14.771 Development Economics: Microeconomic issues and Policy Models

Fall 2008

Does Health Affect Productivity?

Esther Duflo

14.771

The Impact of Health on Productivity?

- Strong biological reasons to think that health (and nutrition) affects productivity: strength, days of illness, etc.
- At the micro-level, some indicators of health show fairly strong relationship with earnings ▸ StraussFigure
- At the macro-level, some have argued extremely high impact of health on GDP per capita (Sachs, Commission on macroeconomics and health)
- E.g. Gallup and Sachs (2003) log(GDP/capita)=-1.3*Population exposed to Malaria
- Potential problems with these estimates (both micro and macro?)
- Today we will focus on both micro and macro estimates of the productivity impact of health, which are trying to go around these problems.
- We will start by taking a step back and think about how to correctly estimate such effects.

The Rubin Causal Model

(Reference: Imbens and Woolridge, 2008).

- Consider a binary treatment W: 1 for treated, 0 for control, and an outcome Y (e.g. the treatment is : received an iron pill, the outcome could be: anemia, or earnings).
- Ex-ante, each individual i has two *potential outcomes*, $Y_i(1)$ if treated, $Y_i(0)$ if non-treated.

$$Y_i = Y_i(1)W_i + Y_i(0)(1 - W_i)$$

- The *treatment effect* for individual i is $Y_i(1) - Y_i(0)$.
- Ex-post, only one of the outcomes is realized: individual is treated or non-treated. Since no individual is observed both in the treated and non-treated state, we will not be able to estimate the treatment effect for each individual. All we can hope to estimate are some statistics concerning the treatment effect for a sample of individual.

Estimand

- We could be interested in the average treatment effect for the population: $E[Y_i(1) - Y_i(0)]$.
- we could want to know the average treatment effect for those who receive the treatment: $E[Y_i(1) - Y_i(0)|W_i = 1]$.
- Could be interested in the average treatment for those who have some characteristics (observed or unobserved): $E[Y_i(1) - Y_i(0)|X_i = x]$
- Or we may want to know other things about the treatment:
  - How the treatment is affecting the distribution in treatment and control groups (quantile treatment effects).
  - The quantile of treatment effects (this is not the same, and it is very hard to know!)

Estimating Average Treatment Effect

Suppose we have a population, with N_1 treated individual, and N_0 non treated individuals. Consider the difference between treated and control population:

$$E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$$= E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 1]$$
$$+E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

$$= E[Y_i(1) - Y_i(0)|W_i = 1] + E[Y_i(0)|W_i = 1] - E[Y_i(0)|W_i = 0]$$

First term: ATT. Second term: difference in the underlying characteristics of the treated and non treated population (selection effect).

Selection mechanisms

Three cases:

- The probability of assignment does not depend on potential outcomes, and is a known function of covariates (*random assignment*). this case, $E[Y_i(0)|W_i = 1] = E[Y_i(0)|W_i = 0]$ and $E[Y_i(1)|W_i = 1] - E[Y_i(0)|W_i = 0]$ is an unbiased estimate of the effect of the treatment on the treated. Health example: Thomas et al (iron); Miguel and Kremer (worms)
- The probability of assignment does not depend on potential outcomes, but is an *unknown* function of covariates .

$$W_i \perp (Y_i(1), Y_i(0)) | X_i$$

(unconfoundness assumption, a.k.a. exogeneity, selection on observables). In this case,
$E[Y_i(0)|W_i = 1, X = x] = E[Y_i(0)|W_i = 0, X = x]$, so the selection bias disappears if we appropriately control for x. *Matching, propensity score matching, regressions*, are various ways to deal with this.

Selection mechanisms (3)

- The probability of assignment depends on potential outcomes: there is a selection bias of unknown size. Program evaluation question is to find ways to deal with that. Leading strategies: *Difference-in-differences*, *Regression Discontinuity*, *Instrumental variables*.

Difference in Differences

Simplest setting:

- Individual i belong to one of groups $G = 1$, treated group, $G = 0$, non treated group.
- and is observed in one of two periods (or cohorts) $T = 1$ (post) and $T = 0$ (pre).
- Group $G = 1$ is treated when $T = 1$, not when $T = 0$.
- Identification Assumption: Potential outcome $Y_i(0)$ can be written:

$$Y_i(0) = \alpha + \beta T_i + \gamma G_i + \epsilon_i$$

  with $\epsilon_i \perp (T, G)$, i.e. ϵ_i is independent of the group indicator and its distribution does not change over time.
- Then: $Y_i(1) = Y_i(0) + \tau_{DID}$
- What is the key identification assumption?

Difference in difference estimator

$$\tau_{DID} = (E[Y_i|G=1, T=1) - E[Y_i|G=1, T=0])$$
$$-((E[Y_i|G=0, T=1) - E[Y_i|G=0, T=0]))$$

Sample equivalent:

- Replace expectation by population averages:

$$\tau_{DID} = (\overline{Y_{11}} - \overline{Y_{10}}) - (\overline{Y_{01}} - \overline{Y_{00}})$$

  where $\overline{Y_{gt}} = \frac{1}{N_{gt}} \sum_{G_i=g, T_i=t} Y_i$
- Or equivalently estimate OLS on

$$Y_i = \alpha_1 + \beta_1 T_i + \gamma_1 G_i + \tau_{DID}(T_i * G_i) + \epsilon_i$$

- Under the identification assumption, it is easy to show that τ_{DID} recovers the average treatment effect.

Example: Malaria Eradication in the Americas (Bleakley, 2007)

- Set-up:
  - Relatively swift ▸ Malaria Eradication Campaigns
  - Intensity of treatment depends on whether there was malaria before or not ▸ Figure
- Diff in Diff
  - Definition of treated and control cohorts in the US:
  - 1920 or later
  - Definition of treated and control regions
  - 1899 or earlier
  - Results: ▸ Regression
  - How would it look in a Diff and Diff table?
- Testing the identification assumption
  - Old versus very Old ▸ Results
  - Young versus very Young ▸ Results

Extension: Multiple groups, multiple Periods, or both

Let T denote the number of periods, and G the number of groups:

$$Y_i(0) = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \epsilon_i$$

and $Y_i(1) = Y_i(0) + \tau_{DID}$
The model can be estimated with OLS regression:

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \tau_{DID} W_i + \epsilon_i$$

Where as before W_i is 1 for treated group for treated periods.

Extension: variable treatment intensity across periods

- Equivalent to have several treatments W_i^t, where W_i^t is equal to 1 for treated groups in year t

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t 1[T_i = t] + \sum_{g=1} \gamma 1[G_i = g] + \sum_{t=2}^{T} \tau_{tDID} W_i^t + \epsilon_i$$

  (alternatively: compute a series of DID relative to one base period)
- Specification check: the treatment effect should follow the pattern of the extension of the program. It should be be 0 for all the periods before the treatment starts; it should equal for all periods where the treatment intensity was the same.
- In the malaria case, exposure depends on cohort of birth in a specific way: ▸ Exposition By Age
- We get this ▸ graph for the coefficients: encouraging?

Extension: Continuous treatment intensity across groups

- Suppose that the intensity of the treatment also depend on the group. We can think about this as if it were several treatments: $Y_i(w)$, for $w = 0, 1, 2, 3$.
- Alternatively, if we define $Wi = 1$ is the unit got any treatment, for some observable variable X, we may want to model $Y_i(Wi * X_i) = g(X_i) + Y_i(0)$
- For example, in Bleakley's case: X is the pre-campaign intensity level in the group, and he assumes linearity: $Y_i(g) = \tau_{CDID} M_g + Y(0)$
- With only two cohorts:

$$Y_i = \alpha + \beta T + \sum_{g=1}^{G} \gamma 1[G_i = g] + \tau_{CDID}(M_g * T_t) + \epsilon_i$$

▸ Bleakley, Table 2

Extension: Continuous treatment intensity across groups

- With more than 2 cohorts.
- Do the two cohorts approach cohort by Cohort, and graph the results: which pattern should it have? ▸ graph.

