Aggregate Conditions, Child Growth & the DHS or Linked Data, Disciplinary Expertise & Statistical Delusions

Joseph R. Cummins
University of California, Riverside
AAEA Virtual Workshop

Friendly Fire: UC Davis Edition

- Tom Beatty told us it is now much harder to publish in Econ Journals just using public data
- But I try a lot and sometimes succeed!
- So – 3 Things in 20 Minutes:
  - Publishing in Econ using Public Data
  - The Demographic and Health Surveys
  - Determinants of Child Growth Faltering

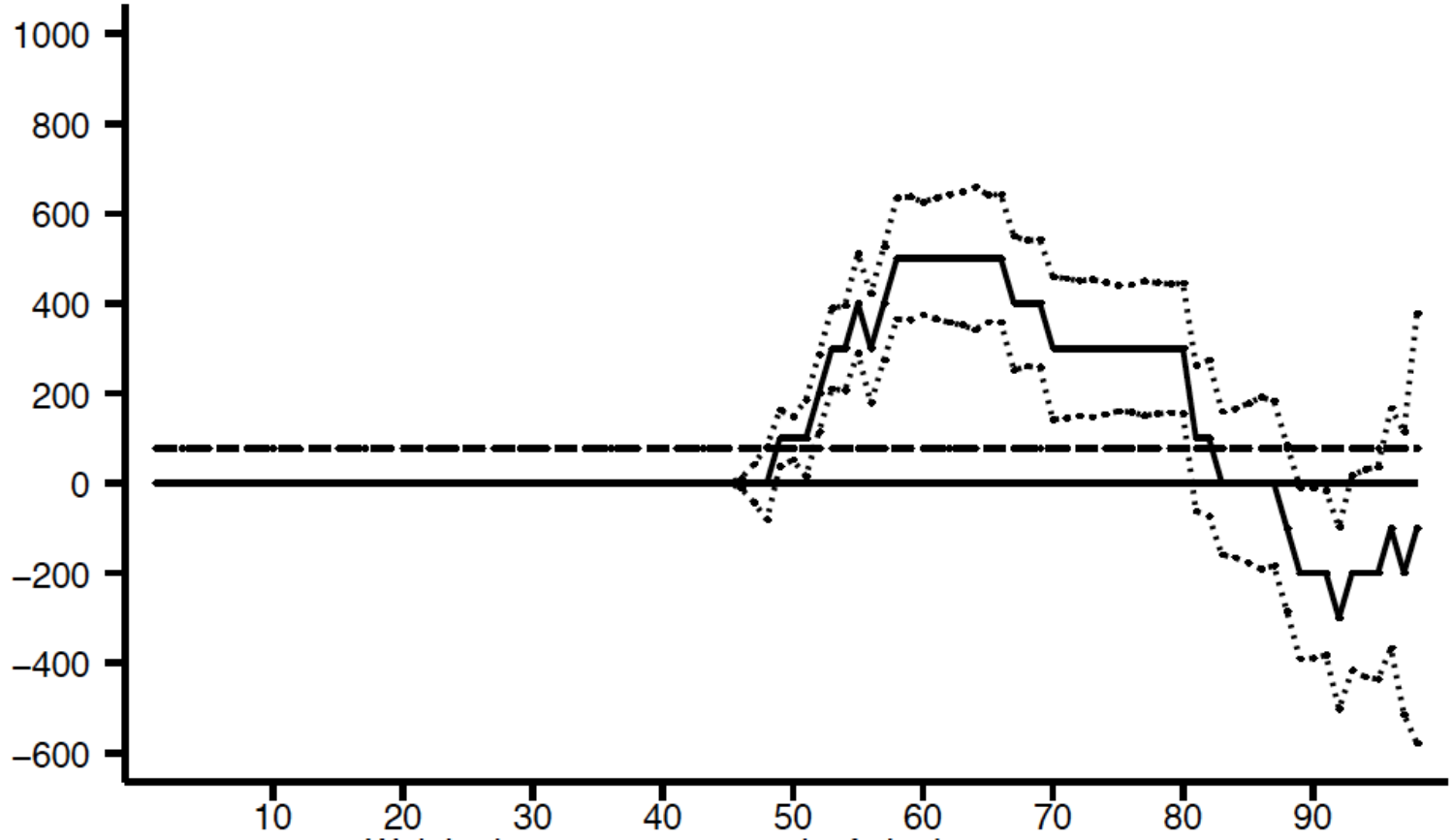
Part 1: Two Things You Need to Publish Well With Public Survey Data

- Document a Really Interesting Pattern:
  - N.b. - Tens of millions of women have exited the labor market in India over the last decade or two…. Why?

