NOWCASTING GLOBAL POVERTY

R. ANDRES CASTANEDA AGUILAR
DANIEL GERSZON MAHLER
DAVID NEWHOUSE

WORLD BANK GROUP

November 7, 2019

Motivation

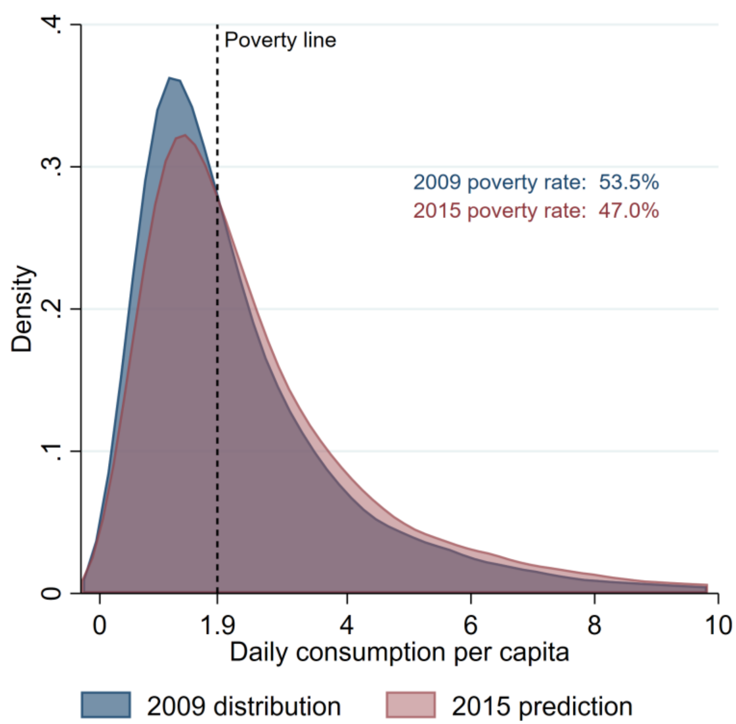
- The first SDG calls for eliminating extreme poverty by 2030
- In the typical developing country, the latest data on extreme poverty is from 2013
- This prohibits tracking the SDG and prioritizing funds and resources to the places most in need
- **Objective**: Predict poverty in every country of the world as of the last year ("nowcast") using cross-country macro data and machine learning techniques.

The World Bank's current nowcasts

$$povertyrate_{t_{survey}} = F\left[welfare_{h,t_{survey}} < 1.9\right]$$

$$povertyrate_{t_{nowcast}} = F\left[welfare_{h,t_{survey}}\left(1 + growth_{t_{nowcast},t_{survey}}\right) < 1.9\right]$$

Nigeria example



WORLD BANK GROUP

Assumptions behind current nowcasts

1. Growth in GDP/capita is fully passed through to growth in welfare observed in household surveys
2. Growth in GDP/capita is the only variable informative of changes in poverty
3. Inequality does not change between the time of the survey and the time of the nowcast

Would relaxing these assumptions improve accuracy?

METHOD
&
DATA

Data

- Use all international poverty estimates available in PovcalNet
- Predict poverty using
  - Other variables in PovcalNet
  - All variables in the World Economic Outlook (WEO)
  - All variables in the World Developing Indicators (WDI)
  - Remote sensing variables (to come)

WORLD BANK GROUP

What should be the predictor?

The <u>level</u> of poverty	<u>Annualized growth</u> in mean consumption
Pro	Pro
	Con

WORLD BANK GROUP

What should be the predictor?

The <u>level</u> of poverty	<u>Annualized growth</u> in mean consumption
Pro	Pro
• Most intuitive option; this is what we ultimately care about • Can yield estimates for countries with no data	• Works with all poverty lines • Takes advantage of previous data
Con	Con
• A new model is needed for each poverty line	• Unlikely to work well when data is very old • Does not yield estimates for countries without data • Requires distributional assumption

WORLD BANK GROUP

Organizing the data

Country	Year	Poverty rate (%)	Mean consumption	GDP per capita	Employment ratio	…
Mali	1994	85.1	37.4	504	67.5	
Mali	2001	58.5	68.2	620	65.7	
Mali	2006	51.2	76.9	675	63.1	
Mali	2010	49.7	72.6	707	65.2	
Mali	2018			777	64.2	
Togo	2006	55.6	76.2	492	79.2	
Togo	2011	54.2	79.9	546	76.5	
Togo	2015	49.2	82.5	622	76.6	
Togo	2018			664	76.4	
…						

WORLD BANK GROUP

Organizing the data

Country	Spell	Annualized growth in mean consumption (y-variable)	Annualized growth in GDP per capita (x-variable)	Annualized change in employment ratio (x-variable)
Mali	1994-2001	9.0%	3.0%	-0.27
Mali	2001-2006	2.4%	1.7%	-0.52
Mali	2006-2010	-1.4%	1.2%	0.54
Mali	2010-2018		1.2%	-0.13
Togo	2006-2011	1.0%	2.1%	-0.54
Togo	2011-2015	0.8%	3.3%	0.04
Togo	2015-2018		2.2%	-0.09

WORLD BANK GROUP

Organizing the data

Country	Spell	Annualized growth in mean consumption (y-variable)	Annualized growth in GDP per capita (x-variable)	Annualized change in employment ratio (x-variable)	Sample
Mali	1994-2001	9.0%	3.0%	-0.27	Train
Mali	2001-2006	2.4%	1.7%	-0.52	Train
Mali	2006-2010	-1.4%	1.2%	0.54	
Mali	2010-2018		1.2%	-0.13	
Togo	2006-2011	1.0%	2.1%	-0.54	Train
Togo	2011-2015	0.8%	3.3%	0.04	
Togo	2015-2018		2.2%	-0.09	

WORLD BANK GROUP

Organizing the data

Country	Spell	Annualized growth in mean consumption (y-variable)	Annualized growth in GDP per capita (x-variable)	Annualized change in employment ratio (x-variable)	Sample
Mali	1994-2001	9.0%	3.0%	-0.27	Train
Mali	2001-2006	2.4%	1.7%	-0.52	Train
Mali	2006-2010	-1.4%	1.2%	0.54	Test
Mali	2010-2018		1.2%	-0.13	
Togo	2006-2011	1.0%	2.1%	-0.54	Train
Togo	2011-2015	0.8%	3.3%	0.04	Test
Togo	2015-2018		2.2%	-0.09	

