How Unequal Is Europe?
Evidence From Distributional National Accounts, 1980–2017

Thomas Blanchet¹,²  Lucas Chancel²,³  Amory Gethin²

OECD
ELS Seminar
September 10th, 2019

¹ Paris School of Economics – EHESS
² World Inequality Lab
³ IDDRI
Unlike some other parts of the world, there is quite a lot of data available on the distribution of income in Europe.

So why write a paper on the topic?

- Not because of a lack of data *per se*...
- ... but because existing data is scattered across a variety of sources, so it can be hard to make sense of it.

Disparate set of indicators:

- hard to compare
- hard to aggregate
- hard to tell consistent stories
A Variety of Sources

- **National Accounts**
  - Income concept of reference
  - Inequality between countries, but not within countries

- **Surveys**
  - Microdata: good geographical coverage, but only in recent years
  - Tabulations: better coverage in older years, but inconsistent concepts and units
  - Covers the whole distribution, but with large biases at the top

- **Tax data**
  - Better at covering the top of the distribution
  - Covers only the top of the distribution
  - Inconsistent concepts and units
Many Questions Lack a Clear Answer

⇒ Literature has struggled to answer simple questions:

- Is Europe as a whole more or less unequal than the United States?
- Is the difference between European and US inequality driven by pre-tax incomes or redistribution?
- Is European inequality driven by the distribution of income between or within countries?
- Which parts of the distribution have benefited the most from European growth?

⇒ Difficulties monitoring of internationally agreed goals

- European institutions’ Pillar of Social Rights (2017)
- Sustainable Development Goals adopted by all European countries (2015)
This paper is an attempt to address this problem by constructing distributional national accounts (DINA) for Europe since 1980.

Part of larger ongoing project in which we construct similar estimates for as many geographical areas as possible, over as long as possible.

This paper contributes to that project in several respects:

- Productions of new estimates of the income distribution within the DINA framework.
- Methodological advances to build such estimates in spite of data limitations.
Conceptual Framework
A Short History of DINA


- Work extended to large number of countries by many researchers, and collected in two collective volumes edited by Piketty & Atkinson (2007, 2010).

- Data collected into the World Top Income Database (WTID), which then became the World Inequality Database (WID.world), maintained by the World Inequality Lab.

- Estimated based on raw tax data have many qualities, but also many drawbacks:
  - Inconsistent concepts
  - Only covers the top of the distribution
  - Does not take redistribution into account
The DINA project was created to address those criticisms.

Set of guidelines for measuring inequality:
- consistent concepts.
- consistent methodology.
- consistent with the framework put forth by the System of National Accounts (SNA).

Combine existing sources (surveys, tax data, national accounts).

Be consistent the macro totals and distribute 100% of national income.

Describe both the top and the bottom of the distribution.

Measure several income concepts, including pre-tax and post-tax inequality.
Comparison With Other Initiatives

• Some papers have used harmonized income statistics (EU-SILC) to estimate the distribution at the European level, but relying solely on survey data (Filauro (2018) and Brandolini and Rosolia (2019)).

• Several initiatives to address gaps between micro estimates and macro totals:
  • EG-DNA since 2011 at the OECD.
  • Various “experimental statistics” published by Eurostat (2018) and other national statistical institutes.

• Mostly rely on survey data.

• Only distribute a fraction of national income.
Comparison With Other Parts of the DINA Project

Several highly detailed country-specific studies:

- United States (Piketty, Saez, & Zucman), France (Garbinti, Goupille-Lebret & Piketty, 2018), . . .
- Very precise work, very detailed decomposition, but takes a lot of time. . .
- Will be very long until we can apply such work universally.

Our work seek to provide roughly comparable estimates:

- At a lower cost, using less precise data and simpler assumptions.
- (On the plus side, using more straightforward and transparent assumptions.)
We construct DINA Estimates for 38 countries over 1980–2017.

