Decomposition of poverty indices

Alain Trannoy (AMSE and EHESS)

Prepared for the 2025 Canazei Winter School on inequality and social welfare theory

The aim

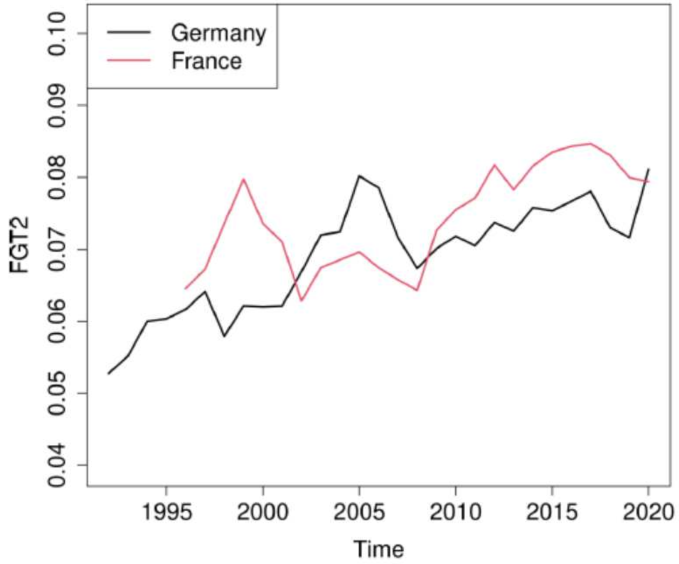
- Decomposition by subpopulations
  - FGT is decomposable
  - Watts is decomposable

- Decomposition by sources
  - Shapley decomposition
  - Nested Shapley (natural order) (Sastre Trannoy 2002)

→ • Decomposition of poverty changes between two periods or two countries

The aim: Understand changes in poverty

- Poverty changes over time or across countries



market incomes (incl. pensions)

- How?
  - Fewer/more poor individuals (poverty line)
  - The poor stay below poverty line, but get closer/further away
  - → Decompose poverty measures to better understand what drives the changes

Decomposing changes in Poverty indices

- Decomposing the variation of a poverty measure dependent on m drivers into m additive parts, each one corresponding to the contribution of one of these drivers.

- P$(d_{1,} \ldots d_{j,} \ldots, d_m)$
- $(\ldots d_{Aj} \ldots) \longrightarrow (\ldots d_{Bj} \ldots)$
- $(\ldots d_{jT} \ldots) \longrightarrow (\ldots d_{jT'} \ldots)$
- ΔP$= \sum_j \; C_j(\Delta\, d_j)$

The idea

- Taking the most from the FGT family
- The FGTs closely related to the moments of the distribution. Indeed, they are partial moments.
- Each moment: a way to look at the distribution specifically. Each first moment tells us something specific.
- Decomposing the FGTs changes into changes in the first moments

Road map

1. Literature review about decomposing poverty changes
2. New results about FGT poverty indices. The 40th anniversary!
3. Application to a comparison of French and German dynamics of poverty rates over 25 years using LIS data
4. Conclusion: Dwarfs and Giants: Application to richness indices

- 2&3 based on work in progress « *Poverty levels and trends France and Germany compared with a nested decomposition of the FGTs* »
  - Julia Baarck, Edwin Fourrier-Nicolas , David Gstrein, Alain Trannoy

1. Literature Review

OutLine

- Growth and distribution
- The Shapley perspective
- The integral-approximation approach
- Discussion and Wrap-up

Some notations

- $F(y); f(y)$
- z: $poverty\ line$
- $P(f(y); z)$
- Share $s = \frac{y}{\mu}$
- Distribution of shares: $F_s(s)$; $f_s(s) = \mu f(y)$

1.1 Growth and distribution

- Dynamic analysis: poverty change between two periods: Datt and Ravallion (1992)

- Two drivers: growth (mean income) and the income distribution (Lorenz curve, or the share distribution)

- $\Delta P(f(y); z) = P(\mu_1, f_1(s); z) - P(\mu_0, f_0\,(s); z)$

- C_{μ}= $P(\mu_1; f_0\,(s), z) - P(\mu_0; f_0\,(s), z)$

- C_I= $P(f_1(s);\; \mu_0, z) - P(f_0(s);\; \mu_0, z)$

Drawbacks

- They don’t add up to the poverty change
- The residual term captures the interaction between growth and redistribution components
- Residual, when too large, spoils the explanatory power of the decomposition
- Not symmetric in treating periods. The decomposition rule must be independent of the base period

The second attempt: Kakwani (2000)

- Properties that should respect a decomposition

i) When the growth (inequality) effect is zero, then the change in poverty must be entirely due to a change in income inequality (mean income)

ii) When both growth and inequality effects are negative (positive), then poverty should decline (increase)

iii) Time reversion consistent : the growth (inequality) effect from the initial to the final date must be the opposite of the growth (inequality) from the final to the initial date

Kakwani exact decomposition

- Averaging the formulae with base periods 0 and 1

- $C_\mu = \frac{1}{2}(P(\mu_1; f_0\,(s), z) - P(\mu_0; f_0\,(s), z))$
  $+ \frac{1}{2}(P(\mu_1; f_1\,(s), z) - P(\mu_0; f_1\,(s), z))$

- $C_I = \frac{1}{2}(P(f_1(s);\ \mu_0, z) - P(f_0(s);\ \mu_0, z))$
  $+ \frac{1}{2}(P(f_1(s);\ \mu_1, z) - P(f_0(s);\ \mu_1, z))$

1.2 The Shapley decomposition

- Shorrocks (1999) introduced it almost as a one-size-fits-all solution

- Contribution of driver j is the weighted average of marginal contributions of driver j when the change of other drivers is removed and the drivers are set to one of the two values in comparison.

- Marginal contribution: $C_j(\Delta\, d_j; \bar{d}_{-j}) =$

$$P(\bar{d}_1 \ldots, d_{jT+1}, \ldots, \bar{d}_m) - P(\bar{d}_1 \ldots, d_{jT}, \ldots, \bar{d}_m)$$

How to weigh?

- What are the weights for all possible marginal contributions of driver j?

