What would it cost to end extreme poverty?

Roshni Sahoo[1] Joshua Blumenstock[2] Leo Selker[2] Paul Niehaus[3]
Stefan Wager[1]

[1]Stanford University [2]University of California, Berkeley [3]University of California, San Diego

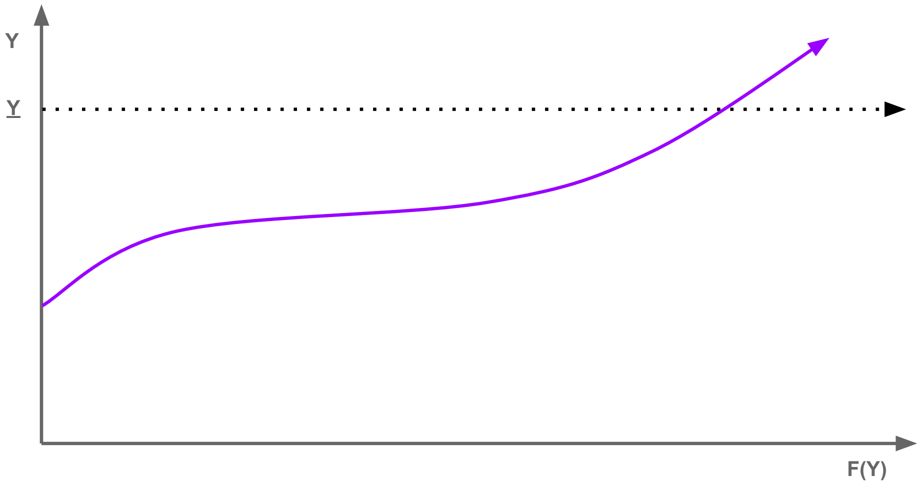
December 2025

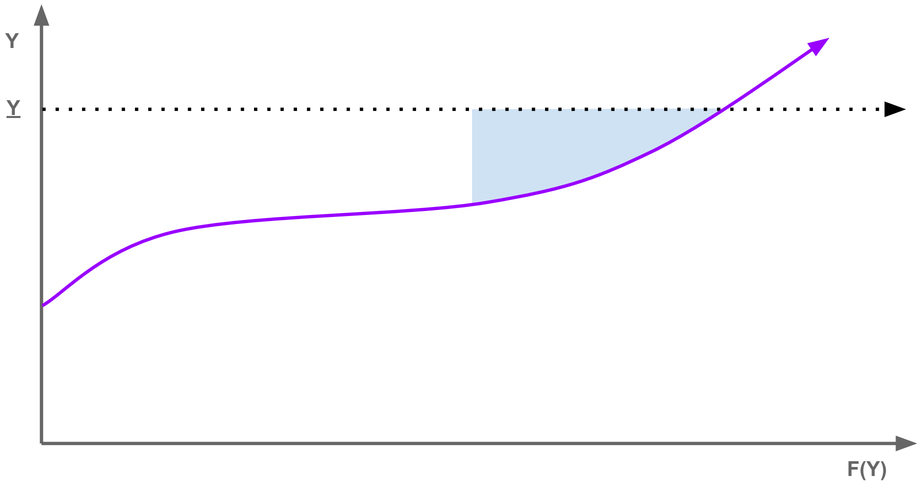
With thanks to Aditi Acharya, Erick Lopez, and Shruthi Ramesh for excellent RA work.