WORLD BANK GROUP

Organizing the data

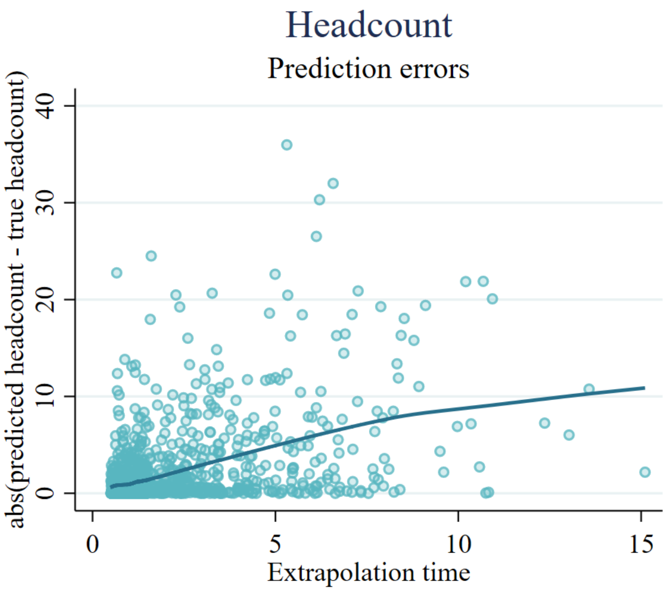
Country	Spell	Annualized growth in mean consumption (y-variable)	Annualized growth in GDP per capita (x-variable)	Annualized change in employment ratio (x-variable)	Sample
Mali	1994-2001	9.0%	3.0%	-0.27	Train
Mali	2001-2006	2.4%	1.7%	-0.52	Train
Mali	2006-2010	-1.4%	1.2%	0.54	Test
Mali	2010-2018		1.2%	-0.13	Nowcast
Togo	2006-2011	1.0%	2.1%	-0.54	Train
Togo	2011-2015	0.8%	3.3%	0.04	Test
Togo	2015-2018		2.2%	-0.09	Nowcast

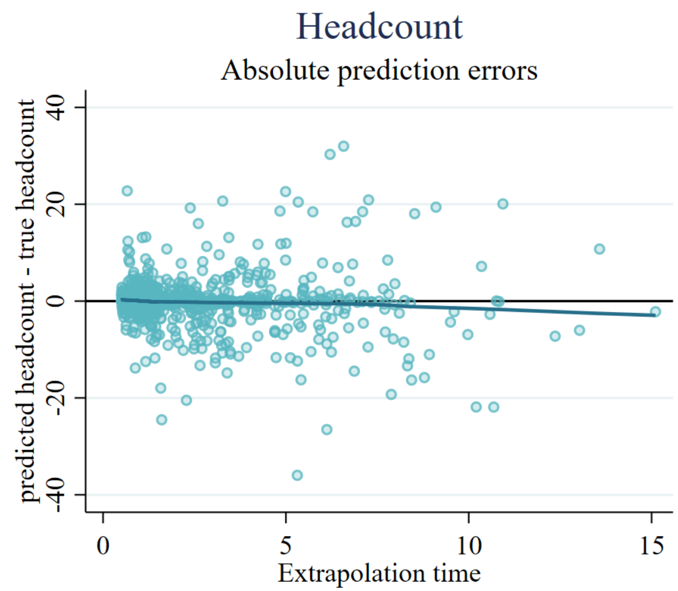
Prediction methods

- The current extrapolation method
- Applying a global passthrough rate
- Estimate growth in the mean using
  - Lasso
  - Post-lasso
  - Random Forest
  - Conditional Inference Random Forest
  - Gradient Boosting