- With 2 drivers, very simple, coincides and generalizes Kakwani

- $C_{s1}(\Delta\, d_1; \bar{d}_2) = \frac{1}{2}\left(P(d_{11}, d_{21}) \ - P(d_{10}, d_{21})\right)$
$$+\frac{1}{2}\left(P(d_{11}, d_{20}) \ - P(d_{10}, d_{20})\right)$$

- $C_{s2}(\Delta\, d_2; \bar{d}_1) = \frac{1}{2}\left(P(d_{21}, d_{11}) \ - P(d_{20}, d_{11})\right)$
$$+\frac{1}{2}\left(P(d_{21}, d_{10}) \ - P(d_{20}, d_{10})\right)$$

The weight formula

- Set of drivers: $D\ of\ cardinality$ |m|
- Subset of drivers B of cardinality |b| such that the base period is identical to that of driver j, |b|$\leq m-1$
- Complement subset of drivers of cardinal |m-1-b| such that the base period is opposite to that of driver j
- The weight = $\frac{(m-1-b)!b!}{m!}$ for the marginal contribution:

Example with 3 drivers

- Poverty line, growth and inequality
- Shapley contribution of the growth driver is the sum
- $|b| = 0;\ \frac{1}{3}(P(\mu_1; f_1(s), z_1) - P(\mu_0; f_1(s), z_1))$
- $|b| = 1;\ \frac{1}{6}(P(\mu_1; f_0(s), z_1) - P(\mu_0; f_0(s), z_1))$
- $|b| = 1;\ \frac{1}{6}(P(\mu_1; f_1(s), z_0) - P(\mu_0; f_1(s), z_0))$
- $|b| = 2;\ \frac{1}{3}(P(\mu_1; f_0(s), z_0) - P(\mu_0; f_0(s), z_0))$

Hierarchical structures

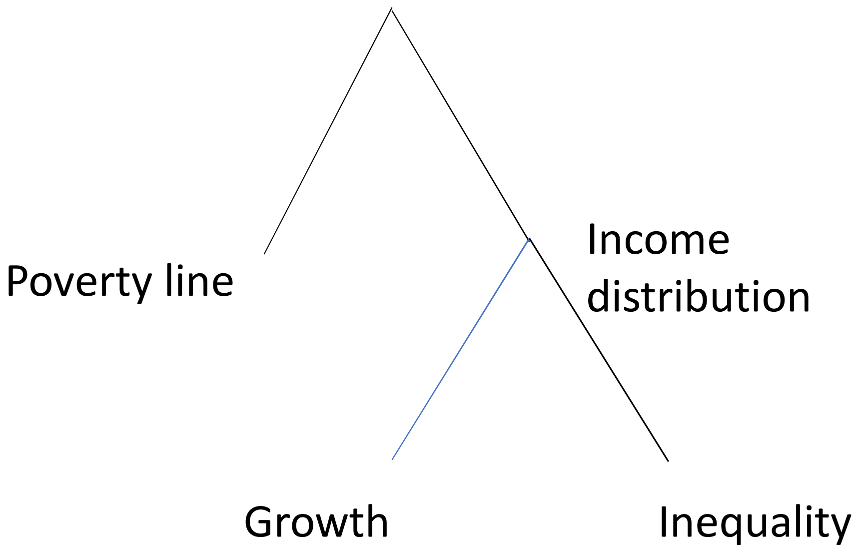
- In many applications (see Sastre and Trannoy (2002), some drivers naturally cluster together

- In the present context, with absolute poverty lines
  - Growth and distribution cluster together
  - Poverty change comes from other factors (new needs,, prices, housing, productivity shocks etc.)

- Less true for relative poverty line (60% of the median) linked to distributional change.

Hierarchical structures: example

- See Aristondo, D'Ambrosio and Lasso de la Vega (2023)



- Use Shapley decomposition at each level between two drivers
- With a 2x2 hierarchical structure, the drivers' weights of stage 2 are ¼.

1.3 The integral-approximation approach (Muller 2006)

- The case of a temporal decomposition between two times
- Assumption that there is a continuous process, a path linking the situation in t and the situation in t+1.
- All functions appearing in the inequality index are a function of time with initial and terminal dates.
- Nice interpretation/extension, but we don’t know the values of these functions except at initial and terminal dates.

The variation of poverty between two dates as a definite integral

- Applying the second fundamental theorem of calculus

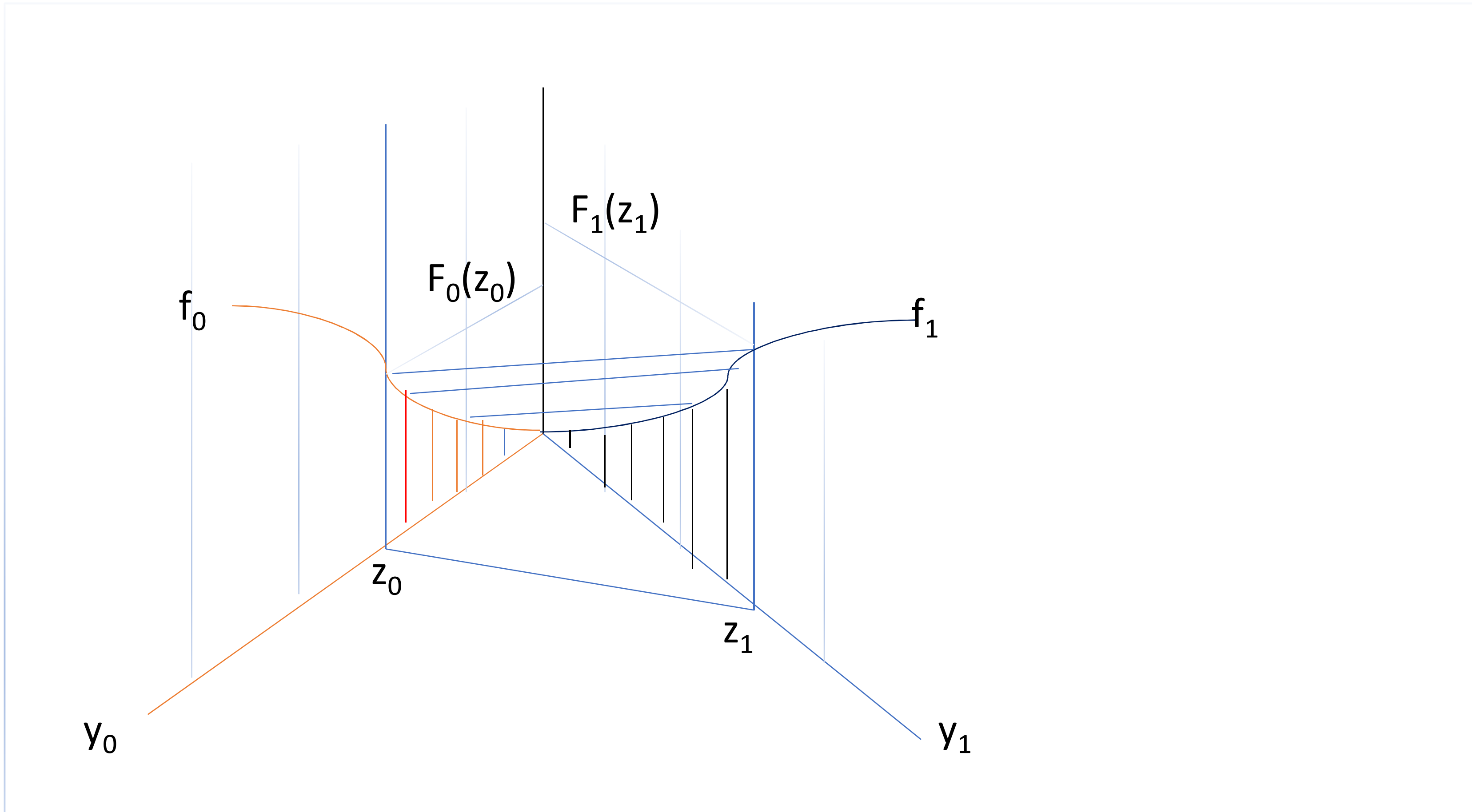
$$\mathrm{P}(d_{T+1}) - \mathrm{P}(d_T) = \int_T^{T+1} \frac{dP}{dt} dt$$

