2024 UPDATE FOR FEMALE LABOR INCOME SHARE

VALENTINA GABRIELLI
THERESA NEEF
ANNE-SOPHIE ROBILLIARD

TECHNICAL NOTE N°2024/13

DECEMBER 2024

WORLD INEQUALITY LAB

Female Labor Income Share

Methodological Note

Valentina Gabrielli[1], Theresa Neef[2], and Anne-Sophie Robilliard[3]

December 2024

Overview

The female labor income share update is based on the methodology by Neef and Robilliard (2022). This update provides estimates of the female labor income share for 1990 –2023. This methodological note explains the data sources and methodology in detail, highlighting data availability and new data that was incorporated.

We provide estimates for 216 countries and jurisdictions. 138 jurisdictions have at least one data point throughout the period 1990–2023 for which we can estimate the female labor income share from original data, i.e. information on female and male wage and self-employment incomes. We predict the female labor income share for 36 additional jurisdictions using an OLS regression with female shares of wage and self-employment as the primary predictors. Employment data for this prediction comes from ILO modelled estimates which cover 176 jurisdictions. Finally, for 42 jurisdictions, we lack information on both income and employment. We impute the female labor income share as the regional average for these jurisdictions.

We classify countries into nine regions: Asia (excl. China) comprising 32 jurisdictions, China, the Former Eastern Bloc (24 jurisdictions), Latin America and the Caribbean (43 countries & jurisdictions), Middle East and Northern Africa (20 countries), Northern America (4 countries), Oceania (16 jurisdictions), Sub Saharan Africa (48 countries), and Western Europe (28 countries).

[1] World Inequality Lab, Paris School of Economics
[2] Gender Coordinator of the World Inequality Lab (theresa.neef@psemail.eu).
[3] West and Central Africa Coordinator of the World Inequality Lab

Data availability and quality

Micro data

We draw on four key micro survey data sets:

1. The EU-SILC (**European Union Statistics on Income and Living Conditions**) release of 2023, which covers 32 European countries for the period 2003–2020 (income year, wave year 2004–2021, except for IE). In this EU-SILC release, only 6 countries (AT, BE, BG, CY, CZ, and DE) release data for the income year 2020 (wave year 2021). For coverage see figure 1.
2. LIS (Luxembourg Income Study) providing individualized wage and self-employment income data for 51 countries for varying time spans. For coverage see figure 2.
3. The ILO Harmonized Microdata Repository[4] includes survey information on monthly pay from wage employment and self-employment separately for at least one year since 1990 for 134 jurisdictions. We corrected survey weights to align them with UN population data as the reference and excluded outliers, i.e. those observations with annual deviations in female (male) wage or self-employment income exceeding ±2 standard deviations from the cross-year mean. Original incomes are in current local currency. For coverage see figure 3.
4. National surveys for India and eight Western African countries. For India, we use the data from the national survey available for 1993, 1999, 2004, and 2009. For eight Western African countries - Burkina Faso, Benin, Cote d'Ivoire, Guinea Bissau, Mali, Niger, Senegal, and Togo - we can draw on harmonized national surveys for 2018 (see https://phmecv.uemoa.int/ for more information).

Aggregate data

For the regression imputations, we additionally draw on two aggregate data sources:

1. ILO modelled estimates of employment by sex, age and status in employment (obtained by the code EMP_TEMP_SEX_AGE_STE_NB_A), comprising 176 countries for our time period 1990–2023 (see https://ilostat.ilo.org/data/bulk/)
2. National accounts aggregates for net national income (B5n, S1), employees compensation (D1, S1), and net mixed income (B3n, S14) as well as population statistics from the World Inequality database. See Moshrif et al. (2024) for documentation. For coverage see figure 4.

[4] We thank the ILO's Data Production and Analysis Unit for kindly providing this data. Information on the dataset can be found in the following: D. Bescond, S. Kapsos, V. Karkee, D. Limani, Q. Mathys, Y. Perardel and M. Sodergren (2023). Unlocking the Power of Microdata: Enhancing International Comparability and Data Availability in ILOSTAT, and ILO Department of Statistics: Data Production and Analysis Unit (2024): ILOSTAT Microdata processing quick guide. Principles and methods underlying the ILO's processing of anonymized household survey microdata.

