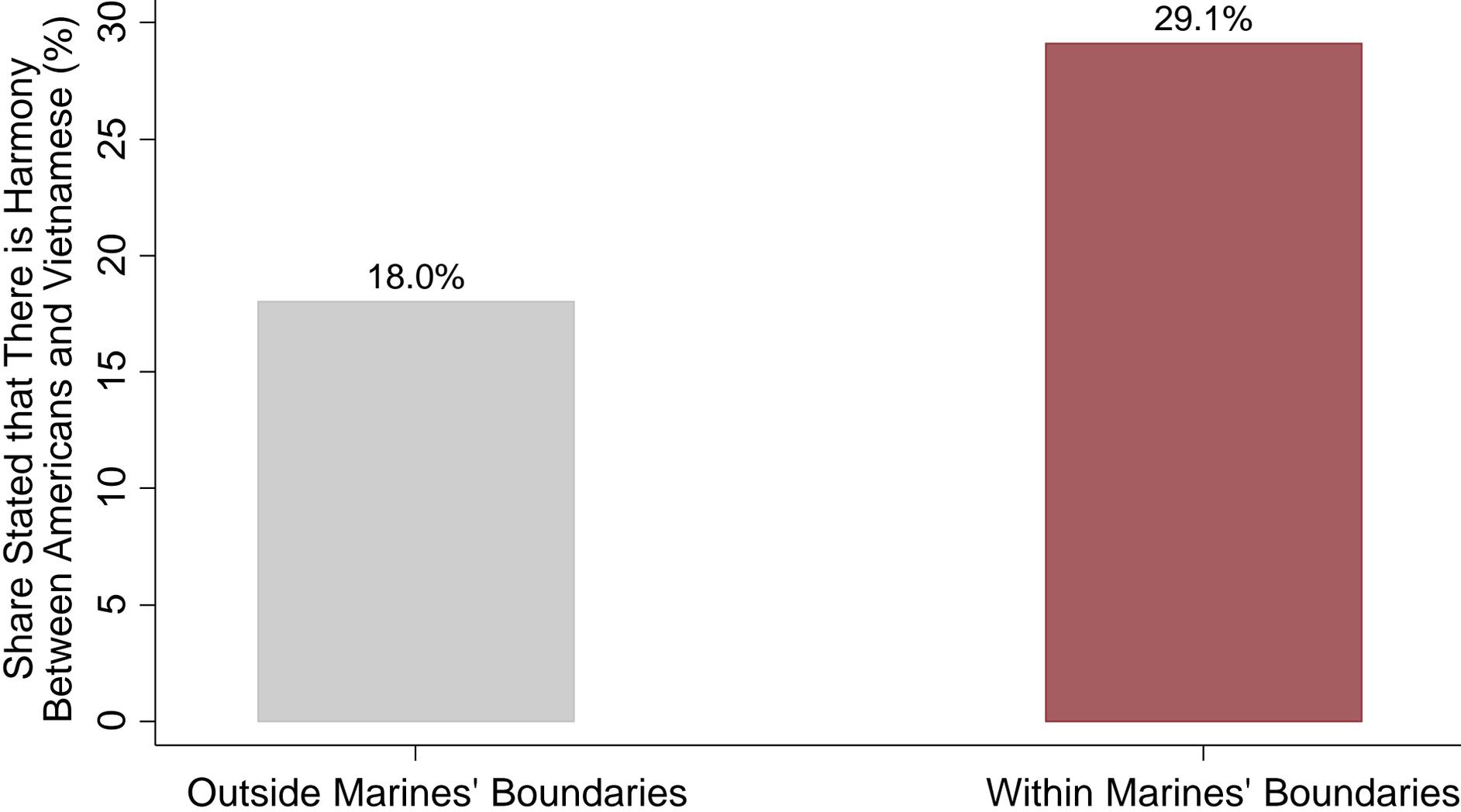
Difference: **56.2**
SE: (19.8)