$$= \sum_{j=1}^{m} \int_T^{T+1} \frac{\partial P}{\partial d_j} \frac{\partial d_{j,}}{\partial t} dt$$

Caveat

- Only for temporal decomposition not for a static decomposition (between countries, regions, cities)
- The notion of a path connecting countries generally does not make sense.
- But even for temporal variation, we don’t know the paths. We only know the end-points
- Solution: resort to a linear approximation

3D visualization for the Headcount



$$\int_0^1 \frac{\partial P}{\partial \mu(t)} \frac{\partial \mu(t)}{\partial t} dt + \int_0^1 \frac{\partial P}{\partial f_s} \frac{\partial f_s(t)}{\partial t} dt$$

We approximate the time derivative of the drivers by the slope of the straight line joining the end-points

$$\frac{\partial \mu(t)}{\partial t} = \frac{\mu_1 - \mu_0}{1}$$

We use a trapezoid approximation given by joining the upper and lower end-points with a straight line

$$\int_0^1 \frac{\partial P}{\partial \mu(t)} dt \approx \frac{1}{2} (\frac{\partial P}{\partial \mu(t)} (\mu_1, f_{s_1}, z) + \frac{\partial P}{\partial \mu(t)} (\mu_0, f_{s_0}, z))$$

Now we approximate each derivative with respect to the average by the slope of the straight line joining the end-points

$$\frac{\partial P}{\partial \mu(t)} (\mu_1, f_{s_1}, z) \approx \frac{1}{2} \frac{P(\mu_1; f_{s_1}, z) - P(\mu_0; f_{s_1}, z)}{\mu_1 - \mu_0}$$

$$\frac{\partial P}{\partial \mu(t)} (\mu_0, f_{s_0}, z) \approx \frac{1}{2} \frac{P(\mu_1; f_{s0}, z) - P(\mu_0; f_{s_0}, z)}{\mu_1 - \mu_0}$$

Back to the future: Kakwani !

Finally

$$\int_0^1 \frac{\partial P}{\partial \mu(t)} \frac{\partial \mu(t)}{\partial t} dt = \frac{1}{2}(P(\mu_1; f_{s_1}, z) - P(\mu_0; f_{s_1}, z) + P(\mu_1; f_{s0}, z) - P(\mu_0; f_{s_0}, z))$$

- Much do about nothing?
- Interesting to derive a quite ad hoc formula from rigorous reasoning
- When more than two drivers?

Shapley cannot be interpreted in that way for m>2

- With a continuous path from T to T+1, base periods for *ceteris paribus* drivers should be the same

- With Shapley, the base period for the « ceteris paribus » drivers may be not the same
  - For the contribution of the poverty line, choosing mean at T and share distributions at T+1

- Shapley marginality (discrete) different from marginality for partial derivative (continuous) when m>2

A drawback: Violation of Subperiod additivity?

- $C_j(\Delta_{02}\ d_j)$=$C_j(\Delta_{01}\ d_j)$+$C_j(\Delta_{12}\ d_j)$

- All previous decompositions do not respect this condition

- Is it a problem? See Kakwani (2000), Bresson (2008) and Fujii (2017)

- In the integral approximation approach, the question should be framed about knowing whether it should be better to use a second-order approximation for two-period interval with using the information about the mid-point than a first-order approximation for each period

Wrap-up

- Shapley and integral approach with first-order approximation coincide with two drivers

- They also coincide with hierarchical structures when each level contains at most two drivers.

- With more than two drivers? Shapley decomposition is a default solution when there is no reason to introduce some natural order between drivers

- All this works for the headcount! Decomposition is done

2. 40th Anniversary of FGT indices (1984)

$$
\begin{aligned}
P_\alpha(y;z) &= \int_0^z \left(\frac{z-y}{z}\right)^\alpha f(y)dy \\
P_0(y;z) &= \int_0^z f(y)dy \\
P_1(y;z) &= \int_0^z \left(\frac{z-y}{z}\right) f(y)dy \\
P_2(y;z) &= \int_0^z \left(\frac{z-y}{z}\right)^2 f(y)dy
\end{aligned}
$$

Roadmap

- Nested writing of the FGTs
  - FGT(α) in function of FGT(0)…FGT(α-1) and other terms
- Decomposition of the FGTs change (two periods or two countries)
  - Using the nested writing
  - The symmetric (both ways) Oaxaca-Blinder decomposition

1. Nested writing of the FGTs

- A story of moments
- FGT 1 in function of FGT0
- FGT 2 in function of FGT1, FGT0 + ?
- FGT 3 in function of FGT0, FGT 1 and FGT 2 + ?

