PART II: Analysis of labour force data

11. Introduction
12. Descriptive analysis of survey variables
13. Analysis of supply and demand of labour
14. Size and composition of the population
15. Labour force participation of men and women
16. Employment-population ratio
17. Unemployment and its duration
18. Youth and school-to-work transition
19. Hours of work, underemployment and labour slack
20. Branch of economic activity and productivity
21. Occupational structure and segregation
22. SE and informal employment
23. Income from employment and earning differentials
24. Low pay and working poor
11. INTRODUCTION

The data collected in a labour force survey are used to:

A. Derive the variables introduced in Part I;

B. Produce classifications such as those relating to occupation, industry and SE; and

C. Describe the socio-demographic characteristics of the population, including age, sex, geographical location, educational attainment and sometimes migration status.

The initial objective is to produce a survey report consisting largely of descriptive analysis of these variables and their inter-relationships. The analysis is usually in the form of tables, diagrams and related statistics. The actual production of these is easily done using one of the many available statistical software packages. The challenge is to understand and use the tables, statistics and diagrams produced to write the survey report. This part of the Guide Book discusses approaches to meet this challenge. It also addresses the key issue of the quality of the survey results in terms of the errors of the estimates and the coherence of these estimates both internally and externally in comparison with other sources.

12. DESCRIPTIVE ANALYSIS OF SURVEY VARIABLES

12.1 ANALYSIS OF A SINGLE VARIABLE

12.1.1 Single population: full distribution

The intention is to describe the observed pattern in a variable in the form of its distribution. This can be: (i) a table of counts of observations having the different distinct values of the variable (a frequency table); (ii) a diagram such as a bar chart or pie chart; and (iii) a distribution function. Note that if the variable is ordinal, the table, bar chart or pie chart

<table>
<thead>
<tr>
<th>Industry</th>
<th>Number ('000)</th>
<th>% share</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agriculture</td>
<td>799</td>
<td>5.9</td>
</tr>
<tr>
<td>Mining</td>
<td>333</td>
<td>2.4</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1,988</td>
<td>14.6</td>
</tr>
<tr>
<td>Utilities</td>
<td>95</td>
<td>0.7</td>
</tr>
<tr>
<td>Construction</td>
<td>1,112</td>
<td>8.2</td>
</tr>
<tr>
<td>Trade</td>
<td>3,156</td>
<td>23.2</td>
</tr>
<tr>
<td>Transport</td>
<td>747</td>
<td>5.5</td>
</tr>
<tr>
<td>Finance</td>
<td>1,667</td>
<td>12.2</td>
</tr>
<tr>
<td>Community and social services</td>
<td>2,564</td>
<td>18.8</td>
</tr>
<tr>
<td>Private households</td>
<td>1,163</td>
<td>8.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13,623</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

*Source:* Quarterly Labour Force Survey, Table C, Quarter 1, Statistics South Africa

27 Recommended tables for analysis of labour force data can be found in

a. UN (2008): Principles and Recommendations for Population and Housing Censuses, Rev. 2, pp. 284–294, UN, New York; and

b. International Household Survey Network, Question Bank Module on labour force surveys.
should retain the order in the values. If the variable is continuous, the bar chart could be replaced by a histogram or a smooth density curve. The analysis is very simple, based on highlighting any key points that can be deduced about the distribution.

Example 1
Table 1 is the distribution of the employed population by industry in South Africa for the first quarter of 2008. It is in the form of both counts (frequencies) and percentages (relative frequencies). In the report, one can conclude that the major industries employing the largest numbers of persons are industry (23.2%), community and social services (18.8%), manufacturing (14.6%) and finance (12.2%).

Figure 1 is a presentation of the same distribution in the form of a bar chart:

Figure 2: Histogram and density curve of an income distribution (artificial data)

From the diagram, we can conclude that most of the employed have incomes in the lowest three income groups, whilst very few receive income in the highest groups.