- Bring In More Data:
  - A lot of things you can do on one DHS you can do on most DHS. And you can do it simultaneously.
  - A lot of things about the world have been measured for a long time. And you can match them to people.

The Problem of Finding Something New and Interesting….

- You have to have a new perspective to see something new:



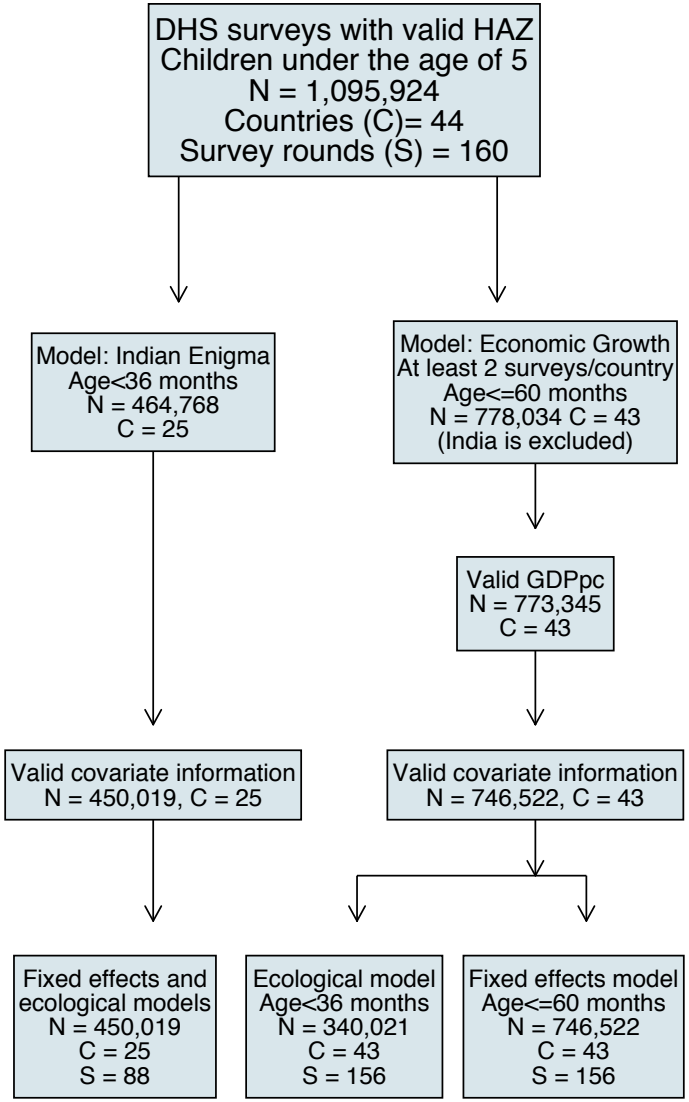
The Problem of Adding More Data…

- Often this Requires Multiple Fields of Expertise:
  - X is not in field: Environmental (Pollution Exposure)
  - Y is not in field: Health (Child Growth)

- Failing to Embrace Expertise Can Be Problematic
  - You can misunderstand magnitudes
  - You can report (physically, biologically) impossible things
  - You can be lead down blind alleys
  - You won't learn as much.

Part 2: Data Linkages in the DHS

- 1. The DHS
- 2. Appending
- 3. Merging



Adapted from: www.consort-statement.org/consort-statement/flow-diagram

How to Create: Appending and Merging

- Appending: adding rows
  - stacking multiple data collection rounds (panel or repeated cross-section)
- Merging: adding columns
  - Adding information to individuals or groups of people that was not in the orginal data via some form of matching
  - E.g.: geography (space) or survey year (time) or exposure timing (cohort)

- You don't even have to Append anymore. ¡Thanks iPums!

