Inequality of Opportunity: measurement issues and empirical results

John Roemer (Yale)
Alain Trannoy (AMSE)

Based on Section 10 for the Chapter prepared for the Handbook of Income Distribution + ε

Responsibility and Equality

- Equality of outcome cannot be maintained as a reasonable social objective in all configurations.

- Because incentives will say the positive economist

- Because you may be responsible of your misery will say political philosophers such as Dworkin, Arneson, Cohen, etc.
  - Illegitimate inequality due to circumstances
  - Legitimate inequality due to effort

Two competing views about responsibility

- The preference view (Dworkin, Fleurbaey)
  - You are responsible for your preferences
- The control view (Cohen, Roemer)
  - You are responsible for what you control
  - Effort should take into account what set of actions a person can *access*,
  - Where access is a question not simply of material constraints but of psychological ones, which may be determined by one's circumstances

Not so different

- A way to encompass the preference view in the control view
- A way to anchor the control view in microeconomics

- *E*: Effort ; *C* circumstances; *M* market conditions; θ preference parameter (*all vectors*).

- $P\ (M, C)$ Possibility Set

- Control view E^*= arg*max* $V(E;\ C,\ \theta)$ *under* $E \in P\ (M, C)$

  $\rightarrow$ Behavourial reaction $E = f(C,\ M,\ \theta)$
- Preference view E^*=arg*max* $U(E;\ \theta)$ *under* $E \in P\ (M, C)$

  $\rightarrow$ Behavourial reaction $E = g(C,\ M,\ \theta)$

The same kind of reduced form for effort

- One of the aim of the control view is to clean effort from circumstances

- You should also do so in the preference view

- The only difference is that you clean the effect of circumstances
  - on both preferences and possibility set in the former
  - and only on the possibility set in the latter

- Can we identify both effects or only the cumulate impact?

Empirics and the link between the two views

- How to proceed to allow computations according to both views and let the reader choose?

- Good Survey by Ramos and van de Gaer (2012)

Outline

- 1. Methodological Issues: General remarks
- 2. Estimation phase
- 3. Measurement phase
- 4. Results

Methodological Issues: General remarks

- A multi-dimensional problem
- EOP as a process
- Lack of relevant information
- Age & Sex

Multi-dimensional problem

- IOP: Conditioning inequality
- Conditioning variable: effort and circumstances
- Applying the Sequential General Lorenz Criterion?
- Effort: needs
- Circumstances: negative needs
- Pb: the theory does not anything about how additional effort should be rewarded across the circumstance dimension.

IOP as a process

- It is not so much the difference in circumstances that matters, but the difference in the impact of circumstances
- Estimation of the impact of circumstances and effort through the best model (likely a dynamic one) should precede the measurement phase (Fleurbaey and Schokkaert 2009)
- The measurement phase should be thought independtly from the measurement phase to allow all ethical possibilities (option value)

What would be the best model?

- A structural model? Of course yes in principle!
  - Reverse causality should not be a problem
  - Endogeneity, unfortunately yes : Omitted variables such as genetic variables
  - « *An IV strategy is unlikely to succeed, since it is difficult to conceive of correlates of the circumstances that would not themselves have an indirect effect on earnings* » Bourguignon et al. 2007
  - Causal interpretation will be lacking.
    - Worrying for economic policy
    - not if we just want to measure the degree of inequality of opportunity

Lack of relevant information

- Illustration for two popular effort variables

- Number of hours of work
- For self-employed, good effort variable for the control view
- For wage-earners?
  - Unvoluntary part-time jobs, overtime, unemployment in a snapshot distribution
  - In the lifespan, better

- Years of education
  - Primary and even secondary take place before the age of consent.
  - Only tertiary education and lilefong education
  - But tertiary education is path-dependent

First Lessons

- Difficult to have accurate measure of effort at least in the income attainment example (better in health with lifestyles)
- Both noise and bias will affect effort variables
- Generally, we have accurate measure of circumstances even if they are incomplete
- Likely, safer to condition wrt circumstances variables (type) than to condition wrt effort (tranche)

Age & Sex

- Matter more for some outcome than for another (Health SAH, 45%)

- Control view: age and sex are circumstances

- Preference view: determinant of preferences and then responsibility variables

- Age as in the responsibility variable because we should be interested in lifetime earnings (Almas et al. 2011)

- Sex as neither responsibility nor circumstances in health matter for all biological differences.