RESULTS

Prediction error with current approach



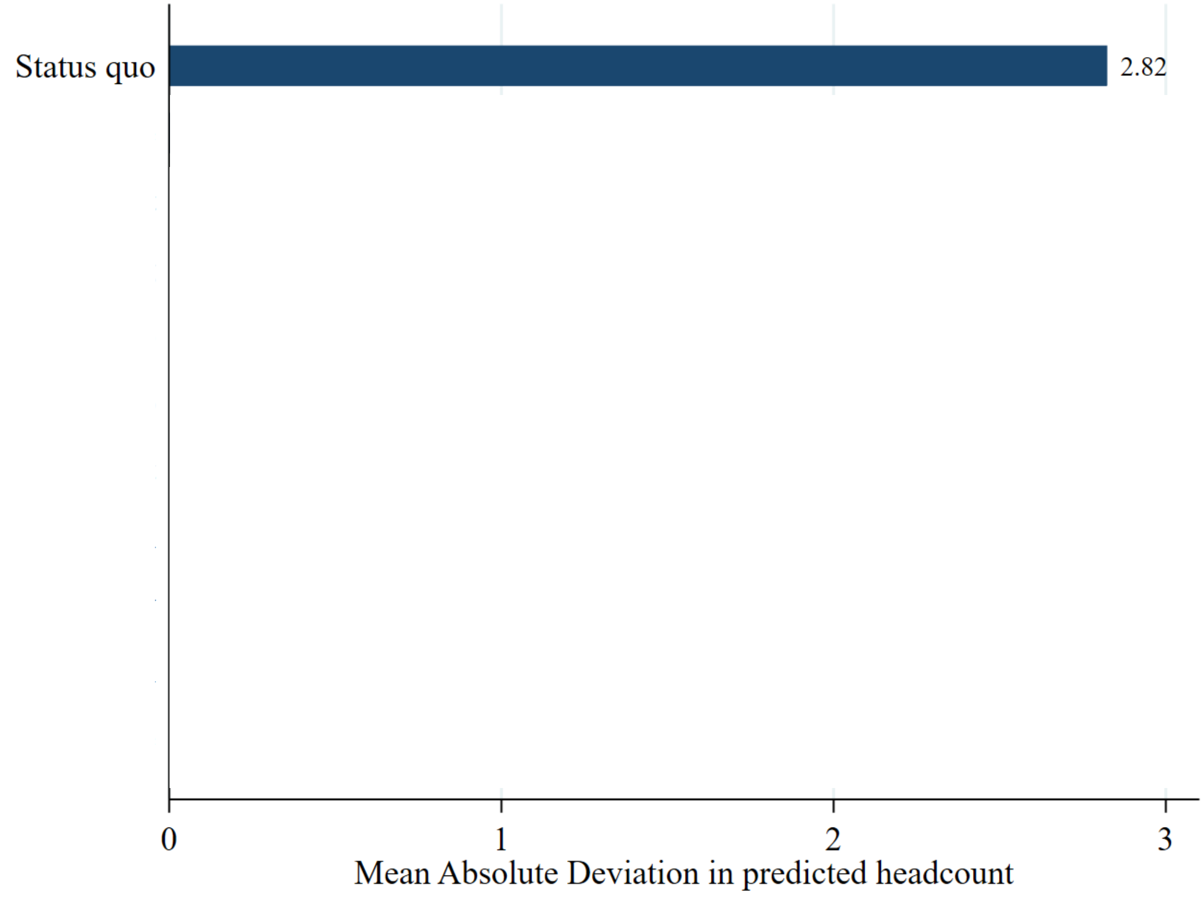


Mean absolute deviation (MAD):

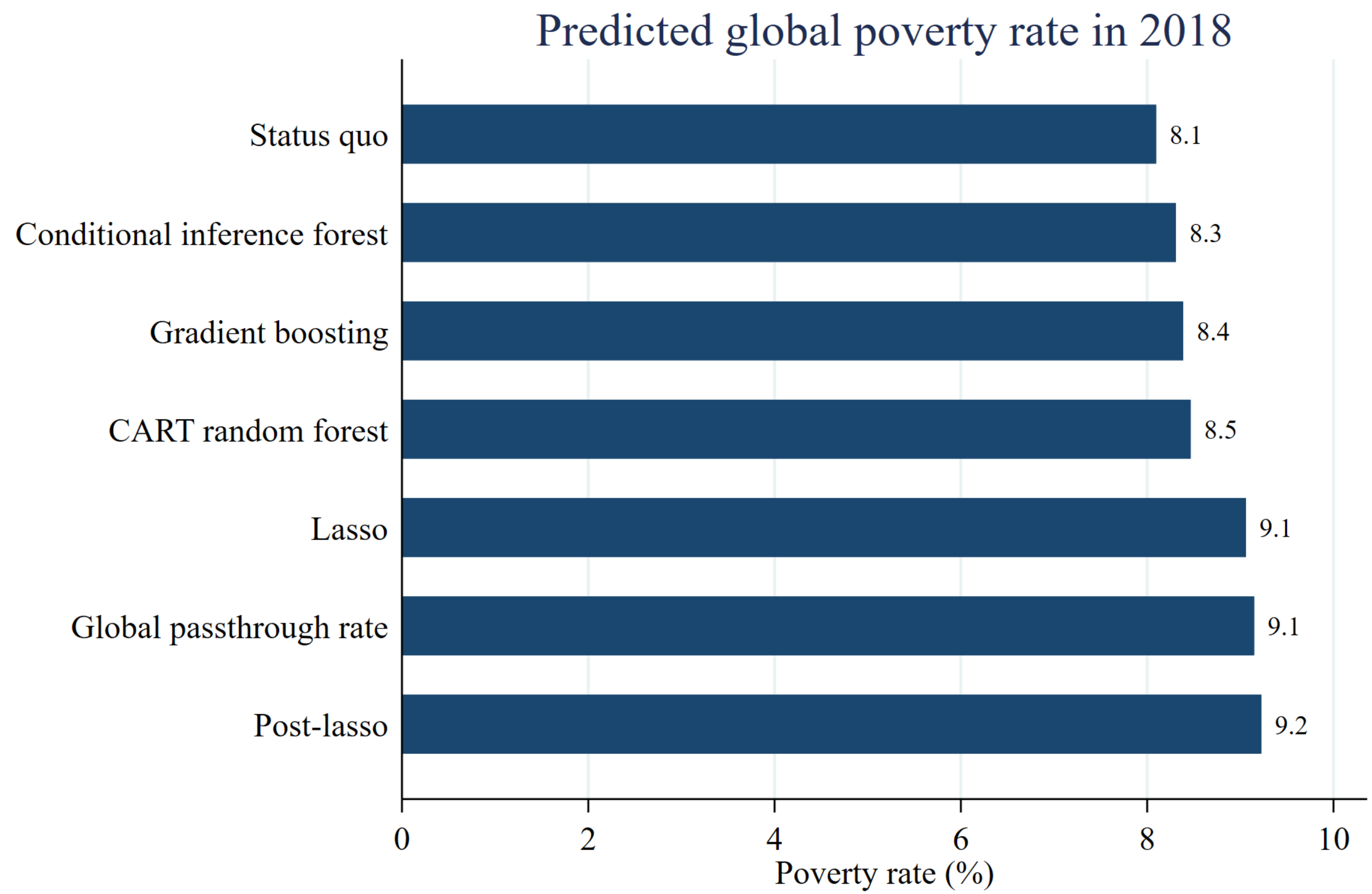
$$MAD_{test} = \frac{\left|headcount_{predicted,test,c} - headcount_{true,test,c}\right|}{N_{test,c}}$$

