# Online appendix

## Section A: IEWB Data

The IEWB database provides annual data on old-age poverty rate and gap computed on after-tax equivalent income data.

Information on these two variables for European countries is obtained from the Luxembourg Income Study for the 1980-1994 period and from Eurostat for the 1995-2014 period. This latter period is coverable by combining EU-SILC data with European Community Household Panel survey (ECHP) information. Data on for Canada are obtained from Statistics Canada’s Canadian Income Survey available through CANSIM. Data for Australia for 1980-2000 are from the Luxembourg Income Study; for 2001-2013, the estimates were calculated using microdata based on the Household, Income and Labour Dynamics in Australia Survey. Data for the United States are from the Luxembourg Income Study for 1980-1994; for 1995-2014, the estimates were calculated using microdata from the Annual Social and Economic Supplement to the Current Population Survey. For each country, missing values are interpolated using the average annual growth rate between observations. Furthermore, the methodology used by Eurostat on EU-SILC data was employed to develop EU comparable estimates for the non-EU countries included in the IEWB dataset, i.e., the United States, Australia, and Canada. For more detail, see:

Andrews, B. and Thomas, J. (2015), Estimation of EU-Comparable Poverty-Related Variables in Australia, 2001-2013, *CSLS Research Report 2015-13, September*.

<http://www.csls.ca/reports/csls2015-13.pdf>.

Andrews, B., Thomas, J. and Palesch, N. (2015), Estimation of EU-comparable Poverty-Related Variables in the United States, 1995-2014, *CSLS Research Report 2015-12, September*.

<http://www.csls.ca/reports/csls2015-12.pdf>.

Thomas, J., and Uguccioni, J. (2016), Equality and Economic Security Take a Hit: The Index of Economic Well-being for Selected OECD Countries, 1980-2014, *CSLS Research Report 2016-06, June*. <http://www.csls.ca/reports/csls2016-06.pdf>.

**Income information**

Main sources for IEWB poverty variables are LIS and EUROSTAT (and, also, national sources for USA, Canada and Australia, however these sources were harmonized to be consistent with the rest of data).

In the case of LIS data, IEWB uses DPI variable which captures cash disposable household income. In LIS documentation, they define this variable in the following way: “*Total monetary current income net of income taxes and social security contributions; this variable is defined (as well as named) in exactly the same way as the old LIS DPI variable*”.

The components of this variable are:

* *Total monetary payments from labour, property, and social or private transfers.*
* *Total value of non-monetary goods and services received from labour and social or private transfers, excluding social transfers in kind such as universal health insurance, universal education benefits, and near cash benefits from public housing.*

Moreover, in the income from rent voice, they add this note:

* *Imputed rent is not considered as current income and hence not included in the income variables; if available in the original dataset, imputed rent is reported in the LIS variables among non-monetary consumption items.*

Therefore, while monetary social benefits, including housing, are part of the income variable in LIS, imputed rents are not.

In the case of EUROSTAT data, the income information is derived from EUSILC. In the EUROSTAT website, they define income as follows: “*The equivalised disposable income is the total income of a household, after tax and other deductions, that is available for spending or saving, divided by the number of household members converted into equalised adults”.* EUSILC documentation, on the other side, reports the following information regarding their income variable: “*All monetary incomes received from any source by each member of a household are added up; these include income from work, investment and social benefits, plus any other household income; taxes and social contributions that have been paid, are deducted from this sum;”*

The components of household income variable (HY010) are the following ones:

 - *gross employee cash or near cash income (PY010G),*

*- company car (PY021G),*

*- gross cash benefits or losses from self-employment (including royalties) (PY050G),*

*- pensions received from individual private plans (other than those covered under ESSPROS) (PY080G),*

*- unemployment benefits (PY090G),*

*- old-age benefits (PY100G),*

*- survivor' benefits (PY110G),*

*- sickness benefits (PY120G),*

*- disability benefits (PY130G),*

*- education-related allowances (PY140G);*

plus gross income components at household level:

*- income from rental of a property or land (HY040G),*

*- family/children related allowances (HY050G),*

*- social exclusion not elsewhere classified (HY060G),*

*- housing allowances (HY070G)*

As in the case of LIS data, imputed rents are available in the following variable (HY030) but they are not part of the household income calculation. Social housing benefits as housing allowances (HY070G) are instead included in the calculation of household income.

