README CONCORDING CN8 OVER TIME

This readme file accompanies the .rar file (zipped) "cn8_over_time.rar". When you extract, you will see that there is a do-file called "CN8_over_time.do". The "README" file in the .rar is useful, there are still issues that are not so clear. Below I write the separate steps below that are involved more clearly.

1) Open the do file "CN8_over_time.do" and inside you change the **beginning** year and the **end** year to the ones that we need. For example, you may wish to concord product codes from the year **1997** to the year **2007**. All you need to do is to supply the **first** and **last** year of your data and the do file will do the concordance for you.
2) Run the do file in the directory where you opened the .rar file
3) The do-file generates a file in \output\cn8_cn8plus_bbbb_eeee.dta, where "bbbb" stands for begin year and "eeee" for end year that you specify in the do_file at the top. This file is sorted at cn8-year level.
4) The .dta data that is created contains the following variables: **cn8**, **cn8plu**s, **year**, **synthetic**.
   Cn8: the code as it is listed in every year in the data. **Cn8plus**:a new code that takes into account whether the CN8 classification code changed over time and that puts these changes in one family code called cn8plus. **Synthetic** is just a dummy variable indicating whether the cn8 code changed over time. If you do tab "synthetic" you will find that about 33% of codes have changed over time (between '97-'07). Of course this does not say anything about the value of trade that is contained in these changed codes in our trade data. We will have to check that.

5) Now turn to the dataset that you want to concord.
   When the data is uploaded in stata make sure that:
   a) The name of the product variable is cn8
   b) Make sure cn8 is numeric (leading zero drops !)
   c) Sort cn8-year
   d) Merge it with file cn8_cn8plus_bbbb_eeee.dta that was created in the do file above
   e) Keep if _m==3
6) At this point the work is not over yet. At the end of the do-file CN8_over_time.do, an example is given on what to do after the merge in order to get a new dataset with concorded product classifications. Basically it boils down to the fact that you still have to collapse the data on the new product code cn8plus. The example at the end of the do-file CN8_over_time.do refers to product-year trade which I adjusted a little to "firm-product-year-land" trade.
   So the additional steps that you need to take are given by the following piece of code, that you perform after the merge:

   ```
   replace cn8plus=cn8 if cn8plus==.

   Sort year cn8plus

   Collapse (sum) value weight units, by (year vat land cn8plus)
   ```

Save your new dataset.

7) One potential problem is however that in the collapse, the dummy synthetic disappears and we may have to create our own dummy “special” again that indicates whether a code has changed over time. There is a piece of code in one of the older do-files for this which I will look up and send you.

   Also, it has to be noted that when making CN8 a numeric variable you lose the leading zero on some of the CN8 codes. You may need to put it back “on” when turning cn8plus into a string again.