Inequality as cholesterol:
Attempting to quantify inequality of opportunity

Francisco H. G. Ferreira
The World Bank and IZA

Policy Research Talk
World Bank, 22 January 2019

Outline

1. Why measure inequality of opportunity? Motivation
2. What is (in)equality of opportunity? Definition and illustrations
3. Can inequality of opportunity be measured?
   a) A 'canonical model' and allocation rules
   b) One approach to measurement (among many)
   c) Empirical examples: "first generation"
   d) "Second generation" studies
4. Extensions
   a) Intergenerational mobility and economic growth
   b) Impact evaluation and algorithmic design
5. Conclusions

1a. Motivation: normative arguments

In pursuing social justice, what inequalities should you seek to eliminate?

- Is information on the distribution of outcomes enough?
  - John Rawls (1971): *A Theory of Justice* (Harvard University Press)
  - Amartya Sen (1980): "Equality of what?" in McMurrin (ed.), *The Tanner Lectures on Human Values*
  - Ronald Dworkin (1981): "What is Equality? Part 1: Equality of Welfare; Part 2: Equality of Resources", *Philos. Public Affairs*, **10**, pp.185-246; 283-345.
  - Richard Arneson (1989): "Equality of Opportunity for Welfare", *Philosophical Studies*, **56**, pp.77-93.
  - Gerald Cohen (1989): "On the Currency of Egalitarian Justice", *Ethics*, **99**, pp.906-944.

1b. Motivation: evidence on preferences

1. It is now well-established that individuals value 'fairness', in the sense that many are prepared to give up private monetary gains to achieve what they perceive as a just allocation.
   - Offers made and rejected in ultimatum and dictator games.
   - Fehr and Gachter (2000); Fehr and Fischbacher (2003); Henrich et al. (2004)

2. There is also evidence that what is deemed to be a just allocation depends on how it is arrived at!
   - E.g. Cappelen, Sorensen and Tungodden (2010) asked Norwegian business students and alumni to propose a fair distribution of earnings from a typing exercise where "wages" (p) were randomly allocated, work duration (q) was chosen; and "productivity" (a) was measured. Respondents were grouped as follows:

Preference groups	Responsibility sets	Frequency in sample
Strict egalitarians	$\mathcal{R}^{SE} = \emptyset$	0.18
Choice egalitarians	$\mathcal{R}^{CE} = \{q\}$	0.05
Meritocrats	$\mathcal{R}^{M} = \{q, a\}$	0.47
Libertarians	$\mathcal{R}^{L} = \{q, a, p\}$	0.30

1c. Motivation: political salience

"We know that equality of individual ability has never existed and never will, but we do insist that equality of opportunity still must be sought"

(Franklin D. Roosevelt, second inaugural address, 20 January 1937)

"The rise in inequality in the United States over the last three decades has reached the point that inequality in incomes is causing an unhealthy division in opportunities, and is a threat to our economic growth"

(Alan Krueger, Center for American Progress, 12 January 2012)

If these concepts matter for policymakers, can they be rigorously defined and measured?

2. What is equality of opportunity?

An "economic definition"

- Often expressed in terms of adherence to two central principles:

  Principle of compensation: outcome differences due to factors beyond an individual's responsibility ("circumstances") are unfair, and are compensated.

  Principle of reward: outcome differences reflecting differential reward to individual responsibility and effort are ethically legitimate, and are preserved.

- Pioneering economists who adopted and built on these ideas include:

  John Roemer (1993, 1998)

  Dirk van de Gaer (1993)

  Marc Fleurbaey (1994, 2008)

- Building primarily on the Arneson / Cohen "control view" of equality of opportunity.

- Formally, equality of opportunity can be defined as a situation in which the outcome of interest (x) is distributed independently of (predetermined) circumstances (C) for which the individual ought not to be held responsible:

$$F(x|C) = F(x)$$

2. What is equality of opportunity?

An “economic definition”

This approach “… performs for egalitarianism the considerable service of incorporating within it the most powerful idea in the arsenal of the anti-egalitarian right: the idea of choice and responsibility” (Cohen, 1989, p.993)

2. What is equality of opportunity?

An "economic definition"

This approach "… performs for egalitarianism the considerable service of incorporating within it the most powerful idea in the arsenal of the anti-egalitarian right: the idea of choice and responsibility" (Cohen, 1989, p.993)

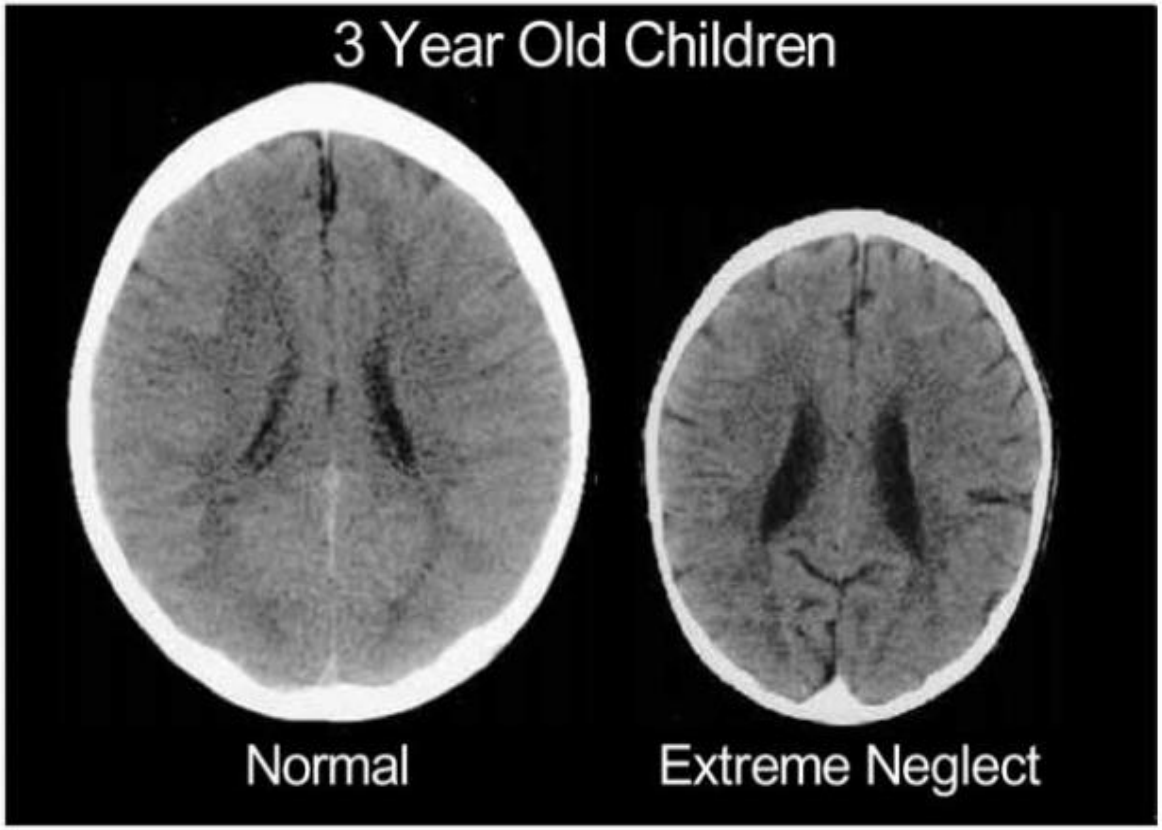
But it also expands the scope for normative judgement inherent in inequality analysis. Recall Atkinson (1970): "…this direct approach […] serves to emphasize that *any* measure of inequality involves judgements about social welfare" (p.257, emphasis in the original)

Beside normative views on how transfers in different parts of the distribution should be weighted against one another, I.Op. requires normative judgment also on what factors should be considered "circumstances".

2. What is (in)equality of opportunity?

Illustrations

Brain scan for a 3 year-old with median cognitive skills



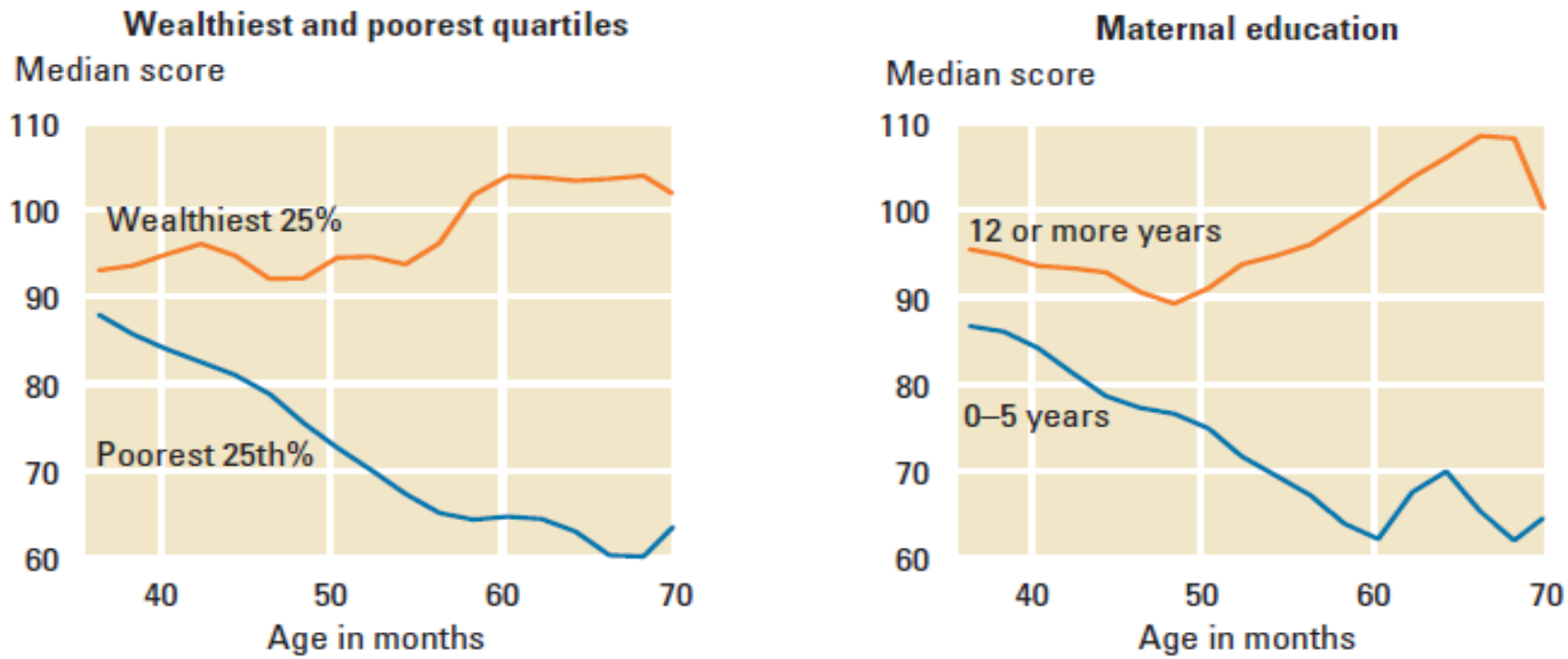
Brain scan of a 3 year-old with acute developmental gaps (language, tact, and social interactions)