- He then tests whether the cohort effects have the right shape: ▸ Cohort pattern, US. ▸ Cohort pattern, Other countries
- Alternatively, we can follow the "multi-cohorts" approach:

$$Y_i = \alpha + \sum_{t=1}^{T} \beta_t^T 1[T_i = t] + \sum_{g=1}^{G} \gamma_g 1[G_i = g] + \sum_{t=2}^{T} \tau_{CDIDt}(M_g * T_t) + \epsilon_i$$

▸ Table Appendix D.

Macro Implication

- Similar identification strategies are used by Cutler et al. (2008) and Lucas (2008) for looking at the impact of malaria on education. Results are quite comparable
- What are the macro-economic implications?
- Estimate that childhood infection decreases income by -0.5. Assuming no general equilibrium effect, this is also the GDP estimate
- Sachs's estimates translate into -2.16 (-1.3/0.6)
- This is much lower, but still significant (malaria would account for 10%-15% of the gap of Brazil and Mexico with the US)

Acemoglu and Johnson: A more macro approach

- A potential issue when going from micro-estimate to macro-estimate is the possibility of equilibrium effects: in the case of malaria, we compare cohorts. Maybe the younger cohorts are richer than the richer cohorts, but everybody is richer (or poorer) than they would be otherwise.
- The problem with macro setting is to find plausible source of variation
- Acemoglu and Johnson (2007) use the same identification strategy as Bleakley, but in a cross-country setting, for the disease against significant progress were made in the post-war period. (mainly turberculosis, pneumonia, malaria)
- Treatment intensity is a function of pre-campaign morbidity from those diseases (as in Bleakley).
- See graphs: Considerable gains in ▸ life expectancy and ▸ population.
- ... but ▸ no gains in GDP.
- so on balance a *loss* in GDP per capita.

Long Term Impact of Low Nutrition on Productivity

- "Barker" hypothesis (or fetal health). What matter in early childhood continue to matter later in life
- Evidence: Doblhammer–long term impact of month of birth, likely linked to nutrition available to mother.
- Almond, Qian: long term impact of famine in China (even on survivors, despite selection)
- Almond: people who were in gestation during 1918 influenza epidemics have lower life expectancy
- Banerjee, Duflo, Postel-Vinay and Watts: impact of shock at birth on height at 20.
- Field: Iodine supplementation

Iodine supplementation in Tanzania

- 1 billion people at risk of iodine shortage (old soil and no seafood)
- Iodine deficiency in first trimester of pregnancy thought to lead to permanent irreversible brain damage, apparently especially for girls.
- Tanzania had an intensive campaign to distribute iodine capsules, ultimately reaching 25% of the population, starting in 1986, and targeted to the 25 districts that had the largest goiter rates.
- In principle, women must receive a capsule every 2 years (duration of the dose). In practice some districts started later, and the distribution was not every 2 years: ▸ table 1

Strategy and results

- Authors calculate the probability that a child born in a given month was covered when in Utero (as a function of when the pills were distributed in the district) and introduce district fixed effect and month fixed effect (compare children born at the wrong time in treated district).
- They also carry out the analysis within households (siblings born a little bit too late or too early).
- Results in ▸ table 3 ▸ table 4 : Large effects, especially for girls.
- Robustness: Given the dose, effect should be highest when IDD is not too high and not too low: compare results across regions which produce more or less cassava: ▸ table 6
- Other than the effect through cognition, what could be the channel through which this intervention affect education? What regression can they run to rule them out?

Conclusion: Back to Das Gupta and Ray

- There is a strong relationship between health and productivity at the micro level (and also between education and productivity)
- Role of Micro-nutrients seems particularly important (iodine, iron: Thomas).
- No very solid estimate of the impact of nutrition on productivity (nobody does that!) but earlier estimates suggest an elasticy of about 0.4 (Strauss).
- Impact of nutrition in-utero and in childhood may be much larger than later in life, since it may cause permanent damage on health (so impact would be multiplied by years of life), and also through amplification impacts through education.
- Need to go back to thinking in more detail about what is happening within the household: if nutrients are indeed shared more unequally in the household when there is a shock (as the Das Gupta Ray model would suggest), this may create a space for a inter-generational poverty trap to emerge.

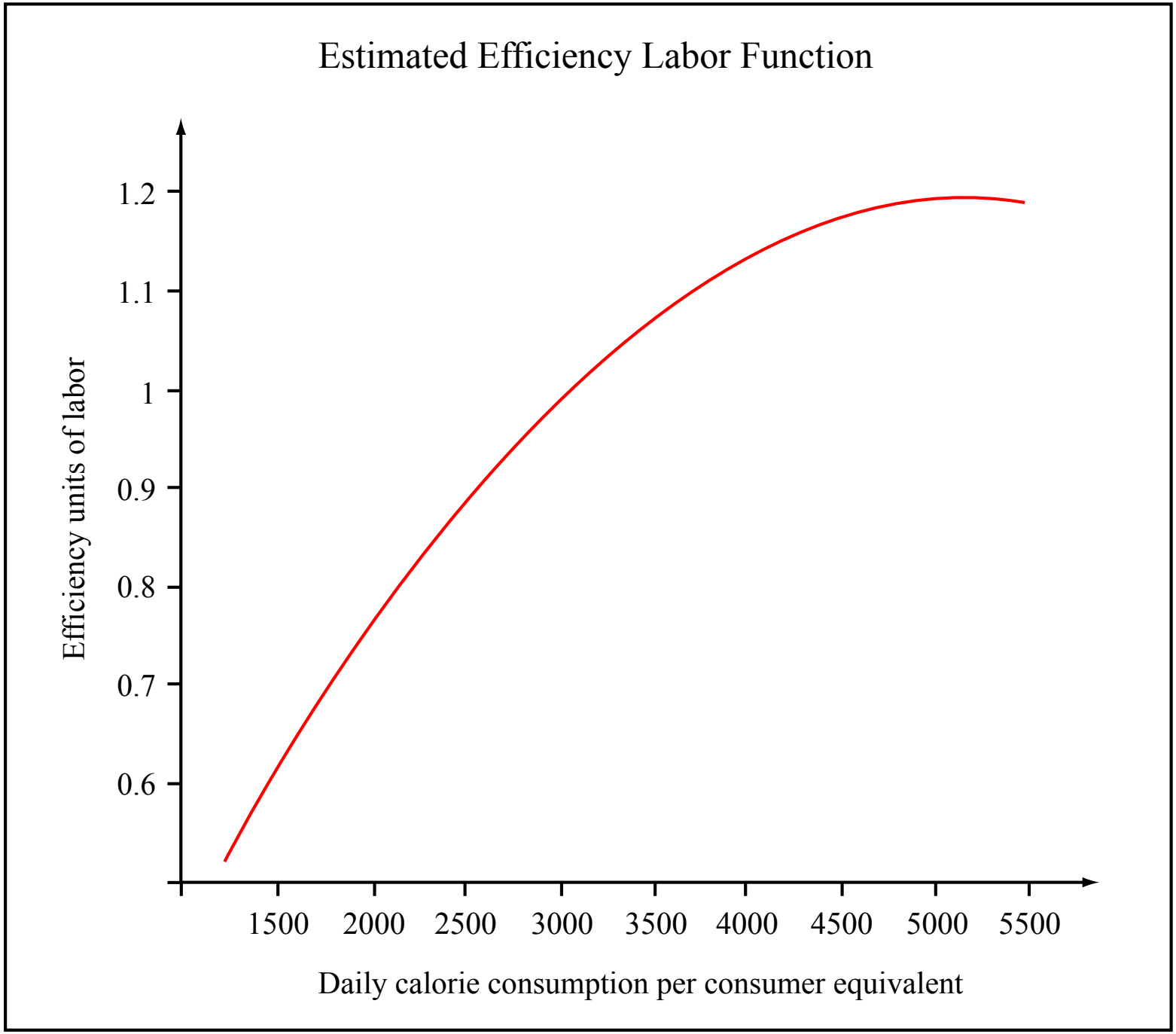
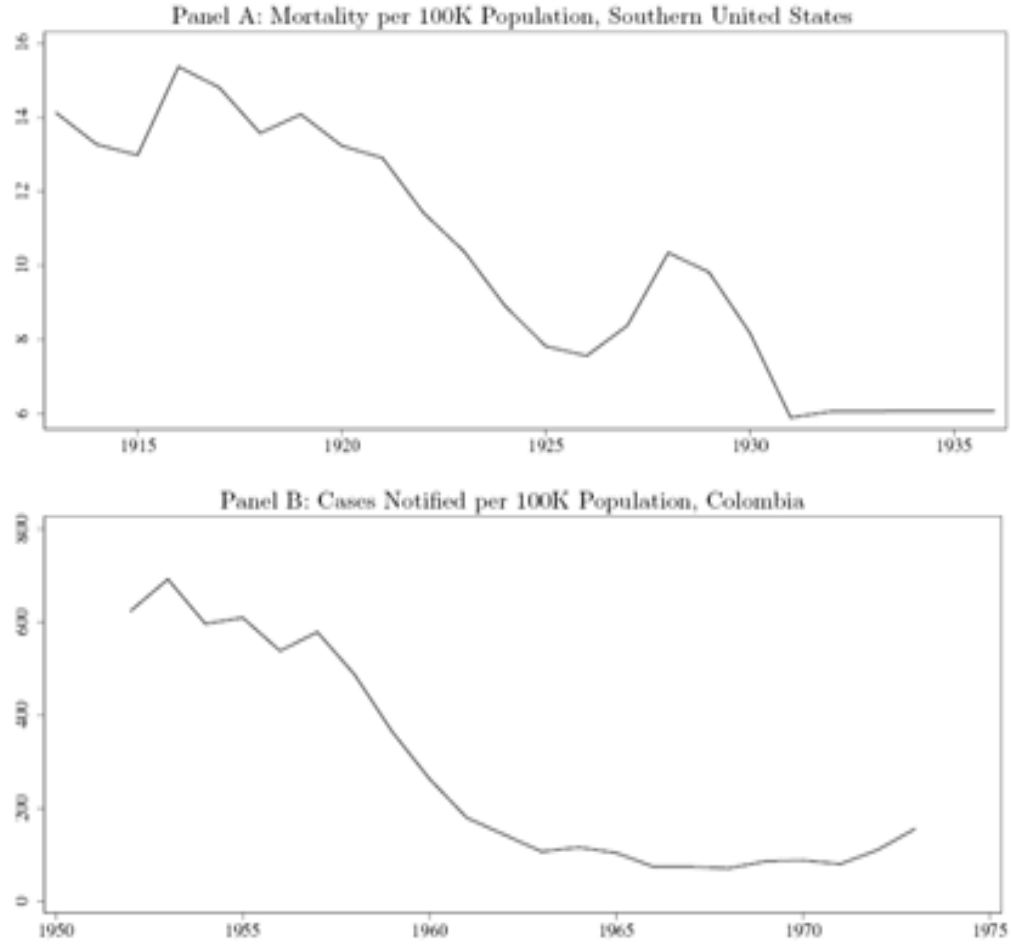


