Consequences of Historical Adaptation for Undernutrition and Diabetes in Developing Countries

Nancy Luke
Pennsylvania State University

Kaivan Munshi
Yale University

Anu Oommen
Christian Medical College, Vellore

Swapnil Singh
Bank of Lithuania

November 2022
Motivation

• Two stylized facts motivate this research
  1. The absence of a clear link between nutritional status and income in developing countries (Deaton, 2007; Swaminathan et al., 2019)
  2. The increased prevalence of diabetes among normal weight individuals in these countries (Narayan, 2017)
• Our primary objective is to develop and test a model with three ingredients – adaptation, mismatch, and a set point – that can explain these seemingly unrelated observations
Adaptation

- Two models of early-life adaptation or developmental plasticity have been proposed by evolutionary biologists
  1. The developmental constraints model in which developing organisms in severely resource limited environments make immediate tradeoffs to protect critical functions and improve survival in early life (Barker, 1995)
  2. The predictive model in which maternal cues *in utero* predict the (normal) adult environment and the organism evolves accordingly in anticipation of future conditions (Gluckman and Hanson, 2006)
Adaptation

• The predictive adaptive model will be relevant when conditions in past generations are an accurate predictor of (average) conditions in the current generation (Burgess and Marshall, 2014; Lind et al., 2020)
  • This is likely to have been the case in the pre-modern economy, which was characterized by wide short-term fluctuations in food supply but had growth rates close to zero for centuries
• The pre-modern population would thus have been adapted physiologically to long-term (low) food supply, with the adaptation varying across space with agroclimatic conditions (Pomeroy et al., 2019)
Mismatch

• With economic development, there is a substantial increase in income, and, with it, consumption.

• It has been hypothesized in the developmental origins of adult disease literature that the mismatch between current and pre-modern consumption (to which the population is adapted) has contributed to the high rates of diabetes in developing countries; e.g. Gluckman and Hanson, 2004; Wells et al., 2016; Narayan, 2017.

• We place additional restrictions on the relationship between diabetes and the mismatch, while simultaneously explaining the persistence of undernutrition in these countries by tying the initial adaptation to a set point.
Many individuals have a stable set point for their bodyweight throughout adult life (Muller et al., 2010).

The set point is part of a homeostatic (stabilizing) system that maintains the body's energy balance against fluctuations in food intake by making metabolic and hormonal adjustments.

We posit that the set point for a given dynasty (family) is determined by food supply in the pre-modern economy.

While the adapted set point would have allowed pre-modern populations to maintain their energy balance and to survive and reproduce in an environment characterized by low and fluctuating food supply, it becomes a liability if it persists for multiple generations after the onset of economic development.
A property of all – physical and biological – homeostatic systems is that they can only self-regulate within fixed bounds and will malfunction when the environmental stresses to which they are subjected exceed a threshold level (Kultz, 2020).

This implies that as long as current and pre-modern (ancestral) consumption or, equivalently, income remain sufficiently close to each other, the body will successfully defend its bodyweight set point.

Once the gap between current and pre-modern income crosses a threshold, however, the body will no longer be able to defend the set point.

Escape from the set point is associated with imbalance in energy regulation, which will be accompanied by imbalance in related (inter-linked) homeostatic systems.

- Failure of glucose homeostasis, in particular, manifests as diabetes.
Implications

• Although it may be appropriate to characterize the set point with respect to weight for a given individual, we account for possible variation in height across generations by specifying a common set point for members of a dynasty with respect to their BMI; i.e. weight conditional on height

• This normalization is especially useful for our analysis because BMI is a standard measure of nutritional status and is also associated with the risk of diabetes

• It follows that there will be two types of individuals in a developing economy:

1. Those who remain at their set point, despite the increase in their consumption, are partly responsible for the weak association between nutritional status, which we measure by BMI, and income

2. Those who have escaped their set point, but are not necessarily overweight, are the primary contributors to the increased prevalence of diabetes and related metabolic disorders
Extensions

- The partition of the population just described is not permanent.
- Many models of developmental plasticity assume that the initial adaptation is epigenetic (changing gene expression) and, hence, will persist for a limited number of generations (Jablonka and Raz, 2009; Lind and Spagopoulou, 2018).
- This explains why European populations, which were also under-nourished historically, no longer display these traits.
- It also explains the multi-generational health experience of migrants from developing countries to advanced economies.
Analysis