Source: Perry (2002): "Childhood Experience and Expression of Genetic Potential", Brain and Mind

2. What is (in)equality of opportunity?

Illustrations

Figure 2 Opportunities are determined early
Cognitive development for children ages three to five in Ecuador differs markedly across different family backgrounds



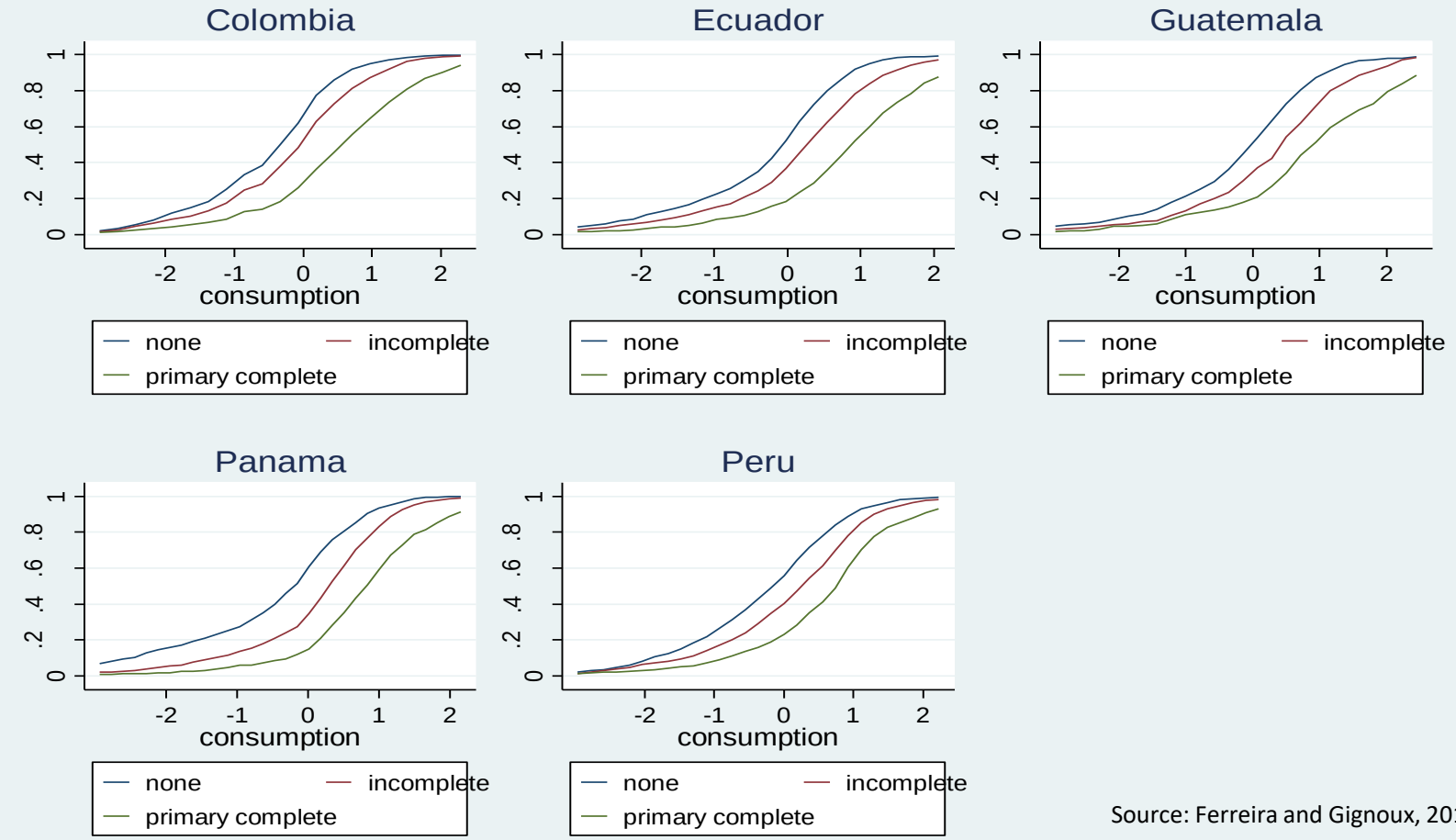
Source: Paxson and Schady (2005a).
Note: Median values of the test of vocabulary recognition (TVIP) score (a measure of vocabulary recognition in Spanish, standardized against an international norm) are plotted against the child's age in months. The medians by exact month of age were smoothed by estimating fan regressions of the median score on age (in months), using a bandwidth of 3.

Source: Paxson and Schady, *JHR*, 2007

2. What is (in)equality of opportunity?

Illustrations

Distribution of p.c.h. consumption conditional on mother's education, in five LAC countries



Source: Ferreira and Gignoux, 2011

3. Can inequality of opportunity be measured?

A simple "canonical" model

- Let each and every individual be fully characterized by the triple *(x, C, e).*
- Let all elements of the vector ***C***, as well as *e*, be discrete.
- Let $x_{ij} = g(C_i, e_j)$
- Let a *type* consist of all individuals with identical circumstances
- Let a *tranch* consist of all individuals with identical effort levels
- Let there be *n* types and *m* tranches
- Then the population can be represented by the *n x m* matrix $[X_{ij}]$ below.
- To $[X_{ij}]$, let there be associated another *n x m* matrix $[P_{ij}]$, whose elements p_{ij} denote the proportion of the total population with circumstances C_i and effort level e_j.

Discussion draws on Ferreira and Peragine (2016)

3a. A ‘canonical model’

Table 1

	e_1	e_2	e_3	…	e_m
C_1	x_{11}	x_{12}	x_{13}	…	x_{1m}
C_2	x_{21}	x_{22}	x_{23}	…	x_{2m}
C_3	x_{31}	x_{32}	x_{33}	…	x_{3m}
…	…	…	…	…	…
C_n	x_{n1}	x_{n2}	x_{n3}	…	x_{nm}

Discussion draws on Ferreira and Peragine (2016)

3a. A ‘canonical model’

A tranch

Table 1

A type

	e_1	e_2	e_3	…	e_m
C_1	x_{11}	x_{12}	x_{13}	…	x_{1m}
C_2	x_{21}	x_{22}	x_{23}	…	x_{2m}
C_3	x_{31}	x_{32}	x_{33}	…	x_{3m}
…	…	…	…	…	…
C_n	x_{n1}	x_{n2}	x_{n3}	…	x_{nm}

Discussion draws on Ferreira and Peragine (2016)

3b. Scalar measurement

In essence, the measurement of inequality of opportunity can be thought of as a two-step procedure:

- first, the actual distribution $[X_{ij}]$ is transformed into a counterfactual distribution $[\widetilde{\boldsymbol{X}}_{ij}]$ that reflects *only and fully* the unfair inequality in $[X_{ij}]$, while all the fair inequality is removed.
- In the second step, a measure of inequality – satisfying the usual axioms - is applied to $[\widetilde{\boldsymbol{X}}_{ij}]$.

- The first step can be taken in many different ways, depending on the specific form of the compensation (e.g. ex-ante vs. ex-post) and reward principles one chooses to adopt.
  - In what follows I focus on a single example, an ex-ante approach known as between-types inequality.

Discussion draws on Ferreira and Peragine (2016)

3b. Scalar measurement

The between-types approach

Between types ***(*** $\tilde{X}_{BT}$ ***)***: *For all* $j \in \{1,\ldots,m\}$ *and for all* $i \in \{1,\ldots,n\}$, $\tilde{x}_{ij} = \mu_i$.

Table 2: Between-types inequality *(n=m=3)*

	e1	e2	e3
C1	μ_1	μ_1	μ_1
C2	μ_2	μ_2	μ_2
C3	μ_3	μ_3	μ_3

Draws on the min of means approach.

Discussion draws on Ferreira and Peragine (2016)

3b. Scalar measurement

The between-types approach

- This approach - $I(\tilde{x}_{BT})$ - has been applied sufficiently widely so as to permit international comparisons.

