

EUKLEMS: Sources of the January 2023 Release

January 2023

Bernadette Biatour, bbi@plan.be

Bart Van den Cruyce, bv@plan.be

Abstract - Cette note présente les principales sources utilisées pour mettre à jour le volet belge de la base de données EUKLEMS.

Abstract - Deze nota presenteert de belangrijkste bronnen voor de update van de EUKLEMS databank voor België.

Abstract - This paper presents the main sources used to update the EUKLEMS database for Belgium.

Introduction

This paper gives an overview of the sources used to update the EUKLEMS database for Belgium. The database has been constructed according to the EUKLEMS methodology which is available on the project website (www.euklems.net).

The basic data are compatible with the National accounts published in October 2022.

1. Gross value added, Gross output, intermediate inputs

1.1. Sources

Published data from the National Accounts (October 2022 release): period 1995-2021, A38 and A64 industry level, current prices and volume (chained Euros, Laspeyres index, reference year 2015).

2. Compensation of employees

2.1. Sources

Published data from the National Accounts (October 2022 release): period 1995-2021, A38 and A64 industry level, at current prices.

3. Employment

3.1. Sources

Published data from the National Accounts (October 2022 release): period 1995-2021, A38 industry level (persons and hours worked) and A64 industry level (persons). Hours worked at A64 industry level: Federal Planning Bureau.

4. Labour and Capital compensation

4.1. Sources

Published and unpublished data from the National Accounts (October 2022 release): gross value added, employee wages, other taxes and subsidies on production, households' mixed income, at current prices, period 1995-2021, A64 industry level.

4.2. Comment

Capital compensation (CAP) is defined as the sum of gross operating surplus, capital part of mixed income and other net taxes on production (employment subsidies are removed from the total subsidies). Labour compensation (LAB) is defined as the sum of labour compensation of employees (employment subsidies included) and the labour part of mixed income.

To compute CAP and LAB, the mixed income of each industry must be decomposed into capital and labour parts. The labour income and the capital income of self-employed workers are first estimated by industry. The labour income of self-employed workers is based on the hourly wage of employees by industry and hours worked by self-employed workers. The capital income of self-employed workers is based on the rate of return on net capital stock per industry of the economy without the household sector (S1-S14) and the net capital stock of household sector (S14). These estimated incomes are summed by industry to create a virtual net mixed income. By industry, the shares of labour and capital are estimated in this virtual net mixed income and then applied to the net mixed income of the national accounts.

5. Labour input data and labour services

5.1. Sources

- Number of persons engaged by gender, age class, labour regime¹ and type² and educational attainment: qualitative employment data from the Federal Planning Bureau for the period 1999-2020 split in 38 Nace rev. 2 industries (A38) and compatible with national accounts in October 2021;
- Labour volumes (hours worked) by gender, age class, labour regime and type and educational attainment: qualitative employment data from the Federal Planning Bureau for the period 1999-2020 split in 38 Nace rev. 2 industries (A38) and compatible with national accounts in October 2021;
- Employee wage costs by gender, age class, labour regime and type and educational attainment: qualitative employment data from the Federal Planning Bureau for the period 1999-2020 split in 38 Nace rev. 2 industries (A38) and compatible with national accounts in October 2021.

Starting point for the labour composition data are the national accounts totals by SUT industry of October 2021. These industry totals have been detailed by gender and age class using administrative data sources. Thus, the number of persons engaged (both employees and self-employed), labour volumes, wage costs and gross wages for employees have been detailed by gender and age class using Social Security data³. The remaining detail was generated using Survey data.

Labour volumes for self-employed were detailed by gender and age class by combining the number of self-employed from above with estimates of their average hours worked using Labour Force Survey data for the period 1999-2020. The distribution of persons engaged as well as the labour volumes by educational attainment were also estimated using the Labour Force Survey data.

Educational attainment separates low-skilled (primary and lower secondary) from medium skilled (higher secondary and prolonged secondary education) and high skilled (higher short type, higher long type and university) workers.

¹ Distinguishing full-time and part-time workers.

² Distinguishing employees (blue or white collar workers or civil servants) and self-employed.

³ Including data from the National Social Security Office (NSSO), the National Social Security Office for Provinces and Local Authorities, and the National Institute for the Social Security of the Self-Employed (NISSE).

NOTE

To be able to yield plausible results for all possible combinations and to avoid year to year shocks caused by a small number of observations, all distributions by education level were estimated using regression techniques. Unless rendered impossible by a small number of observations, separate regressions were estimated per gender, age class and A38 industry. To capture the impact of the covid related lockdowns, a dummy variable for the year 2020 was added to the regressions for persons employed and hours worked⁴.

Labour costs by educational attainment level were obtained in two steps. First, gross wages were split by combining labour volumes per educational attainment level with skill premiums for gross wages. These skill premiums for educational attainment were estimated using data from the Structure and Distribution of Wages Survey for the period 2000-2019. The use of regression techniques allows to obtain skill premiums for the surrounding years 1999 and 2020 as well as annual results for industries that are represented in this survey only every four years^{5 6}.

In a second step labour costs were distributed proportionally over educational attainment levels according to the distribution of gross wages. This distribution was done at the working level combining detailed industry, gender, 5-year age class, labour regime and type. The difference between labour costs and gross wages are the social security contributions paid by employers. These are a concave function of gross wages in Belgium, which implies that a proportional distribution may overestimate the wage costs for workers that receive the lowest wages. Thus, within each narrowly defined cell given by gender, industry, age class ... labour costs for workers with the lowest education levels are likely to be overestimated.