Partial moments and FGTs

- The FGTs are close to partial moments sometines referred as "one-sided moment".

- The alpha-th order lower partial moments with respect to a reference point r are expressed as

$$m_{p\alpha}(y;r) = \int_0^r (r-y)^\alpha f(y)dy$$

- When r is a poverty line, the partial moment is a FGT defined in terms of the absolute poverty gap

Decomposition of FGT1 (Intensity)

$$P_1 = P_0(1 - \frac{\mu_z}{\mu}\frac{\mu}{z})$$

- $\frac{\mu_z}{\mu}$: The growth factor (G), an indicator of pro-poorness
  - If the growth is pro-poor, then intensity ↓
- $\frac{\mu}{z}$: The unchallengingness (U) of poverty policy
  - If the objective of the poverty policy becomes less challenging, then intensity ↓
  - When z is a fraction of the median, U is a crude measure of inequality: the mean median ratio
- Why inequality not important for P_1? Because it is insensitive to within-poor inequality

Normalized distribution

- Let us define the normalized density as the density normalized to the poverty gap

$$f_z(y) = \frac{f(y)}{P_0}$$

- The normalized average poverty gap

$$P_1^* = \frac{P_1}{P_0} = \int_0^Z \left(\frac{z-y}{Z}\right) f_z(y)\,dy$$

A tale of two distributions y and g

- The normalized poverty gap distribution g

$$g(y) = \frac{z-y}{z} \text{ with } f_z(y)$$

- Comparison of the moments of the distribution g and the moments of the distribution y

$$m_g^\alpha = \frac{m_z^\alpha}{z^\alpha}$$

for $\alpha \geq 2$ and

$$m_g^1 = \mu_g = (1 - \frac{\mu_z}{z^\alpha})$$

Decomposition of FGT2 (Severity)

$$P_2 = P_0(P_1^{*2} + \sigma_g^2)$$

- Inequality plays a role through the variance among the poor on an equal footing as the square of the normalized poverty gap.
- Absolute index of inequality because FGT2 is absolute
- P_1^* and σ_g^2 are independent from the headcount
- Neater formula that was obtained by FGT 1984

$$P_1 = P_0(P_1^{*2} + (1 - P_1^*)^2 Cv_z)$$

- Plugging the expression of P_1^*

$$P_1 = P_0((GU)^2 + \sigma_g^2)$$

Decomposition of FGT3

$$P_3 = P_0[P_1^*[P_2^* + 2\sigma_g^2] - \widetilde{\gamma}_g]$$

- Introducing unstandardized skewness among the poor, $\widetilde{\gamma_g}$

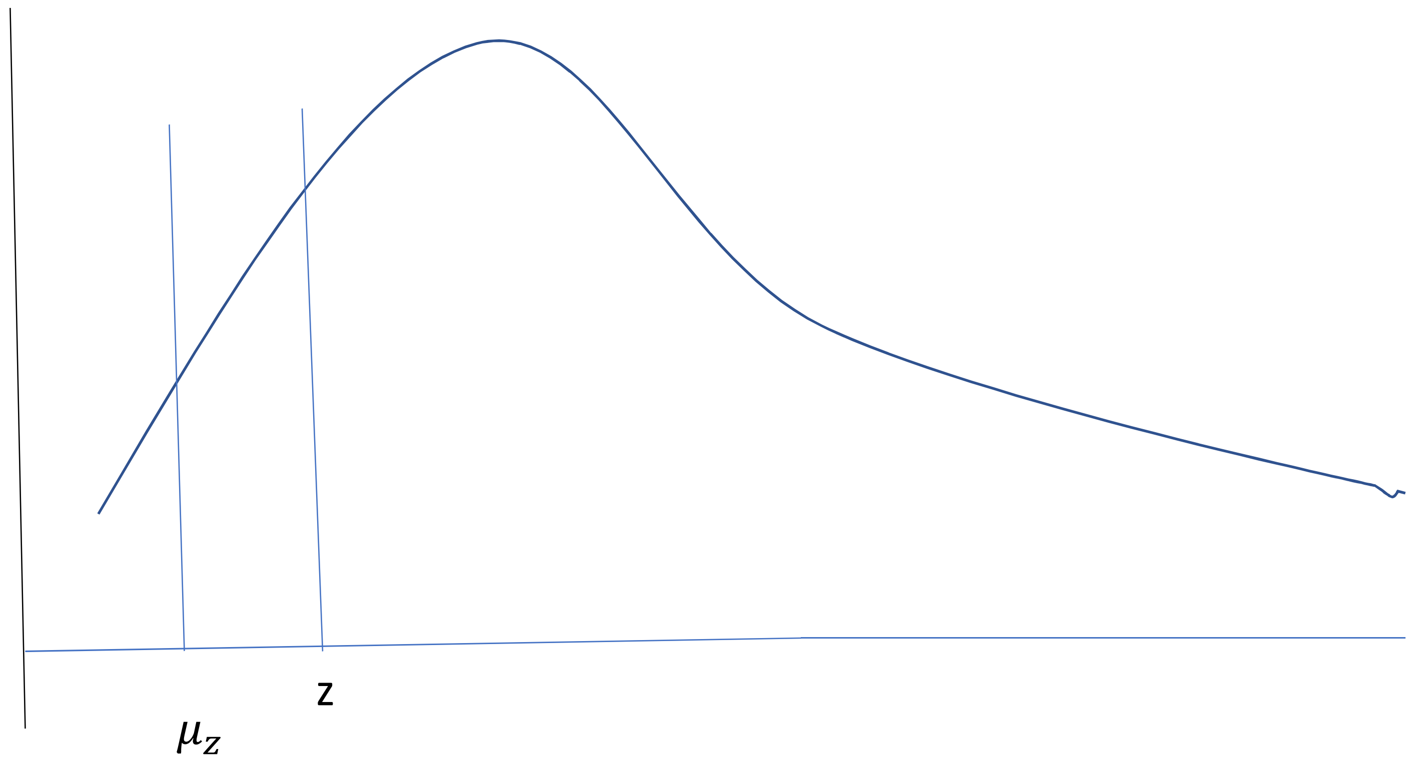
- Lower FGT indices enter into a multiplicative way

- Variance has more weight in addition to that already included in FGT2 (on total a weight three times larger)

- Why a negative sign for skweness?