Effect on Likelihood of Liking Americans



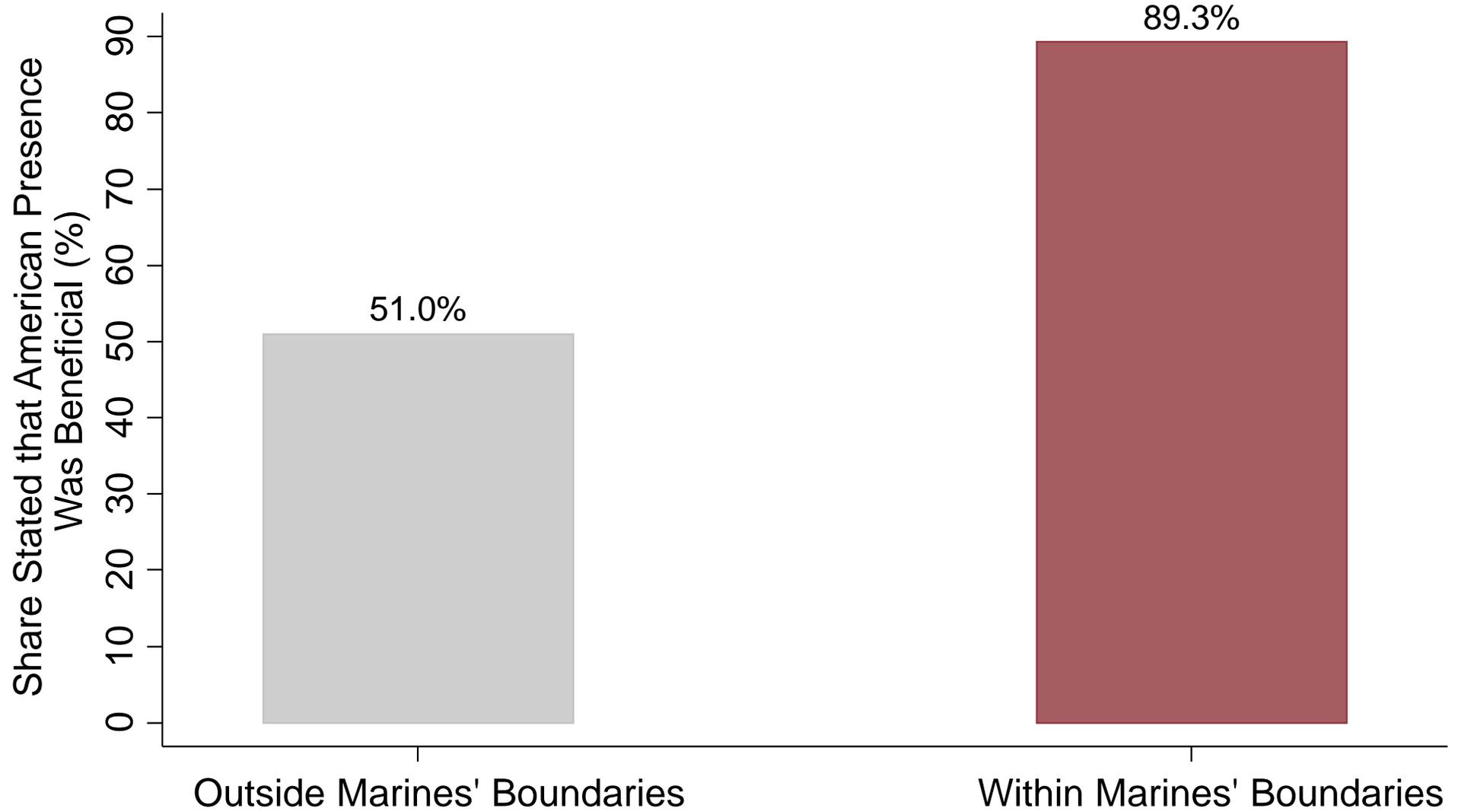
Difference: **15.8 pp**
SE: (8.2)