Figure by MIT OpenCourseWare.

Bleakley (2006)

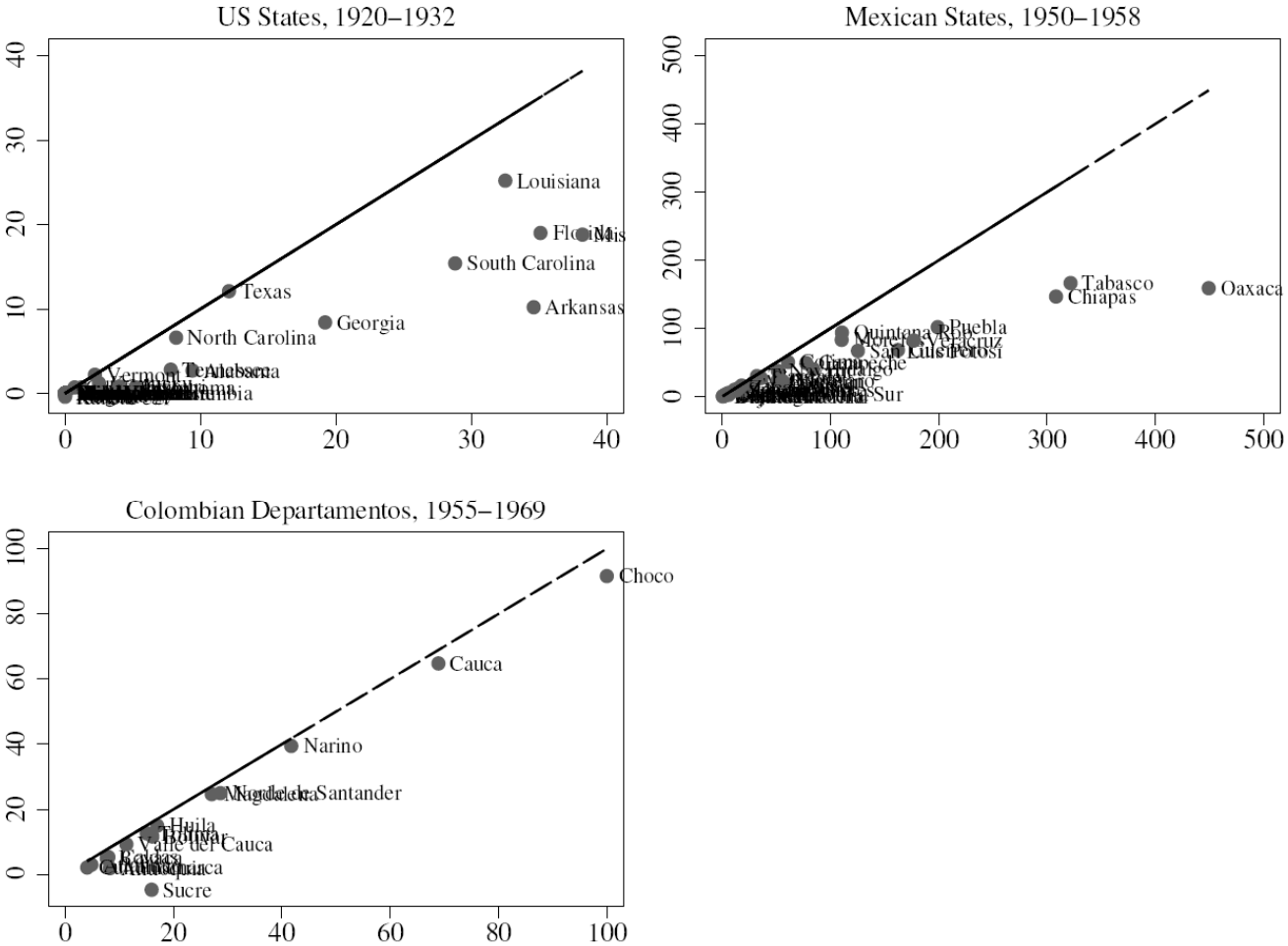
Figure 1: Malaria Incidence Before and After the Eradication Campaigns



Notes: Panel A plots the estimated malaria mortality per capita for the Southern region and bordering states. Because the death registration system was being phased in over the period, a regression model with state fixed effects is used to control for sample changes, and the time series is constructed from the year dummies in the regression, normalized to match the end-of-period data when all states were represented. (Census Bureau *Vital Statistics*, various years, and author's calculations.) Panel B reports data on notified cases of malaria for Colombia (SEM, 1979).

