Approaches to Inequality of Opportunity

Principles, Measures, and Evidence

Xavier Ramos

Universitat Autònoma de Barcelona

Winter School on Inequality and Social Welfare Theory
13th January 2016

Acknowledgement

This lecture builds on joint work with Dirk Van de gaer

- Approaches to Inequality of Opportunity: Principles, Measures, and Evidence, *Journal of Economic Surveys*, in press.

Other recent review papers on Inequality of Opportunity

- Roemer, J.E. and A. Trannoy (2015). Equality of opportunity, in A.B. Atkinson and F.Bourguignon (eds.), *Handbook of Income Distribution* (Ch. 4, pp. 217–300), North-Holland
- Ferreira, F.H.G. and V. Peragine (2015). Equality of opportunity: Theory and evidence, IZA Discussion Paper 8994. Forthcoming in M. Adler and M. Fleurbaey (eds.), *Oxford Handbook of Well-Being and Public Policy*, OUP.
- Pignataro, G. (2012). Equality of opportunity: policy and measurement paradigms. *Journal of Economic Surveys*, 26(5): 800–834.

Motivation I

- Yo soy yo y mi circunstancia
- People's lives should not be conditioned by who their parents are, their gender or ethnicity.
- Those who try equally hard should obtain the same.

Motivation II

- The theory of equality of opportunity puts individual *responsibility* at the forefront when assessing situations of economic advantage.
- Outcomes are determined by `circumstances` and `effort` or responsibility
  - `circumstances`: factors that are beyond individual's control or responsibility
  - `effort`: factors for which individuals are deemed responsible.
- Here the benchmark is not equality per se, but a distribution where effort is rewarded adequately and the effect of circumstances is compensated for, so that only disparities due to effort remain.

Motivation III

- In recent years, we have seen an explosion of empirical literature that tries to measure the extent of inequality of opportunity.
- The measurement of equality of opportunity entails many methodological and empirical questions that are often difficult to resolve.
- We present and discuss the main conceptual issues and outline the solutions that have been put forth in the literature.
- We discuss them in a systematic manner, spelling out the conceptual grounding and implication of each option.
- In doing so, we identify and suggest new possibilities to measuring inequality of opportunity.

Outline I

Outline II

Principles

Equality of Opportunity: Ex-ante vs. ex-post

Useful to think of outcome matrices, e.g.,

$$Y^1 = \begin{bmatrix} 20 & 15 \\ 15 & 10 \\ 30 & 6 \\ 25 & 1 \end{bmatrix}$$

- Ex-post (EOP): differences in income due to differences in circumstances are eliminated and so all elements in each column must be the same.
- Ex-ante (EOA): equal rows.

Conclusion: EOP and EOA are equivalent if efforts and circumstances are independent.

$\mathrm{E0}$: Equality of opportunity holds iff all rows are equal.

Compensation: Ex-ante vs. ex-post

Differences that are due to circumstances should be compensated.

- Ex-post compensation (EPC) requires outcomes to be as equal as possible for individuals with same effort.
  - Pigou-Dalton transfers within columns increase equality of opportunity.
- Ex-ante compensation (EAC) Pigou-Dalton transfers from a type that is unambiguously better-off to a type that is unambiguously worse-off increase equality of opportunity.
- EPC and EAC are incompatible [Fleurbaey & Peragine 2011]

$$Y^2 = \begin{bmatrix} 20 & 15 \\ 15 & 10 \\ 30 & 6 \\ 25 & 1 \end{bmatrix} \text{ and } Y^3 = \begin{bmatrix} 21 & 15 \\ 15 & 9 \\ 30 & 7 \\ 24 & 1 \end{bmatrix}.$$

- EAC: $Y^2 \succ Y^3$, while EPC: $Y^3 \succ Y^2$.

Reward Principles I

Efforts should be adequately rewarded.

- Several reward principles exist
- *Natural Reward* (NR): Policy (tax-transfer systems) should respect the income differences that are due to differences in exerted effort.
  - i.e. all elements in the same row of R should be equal.
- NR and EPC are incompatible [Bossert 1995 and Fleurbaey 1995]
  - Suppose incomes are generated by the function $Y_{ij} = R_{ij} + 10\,|i - j|$

$$R^4 = \begin{bmatrix} 40 & 30 \\ 30 & 40 \end{bmatrix} \quad \textit{and} \quad Y^4 = \begin{bmatrix} 40 & 40 \\ 40 & 40 \end{bmatrix}$$
$$R^5 = \begin{bmatrix} 39 & 31 \\ 31 & 39 \end{bmatrix} \quad \textit{and} \quad Y^5 = \begin{bmatrix} 39 & 41 \\ 41 & 39 \end{bmatrix}.$$

- EPC calls for rich-to-poor transfer within tranche $\left[Y^4 \succ Y^5\right]$, which goes against NR $\left[Y^5 \succ Y^4\right]$.
- NR also conflicts with EOP and EOA.

Reward Principles II

- *Utilitarian Reward* (UR): there should be no inequality aversion with respect to differences in incomes that are due to differences in efforts.
- UR and EPC are incompatible [Fleurbaey and Peragine 2011].