Effect on Stating There is Harmony Between Americans and Vietnamese



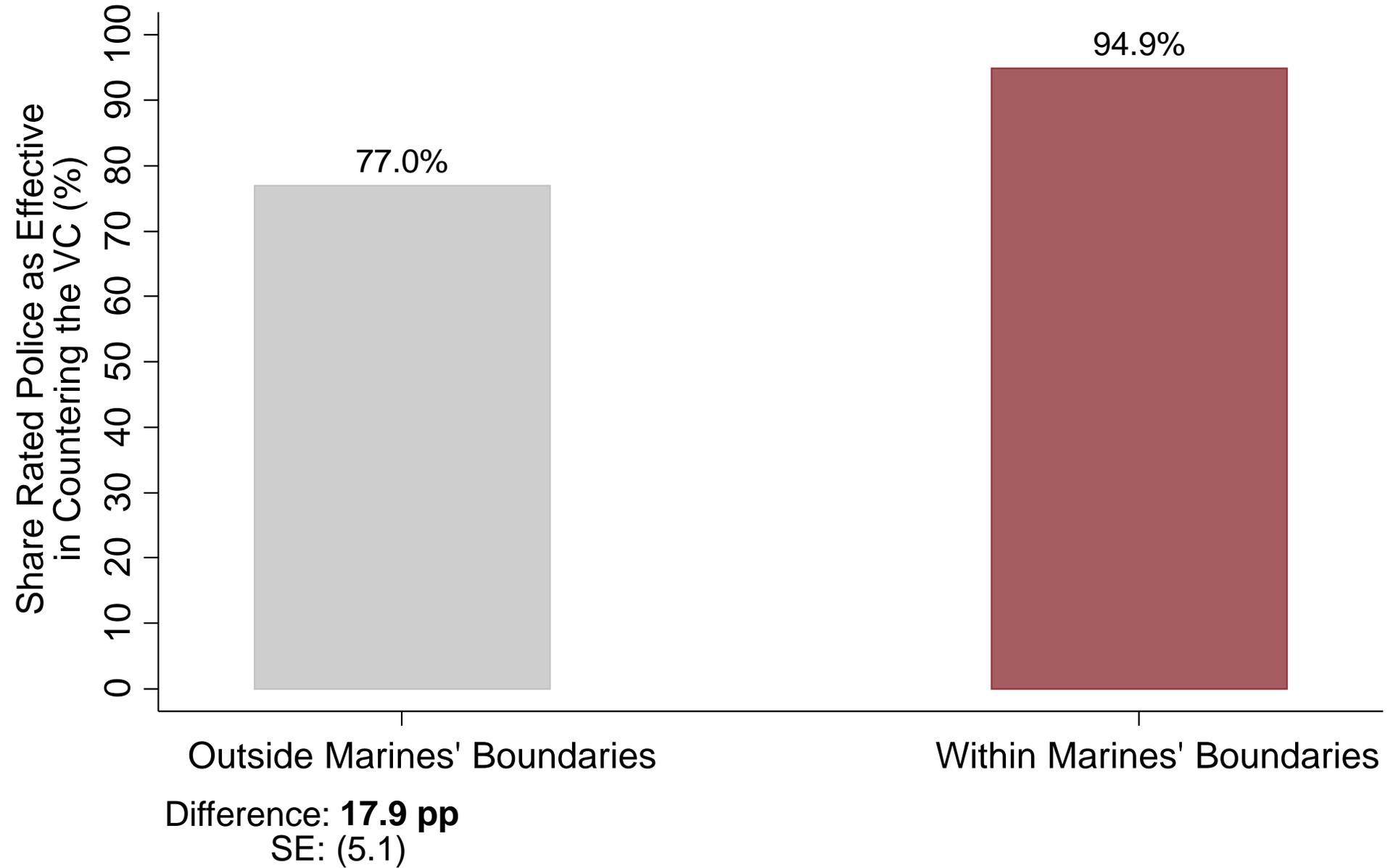
Difference: **11.1 pp**
SE: (5.8)