2.Estimation phase

We should adapt our empirical strategy to the richness of the informational structure

- The case of rich data set
  - Rich set of circumstances + some effort variables
- The case of poor data set
  - Only some circumstances

Case of a rich data set

- *Discussion in a linear model: Outcome y, Demographics, D*

Outcome or return equation

- $y_i = \mu + \alpha_c C_i + \alpha_d D_i + \alpha_e E_{i,} + \varepsilon_{i.},$ (1)

Reaction equation in a reduced form

- $e_{ij} = \mu_j + \beta_c C_i + \beta_d D_i + \beta_m M_{i,} + \gamma_{cd} C_i D_i + \gamma_{cm} M_i D_i + o_{ij.},$ (2) for each effort variable $j = 1,\ldots,k$

Interpretation of the parameters: in red, parameters that could by considered as 'preference shifters' in the preference view (they are correlated to preferences in (2))

Avantages and weaknesses of the super-reduced form

- The estimation of the full system is not necessary is we only want to capture the full impact of circumstances

- The estimation of *super-reduced form* *is enough*

$$y_i = \mu_y + \delta_c C_i + \delta_d D_i + \upsilon_{i.} \qquad (3)$$

$\delta_c = \alpha_c + \alpha_e \beta_c$ direct + indirect effect through effort

- But we cannot isolate the effect of demographics as circumstances from the effect of demographics as preference shifters

Omitted variables in this setting

- δ_c conveys the impact of any unobserved variable correlated with circumstances

- If these ommitted variables are circumstances, fine
- According to Ferreira and Gignoux (2011) δ_c *C* is a lower bound estimate of the impact of circumstances.

- But if they are responsibility variables, it is the other way round.

- Example of IQ

- Not easy solution in a parametric setting

Brian Barry vs John Roemer

- Brian Barry: we should respect the true effort

$$y_i = \mu_1 + \alpha_c C_i + \alpha_d D_i + \alpha_e E_{i,} + \varepsilon_{i.} \qquad (1)$$

- John Roemer: we should respect the true effort disembodied from the impact of circumstances

- An estimate of the true effort is the residual of equation (2). Plugged into (1) gives

$$y_i = \mu_4 + \delta_c C_i + \delta_d D_i + \alpha_e \mathrm{O}_{\mathrm{i}} + \tau_{i.} \qquad (4)$$

What to do with the residual of the return equation?

- In one of the best studies in terms of richness of the dataset Björklund et al. (2012), residuals represent 70% of the variance

- Roemer and co-authors put it on the effort side
- Devooght (2008) and Almas et al.(2010) on the circumstance side.
- Lefranc et al (2009) argue that it is a mixed bag and move on with the explained part.

Non-parametric studies

- To estimate the conditional distribution *F(y|C,E)* and G(E/C) (analogues of (1) and (2)
- O'Neill et al (2000) were the first to use a kernel density estimator to estimate the distribution of income conditional on parental income.
- Pistolesi (2009) borrows a semi-parametric estimation technique from Donald et al.(2000). In a nutshell, since the hazard rate is defined as,

$$H(y) = \frac{f(y)}{1-F(y)} = \frac{f(y)}{S(y|C,E))} ,$$

with S (.|.) the conditional survivor function, one can write :

$$f(y|C,E) = H(y|C,E)\big(S(y|C,E)\big).$$

- The trick is then to estimate a hazard-function-based estimator and introduce covariates using a proportional-hazards model.
- In a second step, the necessary transformations using the above equation are made to obtain an estimate of the associated conditional density function.

The case of a poor data set

- Distinctive feature : no effort variable is available

- The approach here comes from Roemer (1993, 1996, 1998) with his identification axiom.

- It is non-parametric in essence, since effort is deduced from the distribution of outcome for a type, $F(y|C)$.

- Two individuals located at the same quantile of their type-conditional distribution are defined as having exerted the same effort, which will be denoted e_{RO}

Roemer identification axiom (RIA)

- Starting from the income generating process
  $y = g(C, E)$

- $F_y\left(g(C,E)\middle|\, C\right) = F_y\left(g(C',E')\middle|\, C'\right) \Rightarrow e_{RO} = e'_{RO}$

- By construction, this effort is distributed uniformly over [0, 1] for all types.

- This way of identifying effort used by O'Neill (2001), Peragine (2004, 2008)

Not immune to omitted circumstances

- Omitted circumstances induce wrong identification of the Roemerian effort unless the unobserved circumstances, after conditioning on observed circumstances, no longer affect income (Ramos and Van de gaer (2012))

- In addition, it is not clear how multi-dimensional effort can be aggregated into one indicator, (see Fleurbaey (1998))

- Interaction with luck factors

A weaker axiom

- The *type-independent effort distribution(TIED):* the relevant normative effort distribution should be independent of type.

- The view that the *distribution* of effort specific to a type is a circumstance makes sense in the control view and …?. in the preference view

- Examples of application
- Björklund et al. (2012) : parametric setting
- Lefranc et al. (2009): non parametric setting

Example of application: Parametric

- Björklund et al. (2012) estimated a reduced form (3) with υ_i a Gaussian white noise.

- They assimilate the distribution of the residual to the distribution of effort.