Bleakley (2006)

Figure 2: Highly Infected Areas Saw Greater Declines in Malaria



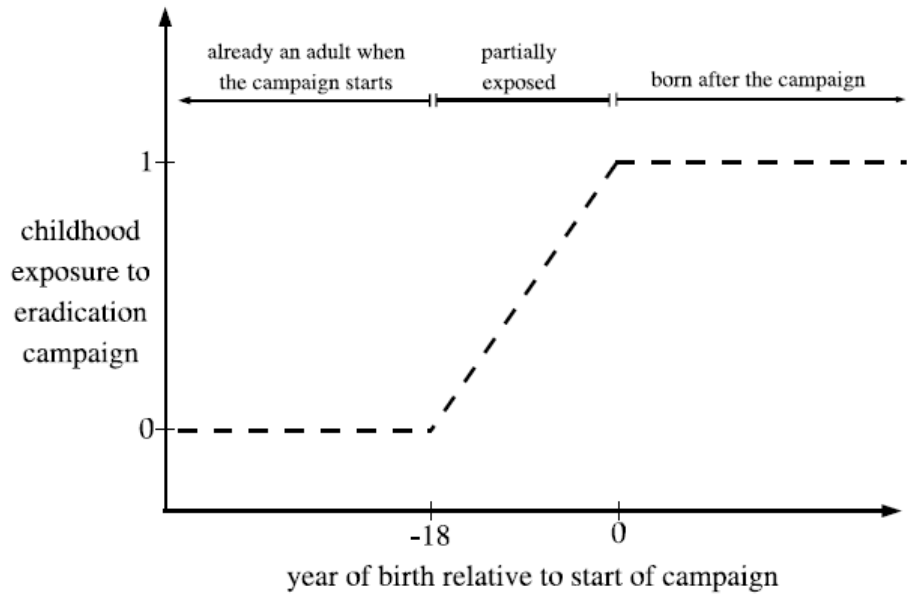
Notes: The y axis displays the estimated decrease in malaria mortality post-intervention. The x axis is the pre-campaign malaria mortality rate. The 45-degree line represents complete eradication. Both variables are expressed per 100,000 population. United States data are reported in Maxcy (1923) and *Vital Statistics* (Census, 1933). Mexican data are drawn from Pesqueira (1957) and from the Mexican *Anuario Estadístico* (Dirección General de Estadística, 1960). SEM (1957) and the Colombian *Anuario de Salubridad* (DANE, 1968-70) are the sources for the Colombian data.

```
.
. ***************** PART 2 *****************
. g byte mhigh_young= mal_high*young

. reg sei mhigh_young mal_high young if (young==1 | old==1) & mal_highlow, r

Linear regression                                      Number of obs =    2044
                                                       F(  3,  2040) =  501.00
                                                       Prob > F      =  0.0000
                                                       R-squared     =  0.3374
                                                       Root MSE      =  .30635

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
 mhigh_young |    .159573   .0231584     6.89   0.000     .1141563    .2049897
    mal_high |  -.2190388   .0200889   -10.90   0.000    -.2584357    -.179642
       young |   .3299299    .014544    22.68   0.000     .3014073    .3584525
       _cons |   3.375146   .0133338   253.13   0.000     3.348997    3.401295
------------------------------------------------------------------------------
```

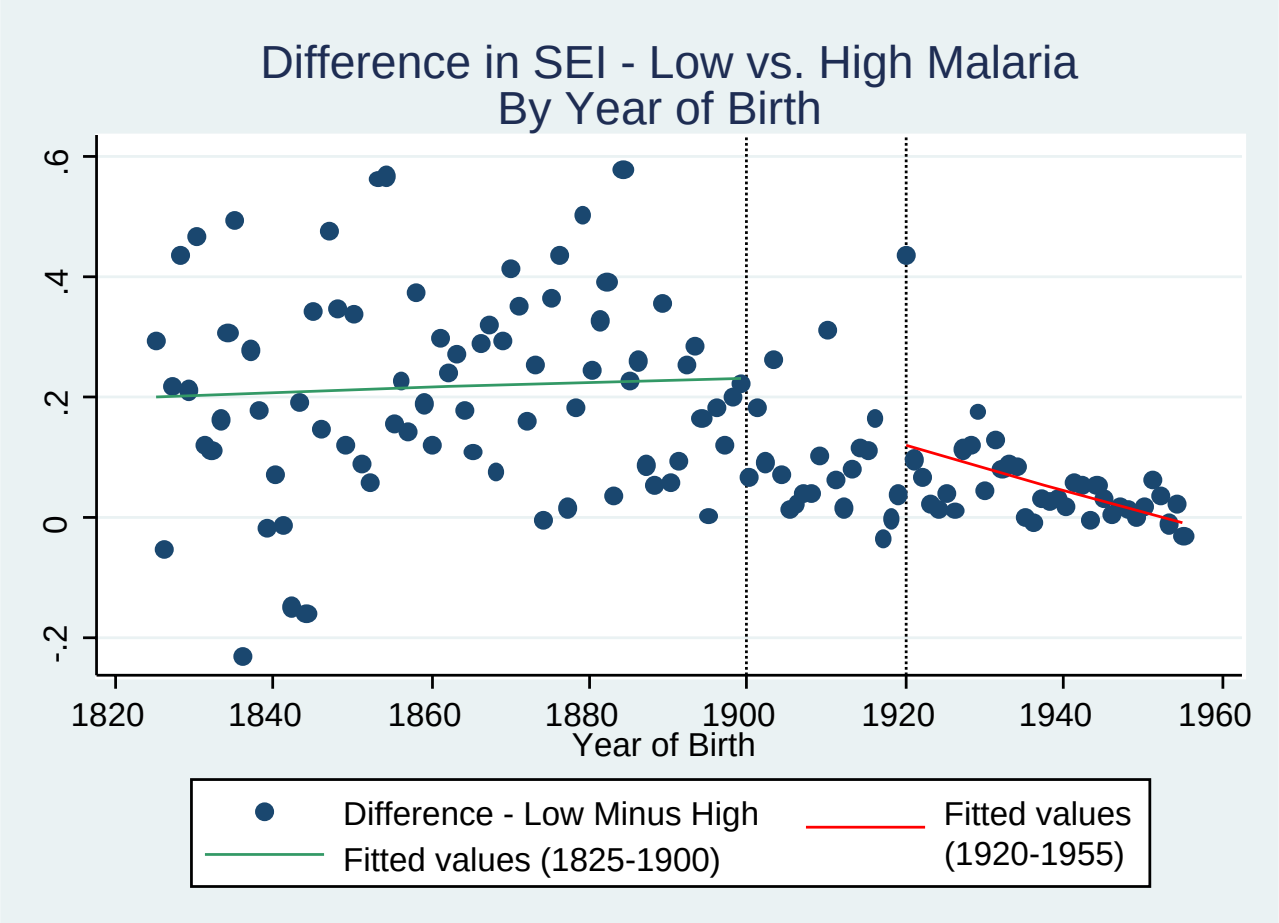
```
.
. reg sei mal_high yold mhigh_yold if old & mal_highlow, r

Linear regression                                       Number of obs =    1279
                                                        F(  3,  1275) =   58.19
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.1257
                                                        Root MSE      =  .35491

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    mal_high |  -.2069026   .0321059    -6.44   0.000    -.2698887   -.1439165
        yold |   .1692435   .0259945     6.51   0.000     .1182468    .2202402
  mhigh_yold |  -.0152891   .0405002    -0.38   0.706    -.0947435    .0641654
       _cons |   3.268976   .0197418   165.59   0.000     3.230246    3.307706
------------------------------------------------------------------------------

. reg sei mal_high vyoung mhigh_vyoung if young & mal_highlow, r

Linear regression                                       Number of obs =     765
                                                        F(  3,   761) =   19.75
                                                        Prob > F      =  0.0000
                                                        R-squared     =  0.0729
                                                        Root MSE      =  .16983

------------------------------------------------------------------------------
             |               Robust
         sei |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    mal_high |  -.0775506   .0165237    -4.69   0.000    -.1099882   -.0451131
      vyoung |   .0440054   .0107842     4.08   0.000     .0228351    .0651756
mhigh_vyoung |   .0519633    .018939     2.74   0.006     .0147845    .0891421
       _cons |   3.689949   .0077427   476.57   0.000      3.67475    3.705149
------------------------------------------------------------------------------
```

Bleakley (2006)

