3. Data and stylised facts

3.1 Data

We have compiled a comprehensive database for nine OECD economies drawing on six publicly available international databases for sectoral data which we augmented by country level data.\(^3\)

We measure the wage share as labour compensation over value added with data obtained from the EU KLEMS database. Labour compensation includes the wage of self-employed workers, imputed based on the assumption that their wage is equal to the average hourly wage of the sector. Different concerns have been raised with regard to this imputation, as it is generally said to overestimate the wage share for sectors of predominantly low skilled workers while it underestimates high skilled sectors’ wage shares. Indeed we find the wage share to exceed 1 in a total of 588 out of 13796 cases (4.26\%) for data at 2 digits and 324 out of 10245 observation (3.16\%) for the 1-digit level.\(^4\) However, wage shares exceeding one are not generally a problem and can naturally arise for mainly two reasons which have nothing to do with overestimations of the imputed wages for self-employed workers: First, if a sector incurs heavy losses and second, if a sector receives significant subsidies (EU KLEMS, 2007). The second case arises because value added in KLEMS is calculated as compensation of employees plus operating surplus plus taxes minus subsidies (on labour and capital), i.e. at basic prices, and therefore can fall short of labour compensation if the subsidies exceed operating surplus and taxes in a particular period.\(^5\) Since data from EU KLEMS is only available until 2009 we extrapolate through splicing. More specifically, we link the wage share from KLEMS with the growth rate of the wage share obtained from the OECD Structural Analysis database (OECD STAN).\(^6\) Both series have a correlation of 0.91. We control for violent swings in the wage share by excluding years where the percentage change

\(^3\) The use of an international database is instructional for making the variables and estimations comparable between countries. See table A3 in the appendix for further information on sector definitions and the skill taxonomy.

\(^4\) This number excludes Agriculture, Fishing and Foresting. These sectors are repeatedly reported to have wage shares bigger than one because of poor data quality and because the imputation for self-employed workers largely overestimates the labour compensation for this low skilled sectors (EU Klems growth and productivity, p. 37).

\(^5\) It would be preferable to use value added at factor cost for the calculation of the wage share. Unfortunately, there are no long series on taxes minus subsidies on production in EU KLEMS.

\(^6\) Since self-employed are not included in the measure of labour compensation in OECD STAN we impute their wages by applying the same technique as in EU KLEMS. We exclude observations where the number of self-employed suddenly falls to zero, assuming that it must be related to a measurement error.