Main concept

Our main concept is the female labor income share at the country level defined as:

$$Female\ Labor\ Income\ Share = \frac{Labor\ income\ received\ by\ women}{Total\ labor\ income}$$

In line with the Distributional National Accounts method, labor income includes wage and salary income as well as the labor share of self-employed income:

$$Labor\ income = Wage\ income\ + 0.7\ self\ employment\ income$$

To compute the female labor income share for each country, we first aggregate labor income by gender within each country. Our inequality indicator thus comprises gender differentials in earnings as well as labor force participation. This value is therefore lower than the gender pay ratio, as it accounts for differences in earnings and labor force participation.

Imputation Methodology

The female labor income share is computed based on country-year aggregates of the female (male) wage (self-employment) income. We build this database in several steps.

1. Combining edited microdata sources EU-SILC, LIS, and ILO by country and year.
2. Selection of data sources in case multiple are available. We set a benchmark series for each country based on the following prioritization:
    a. First, we give priority to EU-SILC data, if available. Exceptions are Germany, Italy, Denmark, France, Great Britain, and Norway, for which we assume LIS to be of higher precision. For Serbia, we rely on ILO data instead of EU-SILC.
    b. For India, and eight Western African countries we use national survey data.
    c. If EU-SILC is not available, we draw on LIS data.
    d. If neither EU-SILC nor LIS are available, we use ILO micro data.

From these data sources, we obtain data for 138 countries (jurisdictions) and 1,214 country year observations.

3. We augment the benchmark data sources with secondary choices when they offer a broader year range. For example, we use EU-SILC data supplemented with LIS or ILO data for years prior to 2003, and LIS data supplemented with ILO data where applicable.
4. We interpolate the share of female (male) wage (self-employment) income in total labor income linearly between original data points and subsequently compute labor income aggregates and the female labor income share for these years.

5. We extrapolate aggregates back to 1990 and forward to 2023, keeping the female labor income share constant. The evolution of the labor income aggregate is benchmarked to the growth of net national income.

This procedure gives us complete times series from 1990–2023 for 138 jurisdictions.

6. For countries not included in this database, we use an imputation approach that proceeds in two steps. First, we estimate the female labor income share as a simple linear function of the female shares in wage- and self-employment and world region indicators using the combined LIS-EU-SILC-ILO database. Second, combining the estimated coefficients with ILO's employment series, we predict the female labor income share for all countries and years for which ILO modelled estimates exist. More precisely, we estimate the following regression model:

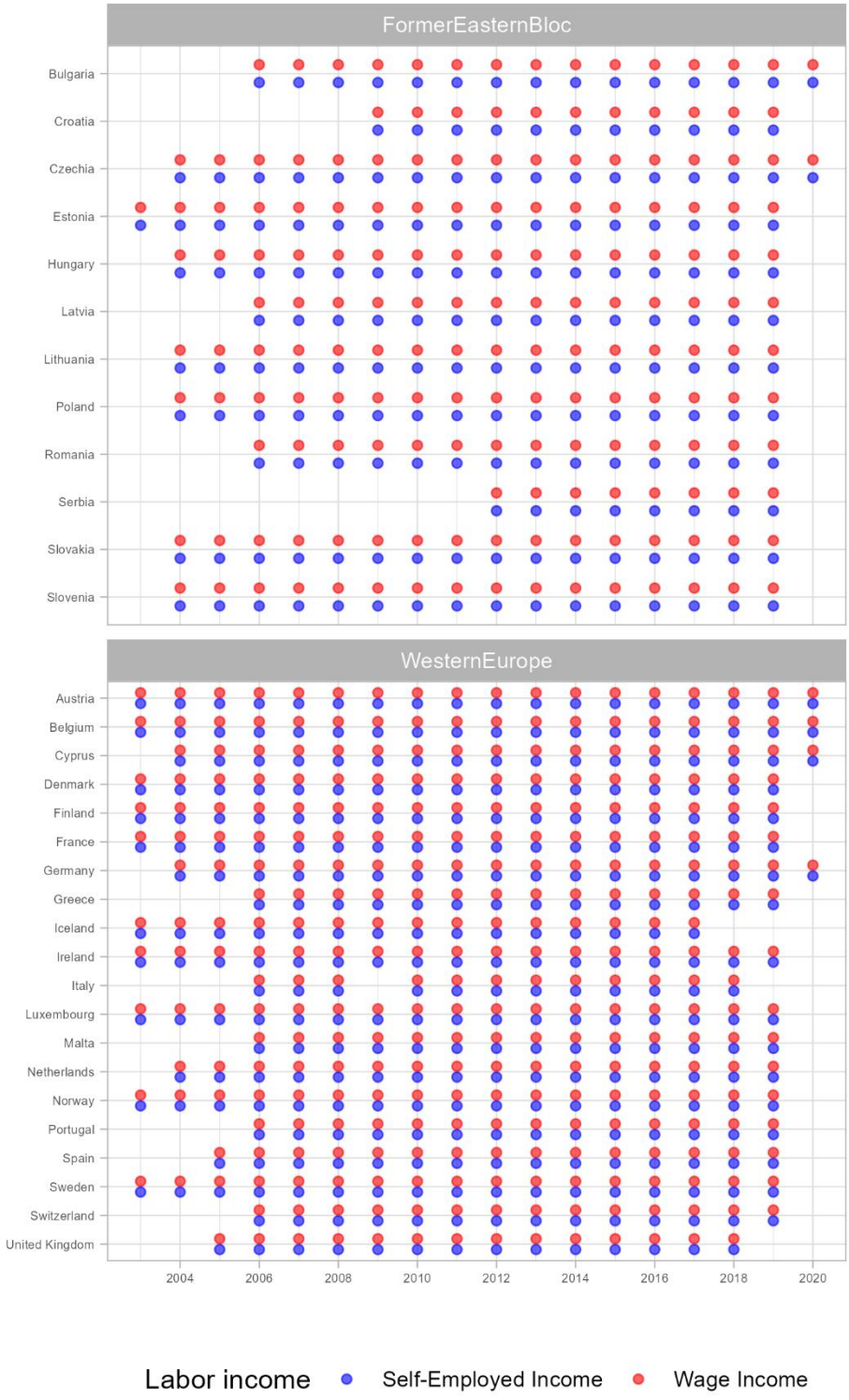
$$\begin{aligned} Female\ Labor\ Income\ Share_{ct} \\ = \ \alpha \ + \ \beta\ Female\ Share\ of\ Wage\ Employment_{ct} \\ + \ \gamma\ Female\ Share\ of\ Self\ Employment_{ct} \\ + \ \delta\ World\ Regions_c \ + \ \varepsilon_{ct} \end{aligned}$$

where *c* indicates countries and *t* years. The variable $Female\ Labor\ Income\ Share_{ct}$ is the female labor income share for country *c* and year *t*, the variables $Female\ Share\ of\ Wage\ Employment_{ct}$ and $Female\ Share\ of\ Self\ Employment_{ct}$ are the female shares of wage- or self-employed among all wage- or self-employed respectively. $World\ Regions_c$ corresponds to fixed effects for institutional and cultural differences of nine world regions. Observations are weighted according to population size.

The model fit is high. Employment variables contribute positively to the female labor income share, with wage employment having a significantly higher coefficient. This could be related to the fact that wage employment is associated with a more skilled labor force and higher earnings on average. Compared to Asia (excluding China) most regions exhibit a positive and significant fixed effect.

