Explaining Income Inequality Trends in Countries: An Integrated Approach

Petra Sauer*, Narasimha Rao‡, and Shonali Pachauri‡

*WU Vienna, Institute for Macroeconomics, Austria;
‡ International Institute for Applied Systems Analysis, Austria

34th IARIW General Conference
Session 7B: New Approaches to Studying the Causes and Consequences of Poverty, Inequality, Polarization, and Social Conflict

Discussed by Roberto Zelli - Sapienza University of Rome

Main goals of the paper

- Comprehensive picture of global and regional within-country inequality trends:
  - three measures of income inequality with a correction for the multitude of data sources and measurement errors.
- Associations with potential drivers that explain common trend of (increasing) income inequality:
  - Focus on identifying systematic *global vs. region specific* drivers of inequality.

Causes of changing inequality in the World: review I

- Skill-biased **technological change** and increasing wage inequality
  - Innovation in advanced economies
  - Imitation and adoption in developing economies
- Technological change might be induced via **trade**
  - Innovation incentives in exporting sectors (e.g. Meschi & Viavarelli, 2009)
  - Technologies embedded in capital imports (e.g. Acemoglou, 2003)
  - Outsourcing and Multinationals
  - Trade effects (Hecksher-Ohlin model - Stolper-Samuelson theorem (SST) theorem):
    - Increasing inequality in high income countries
    - Decreasing inequality in low income countries
- Education: increasing the skill level of people via education should countervail the skill premium.

  "Race" Between Education and Technology
- Capital/labour share and role of finance

Causes of changing inequality in the World: review II

- Political system: States mitigate market risks and perform a more or less extensive redistributive role
  - Labor market institutions
  - Tax and transfer system & the secondary income distribution
  - Education & health policies

Measures of inequality and data source I

- When we move toward an analysis of inequalities in the wider world, we are required to cope with far more complex and uncertain data
- Development of a self-consistent data set derived from the UNU-WIDER World Income Inequality Database (WIID V3.0B).
- WIID combines an updated (unpublished) version of the Deininger&Squire dataset with unit data from a variety of other sources including the Luxembourg Income Study (LIS), Transmonee by UNICEF, SEDLAC (Social and Economic Database for Latin American Countries), World Bank sources and household surveys from national statistical offices, resulting in a total of 7,054 observations (with a time span 1960–2013).
- While the data still originate from different sources, they are transparent with respect to the income- and/or consumption definition, the statistical units to be adopted and the use of equivalence scales and weighting.

Measures of inequality and data source II

- Data selection to give robustness:
  - Only sources with comprehensive population coverage by gender, age and region.
  - Requirement of further consistency with respect to the underlying income concept, and construct per capita-, rather than household-based, Gini indices. Insurmountable, albeit well-known, source of inconsistency in the dataset is the income vs consumption concept used across countries to measure inequality.
  - Minimum of three time observations over one decade required.
- Two measures of Gini indices:
  - **Single Source Gini (SS Gini)**: select one source per country that best meets the selection criteria with regards to time coverage, reliability and regional consistency.
  - **MultiSource Gini (MS Gini)**: allow multiple sources for a country, but exclude sources in two cases: sources have data for only a single year; and sources whose data are unreasonably inconsistent for consecutive years with data from other sources