- Although there is theoretical and empirical support for each ingredient in our argument – adaptation, mismatch and a set point – the specific biological mechanism we propose has not been tested.
- While we plan to examine the biological foundations in the future with an animal model, we make independent progress in the current research by testing the implications of the argument itself.
• If data on income, BMI, and diabetes were available for each household (dynasty) over many generations, going back to the pre-modern period, then:
  • For a given dynasty, we would expect to observe a discrete increase in BMI in a particular generation (in which the gap between current and pre-modern income exceeded the threshold) with an accompanying increase in the risk of diabetes

• In the absence of such data,
  1. We model the evolution of (permanent) income in the population to derive implications of the model with respect to current income
  2. We test these implications with representative data from India, rule out alternative explanations, and internally validate the model
  3. We examine the external validity of the model with data from many countries
  4. Independent measures of pre-modern income are constructed to validate the specified biological mechanism
The Model
The Income Process

- The population consists of a large number of infinitely lived dynasties.
- Each dynasty consists of a single individual in each generation, who is replaced by a single descendant in the next generation.
- There is a fixed return on wealth in each generation, i.e., income flow which is consumed, so that the stock is passed on (without depletion) to the next generation.
- We thus use the terms wealth and (permanent) income interchangeably.
The Income Process

- Income is the same in each generation in the pre-modern era, during which epigenetic adaptation takes place, but subsequently evolves.
- Denote the logarithm of the dynasty’s initial income by $y_0$.
- In each subsequent period or generation the dynasty receives an additive permanent income shock, which is normally distributed.
  - Permanent income in the modern economy is well approximated by the log-normal distribution (Battistin et al., 2009).
- Solving recursively, $y_t = y_0 + U_t$, where $U_t$ measures the accumulation of income shocks and is also normally distributed.
There is a positive and continuous relationship between (food) consumption and income in all time periods.

Nutritional status, which we measure by BMI, is increasing continuously in income (consumption) in the initial period.

- A dynasty’s initial BMI: \( z_0 = a + by_0 \), determines its set point.

In subsequent periods, there is a set-point threshold \( \alpha \) such that

\[
z_t = \begin{cases} 
  a + by_0 & \text{if } U_t \leq \alpha \\
  a + by_t & \text{if } U_t > \alpha 
\end{cases}
\]

Notice that there is no lower threshold; the assumption is that dynasties do not regress with regard to nutritional status during the process of development.
Biological Relationships (Diabetes - Income)

- As long as individuals remain at their set point, variation in income (consumption) has no effect on diabetes.
- Once they escape the set point, the risk of diabetes is increasing in the income mismatch.

\[
P(D_t) = \begin{cases} 
\gamma_1 & \text{if } U_t \leq \alpha \\
\gamma_1 + \gamma_2(y_t - y_0) & \text{if } U_t > \alpha
\end{cases}
\]
BMI Dynamics (dynasty level)

\[ y_t = y_0 + \alpha \]

\[ a + b(y_0 + \alpha) \]

\[ a + by_0 \]

**Note:** this describes the special case of a dynasty that only receives positive shocks.
Cross-Sectional BMI-Income Relationship

- We normalize so that the initial income distribution is bounded below at zero.
- Given that $y_0 \geq 0$ and the assumption that there is no lower threshold, all individuals with $y_t \leq \alpha$ must lie within their set point thresholds.
- Some individuals with $y_t > \alpha$ will have crossed their threshold, while others will remain at their set point.
- Simplifying assumptions, validated below, together with the properties of the normal distribution, allow us to derive the following result:

Model Equations
Cross-Sectional Relationships

- BMI without set point
- BMI with set point
- Probability of disease

Mean BMI vs. current income, $y_t$
Alternative Specifications of the Set Point

• Our model describes the initial phase of economic development in which the set point, $z_0$, is assumed to be fixed

• Alternative specifications of the set point would allow it to vary
  • Weighted average of $y_0$ and $y_t$, with the weight on $y_t$ increasing over time
  • Determined by initial conditions in each generation as in the developmental constraints model, i.e. $y_{t-1}$