- Consistent statistical unit (equal-split adults).
- Pre-tax and post-tax income.
- From the bottom to the top 0.001%.

We face numerous challenges in constructing but we try to overcome them in a meaningful and transparent way.

- All the data is public and available on WID.world.
- All the codes are also available on WID.world.
- Detailed appendix describing the sources and specific assumptions for each country.
Distributing all of national income means accounting for income that may never explicitly appear on the bank account of any household, yet is still part of “economic growth”.

\[
\text{net national income} = \text{GDP} - \text{depreciation} + \text{net foreign income}
\]

We consider income from all sectors of the economy:

- Income of the household sector (incl. imputed rents)
- Income of the government:
  - before taxes: production taxes
  - after taxes: taxes - transfers + government consumption
- Income of corporations (retained earnings)
Standard Inequality Statistics Miss a Large Part of Income

about half of capital income not captured by common inequality statistics

0% 5% 10% 15% 20% 25%
Share of net national income

interests and dividends
capital component of mixed income
imputed rents
retained earnings
corporate tax

Year
Yet “missing income” is still income

Not accounting for income outside of the household sector sometimes yields undesirable results:

- The owners of corporation can choose to distribute income to themselves arbitrarily.
- Some taxes are accounted for (e.g. income tax) but not others (e.g. VAT).
- Countries with strong provision of public goods appear to have poorer households.

We make simple and transparent assumptions to distribute these types of income. Still a lot of room for improvement.
DINA concepts: pre-tax and post-tax income

• **pre-tax income**: income after the operation of social insurance systems (pension and unemployment insurance), but before other types of transfers. We add pensions and unemployment benefits, but remove the social contributions that pay for them. Conceptually similar to “taxable income” in many countries.

• **post-tax**: income after the operation of all government redistribution. We remove all taxes, the remaining social contributions, and then add all transfers and government consumption.
Methodology
A myriad of income inequality datasets in Europe

Macro data:

- UN, OECD, Eurostat

Survey microdata:

- Eurostat surveys: SILC, ECHP, HBS
- Luxembourg Income Study (LIS)
- Imputation of social contributions

Survey tabulations:

- Diverse sources, compiled in WIID by the UN.
- We turn them into complete distributions using generalized Pareto interpolation (Blanchet, Fournier & Piketty, 2018)

Tax data:

- WID.world
- Couple of formerly unused sources: Iceland, East Germany.
Data availability: good coverage for most of the population

- Low (survey tabulations, low time coverage)
- Medium low (survey tabulations, good time coverage)
- Medium (survey tabulations + microdata)
- Medium high (survey tabulations + microdata + tax data)
- High (distributional national accounts)
Three reasons why these data sources yield different results:

- **Sampling error**: survey sample size too small to properly capture top income groups
- **Conceptual differences**: data refers to different concepts:
  - income concepts
  - statistical unit: individuals, households, per capita, per adults, equivalence scales
- **Non-sampling error**: heterogeneous non-response in surveys, misreporting
Sampling error makes it difficult to study the distribution within the top 10%

- extreme value theory provides tools to infer shape of the top tail with limited samples sizes
- top tail modelled using the generalized Pareto distribution

\[
P\{X > x\} = \left[1 - \xi \left(\frac{x - \mu}{\sigma}\right)\right]^{-1/\xi}
\]

- estimated using the method of probability-weighted moments (Greenwood et al., 1979)
Harmonizing concepts using machine learning

- Use existing survey data to establish correspondences between the different concepts used in the literature:
  - 3 income concepts: pre-tax income, post-tax income, consumption
  - 5 statistical units: households, per capita, per adult, OECD equivalence scale, square root equivalence scale
  - $3 \times 5 = 15$ distributions by year and country

- We train a machine learning algorithm to map how all these different concepts relate to one another, and then use it to correct for systematic biases.
Harmonization of Concepts: Example

