

Personal Earnings Inequality and Polarization: The Czech Republic in a Comparison with Austria and Poland

Martina Mysíková

Institute of Sociology of the Czech Academy of Sciences, Prague

34th IARIW General Conference

Session 2E: The Great Recession and the Middle Class I

Discussed by Roberto Zelli - Sapienza University of Rome

Main goal of the paper

- Analysis of personal earnings distribution in the Czech Republic (CR) and two neighbouring countries, Austria and Poland, before the crisis (2003-2007) and after the crisis (2007-2010), using longitudinal data.
- Three perspectives:
 - 1 To verify if the crisis has spurred the hollowing of the middle class by analyzing the whole shape of the distribution.
 - 2 To examine the factors that influenced movements of individuals along the earnings distribution.
 - 3 To examine the structure of earnings inequality by a regression-based decomposition.

Data I

- 1 Longitudinal data from EU Statistics on Income and Living Conditions (EU-SILC)
- 2 Each longitudinal file covers four-year period: 2004-2007 and 2007-2010
- 3 Selected individuals aged initially 25 to 54 who worked at least six months in the first observed year and reported positive annual earnings both in the first and last year
- 4 Annual earnings (labour income both from employment and self-employment)

Information on monthly variation in earnings or working hours each month is not available. A reductions of working hours or hourly wages thus cannot be distinguished; both would be jointly reflected in earnings decrease

1. Changes in the distribution I

- Relative distribution method (Handcock and Morris, 1999). The relative distribution compares two distributions in periods 0 (reference) and 1 (comparison), where the values of period 1 are expressed as positions in the distribution of the reference period.
- The relative probability density function (PDF) is the density ratio at each r^{th} percentile y_r , where the percentile y_r is computed for the distribution at time 0 (reference). The relative PDF $g(r)$ is defined as the ratio:

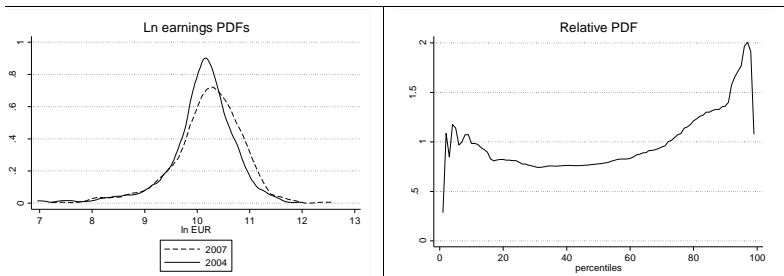
$$g(r) = \frac{f(F_0^{-1}(r))}{f_0(F_0^{-1}(r))} = \frac{f(y_r)}{f_0(y_r)} \quad 0 \leq r \leq 1, y_r \geq 0$$

that can be interpreted as the ratio of the fraction of individuals in the comparison population to the fraction of individuals in the reference population at a given level of the percentile $y_r = F_0^{-1}(r)$.

1. Changes in the distribution II

- When no changes occur between the two samples, $g(r)$ is uniform in $[0, 1]$. A value of $g(r)$ higher (lower) than 1 means that the share of individuals in the comparison population is higher (lower) than the corresponding share in the reference population, at the r^{th} percentile of the reference population.

Earnings distribution functions - Austria. Reference=2004; Comparison=2007



Location and shape I

- This method allows the researcher to detect how much of overall changes stems from changes due to a simple shift of all incomes (**location effect**) and changes due to the redistribution of income along the income scale (**shape effect**).
- First a location adjusted population, Y_{0L} , is constructed to have the same shape of the reference distribution, and the median of the comparison distribution, i.e. $Y_{0L} = Y_0 + \rho$, where $\rho = \text{median}(Y) - \text{median}(Y_0)$.
- then, it is possible to decompose the relative PDF:

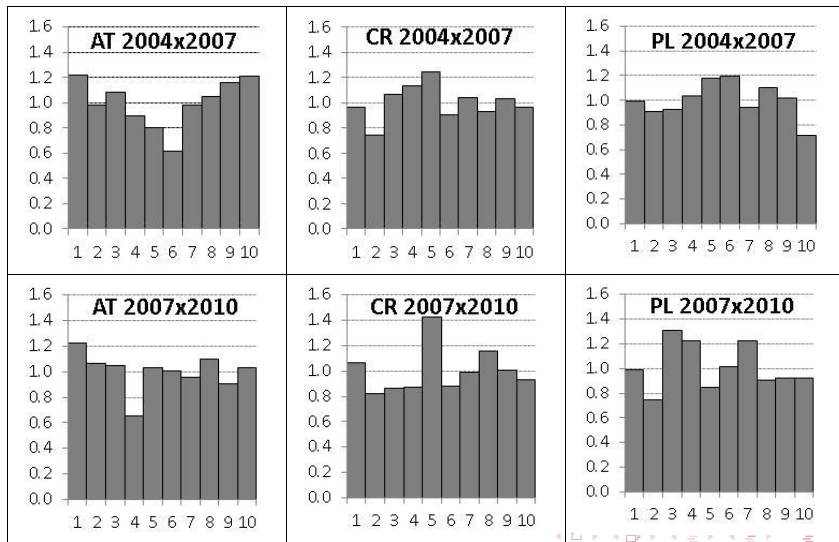
$$\underbrace{\frac{f(y_r)}{f_0(y_r)}}_{g(r)} = \underbrace{\frac{f_{0L}(y_r)}{f_0(y_r)}}_{g_L(r)} \times \underbrace{\frac{f(y_r)}{f_{0L}(y_r)}}_{g_S(p)},$$

where p is the percentile rank in the location-adjusted population which correspond to y_r .