Appending

- Increased Sample Size
  - Smaller standard errors/confidence intervals
- Represents a larger population of people
  - More countries
- Extends over longer period of time
  - More years of data and cohorts of people

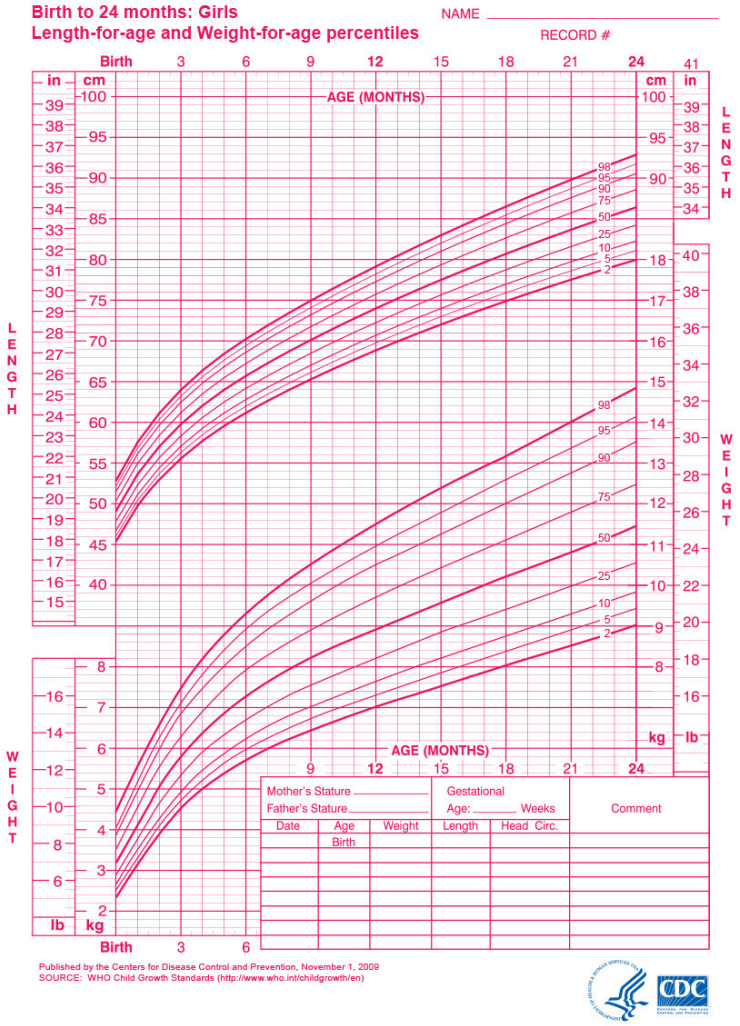
Merging

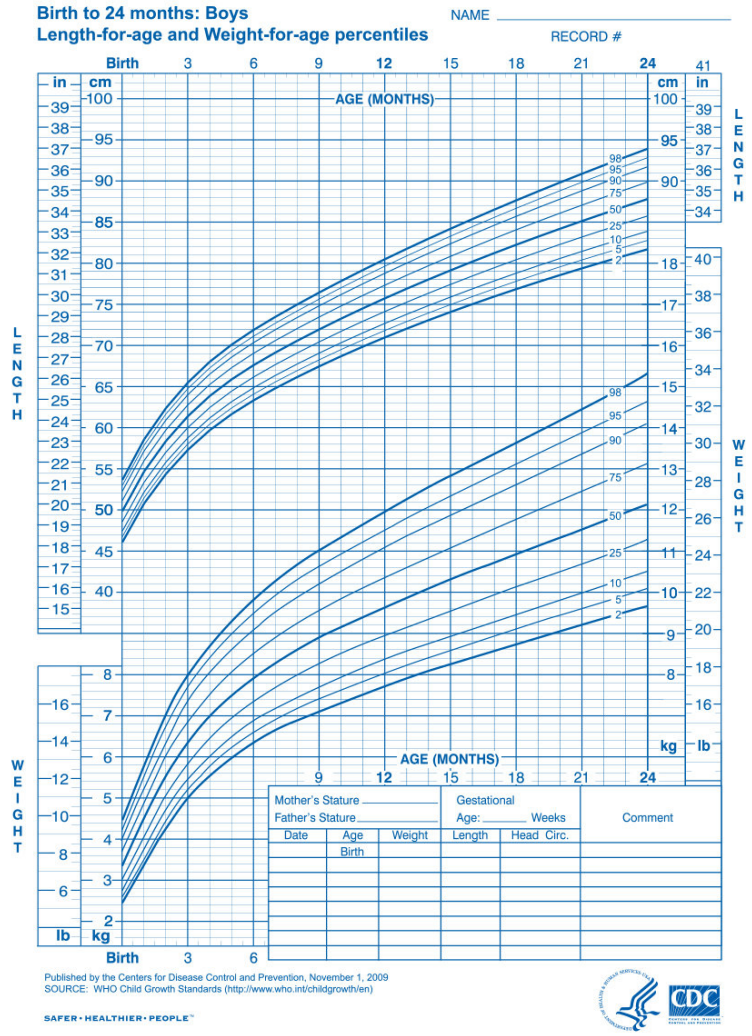
- Geography:
  - Your Country, City, Neighborhood, Address
- Calendar Time:
  - Something happened in year Y
- Cohort Time:
  - Some condition when a person was age A