Skewness of distribution among the poor

When there is positive skewness, it means that the left tail is short (not so many very poor) Then it appears as a mitigating factor



Decomposition of FGTs changes

Over space or time

General formula for changes of product of terms

- The FGT can be written as products of terms or products of sums of terms.

- X=YZ between two countries, two years

- With $\Delta X = \mathrm{X}_a - \mathrm{X}_b$, $\bar{Y} = \frac{\mathrm{Y}_a + \mathrm{Y}_b}{2}$, $\bar{Z} = \frac{\mathrm{Z}_a + \mathrm{Z}_b}{2}$

- Then $\Delta X = \Delta Y \bar{Z} + \Delta Z \bar{Y}$

Decomposition of the poverty gap change

- Headcount change

- Growth change

- Unchallengingness change

- $\Delta P_1 = \Delta P_0(1 + \overline{GU}) - \Delta G\overline{P_0 U} - \Delta U\overline{P_0 G}$

Decomposition of the severity change

- Headcount change

- Intensity change

- Poor Variance

$$\Delta P_2 = \Delta P_0(\bar{P_1^{*2}} + \bar{V_z}) + \Delta P_1^{*2}\bar{P_0} + \Delta V\bar{P_0}$$

Decomposition of FGT3 change

- Headcount change; Poverty gap change
- Severity change; Poor variance change
- Skweness change

$$
\begin{aligned}
\Delta P_3 \quad = \quad & \Delta P_0(\overline{P_1^* P_2^*} + \overline{P_1^*}2\overline{\sigma}_g^2 - \overline{\widetilde{\gamma}_g}) + \Delta P_1^*(\overline{P_0 P_2^*} + \overline{P_0}2\overline{\sigma}_g^2) \\
& + \Delta P_2(\overline{P_0 P_1^*} + \Delta 2\sigma_g^2 \overline{P_0 P_1^*}) - \Delta\widetilde{\gamma}_g \overline{P_0}
\end{aligned}
$$

3. French-German comparison

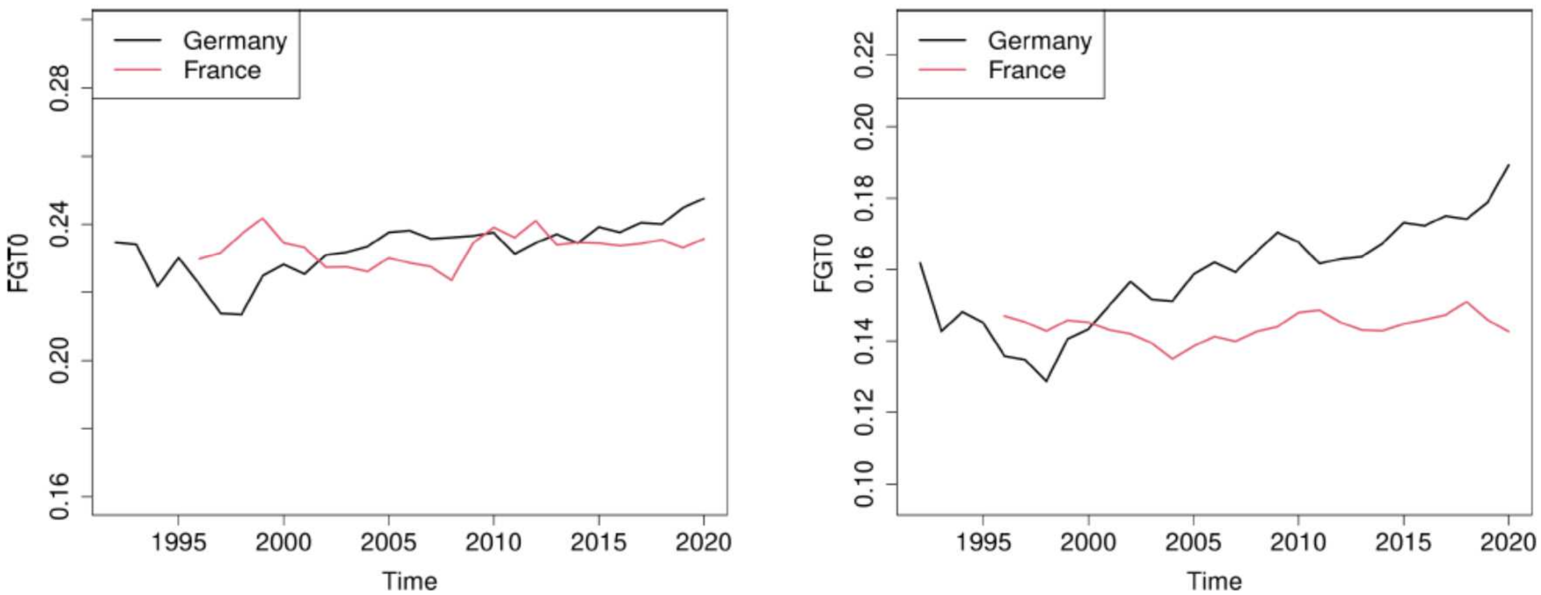
From German Reunification to Covid

Data

- LIS data from 1992-2020
- Exclusion of negative incomes and students
- France: 1.2 million households
- Germany: about 0.4 million
- Equivalized Income
- Inflation adjusted using ECB Harmonized CPI
- Market income (including pensions and unemployment benefits)
- Disposable income accounting for taxes and transfers
- Relative poverty line

