Convergence of income distributions in Europe – a framework for analysis

Zsombor Cseres-Gergely and Virmantas Kvedaras

European Commission, Joint Research Centre

9-10 November 2018, Szirák

Agriculture and global food security

Economic and monetary union

Low carbon economy and resource efficiency

Single market, growth, jobs and innovation

Public health, safety and security

Joint Research Centre

the European Commission's
in-house science service

Joint Research Centre

Motivation

1. Convergence is a central aim of the European Union. Structural funds have been deployed for many years to facilitate (upwards) convergence. Testing convergence of more than the average of income and monitoring the process is important too, now in context of the 'European Pillar of Social Rights'.

2. Mainstream economic growth theories predict convergence. Tests are mostly related to mean income / GDP of countries or regions. Results on convergence in income distribution are few and simple for Europe.

   Aim: create a framework to test and characterise convergence in income distributions and apply it to countries within the EU.

   Warning: work heavily in progress!

Previous work: between-unit convergence - where?

Comparison of areas and countries within the Union over time and over the income distribution.
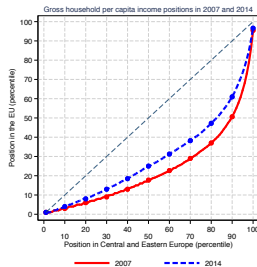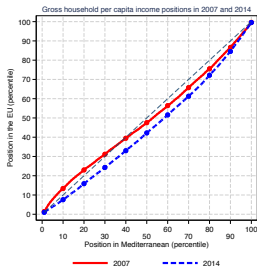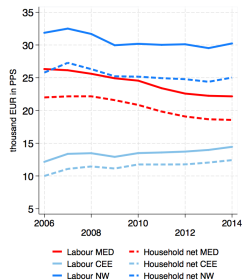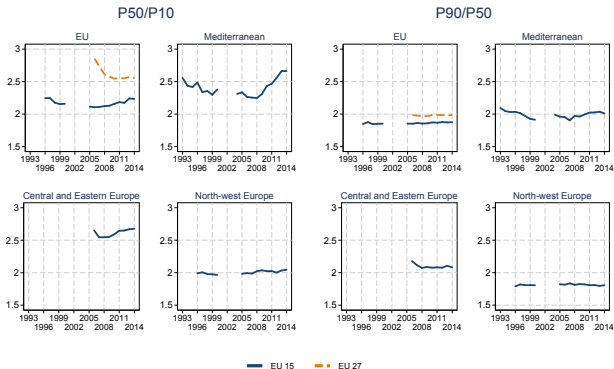

Convergence of large geographic areas: MED and CEE converges in mean

over the whole distribution, MED loses, but more at the bottom

CEE gains mostly in the middle of the distribution

Previous work: within-action - where?

Contribution to changes at bottom/top of income distribution and localised effects.

Inequality decreases in 1990, decreases in 2000s: largely due to lower part (in the Mediterranean).

Also: effect of increased Chinese trade pressure from 2002 on EU15 regions mostly through the lower part of income distribution.

Outline

Sketch of a framework

1. Test convergence: a tendency of a measure across time and entities to get closer to each other or to a reference. Here: a descriptive measure of the distribution/income inequality.

2. Check convergence of entities: using endogenous grouping based on changes in relative (to reference) distribution.

3. Characterise points of the process and group entities based on (levels) of relative (to reference) distribution. Look at dynamics of relative distribution.

4. Extend analysis to other income types: subsets, PPP uncorrected and conditional. Put together big picture.

Methods - convergence - the usual way

The usual way of testing for convergence is to look at β and σ convergence (mean-reversion and decreasing variance, respectively).

Tests of beta-convergence are in reality tests of mean-reversion. In case of income *levels*: β-convergence is a necessary, but not sufficient criterion – see Young, Andrew et.al JMCB(2008).