- There are two versions of this index, both of which are subject to a downward bias due to the partial observability of circumstances. Using a slightly different notation:

- IOL: $\theta_a = I(\tilde{x}_{BT})$

- IOR: $\theta_r = \dfrac{I(\tilde{x}_{BT})}{I(x)} = \dfrac{I(\mu_i)}{I(x_{ij})}$

- These indices can be computed non-parametrically (Checchi and Peragine, 2010), but precision weakens and biases arise when *n* is large relative to the sample size.
- Common parametric approximation:

$$y = C\psi + \varepsilon \qquad \tilde{\mu}_i = C_i\hat{\psi} \qquad IO_p = \frac{I(\tilde{\mu}_i)}{I(y)}$$

3c. An application to Latin America

TABLE 1

HOUSEHOLD SURVEY NAMES, DATES, AND SAMPLE SIZES

	Brazil	Colombia	Ecuador	Guatemala	Panama	Peru
Survey	PNAD 1996	ECV 2003	ECV 2006	ENCOVI 2000	ENV 2003	ENAHO 2001
Sample of 30 to 49 year-olds	85,692	22,517	12,650	6,956	6,339	17,030
Sample of heads and spouses, aged 30 to 49 years	73,847	18,069	10,719	6,067	5,105	13,947
Of those, observations with income/consumption and circumstances	70,521	17,979	10,719	5,988	4,556	13,621
(share of original sample)	0.823	0.798	0.847	0.861	0.719	0.800

Source: Ferreira and Gignoux, 2011

3c. An application to Latin America

TABLE 3

DEFINITION OF CIRCUMSTANCE VARIABLES, BY COUNTRY

	Brazil	Colombia	Ecuador	Guatemala	Panama	Peru
Ethnicity						
Category 1	Self reported white ethnicity	Other	Self-reported ethnicity: white, mixed blood ("mestizo") or other	European maternal language	Other	European maternal language
Category 2	Self reported black ("negro") and mixed blood ("pardo") ethnicity	Self-reported minority ethnicity: "indígena, gitano, archipiélago o palenquero"	Self-reported ethnicity: indigenous, black ("negro" or "mulato")	Indigenous maternal language	Speaks indigenous language	Indigenous maternal language
Father's occupation						
Category 1	Agricultural worker	Missing	Agricultural worker or domestic worker	Agricultural worker	Agricultural worker	Missing
Category 2	Other		Other	Other	Other	
Mother's and father's education						
Category 1	None or unknown	None or unknown	None or unknown	None or unknown	None or unknown	None or unknown
Category 2	Completed grade 1 to 4	Primary incomplete	Primary	Primary incomplete	Primary	Primary incomplete
Category 3	Completed grade 5 or more	Primary complete or more	Secondary or more	Primary complete or more	Secondary or more	Primary complete or more
Birth region						
Category 1	Sao Paulo and Federal district	Departments at the periphery	Sierra and Amazonia provinces	Guatemala City, Northeast departments and El Petén	Cities and intermediate urban centers	Inland non-southern departments
Category 2	South East, Center-West, and South	Central departments(a)	Costa and Insular provinces	North and Northwest departments	Other urban centers	Southern and other costal departments
Category 3	North-East, North or missing	Bogota, San Andres, and Providencia islands and foreign country	Pichincha province (with Quito) and Azuay province	Southeast, Southwest, and Center departments	Rural areas	Arequipa, Callao, and Lima

Note: Central departments are Boyaca, Caldas, Caqueta, Cundinamarca, Huila, Meta, Norte de Santander, Quindio, Risaralda, Santander, Tolima, and Valle del Cauca.

Source: Ferreira and Gignoux, 2011

3c. An application to Latin America

TABLE 6

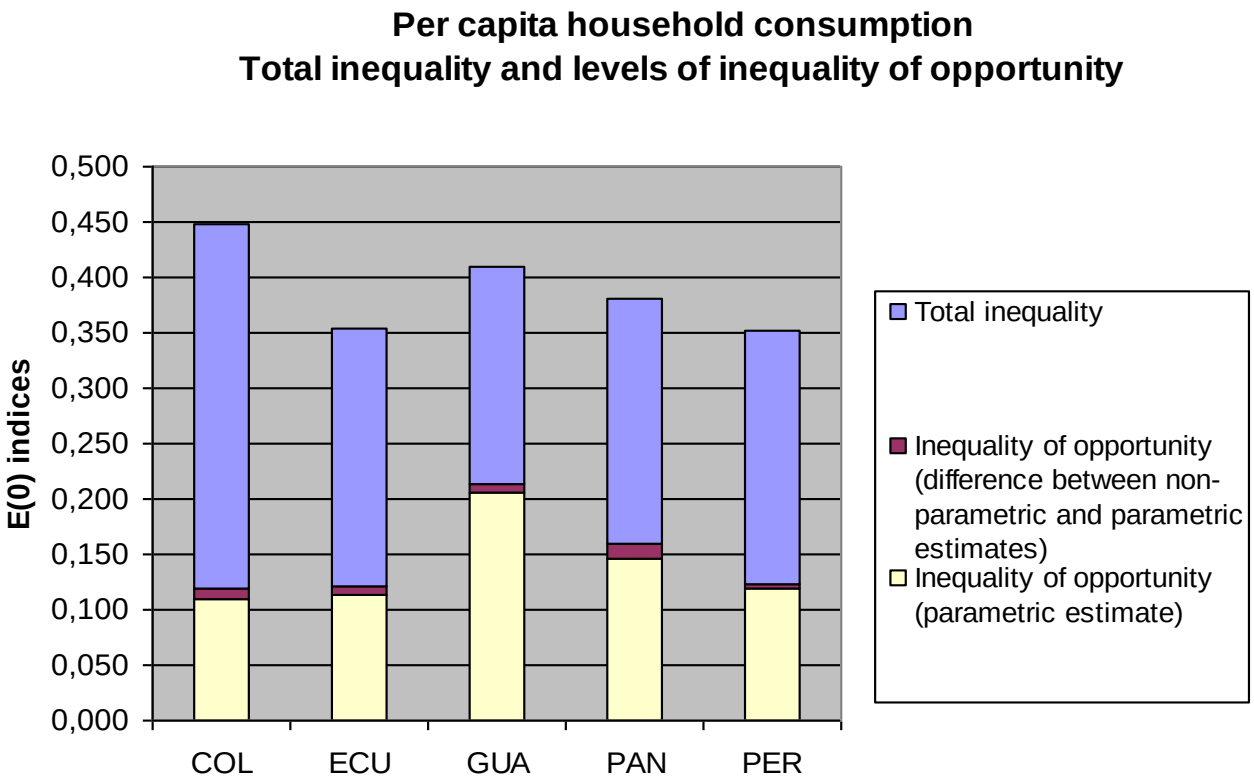
SCALAR INDICES OF INEQUALITY OF OPPORTUNITY

	Brazil	Colombia	Ecuador	Guatemala	Panama	Peru
Panel A: Household income (per capita)						
Total inequality (E_0)	0.692	0.572	0.580	0.593	0.630	0.557
	(0.013)	(0.033)	(0.028)	(0.036)	(0.029)	(0.022)
Non-parametric estimates						
IOL	0.227	0.144	0.164	0.213	0.213	0.163
	(0.008)	(0.023)	(0.022)	(0.031)	(0.024)	(0.015)
IOR	0.329	0.252	0.283	0.359	0.338	0.293
	(0.008)	(0.026)	(0.023)	(0.030)	(0.026)	(0.018)
Parametric estimates						
IOL	0.223	0.133	0.150	0.199	0.190	0.156
	(0.008)	(0.019)	(0.020)	(0.028)	(0.023)	(0.014)
IOR	0.322	0.232	0.259	0.335	0.301	0.279
	(0.008)	(0.023)	(0.023)	(0.030)	(0.028)	(0.018)
Panel B: Household consumption expenditures (per capita)						
Total inequality (E_0)		0.462	0.359	0.415	0.381	0.351
		(0.018)	(0.015)	(0.025)	(0.018)	(0.013)
Non-parametric estimates						
IOL		0.123	0.124	0.221	0.156	0.123
		(0.015)	(0.013)	(0.024)	(0.016)	(0.010)
IOR		0.265	0.346	0.532	0.409	0.351
		(0.021)	(0.021)	(0.023)	(0.025)	(0.018)
Parametric estimates						
IOL		0.114	0.117	0.213	0.144	0.119
		(0.014)	(0.012)	(0.022)	(0.015)	(0.009)
IOR		0.247	0.326	0.514	0.377	0.339
		(0.021)	(0.022)	(0.022)	(0.026)	(0.017)

Notes: Sample: household heads and spouses, aged 30–49, with positive income and information on a set of circumstances; bootstrap standard errors (taking into account stratification and clustering) in parentheses; father's occupation missing for Colombia and Peru.

3c. An application to Latin America

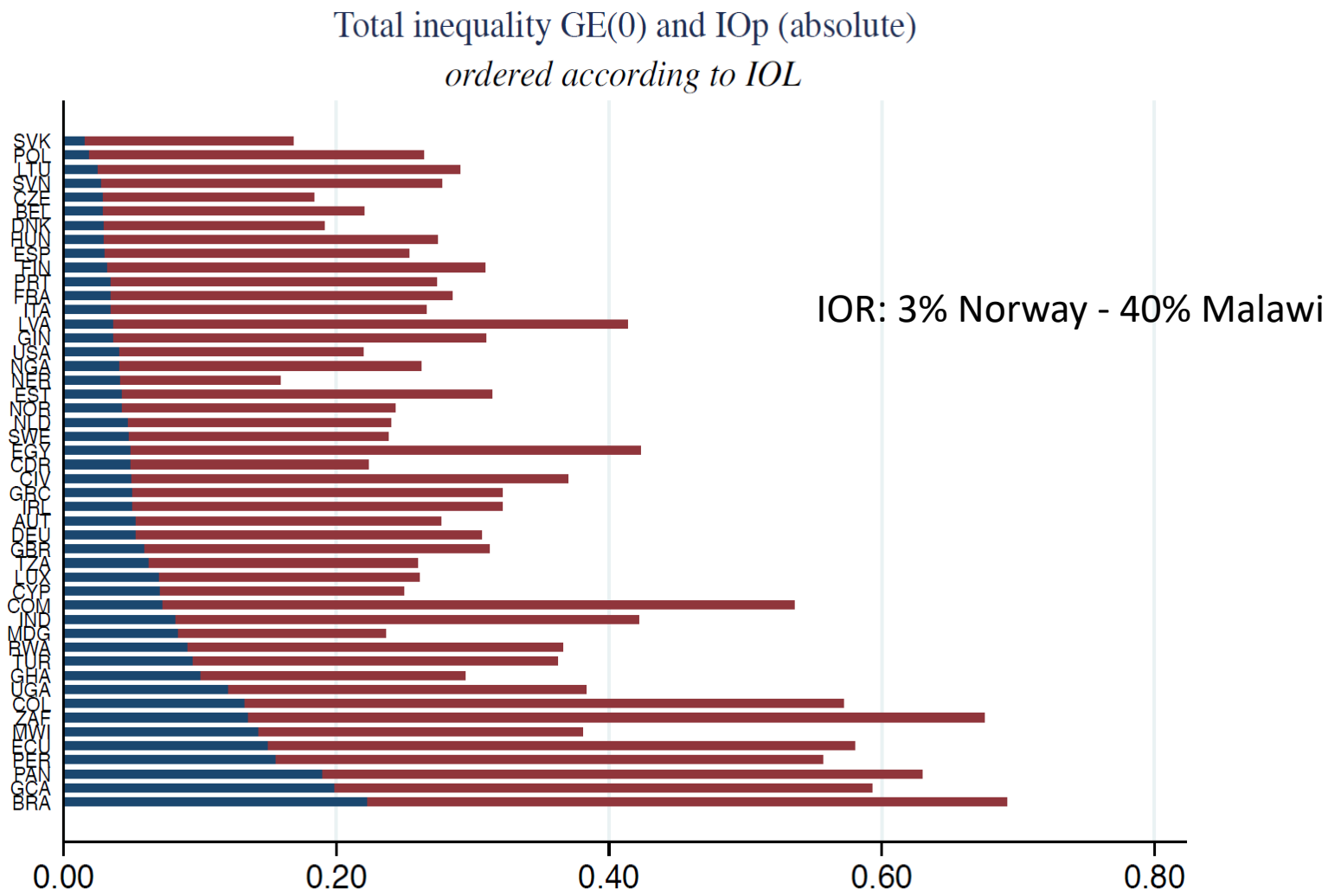
1. In Latin America, inequality of economic opportunity:
   - ranges from 23% to 35% for income per capita.
   - ranges from 24% to 50% for consumption per capita.