Measures of inequality and data source III

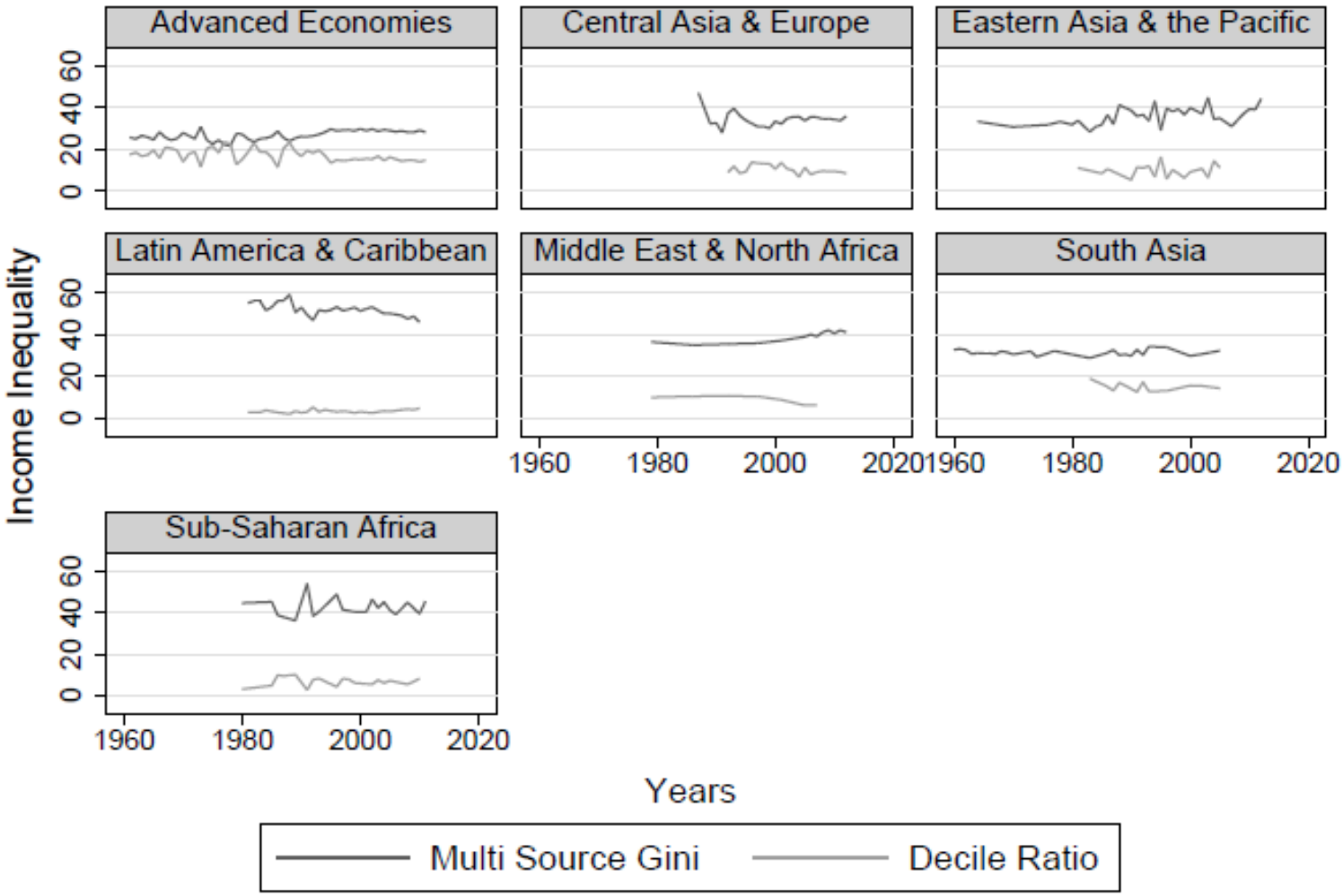
- **Decile Ratio**: the share (percent) of income accruing to the bottom decile in relation to that of the top decile of the income distribution.
- Creating this internally consistent dataset reduces the size of available observations and time span:
  - MS Gini: 945 observations from 74 countries, weighted to the advanced economies (1975–2013, concentrated in the period 1980–2011 for advanced economies and in the period 1988-2006 for developing economies)
  - SS Gini: 649 observations from 60 countries (1985–2004)
  - DecRatio: 691 obs from 71 countries

Table 1: Income Gini Source Inconsistencies

Country	Year	WB1	WB2	WB3
Ghana	1997	32.7		
Ghana	1998		50.7	
Ghana	1999			40.7
Sri Lanka	2000		27	
Sri Lanka	2002			40.2

WB1: World Bank Poverty Monitoring Database 2002; WB2: Deininger and Squire 2004; WB3:World Development Indicators 2004. In each case, WB1 provides the longer time series but WB2 or WB3 appear only once. Source: WIID V3.0B.