Headcount P_0

Figure 1: Poverty headcount

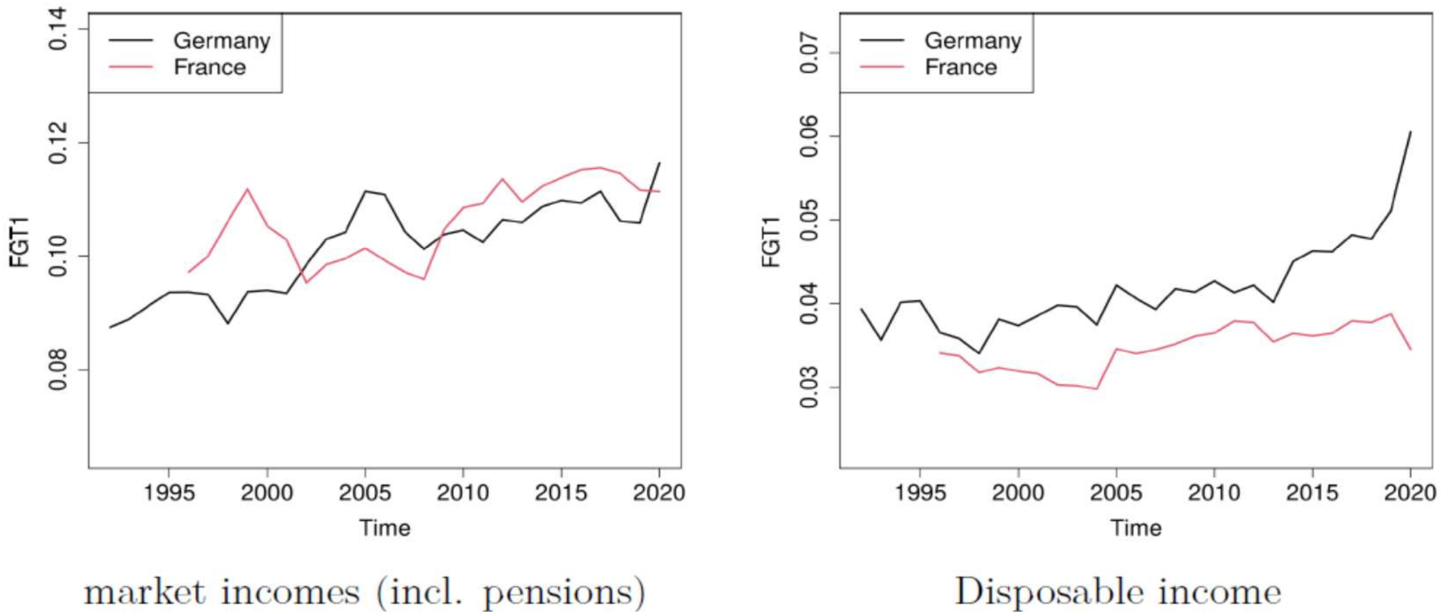


market incomes (incl. pensions)

Disposable income

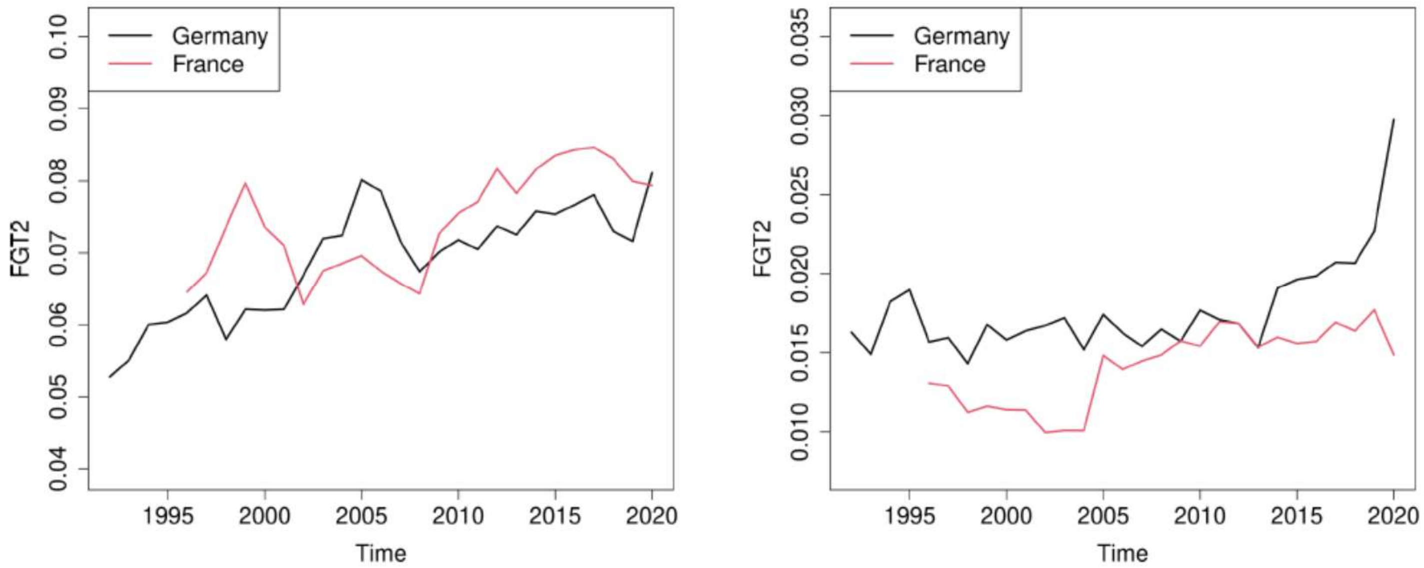
Poverty Intensity P_1

Figure 2: Average poverty gap



Poverty severity P_2

Figure 5: Poverty severity



market incomes (incl. pensions) Disposable income

What have we learnt?

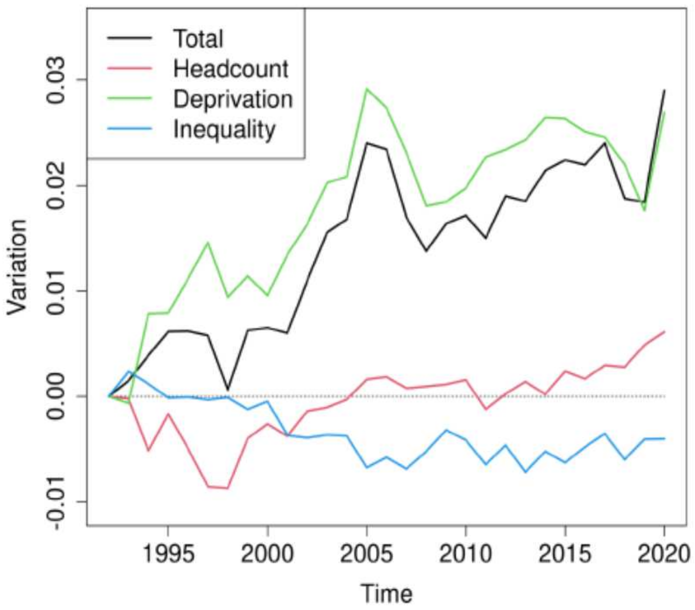
- For market incomes, same evolution in both countries
- Except for the headcount, the poverty is on the rise in both countries
  - *Conjecture*: the same forces are on play in both countries