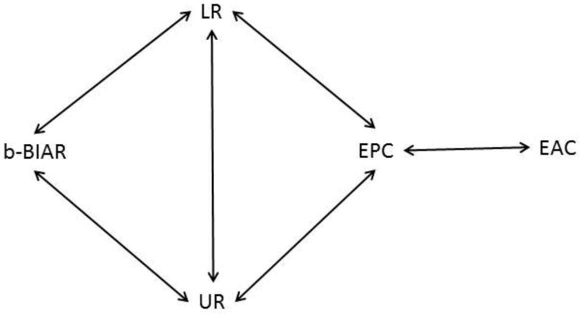
$$Y^6 = \begin{bmatrix} 30 & 5 \\ 20 & 10 \end{bmatrix} \text{ and } Y^7 = \begin{bmatrix} 29 & 6 \\ 21 & 9 \end{bmatrix}.$$

- EPC: $Y^7 \succ Y^6$, while UR: $Y^6 \sim Y^7$, since the sum of the cells of both rows are the same in both matrices.

Reward Principles III

- Some compensation is due even after taking circumstances into account, for three reasons:
  1. the market reward to effort leads to excessive income inequalities. Why should actual property rights be the benchmark to assess equality of opportunity?
  2. we are often unsure about the actual set of circumstances, and some are unobservable. Because of these two issues, Roemer 2012 suggests an increasing concave transformation of incomes as relevant outcome measure.
  3. incomes are stochastic. Since individuals are risk averse, opportunity sets should be evaluated in a risk averse way (Lefranc, Pistolesi, Trannoy, 2009).
- *b-Bounded Inequality Averse Reward* (b-BIAR): calls for inequality aversion with respect to differences in incomes, within types.
- It can be shown (p. 9 in paper) that NR, UR and b-BIAR are incompatible.

Wrapping up: Incompatibilities between Principles



Measures

Measures: Preliminaries I

- When using *parametric methods*, we should incorporate unobserved variables u_k and a random term ς_k, such that

$$y_k = g\left(a_k^C, a_k^R, u_k, \varsigma_k\right).$$

- Since u_k is unobserved and the functional form g is unknown, we are left with

$$\widehat{g}\left(a_k^C, a_k^R, e_k\right)$$

- The error term, e, captures the effect of unobservables and specification errors.
- With omitted variables, estimates may be biased (and no causality).
  - Counterfactual incomes used to measure inequality of opportunity will be over- or under-estimated.
- However, omitted variable bias has been largely ignored in the literature, which has been more concerned with developing lower bound estimates.

Measures: Preliminaries II

- Some measures below use parametric specifications that only include circumstances or efforts, and random variation.

$$\widehat{g}^C\left(a_k^C, e_k\right)$$
$$\widehat{g}^R\left(a_k^R, e_k\right)$$

- Incomes can be estimated by setting $e_k = 0$.
- *Non-parametric methods* typically rely on averaging procedures.
- Non-parametric estimates of the above are

$$\begin{aligned} y_k^{c1} &= \frac{1}{|N_{k.}|}\sum_{i\in N_{k.}} y_i \\ y_k^{EO1} &= \frac{1}{|N_{.k}|}\sum_{i\in N_{.k}} y_i \end{aligned}$$

- Omitted and unobservable variables also bias non-parametric estimates.

Measures: Preliminaries III

Parametric methods impose functional form assumptions and may suffer from specification errors. Three reasons may justify this cost.

1. Multivariate regression framework uses data *more efficiently*. As number of circumstances and efforts increase, types and tranches grow exponentially and small cell sizes is a real problem.
2. This problem is *more severe with continuous circumstances*. Kernel density techniques have been used (e.g. o'Neill et al 2000) but require large datasets to yield reliable estimates.
3. Parametric methods allow to *estimate the partial effect* of one (or a set) of circumstance variables (see on).

Measures: Preliminaries IV

- We distinguish 4 approaches to measuring Inequality of Opportunities
    - Stochastic dominance
    - Direct approach
    - Indirect approach
    - Norm based measures

Stochastic Dominance

- Absence of first order stochastic dominance between type's cumulative distribution functions can be seen as a test for ex-ante equal opportunities.
- Absence of second order stochastic dominance between type's cumulative distribution functions is consistent with the b-BIAR principle $(b = 1)$.
- Absence of equal type mean incomes is consistent with the UR principle.

- Here circumstances and efforts are independent, so rejecting *ex-ante* equality of opportunity implies also rejecting *ex-post*.

Direct measures: Ex-ante non-parametric

- Inequality, $I(\cdot)$, in a counterfactual distribution, y^c, where inequalities due to differences in effort have been eliminated.
- *Ex-ante*: replace every individual's actual income by some evaluation of his opportunity set.
  - The value assigned to his opportunity set should not depend on his own effort level.
- Two non-parametric approaches.
  1. Van de gaer (1993): surface under Pen parade of type

$$y_k^{c1} = \frac{1}{|N_{k.}|} \sum_{i \in N_{k.}} y_i$$

-
  - Inspired by UR since no inequality aversion within type.
  - $I(\cdot)$ with infinite inequality aversion, since inequalities due to circumstances are morally objectionable.