Top 10% income share, Portugal

Pre-tax income per adult (EU-SILC)
Harmonization of Concepts: Example

Top 10% income share, Portugal

Pre-tax income per adult (EU-SILC)
Post-tax income per adult (EU-SILC)
Harmonization of Concepts: Example

Top 10% income share, Portugal

Pre-tax income per adult (EU-SILC)
Post-tax income per adult (EU-SILC)
Post-tax income per adult (ECHP)
Harmonization of Concepts: Example

Top 10% income share, Portugal

Pre-tax income, adults (EU-SILC)
Post-tax income, adults (EU-SILC)
Post-tax income, adults (ECHP)
Post-tax income, square-root scale (Atkinson, Rainwater and Smeeding 1995)
Post-tax income, individuals (Gouveia and Tavares 1995)
Post-tax income, households (Gouveia and Tavares 1995)
Harmonization of Concepts: Methodology

- Discretize the quantile function of distributions, normalized by the average: \( Q_1, \ldots, Q_n \)
- Statistical model between distributions A and B:

\[ Q_k^A = f(p_k, Q_1^B, \ldots, Q_n^B, Z) + \varepsilon_k \]

where \( Z \) is a set of additional variables to improve prediction (regional dummy, average income, demographic composition, tax schedule, etc.)

The problem is hard statistically because:

- Non-parametric functional form.
- Large number of highly correlated predictors.
- Monotonicity constraint on \( Q_k^A \).

\[ \Rightarrow \] machine learning is very good at dealing with this type of problem: we use a standard, state-of-the-art algorithm known as gradient tree boosting
Harmonization of Concepts: Example

Top 10% income share, Portugal

- Pre-tax income, per adult (harmonized)
- Post-tax income, per adult (harmonized)
- Pre-tax income, adults (EU-SILC)
- Post-tax income, per adult (EU-SILC)
- Post-tax income, per adult (ECHP)
- Post-tax income, square-root scale (Atkinson, Rainwater and Smeeding 1995)
- Post-tax income, per capita (Gouveia and Tavares 1995)
- Post-tax income, households (Gouveia and Tavares 1995)
Correcting for misreporting and nonresponse using tax data

- Adapt the standard survey calibration framework (age, gender, ...) to improve the representation of the rich (Blanchet, Flores & Morgan, 2018).
- Adjust weights to enforce constraints on the top income shares while minimizing distortion from the original survey:
  \[
  \min_{f^*} \int \frac{[f^*(x) - f(x)]^2}{f(x)} f(x) \, dx \quad \text{st. income shares from tax data}
  \]
- Empirically, the reweighting depends on the inequality in the survey.
- When no tax data is available, we impute a reweighting profile using the same machine learning method as before.
Distribution of incomes not captured by survey or tax data

Imputed rents:

- Total of imputed rents from the national accounts.
- Statistical matching using the distribution of imputed rents recorded in EU-SILC.

Retained earnings:

- Distribution proportional to stock ownership.
- Using statistical matching with data from HFCS.
Production taxes:

- Taxes on products distributed proportionally to consumption (from HBS).
- Other taxes on production proportional to income.

Government consumption:

- Benchmark assumption: proportional to income (distribution neutral)
- Experiment with other assumption to assess how government consumption affects inequality.
United Kingdom: meaningful survey correction but little impact of missing incomes
Sweden: mild survey correction but significant role for missing incomes
Results
Large inequalities persist between European countries.