For inequality, need to look at distribution. An aggregator can collapse also these into a single measure, reducing the problem to the above. A typical recent analysis in this spirit is Dhongde-Silber JEI(2016), considering σ convergence of income shares at different deciles. Similar alternative: Biancotti JEI(2006), polarisation of Ginis.

Need an approach to handle the distribution as a whole, but with details.

Joint
Research
Centre

Methods - convergence - our preference - theory

We use the nonparametric approach of Székely-Rizzo (2004):
Let $\{\mathcal{A}_i\}_{i=1}^{k}$ denote independent random samples of random vectors (in $\mathcal{R}^d$) from the corresponding distributions $\{F_i\}_{i=1}^{k}$. Let the respective sample sizes be denoted by $\{n_i\}_{i=1}^{k}$ and $n := \sum_{i=1}^{k} n_i$. The k-sample test statistic is given by

$$\xi_n = \sum_{1 \le i < j \le k} e(\mathcal{A}_i, \mathcal{A}_j), \tag{1}$$

where for any pair $\mathcal{A} = \{a_1, a_2, \ldots, a_{n_1}\}$ and $\mathcal{B} = \{b_1, b_2, \ldots, b_{n_2}\}$,

$$e(\mathcal{A}, \mathcal{B}) = \frac{n_1 n_2}{n_1 + n_2} \left(\frac{2}{n_1 n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \|a_i - b_j\| - \frac{1}{n_1^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \|a_i - a_j\| - \frac{1}{n_2^2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \|b_i - b_j\| \right). \tag{2}$$

Eq. (2) satisfies the triangle inequality and therefore $\xi_n = 0$ if all elements in the samples coincide and $\xi_n > 0$ positive otherwise. Székely-Rizzo (2004) show that under the null

$$H_0: \ F_1 = F_2 = \cdots = F_k,$$

ξ_n has a well defined limiting distribution, whereas under the composite alternative

$$H_1: \ \exists i, j, \ F_i \ne F_j,$$

$\mathrm{E}[\xi_n]$ is asymptotically a positive constant times n, provided that $n_i / n \to c_i \in (0, 1)$.

Methods - convergence - our preference - estimation

The limiting distribution under the H_0 depends on $\{F_i\}_{i=1}^k$, thus critical values $\kappa_{1-\alpha}^{(b)}$ are produced with bootstrap draws from the pooled sample. For better visualisation, define $\tau_\alpha = \xi_n / \kappa_{1-\alpha}^{(b)}$ as a scaled indicator of 'distance from rejection', $\tau_\alpha < 1$ indicating that the deviation is insignificant.

Computing $\widehat{\xi}_n$ as the direct sample analog of 2 is computationally prohibitive for real-life sample sizes. Our alternative approach is the following:

1. estimate the distribution functions nonparametrically, invert to obtaining the quantile function,

2. draw random samples of smaller size (500-2000 observations) for each country and the rest of the population using the quantile-based procedure as in Hudson-Ernst (2000).

3. do bootstrap (100+ repetitions) to get both the critical value $\kappa_{1-\alpha}^{(b)}$

4. repeat (200+ times) to get an estimate of $\mathrm{E}[\widehat{\xi_n / \kappa_{1-\alpha}^{(b)}}]$

5. repeat for each year, infer convergence based on time-path

Methods - details of convergence

One can compare income distributions $F(x)$ and $G(x)$ based 4on:

1. *Monetary transfer* to shift a percentile of $F(x)$ to the same percentile of $G(x)$. Support of $F(x)$ must be a subset of $G(x)$ or need arbitrary scaling. Natural interpretation. $[0, 1] \rightarrow [x_{min}, x_{max}]$

2. *Differences in quantile* $F(x)^{-1}$ to reach the corresponding quantile $G(x)^{-1}$. Support of $F(x)$ must be a subset of $G(x)$ or need arbitrary scaling. Natural interpretation. $[0, 1] \rightarrow (-1, 1)$

3. *Relative quantiles* $F(x)^{-1}$ to reach the corresponding quantile $G(x)^{-1}$. No need for supports even to overlap. Less natural, but still fine interpretation. $[0, 1] \rightarrow (0, \infty)$

Our choice: 3) relative quantiles.
Note: compares both 'shape' and 'location' - levels matter!