Multi-source Gini coefficients and Decile ratios



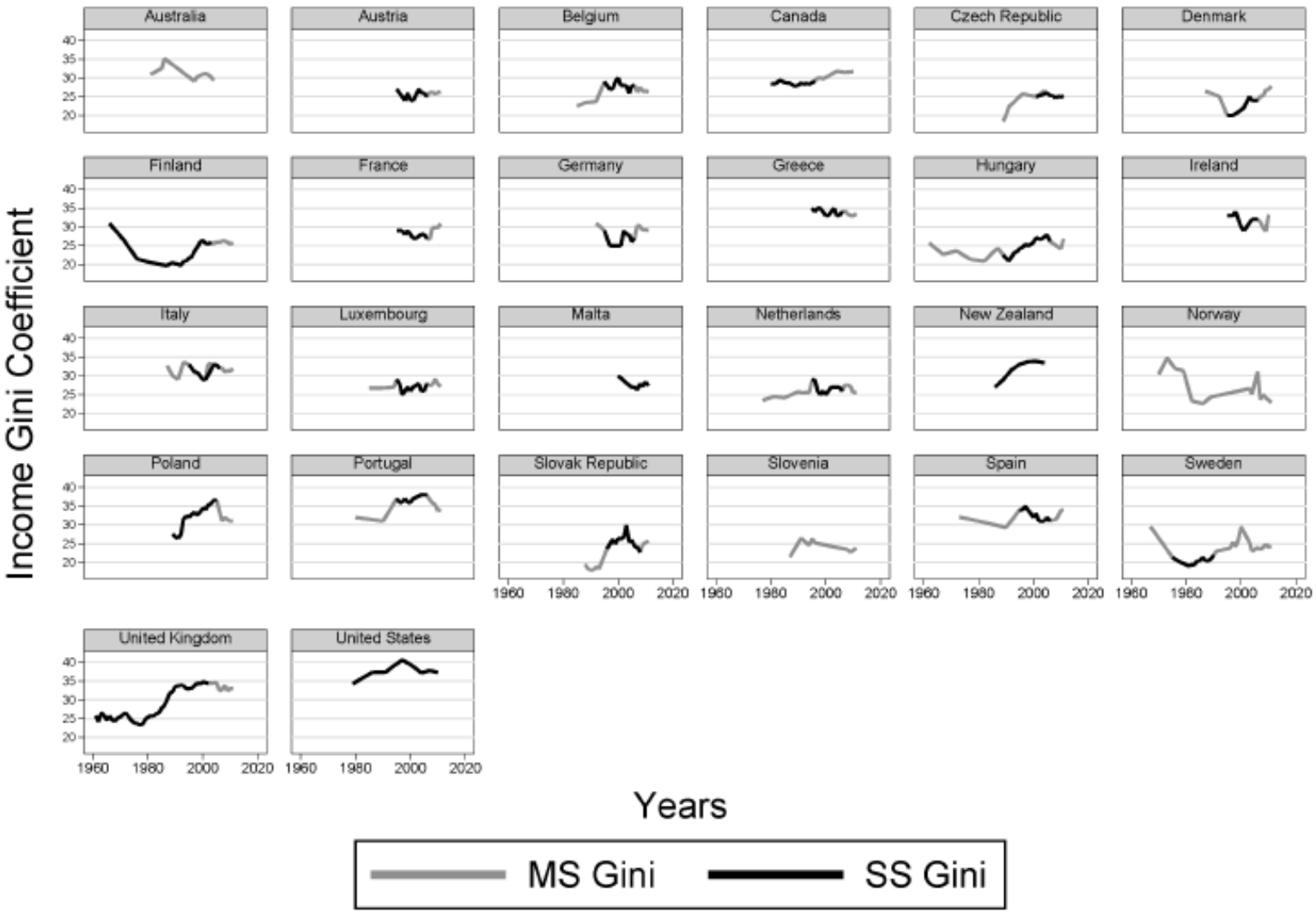
Summary Statistics of Income Inequality Series