Figure 3: Childhood Exposure to Eradication Campaign

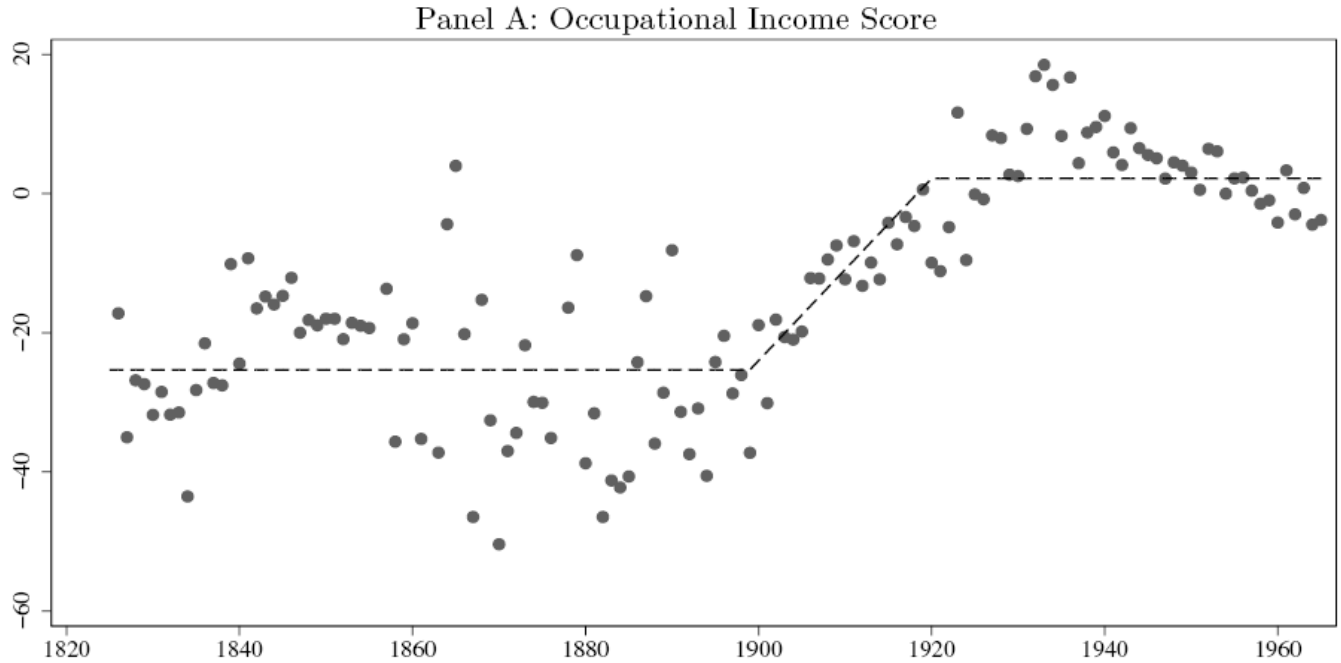


Notes: This graph displays on the fraction of childhood that is exposed to a hypothetical (and instantaneous) campaign as a function of year of birth minus the start year of the campaign.


Difference in SEI - Low vs. High Malaria
By Year of Birth
-.2
0
.2
.4
.6
1820
1840
1860
1880
1900
1920
1940
1960
Year of Birth
Difference - Low Minus High
Fitted values (1825-1900)
Fitted values (1920-1955)

Bleakley (2006)

Figure 4: Cohort-Specific Relationship: States in the U.S.



Notes: These graphics summarize regressions of income proxies on pre-eradication malaria-mortality rates (measured by the Census in 1890). The y axis for each graphic plots the estimated cohort-specific coefficients on the state-level malaria measure. The x axis is the cohort's year of birth. Each cohort's point estimate is marked with a dot. The dashed lines measure the approximate number of years of potential childhood exposure to the malaria-eradication activities in the South. For each year-of-birth cohort, OLS regressions coefficients are estimated on the cross section of states of birth. The state-of-birth average outcome is regressed onto malaria, Lebergott's (1964) measure of 1899 wage levels, a dummy for the Southern region, and the various control variables described in Appendix C. Appendices A and B describe, respectively, the outcome variables and the malaria measure.

Bleakley (2006)

Table 1: Exposure to Malaria Eradication versus Trends

Degree of Polynomial-Trend Control:	0	1	2	3
Outcome Variables:				
		Panel A: United States		
Occupational Income Score	28.684 ***	33.802 ***	34.611 ***	34.235 ***
	(1.509)	(3.664)	(4.105)	(5.412)
	{0.109}	{0.129}	{0.132}	{0.130}
Duncan's Socioeconomic Index	52.549 ***	48.862 ***	57.078 ***	55.248 ***
	(2.956)	(6.654)	(7.485)	(9.782)
	{0.158}	{0.147}	{0.172}	{0.166}
		Panel B: Brazil		
Literacy	0.029 ***	0.018 ***	0.017 ***	0.002
	(0.002)	(0.004)	(0.004)	(0.006)
	{0.152}	{0.094}	{0.089}	{0.010}
Years of Schooling	0.214 ***	0.116 *	0.349 ***	0.179 **
	(0.025)	(0.070)	(0.057)	(0.090)
	{1.120}	{0.607}	{1.827}	{0.937}
Log Total Income	0.073 ***	0.094 ***	0.104 ***	0.084 ***
	(0.005)	(0.011)	(0.011)	(0.019)
	{0.382}	{0.492}	{0.544}	{0.440}
Log Earned Income	0.056 ***	0.080 ***	0.082 ***	0.048
	(0.008)	(0.022)	(0.025)	(0.054)
	{0.293}	{0.419}	{0.429}	{0.251}

Bleakley (2006)

Table 1: Exposure to Malaria Eradication versus Trends

Degree of Polynomial-Trend Control:	0	1	2	3
Outcome Variables:				
		Panel C: Colombia		
Literacy	0.023 **	0.047 *	0.058 **	-0.019
	(0.011)	(0.026)	(0.028)	(0.052)
	{0.009}	{0.019}	{0.023}	{-0.008}
Years of Schooling	0.800 ***	0.854 **	0.683 **	0.673
	(0.131)	(0.358)	(0.340)	(0.601)
	{0.317}	{0.338}	{0.270}	{0.267}
Industrial Income Score	0.170 ***	0.104 **	0.121 ***	-0.146
	(0.016)	(0.047)	(0.044)	(0.090)
	{0.067}	{0.041}	{0.048}	-{0.058}
		Panel D: Mexico		
Literacy	0.008 ***	-0.006	-0.009 *	0.008
	(0.003)	(0.004)	(0.005)	(0.007)
	{0.026}	{-0.019}	{-0.029}	{0.026}
Years of Schooling	-0.087 ***	-0.194 ***	-0.178 ***	-0.021
	(0.020)	(0.051)	(0.046)	(0.077)
	{-0.279}	{-0.623}	{-0.571}	{-0.067}
Log Earned Income	0.067 ***	0.021	0.063 **	-0.050
	(0.016)	(0.035)	(0.026)	(0.070)
	{0.215}	{0.067}	{0.202}	{-0.160}

Notes: This table reports estimates of the childhood-exposure variable in equation 2 using OLS. The outcome variables used to construct the time series of $\hat{\beta}_k$ are as indicated in each row. Robust (Huber-White) standard errors in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Observations are weighted by the inverse of the coefficient's standard error. Reporting of additional terms suppressed. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity. For the United States, this number is also normalized by the average value of the relevant income proxy for white males born in the South between 1875

Bleakley (2006)

Appendix D: Panel Estimates of Childhood Exposure

Dependent Variables: Income Measures by Cohort

Panel A: United States

Occupational Income Score	Duncan Index
27.936 ***	52.040 ***
(5.528)	(13.176)
{0.106}	{0.157}

Panel B: Brazil

Literacy	Education	Log Total Income	Log Earned Income
0.030 ***	0.171 ***	0.070 ***	0.046 ***
(0.005)	(0.051)	(0.021)	(0.017)
{0.156}	{0.894}	{0.366}	{0.242}

Panel C: Colombia

Literacy	Education	Income Index
0.209 ***	1.751 ***	0.149 ***
(0.032)	(0.425)	(0.051)
{0.083}	{0.694}	{0.059}