12.1.2 Single variable:
Using summary statistics

As an alternative to the complete description of the distribution in the form of a table or diagram, its essential characteristics (i.e. its summary statistics) can be used instead. Summary statistics characterise a distribution in terms of:

- Some measure of central location such as mean, median or mode;
- Some measure of other locations such as quartiles, deciles, percentiles;
- Some measure of the spread of the values such as standard deviation and inter-quartile range; or

**FIGURE 1**

**EMPLOYMENT BY INDUSTRY (1ST QTR 2008)**

Source: Quarterly Labour Force Survey, Table C, Quarter 1, Statistics South Africa
Part II: Analysis of labour force data

E. Using data on weekly days worked by employed persons in Switzerland in 2004,29 we derive the following summary statistics:

\[
\text{Mean} = 3.99 \text{ days}; \quad \text{SD} = 1.56 \text{ days}; \\
\text{Skewness} = -0.65
\]

We can deduce that on average, employed persons worked about four days a week, with most persons working for many days and a few who worked for only a few days (negatively skewed).

12.1.3 Comparison across groups

12.1.3.1 Full distributions

More often than not, the analysis actually done in a survey report involves comparing the distribution of a variable for different subgroups of the population or over time. In this case, using percentages instead of actual counts for

> Some measure of the shape of the distribution such as skewness coefficient.

Using summary statistics to describe the distribution of a variable is particularly useful when the variable is continuous.

Example 3

Data on annual unemployment rates of a selection of countries in 200228 yield the following summary statistics of the distribution of annual unemployment rates across countries:

\[
\text{Mean} = 9.5\%; \quad \text{SD} = 6.7\%; \quad \text{Skewness} = 1.3
\]

We can conclude that the mean unemployment rate is 9.5%. The values are spread out widely (SD = 6.7%), with most countries having relatively low values and a few having very large unemployment rates (positively skewed).

Source: Quarterly Labour Force Survey, Table C, Quarter 1, Statistics South Africa

Source: LABORSTA, ILO statistical database.

28 Source: LABORSTA, ILO statistical database.
the distributions is advisable. The analysis comprises describing the essential differences and similarities between the distributions.

Example 4
Table 2 gives the age distributions of the employed population, the unemployed population and the population outside the labour force for South Africa in the first quarter of 2008. The predominant age group for those in the labour force (employed or unemployed) is 25-34 years. For those outside the labour force it is, not surprisingly, the age group 15-24 years as a substantial number of those in this age group would have been in education. It is also interesting to note the different shapes of the distributions. The distributions of the unemployed population and those outside the labour force are lop-sided, with most of the persons in the lower age ranges, 15-34 years, whilst that of the employed population although also lop-sided is more of an upturned U-shape.

12.1.3.2 Using summary statistics
We can use summary statistics to make comparisons between groups or across time either directly or using diagrams such as box plots. Since the value of the standard deviation is affected by the unit of measurement, the preferred measure of dispersion when comparing distributions is the coefficient of variation (CV), which is obtained by dividing the standard deviation by the mean.

Example 5
Table 3 gives the summary statistics for the distributions in Table 2. The employed population has the highest mean age compared to the other two populations, indicating that

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Population (in thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Employed</td>
</tr>
<tr>
<td></td>
<td>Number</td>
</tr>
<tr>
<td>15–24</td>
<td>1,657</td>
</tr>
<tr>
<td>25–34</td>
<td>4,645</td>
</tr>
<tr>
<td>35–44</td>
<td>3,604</td>
</tr>
<tr>
<td>45–54</td>
<td>2,589</td>
</tr>
<tr>
<td>55–64</td>
<td>1,128</td>
</tr>
<tr>
<td>Total</td>
<td>13,623</td>
</tr>
</tbody>
</table>

Source: Quarterly Labour Force Survey, Table 6, Quarter 1, Statistics South Africa

<table>
<thead>
<tr>
<th></th>
<th>Employed population</th>
<th>Unemployed population</th>
<th>Population outside labour force</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age</td>
<td>37.7</td>
<td>30.4</td>
<td>31.3</td>
</tr>
<tr>
<td>St. Dev.</td>
<td>11.4</td>
<td>10.0</td>
<td>14.5</td>
</tr>
<tr>
<td>CV</td>
<td>0.30</td>
<td>0.33</td>
<td>0.46</td>
</tr>
<tr>
<td>Skewness</td>
<td>-0.3</td>
<td>-0.9</td>
<td>-0.9</td>
</tr>
</tbody>
</table>