• While the alternative models generate the same qualitative predictions, the distinguishing feature of our model is that BMI below the threshold is determined exclusively by $y_0$
Diabetes - BMI Relationship

• The model implies that
  1. BMI is increasing with income at all levels, more steeply above a threshold
  2. Diabetes is increasing in income above the same threshold
• Bringing the two implications together
  • There is no association between the risk of diabetes and BMI up to a BMI threshold, and a positive association thereafter
Testing the Model
Cross-sectional Analysis

- The core dataset that we use to test the model is the India Human Development Survey (IHDS)
- Income is measured at the household level, as the average over the 2004-2005 and 2011-2012 rounds
- Nutritional status for adults, in each round, is measured by BMI
- Diabetes is measured by a composite variable, which we refer to as “metabolic disease,” that indicates whether the household head and his spouse have been diagnosed with diabetes, hypertension, or cardiovascular disease in each survey round
- The following covariates are partialled out prior to nonparametric estimation of the BMI-income and metabolic disease-income relationships:
  - age (linear, quadratic and cubic terms) and dummies for gender, caste group, rural area, district and survey-round
BMI and Metabolic Disease with respect to Income

(a) Nonparametric relationship

(b) Threshold test

Source: India Human Development Survey (IHDS)
### Piecewise Linear Equation Estimates

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BMI (1)</th>
<th>metabolic disease (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline slope ($\beta_1$)</td>
<td>0.239** (0.057)</td>
<td>0.002 (0.002)</td>
</tr>
<tr>
<td>Slope change ($\beta_2$)</td>
<td>0.940** (0.066)</td>
<td>0.028** (0.003)</td>
</tr>
<tr>
<td>Threshold location ($\tau$)</td>
<td>1.65 [1.55, 1.75]</td>
<td>1.90 [1.80, 2.05]</td>
</tr>
<tr>
<td>Threshold test $p$–value</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Mean of dependent variable</td>
<td>22.002</td>
<td>0.074</td>
</tr>
<tr>
<td>N</td>
<td>76,949</td>
<td>148,928</td>
</tr>
</tbody>
</table>
Robustness Tests

1. Include measures of household composition as additional covariates in the estimating equation and construct a nonparametric shift-share instrument for current income.

2. Separate men and women.

3. Separately examine the components of BMI (height, weight) and metabolic disease (diabetes, hypertension, cardiovascular disease) with respect to income.
Alternative Explanations
Alternative Explanations

• Additional covariates in the estimating equations are included to account for independent determinants of nutritional status and metabolic disease, but such controls may not be complete

• Any alternative mechanism must first explain the discontinuous association between BMI and income
  • We verify that two important proximate determinants of nutritional status in developing countries: nutrient intake and children’s illness do not satisfy this requirement
  • However, selective child mortality (Deaton, 2007) or specific poverty trap models (Dasgupta and Ray, 1987) could generate a discontinuity
Alternative Explanations

- The second requirement for any alternative mechanism is that it must explain why
  1. BMI and metabolic disease do not initially track together with respect to income
  2. Why both outcomes increase discontinuously at a particular income threshold
  3. Why a marginal increase in BMI, starting from a base level below 22, should be associated with an increase in the risk of metabolic disease

- In our model, BMI and the risk of diabetes change simultaneously because they are independently impacted by the failure of an underlying homeostatic system, which is specific to developing country populations
  - The tests of external validity and the biological mechanism provide additional support for the model
Internal Validity
Structural Estimation

\[ \text{mean BMI} = b \times \text{current income}, y_t \]

\[ \alpha \]

\[ b e^L(y_t) \]

\[ b e^H(y_t) \]

BMI without set point

BMI with set point
Structural Estimation

• The model implies the following cross-sectional relationships, below and above the threshold

\[
E(z_t | y_t) = a + b(y_t - e^L(y_t)) \\
E(z_t | y_t) = a + b(y_t - e^H(y_t))
\]

• Closed-form solutions for \(e^L(y_t), e^H(y_t)\) are derived as functions of the model’s parameters
  • \(\alpha\) can be estimated from the location of the threshold and \(\mu_t, \sigma^2_t\) are obtained from the World Inequality Database (Chancel and Piketty, 2017)
  • If we replace \(y_t\) with \(y_t - e^L(y_t)\) below the threshold and \(y_t - e^H(y_t)\) above the threshold, then the slope of the BMI-income relationship will be \(b\)