	Mean		Sd		Trend[a]	
Region	MS Gini[b]	DecRatio	MS Gini	DecRatio	IncGini	DecRatio
AdvEcon	28.10	0.20	2.23	0.03	⇑	⇓
CAEE	34.02	0.11	3.01	0.01	**None**	**None**
LAC	50.86	0.03	2.44	0.01	⇓	⇑
EAP	36.79	0.10	4.01	0.01	⇑	**None**
SA	31.42	0.15	2.18	0.02	**None**	**None**
MENA	38.78	0.09	1.37	0.02	⇑	⇓
SSA	42.05	0.07	4.45	0.02	**None**	**None**

[a]Statistically significant time trend (> 1975) from a fixed effects regression of inequality against time.
[b] *Multi-source Gini*

AdvEcon (Advanced Economies), CAEE (Central Asia & Eastern Europe), LAC (Latin America & Carribean), EAP (Eastern Asia & the Pacific), SA (South Asia), MENA (Middle East & North Africa, SSA (Sub Saharan Africa).

Multi-source vs Single-source income Gini coefficients



Measures of covariates I

- Technological change represented as total factor productivity (TFP), measured using a conventional growth accounting framework.
  - The growth rate of real GDP per worker, the growth rate of physical capital per worker and the capital and labor shares are obtained from Penn World Tables (PWT8.0).
  - Human capital by worker as a function of mean years of schooling taken from the demographic dataset of the International Institute for Applied Systems Analysis and the Vienna Institute of Demography (IIASA/VID).
- Trade flows from Correlates of War (COW v3.0) bilateral trade database.
  - Imports flows categorized into those from high-income and low-income countries, as a proxy for high-skilled and low-skilled (manufacturing) imports respectively (exclusion of countries whose exports are predominantly natural resources or certain plantation crops)
  - Total exports in GDP