Effect on Stating that American Presence was Beneficial

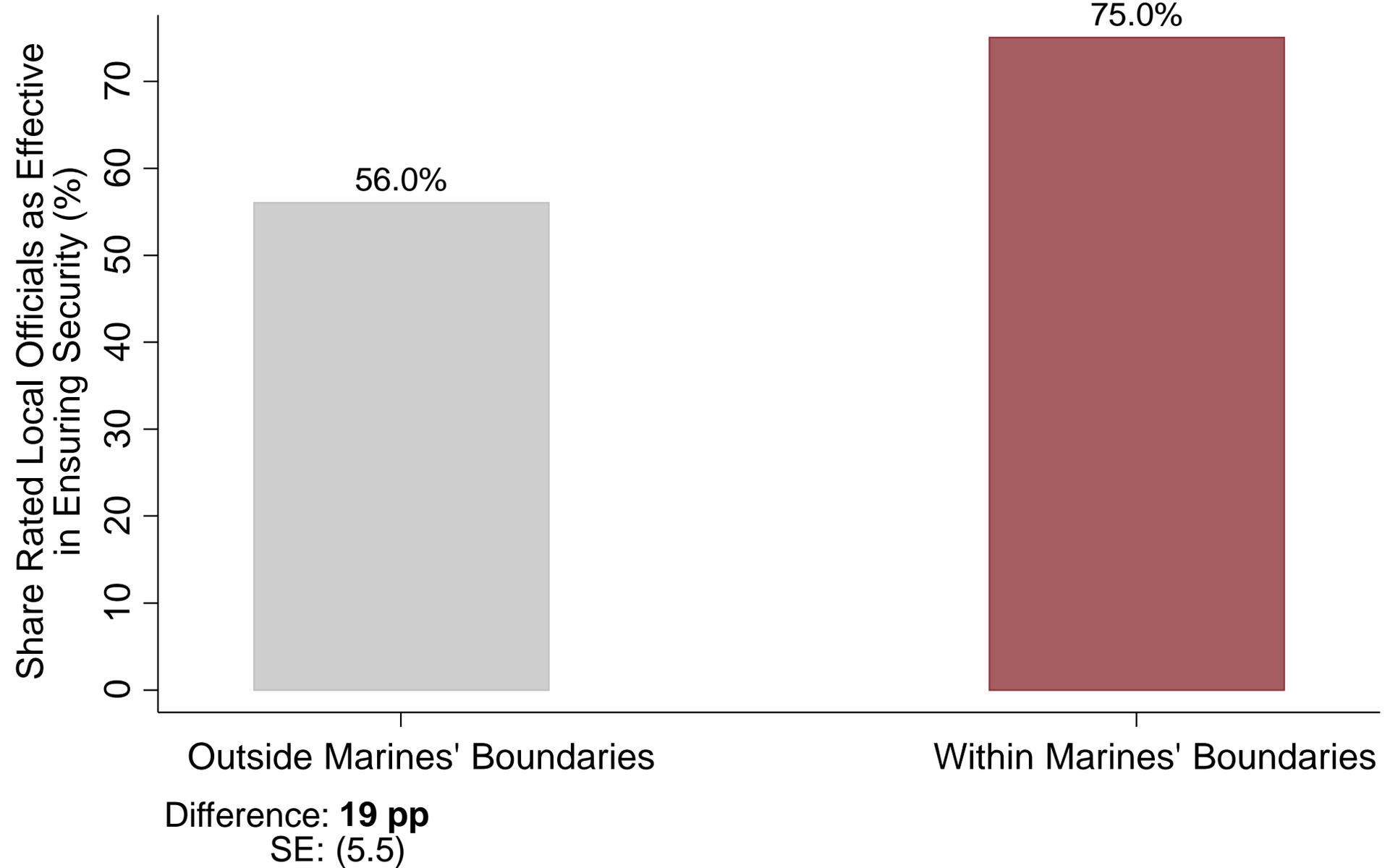


Difference: **38.3 pp**
SE: (7.9)