Part 3: Know Your (New) Data

- Height-for-Age Z-score
- Example 1: GDP Growth and Child HAZ
- Example 2: The Indian Enigma
- Example 3: Anaka Aiyar on Examples 1 & 2

Know Your Y


Birth to 24 months: Girls
Length-for-age and Weight-for-age percentiles
NAME
RECORD #
AGE (MONTHS)
LENGTH
WEIGHT
Mother's Stature
Father's Stature
Gestational Age: Weeks
Date
Age
Weight
Length
Head Circ.
Comment
Birth
Published by the Centers for Disease Control and Prevention, November 1, 2009
SOURCE: WHO Child Growth Standards (http://www.who.int/childgrowth/en)


Birth to 24 months: Boys
Length-for-age and Weight-for-age percentiles
NAME
RECORD #
AGE (MONTHS)
LENGTH
WEIGHT
Mother's Stature
Father's Stature
Gestational Age: Weeks
Date
Age
Weight
Length
Head Circ.
Comment
Birth
Published by the Centers for Disease Control and Prevention, November 1, 2009
SOURCE: WHO Child Growth Standards (http://www.who.int/childgrowth/en)
SAFER · HEALTHIER · PEOPLE™

Vollmer et al. 2014, Lancet Global Health

Weak to null association between GDP per capita and child stunting rates

	Adjusted			Unadjusted		
	Stunted	Wasted	Underweight	Stunted	Wasted	Underweight
Full sample						
OR (95% CI)	0·996 (0·993–1·000)	0·983 (0·979–0·986)	0·989 (0·985–0·992)	0·993 (0·989–0·995)	0·984 (0·981–0·986)	0·986 (0·982–0·990)
p value	0·021	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001
N	462 854	459 538	485 152	462 854	459 538	485 152
Poorest wealth quintile						
OR (95% CI)	0·997 (0·990–1·004)	0·991 (0·978–1·004)	0·999 (0·991–1·008)	0·995 (0·992–0·998)	0·985 (0·982–0·987)	0·988 (0·983–0·993)
p value	0·367	0·153	0·784	0·002	<0·0001	<0·0001
N	104 040	103 473	109 329	104 040	103 473	109 329
Richest wealth quintile						
OR (95% CI)	0·997 (0·992–1·001)	0·984 (0·981–0·987)	0·990 (0·987–0·993)	0·990 (0·987–0·993)	0·983 (0·981–0·986)	0·985 (0·980–0·989)
p value	0·086	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001
N	74 575	73 902	77 883	74 575	73 907	77 888
Children aged 0–11 months						
OR (95% CI)	0·991 (0·981–1·001)	0·998 (0·988–1·008)	0·994 (0·987–1·001)	0·989 (0·985–0·993)	0·985 (0·982–0·989)	0·987 (0·984–0·989)
p value	0·071	0·593	0·054	<0·0001	<0·0001	<0·0001
N	162 048	158 770	170 633	162 048	158 770	170 633
Children aged 12–23 months						
OR (95% CI)	0·989 (0·979–1·000)	0·989 (0·977–1·002)	0·996 (0·985–1·006)	0·993 (0·990–0·997)	0·983 (0·981–0·986)	0·985 (0·979–0·991)
p value	0·035	0·085	0·372	0·0002	<0·0001	<0·0001
N	155 071	155 437	162 378	155 071	155 437	162 378
Children 24–35 months						
OR (95% CI)	0·997 (0·994–1·000)	0·982 (0·979–0·986)	0·988 (0·984–0·992)	0·995 (0·993–0·997)	0·982 (0·979–0·985)	0·986 (0·982–0·990)
p value	0·008	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001
N	145 734	145 330	152 140	145 735	145 331	152 141
Sub-Saharan Africa						
OR (95% CI)	0·996 (0·994–0·997)	0·984 (0·981–0·986)	0·992 (0·990–0·994)	0·995 (0·994–0·996)	0·988 (0·986–0·991)	0·993 (0·992–0·995)
p value	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001
N	241 448	239 546	250 507	241 448	239 547	250 508
Asia						
OR (95% CI)	0·999 (0·999–1·000)	0·979 (0·978–0·979)	0·984 (0·984–0·985)	0·992 (0·992–0·992)	0·978 (0·978–0·978)	0·980 (0·980–0·980)
p value	0·0002	<0·0001	<0·0001	<0·0001	<0·0001	<0·0001
N	112 342	111 587	123 619	112 342	111 587	123 619