Source: Ferreira and Gignoux, 2011

3c. A broader international comparison

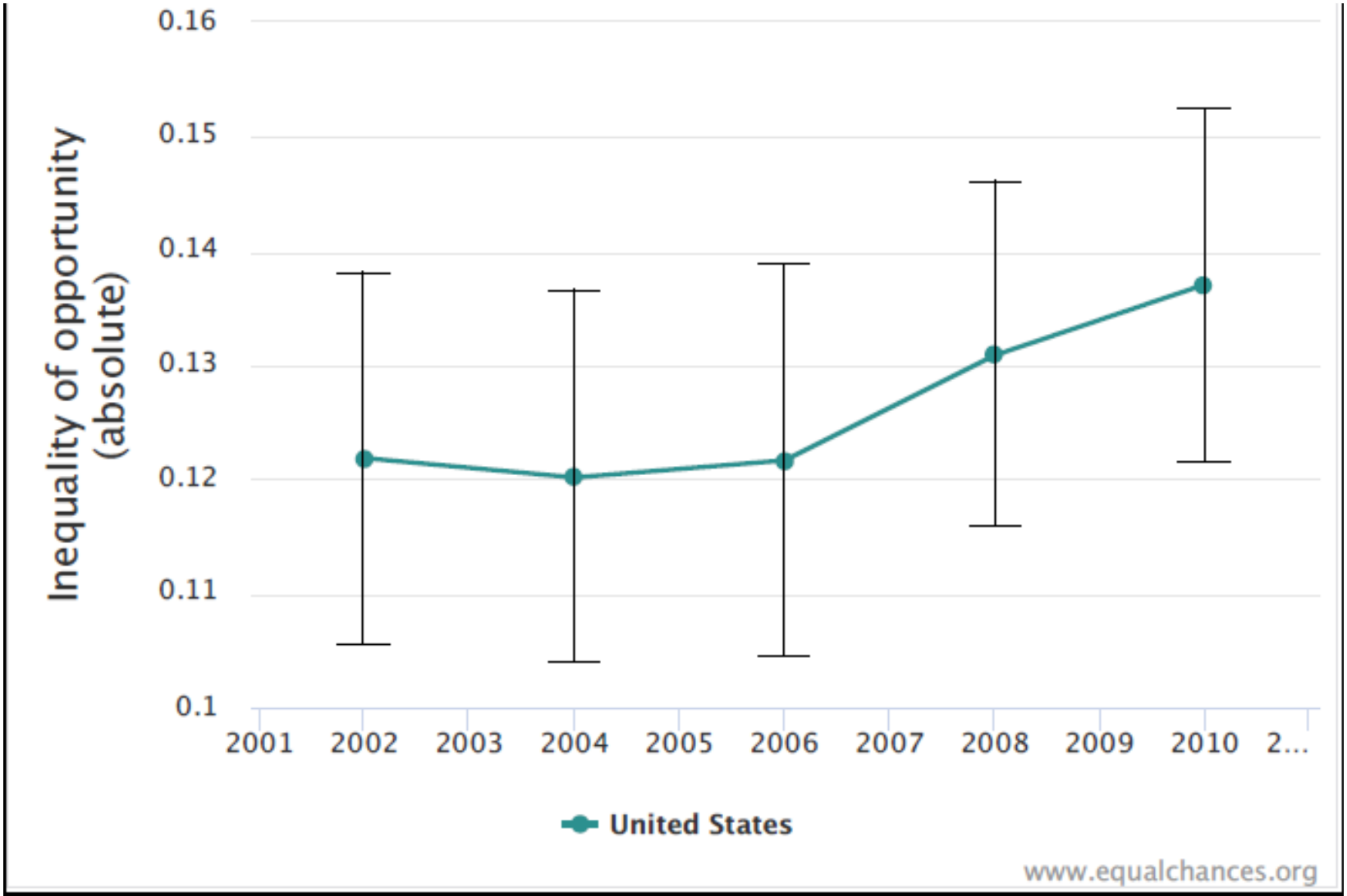
Drawing on 8 "first-generation" studies covering 51 countries; x = income



Source: Brunori, Ferreira, Peragine (2013) – with updates

3c. Recent evolution of I.Op. in the US

An example from the Equalchances database



3d. "Second generation" studies

1. 'Second-generation' between-types approach: looking for upper-bound estimates (Niehues and Peichl, SCW 2014)

- Two-stage estimator using panel data:

  i. Estimate $\ln w_{it} = \beta E_{it} + c_i + u_t + \varepsilon_{it}$

  ii. Back in cross-section, estimate $\ln w_{is} = \varphi \hat{c}_i + v_{it}$

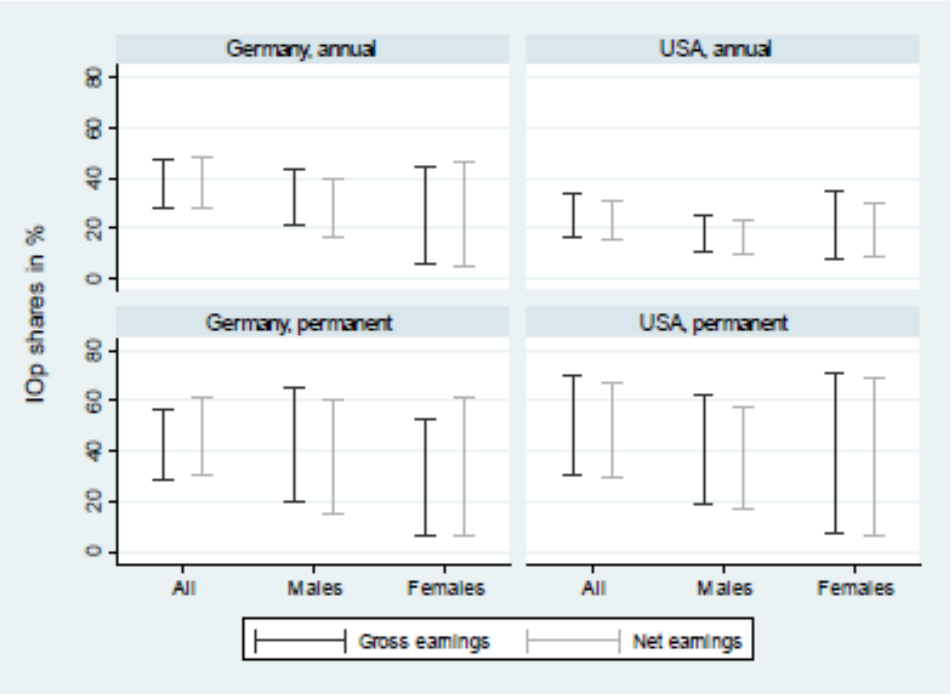
  Construct $\tilde{\mu}^{UB} = \exp(\hat{\varphi}\widehat{c_i} + \sigma^2/2)$

  – Application to Germany (SOEP) and the US (PSID), for both current and permanent incomes

3d. “Second generation” studies

1. ‘Second-generation’ between-types approach: looking for upper-bound estimates (Niehues and Peichl, SCW 2014)

Figure 2: IOp shares in outcome inequality



Source: Own calculations based on SOEP and PSID. The two graphs on the top ustrate IOp shares in annual incomes (2009 for Germany, 2007 for the US); the grapl at the bottom IOp shares in permanent incomes.

3d. “Second generation” studies

2. ‘Second-generation’ between-types approach: enlarging the circumstance set through admitting an “age of consent” (Hufe, Peichl, Roemer and Ungerer; SCW 2017)
   - Use National Longitudinal Survey of Youth (NLSY -79) for the US and British Cohort Study (BCS – 70) for the UK