Direct measures: Ex-ante non-parametric

2. Lefranc, Pistolesi, Trannoy (2008): surface under generalized Lorenz curve of type.

$$y_k^{c2} = \frac{2}{|N_{k.}|\,|N_{k.}+1|}\sum_{i\in N_{k.}} i\widetilde{y_i}$$

-
  - It embodies the b-BIAR principle.
  - Take Gini coefficient as measure of $I(\cdot)$.

Direct measures: Ex-ante parametric

- The parametric approach can estimate the income associated with efforts not chosen by sample members.
  1. Ferreira and Gignoux (2011): mean income conditional on circums.

$$y_k^{c3} = \widehat{g}^C \left(a_k^C, 0 \right).$$

  2. Parametric analogue to Van de gaer's y_k^{c1}.

$$y_k^{c3b} = \frac{1}{|N_{k.}|} \sum_{i \in N_{k.}} \widehat{g} \left(a_i^C, a_i^R, 0 \right).$$

- y_k^{c3b} is new:
  - it deals with the covariance between a^C and a^R in a more flexible type-dependent way.
  - but it needs observations on a^R.

Direct measures: Ex-post non-parametric

- *Ex-post*: counterfactual should eliminate inequalities due to efforts.
- Checchi and Peragine (2010): scale up or down incomes so that inequalities within tranche are preserved but eliminates inequality of average income between tranches.

$$y_k^{c4} = y_k \frac{\mu(y)}{y_k^{EO1}}.$$

- Recall y_k^{EO1} is mean tranche income, i.e. mean income of individuals with same effort.

Direct measures: Ex-post parametric

- Pistolesi (2009) and Schokkaert (1998) estimate y^c parametrically as:

$$y_k^{c5} = \widehat{g}\left(a_k^C, \overline{a}^R, e_k\right).$$

- The parametric approach always yields meaningful estimates for y_k^{c5}, even when the combination $\left(a_k^C, \overline{a}^R\right)$ does not occur in the sample.
- What value of e_k to take?
  - zero: amounts to treating e_k as effort.
  - estimated value: amounts to treating e_k as a circumstance.
- Most authors take mean value for effort in the sample as reference value $\overline{a}^R$.This is arbitrary.
- Alternative: use averaged inequality measure $\frac{1}{N}\sum_{l=1}^{N} I\left(y^{c5}(a_l^R)\right)$.

Indrect measures

- Compare the inequality in the actual distribution of income, $I(y)$, to the inequality in a counterfactual income distribution where there is no inequality of opportunity $I\left(y^{E0}\right)$

$$\Theta_I(y, y^c) = I(y) - I\left(y^{EO}\right).$$

- To estimate y^{EO}, most applications construct a counterfactual that eliminates inequalities between individuals with same effort.
- Thus, they are ex-post measures, but remember that when effort is distributed independently of type, EOp ex-post implies EOp ex-ante.
- We show that for each counterfactual y^c, there exist a *dual* counterfactual in the indirect approach.

Indrect measures: Ex-post non-parametric
(dual of direct ex-ante)

- Dual of y_k^{c1}: Peragine and Checchi (2010) replace individuals' income by the average income of tranche:

$$y_k^{EO1} = \frac{1}{|N_{.k}|} \sum_{i \in N_{.k}} y_i$$

- y_k^{EO1} is close to the notion of Utilitarian Reward.
- Alternative based on Inequality Averse Reward:

$$y_k^{EO2} = \frac{1}{|N_{.k}|\,|N_{.k}+1|} \sum_{i \in N_{.k}} i\widetilde{y}_i$$

- which is new and the dual of y_k^{c2}.

Indrect measures: Ex-post parametric
(dual of direct ex-ante)

- Dual of y_k^{c3}: obtain mean income conditional on effort

$$y_k^{EO3} = \widehat{g}^R\left(a_k^R, 0\right).$$

- y_k^{EO3} is new.

Indrect measures: Ex-ante non-parametric

(dual of direct ex-post)

- All above approaches ensure ex-post EOp and entail ex-ante EOp when efforts are distributed independently of type.
- Checchi and Peragine (2010) is the only one ex-ante proposal that does not impose ex-post. It evaluates individual's opportunity set by y_k^{c1}, and builds the conunterfactual

$$y_k^{EO4} = y_k . \frac{\mu(y)}{y_k^{c1}}.$$

- where mean average within type income equals overall mean income for all types; i.e. all individuals face the same opportunity set.
- Opportunity sets can be valued differently. For instance, y_k^{EO5} uses y_k^{c2}

$$y_k^{EO5} = y_k . \frac{\mu(y)}{y_k^{c2}}.$$

Indrect measures: Ex-post parametric
(dual of direct ex-post)

- Dual of y_k^{c5}: Bourguignon, Ferreira, Menéndez (2007) fix a reference value for the circumstances:

$$y_k^{EO6} = \widehat{g}\left(\overline{a}^C, a_k^R, e_k\right).$$

- What value of e_k to take?
  - zero: amounts to treating e_k as circumstance.
  - estimated value: amounts to treating e_k as effort.
- Most authors take mean value for circumstances in the sample as reference value $\overline{a}^C$.
- Could use instead the aggregate inequality measure

$$\frac{1}{N}\sum_{l=1}^{N} I\left(y^{EO6}(a_l^C)\right).$$

Norm Based Measures

- Estimate a norm distribution, y^n, as a function of individual's circumstances and efforts, according to some fair allocation rule.
- Compare actual distribution to the norm distribution

$$I\left(y, y^n\right)$$

- $I(\cdot, \cdot)$ must satisfy at least two requirements:
  - It must satisfy *partial* (not full) *symmetry*, i.e. be invariant to permutations of $\left(y_k, y_k^n\right)$ pairs
  - Due to the heterogeneity of the population in terms of compensation and responsibility characteristics, the usual *transfer principle does not apply*.
    - Devooght (2008) uses Cowell's (1985) measure of distributional change.
    - Almas *et al.* (2011) define unfair treatment of each individual as $\left|y_k - y_k^n\right|$, and propose an unfairness Gini to aggregate these differences.