WORLD BANK GROUP

Prediction error of different methods

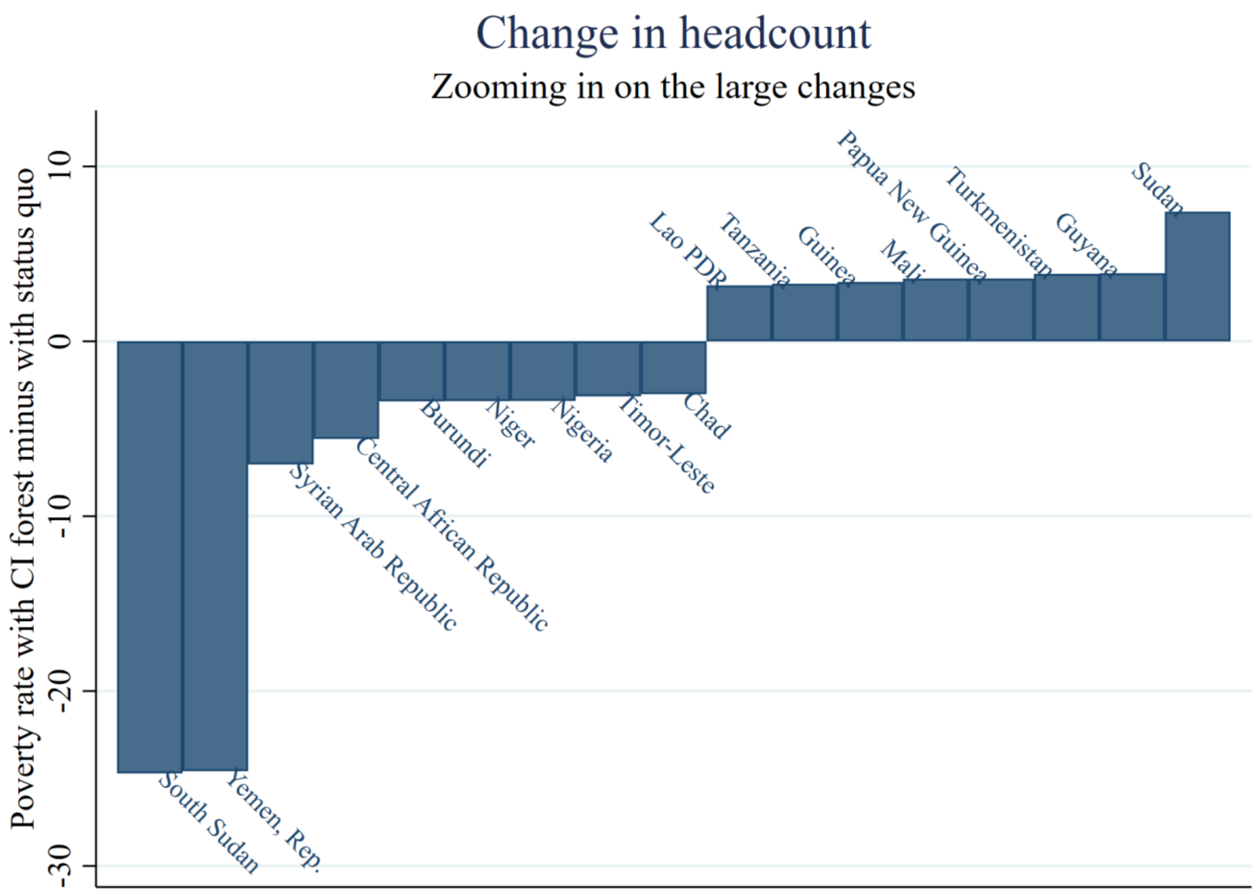


Impact on global poverty rates

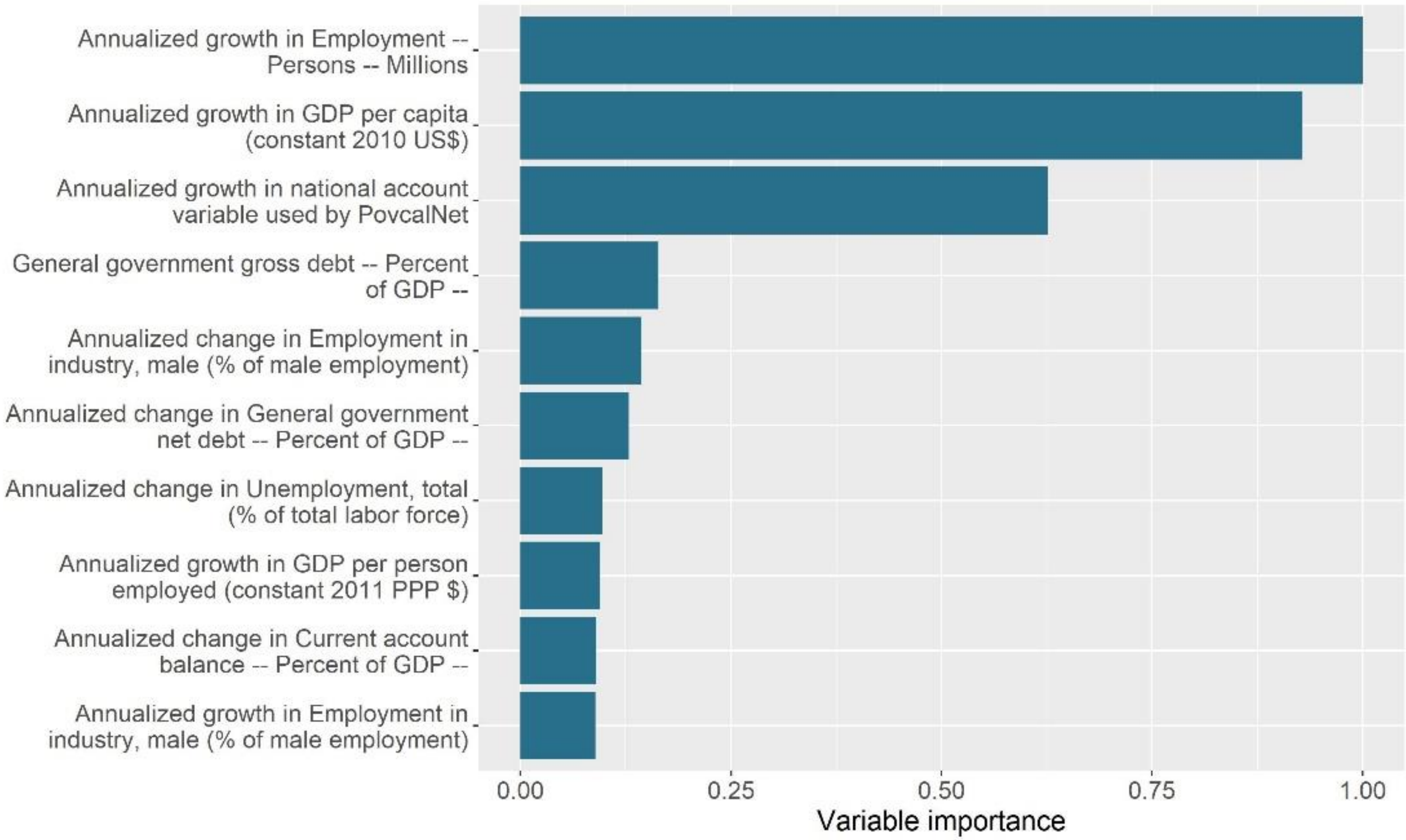


WORLD BANK GROUP

Impact on country-level poverty rates



Variables important for predictions



WORLD BANK GROUP

Conclusion

- Applying machine learning methods to WDI indicators only mildly improves on current nowcasts
  - But can noticeably affect extreme poverty estimates
- Yet there are avenues which could lead to improvements in the predictions
  - Add remote sensing indicators
  - Explore predicting inequality
  - Explore different methods for different types of cases