Methods - endogenous grouping

We want to look at

- ▶ details of the (perhaps convergent) transition based on the *change* in relative quantiles and
- ▶ the states between which the transition took place based on the *level* of relative quantiles

We perform hierarchical clustering based on relative quantiles to form the groups.

[Need to think about sensible alternatives to clustering. Some machine learning approaches perhaps.]

Data

Microdata from European household surveys, the European Community Household Panel for 1995-2000 and the European Union Statistics on Income and Living Conditions for 2003-2014. Overall, data for 1995-2014 with gaps.

Three aspects: 1) household net income for EU15: 1995-2014. 2) same for EU27: 2006-2014. 3) gross wage rates for AT, UK, EL, IT, HU, PL: 2006-20015.

To study composition effects: conditional wages as residuals from year-, country- and sex-wise regressions with log wags on the LHS, education, age dummies on the RHS (weighted, exp. of residuals).

Calculations use the subsample of 16+ for household equivalised income and the working 25-64-year-olds for wages.

All income expressed in 2015 EUR, PPS corrected. Household income equivalised using the modified OECD scale, trimmed at the lower 1 per cent. Wage rates Winsorised at 1 per cent and 0.01 per cent.

Outline

Simple test for convergence

[Insert beta and sigma convergence results for net HH eq income.]

Convergence - relations

In four domains:

1. household net income: 1996-2014, EU15
2. household net income, individual gross labour income: 2007-2014, EU27
3. gross wage rate: 2007-2014, selected countries only
4. residual gross wage rate: 2007-2014, selected countries only

Within the above: a) areas to EU, b) countries to areas.
Reference: relevant EU-concept (EU27 or EU15). Three areas:

- North-west (AT, BE, DE, DK, FI, FR, IE, LU, NL, SE, UK)
- Mediterranean (CY, EL, ES, IT, MT, PT)
- Central and Eastern Europe (BG, CZ, EE, HU, LT, LV, PL, RO, SI, SK)

Testing convergence - household net income

EU15 1996-2014

EU27 2007-2014
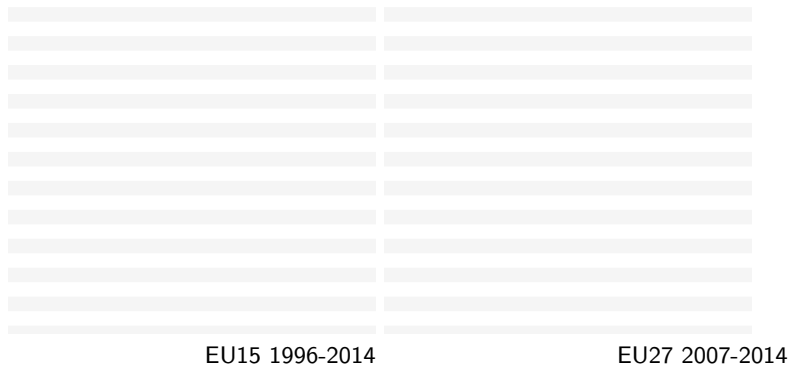
Being computed...

Convergence - wages

Wages available only for a set of countries: AT, UK, IT, EL, HU, PL
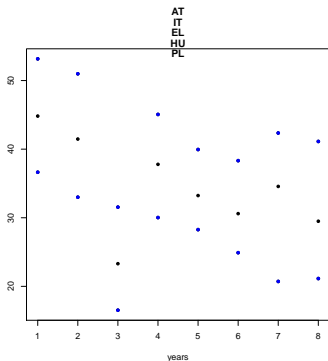
wage rates   residual wage rates

Unconditional wage rates are well above the 1 threshold, but show a secular decreasing trend between 2007 and 2014. Significant drop in 2010.