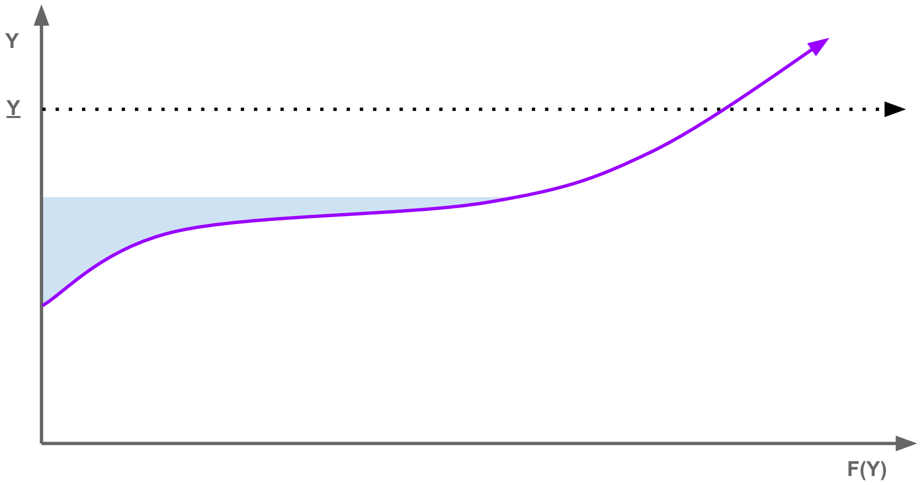
- Retain the information environment governments actually face (as in PMTs)
  - Observe living standards Y_i in a small representative sample (typically 0.001% to 0.01% of national population)
  - Observe predictors X_i in the population
- Formulate poverty targeting as an optimization & statistical learning problem

$$\min_{t:\mathcal{X}\to\mathbb{R}_+} \left\{\mathbb{E}_F\left[L(Y_i + t(X_i))\right] : \mathbb{E}_F\left[t(X_i)\right] \leq B, \quad t(x) \geq 0 \quad \forall x \in \mathcal{X}\right\}. \tag{1}$$

where F is the population distribution, L a loss function, t a transfer function, and B a budget


Y
Y
F(Y)


Y
Y
F(Y)


Y
Y
F(Y)

$$\min_{t:\mathcal{X}\to\mathbb{R}_+} \left\{\mathbb{E}_F\left[L(Y_i + t(X_i))\right] : \mathbb{E}_F\left[t(X_i)\right] \leq B, \quad t(x) \geq 0 \quad \forall x \in \mathcal{X}\right\}.$$

- Implementation requires analytical characterizations of solutions sufficient to

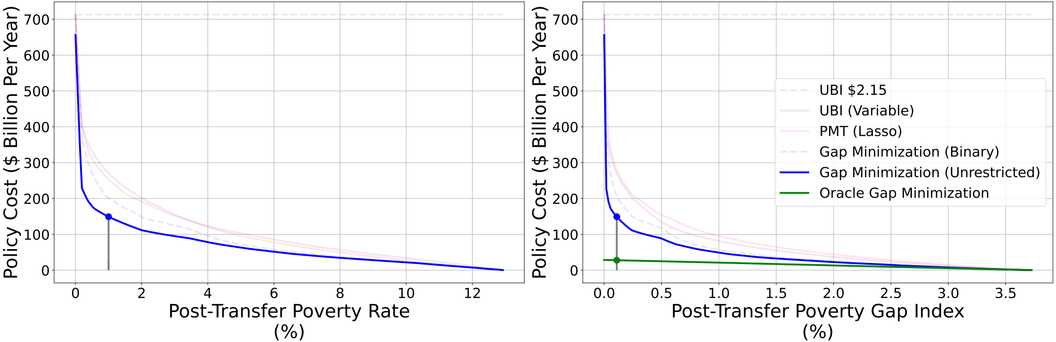
  (a) Reduce dimensionality of space of t's to search
  (b) (Hopefully) simplify functionals of F we need to learn

- This too turns out to depend on the poverty metric L used

  - The poverty gap (which is weakly convex) yields FOCs and requires learning conditional *quantiles* of F $\to$ can use neural networks and large predictor sets
  - The poverty rate (which is not) does not, and requires learning conditional *densities* of F $\to$ requires semiparametric methods and smaller predictor sets