Panel D: Mexico

Literacy	Education	Log Earned Income
0.047 ***	0.229 *	0.151 ***
(0.011)	(0.140)	(0.036)
{0.150}	{0.736}	{0.486}

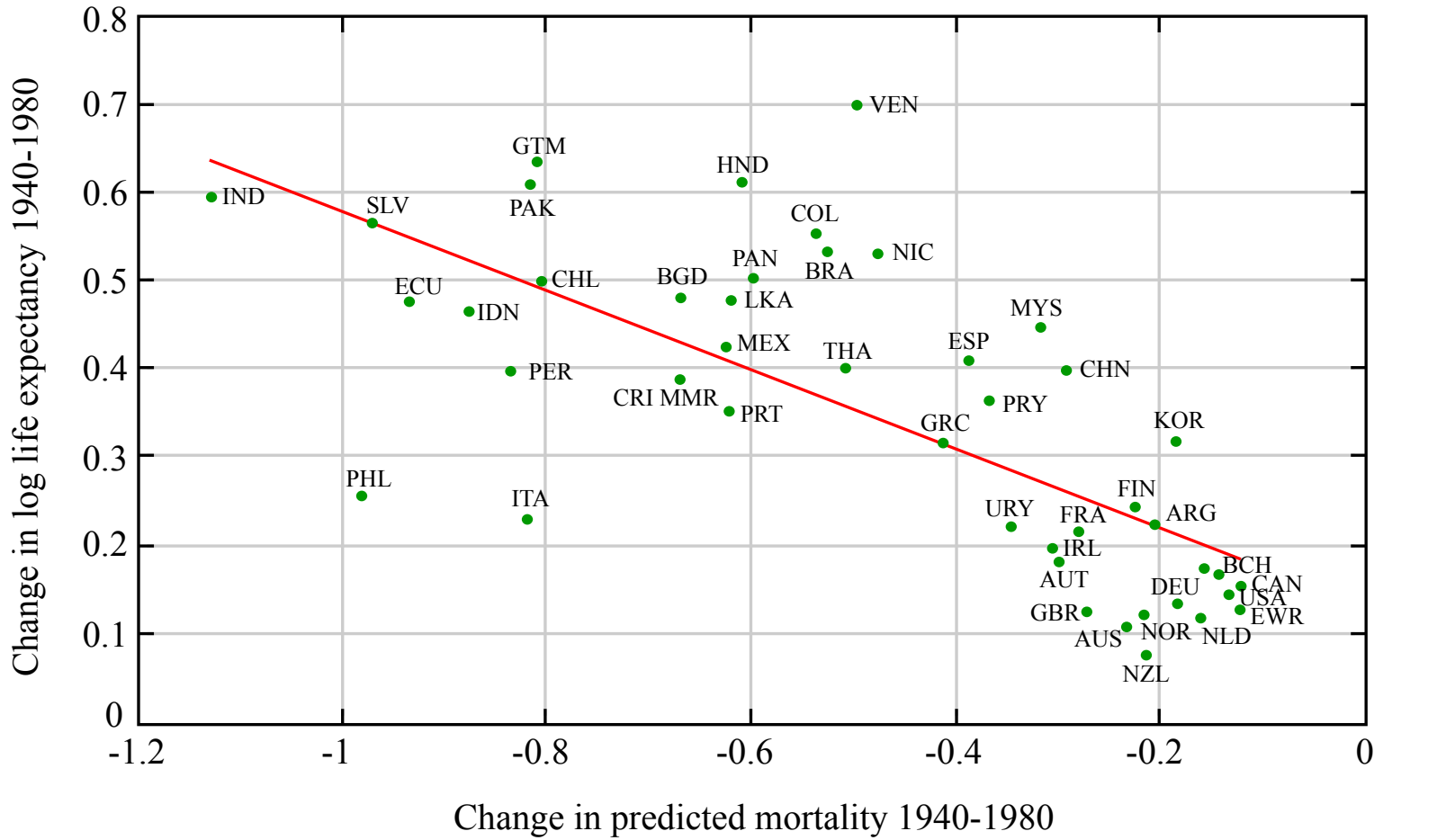
Bleakley (2006)

Table 2: Cross-Cohort Differences and Malaria: United States

	Malaria Mortality (Fraction of Total), 1890		Malaria Ecology (Mellinger)		Malaria Ecology (Hong)		Malaria Mortality (per 100K Population), 1920	
Dependent Variable:								
Occupational Income Score	X		X		X		X	
Duncan's Socioeconomic Index		X		X		X		X
Specification:	*Panel A: Basic Results*							
OLS, Basic Specification	37.927 ***	60.316 ***	0.570 **	1.191 **	16.278 ***	19.608 ***	0.030	0.057
	(11.101)	(21.311)	(0.267)	(0.535)	(2.040)	(4.737)	(0.021)	(0.036)
	{0.144}	{0.182}	{0.032}	{0.052}	{0.265}	{0.253}	{0.050}	{0.074}
2SLS, Using the Other Three Proxies as Instruments	44.367 ***	71.573 ***	1.312 *	2.064 *	15.133 ***	23.345 ***	0.074 **	0.110 **
	(14.238)	(24.199)	(0.748)	(1.075)	(3.813)	(8.205)	(0.031)	(0.053)
	{0.169}	{0.216}	{0.073}	{0.091}	{0.247}	{0.301}	{0.122}	{0.144}
Additional Controls:	*Panel B: Alternative Control Sets*							
Health	33.897 ***	63.480 ***	0.483 ***	1.078 ***	15.171 ***	24.580 ***	0.038	0.066
	(9.733)	(20.610)	(0.183)	(0.346)	(2.506)	(5.092)	(0.025)	(0.044)
Education	44.825 ***	59.306 **	0.552 **	1.080 ***	14.119 ***	16.543 ***	0.062 **	0.063
	(12.240)	(23.279)	(0.268)	(0.412)	(2.093)	(5.067)	(0.024)	(0.046)
Other	30.118 ***	45.827 **	0.388 **	1.050 ***	12.423 ***	13.082 ***	0.029	0.006
	(11.400)	(18.134)	(0.162)	(0.367)	(2.083)	(4.751)	(0.038)	(0.045)
Full Controls	33.392 **	59.257 **	0.385	0.985 **	15.564 ***	24.357 ***	0.048	0.060
	(13.844)	(29.103)	(0.236)	(0.473)	(3.280)	(7.088)	(0.030)	(0.056)

Notes: This table reports estimates of equation 3 using OLS and 2SLS. The units of observation are U.S. states. The dependent variables are as indicated in the column headings. Robust (Huber-White) standard errors in parentheses. Single asterisk denotes statistical significance at the 90% level of confidence; double 95%; triple, 99%. Reporting of constant term suppressed. Unexposed cohorts are those born before 1890 and fully exposed cohorts are those born after 1920. Cohorts are determined based on state of birth. The universe for the base sample consists of the native-born white population between the ages of 25 and 55 (15–55 for literacy) in the 1880–2000 census microdata from the IPUMS and NAPP databases. The terms in curly brackets report the point estimate multiplied by the difference between 95th and 5th percentile malaria intensity and normalized by the average value of the relevant income proxy for white males born in the South between 1875 and 1895. The specification for the basic results includes the malaria variable, a dummy for Southern birthplace, and the Lebergott (1964) measure of average unskilled wage in the state of birth. Appendices A and B describe, respectively, the outcome variables and malaria measures. The additional controls are described in the text and Appendix C.