The basis for comparison – example: Hungary

Graph: income level in HU @ percentile / income level in REF @ percentile
Horizontal line: location shift only. Non-horizontal: shape difference.

Income relative to EU
(PPS adjusted equivalized household net income)

Income relative to CEE
(PPS adjusted equivalized household net income)
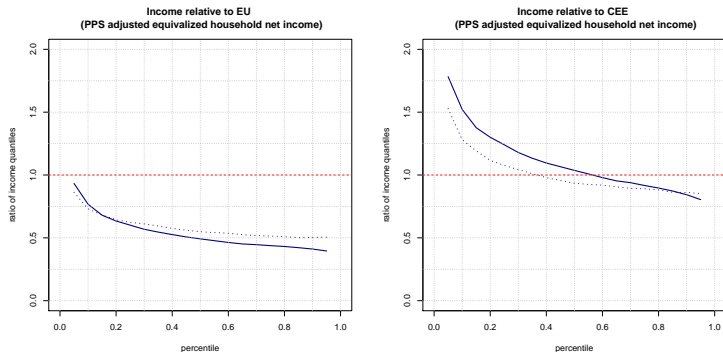
Legend: continuous line = 2007, dotted line = 2014

Bottom always relatively higher to reference than top. Change from 2007 to 2014: bottom loses, top gains.
To CEE: clear convergence. To EU: divergence up to 15 percentile, convergence after that.
(Results similar without PPP adjustment)

Zsombor Cseres-Gergely and Virmantas Kvedaras

Endogenously set groups: hierarchical clustering, changes
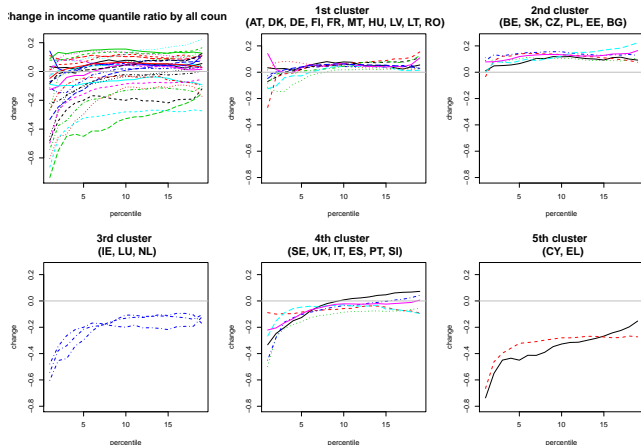
Clustering on the basis of quantile ratios: country vs. EU27.

1: moderate gain, some loss at lower end (NW-CEE mix), 2: serious gain (mostly CEE), 3: the losing NW, 4: serious loss at the bottom (NW-MED mix), 5: Greece

Distance from convergence – household income vs. EU
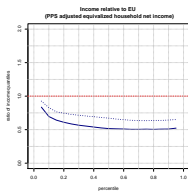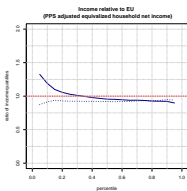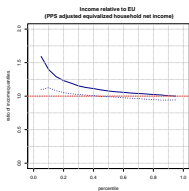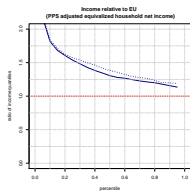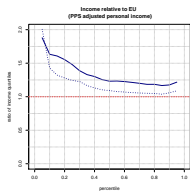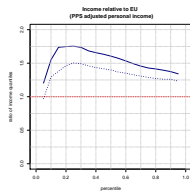

