

Gender (Person Level)

Goal

In doing any estimation, it is important to be careful about the unit of analysis. In research focusing on women, you must take into account whether the data you are using are individual- or household-specific.

Using individual-level data allows you to identify individual-specific income, but problems may arise in estimation depending on your research question. Some of these issues are general, but others are specific to the LIS database. First, some income sources are common to the household (such as child benefits or housing allowances) and are not available at the individual level. In LIS, certain individual income sources (invalidity and work accident pensions, sickness and maternity allowances, means-tested benefits, social transfers) are reported in detail only in the household file. The information is present in the person-level file in an aggregated form.

In this exercise, we introduce income analysis by gender. Using the person-level file, we will focus only on earned income amounts, not considering social transfers.

Activity

Examine the working-aged population (25 to 60, inclusive) in the UK in 1999 and the US in 2000. Compare the percentage of working men to that of working women (defined as those with positive earnings from any employment). Calculate the average total income by gender of both the total working-aged population and the working population. Estimate the gender earnings gap for both the working-aged population and for those who work.

Guidelines

- For this exercise, define the “working-aged” population as those aged 25 to 60, inclusive, and the “working” population as those with positive earnings from paid and/or self-employment (*pgwage* + *pself*).
- The gender income gap is defined as the ratio of average total earnings (*pgwage+pself*) of males to females.
- To simplify the analysis in this exercise, set negative values of *pself* to missing before calculating the “working” population. Failure to do so may result in self-employed with negative incomes being counted as not working, or negative incomes being considered in the average for the population.
- **SAS reminder:**
 - Only keep the variables you will be using. This avoids unnecessary burden on the machine so that submitted jobs will run faster.

```
(KEEP=casenum d6 d27 married d1 d3 dpi);
```

- To keep observations that pertain to a certain group of persons, use the statement **IF <expression>**. The expression is assessed by SAS as a Boolean (getting the value either true or false). Only observations that receive a true assessment are kept.
- **IF ((page>15) AND (pgwage>0));** this expression allows you to select earners.
- **IF condition THEN action;**
<ELSE action;>

This piece of code tests whether the condition is true; if so, the action in the **THEN** clause is carried out. If the condition is false and an **ELSE** statement is present, then the **ELSE** action is carried out. If the condition is false and no **ELSE** statement is present, then the next statement in the **DATA** step is processed. The condition is one or more numeric or character comparisons. The action must be an executable statement; that is, one that can be processed in an individual iteration of the **DATA** step. In SAS processing, any numeric value other than 0 or missing is true; 0 and missing are false. Therefore, a numeric value can stand alone in a comparison. If its value is 0 or missing, then the comparison is false; otherwise, the comparison is true.

- Useful **PROC MEANS** statement options:

SUMWGT Calculate separate statistics for each BY group
MISSING Use missing values as valid values to create combinations of class variables

- Useful **PROC MEANS** statements:

BY Calculate separate statistics for each BY group
CLASS Identify variables whose values define subgroups for the analysis
TYPES Identify specific combinations of class variables to use to subdivide the data

```
PROC MEANS DATA=youroutput MEAN SUMWGT MISSING;
  CLASS wap wp psex;
  TYPES wap*psex wp*psex ;
  VAR yourtotalincome ;
  WEIGHT pweight;
RUN ;
```

Where the dummy variable **wap** represents the “working-aged population” and **wp** the “working population” for instance.

Program

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

```
%LET pi = uk99 us00;
```

```
%MACRO gender;
```

```
  %LET I = 1 ;
```

```
  %DO %UNTIL (&I > 2);
```

```
    %LET country = %SCAN(&pi,&i) ;
```

```
    DATA ex4          (DROP=pgwage page pself) ;
```

```
    SET &&&country.p (KEEP=pweight psex page pgwage pself);
```

```
    totinc = pgwage+pself ;
```

```
    IF ((24<page<61) AND (pself=>0)) THEN a = 1;
```

```
    IF ((24<page<61) AND (pself=>0) AND (totinc>0)) THEN b = 1;
```

```
  RUN;
```

```
  TITLE "Country - %scan(&pi,&i)" ;
```

```
  PROC MEANS DATA=ex4 MEAN SUMWGT MISSING;
```

```
    CLASS      a b psex;
```

```
    TYPES      a*psex b*psex ;
```

```
    VAR      totinc ;
```

```
    WEIGHT pweight;
```

```
  RUN ;
```

```
  %LET I = %EVAL(&i+1) ;
```

```
%END ;
```

```
%MEND gender;
```

```
%gender
```

Results

	<i>UK99</i>		<i>US00</i>	
	Males	Females	Males	Females
Percentage of individuals with positive earnings in working-aged population	<i>11,359,337 / 13,906,081 = 81.7%</i>	<i>9,429,261 / 13,906,112 = 67.8%</i>	<i>60,066,557 / 66,098,374 = 90.9%</i>	<i>53,819,001 / 68,821,548 = 78.2%</i>
Average total earnings (working-aged population)	<i>£ 19,101</i>	<i>£ 8,543</i>	<i>\$ 44,133</i>	<i>\$ 22,062</i>
Average total earnings (working population)	<i>£ 23,384</i>	<i>£ 12,599</i>	<i>\$ 48,565</i>	<i>\$ 28,212</i>
Gender earnings gap (working-aged population)	<i>19,101 / 8,543 = 2.24</i>		<i>44,133 / 22,062 = 2.00</i>	
Gender earnings gap (working population)	<i>23,384 / 12,599 = 1.86</i>		<i>48,565 / 28,212 = 1.72</i>	

Comments

- Please note that this exercise examines only individual earnings. Allocation of earnings (and other income) among household members is not considered. Income gender analysis becomes much more demanding and requires many more assumptions about the allocation of total household income when other household members are present.
- It is interesting to note that the earnings gap is lower when the employment rate is higher. While a two-country statistical snapshot does not provide enough information to draw conclusions, these types of summary statistics often provide researchers with new questions to investigate.