Overview I

- We propose several new measures
  - indirect ex-post measures as duals of existing ex-ante measures (y^{EO2}, y^{EO3})
  - Checchi and Peragine's y^{EO4}, adjusted to deal with inequality averse reward (y^{EO5})
  - approaches that require the choice of a reference value for either efforts (y^{c5}) or circumstances (y^{EO6}), an average inequality index could be used.
- Many inequality measures have been used, often without much justification
  - Norm based: index that does not satisfy the standard transfer principle and satisfies partial symmetry.
  - Direct approach: infinite inequality aversion index since all inequalities that are due to differences in circumstances are unacceptable.
  - MLD since it is the only path independent decomposable measure, which means that y^{c1} and y^{EO4} yield the same result.

Overview II

- Stochastic dominance approach is by nature ex-ante, but if circumstances and efforts are independent, rejecting ex-ante is equivalent to rejecting ex-post.
- Norm-based approaches have only been applied with *ex-post* type allocation rules (e.g. conditional egalitarian, egalitarian equivalent). Using y^{EO4} or y^{EO5} yields a norm distribution based on *ex-ante* equality of opportunity.
- The indirect approach may be questioned, as can be seen as a norm based approach with indices satisfying full symmetry.

Identification of Circumstances and Effort

Circumstances

What should individuals be held responsible for?

- Hard determinists deny the existence of free will. The set of responsibility is empty.
- Individuals ought to be held responsible only for what *lies within their control* (Arneson, Cohen, Roemer).
  - Classify as circumstance family background variables (parental education), individual characteristics (gender, ethnicity), innate characteristics (IQ), and contextual variables (access to basic services).
- Individuals ought to be held responsible for their *preferences and the ensuing choices* (Rawls, Dworkin, Van Parijs).
  - Minimal set of circumstances including innate characteristics or traits (talent or beauty)
  - Variables such as gender or ethnicity, should belong to the realm of responsibility if the differential effect they bring about reflects exclusively differences in behaviour.
- Individuals are *entitled to the products of all personal characteristics*, including genetic ones such as innate talent (self-ownership argument by Nozick).

- Partial observation of circumstances (and complete observation of efforts) lead to *downward bias* in ex-post and ex-ante inequality of opportunity.
- Ex-post: inequality within columns (tranches) based on observed types is smaller than that based on true types, since outcome of observed types is a weighted average of outcomes conditioned on true types.
- Ex-ante: the rows associated with the observed types are weighted averages of the rows associated with the true types.
- Ferreira and Gignoux (2011) argue estimates based on partial observation should be interpreted as lower bound estimates of Inequality of Opportunity.
- Niehues and Peichl (2014) argue that treating estimated individual fixed effects as circumstance in the direct approach yield an upper bound.

- Consider the indirect parametric approach and let

$$\ln y_k = \beta^C a_k^C + \beta^R a_k^R + e_k.$$

- We can estimate the partial effect of (a set) of circumstance variables J, controlling for the others $(j \neq J)$, by constructing alternative counterfactual distributions

$$y_k^{EO(J)} = \exp\left[\widehat{\beta}^J \overline{a^{CJ}}_k + \widehat{\beta}^{j \neq J} a_k^{Cj \neq J} + \widehat{\beta}^R a_k^R + \widehat{e}_k\right],$$

- $\overline{a^{CJ}}_k$ is the vector of reference values of the circumstances in set J.

Measuring effort

Unobservable effort, non-parametric identification

- RIA (Roemer's Identification Assumption): those that are at the same percentile of the distribution of income conditional on their type have exercised the same degree of effort.
- RIA assumes:
  1. multi-dimension effort variables, a_i^R, can be aggregated into a scalar measure of responsibility a_i^r, and income is a stricly increasing function of a_i^r.
  2. a_i^r is distributed independently of a_i^C.
- This is a very powerful assumption, as it allows estimating inequality of opportunity even when effort is unobservable.

- The inverse of the cumulative income distributions conditional on types gives, for each percentile, the corresponsing income level.
- Fixing a percentile value and looking at corresponding incomes for all types is like looking at a column of a matrix, i.e. ex-post perspective.
- If the plots for two types differ at some percentile, we have ex-post IOp.
- *Ex-post* equality of opportunity requires Equal Conditional Cumulative Distribution Functions.
- Looking at the CDF for each type is like looking at the rows of the matrix, i.e. ex-ante perspective.
- We need all types' CDFs to be the same for equality of opportunity.
- *Ex-ante* equality of opportunity requires Absence of First Order Stochastic Dominance between types' cumulative distribution functions.
- Hence, accepting RIA, ex-post equality of opportunity implies ex-ante equality of opportunity.