• This test validates the structure we have imposed on the model and the simplifying assumption that are needed to derive \(e^L(y_t), e^H(y_t)\)
## Structural Parameter Estimates

<table>
<thead>
<tr>
<th>Dep. variable:</th>
<th>adult BMI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specification:</td>
<td>without adjustment</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
</tr>
<tr>
<td>Slope below threshold ($\beta_L$)</td>
<td>0.223*** (0.048)</td>
</tr>
<tr>
<td>Slope above threshold ($\beta_H$)</td>
<td>1.140*** (0.035)</td>
</tr>
<tr>
<td>$F$–statistic ($\beta_L = \beta_H$)</td>
<td>234.45 [0.000]</td>
</tr>
<tr>
<td>Imposed threshold</td>
<td>1.65</td>
</tr>
<tr>
<td>N</td>
<td>76,949</td>
</tr>
</tbody>
</table>

Source: India Human Development Survey (IHDS)
Counter-factual Simulations

Sensitivity of slope coefficients

![Graph showing data, predicted, and counter-factual lines for BMI vs log household income.](image-url)
External Validity
Tests With Respect to Income

- The presence of a set point is not unique to India
- To test the external validity of the model, data that are comparable to IHDS are required
  - Two datasets that satisfy the requirements are the Indonesia Family Life Survey (IFLS) and the Ghana Socioeconomic Panel Study (GSPS)
- Results for other countries will depend on their stage in the development process
  - In particular, the gap between current and pre-modern income
  - We measure pre-modern income by adult height in 1900; height is increasing continuously with income in the pre-modern period in our model and as noted by Deaton (2007) genes are important determinants of individual heights but cannot explain variation across populations
Historical Height and Current Income

(a) GDP 2010.

(b) GDP 1960 and 2010.

Source: Penn World Table and NCD-RisC
BMI and Metabolic Disease (Indonesia and Ghana)

(a) Indonesia: BMI and Disease

(b) Ghana: BMI

Source: Indonesia Family Life Survey (IFLS), Ghana Socioeconomic Panel Survey (GSPS)
• Many datasets, including DHS and WHO-STEPS, provide information on BMI and diabetes (with biomarkers)
• Based on the model, there should be no association between diabetes and BMI up to a threshold and a positive association thereafter
Tests With Respect to BMI

(a) India

(b) Asia and Sub-Saharan Africa

Source: India Human Development Survey (IHDS), Indonesia Family Life Survey (IFLS), Demographic and Health Survey (DHS), WHO STEPS
Cross-Country Analysis

- Assume that African populations are largely at their (higher) set points, whereas some fraction of Asian populations have escaped.
- Consider an African country and an Asian country with the same current income.
- African BMI is determined by ancestral income, whereas Asian BMI is a weighted average of (lower) ancestral income and (higher) current income.
  - If a sufficient fraction of the Asian population remains at its set point, then BMI conditional on current income will be higher in Africa.
- Next consider an African country and an Asian country with the same BMI.
  - Based on the preceding figure, the Asian country will have a higher prevalence of diabetes.
Diabetes and BMI

(a) BMI - current income

(b) Diabetes - BMI

Source: NCD-RisC and Penn World Table 9.0
The Mechanism

- The model is based on two biological relationships:
  1. BMI is determined by ancestral income below a threshold and by current income above the threshold
  2. The risk of diabetes is constant below the threshold and increasing in the income mismatch above the threshold
- We now proceed to verify these relationships directly by constructing measures of ancestral income
- The threshold location for this exercise is derived from the cross-sectional tests of the model; recall that households below the current income threshold remain at their set point
- The analysis is restricted to rural households, who would have remained in their place of residence for many generations
District-Level Evidence

• Our first, district-level measure of $y_0$ is based on historical food supply given that agriculture was the dominant activity in the pre-modern economy
• Galor and Özak (2016) convert potential yields from the FAO-GAEZ database to caloric production and then average across crops to construct a Caloric Suitability Index (CSI) which they argue is a good indicator of pre-modern economic development
• We do the same, except that our variant of the CSI for India is based on two staple crops - wheat and rice - that dominated agricultural production
District-Level Evidence