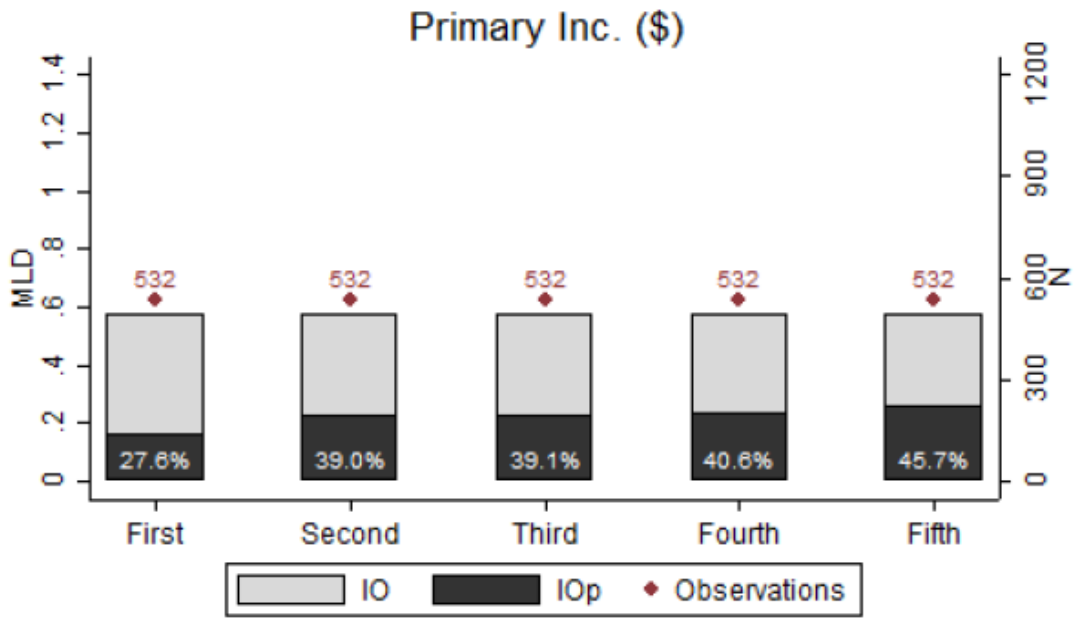
Scenario						Circumstance Set	Circumstance Var.
Sixth	Fifth	Fourth	Third	Second	First	Base	Sex, Country of Birth, Ethnic Affiliation, Cohort, Age, Academic Achievement Mother, Occupation Code Mother, Rural/Urban, Height (16), Family Income
						Ability	PIAT Math, PIAT Reading
						Behavioral Problems	Behavioral Problems Index (BPI)
						Child-Parent Relationship	Play/Schoolwork w/ Parents, Perceived Quantity of Time w/ Parents, Parents Split, Parental Income
						Health-Related Behavior	Smoking Habits Mother, Drinking Habits Mother, Health Restrictions Child
						Survey Specifics	Specific to NLSY79 and BCS70. See text for more information.

Table 1: Overview of Circumstance Scenarios

3d. "Second generation" studies

2. Hufe, Peichl, Roemer and Ungerer (2017) find that the lower-bound IOR can be as high as 45% in the US and 31% in the UK when using this extended circumstance set.

Figure 2: IOp with varying circumstance sets (NLSY79), comparable sample, average income



Note: The overall bar yields the extent of outcome inequality IO. The black colored share of each bar yields inequality attributed to circumstances, i.e. the lower bound absolute measure of inequality of opportunity IOp. The residual gray colored share of each bar can be interpreted as an upper bound measure of inequality attributed to differential efforts. The white labels at the bottom of each bar indicate the share of IOp in IO, i.e. the relative measure of inequality of opportunity r.

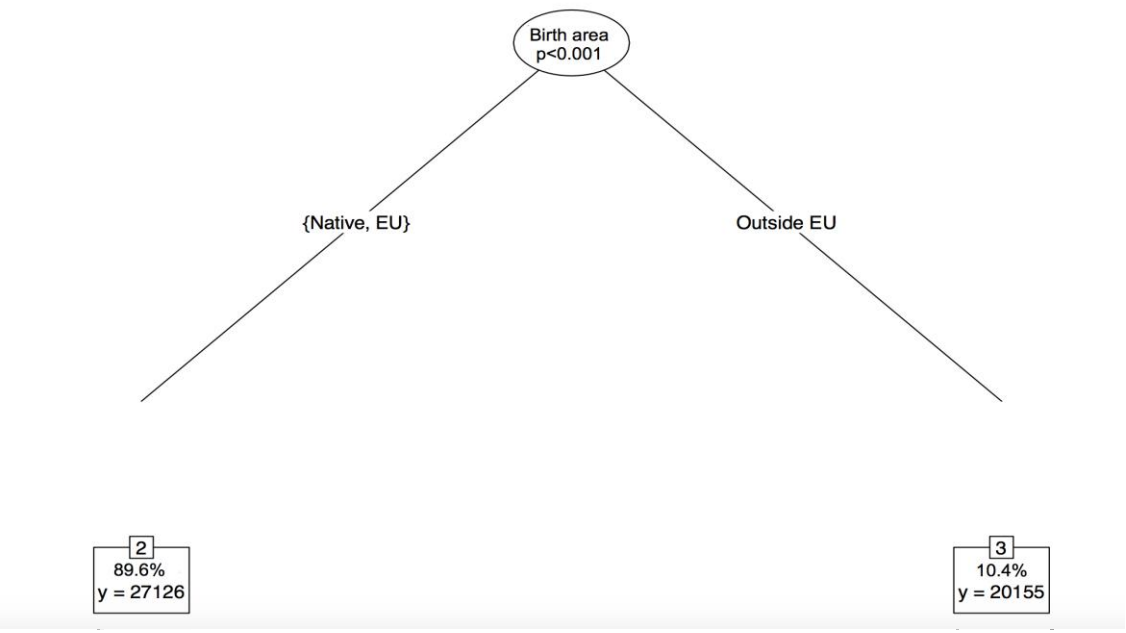
3d. "Second generation" studies

3. An alternative approach to "let the data choose the model specification" is proposed by Brunori, Hufe and Mahler (2018), using conditional inference trees and forests.
   - A conditional inference tree consists of a set of terminal nodes (leaves) obtained by recursive binary splitting, as follows.
   - Given a set of circumstance variables and categories, the algorithm splits the sample in all possible partitions [C], and computes the p-value for the null hypothesis that the statistic of interest (e.g. the mean) in the two sub-samples is identical.
   - [C]* is chosen as $[C]^* = \{[C]: argmin\ p_{adj}^{[C]}$, where the adjustment is a Bonferroni correction (for multiple hypothesis testing).
   - A critical significance level α can be chosen so that if $p_{adj}^{[C]*} > \alpha$, one exit the algorithms, and otherwise [C]* is chosen as splitting variable.
   - Repeat the algorithm for each node (sub-sample), until one has exited everywhere.
   - A conditional inference forest is basically a set of trees estimated on random subsamples of the original data, in each case using a different subset of circumstance variables. The size of the subsets of circumstances is chosen by minimizing the "out-of-the-bag" MSE.

3d. “Second generation” studies

3. Although forests outperform trees in terms of out-of-sample prediction, trees can be visually informative of the ‘structure’ of inequality of opportunity in different countries.

Figure 3: Opportunity Tree: Sweden

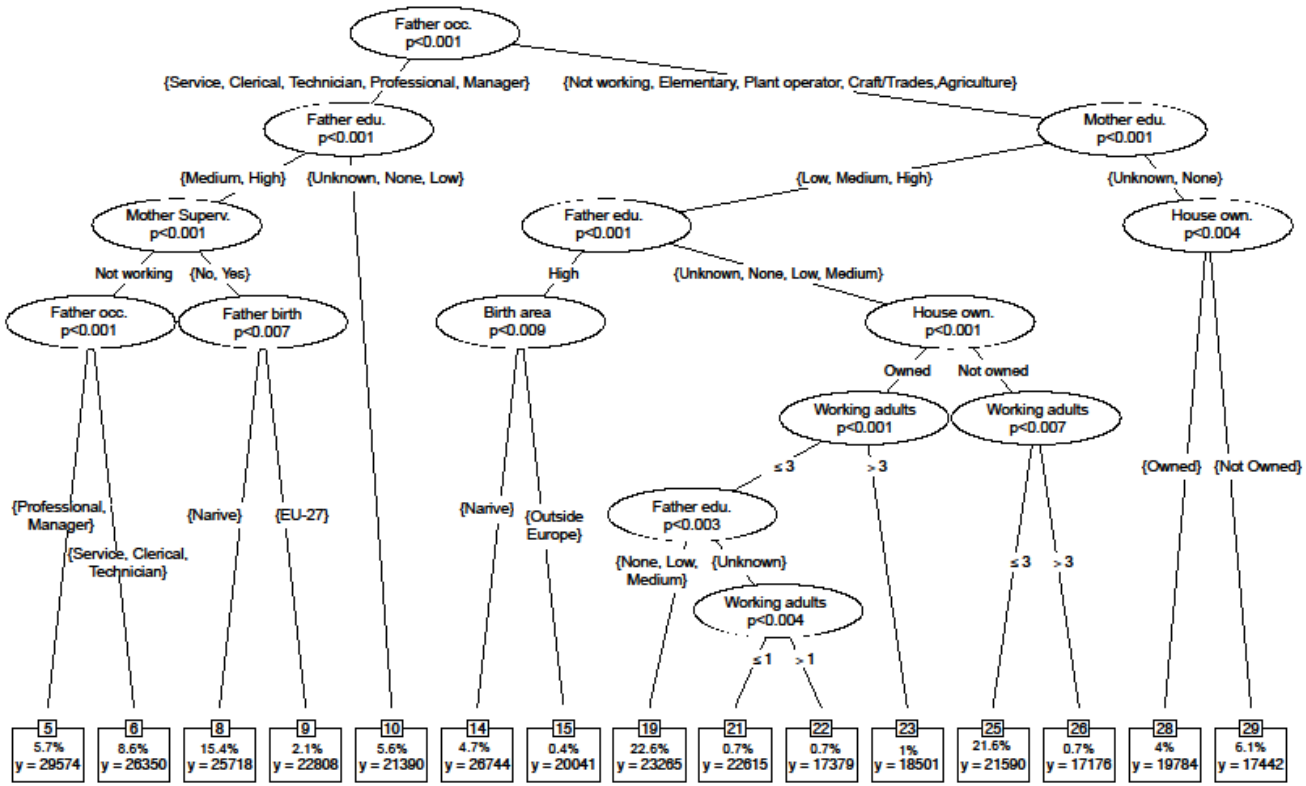


Note: Opportunity tree for Sweden. White rectangular boxes indicate terminal nodes. The first number inside the rectangular boxes indicates the share of the population belonging to this group, while the second number indicates the predicted income.

3d. “Second generation” studies

3. Although forests outperform trees in terms of out-of-sample prediction, trees can be visually informative of the ‘structure’ of inequality of opportunity in different countries.

