

Calculating Employment Rates

Goal

The employed population can be defined differently depending on the research purpose. In labour market analysis, you may wish to calculate employment rates or run wage regressions. In this case, you may want to identify persons working at a certain moment in time. Within the framework of this strict reference period, a person may be considered as employed as soon as he/she has carried out any work (i.e., the International Labour Organisation (ILO) definition of employment, LIS variable *pclfs*). Those more interested in the primary employment status at a certain moment in time should not use the ILO definition, but rather look at those whose main activity is employment (LIS variable *pcmas*).

Those wishing to perform income distribution analyses according to the activity status of individuals (e.g., calculation of poverty rates for the working poor), may choose to focus on employment over a longer reference period. In this case, you need to be able to identify an individual's primary activity over the income reference period (which in LIS is normally one year, LIS variable *pumas*).

In this exercise, we identify the employed population according to these three different concepts of employment.

Activity

Calculate employment rates in the US and Germany in 2000 using the three different employment measures. Use the sample of individuals 18 years of age and older. Recode all of those with unclear employment attachment (coded in 900s) to missing.

Guidelines

- The employment rate is the percentage of employed persons in the adult population, which we define here as 18 years of age or older.
- There are three different LIS variables that coincide with the definitions of employment described above:
 - persons carrying out any work at present can be identified using *pclfs*;
 - persons for whom work is the main activity at present are in *pcmas*;
 - persons who are mainly employed during a longer reference period can be found in *pumas*.
- By definition, those coded in the 900s may have some employment attachment, but the information is too ambiguous to either define them either as “employed” or “not employed”. These individuals were kept in separate detailed categories to allow users to redefine them based on the specific research project. Always check the country-level documentation to determine how you wish to recode those assigned to the 900 (and 400, in *pclfs*) category.

➤ In order to measure employment, create dummy variables for employment status, by recoding the original variables into indicators of dichotomous employment status (yes=1 & no=0).

- The generic syntax for this statement is as shown below. NOTE that here the code is kept as simple as possible. In the long run, try to use this statement more often:

```
ARRAY emp {*} emp1-emp3;
```

- Use a loop statement to apply similar code on each array elements at the same time.

```
DO i = 1 TO 3;  
  IF (900 <= emp(i) <= 999) THEN emp(i) = 1;  
END;
```

- The index *i* starts with the value 1, is then incremented by the increment value each time. The iterations continue as long as the index *i* is less than or equal to the stop value (here the dimension of the array). If a negative increment is used, then the rules reverse so that the index *i* decrements to a lower bound. If you don't use the **BY** statement the increment is by default equal to 1.
- Note that the *index* is stored as a variable in your output dataset. Don't forget to drop it in your output dataset.

```
DATA emp (DROP=page pclfs pcmass pumas i);
```

Program

```
OPTIONS NOSOURCE NONOTES NOFMterr NODATE NOCENTER LABEL NONUMBER LS=MAX
PS=MAX ;

%MACRO emp(pi);

    DATA emp    (DROP=page pclfs pcmas pumas i);
    SET &&pi.p (KEEP=country pweight page pclfs pcmas pumas);
    * Keep adult (18+) ;
    IF page > 17 ;
    * create variables emp1 emp2 emp3 ;
    emp1 = pclfs ;
    emp2 = pcmas ;
    emp3 = pumas ;
    * create a table to loop into the mapping process ;
    ARRAY emp {*} emp1-emp3 ;
    * create recoded pclf, pcmas, pmas with strict definition ;
    DO i = 1 TO 3;
        IF          (900 <= emp(i) <= 999) OR emp(i) = -1 THEN emp(i) = .
;
                ELSE IF (100 <= emp(i) <= 199)                THEN
emp(i) = 1 ;
        ELSE IF  emp(i) NOT IN (.,0,1)                THEN emp(i)
        = 0 ;
        END ;
    RUN ;

    * get the ratios;
    PROC FREQ DATA=emp ;
        TABLES emp1-emp3 / NOCUM;
        WEIGHT pweight ;
    RUN ;

%MEND emp ;
TITLE " US00 - Employment rates"; %emp(us00)
TITLE " DE00 - Employment rates"; %emp(de00)
```

Results

	US00	DE00
At present (any work)	65.49	59.64
At present (main activity)	59.27	45.76
During the calendar year (main activity)	63.15	54.38

Comments

- As expected, the employment rate is much higher if employment is defined as any work (*pclfs*) than if it is defined as the main activity held (*pcmas* or *pumas*).
- Be sure to look at the Value Labels and read the Comments/Warnings in the Lissification tables! This is where you will find information about the 900 category.
 - In most cases, the difference between including the 900s or setting them to missing in *pclfs* is minor or non-existent. Determining whether an individual held a job in a specified period is fairly straightforward and the 900 category is reserved for truly marginal cases.
 - More ambiguity arises when defining primary activity (*pcmas* and *pumas*). This happens most often when questions about primary activity focus on employment. In some surveys, questions about non-employment are only asked of those who were not employed in *pclfs*, so those who appear to have marginal attachment to the labour force (e.g., part-time or irregular workers), but who have specified no other activity are coded in the 900s.