(Table 2 continues on next page)

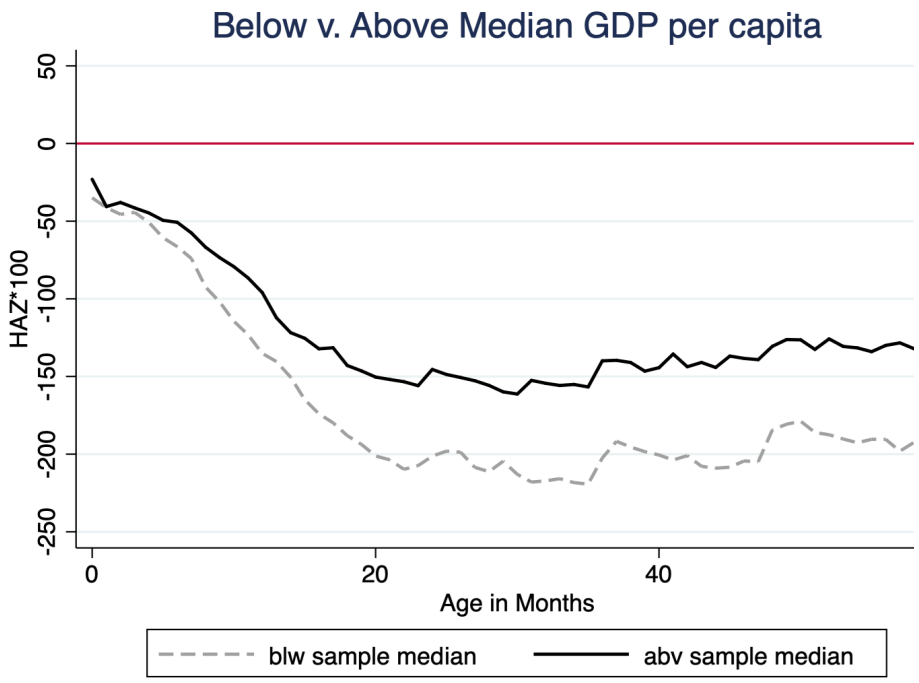
Jayachandran & Pande, 2014 American Economic Review