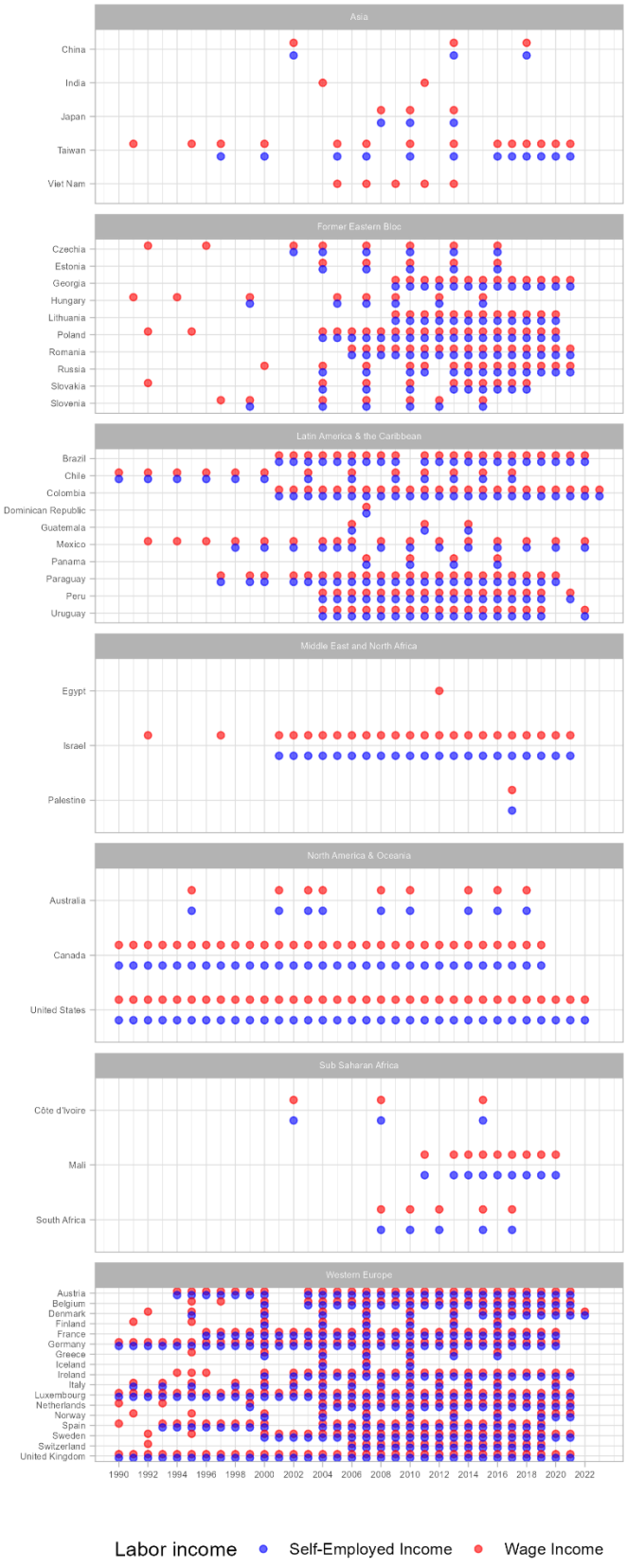
This imputation provides estimates for 36 additional jurisdictions, bringing our final dataset to include female labor income share estimates for 176 jurisdictions covering the period 1990–2023. For 43 jurisdictions without any labor market related information due to a lack of consistent data sources, we provide the regional average.

Figure 1: EU-SILC: Coverage individualized wage and self-employment income data.



Notes: This figure shows country-year observations with information on wage income (red) and self-employed income (blue) available in the European Union Statistics on Income and Living Conditions (EU-SILC) dataset. Income data spans 2003–2020 for 32 countries.