Source: Based on the Quarterly Labour Force Survey, Table 6, Quarter 1, Statistics South Africa
persons who are unemployed or outside the labour force tend to be younger. The ages of those outside the labour force are more spread out in comparison with the others (i.e. they have a CV of 0.46 compared to that of 0.3 for the two other groups). All three distributions are negatively skewed, but less so for the employed population. The extent of skewness in the unemployed population and that of the population outside the labour force again confirms that persons in these populations tend to be at the younger ages.

12.1.4 Comparisons across time periods

The time element in the data set means that the situation can be examined both in terms of distributions as well as the individual changes that have taken place in the numbers over time (i.e. some form of basic time series analysis).

Example 6

The distributions of employment by industry in South Africa for the first two quarters of 2008 are given in Table 4. The distributions in the two quarters are quite similar (columns 3 and 5), indicating that the pattern of employment has not changed significantly over the quarters. However, in column 6, we observe a substantial increase in employment in the community and services industry (up by 71,000) as compared to the equally substantial decrease in employment in the trade industry (51,000) over the two quarters. There was a slight increase in overall employment between the quarters (106,000).

12.2 ANALYSIS OF TWO OR MORE VARIABLES

12.2.1 Bivariate distribution

The distribution of two variables, referred to as the bivariate distribution of the variables, is displayed in a two-way table for nominal/ordinal variables. The distinct values of one variable are placed as rows (and called the row variable) whilst those of the other variable are placed as columns (the column variable). The number of observations having the same row and column values is placed in the corresponding cell, constituting the bivariate distribution. A substantial number of tables used in survey reports are of this type. When one or both variables are continuous with a large number of values, these have to be grouped into intervals to construct the two-way table. Otherwise, the bivariate distribution of two continuous variables is represented by a distribution function of two variables.

Example 7

Table 5 shows the bivariate distribution of occupation and SE in the US in 2006, with occupation as the row variable and SE as the column variable. The cell entries are the number of employed persons having the same pair of values of occupation and SE, i.e. the frequency with which the pair occurs in the dataset.

From Table 5, we can see that:

- 19,961 employed persons are both clerks and employees, i.e. they have the value ‘clerks’ for occupation and the value ‘employees’ for SE. Thus, they have the same observation (clerks, employees).

- Not much can be said about the pattern shown in the table, except that consistently for all occupation groups, employees are in the majority. In short, the bivariate distribution does not easily portray any relationship that may exist between the variables.

- The row totals show the marginal distribution of occupation and the column totals show the marginal distribution of SE. We can conclude that the most frequently occurring occupation is that of ‘service workers & shop and market sales workers’, and that of SE is ‘employees’.

The most important reason for analysing several variables together is to examine and describe any relationship that may exist between the variables. For example, they may be associated with each other in an interdependent way, meaning the values of one are affected by those of the other variable(s) and vice versa. A par-
### Table 4
**Employment by Industry**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Number ('000)</td>
<td>Number ('000)</td>
<td>Change ('000)</td>
</tr>
<tr>
<td>Agriculture</td>
<td>799</td>
<td>790</td>
<td>-9</td>
</tr>
<tr>
<td>Mining</td>
<td>333</td>
<td>346</td>
<td>13</td>
</tr>
<tr>
<td>Manufacturing</td>
<td>1,988</td>
<td>1,968</td>
<td>-20</td>
</tr>
<tr>
<td>Utilities</td>
<td>95</td>
<td>97</td>
<td>2</td>
</tr>
<tr>
<td>Construction</td>
<td>1,112</td>
<td>1,138</td>
<td>26</td>
</tr>
<tr>
<td>Trade</td>
<td>3,156</td>
<td>3,105</td>
<td>-51</td>
</tr>
<tr>
<td>Transport</td>
<td>747</td>
<td>774</td>
<td>27</td>
</tr>
<tr>
<td>Finance</td>
<td>1,667</td>
<td>1,687</td>
<td>20</td>
</tr>
<tr>
<td>Community and social services</td>
<td>2,564</td>
<td>2,635</td>
<td>71</td>
</tr>
<tr>
<td>Private households</td>
<td>1,163</td>
<td>1,185</td>
<td>22</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>13,623</strong></td>
<td><strong>13,729</strong></td>
<td><strong>106</strong></td>
</tr>
</tbody>
</table>

