Tax-benefit systems, employment structures and the distribution of household disposable income: the UK versus Ireland

(preliminary results)

Denisa Maria Sologon\textsuperscript{1,2}, Philippe Van Kerm\textsuperscript{1} & Cathal O’Donoghue\textsuperscript{2,3}

\textsuperscript{1} LISER (formerly CEPS/INSTEAD), \textsuperscript{2} Maastricht University, \textsuperscript{3} TEAGASC

4th European User Conference for EU-Microdata, Mannheim, March 5-6

The financial support of Fonds National de la Recherche is gratefully acknowledged.
Introduction

Context

- large international variation in income inequality across countries - relatively little is known about the sources of the differences in household disposable income inequality.
- "SIMDECO" project aimed
  - to examine the nature of the cross-national differences in income inequality, focusing on the role of differences in tax-benefit systems, employment behaviour and labour market structures across countries
  - to assess how much these factors can help understand international differences in how household incomes have been hit during the crisis.
Project objectives & contribution

▶ our approach: a unified framework
▶ (i) comparable data
Project objectives & contribution

- Overall, the result should be an integrated framework across countries for generating and simulating the distribution of household disposable income.
- Today: a small part of the project
- the methodology we use to explore the nature of the cross-national differences in household disposable income inequality between countries.
- the validation and decomposition results for the UK versus Ireland (pre-crisis)
- (next step) pre-post crisis with labour supply responses
Data

- Outcome income measure: equivalized household disposable income
Methodology

- micro-simulation micro-econometric approach
- the contribution of tax-benefit systems and labour market structures to the differences in inequality is assessed using (a sequence of) simulated counterfactual distributions of household disposable incomes that would prevail in each country, if these factors were swapped between countries.
Methodology

- simulated counterfactual distributions of household disposable incomes
- the logic of the Generalized Oaxaca-Blinder decompositions - extended to the entire distribution

\[ f(D_1 LM_1 Y_1 TB_1) - f(D_2 LM_2 Y_2 TB_2) = \]
\[ f(D_1 LM_1 Y_1 TB_1) - f(D_1 LM_2 Y_1 TB_1) (\text{Labour market effect}) \]
\[ + f(D_1 LM_2 Y_1 TB_1) - f(D_1 LM_2 Y_2 TB_1) (\text{Income effect}) \]
\[ + f(D_1 LM_2 Y_2 TB_1) - f(D_1 LM_2 Y_2 TB_2) (\text{Tax - benefit effect}) \]
\[ + f(D_1 LM_2 Y_2 TB_2) - f(D_2 LM_2 Y_2 TB_2) (\text{Residual effect}) \]

- \(2^4\) counterfactual distributions (including original)
- 12 possible decomposition paths
Methodology

For generating the counterfactual distributions of household disposable income we need:

- a tax-benefit simulator (EUROMOD) and
- a household income distribution model for each country.
Methodology

Building a household income distribution model:

- Labour Market Module
- Income Module
- Pre-Euromod Data Standardization Module
- Tax-Benefit simulator (Euromod)
**Methodology**

**Labour Market Module:**
- In work/Out of work
- Part time
- Employee/Self-employed
- Occupation/Industry/Sector
- Retired
- Unemployed
- Has income sources (employment income, self-employment income, investment income, property income, private pension, other income)
- Pays for housing (Home owner, mortgage, rent)
- Pays contributions (private pensions)

Estimate the LM structure for country X => Store the parameter estimates for country X

Simulate LM on data of country 1 (apply the parameter estimates of country 2 on data 1, while importing the residuals of country 1)

\[ f(D_1LM_2Y_1TB_1) \]

**Income Module**
- Wage rate (correction for endogenous participation for women)
- Self-employment income, investment income, property income, private pension, other income
- Housing payments (mortgage, rent)
- Contributions (private pensions)

Estimate incomes for country X => Store the parameter estimates for country X (x = 1, 2)

Simulate Incomes of country 2 on data of country 1 (apply the parameter estimates of country 2 on data 1, while importing the residuals of country 1)

\[ f(D_1LM_1Y_2TB_1) \]
Methodology

Estimation techniques:

- labour market structure & presence of income sources
  - logit/mlogit models
- income sources
  - wage rate - Singh-Maddala distribution regression (without and with endogenous selection) (Van Kerm, 2013)
    - Cumulative distribution function:
      \[ F(y) = SM(y; a, b, q) = 1 - \left[ 1 + \left( \frac{y}{b} \right)^a \right]^{-q}. \]
      - where:
        - \( q \) is a shape parameter ('upper tail')
        - \( a \) is a shape parameter ('spread')
        - \( b \) is a scale parameter
  - other - Ordinary Least Squares + residual distribution (Juhn et al. 1993)
Methodology