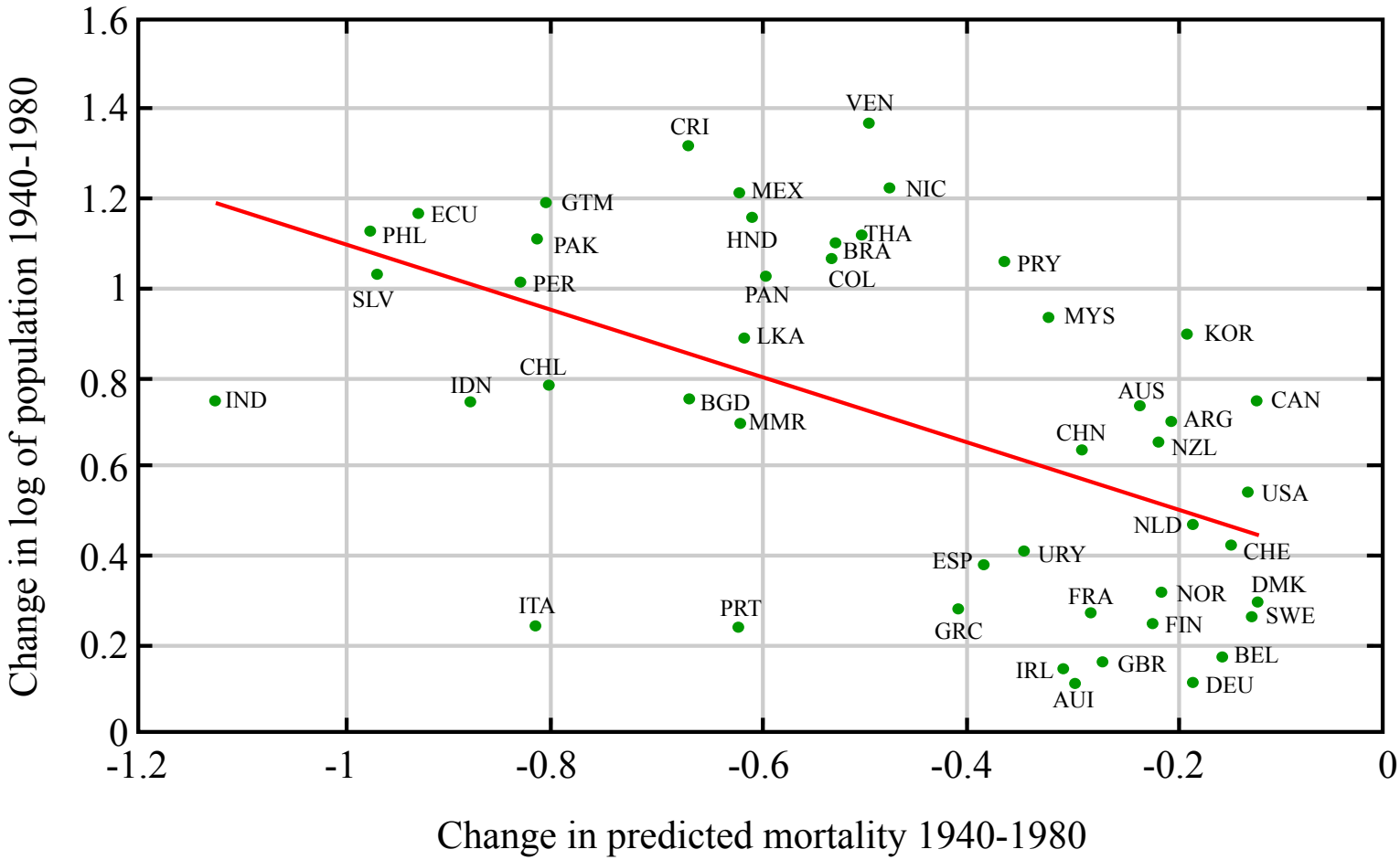
Acemoglu and Robinson (2007)



Change in log life expectancy and change in predicted mortality, 1940-80, base sample.

Figure by MIT OpenCourseWare.

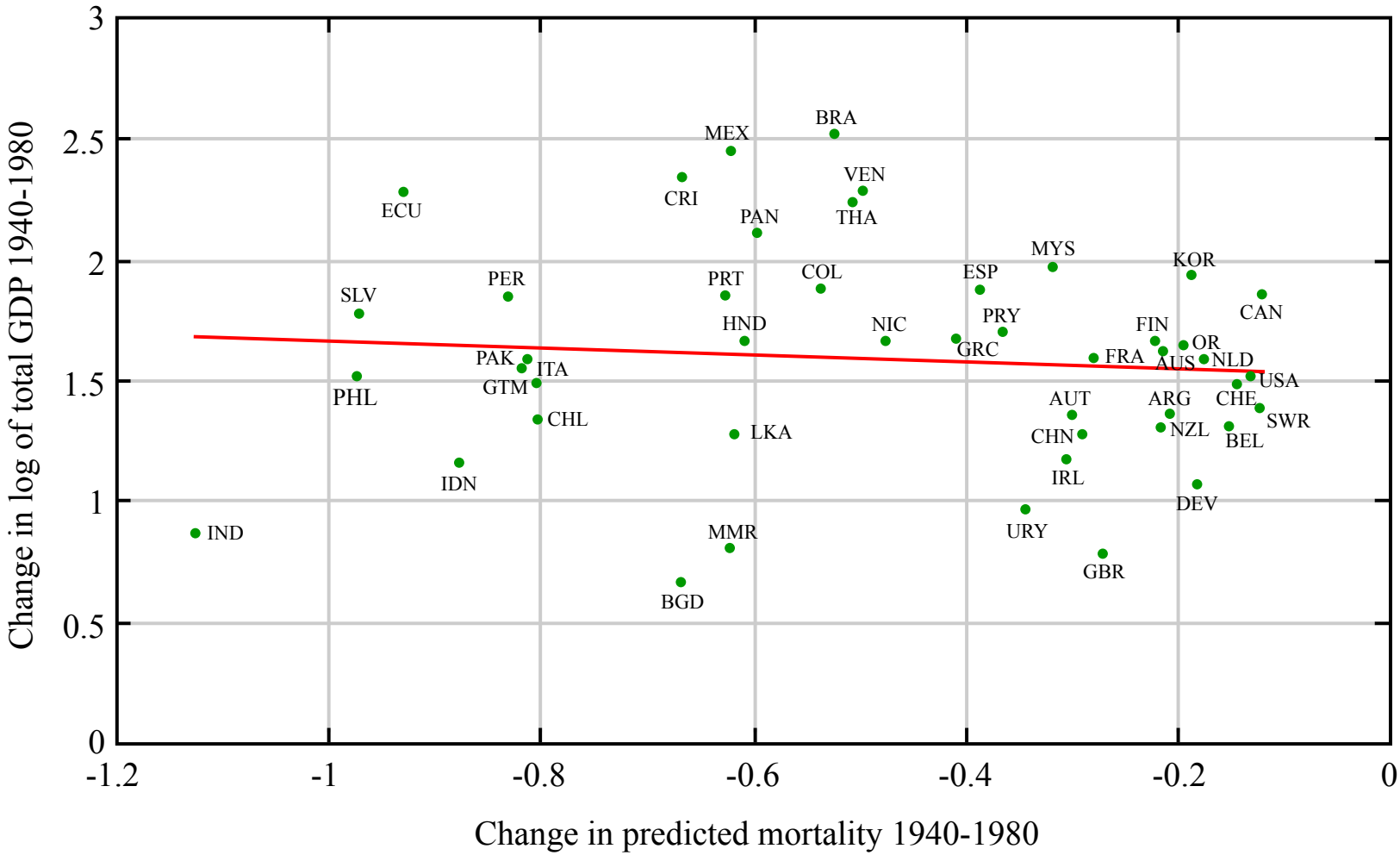
Acemoglu and Robinson (2007)



Change in log of population and change in predicted mortality, 1940-80, base sample.

Figure by MIT OpenCourseWare.

Acemoglu and Robinson (2007)



Change in log of total GDP and change in predicted mortality, 1940-80, base sample.

Figure by MIT OpenCourseWare.