WORLD BANK GROUP

THANK YOU

DISCUSSION

Why are the errors still high?

1. The status quo works well on average
2. The surveys are not comparable
3. Predicting changes rather than poverty levels directly is not the way to go
4. We are very poor at predicting the mean
5. The distribution-neutral assumption does not hold

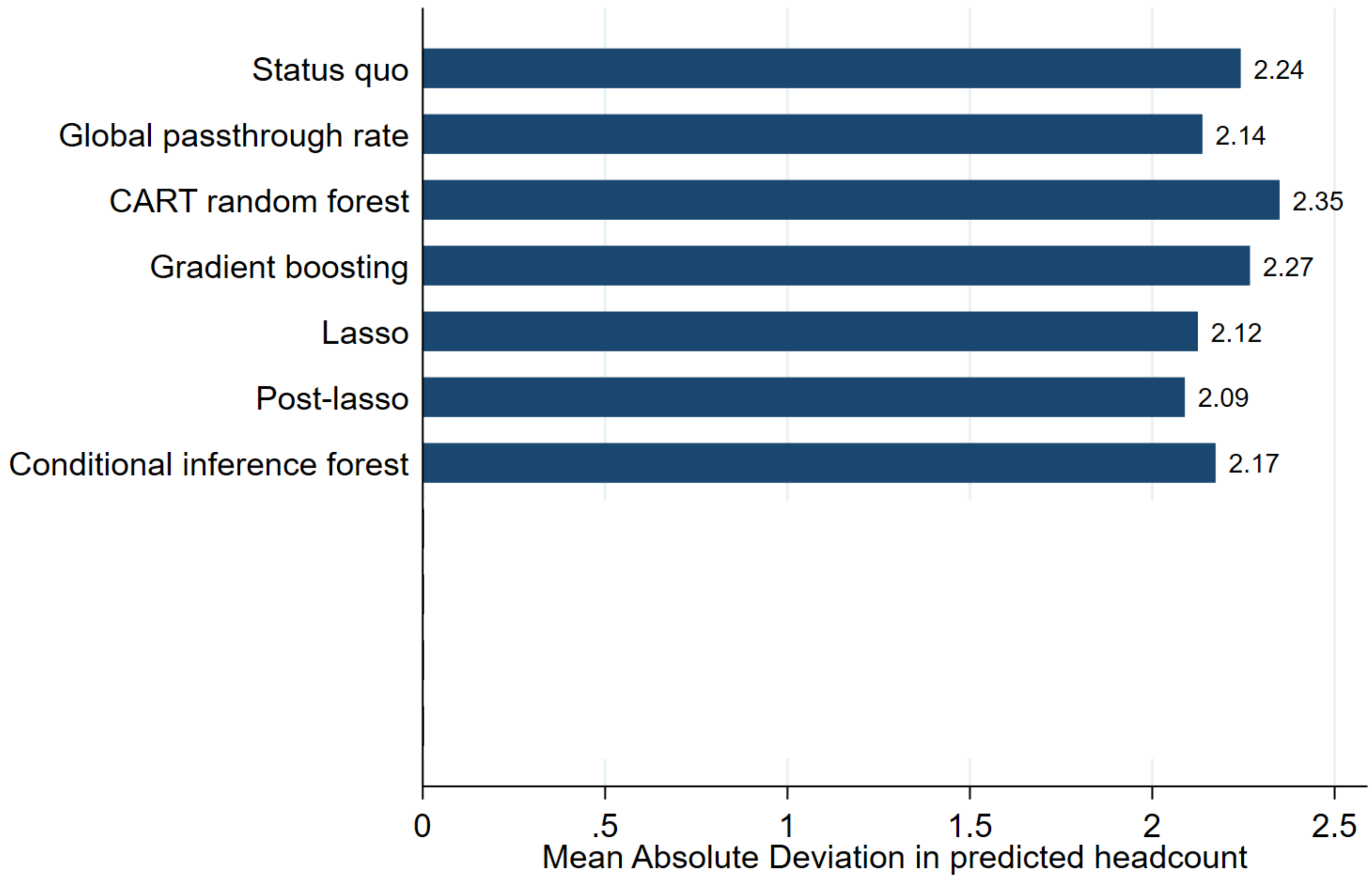
WORLD BANK GROUP

Why are the errors still high?

1. The status quo works well on average
2. The surveys are not comparable
3. Predicting changes rather than poverty levels directly is not the way to go
4. We are very poor at predicting the mean
5. The distribution-neutral assumption does not hold