*Source: Quarterly Labour Force Survey, Table C, Quarter 1, Statistics South Africa*

### Table 5
**Employed Population by Occupation and SE (‘000s) (US, 2006)**

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Employers &amp; Own-account workers</th>
<th>Employees</th>
<th>Unpaid family workers</th>
<th>All workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislators, senior officials &amp; managers</td>
<td>2,485</td>
<td>19,168</td>
<td>6</td>
<td>21,659</td>
</tr>
<tr>
<td>Profs., Technicians &amp; Assoc. profs.</td>
<td>1,664</td>
<td>28,160</td>
<td>2</td>
<td>29,826</td>
</tr>
<tr>
<td>Clerks</td>
<td>356</td>
<td>19,961</td>
<td>38</td>
<td>20,355</td>
</tr>
<tr>
<td>Service workers &amp; shop and market sales workers</td>
<td>3,625</td>
<td>39,095</td>
<td>30</td>
<td>42,750</td>
</tr>
<tr>
<td>Skilled agricultural and fishery workers</td>
<td>63</td>
<td>984</td>
<td>16</td>
<td>1,063</td>
</tr>
<tr>
<td>Craft &amp; related trade workers; Plant &amp; machine operators</td>
<td>2,686</td>
<td>32,425</td>
<td>14</td>
<td>35,125</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>10,879</strong></td>
<td><strong>139,793</strong></td>
<td><strong>106</strong></td>
<td><strong>150,778</strong></td>
</tr>
</tbody>
</table>

*Source: LABORSTA, Web database, ILO*
### TABLE 6
CONDITIONAL DISTRIBUTIONS OF SE GIVEN OCCUPATION (US, 2006)

<table>
<thead>
<tr>
<th>Occupation</th>
<th>SE</th>
<th>Employers &amp; Own-account workers</th>
<th>Employees</th>
<th>Unpaid family workers</th>
<th>%</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislators, senior officials &amp; managers</td>
<td></td>
<td>11.5</td>
<td>88.5</td>
<td>0.0</td>
<td>100.0</td>
<td>21,659</td>
</tr>
<tr>
<td>Profs., Technicians &amp; Assoc. profs.</td>
<td></td>
<td>5.6</td>
<td>94.4</td>
<td>0.0</td>
<td>100.0</td>
<td>29,826</td>
</tr>
<tr>
<td>Clerks</td>
<td></td>
<td>1.7</td>
<td>98.1</td>
<td>0.2</td>
<td>100.0</td>
<td>20,355</td>
</tr>
<tr>
<td>Service workers &amp; shop and market sales workers</td>
<td></td>
<td>8.5</td>
<td>91.5</td>
<td>0.1</td>
<td>100.0</td>
<td>42,750</td>
</tr>
<tr>
<td>Skilled agricultural and fishery workers</td>
<td></td>
<td>5.9</td>
<td>92.6</td>
<td>1.5</td>
<td>100.0</td>
<td>1,063</td>
</tr>
<tr>
<td>Craft &amp; related trade workers; Plant &amp; machine operators</td>
<td></td>
<td>7.6</td>
<td>92.3</td>
<td>0.0</td>
<td>100.0</td>
<td>35,125</td>
</tr>
<tr>
<td><strong>All workers</strong></td>
<td></td>
<td><strong>7.2</strong></td>
<td><strong>92.7</strong></td>
<td><strong>0.1</strong></td>
<td><strong>100.0</strong></td>
<td><strong>150,778</strong></td>
</tr>
</tbody>
</table>