Figure 4: Opportunity Tree: Germany

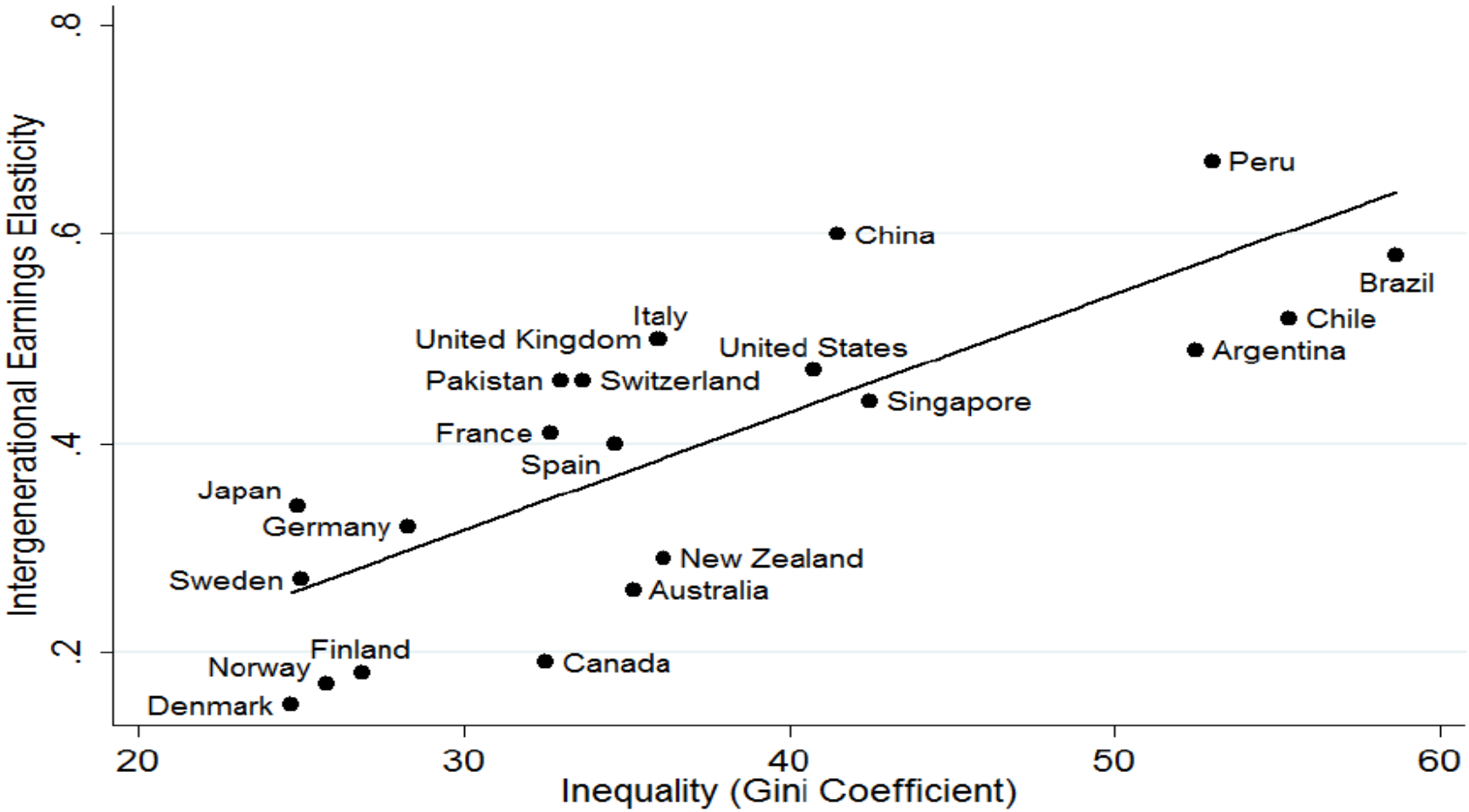


4. Extensions: I.Op. and IGM

- IGM: $y_c = \beta y_p + \varepsilon$ R^2
- IOp: $y = C\psi + \varepsilon$ $IOR = \frac{I(\tilde{\mu}_i)}{I(y)}$
- Inequality of opportunity (at least in the ex-ante approach) is very close to origin-independent measures of IGM.
  - Recent literature on the US (building on Chetty, Hendren, Kline and Saez, QJE 2014) almost seems to equate the two concepts.
  - Difference: more circumstances
    - In "steady state" $(Var\ y_c = Var\ y_p)$, Chetty et al.'s IGE of 0.45 implies $R^2 = 0.2$
  - Omitted variables: I.Op. is explicitly not a causal estimate for any individual circumstance.

4. Extensions: I.Op. and IGM

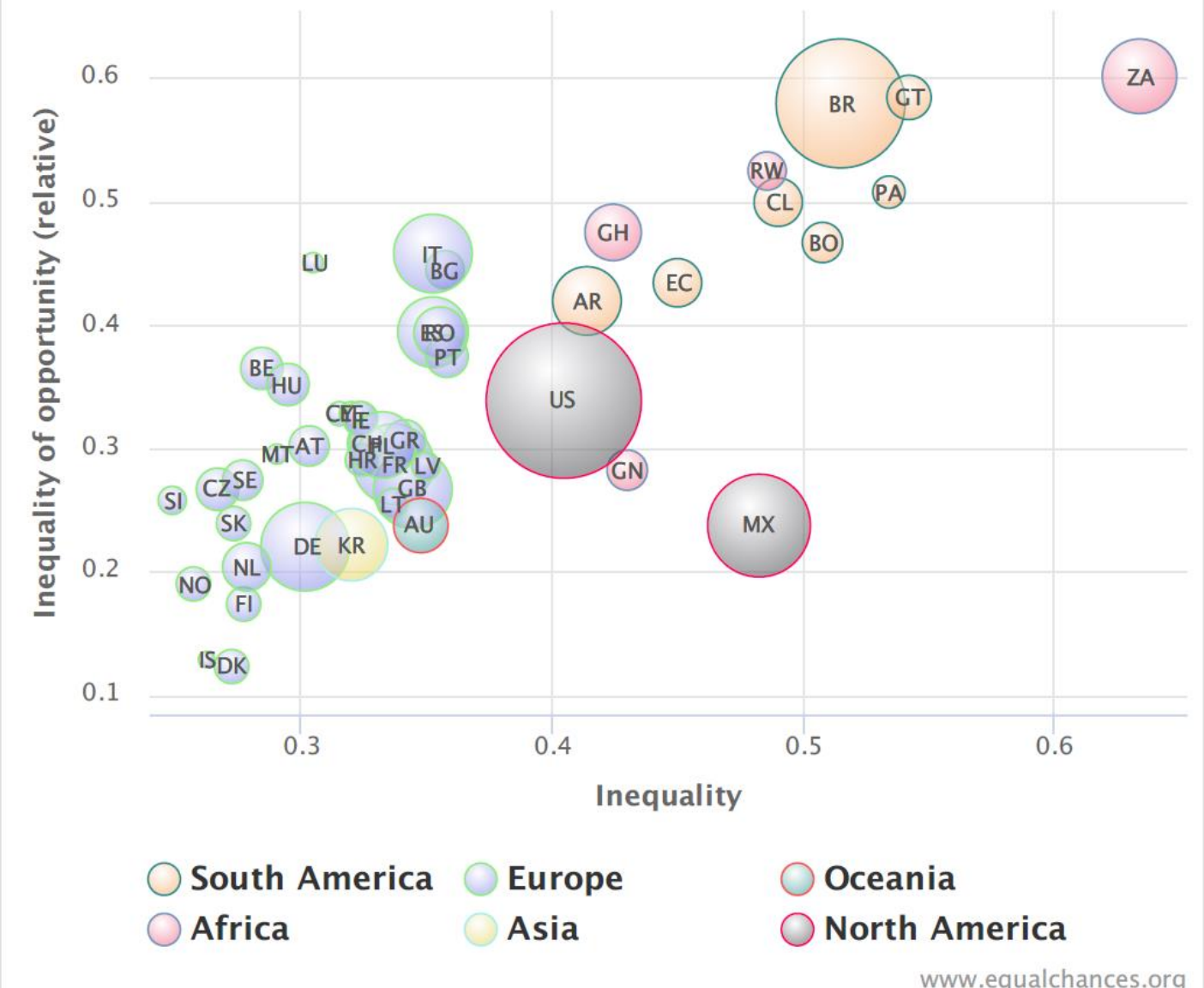
The Great Gatsby Curve



Source: Corak (2012)