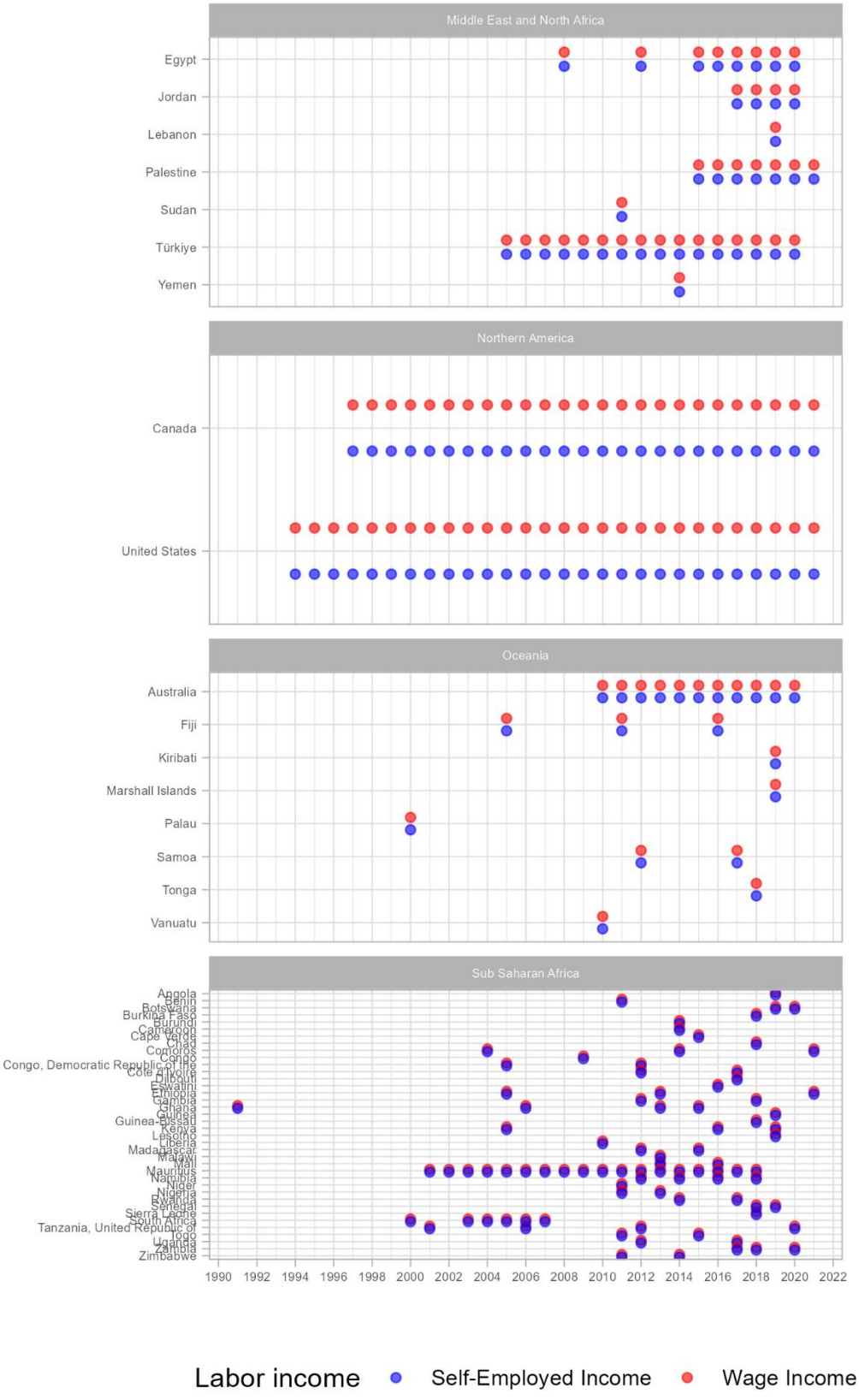
Figure 2: Luxembourg Income study: Coverage individualized wage and self-employment income data



Notes: This figure shows country-year observations with information on wage income (red) and self-employed income (blue) available in the Luxembourg Income Study (LIS). Individualized income data is available for 51 countries, covering various years between 1990 and 2022.

Figure 3: ILO Harmonized Microdata Repository: Coverage individualized wage and self-employment income data

Panel A.