Location and shape II

- $g_S(p)$ represents the relative distribution net of the location effect.
- $g_S(p)$ would take a (reverse) U-shape, if the comparison population is relatively (less) more spread around the median.

Shape shifts of earnings distributions by decile



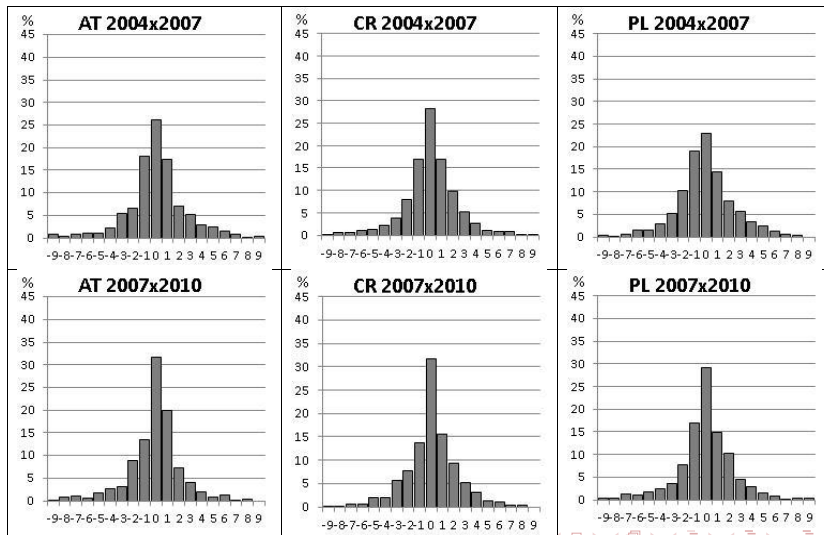
Main results I

- During the period 2004-2007 (before crises):
 - Clear shrink of the middle mass of the distribution towards both the tails of the distribution in Austria.
 - In the CR and Poland, there was instead a concentration of the mass towards the median-income range, though from different directions.
 - In the CR: from the bottom deciles
 - In Poland: from the top deciles
- During the period 2007-2010 (after crises):
 - Austria: movements from the middle to the bottom and no movements in the top half (upsurge of polarization due to movements below the median))
 - Czech Republic: movements towards the 5th decile from the bottom (convergence?).
 - Poland: no clear pattern of polarization

2. Individual movement analysis I

- Since data are longitudinal, it is possible to analyze movements of individuals between deciles.
- The deciles are defined as the cut points of the first-year median-adjusted distribution.
- Focus on factors that influenced movements of individuals along the earnings distribution, by a regression where the dependent variable is the movement between deciles (that ranges from -9 to +9).

Distribution of movements between deciles



OLS regression of movements to a higher decile

	AT		CR		PL	
	2004x2007	2007x2010	2004x2007	2007x2010	2004x2007	2007x2010
Original decile	-0.59***	-0.51***	-0.49***	-0.44***	-0.56***	-0.50***
Male	0.94***	0.58***	0.95***	0.55***	0.79***	0.39***
Work experience	0.07	0.05	-0.02	0.02	0.06**	-0.02
Work experience ²	0.00	0.00	0.00	0.00	0.00**	0.00
Medium education	1.07***	0.17	-0.04	0.61	0.42**	-0.14
High education	1.20**	0.92***	0.34	1.62***	1.06***	0.40
Months worked originally	0.23***	0.23***	0.34***	0.11*	0.16***	0.20***
Change of months worked	0.34***	0.29***	0.35***	0.32***	0.31***	0.32***
Changed job for better	-0.06	-0.49	0.53***	0.33	0.21	0.22
Changed job – forced to	0.39	-0.23	0.01	0.02	-0.32	-0.79**
Changed job – bc. of family	-1.91***	-0.64	0.09	-0.50	0.05	-0.24
Experienced unemployment	-0.08	-0.42	-0.36	-0.67**	-1.22***	-0.73*
ISCO – up	0.40	-0.10	0.41***	0.28	0.57***	0.10
ISCO – down	-0.43*	-0.29	-0.24*	-0.04	-0.40*	-0.55**
Densely pop. area	0.53**	-0.13	0.26**	0.17	0.25**	0.27**
Medium pop. area	0.60***	-0.15	0.08	-0.26	0.15	0.01
ISCO 1-8	controls	controls	controls	controls	controls	controls
Constant	-2.92***	-1.35**	-2.45***	-0.74	-1.40**	-0.58
R ²	0.36	0.38	0.31	0.31	0.37	0.32
N (unweighted)	702	768	2377	1100	1682	1677