For the year 2021, data on employees by industry, gender and age class are available from administrative data sources. The distribution by educational attainment for each category of workers (each industry, gender and age class) is estimated on the basis of the average 2019-2020 ratios.

For self-employed data, labour volume by gender, age class and educational attainment for 2021 is estimated by applying average 2019-2020 ratios to the data by industry (A38) from the national accounts. The distribution of the self-employed compensation (labour part of the mixed income) by gender, age class, and educational attainment on the period 1999-2021 is based on employees' hourly wages by gender, age class, and educational attainment.

⁴ The dummy for 2020 can pick up a possible impact of the covid lockdowns on the distribution of persons employed over education levels. This dummy was only included for private sector employees though (not for self-employed and public sector employees) and may lead to instable results in small groups. Since only 1% of the persons employed are represented in the labour force survey, a combination of age class, industry, sex and skills level that totals 500 persons in the EUKLEMS database should still be considered as small, since the results for the year 2020 will be based on 5 observations on average.

⁵ For the industries Education (P), Human Healthcare (QA), Residential Care and Social Work activities (QB), Arts, Entertainment and Recreation (R) and a part of Other Services (S), the Structure and Distribution of Wages Survey only has observations for the years 2006, 2010, 2014 and 2018.

⁶ Three A38 industries are not covered in the Wages Survey. These are Agriculture, Forestry and Fishing (industry A), Public administration and defence & compulsory social security (O) and Households as employers (T). For these three industries skill premiums for educational attainment level were obtained using data on net wages from the Labour Force Survey. The skill premiums for net wages obtained for the large industry O were transformed into skill premiums for gross wages by assuming that, given the gender, age class and type of worker, the shift from net to gross wages in industry O is similar to that in industry P (Education), for which we have both gross and net wages per education level. For industries A and T this adjustment has not been done.

5.2. Comments

Labour services reflect the changes in the amount and quality of labour input over time. It is assumed that the aggregate measure is a translog function of services of individual labour types. It is further assumed that the flow of labour services for each labour type is proportional to hours worked, and workers are paid their marginal productivities (see EUKLEMS Methodology Manual for more details).

For each industry, the index of labour services input is a translog quantity index (Törnqvist index) of individual types, indexed by i

$$\ln(L_{t+1}/L_t) = \sum_i 0.5(v_{t+1}^i + v_t^i) \ln(H_{t+1}^i/H_t^i)$$

where weights are given by the average shares of each labour type in the value of labour compensation $v_t^i = \frac{P_t^i H_t^i}{\sum_i P_t^i H_t^i}$ with P_t^i the price of one hour work of labour type i . In this way, aggregation takes into account the changing composition of the labour force. Typically, a shift in the share of hours worked by low-skilled workers to high-skilled workers will lead to a growth of labour services (variable LAB_QI in the database) which is bigger than the growth in total hours worked (H_EMP in the database). The labour composition is measured as the difference in growth of labour services and hours worked.

6. Capital input data and capital services

6.1. Sources

- Gross fixed capital formation: Published data from the National Accounts (October 2022 release): period 1995-2021, A38 and A64 industry level, current prices and volume (chained Euros, Laspeyres index, reference year 2015) ;
- Net capital stock: Published and unpublished data from the National Accounts by asset (AN) and by industry (A38), (October 2022 release): period 1995-2021, A38 industry level current prices and volume;
- Consumption of fixed capital: Unpublished data from the National Accounts by asset (AN) and by industry (A38), (October 2022 release): period 1995-2021, current prices.

6.2. Comments

Capital services

Capital services are calculated based on official data of capital stocks. In the Belgian National Accounts, the capital stocks are estimated using the perpetual inventory method with a log-normal survival function and a straight line depreciation function.

For the aggregation of capital services over the different asset types it is assumed that aggregate services are a translog function of the services of individual assets. It is further assumed that the flow of capital services for each asset type is proportional to its stock, independent of time. Hence the corresponding

NOTE

index of capital input K is a translog quantity index of individual assets in a particular industry given by:

$$\ln(K_{t+1}/K_t) = \sum_i 0.5(v_{t+1}^i + v_t^i) \ln(K_{t+1}^i/K_t^i)$$

where weights are given by the average shares of each component in the value of capital compensation

$v_t^i \equiv \frac{uc_t^i K_t^i}{\sum_i uc_t^i K_t^i}$ and uc_t^i the price of capital services from asset type i . The price of user cost of capital is

determined by the nominal rate of return, the rate of economic depreciation (standard EU KLEMS depreciation rates) and the asset specific capital gains (see EUKLEMS Methodology Manual for more details).

Contribution of tangible and intangible capital services to value added growth

The growth accounting model allows value added growth to be broken down into the contributions of labour, capital and total-factor productivity (TFP).

Due to the EUKLEMS hypothesis setting the negative user costs to zero, the sum of the contribution of tangible and intangible capital services to value added growth (based on CAP_{Tang} and CAP_{Intang}) can be different from the contribution of total capital (based on the "correct" CAP). The contribution of TFP to value added growth is estimated as residue with the contribution of total capital in the equation.