- However, the distribution of the residual can vary across types and this variation violates the *type-independent effort distribution.*

- They have corrected for variation in the second moment by adding and subtracting to the regression equation a residual term that has the overall variance.

Hence the relevant effort in each type is renormalized to have the same variance.

Example of application: Non-Parametric

- Effort is non-observed
- We only observe the conditional distribution of income on circumstances $F(./C)$
- $We\ can\ only\ check$ *conditional-distribution equality* (CDN): $F(./C) = F(./C')$ for any C,C'
- Lefranc et al.(2009) show that CDN is a necessary condition to satisfy *equal-luck opportunity*

$F(./C, e) = F(./C', e)$ *for any C,C' for any e* if we retain as ethical effort:

$$e = e_r = G(e|C)$$

Omitted circumstances?

- What's cost?
- The previous proposition is immune to omitted circumstances
- Which was not the case for RIA (see below)

Types versus tranches

- A matrix m in which rows are types and columns effort.
- An element of the matrix m_{ij} is the outcome for type *i* and effort level *j*.
- It means that, with respect to the decomposition of the process allowed by the regression, the residual is assigned to either effort or circumstances, unless the outcome is replaced by the predicted outcome.
- A *tranche (column)* as the set of individuals who expend the same degree of effort.
- A *type (row)* as the set of individuals who share the same circumstances.

3.Measurement phase

- Direct unfairness vs Fairness gap
- Types versus tranches
- Choice of index

Direct unfairness vs Fairness gap

- Direct unfairness (Fleurbaey-Schokkaert 2009 and Pistolesi (2009)
- *Direct unfairness* (DU) is computed as the inequality of the counterfactual distribution when one has removed the effect of effort variables
  - either by suppressing them, or
  - by imputing to each individual a reference value of effort such as the average value.
- *For the super-reduced form* = I (E$(y|C_i, D_i))$ or $I(\hat{\mu}_3 + \hat{\delta}_c C_i + \hat{\delta}_d D_i)$
- For other specifications I (E$(y|C_i, D_i, \overline{E}))$ or $I(\hat{\mu}_1 + \hat{\alpha}_c C_i + \hat{\alpha}_d D_i + \hat{\alpha}_e \overline{E}_{i,})$
- Depending on the assumption on type-mean residual

Fairness gap

- The *fairness gap* (FG) measures the gap between the inequality of the actual distribution and the inequality of a counterfactual distribution in which all the effects of circumstantial variables have been removed,
  - either by suppressing them,
  - or by imputing to each individual a reference value of circumstances such as the average one.

- From raw data $I(y)$ - I (E$(y|E_i)$) *(Checchi and Peragine 2010)*

- *For estimations (1) or (4) depending assumptions on the residual*

- $I(y) - I\,(\mathrm{E}(y|\bar{C}_i, \bar{D}_i, E_i))$ *type-mean residual as effort*
- $I(y) - I(\hat{\mu}_1 + \hat{\alpha}_c \bar{C}_i + \hat{\alpha}_d \bar{D}_i + \hat{\alpha}_e E_i)$. *type-mean as circumstance*

- $I(\hat{y}_i) - I(\hat{\mu}_1 + \hat{\alpha}_c \bar{C}_i + \hat{\alpha}_d \bar{D}_i + \hat{\alpha}_e E_i)$. *without residual*

Two incompatible compensation principles

- The *tranche-compensation principle* states that the closer each column is to a constant vector, the better.

- The *type-compensation principle* states that it is good to transfer from an advantaged type to a disadvantaged type

- They are incompatible in a sufficiently rich domain

- While they are not if we introduce luck
  - Tranche compensation principle = *equal-luck opportunity*
  - *Type-compensation principle leads to conditional–distribution equality*

- The tranche-compensation principle clashes in a rich domain with two principles of reward

Link between compensation principles and DU/FG

- Using direct unfairness is linked to the *tranche-compensation principle*

- If DU(m) < DU (m'), then m is preferred according to the tranche-compensation principle where the considered transfers are of the Pigou-Dalton sort.

- Similarly, there is a link between the *type-compensation principle* and the fairness gap.

- If m is preferred to m' according to the type-compensation principle, then FG (m) < FG (m') for all inequality indices when the reference type is different from the two types involved in the Pigou-Dalton transfer.

Choice of index depending on the context

- Health: variance is good
  - Absolute makes sense in the context of health
  - Covariance of circumstances with outcome (here health) is the arithmetic mean of the DU and FG when the other source is removed in the computations (not put at a reference level). Ferreira and Gignoux (2011) Jusot and al. (2013)

- Income: MLD is good
  - Inequality measured by DU = inequality measured by DG (a consequence of path-independence)
  - Qualification: true for a specific rescaling procedure to nullify the impact of effort.

Results

Look at the paper!