Average national incomes in Europe, 2017

- 850€ / month
- 1500€ - 2000€ / month
- 2500€ - 4000€ / month
- 5000€ / month
- 5000€ / month
- 5000€ / month

Norway
Luxembourg
Switzerland
Iceland
Ireland
Netherlands
Germany
Denmark
Austria
Belgium
United Kingdom
Sweden

850€ / month
1500€ - 2000€ / month
2500€ - 4000€ / month
5000€ / month
In the long run, no clear trend towards convergence

Average national incomes of European regions, 1980-2017
Inequalities within European countries

Top 10% income shares across European countries: 1980 vs. 2017

- Decreasing inequality
- Increasing inequality
Inequalities increased in nearly all countries

Top 10% income shares across European countries: 1980 vs. 2017
Inequality rose more in Eastern and Northern Europe

Share of national income received by top 10%, 1980
Inequality rose more in Eastern and Northern Europe

Share of national income received by top 10%, 1990
Inequality rose more in Eastern and Northern Europe

Share of national income received by top 10%, 2000
Inequality rose more in Eastern and Northern Europe

Share of national income received by top 10%, 2007

[Map showing the distribution of income inequality across Europe, with a color scale ranging from green (low inequality) to red (high inequality).]
Inequality rose more in Eastern and Northern Europe

Share of national income received by top 10%, 2017
Inequalities increased in all regions, but at different speeds.

Top 10% income shares in European regions, 1980-2017
Inequalities between European citizens
Inequality has risen between European citizens since 1980...

Top 10% vs. Bottom 50% income shares, 1980-2017
...especially at the top of the distribution

Total income growth by percentile, 1980-2017

Bottom 50% captured
15% of growth

Top 1% captured
17% of growth
European inequalities remain geographically concentrated…

**Geography of European inequality, 2017**

Population share (%) vs. Income group (percentile) for Eastern Europe, Southern Europe, Other Western, UK, France, Germany, and Scandinavia.
... but within-country inequality remains much more important

Between-country vs. within-country inequality in Europe

- Top 10% income share
- ... assuming perfect equality between countries
- ... assuming perfect equality within countries
... but within-country inequality remains much more important.

Theil decomposition of European income inequality, 1980-2017

Theil index

Inequality within countries

Inequality between countries
Redistribution in Europe
Taxes and transfers reduce inequality less in the East where pretax poverty is higher.

Pre-tax vs. post-tax ratio of top 10% to bottom 50% average incomes

Western Europe: -29%
Eastern Europe: -15%
Southern Europe: -23%
Northern Europe: -23%
The role of redistribution remains limited

2017 (before taxes)
Top 10% make 8x more than bottom 50%

2017 (after taxes)
Top 10% make 6x more than bottom 50%

Ratio of top 10% to bottom 50% average income

Pre-tax inequality
Post-tax inequality
Is Europe more unequal than the US?
Spatial inequalities remain stronger in Europe than in the US...

US vs. Europe ratio of top 10% to bottom 50% average state / country incomes, 1929-2017

Europe (2017)
Top 10% countries earn 2.8x more than bottom 50%

United States (2017)
Top 10% states earn 1.5x more than bottom 50%
... yet overall inequalities are much higher and have increased much more in the US
... yet overall inequalities are much higher and have increased much more in the US
Poor Europeans have benefited much more from growth

US vs. Europe Bottom 50% income growth, 1980-2017

- US Bottom 50% growth: +3%
- Europe Bottom 50% growth: +40%
These differences have little to do with redistribution.

Europe vs. US ratios of top 10% to bottom 50% average incomes

(a) Pre-tax income

(b) Post-tax income
Conclusion

- A new public database to study income inequality in Europe fully consistent with national accounts.
- Inequality increased in nearly all EU countries since 1980, but at different speeds (North & East vs. South & West).
- Despite rising income disparities, EU still much more equal than the US, before and after taxes.
- Inequality in Europe is much more about within-country than about between country inequality.
  - Yet EU policies essentially about between-country convergence (with relatively small impact).
Compared to existing databases

<table>
<thead>
<tr>
<th>Year</th>
<th>Eurostat (weighted average)</th>
<th>DINA (weighted average)</th>
<th>DINA (aggregated distribution)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1980</td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>1985</td>
<td>3</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>1990</td>
<td>4</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>1995</td>
<td>5</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>2000</td>
<td>6</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>2005</td>
<td>7</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>2010</td>
<td>8</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>9</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
⇒ https://wid.world/