Accounting for structural change

Due Feb 19th

For this assignment, you will evaluate the contribution of structural change to aggregate growth. To be clear, there is no *model* involved in this assignment. This is pure accounting work. There are two main parts to the assignment.

Part 1 - Derivation

Total output is the sum of output in separate J industries

$$Y_t = \sum_j^J Y_{jt}$$

where each individual Y_{jt} term is in real terms (meaning the values are comparable, and can be summed up). In addition, you know that L_{jt} workers are in each industry j, and there are a total of L_t workers. Let $s_{jt} = L_{jt}/L_t$ be the share of workers in industry j, and let $y_{jt} = Y_{jt}/L_{jt}$ be output per worker in industry j. Similarly, $y_t = Y_t/L_t$ is aggregate output per worker. Let $\Delta x_t = x_{t+1} - x_t$ for any given variable x.

Show how to derive the following accounting identities

$$\Delta y_t = \sum_j^J y_{j,t+1}\Delta s_{jt} + \sum_j^J s_{jt}\Delta y_{jt}.$$

The first term on the right-hand side here is the "between" growth in output per worker, meaning it comes from the shift of workers between industries. The second term on the right is the "within" growth, as it captures changes in output per worker coming from industries, holding the share of workers in an industry constant.

You can also derive the following, similar relationship.

$$\Delta y_t = \sum_j^J y_{jt}\Delta s_{jt} + \sum_j^J s_{jt}\Delta y_{jt} + \sum_j^J \Delta s_{jt}\Delta y_{jt}$$

Here the first term is again the "between", the second is again "within" (but notice the subscript differences), and the last term is the "covariance" term. This last term captures whether industries that were growing in output per worker were also growing in their share of labor.

Part 2 - Calculation

This is data intense. By which I mean you will need to pull down raw data and manipulate it yourself in order to put it into a form that you can use to calculate the above breakdowns. You should got to the GGDC website, and the 10-sector database. Download it (it comes in either Excel or Stata).

For each country in that database, you should calculate the right-hand side terms of both accounting identities above for ten year periods, 1960-1970, 1970-1980, 1980-1990, 1990-2000, 2000-2010. Not every country will have data in all those periods, but you'll have roughly five observations for each country.

Calculate for each observation the percentage of total growth in output per worker that was due to "within" growth and "between" growth using the first decomposition

Calculate for each observation the percentage of total growth in output per worker that was due to "within", "between", and "covariance" growth using the second deomposition

Write-up

You should write up a short paper/report discussing your results. Specifics 1. Explain what it means for a "between" term to be negative 2. Explain what it means for the "covariance" term to be negative 3. Discuss whether there is any general time trend to the percentages due to the different sources of growth 4. Document/discuss any countries or time periods in which "between" growth was of particular importance (positive or negative). Was this due to a shift into or out of any paricular sector? 5. Do some simple analysis (some OLS regressions and scatter plots) of the contribution of "within" and "between" and "covariance" growth and how it relates to the absolute size of output per worker. Are there any noticeable relationships? (I don't ask this because I know the answer, I'm curious)

Your analysis should all be written in Latex, and include any figures or tables you produce to show your results. In addition, you *must* turn in the code you used to do the data analysis.