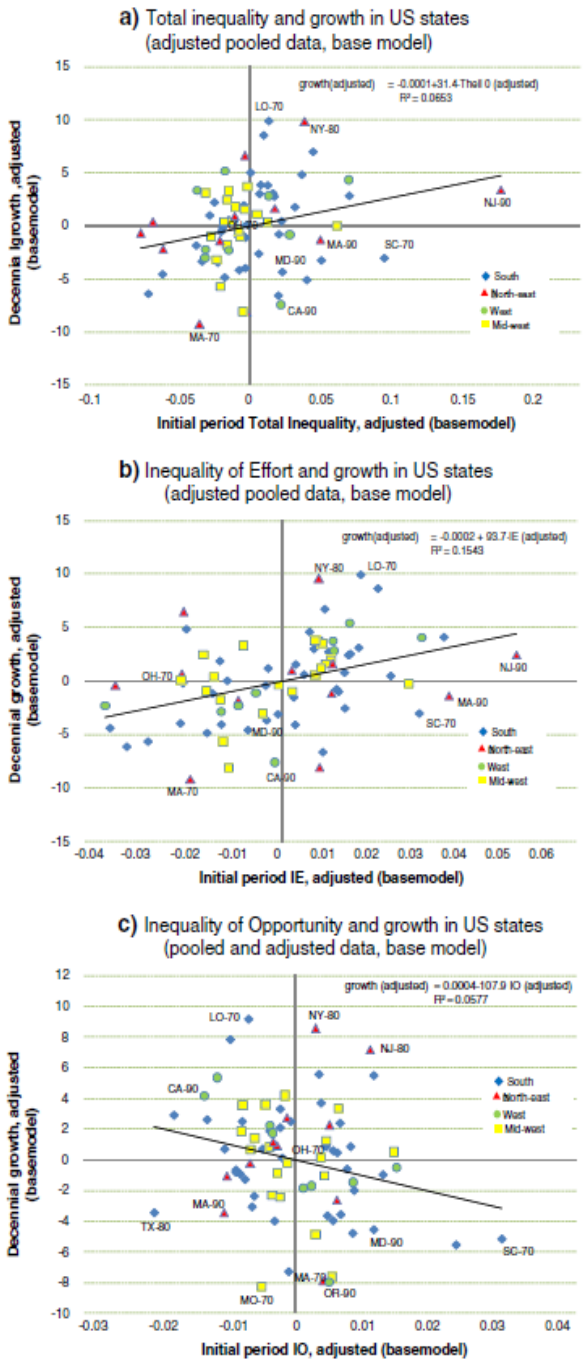
4. Extensions: I.Op. and IGM

Inequalities of outcome and opportunity: strong correlation



4. Extensions: I.Op. and economic growth

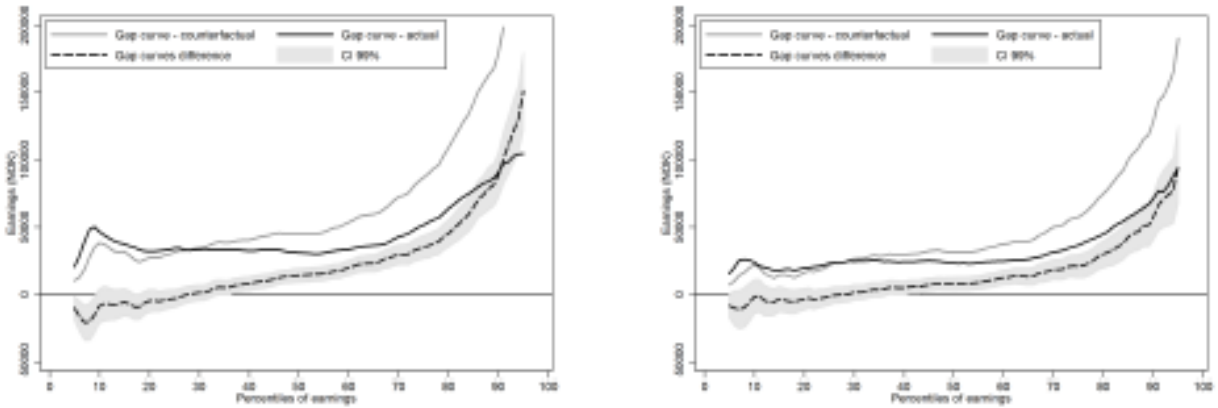
- Beyond measurement: once quantified, the variable can be used in causal analysis.
- A possible answer to the lengthy and inconclusive literature on inequality and growth.
  - Marrero and Rodriguez (2013) for the US.
  - Ferreira, Lakner, Lugo, Özler (2018) – cross country
- Key challenge: comparable data on advantages and circumstances



4. Extensions: I.Op. and policy evaluation

- A growing number of studies evaluates policies and programs in terms of their impacts not only on means, or on the outcome distribution, but also on the distribution of opportunities (e.g. on $[\widetilde{\boldsymbol{X}}_{ij}]$).
  - i. Impact of Progresa on family types defined by indigenous status and parental education (van de Gaer et al., *WBER* 2014)
  - ii. Impact of Progresa on ECD indicators (e.g. PPVT, child behavior checklist, etc.) for children from different family types (Figueroa, *Soc. Sci.& Med*. 2014)
  - iii. Impact of a child care reform in Norway, using gap curves and QTEs. (Andreoli et al., *REStat*, 2018)

B - Gap curves



(e) Lower vs upper class (f) Middle vs upper class

4. Extensions: I.Op. and algorithmic design in machine learning

Digital Society Initiative, University of Zurich

Algorithmic fairness, discrimination, and equality of opportunity

Michele Loi

Fourteenth Winter School
Inequality and Social Welfare Theory
Canazei
University of Verona

algorithms can be 'sexist'

CNBC

MENU | MARKETS | BUSINESS | INVESTING | TECH | POLITICS | CNBC TV

RETAIL

APPAREL | DISCOUNTERS | DEPARTMENT STORES | E-COMMERCE | FOOD AND BEV

Amazon scraps a secret A.I. recruiting tool that showed bias against women

- Amazon.com's machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.
- The team had been building computer programs since 2014 to review job applicants' resumes with the aim of mechanizing the search for top talent, five people familiar with the effort told Reuters.
- The company's experimental hiring tool used artificial intelligence to give job candidates scores ranging from one to five stars — much like shoppers rate products on Amazon, some of the people said.

Published 6:15 AM ET Wed, 10 Oct 2018 | Updated 2:25 PM ET Thu, 11 Oct 2018

REUTERS

models from machine learning can be 'racist'

arXiv.org > cs > arXiv:1301.6822

Computer Science > Information Retrieval

Discrimination in Online Ad Delivery

Latanya Sweeney

(Submitted on 29 Jan 2013)

Google personalised ad for public records

Trevor John **Trevor John, Arrested?**

Source: Michele Loi's lecture at ISWT14, Canazei, January 8, 2019

A Moral Framework for Understanding Fair ML through Economic Models of Equality of Opportunity

Hoda Heidari
ETH Zürich
hheidari@inf.ethz.ch

Michele Loi
University of Zürich
michele.loi@uzh.ch

Krishna P. Gummadi
MPI-SWS
gummadi@mpi-sws.org

Andreas Krause
ETH Zürich
krausea@ethz.ch

The equality of opportunity idea.

1. Moritz Hardt

Hardt, Moritz, Eric Price, and Nathan Srebro. 2016. "Equality of Opportunity in Supervised Learning." *ArXiv:1610.02413 [Cs]*, October. http://arxiv.org/abs/1610.02413.

Equal odds for binary classifiers:

$$\Pr\left\{\widehat{Y} = 1 \mid A = 0, Y = y\right\} = \Pr\left\{\widehat{Y} = 1 \mid A = 1, Y = y\right\}, \quad y \in \{0, 1\}$$

Prediction and protected variable (A) are independent conditional on Y (actual label)

Source: Michele Loi's lecture at ISWT14, Canazei, January 8, 2019

5. Conclusions

- There is reason to think of IOp as the "active ingredient of inequality", both intrinsically and instrumentally.
  - Arguably it focuses on the inequalities that matter most.
- Inequality of opportunity can be measured rigorously, but:
  - Not precisely: lower and upper bounds, because of partial observability of circumstances.
  - The broader information requirements place considerable demands on data (particularly on circumstance variables)
  - Contingency on normative judgments (Atkinson, 1970) is greatly amplified: Not only can different indices be used to measure inequality on $[\widetilde{\boldsymbol{X}}_{ij}]$, but the matrix itself can be constructed in different ways.
- Current empirical estimates range from 3% of total inequality to lower bounds as high as 51% (in Guatemala), and upper bounds above 60% (in the US and Germany!)
- Increasingly used in empirical analysis, and apparently also in the design of machine learning tools used both by governments and companies

5. Conclusions

A number of challenges remain…

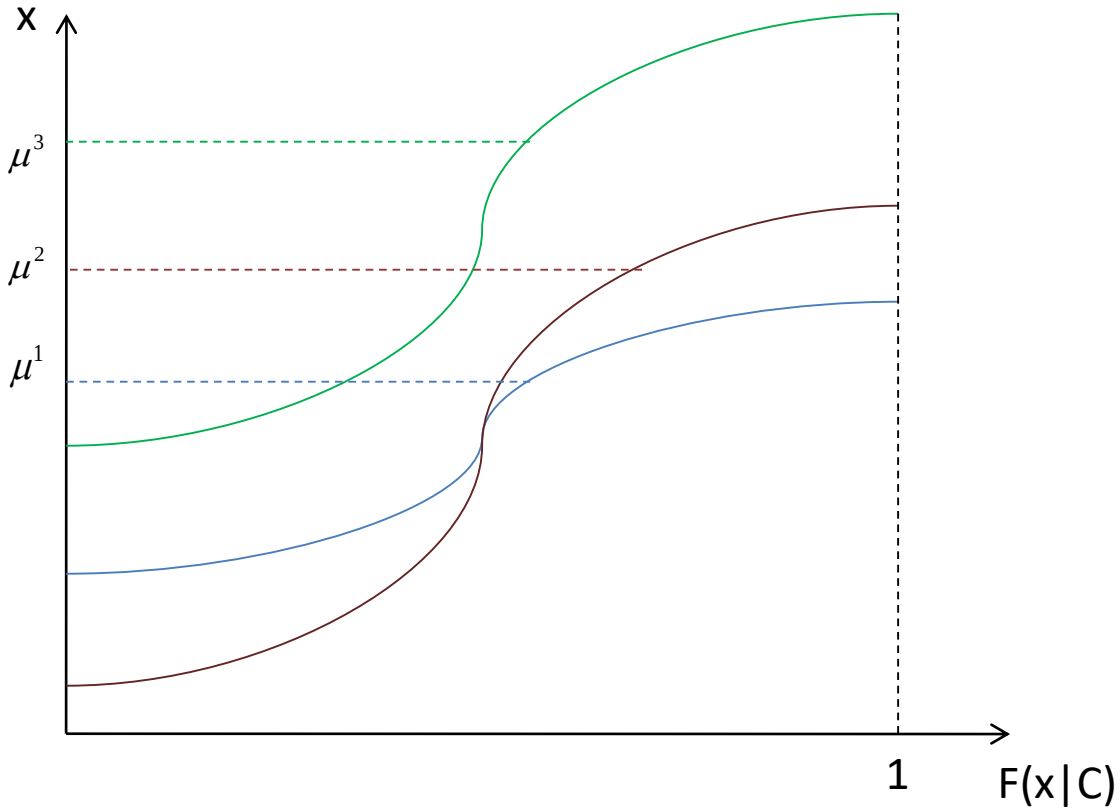
1. Conceptually:
   - Clashes between different versions of the compensation and reward principles (largely absent from our discussion today)
   - The materialist causal thesis and the "incompatibilist" view: "There is no such thing as free will; hence all inequality is I.Op."
   - Requires a discussion of what society chooses to classify as circumstances, particularly as data on (epi)genetics become more widely available.
     - Ways forward: 'Bergen experiments'? Rawlsian 'Reflective equilibrium"?
2. Empirically, because of data limitations:
   - Partial observability of circumstances (downward bias)
   - Sample size limits and sampling variation (upward bias)

But progress is being made on many fronts!