Measuring effort

Unobservable effort, non-parametric identification

- Accepting RIA, with omitted circumstances induces wrong identification of effort unless the unobserved circumstances, after conditioning on observed circumstances, no longer affect income
- Reason: the estimated CDF is a weighted average of the true CDFs (i.e. the CDF of the true, "finer" type partitioning):

$$F\left(y \mid a_i^{CO}\right) = F\left(y \mid a_i^{CO}, \underline{a}^{CU}\right) p_i\left(\underline{a}^{CU}\right) + F\left(y \mid a_i^{CO}, \overline{a}^{CU}\right) p_i\left(\overline{a}^{CU}\right)$$

- The percentile corresponding to $F\left(y \mid a_i^{CO}\right)$ does not correspons to the percentiles in the true type partitioning.

- Björklund et al. (2011) allow the distribution of effort conditional on type to have different variances, as initially suggested by Roemer (1998).
- They assume that effort has two components:
  - a type specific component, η^i_k, whose variance (σ^2_i) differs across types i and which captures the part of effort that is correlated with circumstances,
  - a general component, ω_k, with a homogeneous variance, σ^2.
- Define ω_k as a standarization of η^i_k:

$$\omega_k = \eta^i_k \frac{\sigma^2_i}{\sigma^2}$$

Measuring effort

Unobservable effort, parametric identification (II)

- Income generating process:

$$\ln y_k = \beta^C a_k^C + \eta_k^i = \beta^C a_k^C + \tilde{\eta}_k^i + \omega_k,$$

- where

$$\tilde{\eta}_k^i = \eta_k^i - \omega_k$$

- $\tilde{\eta}_k^i$ captures the indirect effect of circumstances
- ω_k is assumed to capture 'pure' effort.
  - note that error terms (i.e. specification error and omitted vble bias) are lumped together with effort

- Panel data allows to distinguish between time-invariant and time-varying efforts and circumstances
- Time-invariant effort, a_k^R, would include skills, preferences, aspirations.
- Time-varying effort, a_{kt}^R, would include exertion of effort, such as hours worked.
- The income generating process can be modelled as

$$\ln y_{kt} = \beta_1^C a_{kt}^C + \beta_2^C a_k^C + a_k^R + a_{kt}^R + \upsilon_{kt}$$

- a^R_{kt} cannot be distinguished from the idosyncratic error term, υ_{kt}. Thus, we have

$$\ln y_{kt} = \beta^C_1 a^C_{kt} + \beta^C_2 a^C_k + a^R_k + \varepsilon_{kt}$$

- a^R_k are allowed to be correlated with circumstances, *i.e.* circumstances may affect preferences and aspirations but not exerted effort.
- Salvi (2007) uses the indirect approach and estimates the counterfactual y^{EO} as

$$y^{EO}_{kt} = \exp\left[\widehat{\beta}^C_1 \overline{a}^C_{kt} + \widehat{\beta}^C_2 \overline{a}^C_k + \widehat{a}^R_k + \widehat{\varepsilon}_{kt}\right]$$

- Note that $\widehat{\varepsilon}_{kt}$ is lumped together with efforts, as in Björklund et al. (2011).
- To identify $\widehat{\beta}^C_2$ and $\widehat{a}^R_k$ we can use Random Effects and the transformation proposed by Mundlack (1978) to take account of the correlation between individual specific effects and other time-varing covariates.

Measuring effort

Observable effort correlated with circumstances (I)

- Bourguignon et al (2007), model earnings, y_k^a, as function of effort $\left(a_k^R\right)$ and circumstance $\left(a_k^C\right)$ variables

$$\ln y_k = \beta^C a_k^C + \beta^R a_k^R + e_k$$

- and let endogenous effort depend on circumstances:

$$a_k^R = H a_k^C + v_k$$

- We can estimate a reduced form of these two equations

$$\ln y_k^a = \psi a_k^C + \varepsilon_k$$

- This allows the estimation of direct and indirect effects of circumstances on earnings.

- *Overall* effect can be obtained from the conterfactual

$$y_k^{EODT} = \exp\left[\widehat{\psi}\overline{a}^C + \widehat{\varepsilon}_k\right]$$

- *Direct* effect can be obtained from the conterfactual

$$y_k^{EOD} = \exp\left[\widehat{\beta}^C \overline{a}^C + \widehat{\beta}^R a_k^R + \widehat{e}_k\right]$$

- *Indirect* effect:

$$y_k^{EODT} - y_k^{EOD}$$

- Inequality of Opportunity $= I\left(y\right) - I\left(y^{EODT}\right)$

Error terms

- Most sudies include a limited set of circumstances in the analysis.
- Most include social background (parental income or education)
- Very few surveys have observations on genetic luck. Björklund et al. (2011) find IQ to be the most influential factor behind inequality of opportunity in Sweden.
- Genetic luck can be an important contributor to the error term.
- We are unaware of forms of brute luck or option luck being included in the list of circumstances such that they always enter the error terms.
- It is usually claimed that genetic luck should be fully compensated and some compensation is due for brute luck.
- Then, the principle of UR (using a full list of circumstances) has to be replaced by b-BIAR (since one is typically using only a limited list of circumstances).