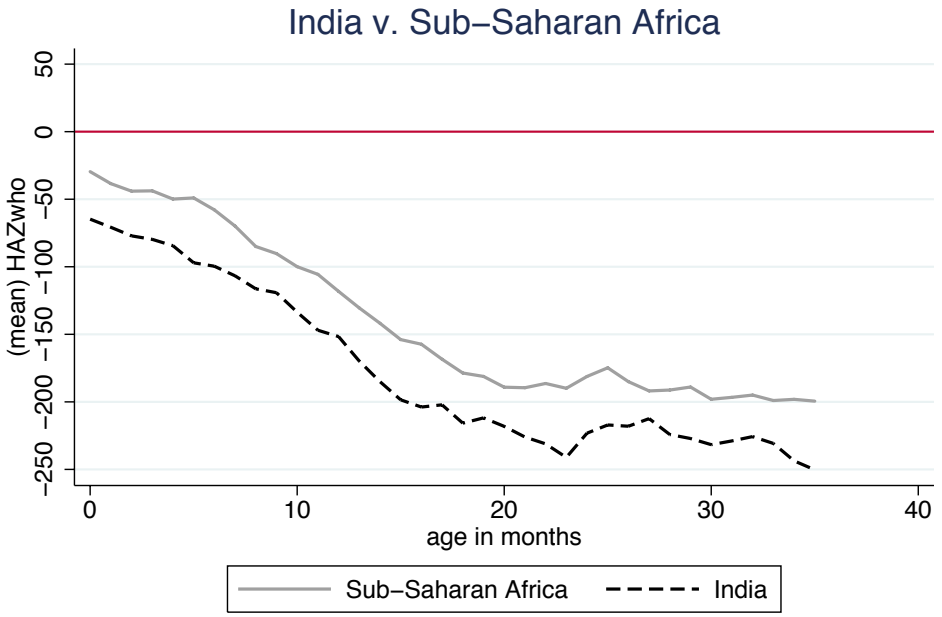
- Indian Enigma
- Partly result of Son Preference in India

TABLE 2—INDIA'S DIFFERENTIAL BIRTH ORDER GRADIENT IN CHILD HEIGHT AND RELATED OUTCOMES

	HFA z-score					Stunted	WFA z-score	Hb level	Deceased
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
India	−0.082	0.092							
	[0.011]	[0.018]							
India × 2nd child		−0.144	−0.161	−0.110	−0.243	0.051	−0.146	−0.094	0.003
		[0.025]	[0.027]	[0.063]	[0.048]	[0.007]	[0.020]	[0.030]	[0.004]
India × 3rd+ child		−0.377	−0.227	−0.193	−0.436	0.064	−0.198	−0.159	0.002
		[0.024]	[0.032]	[0.092]	[0.085]	[0.009]	[0.024]	[0.036]	[0.004]
2nd child		0.023	−0.011	−0.097	−0.167	0.009	0.009	−0.011	−0.014
		[0.015]	[0.017]	[0.053]	[0.027]	[0.004]	[0.012]	[0.022]	[0.002]
3rd+ child		−0.066	−0.118	−0.169	−0.334	0.036	−0.063	−0.037	−0.011
		[0.013]	[0.019]	[0.074]	[0.044]	[0.005]	[0.014]	[0.025]	[0.003]
Africa mean of outcome	−1.351	−1.351	−1.351	−1.351	−1.351	0.375	−0.877	10.150	0.071
Child's age dummies × India	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Mother's literacy × India	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Mother's age at birth × India	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes
PSU fixed effects	No	No	Yes	Yes	No	Yes	Yes	Yes	Yes
Mother fixed effects	No	No	No	No	Yes	No	No	No	No
Completed fertility sample	No	No	No	Yes	No	No	No	No	No
Observations	168,108	168,108	167,737	66,566	83,228	167,737	167,737	88,838	199,514