Source: LABORSTA, Web database, ILO

### TABLE 7
CONDITIONAL DISTRIBUTIONS OF OCCUPATION GIVEN SE (US, 2006)

<table>
<thead>
<tr>
<th>Occupation</th>
<th>SE</th>
<th>Employers &amp; Own-account workers</th>
<th>Employees</th>
<th>Unpaid family workers</th>
<th>All workers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Legislators, senior officials &amp; managers</td>
<td></td>
<td>22.8</td>
<td>13.7</td>
<td>5.7</td>
<td>14.4</td>
</tr>
<tr>
<td>Profs., Technicians &amp; Assoc. profs.</td>
<td></td>
<td>15.3</td>
<td>20.1</td>
<td>1.9</td>
<td>19.8</td>
</tr>
<tr>
<td>Clerks</td>
<td></td>
<td>3.3</td>
<td>14.3</td>
<td>35.8</td>
<td>13.5</td>
</tr>
<tr>
<td>Service workers &amp; shop and market sales workers</td>
<td></td>
<td>33.3</td>
<td>28.0</td>
<td>28.3</td>
<td>28.4</td>
</tr>
<tr>
<td>Skilled agricultural and fishery workers</td>
<td></td>
<td>0.6</td>
<td>0.7</td>
<td>15.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Craft &amp; related trade workers; Plant &amp; machine operators</td>
<td></td>
<td>24.7</td>
<td>23.2</td>
<td>13.2</td>
<td>23.3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
<td><strong>100.0</strong></td>
</tr>
<tr>
<td><strong>Numbers</strong></td>
<td></td>
<td><strong>10,879</strong></td>
<td><strong>139,793</strong></td>
<td></td>
<td><strong>106</strong></td>
</tr>
</tbody>
</table>

Source: LABORSTA, Web database, ILO
A particular type of relationship is when the values of one variable are expected to be affected by changes in the values of the other variable(s), but not the other way round. This is referred to as a dependent relationship. The first variable is called the dependent or endogenous variable and the latter are the independent or exogenous variable(s). Again, the variables may have no relationship with each other. They are then said to be independent.

12.2.2 Using the full conditional distribution

One way of investigating inter-relationships between variables is to compare the conditional distributions of one of the variables for given values of the other variable(s). For example, to study the relationship between occupation and SE, the conditional distributions of occupation for given values of SE could be compared. The conditional distribution of occupation for a given value of SE, say for employees, is that obtained looking at the distribution of occupation amongst the sub-population of employees (i.e. the third column of Table 5). Thus, each category of the SE variable will have its own distribution of occupation.

A relationship exists between the variables if two or more of the conditional distributions are different. If there are no differences, then the conclusion is that the two variables are not related. The survey report should also draw attention to any of the conditional distributions that looks interesting, in the same way as in the previous section on single-variable analysis.

Example 8
The conditional distributions of SE for given categories of occupation are given in Table 6. Note the differences amongst the six conditional distributions. For example, in the conditional distribution of SE for the occupational group ‘Legislators, senior officials & managers’, the share of ‘employers’ is 11.5%, whilst for the occupational group ‘clerks’, the share of ‘employers’ is only 1.7%. Thus, the conclusion is that occupation and SE are interdependent.

We could also have used the conditional distributions of occupation for values of SE (Table 7). Note that the three distributions are different. In the conditional distribution of occupation for the SE group ‘Employers & Own-account workers’, the share of ‘clerks’ is only 3.3%, whilst for the SE group ‘Unpaid family workers’, the share of ‘clerks’ is 35.8%. So, once again the conclusion is that occupation and SE are interdependent.

12.2.3 Using diagrams

Diagrams can also be used to portray the conditional distributions and identify relationships. When both variables are categorical (e.g. occupation and SE), we represent the conditional distributions using a component bar chart in which the bar is divided up into sections representing the different frequencies in the distribution. The bars can then be compared to establish the existence of a relationship. If both variables are continuous, the bivariate distribution can be illustrated using a scatter diagram (see Example 10). In it, the pairs of values of the variables for each observation is plotted on an x-y- graph.