Luck

- Most forms of luck thought to deserve compensation
- *Social background luck*: Rawls' social lottery; family background
- *Genetic luck*: Rawls' natural lottery; predetermined constituent characteristics of the individual, e.g. talent.
- *Brute luck* (Dworkin): situations where the individual cannot alter the probability of an event taking place.
  - Full compensation may entail large redistribution and requires a lot of information: Vallentyne (2002) defends compensating only for *initial* brute luck (before individuals are deemed responsible).
- *Option luck* (Dworkin): when individuals *deliverately* take risks. Does not deserve compensation.
  - Fleurbaey (2008) argues for partial compensation since the outcome of the lottery is uncertain.

Empirical Applications

Some relevant questions I

- Are the different approaches and methods outlined above important in practice?
- How sensitive are the findings to the various modelling options implemented in the literature?
- There is no empirical study that applies in a systematic manner the various approaches put forth in the literature to the same data –we're working on it.
- We draw mostly on studies that implement more than one approach to the same data to address seven relevant questions.

Some relevant questions II

1. Is the stochastic dominance approach able to detect inequality of opportunity?
2. Does the ex-post versus the ex-ante dilema matter in practice?
3. Does it make a difference whether we use direct or indirect measures?
4. Do norm-based approaches yield different results than non-norm based approaches?
5. What is the importance of indirect effects of circumstances?
6. What can we learn from the different treatment of the error term in parametric aproaches?
7. What are the most important circumstances?

Stochastic dominance

- Lefranc, Pistolesi and Trannoy (2008) compare 9 Western countries.
- How sensitive are the findings to the various modelling options implemented in the literature?
- Use pre-tax and net disposable household income
- Sample: male-headed households aged 25-40,
- Circumstance vble.: three levels of social background.
- They compare pairwise the cumulative conditional distributions within each country by means of first and second order stochastic dominance
- Sweden is the only country for which equality of the conditional cumulative distribution functions cannot be rejected.
- It is remarkable that, even though only 3 types are distinguished by Lefranc et al., the stochastic dominance approach is able to detect inequality of opportunity.

Ex-ante vs. ex-post

- Ex-ante and ex-post approaches yield different results.
- All papers use RIA when measuring ex-post inequality of opportunity.
- Recall that if this assumption is not valid, RIA leads to an underestimation of ex-post inequality of opportunity.
- All reviewed papers find that ex-post inequality of opportunity is larger than ex-ante inequality of opportunity.
- Tthis could imply that they underestimate the difference between ex-ante and ex-post approaches.

Direct vs. indirect measures

- Two studies: Pistolesi (2009) and Ferreira and Gignoux (2011).
- Pistolesi (2009) takes an ex-post view while Ferreira and Gignoux (2011) adopts an ex-ante view.
- This suggest that direct and indirect measures yield similar results irrespective of the view taken.
- Likewise, the similarity appears to be rather robust to different inequality measures.
    - Pistolesi (2009) finds that the similarity of the results holds for several inequality indices.

Norm vs. non-norm based measures

- Two studies: Devooght (2008) and Almas *et al.* (2011).
- Norm based measures seem to yield much larger estimates of inequality of opportunity than other approaches.
- This conclusion has to be taken with caution, as there are no empirical studies that directly compare estimates of norm based and other approaches, which means that such differences may also be due to differences in other methodological options, or simply because they use different datasets.
- Differences however are sufficiently large making it hard to believe that they would disappear.

The role of indirect effects of circumstances

- Bourguignon et al. (2007): indirect effect of five circumstances (father's and mother's education, father's occupation, race, and region of birth) through their impact on three observed effort variables (own education, migration out of hometown, and labor market status) accounts for 40% of the overall effect of circumstances.
- Björklund *et al.* (2012) measure the indirect effect of circumstances by the heterogeneous type-specific variances.
- Find that type heterogeneity accounts for 20 to 50% of the overall effect of circumstances, depending on the inequality index.
- Thus accounting for the indirect effect of circumstances on efforts makes a big difference in the assessment of inequality of opportunity.

Treatment of residuals

- Parametric approaches leave a substantial part of the variation unexplained, which goes to the residual.
- The decision to treat residuals as circumstances or efforts, is thus important for the analysis.
- Hence, checking the robustness of the results with respect to this choice is imperative.
- With the norm based approach one can choose whether to include the residual in the circumstance (upper bound estimate of unfairness) or in the effort set (lower bound estimate of unfairness).
- Studies that use the norm based approach find substantive differents.
- When effort is not observable and the non parametric method RIA is applied, the error term is *de facto* treated as an effort variable, such that inequality of opportunity estimates should be considered as lower bound estimates.

Most important circumstances

- There is little consensus about the most important circumstance variable: different circumstances account for the largest share of income or consumption inequality in regions with different economic conditions and degree of economic development.
- Björklund *et al.* (2012), using the largest set of circumstances of all studies to date, find IQ to be the most influential circumstance for Sweden.
- Bourguignon et al. (2007) find parental education to be the most influential circumstance in Brazil.
- Salvi (2007) finds infrastructures and ethnicity to be the most influential circumstances in Nepal.