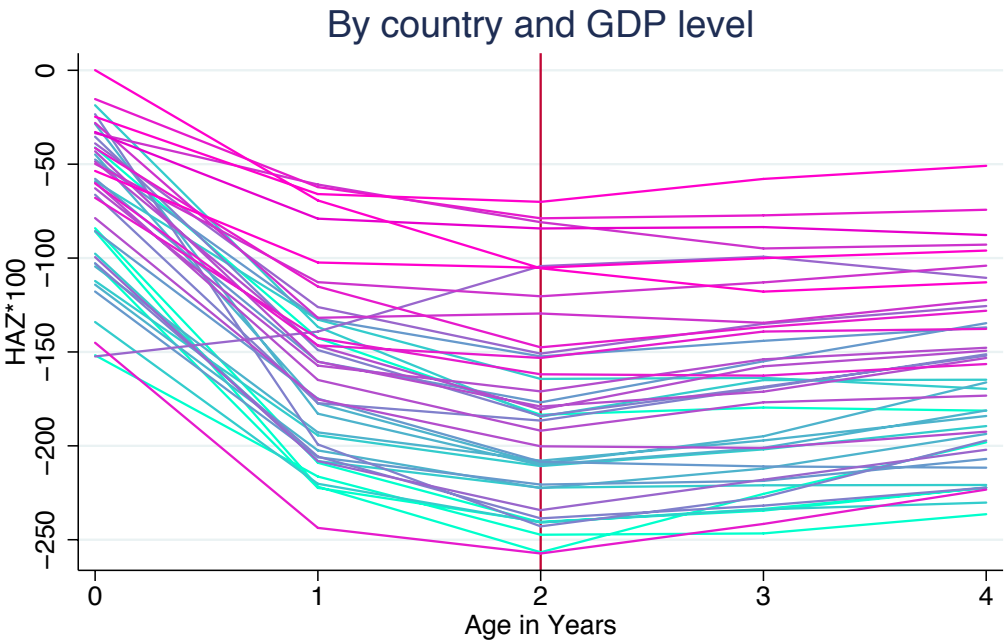
Aiyar & Cummins, 2020
Journal of Development Economics



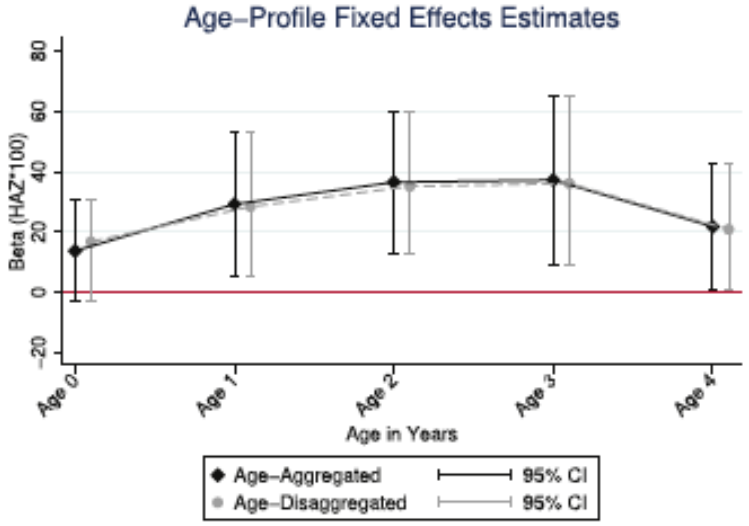


Each survey round given equal weight.
All countries in the sample are developing countries

Figure GDP: Data + Model



Weighted mean HAZ by age−in−years
Cyan to Magenta moves from lowest decile (Cyan) to highest decile (Magenta)
of within−sample GDP/cap averaged over the sample period.