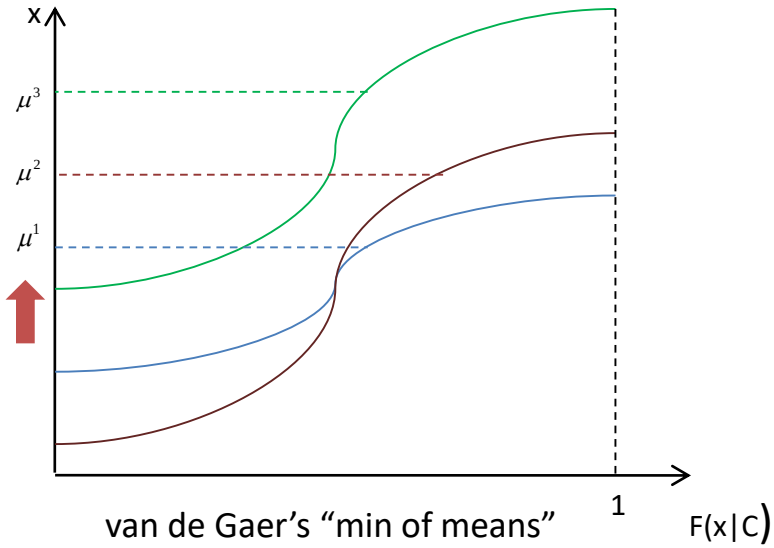
Thank you!

Appendix: A ‘canonical model’

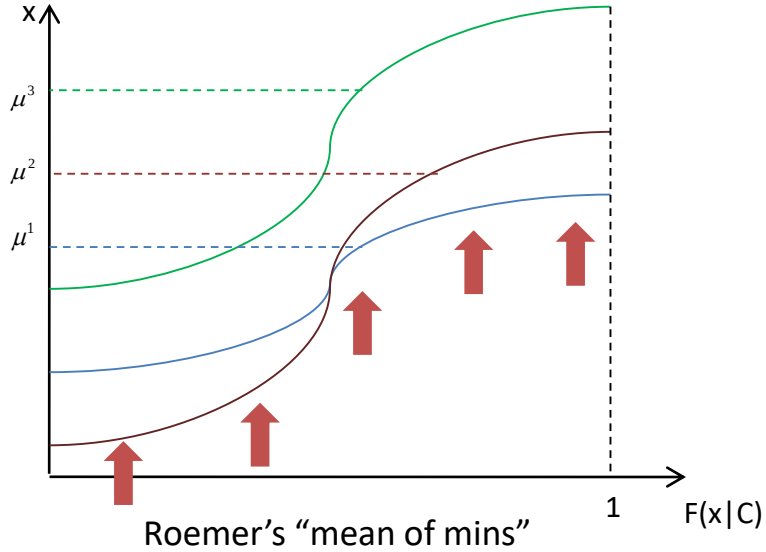
When effort is continuous, n=3



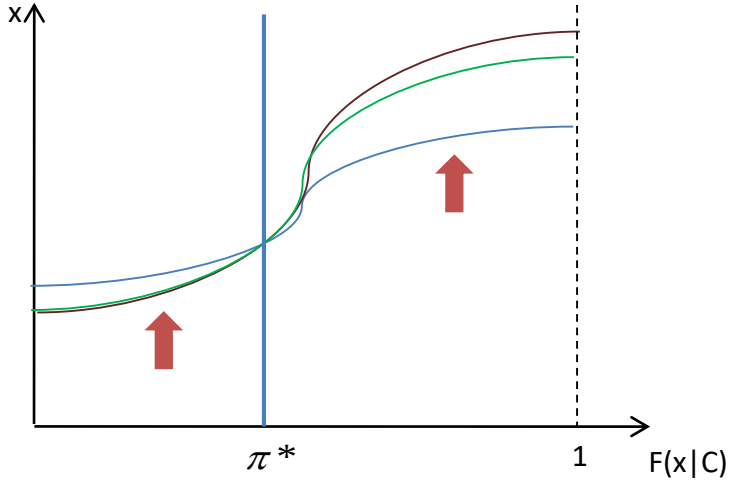
Appendix: Allocation rules



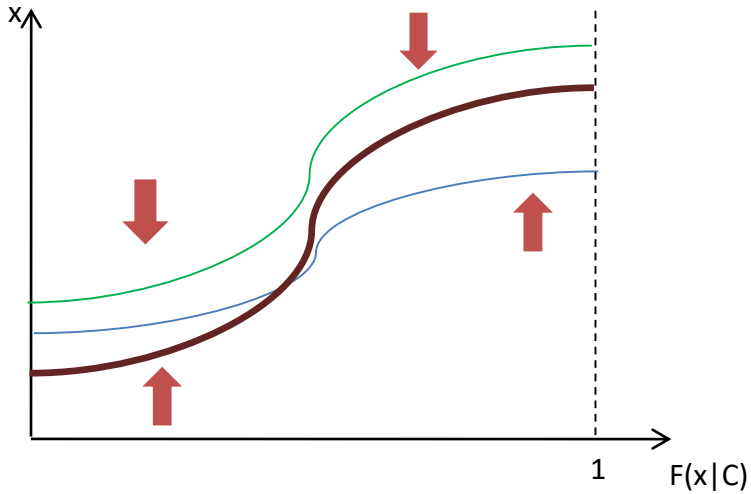
van de Gaer's "min of means"



Roemer's "mean of mins"



Conditional equality



Egalitarian equivalence

5. Extensions: Development objectives

- What is the policy objective for opportunity egalitarians?

$$Max_{\phi \in \Phi} min_i \int_t^{\infty} e^{\delta(t-s)} \mu_{i,s} ds$$

$$s.t. x_{ij,s} \geq z_s, \forall i, j, s$$

- The choice of policies from a feasible set so as to maximize the future stream of 'advantage' for the most disadvantaged type, subject to a no-deprivation constraint and to a policy acceptability constraint.

Source: Bourguignon, Ferreira and Walton, JEI 2007.

5. Extensions: Development objectives

- ‘Deconstructing’ the equitable development policy problem:

$$Max_{\phi \in \Phi} min_i \int_t^{\infty} e^{\delta(t-s)} \mu_{i,s} ds$$

$$s.t. x_{ij,s} \geq z_s, \forall i, j, s$$

“Growth matters”

Permissible Policy Set: Technical feasibility and social acceptability

Poverty eradication as a ‘constraint’.

“Rawlsian” criterion. All weight on the least advantaged.

Source: Bourguignon, Ferreira and Walton, JEI 2007.

3d. "Second generation" studies

3. Brunori, Peragine and Serlenga (2018) note that sampling variation can produce an upward bias in IOp estimates when cell partition is too fine for a given sample size (or regression model is overfitted).
   - They propose choosing the specification $\hat{g}(C)$ that minimizes the mean squared error of out-of-sample predictions:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{g}(C_i)\right)^2$$

   - Which is decomposable as follows:

$$E\left(y_0 - \hat{g}(C_0)\right)^2 = Var(\hat{g}(C_0)) + \left[Bias\left(\hat{g}(C_0)\right)\right]^2 + Var(u)$$

Captures the upward bias from sampling variation

Captures the downward bias from mis-specification

3d. “Second generation” studies

3. The procedure uses k-fold cross-validation. The average MSE for the k test samples is computed for each model specification, and the specification with the lowest MSE is chosen.

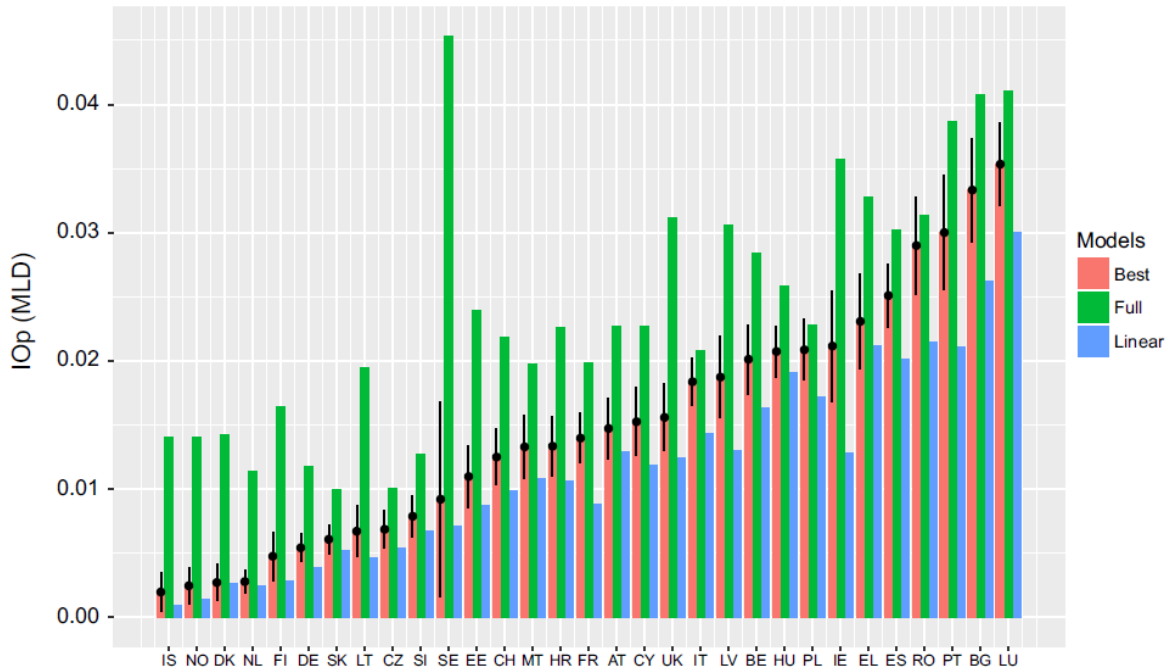


Fig. 1 IOp in 31 European countries under different model specifications. The Figure shows each country's IOp measure obtained with the three alternative methods: (i) the linear, most parsimonious case (*linear*), (ii) the fully interacted model (*full*); (iii) the best model selected (*best*). Countries are ordered according to the IOp level based on the *best* model specification with 95% confidence intervals. Table 2 in the Appendix contains IOp estimates and relative bootstrapped standard errors based on 500 replications for the three alternative model specifications . *Source: EU-SILC, 2011*