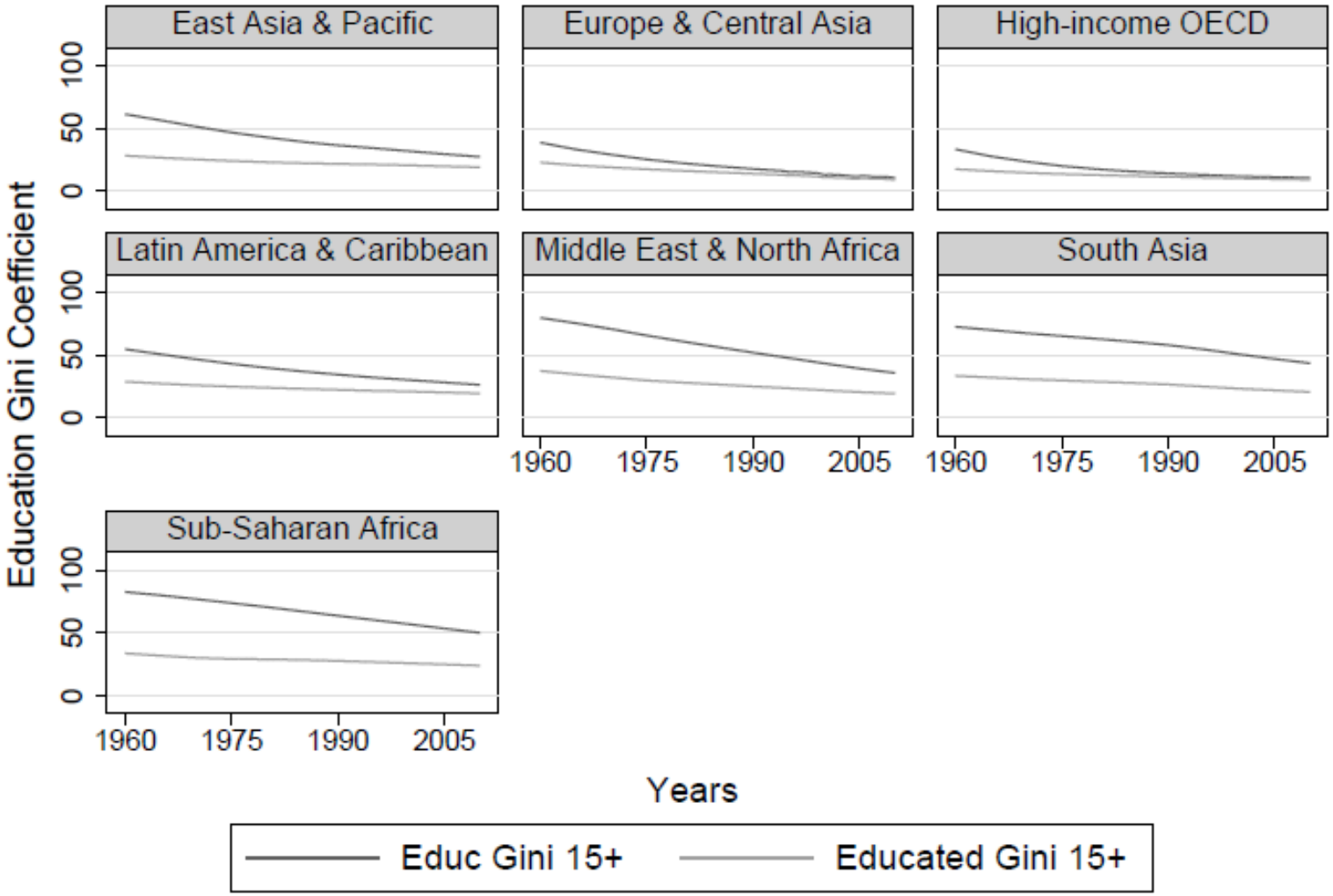
Measures of covariates II

- Education (data consistent with human capital estimation in TFP equation).
  - Source: demographic dataset from the International Institute for Applied Systems Analysis and the Vienna Institute of Demography (IIASA/VID).
  - Measures of attainment: mean years of schooling and population shares by education category: no schooling, primary, secondary and tertiary attainment
  - Measures of inequality: education Gini coefficient measuring the degree of education inequality in the population older than 15. Moreover decomposing the education Gini of the total population, into the share of unschooled people and an education Gini for those with at least some formal education (categories 2-4).
- Labor vs. Capital Income.
  - Share of labor income from PWT 8.0. Data are limited to after 1980 for Advanced Economies and after 1990 for Developing Economies

Measures of covariates III

- Governance. Four measures in order to capture the redistributive capacity of governments:
  - Relative weight of public social spending categories by using data on the shares in total spending from the Statistics of Public Expenditure for Economic Development (SPEED) database of the International Food Policy Research Institute (IFPRI) + robustness checks
    - education expenditures
    - health expenditures
    - social protection expenditures
  - Political orientation of the chief executive's party taken from the World Bank's Database of Political Institutions (DPI): an increasing score from 1 to 3, indicating right, center and left, respectively.

Education Gini coefficient



Summary statistics and trend of the variables in the model by Region

	Advanced Economies			Developing Economies		
Variable[a]	Mean	Within sd	Trend[b]	Mean	Within sd	Trend
$MSGini$	28.09	2.23	⇑	42.18	2.99	⇑
$SSGini$	28.39	2.03	⇑	42.90	2.65	None
$DecRatio$	15.61	2.88	⇓	6.50	1.21	None
TFP	0.94	0.07	⇑	0.95	0.10	⇓
Imp^{high}	25.71	5.62	⇑	20.26	7.82	⇑
Imp^{low}	4.14	2.29	⇓	6.04	3.84	⇑
Exp	30.05	5.89	⇑	26.80	7.32	⇑
L	61.32	2.81	⇓	50.52	3.70	⇑
PS^{Educ}	10.28	2.43	⇑	13.39	3.14	⇑
PS^{Health}	12.07	2.56	⇑	6.56	3.47	⇑
PS^{SP}	35.19	4.21	⇑	11.48	4.75	⇑
MYS_{15+}	12.13	0.63	⇑	8.50	0.61	⇑
$EducGini_{15+}$	12.76	2.20	⇓	27.72	3.18	⇓
$EducGini^{E}_{15+}$	10.70	1.28	⇓	18.01	1.24	⇓
p^{1}_{15+}	2.39	1.34	⇓	12.79	3.06	⇓
p^{2}_{15+}	18.69	4.50	⇓	35.07	2.69	⇓
p^{3}_{15+}	59.53	3.26	⇑	35.07	2.69	⇑
p^{4}_{15+}	19.41	3.09	⇑	11.96	1.54	⇑