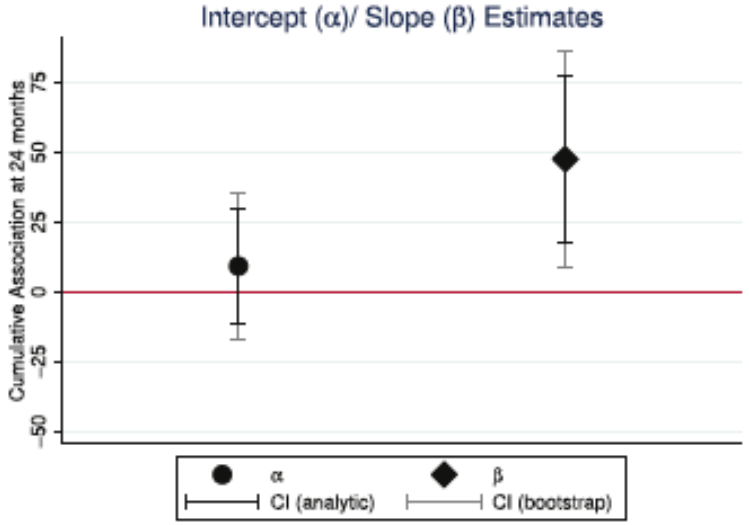
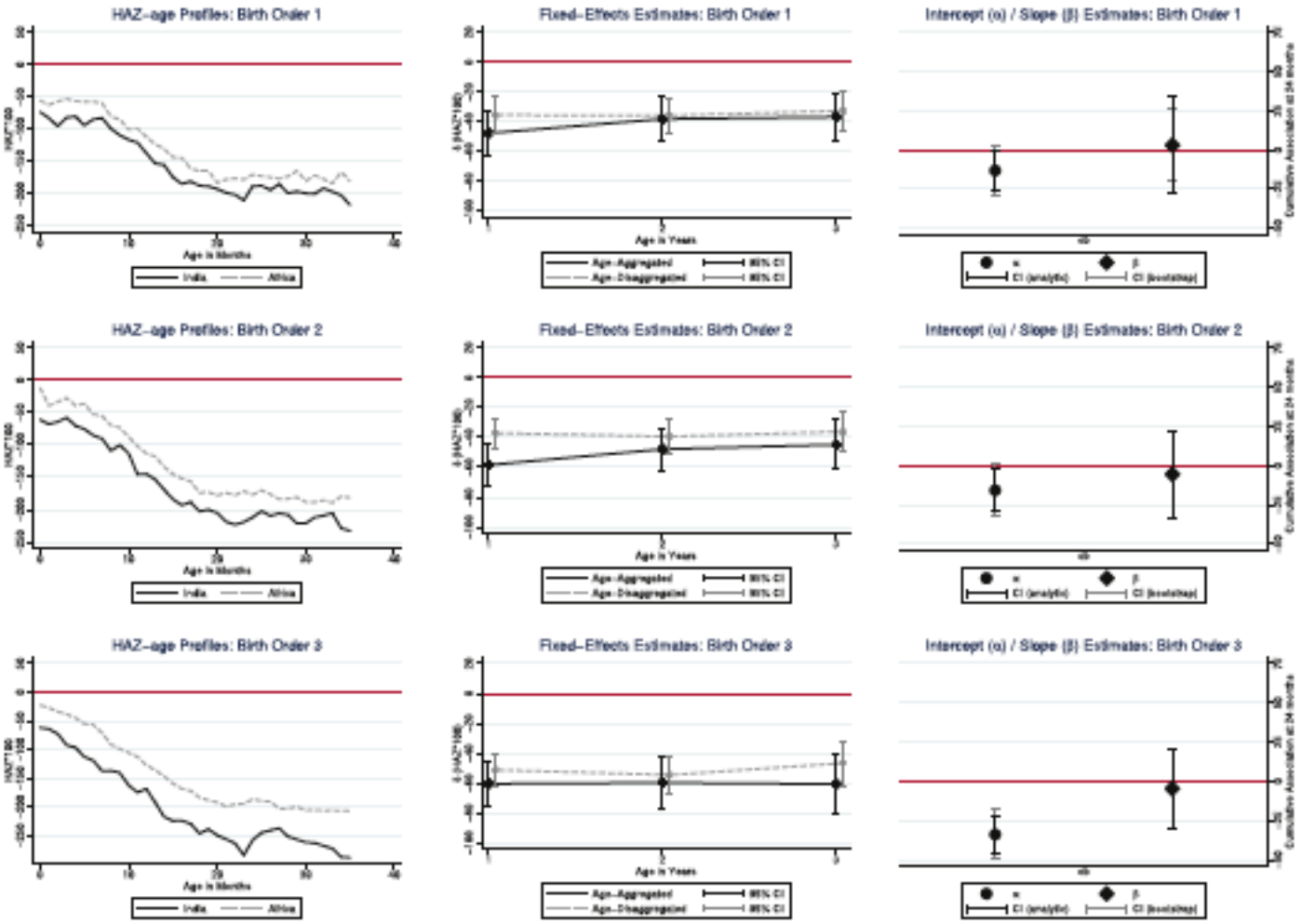


Figure Indian Enigma: Extra Dimensions of Model + Data



Conclusion: (My) Lessons

- GDP Estimates Were Needlessly Imprecise
  - Personal Takeaway: I think improved incomes lead to improvements in the stream of health inputs to children
- Indian Enigma Was Under-Explained
  - Personal Takeaway: I think it's largely about the moms and their early life streams of health inputs.