

The four do files run the analysis in the paper for the US KLEMS, EU KLEMS, and the NIACS & SIC versions of the EU KLEMS US data. These produce the tables and figures in the paper.

The Stata files USKLEMS, EUKLEMS, USNAICSEUKLEMS, USSICEUKLEMS contain the USKLEMS and EUKLEMS data used in the analysis (publicly available on the BLS multifactor productivity webpage and University of Groningen EU KLEMS web pages – please cite those sources if using their data). The EUKLEMS data I use are the March 2008 release (chosen, in my initial review of their data, on the basis of the range of estimates and data available).

The Stata files Instruments and EUInstruments are instruments and controls (unemployment, capacity utilization) developed using FRED, SIPRI & Stock & Watson data, as explained in my paper (please cite those sources if using their data). The Stata file CapUtilization contains the Fed's capacity utilization index by industry.

The Excel files Chain47on.xls, VA47on.xls, GDPbyInd_VA_SIC.xls are the BEA files that I use to compare EUKLEMS SIC goods & services quantity growth to official US data in a footnote in the paper. I also use the first two of these files near the end of the paper to calculate long run US NIPA growth in the relative quantity and price of goods and services.

The Excel file USKLEMSAdjusted.xls presents my estimates of TFP by sector adjusted for labour composition (based on unadjusted BLS TFP numbers). Along with these are the Stata files USWagesData.dta & USWorkersData.dta, which are my estimates of wage per hour, workers and hours by sector, as described in part B of the on-line appendix.

Variables in the Stata files contain descriptions and should be self explanatory. I should clarify, however, one item of variable notation. I use YYDif and XXDif to denote my main estimates (i.e. used in the main analysis) of total factor productivity growth and labour-share-weighted changes in employment shares. For both the EUKLEMS and the USKLEMS YYDif is total factor productivity growth adjusted for the changing composition of the labour force (in the case of the US, adjusted by me). For XXDif, for the USKLEMS it is the labour-share-weighted change in sectoral employment shares by worker type:

$$(a) \quad \sum_j \Theta_{Li}^j \hat{\pi}_i^j$$

but for the EUKLEMS and SIC & NAICS EUKLEMS based US analysis it is the labour-share-weighted change in total sectoral employment:

$$(b) \quad \Theta_{Li} \hat{\pi}_i$$

See equations (7) and (15) in the paper. As explained in footnote 20 in the paper (and also noted again in Table 1), the EUKLEMS data only allow for the calculation of the distribution of total workers by sector (not workers by type - the data frequently contains distribution of hours by worker type, but not workers). (a) and (b) are identical if the distribution of workers by type is proportional to the industry share of total employment, i.e. $L_i^j = L^j(L_i/L)$. Finally, I note that in the USKLEMS data the variables YDif and XDif are productivity growth without adjustment by worker type and the labour-share-weighted change in total sectoral employment, based on the original BEA data (without my computations of changing worker composition). These produce similar results (see footnote 33 in the paper).