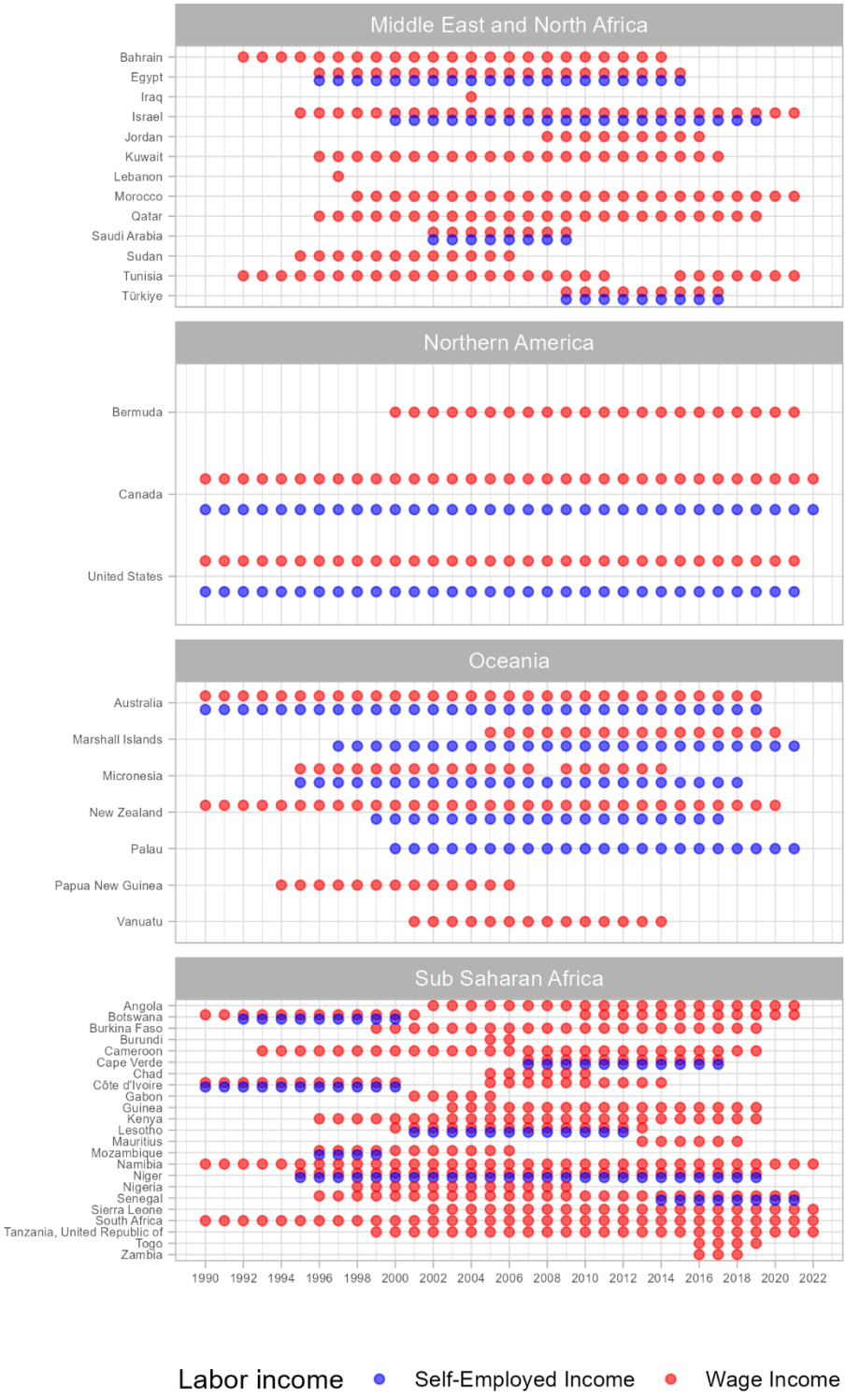


Panel B.