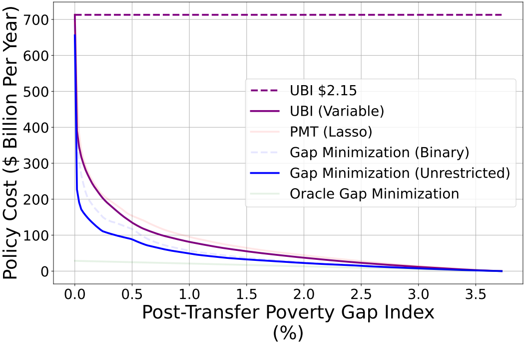
- **Data:** Nationally representative surveys (mostly LSMS) from 19 countries that collectively account for 43% of the world's extreme poor
    - Extrapolate globally using now-cast poverty rate and gap estimates from World Poverty Clock
- **Living standards:** consumption aggregate as provided in data release (except in India). Match WB PIP rates exactly in most cases, approximately in a few.
- **Predictor selection:** characteristics that have been used, or are analogous to others that have been used, by existing real-world PMT exercises
- **Data use plan:** pre-specifies approach including e.g. hyperparameter tuning, to avoid human-in-the-loop overfitting

Feasible versus oracle



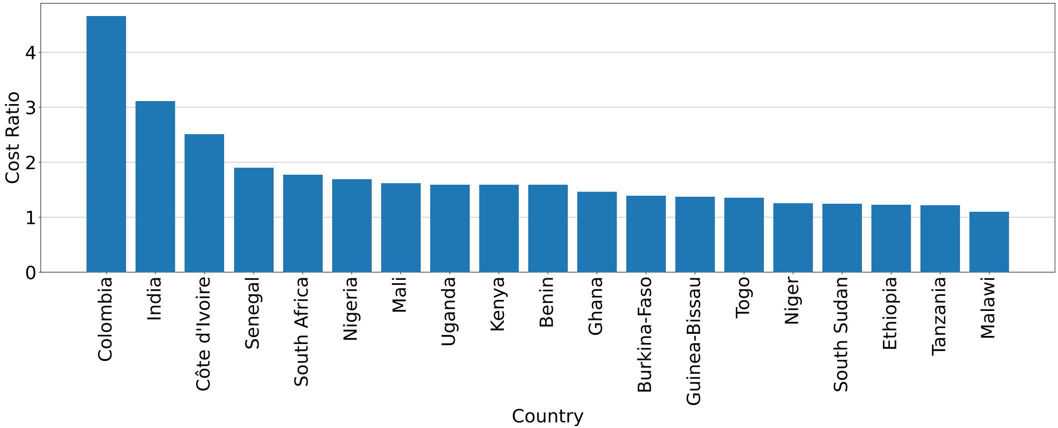
A feasible gap-minimizing policy achieving a 1% poverty rate in every country costs 5.4 times reduction in the poverty gap it achieves (i.e., oracle cost)

Feasible versus universal transfers



A feasible gap-minimizing policy achieving a 1% poverty rate in every country costs 55% of the cost of Universal Supplemental Income

Universal-to-feasible cost ratios, by country



Gap minimization (achieving a 1% poverty rate) is much more cost efficient than USI in most countries, less so in some of the poorest

The cost of achieving 1% national poverty rates increases...

...by 35% if we restrict to "binary" policies, like many current programs

...by 50% if we use the new \$3.00 2021 PPP poverty line

...by 43% (in Togo) if we use only satellite imagery as predictors

...slightly, if we directly minimize the rate, v.s. minimizing the gap

Estimating (via extrapolation) the cost of achieving a *global* poverty rate of 1%,

0.30% **cost / GDP**

Estimating (via extrapolation) the cost of achieving a *global* poverty rate of 1%,

0.30%	**cost / GDP**
0.21%	**ODA / GDP**

Estimating (via extrapolation) the cost of achieving a *global* poverty rate of 1%,

0.30%	**cost / GDP**
0.21%	**ODA / GDP**
2.19%	**alcoholic beverages / GDP**