Example 9
The differences between the three bars are clear, establishing that a relationship exists between occupation and SE, i.e. as SE changes, the distribution of occupation also changes. In particular, the shares of clerks in the three bars are markedly different, the largest being amongst the ‘Other SE’ category. For the group ‘Legislators, etc.’, the order is reversed, with the share largest for the ‘Employers’ category.

Example 10
The scatter diagram in Figure 4 below is derived from the observations of two variables X and Y. The points are closely clustered together with the values of both variables tending to move together in the same direction. We can therefore conclude that the variables are inter-related.

12.2.4 Using summary statistics

Instead of using the full conditional distributions, the comparisons can also be made using
**Figure 3**

Conditional Distributions of Occupation Given SE (US, 2006)

**Figure 4**

Scatter Diagram of Variables X and Y (Artificial Data)
their summary statistics. This method is particularly useful when one of the variables is continuous. A difference in any of these statistics between two or more of the distributions indicates that a relationship exists.

Example 11
Table 8 presents the summary statistics of the conditional distributions of income for a given industrial sector. We see that:

- Means (median) are different – ‘Services’ has the largest mean income (and median income as well) amongst the four sectors;
- CVs are also different, showing a greater spread for ‘Agric’ & ‘Others’; and
- The distributions of ‘Agric’ and ‘Others’ are more skewed than for each of the other two sectors, as shown by their skewness coefficients.

From subject-matter knowledge, one would expect a person’s income to be affected by their industrial sector rather than the other way round. We can therefore conclude that there is a dependent relationship of income on industrial sector, with persons working in the services sector tending to have larger incomes than those in other sectors. Not surprisingly, all the income distributions are positively skewed.

12.2.5 Issues in using conditional distributions

12.2.5.1 Interpreting relationship

In interpreting the results of the analysis, an assumption has to be made about the type of relationship (interdependent or dependent). The choice comes from subject-matter knowledge, hypothesis of interest, research purpose, logical connections, etc.:

A. Logically, in the relationship between the weight and height of workers the former depends on the latter, i.e. it is a dependent relationship with weight as dependent variable and height as the independent variable.

B. From subject-matter knowledge, age and sex tend to influence the level of income, so income has a dependent relationship on

<table>
<thead>
<tr>
<th>TABLE 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMMARY STATISTICS FOR CONDITIONAL DISTRIBUTIONS OF INCOME BY INDUSTRIAL SECTOR</td>
</tr>
<tr>
<td>Agric</td>
</tr>
<tr>
<td>Mean</td>
</tr>
<tr>
<td>Median</td>
</tr>
<tr>
<td>Standard Deviation</td>
</tr>
<tr>
<td>CV</td>
</tr>
<tr>
<td>Skewness</td>
</tr>
<tr>
<td>Minimum</td>
</tr>
<tr>
<td>Maximum</td>
</tr>
<tr>
<td>1st Quartile</td>
</tr>
<tr>
<td>3rd Quartile</td>
</tr>
</tbody>
</table>
In the case of occupations and SE, one can only assume that the relationship, if it exists, has to be interdependent.

The methods described above are directed at establishing the existence of a relationship on the assumption that the type is already known.

12.2.5.2 Choosing variable of conditional distribution

Also, in carrying out analysis of this type, it is necessary to identify which of the two variables is to be used as the conditioning variable. If the intention is to explore only an interdependent relationship, it does not matter which variable is used as the conditioning variable. However, one set of these conditional distributions could be more interesting depending on the hypothesis of interest in the analysis. The choice should then clearly respect this interest. For example, in gender analysis it is more relevant to look at the conditional distributions of sex for a given variable (e.g. occupation) rather than the conditional distributions of occupation given sex. The latter set may look quite similar across the sexes and so fail to identify key sex differences in the individual occupation groups.

Example 12

The conditional distributions of occupation given SE portrayed in Table 7 are more interesting to analyse than those of SE given occupation in Table 6. There are easily perceptible differences in the former whilst the latter look rather similar.

In situations when one variable is continuous and the other is not, it is preferable to produce conditional distributions of the continuous variable for given values of the other variable. For example, conditional distributions of income for given levels of occupation. If both variables are continuous, then one alternative is to convert one of them to a grouped variable.