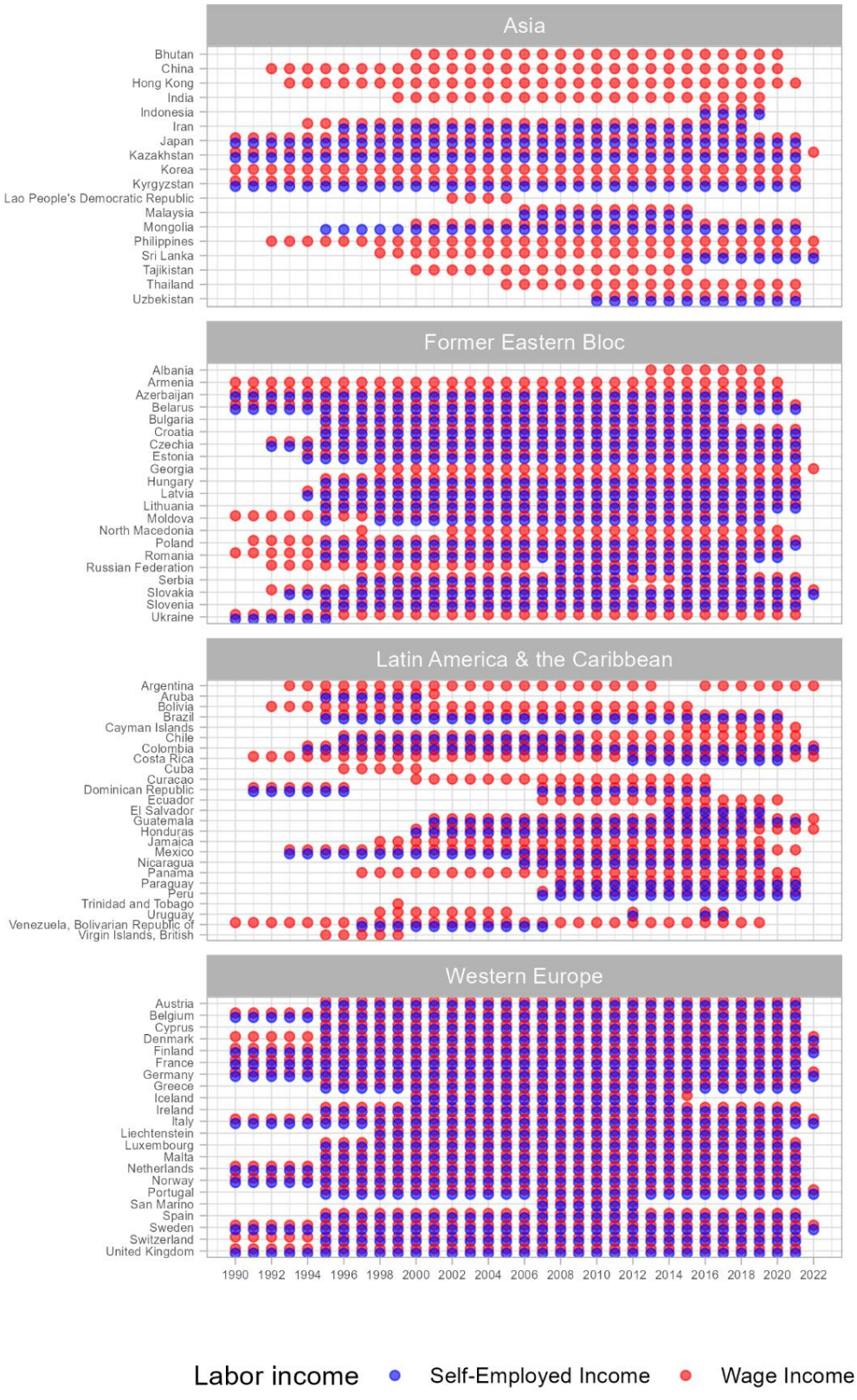
Notes: This figure shows country-year observations with information on wage income (red) and self-employed income (blue) available in the ILO Harmonized Microdata Repository. Individualized income data is available for 134 countries, covering various years between 1990 and 2022.

Figure 4: Wid.world national accounts aggregates: Coverage employees compensation (D1, S1) and net mixed income (B3n, S14).

Panel A.



Panel B.



Notes: This figure shows country-year observations with information for the national accounts aggregates employees compensation (red, D1, S1), and net mixed income (blue, B3n, S14) on the World Inequality Database. Aggregate data is available for 133 countries, covering various years between 1990 and 2022.