- For disposable incomes, France almost succeeds in containing poverty,
- Not the case for Germany
  - More than that, the headcount increases for disposable incomes whereas it is constant for primary incomes

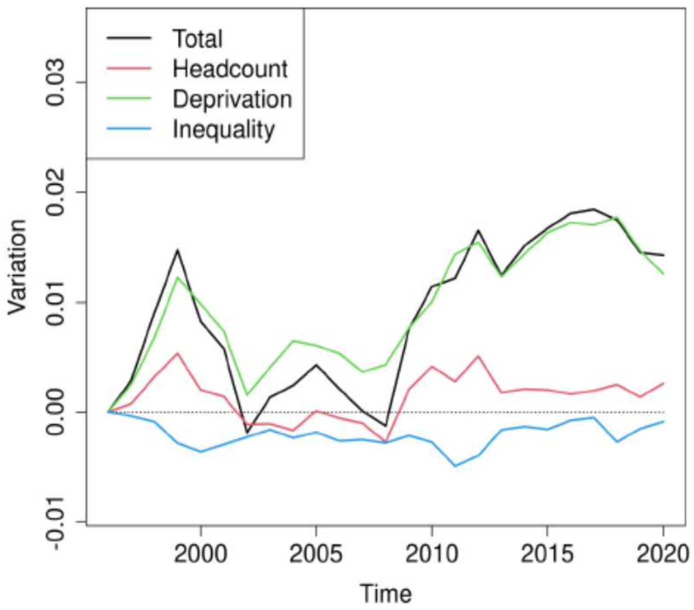
  - *Conjecture*: the implemented policies/budgets followed by the welfare state are different in both countries

Decomposition of the temporal change of poverty intensity for both countries (market)

Figure 3: Decomposition of the average poverty gap (market income)



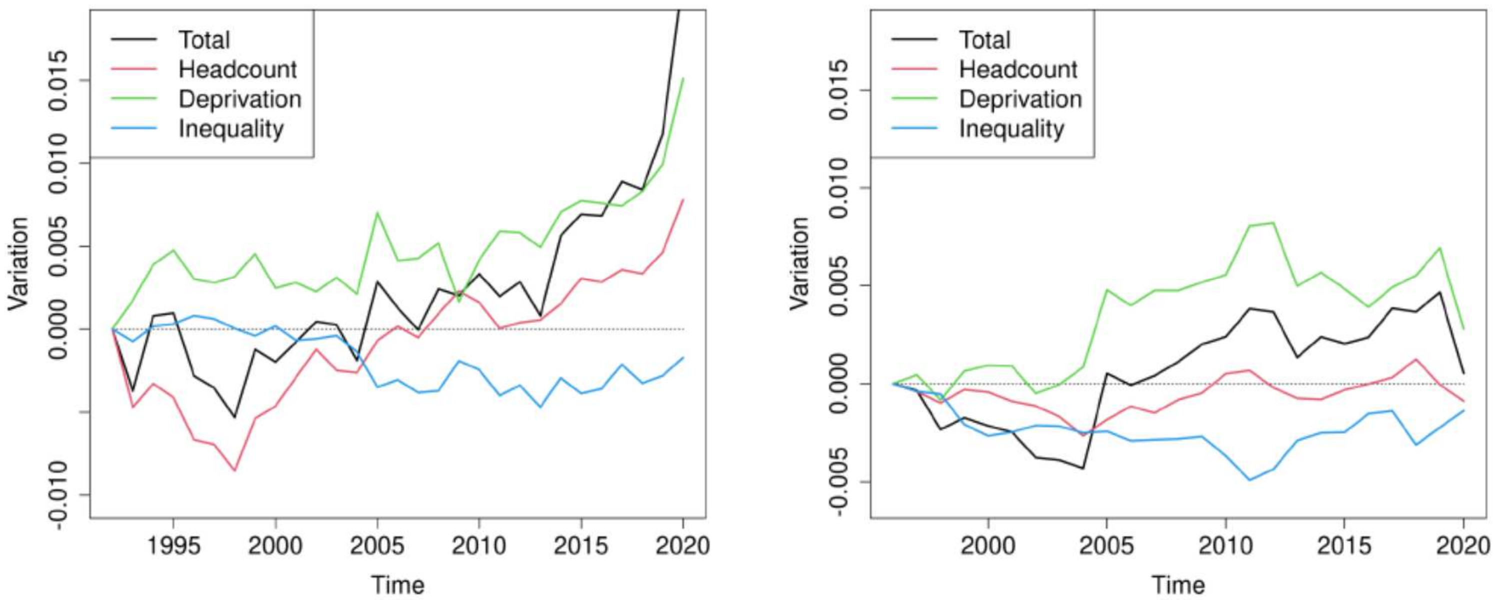
Germany



France

Decomposition of the temporal change of poverty intensity for both countries (disposable income)

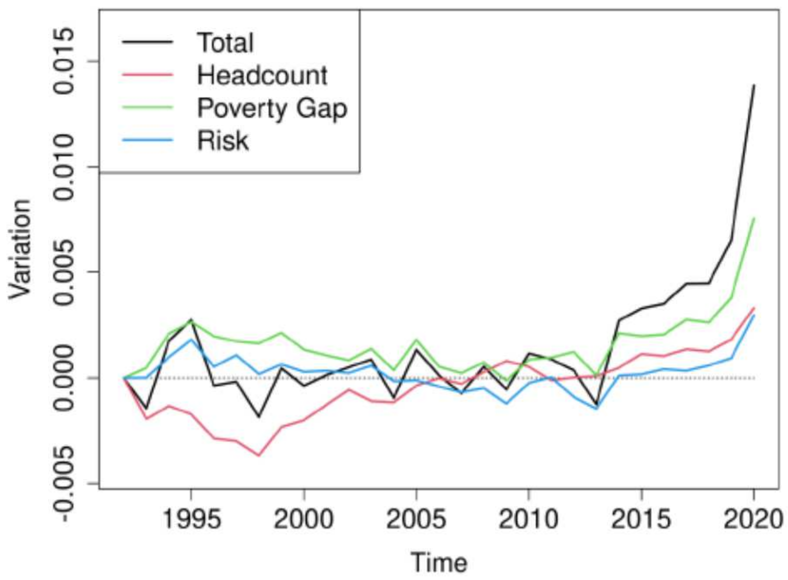
Figure 4: Decomposition of the average poverty gap (disposable income)