- If the CSI is a good measure of pre-modern wealth, then it should be closely related to historical (pre-modern) population density.
- We verify this hypothesis at the district level by estimating a positive association between population density in 1951, when the Indian economy was just starting to develop, and CSI.
District-Level Evidence

- The preceding result indicates that the positive relationship between CSI and population must be accounted for when constructing measures of per household historical wealth.
- We do this by estimating the following equation:

\[ y_t = f(CSI) + \varepsilon_t \]

- This equation can be compared with the income equation in the model:

\[ y_t = y_0 + U_t \]

- Predicted income corresponds to \( y_0 \) and the residual corresponds to \( U_t \equiv y_t - y_0 \).
- The objective when specifying the \( f(CSI) \) function is to capture that part of the variation in current income that is captured by historical conditions.
- Our preferred measure of \( y_0 \) will thus be predicted household income based on the most flexible nonparametric specification of the \( f(CSI) \) function.
BMI - Income Relationship (below and above the threshold)

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>BMI</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Country:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sample:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ancestral income</td>
<td>0.899***</td>
<td>0.165</td>
</tr>
<tr>
<td>(0.243)</td>
<td>(0.283)</td>
<td>(0.254)</td>
</tr>
<tr>
<td>Current income</td>
<td>0.185***</td>
<td>0.852***</td>
</tr>
<tr>
<td>(0.040)</td>
<td>(0.047)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Threshold location</td>
<td>1.65</td>
<td>1.65</td>
</tr>
<tr>
<td>Dep. var. mean</td>
<td>20.482</td>
<td>21.851</td>
</tr>
<tr>
<td>N</td>
<td>27,164</td>
<td>20,296</td>
</tr>
</tbody>
</table>

Source: India Human Development Survey (IHDS), Indonesia Family Life Survey (IFLS)
## Metabolic Disease - Income Relationship

<table>
<thead>
<tr>
<th>Country:</th>
<th>Dependent variable: Pr(metabolic disease)</th>
</tr>
</thead>
<tbody>
<tr>
<td>India:</td>
<td>Income component:</td>
</tr>
<tr>
<td></td>
<td>income mismatch</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>ancestral income</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

| Income component × 1{current income > \( \tau \)} | 0.018*** | -0.002 | 0.032** | 0.001 |
| (0.004) | (0.002) | (0.011) | (0.008) |

| Joint significance | F-statistic [\( p \)-value] | 14.983 | 1.889 | 13.811 | 0.170 |
|                   | [0.000] | [0.153] | [0.000] | [0.844] |

| Threshold location (\( \tau \)) | 1.90 | 1.90 | 6.00 | 6.00 |

| Dep. var. mean | 0.054 | 0.054 | 0.162 | 0.162 |

| N | 90,879 | 90,879 | 11,001 | 11,001 |

Source: India Human Development Survey (IHDS), Indonesia Family Life Survey (IFLS)
Village Level Evidence

- We use data from the South India Community Health Study (SICHS) which covers a rural population of 1.1 million people in Vellore district, Tamil Nadu for the village-level analysis.
- The SICHS survey covers all variables included in the district-level analysis and is supplemented with:
  - Historical records on the agricultural revenue tax per acre of cultivated land that was collected from each village in 1871.
- Like the CSI, the revenue tax is based on a detailed assessment (by the colonial government) of crop suitability and growing conditions:
  - However, it is (i) defined at the village level, (ii) based explicitly on pre-modern growing conditions, and (iii) provides a direct measure of pre-modern income.
**Village-Level Analysis**

1. Establish that the cross-sectional relationships estimated with nationally representative IHDS data are obtained with SICHS data.