Pre-Euromod Data Standardization Module

- Run Euromod $\Rightarrow f(D_1LM_1Y_1TB_2)$
Validation UK: original vs. fully simulated LM, Income, TB
Results

Validation IE: original vs. fully simulated LM, Income, TB

Pen parade - 2008 IE Data

Quantiles of equivalized household disposable income - 2008 data
Quantiles of equivalized household disposable income - simulated
Equivalized household disposable income - UK vs. Ireland
Decomposition results - Equivalized household disposable income (UK vs. UK with IE LM)
Decomposition results - Equivalized household disposable income (UK vs. UK with IE LM & Income)

Pen parade
Decomposition results - Equivalized household disposable income (UK vs. UK with IE LM & Income & TB)
Distributional differences in the equivalized household disposable income due to LM, Income and TB effects (UK-IE)
Another approach: compare the effects of each factor at the 1st stage.
Relative Contribution of LM structure, Income structure, Tax-Benefit system to the difference in Gini (IE-UK), ceteris paribus
Behind the Shapley Index - variation in the Gini index
Main findings

- Ireland more unequal and slightly richer on average, but the poorest 40% are better off in the UK than in Ireland.

- Swapping first the Irish LM structure in the UK - everyone is worse off (the “top” to a larger extent than the “bottom”).

- Swapping first the Irish income structure in the UK - the bottom 70% of the distribution is worse off (at a decreasing rate); top 30% better off at an increasing rate \(\Rightarrow\) the top is favoured.

- Swapping the Irish TB system - the bottom 35% is worse off at a decreasing rate; the rest are better off (between the 50th - 80th quintile the best off).

- The largest contribution to difference in Gini: \(\text{LM} > \text{TB} > \text{Income}\).
Conclusion

Next

- extend the analysis pre-post crisis
- account for labour supply behavioural responses
- more countries
Thank you
Pre-Euromod Data Standardization Module

- Identify the components of household disposable income in each country
  - Simulated in Euromod => INSTRUMENTS
    - Check the explanatory factors needed to simulate these instruments in each country
      - (A) overlapping between countries
      - (B) non-overlapping between countries

  (A) overlapping between countries (INPUT dataset)
  - Estimate the presence of the income source and the level (if necessary)
  - Swap the parameter estimates between countries => simulate the countrefactual distributions

  (B) Non-overlapping between countries
  - Estimate presence/level in country 1 => then simulate in country 2 (and draw residuals from distribution 1)

Output: Fully standardized datasets containing all variables required to run the TB system of country 1 on Data of country 2 and vice versa.

Run Euromod => $f(D_1 LM_1 Y_1 TB_2)$
Validation UK: original vs. fully simulated LM, Income, TB

**Table: Validation Statistics Equivalized Disposable Income - the UK**

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Mean</th>
<th>SD</th>
<th>GE0</th>
<th>GE1</th>
<th>GINI</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2404.968</td>
<td>4778.635</td>
<td>0.135</td>
<td>0.197</td>
<td>0.281</td>
</tr>
<tr>
<td>1111</td>
<td>2407.481</td>
<td>4778.190</td>
<td>0.134</td>
<td>0.197</td>
<td>0.280</td>
</tr>
</tbody>
</table>

Notes: 1 = original distribution (Euromod output); 1111 = fully simulated
Validation Ireland: original vs. fully simulated LM, Income, TB

Table: Validation Statistics Equivalized Disposable Income - Ireland

<table>
<thead>
<tr>
<th>Distribution</th>
<th>Mean</th>
<th>SD</th>
<th>GE0</th>
<th>GE1</th>
<th>GINI</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2427.968</td>
<td>1555.006</td>
<td>0.141</td>
<td>0.152</td>
<td>0.294</td>
</tr>
<tr>
<td>2222</td>
<td>2427.363</td>
<td>1555.051</td>
<td>0.140</td>
<td>0.152</td>
<td>0.294</td>
</tr>
</tbody>
</table>

Notes: 2 = original distribution (Euromod output); 2222 = fully simulated
Decomposition results - Equivalized household disposable income (IE vs. IE with UK LM)
Decomposition results - Equivalized household disposable income (IE vs. IE with UK LM & Income)
Decomposition results - Equivalized household disposable income (IE vs. IE with UK LM & Income & TB)
Distributional differences in the equivalized household disposable income due to LM, Income and TB effects (IE-UK)
Various paths UK (all possible swaps)
Various paths IE (all possible swaps)