WORLD BANK GROUP

Prediction error with comparable data

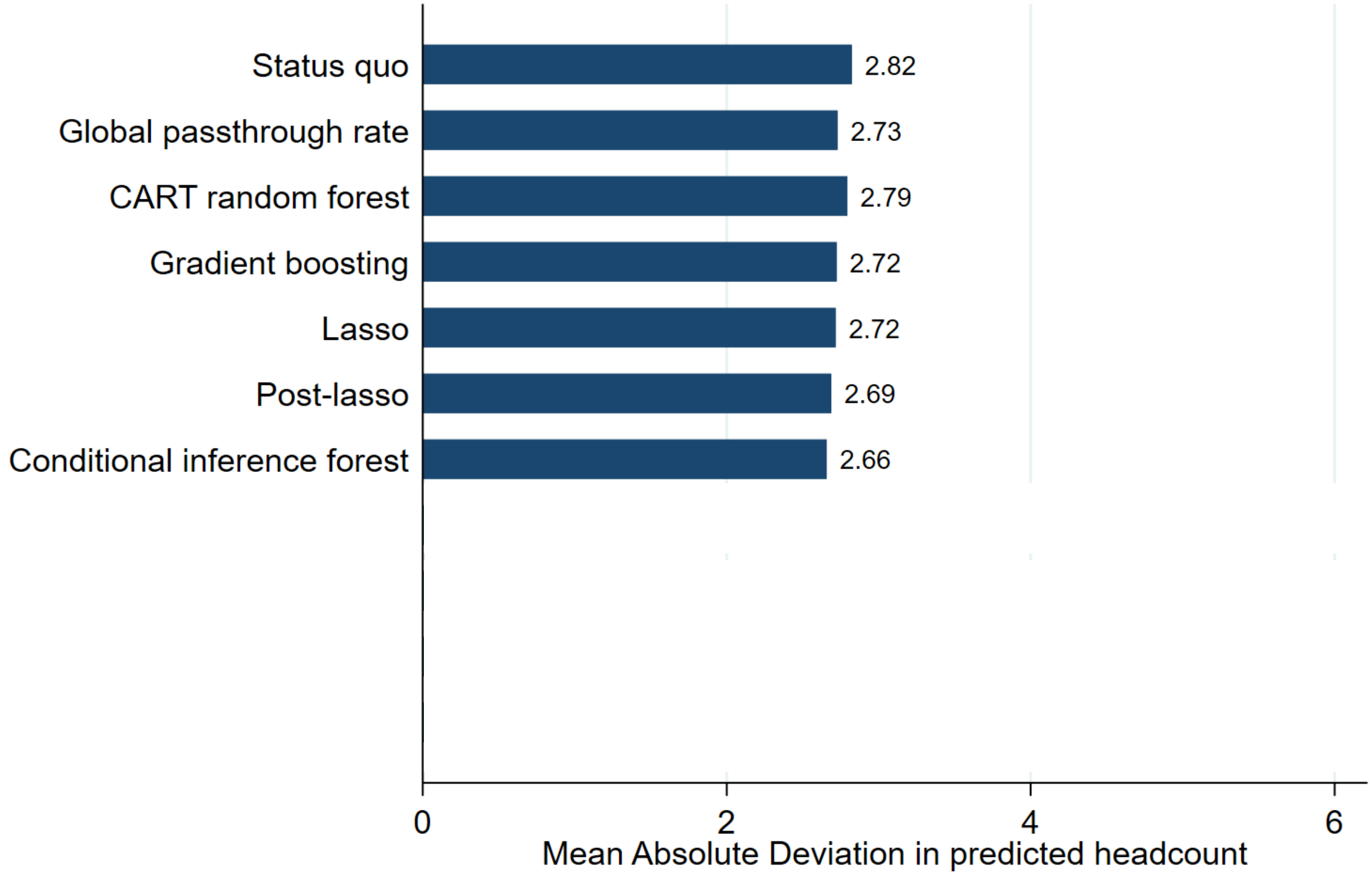


WORLD BANK GROUP

Why are the errors still high?

1. The status quo works well on average
2. The surveys are not comparable
3. Predicting changes rather than poverty levels directly is not the way to go
4. We are very poor at predicting the mean
5. The distribution-neutral assumption does not hold

