

Comparing Educational Outcomes

Goal

When comparing educational levels across countries, it is necessary to carefully look at the labels of those variables for each country, and recode them to make them comparable across the countries you are investigating. Or you may wish to use the education routine created by LIS (as seen in the previous exercise). In this exercise, the educational population structure of different countries is compared with the use of the LIS educational routine.

Activity

Compare the educational composition of the total adult populations (16+) of the US, Luxembourg, and Italy in 2000 by gender. Repeat the exercise for the wage-earning population. Calculate average wages of wage earners for each country by education level.

Guidelines

- Please note that for some datasets, including Italy and Luxembourg, income is reported net of taxes and social contributions.
(For more information, see <http://www.lisproject.org/techdoc/netdatasets.htm>). For those datasets, use net wages instead of gross wages. Net wage is reported in a different variable (*pnwage* rather than *pgwage*).
- One way to ensure that you choose the correct wage is by creating a new variable with the wage you want to use. In this case, you want *pgwage* if it is available and *pnwage* otherwise. You can create the new wage variable immediately by adapting the following code:

```
DATA youroutput (DROP=avarlist);
  IF _N_ = 1 THEN
    DO UNTIL (last);
      SET yourinput END=last;
      cpt + pgwage ;
    END;
  SET yourinput (KEEP=yourvarlist);
  IF cpt = 0 THEN wage = pnwage ;
  ELSE          wage = pgwage ;
  ... ;
RUN;
```

In a data step, SAS can read an input dataset into multiple-stages. In the example above,

- SAS first read entirely the *yourinput* dataset within a loop (`IF _N_ = 1 THEN DO UNTIL(last) ... END;`) monitored by the index *last*. The index is automatically generated by the `END` option of the `SET` statement that allows SAS to know when the end-of-file has been reached. During that stage SAS aggregates the sum of *pgwage*.
- SAS then read again the *yourinput* dataset that additionally includes the sum of *pgwage* as a new variable (*cpt*)

Keep in mind that in some countries, both *pgwage* and *pnwage* exist. In that case, this code will always select *pgwage* over *pnwage*. If you prefer to use *pnwage* when available, adapt the program accordingly.

- Remember to include the variables necessary for the standardisation routine in your keep statement in addition to *pweight* and the applicable wage variable.
- The `PROC FREQ` procedure allows you to requests one-way to n-way frequency and cross-tabulation tables and statistics for those tables.
- To request a two-way cross-tabulation table, use an asterisk between two variables. To request a multi-way table (an n-way table, where $n > 2$), separate the desired variables with asterisks. The unique values of these variables form the rows, columns, and strata of the table. `PROC FREQ` produces a separate cross-tabulation table for each stratum.

```
PROC FREQ ...;
  TABLES var1*var2;
RUN;
```

Program

```
OPTIONS NOSOURCE NONOTES NOFMterr NODATE NOCENTER LABEL NONUMBER LS=200
PS=MAX ;

%MACRO educ;
%LET i = 1 ;
%DO %UNTIL (&i > 3) ;
    %LET country = %scan(&pi,&i) ;
    DATA educ (DROP=peduc country pgwage pnwage ptocc) ;
        IF _N_ = 1 THEN
            DO UNTIL (last) ;
                SET &&&country.p END=last ;
                cpt + pgwage ;
            END ;
        SET &&&country.p (KEEP=psex pweight country page pgwage pnwage peduc
ptocc);

        IF cpt = 0 THEN wage = pnwage ;
        ELSE wage = pgwage ;
        %INCLUDE "i:\educrcodepp.sas";
    RUN;
    PROC SORT DATA=educ ;
        BY psex educ ;
    RUN ;
    TITLE "COUNTRY : &country - Total Population" ;
    PROC FREQ DATA=educ FORMCHAR (1,2,7)='|+ ';
        WHERE page>15 ;
        TABLES educ*psex / FORMAT=5.2 MISSING NOFREQ NOROW NOCUM ;
        WEIGHT pweight ;
    RUN ;
    TITLE "COUNTRY : &country - Earners" ;
    PROC FREQ DATA=educ FORMCHAR (1,2,7)='|+ ';
        WHERE (page>15 AND wage>0) ;
        TABLES educ*psex / FORMAT=5.2 MISSING NOFREQ NOROW NOCUM ;
        WEIGHT pweight ;
    RUN ;
    TITLE "COUNTRY : &country - Average Wage" ;
    PROC MEANS DATA=educ MEAN ;
        WHERE (page>15 AND wage>0) ;
        VAR wage ;
        WEIGHT pweight ;
        BY psex educ;
    RUN ;
    %LET i = %eval(&i+1) ;
%END ;
%MEND educ;
%LET pi = us00 lu00 it00;
```

%educ

Results

	<i>Educ Level</i>	US00		LU00		IT00	
		Males	Females	Males	Females	Males	Females
<i>Percent of total population</i>	low	20.3	19.2	30.4	45.2	57.9	63.3
	medium	49.0	51.0	36.4	28.2	33.6	29.6
	high	30.7	29.8	26.1	19.0	8.5	7.1
<i>Percent of wage earning population</i>	low	14.8	11.9	26.0	32.8	47.7	33.0
	medium	50.6	52.1	39.1	34.9	41.3	50.4
	high	34.6	36.0	34.4	31.8	11.0	16.7
<i>Average wage of wage earners</i>	low	17,744	10,364	883,626	514,064	23,334	17,142
	medium	33,026	20,215	1,204,372	681,842	28,717	22,112
	high	64,344	35,711	1,747,091	1,107,552	40,286	27,220
<i>Returns to education</i>	low →med	86%	95%	36%	33%	23%	29%
	med →high	95%	77%	45%	62%	40%	23%
<i>Gender wage gap</i>	low	1.7		1.7		1.4	
	medium	1.6		1.8		1.3	
	high	1.8		1.6		1.5	

Comments

- Please note that education was not recoded for some countries in certain years. Check documentation on-line (lissification tables and descriptives) for more precise information about education levels.
- Please note that in the results above, the percentage of population by level of education may not add up to 100% because the category *educ* =9 (missing or not defined) was not included in the table, but was included in the calculations.
- The education composition across countries varies considerably. (Italy has the least-educated population of the three countries chosen for this comparison.)
- In all countries, wage earners are more educated than the total population.
- As expected, wages increase with the level of education, but to a different extent in each country. In the US, returns to education are substantially higher than in Luxembourg or Italy.
- Net versus Gross wages:

Please note that, even when considering exchange rates (or PPPs), it is not possible to directly compare the level of wages between countries that report either net or gross wages. In these cases, it is only possible to compare ratios. Even then, a progressive taxation system might affect the ratios. If high wage earners (i.e., the most educated) face higher tax rates than low earners, the returns to higher education will be lower than if the returns had been measured using gross wages. The higher gross returns, therefore, are partly offset by higher taxes.