Austria   Italy   Spain   Poland
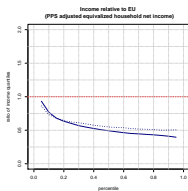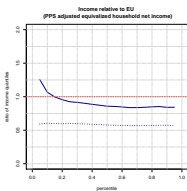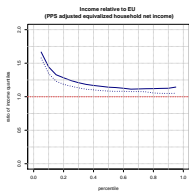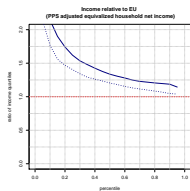
Netherlands   UK   Greece   Hungary
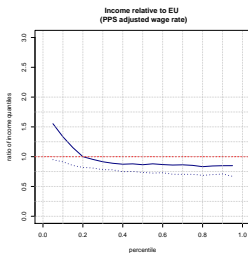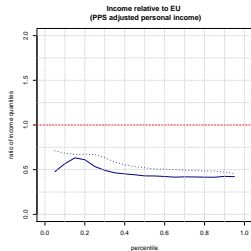
Solid line: 2007, dotted line: 2014.

Dynamics of selected relative quantiles – household income vs. EU
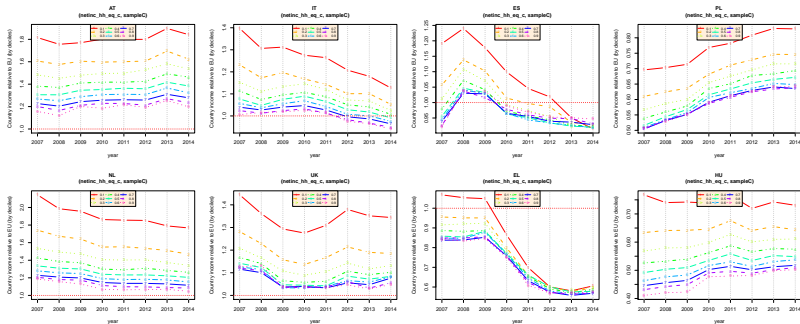
Outline

Groups based on percentile ratios: the result

Exogenously set groups: hierarchical clustering, levels

**Dendogram of income ratios – 2007
(netinc_hh_eq_c,sampleC)**
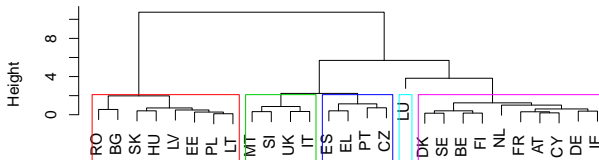
Class of eah leaf
hclust (*, "ward.D2")

**Dendogram of income ratios – 2014
(netinc_hh_eq_c,sampleC)**
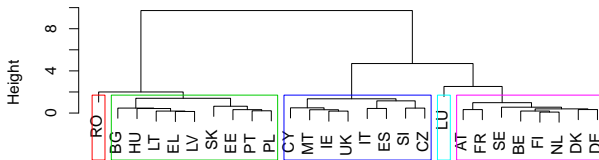
Quantile ratios of endogenously set groups in 2007, 2014
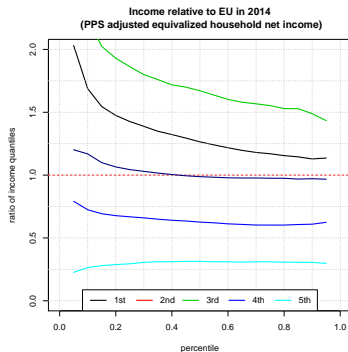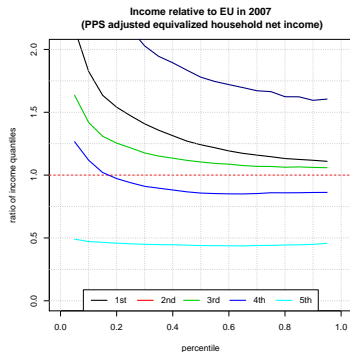
Quantile ratios of the groups formed on their basis: country vs. EU27.