[a]For an explanation of variable abbreviations see Section 6.
[b]Statistically significant time trend (> 1975) from a fixed effects regression of inequality against time.

Model specification

$$IGini_t = \beta_1 TFP_{t-1} + \boldsymbol{\beta_3}\mathbf{T}_{t-1} + \boldsymbol{\beta_4}\mathbf{E}_{t-1} + \beta_5 L_{t-1} + \boldsymbol{\beta_5}\mathbf{P}_{t-1} + \gamma Year + \alpha_i +$$

$$\mathbf{T} = [Imp^{high}, Imp^{low}, Exp]$$

$$\mathbf{E} = [MYS^{15+}; EducGini_{15+}; p^1_{15+}, EducGini^E_{15+}; p^2_{15+}, p^3_{15+}, p^4_{15+}]$$

$$\mathbf{P} = [PO, PS^{Educ}, PS^{Health}, PS^{SP}]$$

Estimation I

- Modelling: dynamic panel models estimated with a method that controls for country-specific effects and corrects for error disturbances, and test the robustness of the findings with respect to income inequality measures, model specifications and regions.
- Feasible Generalized Least Squares (FGLS) estimator combined with unobserved country fixed effects.
- All explanatory variables lagged one period in order to account for reverse causality.
- For tackling the issue of stationary, at least with respect to the dependent variables, TFP and the labor share, a time trend (Year) is included.
- Results for each of three income inequality measures and education variables.

Estimation II

- **Regressions run separately for the global level, for Advanced and for Developing Economies** as well as excluding and including political factors as doing so reduces the number of countries and time periods.

Main results I

Broad regional splitting (Advanced economies vs Developing economies) reveals differentials in drivers, magnitude and direction of effects

- No effect of TFP in AE but marginal effect in DE
- Trade with low-income countries decreases inequality in *all regions*
  - The equalizing effect of imports from low-income countries contradicts the implication of the Stolper-Samuelson theorem that imports from low-income countries compete with goods produced in low-skilled sectors of high-income countries, thereby increasing income inequality.
- Trade with high-income countries increases inequality only in DE
- Support for equalizing role of education in AE, mixed results for DE (increasing population shares with primary and tertiary education have disequalizing effects)
- Inequality increasing effect of education expenditures in AE, no significant effect in DE

Main results II

- Equalizing effect of health expenditures in *all regions*
- Relation between functional and personal income distribution in AE (the general decline in the share of labor incomes added to increase income inequality.

Remarks I

- Importance of dealing with issues of data quality in an empirical study of income inequality.
- Tremendous and valuable effort to create a consistent data-set of income inequality across countries and over time along with a coherent set of potential drivers of inequality.
- Trade-off between quality and coverage (actually, around 30% of the original WIDER data "survived"; Advanced Economies take 50% of the data)
- Exclusion vs imputation (to keep representativeness)
- Model:
  - Meaning of $(t-1)$ in the model when data are not equispaced. Does it have an effect on estimation?
  - FGLS with FE is a reasonable solution in estimation. Comparison with estimates obtained using other consistent estimators (e.g. GMM Arellano and Bond estimator, even if N is fixed)

Remarks II

- Some interactions in the model could give more insights (e.g. TFP and trade)
- The division between Advanced and Developing Economies is quite rough (aknowledged by the authors) and leaves further investigation of differential regional splittings.
- A sort of hierarchical approach may be useful: test the hypothesis that the coefficients of each explanatory variable are jointly equal across all the regressions
- Omitted factors? (Dynamics of top incomes, Financialization, Informal Markets)