Conclusions

Conclusions: Principles

- Inequality of Opportunity theories attempt to combine a compensation principle with a reward principle.
- Compensation may be ex-ante or ex-post.
- Reward also has 3 flavours: utilitarian, natural and inequality-averse, that we suggest.
- Ex-post compensation and reward principles are incompatible, so we have to make choices.
- With unobserved circumstances or arbitrary property rights, a reasonable option is to have infinite inequality aversion between types and moderate inequality aversion within types.

Conclusions: Empirical Approaches

3 empirical approaches

- (Differences in) standard inequality indices.
  - we show the duality between counterfactuals used in direct and indirect approaches, and use it to formulate new indirect measures.
  - all indirect measures but one imply use counterfactuals of ex-post EOp that imply ex-ante EOp if effort is distributed independently of type.
- Stochastic dominance (ex-ante or ex-post with RIA).
- Difference between actual distribution and norm income vector.
- The Direct and Norm based approaches are more suited than the Indirect approach to measuring IOp
- The Indirect approach is useful to decompose inequality of outcome into circumstances and effort.
- Parametric approaches rely on econometric techniques to estimate the counterfactual or norm distribution.
  - If error term is random, should be compensated since it is brute luck.
  - However, the error term often contains missing circumstances and effort.

Conclusions: Empirical Applications

Only few studies compare the performance of different approaches and methods.

1. Norm based approaches yield substantially different results than non-normed based methods (Devooght (2008): Almas et al (2011)).
2. Ex-ante or ex-post matters: C&P (2011) find lower ex-ante EOp in Italy; Aaberge et al (2011) find similar results.
3. Direct and indirect methods yield similar results (Pistolesi (2009), F&G (2011)).
4. No consensus about the most important circumstance: IQ (Sweden); family background (Brazil); acces to infrastructure and ethnicity (Nepal).
5. Indirect effects account for a substantial part of overall opportunity inequality (40% in BFM (2007), 25% in Björklund et al (2011)).
6. Treating error terms as circumstance or as effort makes a whole difference (Almas (2008), Almas et al. (2011)).

A closer look at two examples

Norm-based Approach: Almas et al (2011)

Almas, I., A.W. Cappelen, J. Thori Lind, E. Sorensen and B. Tungodden (2011). Measuring unfair (in)equality, Journal of Public Economics, 95: 488-499.

- Analyse pre-tax income distribution in Norway (1986-2005).

1. What is the measure they propose?
2. What responsibility-sensitive fairness principle do they use?
3. Where do they draw the responsibility cut?

$$G = \frac{1}{2n\left(n-1\right)\mu}\sum_i \sum_j \left|u_i - u_j\right|$$

where

$$u_k = \left|y_k - y_k^n\right|$$

- It satisfies partial symmetry, *i.e.* invariant to permutations of $\left(y_k, y_k^n\right)$ pairs.
- A transfer from a person less unfairly treated (lower u) to a person who is more unfairly treated lowers G.
- Maximum value is 2, when all ind. but one have zero income, the ind. having all the income has a fair income of zero, and one other ind.'s fair income is total income.

- Income is assigned according to each individual's claim
- Claim of individual i, $m\left(a^R;\cdot\right)$: average income of a counterfactual distribution where everyone has the same responsibility vector as individual i.

$$m\left(\cdot;a_i^R\right)=\frac{1}{n}\sum_j g\left(\overline{a}_i^R,a_j^C\right)$$

- Individual's i fair income, the is

$$y_i^{GPP}=\frac{m\left(\cdot;a_i^R\right)}{\sum_j m\left(\cdot;a_j^R\right)}\sum_j y_i$$

- GPP eliminates all *unfair* inequalities arising from *non*-responsibility factors.
- GPP preserves *fair* inequalities due to responsibility factors.

- Use a linear model of the log of labour earnings

$$\log y_i = \gamma a_i^C + \beta a_i^R + \varepsilon_i$$

- Estimate individual's i fair income as

$$\widehat{y}_i^{GPP} = \frac{\exp\left(\widehat{\beta} a_i^R\right)}{\sum_j \exp\left(\widehat{\beta} a_j^R\right)} \sum_j y_i$$

- GPP eliminates all *unfair* inequalities arising from *non*-responsibility factors.
- GPP preserves *fair* inequalities due to responsibility factors.

- This sets circumstances to zero. More faithful to the GPP would be to use the average of the predictions when circumstances are not set to zero

$$\widehat{y}_i^{GPP} = \frac{\frac{1}{n}\sum_j \exp\left(\widehat{\gamma} a_j^C + \widehat{\beta}\bar{a}_i^R\right)}{\sum_j \frac{1}{n}\sum_j \exp\left(\widehat{\gamma} a_j^C + \widehat{\beta}\bar{a}_j^R\right)} \sum_j y_i$$

- This sets ε to zero, i.e. treats it as effort.

- They try several responsibility sets, starting with:
  - a^R: (hours worked, yrs. educ., public/private sector, region resid.).
  - a^C: (field of education, age, gender, u, e).

Responsibility set	G pre-tax		G post-tax	
	1986	2005	1986	2005
$\varnothing$ (standard Gini)	0.270	0.262	0.205	0.219
$\{H\}$	0.223	0.235	0.159	0.192
$\{H, E\}$	0.206	0.229	0.158	0.192
$\{H, E, P\}$	0.206	0.221	0.157	0.184
$\{H, E, P, D\}$	0.204	0.220	0.158	0.184
$\{H, E, P, D, F\}$	0.201	0.217	0.153	0.181
$\{H, E, P, D, F, A\}$	0.200	0.214	0.152	0.178
$\{H, E, P, D, F, A, \varepsilon\}$	0.120	0.076	0.098	0.069

- The inclusion of ε into the responsibility set makes a difference.
- The inclusion of H and E reduce inequality of opportunity to a lesser extent.