Overall, we think that social housing benefits should not be our main source of concert since they are already monetized in the income variable. However, we cannot account for different rates of retirees’ homeownership across countries. Further, it is also possible that there are other sources of capital income that are not captured by survey data. For this reason, we tried to tackle potential unobserved heterogeneities across countries, employing an instrumental variable approach.

Consulted documentation:

LIS

<https://www.lisdatacenter.org/documentation/>

our-lis-documentation-variables-definition.xlsx

<https://www.lisdatacenter.org/wp-content/uploads/files/data-lis-variables.pdf>

EUROSTAT

<https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Glossary:Equivalised_disposable_income>

EUSILC data documentation:

EU-SILC Description Target Variables Household Data (H-file), pag. 1-27

## Section B: List of controls

Table B.1: Description of control variables

|  |  |  |  |
| --- | --- | --- | --- |
| **Dimension** | **Variable** | **Source** | **Description** |
| Pension system | Survivor ‘pensions expenditure as share of total pension expenditure  | OECD  | Reflects the relative importance of survivors’ pension benefits in total pension provision |
| Pension system | Public pension expenditure (old-age and survivors) as a % of GDP | OECD  | Measures the overall public financial effort of a country regarding pensions. |
| Pension system | Retirement age (average women/men) | CWED2 | The higher the pension system retirement age/the longer people have to work before retiring, the higher their income in later years and therefore, the lower the poverty rate of 65+ should be. |
| Labour market and income | Share of over 65 employed  | OECD | The higher the labour market participation rate, the higher the income in the population 65+, and the lower the poverty. |
| Labour market and income | Gini index (after taxation) | IEWB | Level of income inequality across countries. Even if the pension system features redistributive elements, overall high-income inequality may outweigh those effects for pension income. |
| Labour market and income | GDP per capita (log) | PENN table | The higher the overall wealth, the lower the poverty. |
| Labour market and income | Annual growth of GDP per capita  | PENN table | Economic cycles affecting the poverty rate via fluctuations in working age incomes.  |
| Labour market and income | Average production worker wage (log)  | CWED2 | Average level of wages across countries. |
| Demographic | The share of women in the older population  | CWS | Women tend to have shorter employment careers and are thus at risk of getting lower pensions. They could be considered the population with higher risk of poverty in old age.  |
| Demographic | Dependency ratio in the period 1980-2010 (i.e. the ratio between persons aged over 64 years and the number of employees) | CWS | A high dependency ratio puts pressure on the pension system to lower the generosity to keep the system financially sustainable.  |
| Demographic | Male life expectancy at age 65 | CWS | The longer the life expectancy, the higher the potential share of retirees with lower pensions due to relatively low incomes counting for their pension accrual.  |
| Demographic | Female life expectancy at age 65 | CWS | The longer the life expectancy, the higher the potential share of retirees with lower pensions due to relatively low incomes counting for their pension accrual. Women tend to live longer than men and have lower pensions. |
| Health system  | Public expenditure on health, as a percentage of GDP | CWS | It captures health benefits which might be particularly important for older people. |

**Sources:**

**Comparative Welfare Entitlements Dataset (CWED):**

Scruggs, L., Jahn, D. and Kuitto, K. (2018), *Comparative Welfare Entitlements Dataset 2, version 2018*, University of Connecticut & University of Greifswald, Available for download at <http://cwed2.org/>

**Comparative Welfare State (CWS):**

Brady, D., Huber, E., and Stephens, J.D. (2020). *Comparative Welfare States Data Set*. University of North Carolina and WZB Berlin Social Science Center. Available for download at <https://www.lisdatacenter.org/news-and-events/comparative-welfare-states-dataset-2020/>

**Index of Economic Well-being (IEWB):**

Osberg, L. and Sharpe, A. (2016), *Database of the Index of Economic Well-being for selected OECD countries and Alberta, 1980–2014*. Available for download at

<http://www.csls.ca/iwb/oecd.asp.>

**Penn World Table:**

Feenstra, R.C., Inklaar, R. and Timmer, M.P. (2015). *The Next Generation of the Penn World Table*. American Economic Review, 105(10), 3150-3182, available for download at [www.ggdc.net/pwt](http://www.ggdc.net/pwt)

**OECD**

Labour Force Statistics in OECD countries: <https://data.oecd.org/emp/labour-force.htm>

Social Protection Statistics in OECD countries: <https://stats.oecd.org/Index.aspx?DataSetCode=SOCR>

## Section C: Gaps

There are 6 countries in our panel data that have missing information on the pension benefit measures. We have listed below the length of the time-series for these countries.