Further details

The framework allows exploring further details:

▶ Convergence patterns to own area

▶ Convergence patterns without PPP adjustment

▶ Convergence patterns in terms of different income concepts:
  ▶ Labour income
  ▶ Wage rates

Distance from convergence – household income vs. own area
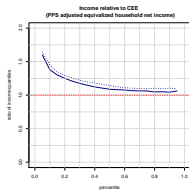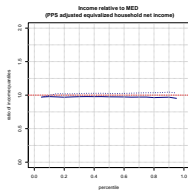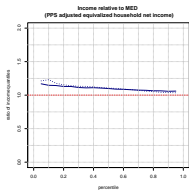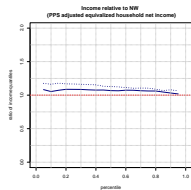
Austria   Italy   Spain   Poland
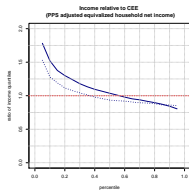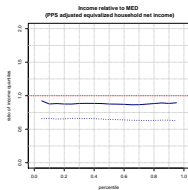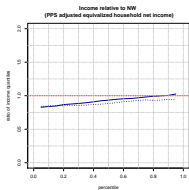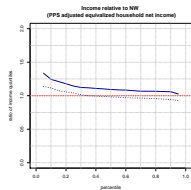
Netherlands   UK   Greece   Hungary

Solid line: 2007, dotted line: 2014.

Zsombor Cseres-Gergely and Virmantas Kvedaras

Distance from convergence – labour income vs. EU
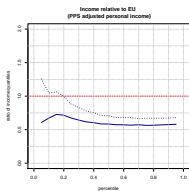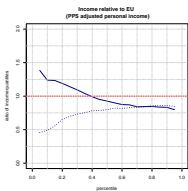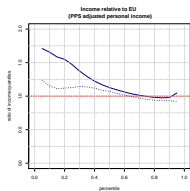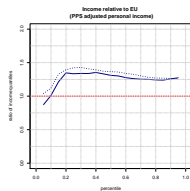
Austria · Italy · Spain · Poland
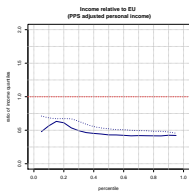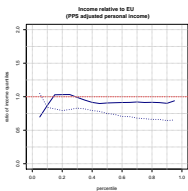
Netherlands · UK · Greece · Hungary

Solid line: 2007, dotted line: 2014.

Distance from convergence – wage rates vs. selected countries
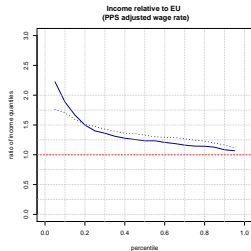
Austria
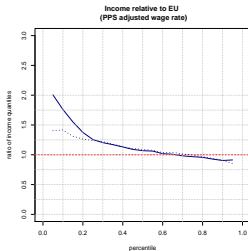
Italy
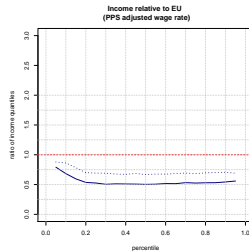
Poland

UK

Greece

Hungary

Solid line: 2007, dotted line: 2014.