Field, Robles, and Torero (2007)

Table 1: Summary of Timing and Coverage of Intervention Across Districts

	Region	District	Year of Intervention (Coverage - %)*					Average Frequency (yr)
			1	2	3	4	5	
1	Dodoma	Mpwapwa	1990 (65)	1992 (58)				2.00
2	Arusha	Monduli	1992 (71)					n/a
3	Arusha	Arumeru	1991 (89)					n/a
4	Kilimanjaro	Rombo	1990 (68)					n/a
5	Morogoro	Ulanga	1988 (73)	1991 (61)	1992 (34)			1.33
6	Ruvuma	Songea Rural	1987 (91)	1991 (74)	1995 (85)			2.67
7*	Ruvuma	Mbinga	1995 (92)					n/a
8	Iringa	Mufindi	1986 (41)	1991 (63)	1995 (54)			3.00
9	Iringa	Makete	1986 (20)	1991 (62)	1993 (62)	1996 (49)		2.50
10	Iringa	Njombe	1989 (76)	1992 (68)	1995 (64)			2.00
11	Iringa	Ludewa	1989 (59)	1992 (62)	1995 (47)			2.00
12	Mbeya	Chunya	1990 (49)					n/a
13	Mbeya	Mbeya Rural	1986 (44)	1989 (84)	1990 (90)	1993 (53)	1997 (53)	1.75
14	Mbeya	Kyela	1989 (91)	1993 (57)				4.00
15	Mbeya	Rungwe	1986 (35)	1990 (73)	1993 (49)			2.33
16	Mbeya	Ileje	1989 (94)	1992 (71)				3.00
17	Mbeya	Mbozi	1989 (67)	1991 (63)				2.00
18	Rukwa	Mpanda	1987 (79)	1991 (60)	1993 (72)			2.00

Courtesy of Erica Field, Omar Robles, and Maximo Torero. Used with permission.

Field, Robles, and Torero (2007)

Table 3. Grade Attainment and IOC Supplementation in Utero

	Boys and girls	Boys	Girls	Boys and girls	Boys	Girls	Boys	Girls
IOC in utero	0.357	0.315	0.77	0.157	-0.027	0.383		
(IOC in utero =born 1-3 years after program)	[0.142]*	[0.267]	[0.298]*	[0.112]	[0.154]	[0.166]*		
IOC in utero							-0.023	0.408
(IOC in utero =born 1-2 years after program)							[0.177]	[0.187]*
Age 11	0.558	0.571	0.743	0.401	0.45	0.296	0.451	0.251
	[0.140]**	[0.272]*	[0.256]**	[0.091]**	[0.125]**	[0.135]*	[0.126]**	[0.135]
Age 12	1.293	1.237	1.531	1.18	1.206	1.148	1.209	1.08
	[0.118]**	[0.216]**	[0.234]**	[0.086]**	[0.120]**	[0.124]**	[0.121]**	[0.125]**
Age 13	2.049	1.952	2.657	1.866	1.714	2.015	1.719	1.941
	[0.148]**	[0.278]**	[0.293]**	[0.096]**	[0.132]**	[0.141]**	[0.127]**	[0.134]**
Female	0.247			0.213				
	[0.090]**			[0.063]**				
Mother < age 23 at birth	-0.195	0.128	-0.31	0.071	0.022	0.11	0.024	0.093
	[0.202]	[0.356]	[0.501]	[0.070]	[0.096]	[0.103]	[0.095]	[0.101]
Number same sex siblings				0.208	0.167	0.25	0.167	0.258
				[0.076]**	[0.104]	[0.113]*	[0.104]	[0.113]*
Fixed effects	*House-hold*	*House-hold*	*House-hold*	*District*	*District*	*District*	*District*	*District*
Observations	*2251*	*1154*	*1097*	*2251*	*1154*	*1097*	*1154*	*1097*

Notes: Data from the 2000 Tanzanian Household Budget Survey, sample restricted to children ages 10-13 in 27 districts targetted for iodized oil capsule (IOC) distribution between 1986 and 1995. In columns 1-6, IOC in utero is equal to 1 or 0.5 (depending on mother's age at birth) if a child was born 1-3 years after IOC was distributed in the district; in columns 7-8, IOC in utero is equal to 1 or 0.5 (depending on mother's age at birth) if a child was born 1-2 years after IOC was distributed in the district. All regressions control for birth order and sex-specific birth order. * significant at 5%; ** significant at 1%

Field, Robles, and Torero (2007)

Table 4: Grade Attainment and IOC Supplementation in Utero, 2004

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Grade attainment, ages 10-14			*Grade attainment, ages 10-12*			*Enter secondary school, ages 10-14*		
	All	Girls	Boys	All	Girls	Boys	All	Girls	Boys
IOC in utero	0.487	0.870	0.403	0.189	1.604	0.276	0.081	0.139	0.096
	[0.184]**	[0.424]*	[0.331]	[0.150]	[0.606]*	[0.662]	[0.031]**	[0.081]+	[0.058]+
Age 11	0.317	0.302	0.405	-0.191	-0.593	0.585	-0.016	-0.045	0.02
	[0.108]**	[0.247]	[0.195]*	[0.098]+	[0.454]	[0.420]	[0.018]	[0.047]	[0.034]
Age 12	0.94	0.884	1.135	-0.124	0.644	1.203	-0.001	-0.098	0.065
	[0.124]**	[0.300]**	[0.226]**	[0.086]	[0.547]	[0.441]**	[0.021]	[0.058]+	[0.040]
Age 13	1.349	1.682	1.267				0.035	-0.03	0.102
	[0.166]**	[0.428]**	[0.307]**				[0.028]	[0.082]	[0.054]+
Age 14	2.036	2.185	2.152				0.117	0.005	0.234
	[0.202]**	[0.507]**	[0.375]**				[0.034]**	[0.097]	[0.066]**
Month of birth	-0.027	-0.051	-0.024	0.003	-0.030	0.018	-0.002	-0.005	-0.004
	[0.010]**	[0.023]*	[0.017]	[0.009]	[0.031]	[0.029]	[0.002]	[0.004]	[0.003]
Female	0.352			0.309			0.015		
	[0.062]**			[0.060]**			[0.010]		
Household fixed effects	yes	yes	yes	yes	yes	yes	yes	yes	yes
Observations	*3672*	*1797*	*1875*	*4984*	*1147*	*1178*	*3672*	*1797*	*1875*

Notes: Data from the 2004 Tanzanian Demographic and Health Survey, sample restricted to children ages 10-14. IOC in utero is equal to (birth month/12) if a child was born 1-3 years after IOC was distributed in the district. All regressions control for dummy indicators of birth order and sex-specific birth order. + significant at 10%; * significant at 5%; ** significant at 1%.

Field, Robles, and Torero (2007)

Table 6: Variation in Effect on Schooling of IOC Supplementation in Utero

	Rate of Cassava Consumption in District			*Amount of IOC*	
	High (0.41-0.62)	Medium (0.10-0.40)	Low (< 0.10)	Mother>22 at birth (380 mg)	Mother<23 at birth (200 mg)
IOC in utero	0.046	0.508	-0.02	0.431	0.066
(IOC in utero =born 1-3 years after program)	(0.391)	(0.165)**	(0.252)	(0.198)*	(0.199)
Female	0.417	0.172	0.154	0.252	0.304
	(0.188)*	(0.146)	(0.148)	(0.146)	(0.156)
Age 11	0.68	0.46	0.451	0.783	0.401
	(0.281)*	(0.222)*	(0.245)	(0.225)**	(0.242)
Age 12	1.64	1.224	0.967	1.524	1.232
	(0.241)**	(0.179)**	(0.205)**	(0.195)**	(0.210)**
Age 13	2.086	2.051	1.724	2.27	1.684
	(0.298)**	(0.232)**	(0.295)**	(0.253)**	(0.263)**
Household fixed effects	yes	yes	yes	yes	yes
Observations	*669*	*804*	*778*	*983*	*799*

Notes: Data from the 2000 Tanzanian Household Budget Survey, sample restricted to children ages 10-13 in 27 districts targetted for iodized oil capsule (IOC) distribution between 1986 and 1995. Children and women below age 23 were given IOC containing 200mg of iodine and women over 22 were given IOC containing 380 mg of iodine. In all regressions, IOC in utero is equal to one if a child was born 1-3 years after IOC was distributed in the district. Regressions also control for birth order and sex-specific birth order. Rate of cassava consumption defined as fraction of THBS households in district that report growing cassava.