ReadMe

You will find the following files:

data.pdf: list of variables in the Gauss data set **data.dat** (data set not included)

DESCRIPA.prg and **DESCRIPB.prg:** Gauss code to get the descriptive statistics in Table 1, Table 2, Table OA1 and Table OA2. In the first program you can get the different statistics by choosing sets of variables commenting them on and out using the symbol @.

OLS.prg, E12GMM.prg and E15GMM.prg: Gauss code to run the OLS regression of Table 3 and the nonlinear GMM estimations of equation 12 and equation 15 reported in Table 4 and Table 7 respectively. **E15GMM.prg** generates productivity estimates in differences and in levels that are used by the next set of programs. Using the GMM code, appropriately modified, you can run any other estimation in the paper.

ratedescrip.prg, **arates.prg, ratecorr.prg, levcorr.prg, levcomp.prg and opmp.prg**: generate the descriptive statistics of the differences and levels of productivity included in Table 5, Table 8, Table 9, Table OA5 and Table OA6.

prodd.prc, prodl.prc and OPprodA.prc: are procedures used by the code.

share_decomposition.xlsx: excel file with the numbers of T6 and Figure 2.

The Gauss data set **data.dat** should be requested from Fundacion SEPI in Spain by emailing desee@fundacionsepi.es with an explanation that the data is required for replication purposes.

To run the programs, you need to have Gauss and the optimization routine called Optmum (version 3.1.7). Any recent Gauss version should work (code was written in Gauss 14 but has also been run in Gauss 16), but please note that Gauss Light is insufficient to run most of the features.

To run the code, first create the following directory structure. Use a directory of your choice, but create eight subdirectories: "data," "procs," "matrices12," "matrices15," "cvars12," "sdcorr," "dprod" and "lprod". Place the data file **data.dat** in the "data" subdirectory and the three procedures in the "procs" subdirectory. The subdirectories "matrices12," "matrices15," "cvars12" and "sdcorr" will automatically collect all the intermediate inputs to be kept as you run the GMM code. The first two keep the optimal weighting matrices, the other two the components needed for the correction of the standard errors when running **E15GMM.prg.** The folders "dprod" and "lprod" will collect the estimates of productivity in differences and levels.

Leave all the files with ".prg" extension in the main directory. You are ready to reset Gauss to the main directory and run the programs.

When running **E12GMM.prg and E15GMM.prg** you must set the **industry** and **optimal** options. First stage (**optimal**=0) estimates use the standard consistent weighting matrix based on the instruments and automatically stores the (inverse of the) optimal weighting matrix in the corresponding subdirectory "matrices." This matrix is automatically read and inverted when you perform the second stage (**optimal=1**).

You should first run the first and second stage estimation of equation 12 in order to have the appropriate inputs ready to run the first stage estimation of equation 15. Then the first stage estimation of equation 15 with **E15GMM.prg** will generate the productivity estimates. All the code describing the productivity estimates is written assuming that the researcher has run initially the first (optimal=0) and the second (optimal=1) stages of **E12GMM.prg,** and then the first stage of **E15GMM.prg** for all industries**.** If you miss an industry, the code will give an error.