Table C.1: List of countries with gaps in CWED

|  |  |  |  |
| --- | --- | --- | --- |
|   | SP-mid0 | SP-low0 | MP-mid0 |
| Belgium  | 1984-2010 | 1984-2010 | 1992-2010 |
| Canada  | 1983-2010 | 1983-2010 | 1983-2010 |
| Finland  | 1981-2010 | 1981-2010 | 1981-2010 |
| Germany | 1982-2010 | 1982-2010 | 1982-2010 |
| Italy  | 1986-2010 | 1986-2010 | 1986-2010 |
| Norway  | 1983-2010 | 1983-2010 |  No gaps  |

## Section D: Descriptive figures

Figure C.1: Percentage of poverty rate and gap, 1980-2010 period



Source: our own calculations based on IEWB 2016.

Note: the red line identifies the year 2010, the last period in which we observe both poverty outcomes and the pension replacement rates.

Figure C.2: Replacement rates, 1980-2010 period



Source: our own calculations based on CWED2 Scruggs et al. 2018

Figure C.3: Associations between poverty and pension generosity, period 1980-2010 – full sample





Source: our own calculations based on IEWB 2016 and CWED2 Scruggs et al. 2018.

## Section E: IV tests

Table D.1 reports the results for the Breusch-Pagan test which demonstrate the presence of heteroskedasticity in our data This is then exploited for constructing the IV Lewbel internal instrument that complements the dependency ratio. A rejection of the null demonstrates the present of no constant variance in the distribution of the error term.

|  |
| --- |
| Table C.1: Breusch-Pagan / Cook-Weisberg test for heteroskedasticity |
|   | **Poverty rate** |
|  | sp\_mid0 | sp\_low0 | mp\_mid0 |  low\_mid |
| chi2(1) | 23.33 | 26.46 | 17.25 | 17.2 |
| *Prob > chi2* | 0.000 | 0.000 | 0.000 | 0.000 |
|   | **Poverty gap** |
| chi2(1) | 38.89 | 39.84 | 33.39 | 37.88 |
| *Prob > chi2* | 0.000 | 0.000 | 0.000 | 0.000 |
| Note: Ho: Constant variance |

As recommended by Baum and Lewbel (2019) in table C.2 we report the Sargan–Hansen or Sargan's J test, i.e. “overidentification test”, associated with the IV Lewbel estimator which allows us to check whether the excluded instruments are appropriately independent of the error process. The joint null hypothesis in this case is that the instruments are valid instruments while the excluded instruments are correctly omitted from the model. As expected, we cannot reject the null confirming the validity of our instruments.