3. Drawing the responsibility cut

- They include family background and IQ in the analysis and examine the effects of these 2 circumstances on equality of opportunity through the responsibility variables.
- FB and IQ are correlated with years of education (E) but not with hours worked (H).
- They control for FB and IQ by building a new eduaction variable, $\widetilde{E}$, which is the difference between actual E and predicted $\widehat{E}$ from FB and IQ.

$$\widetilde{E} = E - \widehat{h}(FB, IQ)$$

- Close in spirit to a direct effect of BFM (2007), i.e. include effort net of circumstances.

- Results don't change much:

Responsibility set $[FB, IQ]$	Baseline 1986	Baseline 2005	With $\widetilde{H}$ 1986	With $\widetilde{H}$ 2005
$\varnothing$ (standard Gini)	0.181	0.241	0.181	0.241
$\{H\}$	0.179	0.236	0.179	0.236
$\{H, E\}$	0.179	0.236	0.177	0.235
$\{H, E, P\}$	0.180	0.237	0.176	0.232
$\{H, E, P, D\}$	0.173	0.233	0.171	0.229
$\{H, E, P, D, F\}$	0.169	0.230	0.170	0.228
$\{H, E, P, D, F, A\}$	0.169	0.229	0.170	0.225

Direct vs. Indirect / Parametric vs. Non-parametric

- Ferreira, F.H.G. and J. Gignoux (2011). The measurement of ineuqality of opportunity: Theory and application to Latin America, Review of Income and Wealth, 57(4): 622-657.
- Examine income and consumption inequality in 7 LA countries.
- Set of circumstances: father's and mother's education and father's occupation, ethnicity, region of birth.
- No effort variables available; regression residual is lumped together with effort.

Direct vs. Indirect approaches - non-parametric (I)

- Non-parametric direct and indirect approaches yield the same result if $I(\cdot)$ is path independent: MLD
- Counterfactuals:
  - Direct approach: $y^c = y_k^{c1}$
  - Indirect approach: $y^{EO} = y_k^{EO7}$
- Estimate Inequality of Opportunity relative to overall inequality
  - Direct approach: $\Theta_D = \frac{I\left(y_k^{c1}\right)}{I(y_k)}$
  - Indirect approach: $\Theta_I = \frac{I(y_k)-I\left(y_k^{EO7}\right)}{I(y_k)}$

Direct vs. Indirect approaches - non-parametric (II)

- Magnitudes change for $E(1)$ and $E(2)$, but country rankings are pretty much preserved

	Brazil			Colombia			Ecuador		
$I(\cdot)$	$E(0)$	$E(1)$	$E(2)$	$E(0)$	$E(1)$	$E(2)$	$E(0)$	$E(1)$	$E(2)$
Θ_D	.329	.337	.191	.250	.261	.157	.290	.287	.187
Θ_I	.329	.319	.416	.250	.287	.397	.290	.315	.421
	Guatemala			Mexico			Panama		
Θ_D	.373	.386	.209	.208	.208	.099	.346	.335	.213
Θ_I	.373	.419	.587	.208	.260	.454	.346	.322	.304
	Peru								
Θ_D	.292	.271	.124						
Θ_I	.292	.337	.418						

- Ranking: Mexico, Colombia, Ecuador, Peru, Brazil, Panama, Guatemala.

Parametric vs. Non-parametric - indirect approach (I)

- Parametrics help with small cell sizes, e.g. in 4 countries more than 1/5 of cells have less than 5 observations.
- Counterfactuals:
  - Non-parametric: $y^{EO} = y_k^{EO7}$
  - Parametric: $y^{EO} = y_k^{EO6}$
- Estimate Inequality of Opportunity relative to overall inequality
  - Non-parametric: $\Theta_I^{NP} = \frac{I(y_k) - I\left(y_k^{EO7}\right)}{I(y_k)}$
  - Parametric: $\Theta_I^{P} = \frac{I(y_k) - I\left(y_k^{EO6}\right)}{I(y_k)}$

Parametric vs. Non-parametric - indirect approach (II)

- Non-parametric estimates yield larger Inequality of opportunity

	Brazil			Colombia			Ecuador		
$I(\cdot)$	$E(0)$	$E(1)$	$E(2)$	$E(0)$	$E(1)$	$E(2)$	$E(0)$	$E(1)$	$E(2)$
Θ_I^P	.322	.305	.382	.233	.259	.350	.269	.284	.365
Θ_I^{NP}	.329	.319	.416	.250	.287	.397	.290	.315	.421
	Guatemala			Mexico			Panama		
Θ_I^P	.345	.371	.498	.172	.193	.342	.315	.274	.233
Θ_I^{NP}	.373	.419	.587	.208	.260	.454	.346	.322	.304
	Peru								
Θ_I^P	.279	.302	.321						
Θ_I^{NP}	.292	.337	.418						

- Country rankings though are pretty much preserved.