Zsombor Cseres-Gergely and Virmantas Kvedaras
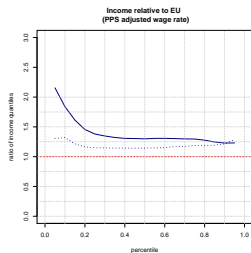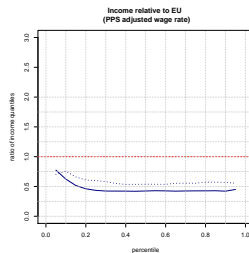
Putting all together - example of Hungary
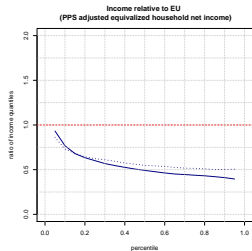
Base: household i. 16+, EU   Household i. 16+, CEE   Household i. 25-64, working, EU
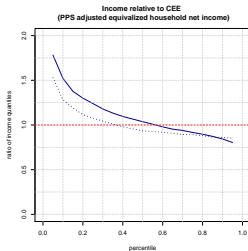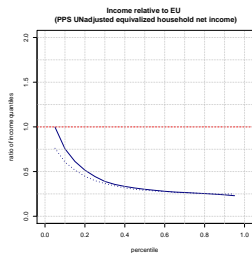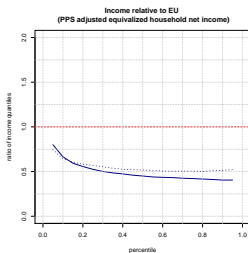
base -PPS, EU   Base, 25-64, working, EU   Wage rate, 25-64, working, EU
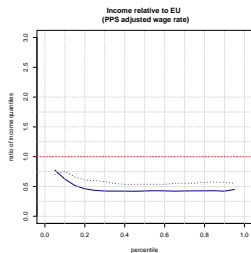
Solid line: 2007, dotted line: 2014.

Next steps

- Make the convergence test work in a real-life situation (N=2000, 20 iterations at least for EU27 2006-2014)
- Look at the EU15 for 1996-2014 (net household income)
- Isolate the effect of PPP correction - convergence with and without
- Wage rates: simulate effect of compositional changes for interesting scenarios (what if education levels would have risen in MS X)

Distance from convergence – household income vs. EU - sample C and D

Solid line: 2007, dotted line: 2014.

Distance from convergence – household income vs. EU - sample C and D
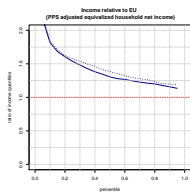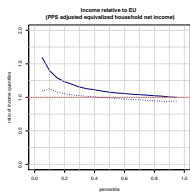
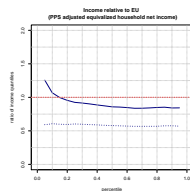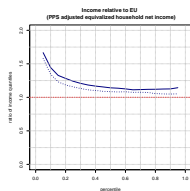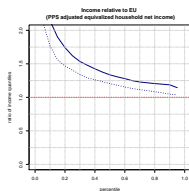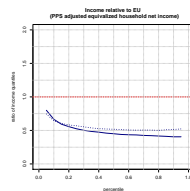Netherlands   UK   Greece   Hungary
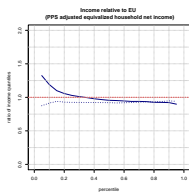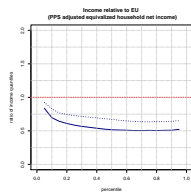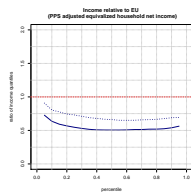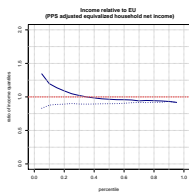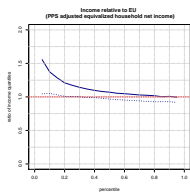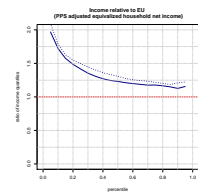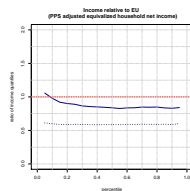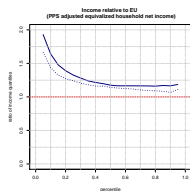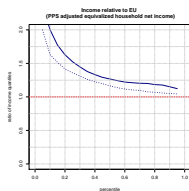
Every 16+ year-old
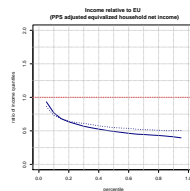
Every 25-64 year-old, labour income > 0

Solid line: 2007, dotted line: 2014.