2. Account for patrilocal marriage by documenting that there is positive assortative matching on pre-modern income.

3. Specify pre-modern per household income as $g(R)$ and then construct a measure of $y_0$ as above.

4. Show that adult BMI is determined by pre-modern income below the threshold and by current income above the threshold.
   - Pre-modern income could be correlated with independent factors that determine nutritional status and current income could be measured with error, but neither is associated with a discontinuity.
A model with three ingredients – adaptation, mismatch, and a set point – can explain two seemingly unrelated facts that have been documented in developing countries:

1. the weak association between nutritional status, which we measure by BMI, and income
2. the elevated risk of diabetes among normal weight individuals

The postulated mechanism is validated with micro data from many developing countries.
Conclusion

- Our structural estimates and counter-factual simulations indicate that the fraction of underweight adults in India, who account for 20% of the population, would decline by 24% in the absence of a set point.
  - At the same time, half the Indian population who remain at their set point are protected from diabetes.
- While the health consequences of the set point are thus ambiguous, an increasing fraction of developing-country populations will inevitably escape their set points in the coming decades.
  - Our estimates of the BMI threshold (below 22 for India) indicate that much of the adult population in these countries will need to be screened for diabetes.
  - The flip-side of this finding is that many individuals detected with diabetes and related disorders will have relatively low BMIs; for example, 55% of diabetics in the India DHS have BMI < 23, whereas 9.5% of U.S. diabetics have BMI < 25.
  - They will thus be relatively close to their individual-specific set points, which implies that diabetes reversal programs may be especially effective in these populations.
Evolution of Income in India

GDP per capita is measured in 2011 US dollars.
Nutritional Status and Metabolic Disease with respect to Income (South India)

(a) BMI

(b) Metabolic disease

Source: India Human Development Survey (IHDS), South India Community Health Study (SICHS)
Threshold Tests (adult BMI)

Source: India Human Development Survey (IHDS); South India Community Health Study (SICHS)
### BMI - Income Relationship (below and above the threshold)

Dependent variable: adult BMI

<table>
<thead>
<tr>
<th>$g(R)$ specification:</th>
<th>nonparametric</th>
<th>quadratic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample:</td>
<td>below (1)</td>
<td>above (2)</td>
</tr>
<tr>
<td></td>
<td>below (3)</td>
<td>above (4)</td>
</tr>
<tr>
<td>Ancestral income</td>
<td>0.334***</td>
<td>0.170</td>
</tr>
<tr>
<td></td>
<td>(0.124)</td>
<td>(0.150)</td>
</tr>
<tr>
<td>Current income</td>
<td>0.012</td>
<td>0.834***</td>
</tr>
<tr>
<td></td>
<td>(0.190)</td>
<td>(0.119)</td>
</tr>
<tr>
<td>Threshold location</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>Dependent var. mean</td>
<td>23.033</td>
<td>23.755</td>
</tr>
<tr>
<td>N</td>
<td>1810</td>
<td>3844</td>
</tr>
</tbody>
</table>

*Source: South India Community Health Study (SICHS)*
Metabolic Disease-BMI Relationship

![Graph showing the relationship between BMI and the probability of disease for All India and South India. The x-axis represents BMI values from 15 to 35, and the y-axis represents the probability of disease from 0.04 to 0.18. The graph includes two lines: one for All India (blue) and one for South India (red). The relationship shows an increase in probability of disease as BMI increases.](image-url)
Given that $y_0 \geq 0$ and the assumption that there is no lower threshold, all individuals with $y_t \leq \alpha$ must lie within their set point thresholds

$$\mathbb{E}(z_t | y_t) = \int_{-\infty}^{y_t} [a + b(y_t - U_t)] \frac{\phi(U_t; \mu_t, \sigma^2_t)}{\Phi(y_t; \mu_t, \sigma^2_t)} \ d U_t$$

$$= a + b(y_t - e^L(y_t))$$

where

$$e^L(y_t) = \frac{1}{\Phi(y_t; \mu_t, \sigma^2_t)} \int_{-\infty}^{y_t} U_t \phi(U_t; \mu_t, \sigma^2_t) \ d U_t$$

$$= \mu_t - \sigma_t \Lambda \left( \frac{y_t - \mu_t}{\sigma_t} \right)$$
BMI-Income Relationship (cross-section)