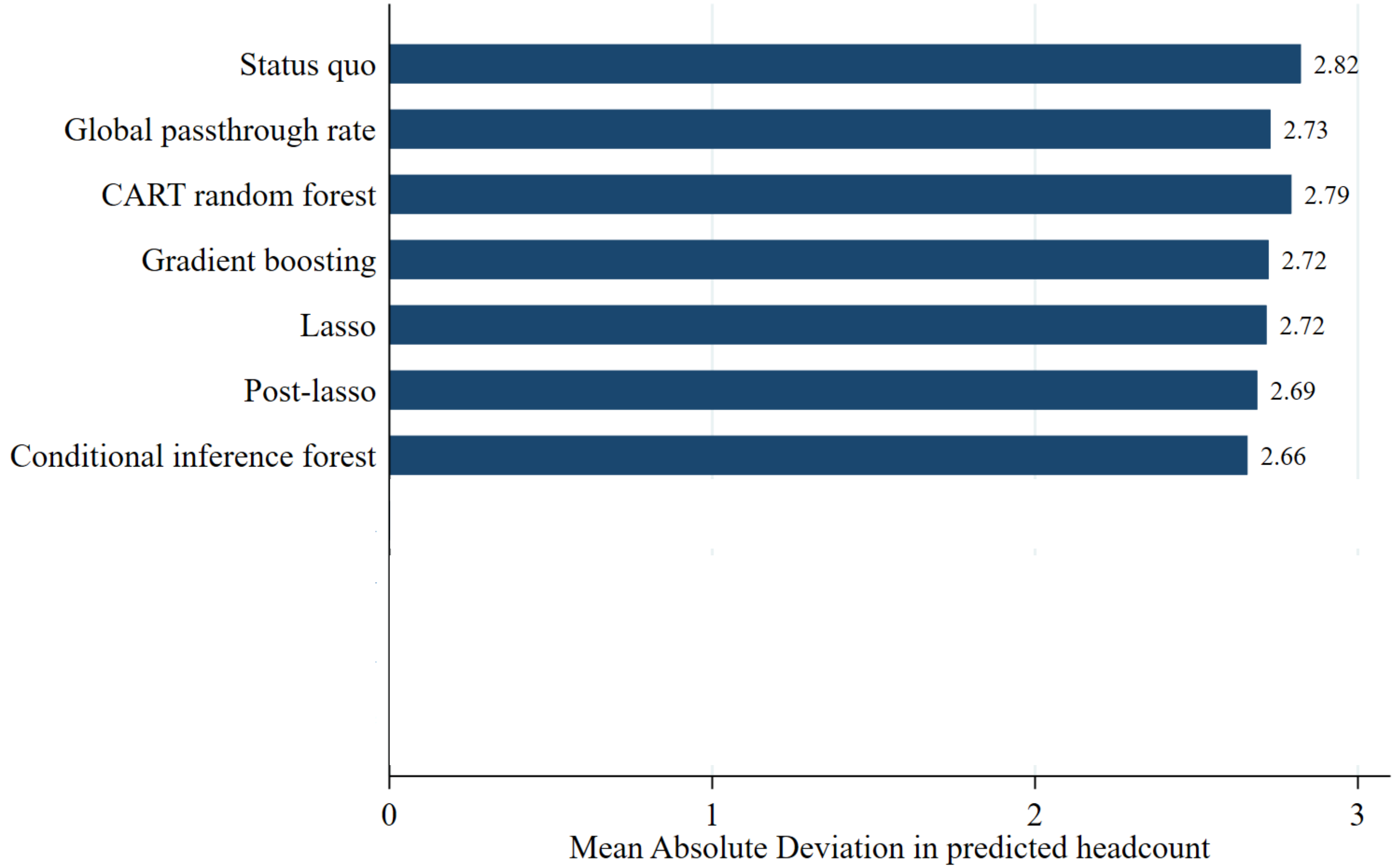
Predicting the headcount directly



Why are the errors still high?

1. The status quo works well on average
2. The surveys are not comparable
3. Predicting changes rather than poverty levels directly is not the way to go
4. We are very poor at predicting the mean
5. The distribution-neutral assumption does not hold

Prediction the mean perfectly



WORLD BANK GROUP

Why are the errors still high?

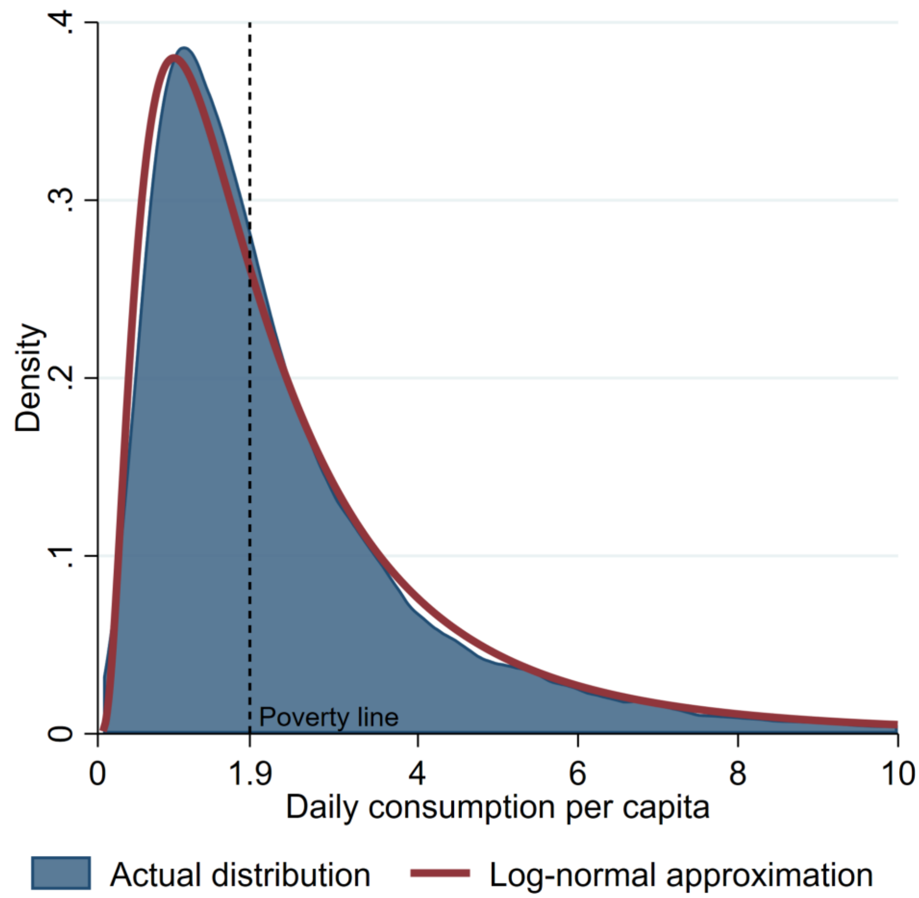
1. The status quo works well on average
2. The surveys are not comparable
3. Predicting changes rather than poverty levels directly is not the way to go
4. We are very poor at predicting the mean
5. The distribution-neutral assumption does not hold

WORLD BANK GROUP

How large would errors be if we could predict growth in the mean and changes in the Gini perfectly?