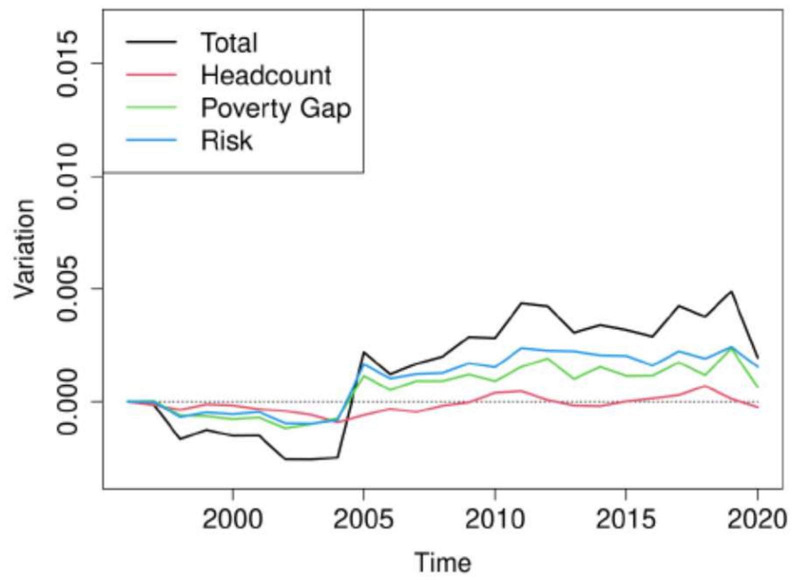
Germany France

Decomposition of the dynamic change of poverty severity for both countries (disposable income)

Figure 7: Decomposition of poverty severity (disposable income)



Germany



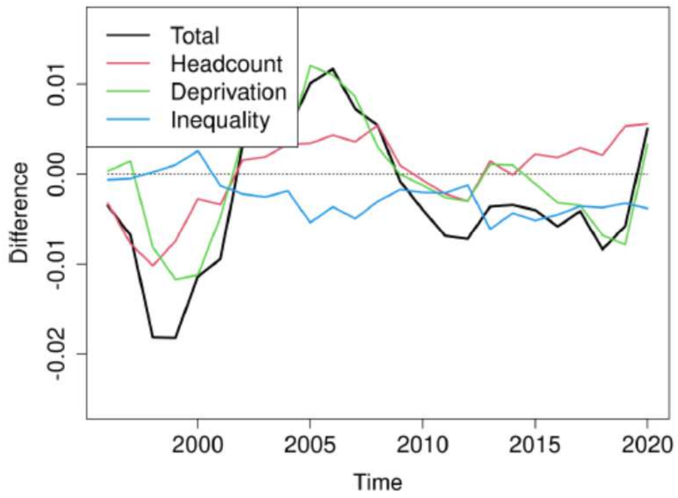
France

Lessons

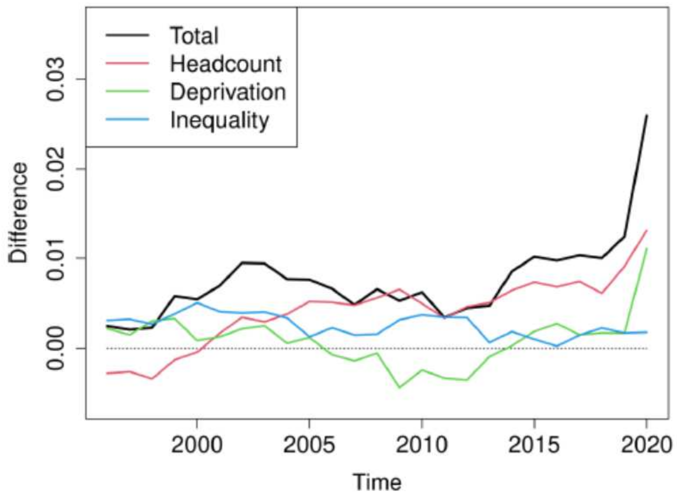
- *For intensity*
  - *Market income:* The main driver is growth. The growth has been antipoor
  - *Disposable income:* same conclusion, but on top of that in Germany an additional factor with the increase of the headcount.
  - *Mean-median gap has not increased, quite the contrary!*
- *For severity, the same pattern for both market and disposal*
  - In Germany, since 2014, the Euro crisis, all indicators are on the rise to explain the surge
  - In France, nothing to say except the break in the series in 2004-2005

Decomposition of the French-German deviation each year for P_1

Figure 10: Decomposition of FGT(1) difference in Germany vs. France



Market income



Disposable income

Pos. values: France has lower poverty gap
Neg. values: Germany has lower poverty gap

Lessons

- *Market income*
  - In the first decade of the 20th century, the growth factor was the main driver (Harz reforms?)
  - In the second decade, the headcount was the main driver (immigration wave in Germany?)

- *Disposable Income*
  - Over the years, the evolution of headcount was the main driver.
  - Further investigation to see the role of new immigrants, particularly in Germany

4. About Dwarfs and Giants

Cousin family of richness indices

- The affluent line: a

$$R_\alpha(y;a) = \int_a^{+\infty} \frac{(y-a)^\alpha}{a} f(y)dy$$

A change for R_3

$$R_3 = R_0[R_1^*(R_2^* + 2\sigma_g^2) + \tilde{\gamma}_g^3]$$

- Positive skweness is a deepening factor

Endogenizing the affluent line

- What is a rich: a rich is someone sufficiently rich to contribute to reduce poverty

- The rich are such that if their income is erased to the affluent line and redistributed to the poor, poverty disappears according to the absolute poverty gap.

a root of the

$$R_1(y;a) = \int_a^{+\infty} (y-a) f(y) dy = \int_0^z (z-y) f(y) dy$$

- a/z: an indicator of rich-poor disparity