Effect on Rating Police as Effective in Countering Viet Cong



Effect on Rating Local Officials as Effective in Ensuring Security



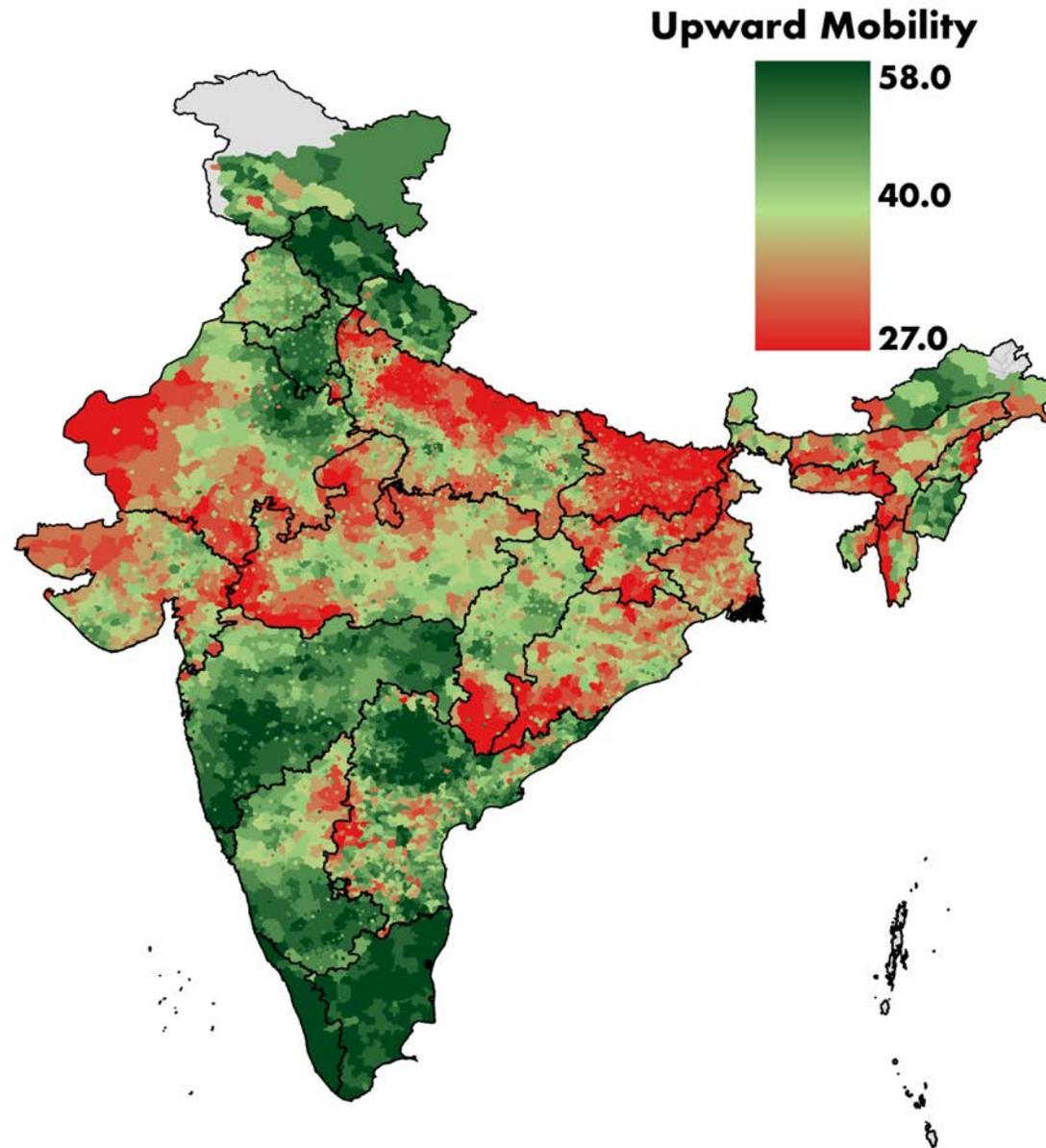
Implications for Nation Building Strategies

- Military force may worsen security and undermine local government relative to a hearts and minds oriented approach
 - Citizens have many ways to undermine a state they don't support
- Broader lesson: careful empirical evaluation is feasible and valuable even for challenging questions like military strategy

Economic Development in the Era of Big Data

- Overarching theme of recent work: policies and institutions vary at a sub-national level in important ways
- Modern data is permitting researchers to study key questions in developing economics at much finer granularity, as in developed countries
 - Asher and Novosad (2018) analyze geography of economic opportunity in India, using methods analogous to those used in the U.S.

The Geography of Intergenerational Mobility in India





EC 1152

EC CONCENTRATION

- Economic Theory (Ec 10 **and** 1010/11)
- Calculus (Math 1a)
- Stat (Stat 100, Stat 104, **or** Stat 110)
- Econometrics (Ec 1123 **or** 1126)
- Sophomore Tutorial (Ec 970)
- Three additional Economics electives

EC SECONDARY FIELD

- Economic Theory (Ec 10 **and** either 1010a/11a **or** 1010b/11b)
- Three additional Economics electives

SAMPLE ELECTIVES:

- Ec 1033 Golub: Social Networks
- Ec 1078 Goodman: Analyzing Education Policy
- Ec 1393 Nunn: Poverty and Development
- Ec 1410 Feldstein: Public Economics
- Ec 1800 Glaeser: The Economics of Cities
- EMR 20 Cutler: Health Care Policy
- Ec 1017 Miron: A Libertarian Perspective on Economic and Social Policy

SAMPLE JUNIOR SEMINARS:

- Goldin: Economics of Work and Family
- Hendren: Market Imperfections and Implications for Government Intervention