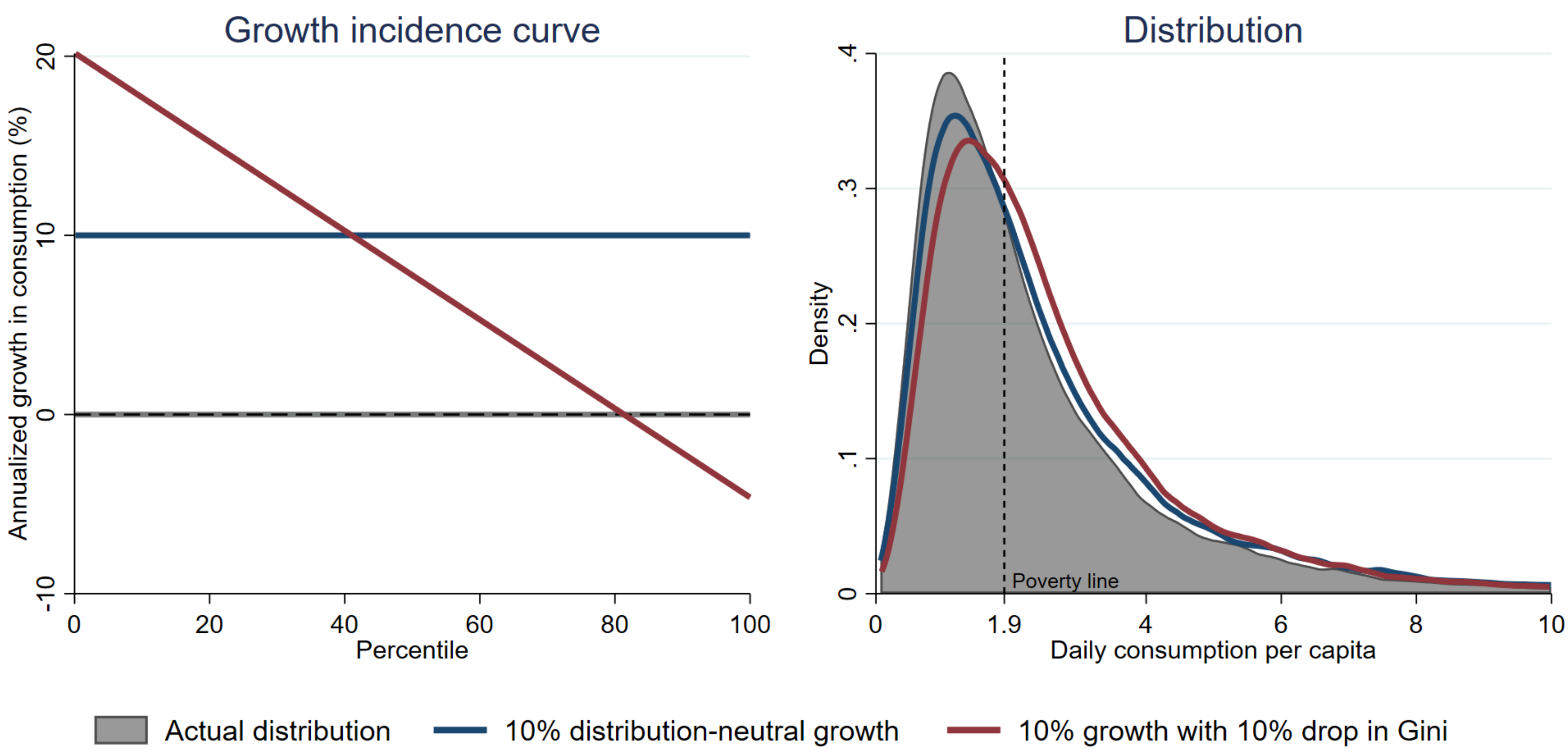
WORLD BANK GROUP

Using distributional approximations

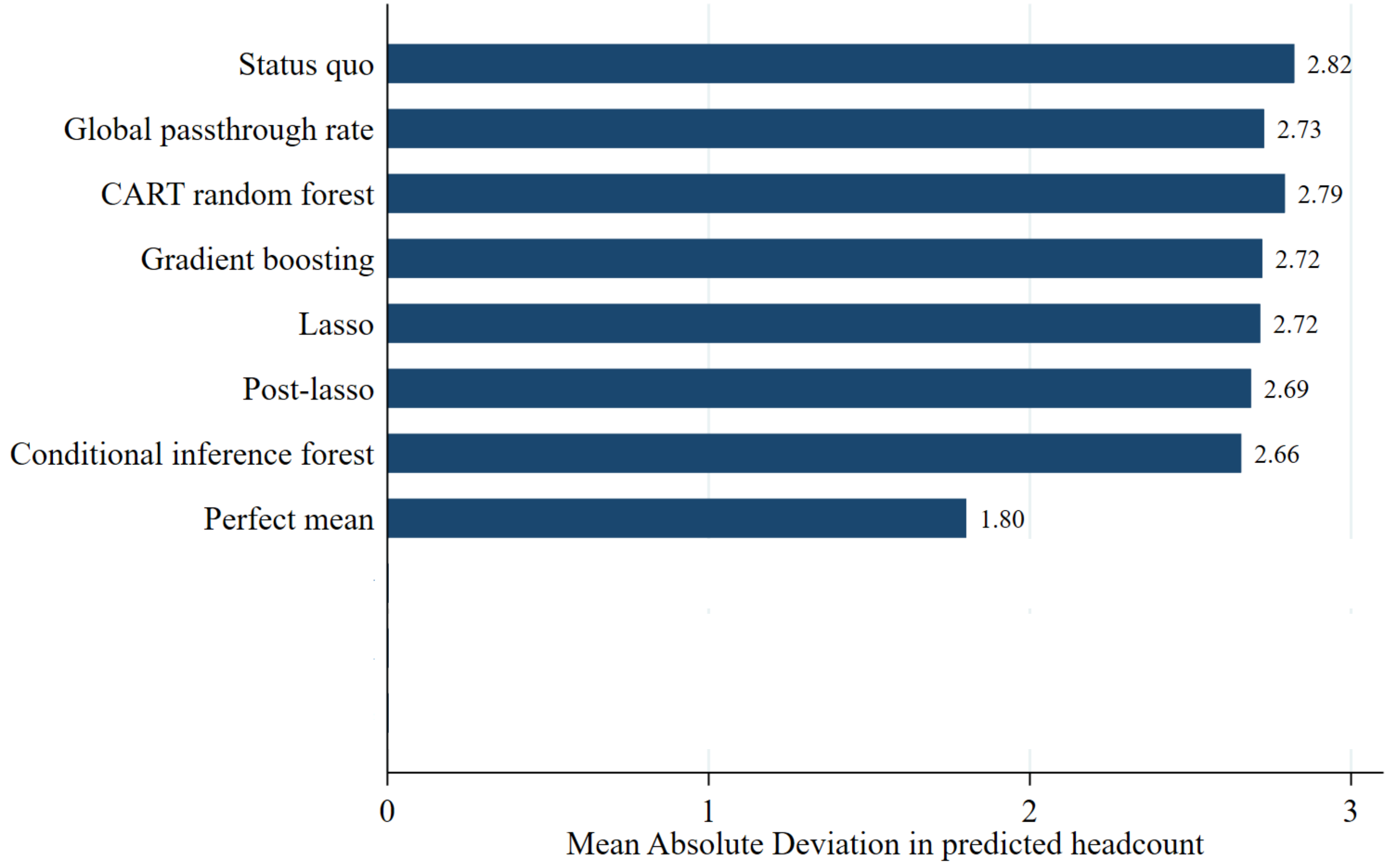


WORLD BANK GROUP

Using growth incidence curves



Prediction the mean and Gini perfectly



WORLD BANK GROUP