|  |
| --- |
| Table C.2: Overidentification test  |
| **Country FE** | **Country and year FE** | **Country, year FE and trend** |
| **Poverty rate** | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  |
| **Coef.** | -0.20\*\*\* | -0.14\*\* | -0.17\*\* | -0.10 | -0.28\*\*\* | -0.15\*\*\* | -0.15\*\* | 0.02 | -0.28\*\*\* | -0.15\*\*\* | -0.15\*\* | 0.02 |
| **se** | (0.08) | (0.07) | (0.07) | (0.09) | (0.06) | (0.06) | (0.06) | (0.08) | (0.06) | (0.06) | (0.06) | (0.08) |
| **N** | 406 | 406 | 399 | 406 | 406 | 406 | 399 | 406 | 406 | 406 | 399 | 406 |
| **J-statistic** | 14.7 | 17.1 | 9.05 | 14.8 | 44.7 | 45 | 54.4 | 49.7 | 44.7 | 45 | 54.4 | 49.7 |
| **Jdf**  | 11 | 11 | 11 | 11 | 44 | 44 | 44 | 44 | 44 | 44 | 44 | 44 |
| ***p-value***  | 0.196 | 0.105 | 0.617 | 0.194 | 0.443 | 0.429 | 0.136 | 0.258 | 0.443 | 0.429 | 0.136 | 0.258 |
|  |  |  |  |   |   |  |  |   |  |  |  |  |
| **Poverty gap** | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  | **SP\_MID** | **SP\_LOW** | **MP\_MID** | **SP\_LOW–SP\_MID**  |
| **Coef.** | -0.09 | 0.00 | 0.17 | 0.29\* | -0.17 | 0.02 | 0.12 | 0.27\*\* | -0.17 | 0.02 | 0.12 | 0.27\*\* |
| **se** | (0.15) | (0.13) | (0.13) | (0.16) | (0.11) | (0.11) | (0.11) | (0.14) | (0.11) | (0.11) | (0.11) | (0.14) |
| **N** | 406 | 406 | 399 | 406 | 406 | 406 | 399 | 406 | 406 | 406 | 399 | 406 |
| **J-statistic** | 16.9 | 16.8 | 9.16 | 10.7 | 44.4 | 60.2 | 61.3 | 58.3 | 44.4 | 60.2 | 61.3 | 58.3 |
| **Jdf**  | 11 | 11 | 11 | 11 | 44 | 44 | 44 | 44 | 44 | 44 | 44 | 44 |
| ***p-value***  | 0.111 | 0.114 | 0.607 | 0.467 | 0.456 | 0.053 | 0.043 | 0.073 | 0.456 | 0.053 | 0.043 | 0.073 |

Note: The table reports the estimated coefficients, se, N J-statistic, degree of freedom and the p-value associated with the J-statistics. With respect to the country-year FE, the results do not change for the model specification which add the time trend. For OLS models, we use panel-corrected standard errors (PCSEs, Beck and Catz 1995); for IVs models, we used Heteroskedasticity and Autocorrelation Consistent Standard Errors (HAC se); \* p < 0.10, \*\* p < 0.05, \*\*\* p < 0.01.

**References**

Baum, C. F., and Lewbel, A. (2019), Advice on Using Heteroskedasticity-Based Identification. *The Stata Journal,* 19, 4, 757–67.

Baum, C. F., Schaffer, M. E. (2012), ivreg2h: Stata module to perform instrumental variables estimation using heteroskedasticity-based instruments. Statistical Software Components S457555, Department of Economics, Boston College. <https://ideas.repec.org/c/boc/bocode/s457555.html>.

Beck, N., and Katz, J. (1995), What to do (and not to do) with Time-Series Cross-Section Data, *The American Political Science Review*, 89,3, 634-647.

**Section F:** **Relationship between poverty rate and poverty gap**

In order to clarify how the poverty gap might increase as consequence of a change in the pension benefits define poverty rate as:

|  |  |
| --- | --- |
| $$\frac{\sum\_{j=1}^{g}P}{N}$$ | [A] |

And define poverty gap as:

|  |  |
| --- | --- |
| $$\frac{1}{N}\sum\_{j=1}^{g}\frac{Z-y\_{j}}{Z} $$ | [B] |

Where $P$ is the number of poor people, $Z$ is the poverty line (threshold) and $y$ is the income for country $g$ and household $j$. N is the total number of households/units. Both numerators (in A and B) are directly linked via: $P\_{j}= (Z-y\_{j})$.

This implies that the numerator of the poverty gap would be very sensible to changes in the poverty rate. Suppose the scenario where we move households that are considered as poor (but very close to the threshold, so the amount of income that they need to avoid poverty is considerable low) to a position above that threshold. For example, a policy intervention that reduces the poverty risks by increasing monetary transfers to poor households (i.e., an increase in pension benefits). If we move these households above the poverty line, it means that they would have a total gap $(Z-y\_{j})$ of $0$, leaving only households with a bigger total gap in the numerator ($Z-y\_{j}$) where this expression would be ($y\_{j}<Z)$. If the denominator remains fixed in the equation, this will automatically increase the value of the poverty gap in the considered country. Given this, in a counterfactual scenario in which households that are placed close to threshold are systematically moved above the poverty line, the head count of poor people would decrease (i.e., a reduction in the poverty rate) but the poverty gap might increase because only the poorest remained in the numerator. That is, the average distance from the threshold would increase simply because only values at the bottom of the income rank would remain.