Notes: * statistically significant at the 10% level, ** statistically significant at the 5% level, *** statistically significant at the 1% level

Main results I

- Before the crisis, men had better prospects than women in all the countries (see Table 1). Compared to women, men moved up the distribution (or dropped less) by almost one decile in the pre-crisis period (2004-2007). During the crisis (2007-2010), men's advantage lowered almost to a half and gender played a less important role.
- One of the reasons was gender labour market segregation with female over-representation in sectors (e.g., services) less hit by the crisis and their under-representation in male-dominated sectors (manufacturing, construction, financial sector) that were hit more
- Education had a significant positive impact on movements along the distribution before the crisis while its effect weakened substantially or lost significance during the crisis in Austria and Poland.
- However, the CR represents an opposite development: the impact of tertiary education gained significance and strongly influenced movements along distribution only during the crisis.

Main results II

- Change of job by reason of change (a reference group consists of individuals who have not changed a job between the first and fourth year): the CR is the only country where a change of job for a better one helped individuals to move up along the distribution, but only in the first period.

3. Earnings inequality decomposition I

- Quantifying the contribution to the inequality of a set of factors, in two different periods: 2007 and 2010.
- Regression-based inequality decomposition (Fields, 2003; Fiorio and Jenkins, 2007 implemented in Stata):
 - The factors are introduced as explanatory variables in an earning-generating model that is estimated through a linear regression model:

$$y_i = \beta_0 + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \cdots + \beta_k x_{k,i} + \epsilon_i$$

- The linear model can be rewritten as:

$$y_i = \beta_0 + z_{1,i} + z_{2,i} + \cdots + z_{k,i} + \epsilon_i$$

where $z_{k,i} = \beta_k x_{k,i}$ is the 'composite' variable.

3. Earnings inequality decomposition II

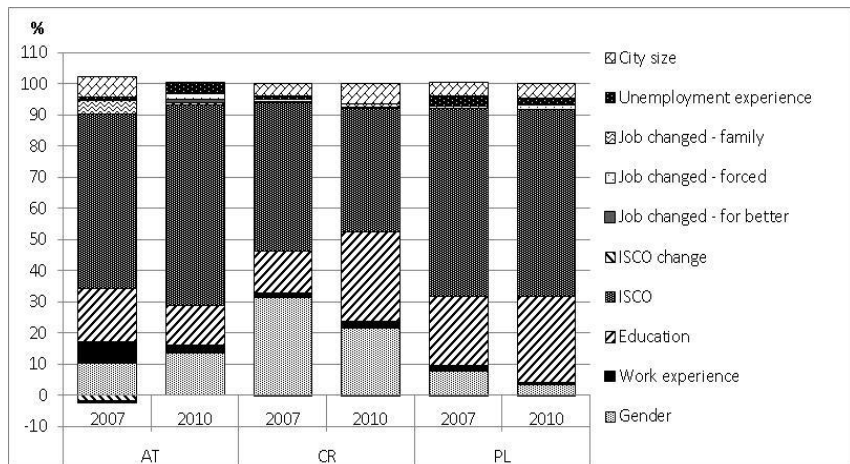
- OLS estimates of the linear regression model can be used for decomposition:

$$y_i = \hat{\beta}_0 + \hat{z}_{1,i} + \hat{z}_{2,i} + \cdots + \hat{z}_{k,i} + e_i$$

where $\hat{z}_{k,i} = \hat{\beta}_k x_{k,i}$ is the estimated 'composite' variable.

- Shorrocks's decomposition by factor components can be applied once one treats each z as an income source (variance as a measure of inequality).
- Earnings function estimated by including similar regressors as in the previous analysis.
- Decomposition applied to predicted earnings (\hat{y}_i), i.e. net of residual terms.

Regression-based earnings inequality decomposition



Main results I

- Dominant role is type of occupation (ISCO classification)
- Other important factors: gender and human capital (education and experience)
- Gender gap causes the highest share of overall earnings inequality in the CR, but decreasing (more than 30% in 2007 and 20% in 2010)
- Role of education increased in the CR and in Poland, while decreased in Austria.
- Role of work experience is relevant only in Austria (but decreasing) and marginal in the CR and Poland.
- The other factors contribute markedly less to the overall earnings inequality.

Remarks I

- What population does the selected sample represent? People who were initially of working age (25-54) and in employment in the initial year AND in the final year. is it sufficient to conclude that the crises did not spur earnings polarization?
- Better to use longitudinal weights?
- Estimation of relative distribution does not need longitudinal-data. Comparison with relative distribution using cross-sectional data.
- RD allows one to measure the relative polarization (changes in polarization) and the contribution to the overall polarization made by observations above and below the median of the relative distribution.
- Interesting to verify if indices of (relative) polarization are in accordance with indices of (relative) inequality.
- Location shifts: due to price effects or real (median) growth?
- Peak in the 5th decile: check robustness to data rounding and heaping, that are potentially present in survey data on earnings.

Remarks II

- Decomposing *predicted* earnings instead of earnings makes a difference!
- Are the explanatory variables included as determinants broad enough to account for the main factors that are likely to explain earnings inequality?