- Some individuals with $y_t > \alpha$ will have crossed their threshold, while others will remain at their set point

$$\mathbb{E}(z_t | y_t) = \int_{-\infty}^{\alpha} \left[ a + b(y_t - U_t) \right] \frac{\phi(U_t; \mu_t, \sigma^2_t)}{\Phi(y_t; \mu_t, \sigma^2_t)} \, dU_t + \int_{\alpha}^{y_t} \left[ a + by_t \right] \frac{\phi(U_t; \mu_t, \sigma^2_t)}{\Phi(y_t; \mu_t, \sigma^2_t)} \, dU_t = a + b \left( y_t - e^H(y_t) \right)$$

$$e^H(y_t) = \frac{1}{\Phi(y_t; \mu_t, \sigma^2_t)} \int_{-\infty}^{\alpha} U_t \phi(U_t; \mu_t, \sigma^2_t) \, dU_t$$

$$= \frac{\mu_t \Phi \left( \frac{\alpha - \mu_t}{\sigma_t}; 0, 1 \right) - \sigma_t \phi \left( \frac{\alpha - \mu_t}{\sigma_t}; 0, 1 \right)}{\Phi \left( \frac{y_t - \mu_t}{\sigma_t}; 0, 1 \right)}$$
Additional covariates and nonparametric shift-share instrument

Source: India Human Development Survey (IHDS)
Nonparametric Shift-Share Instrument

Individual crop shares

(a) BMI

(b) Metabolic Disease

Source: India Human Development Survey (IHDS)
Outcomes by gender

(a) Men

(b) Women

Source: India Human Development Survey (IHDS)
Alternative measures of nutritional status and separate metabolic diseases

(a) Height and weight

(b) Diabetes, hypertension and cardiovascular disease

Source: India Human Development Survey (IHDS)
Nutrient Intake and Children’s Illness with respect to Income

(a) Nutrient intake

(b) Children’s illness
Threshold Tests (Nutrient Intake and Children’s Illness)

(a) Nutrient intake

(b) Children’s illness
Selective mortality: BFA quantile regressions

(a) Baseline slope

(b) Slope change
Sensitivity of slope coefficients

(a) Threshold

(b) Mean of the income shock

(c) Standard deviation of the income shock

(d) Number of generations
Predicted Income and CSI Relationship

![Graph showing the relationship between predicted income and log CSI. The graph includes two curves: one for non-parametric and another for quadratic models. The x-axis represents log CSI, and the y-axis represents predicted income. The non-parametric curve is a dashed blue line, while the quadratic curve is a dotted red line.](image-url)
Assortative Matching

(a) Current generation

(b) Parent's generation

Source: South India Community Health Study (SICHS)
Historical income is measured by tax revenue per acre of cultivated land in 1871.
Threshold Tests (Indonesia and Ghana)

(a) Indonesia

(b) Ghana

Source: Indonesia Family Life Survey (IFLS), Ghana Socioeconomic Panel Survey (GSPS)
### Piecewise Linear Equation Estimates (Indonesia and Ghana)

<table>
<thead>
<tr>
<th>Sample country:</th>
<th>Indonesia</th>
<th>Ghana</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample country:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dependent variable:</td>
<td>adult BMI</td>
<td>metabolic disease</td>
</tr>
<tr>
<td>Slope below ((\beta_L))</td>
<td>0.067</td>
<td>-0.001</td>
</tr>
<tr>
<td>(SE)</td>
<td>(0.065)</td>
<td>(0.010)</td>
</tr>
<tr>
<td>Slope above ((\beta_H))</td>
<td>0.398**</td>
<td>0.017</td>
</tr>
<tr>
<td>(SE)</td>
<td>(0.069)</td>
<td>(0.011)</td>
</tr>
<tr>
<td>Threshold location</td>
<td>6.10</td>
<td>6.00</td>
</tr>
<tr>
<td>[5.80, 6.65]</td>
<td>[4.10, 6.75]</td>
<td></td>
</tr>
<tr>
<td>Threshold test (p - value)</td>
<td>0.000</td>
<td>0.006</td>
</tr>
<tr>
<td>Dep. var. mean</td>
<td>23.532</td>
<td>0.160</td>
</tr>
<tr>
<td>N</td>
<td>30,812</td>
<td>24,164</td>
</tr>
</tbody>
</table>

*Source:* Indonesia Family Life Survey (IFLS), Ghana Socioeconomic Panel Survey (GSPS)
Population Density - CSI Relationship

Source: FAO Global Agro-Ecological Zones (GAEZ) dataset