12.2.5.3 Possible misinterpretation

When using conditional distributions to explain relationships, it is important to present and describe all such distributions. The use of one to the exclusion of others can lead to fallacious conclusions about the relationship.

Example 13

In the well-known example of the relationship between being drunk and driving, the statement that 30% of road accidents is caused by drunk drivers can be misused to conclude that it is better to drive whilst drunk, as sober drivers are responsible for 70% of road accidents. This is as a result of using only the conditional distribution of the state of soberness of drivers given road accidents, ignoring the other conditional distribution of state of soberness of drivers given safe driving (no road accidents). The latter could have shown, for example, that only 1% of safe driving is done by drunk drivers compared to 99% done by sober drivers. The correct conclusion would then be that drunk drivers are responsible for a highly disproportionate percentage of road accidents compared to their percentage of safe driving, i.e. compared to their percentage amongst all drivers.

12.2.6 Using measures of interdependence

These measures are statistics derived from the joint distribution of two (or more) variables which assess the strength (sometimes also the direction) of the relationship between the variables.

12.2.6.1 Both variables continuous

(\textit{Correlation coefficient} (r))

The correlation coefficient (r) is a measure of (linear) interdependence between the variables. Its values range from -1 to +1. The magnitude shows the strength of the relationship:

\begin{itemize}
  \item Value 0 means no linear relationship whatsoever;
  \item Value 1 means perfect linear relationship (i.e. an exact mathematical relationship, e.g. Circumference of circle = 2 times diameter); and
\end{itemize}
The measure of interdependence is the Chi-square statistics. Its value is always positive and the larger it is, the stronger is the relationship. A value of 0 indicates complete independence between the variables. The value should actually be compared to some critical cut-off determined by theory in order to assess if a relationship exists statistically. This critical value increases with the product of \((r – 1)\) and \((c – 1)\), where \(r\) and \(c\) are respectively the number of rows and number of columns of the two-way distribution table. Thus, the bigger the two-way table the higher should be the Chi-square value for it to indicate that a relationship exists. If one of the variables is continuous or discrete, its values should be grouped into intervals to carry out this analysis.

**Example 15**

We can examine the existence of a relationship between occupation and SE using data in Table 5. The Chi-square value from the bivariate table is 2,103.65. The table has six rows and three columns, so the theoretical value to be expected if the variables were independent is 18.31 at the 5% level. The huge difference between these two numbers indicates that the variables are interdependent, and strongly so. This is, of course, the same conclusion we got above using conditional distributions.

12.2.7 Using measures of dependence

12.2.7.1 Both variables continuous (regression analysis)

The measure used to assess (linear) dependence of a continuous variable \(Y\) (e.g. income) on another continuous variable \(X\) (e.g. age) is the regression coefficient. The method for establishing the linear dependent relationship of \(Y\) on \(X\) is called the ‘Simple linear regression of \(Y\) on \(X\)’. It can be formulated in the form \(Y = mX + c\), where \(m\) and \(c\) are unknown constants to be estimated from the data. The equation \(Y = mX + c\) is called the regression equation and the graph of \(Y = mX + c\) is a line called the regression line. ‘\(m\)’ is the regression coefficient.

---

30 Actually, from experience, this linear relationship exists only up to a certain age, after which the curve flattens out.
The sign of \( m \) gives the nature of the dependence of \( Y \) on \( X \), whilst its magnitude reflects the extent to which the value of \( Y \) changes as \( X \) changes. If \( m = 0 \), \( Y \) does not depend linearly on \( X \). If \( m \) is positive, as \( X \) increases \( Y \) also increases. However, if \( m \) is negative, as \( X \) increases, \( Y \) decreases. The magnitude of \( m \) is the extent to which the value of \( Y \) changes for every unit increase in the value of \( X \). Note in fact that the regression equation applies to the (conditional) mean value of \( Y \) for a given \( X \) value. For example, the equation indicates that the mean of the conditional distribution of income given age changes by \( m \) for each unit increase in the value of age. So the correct vector of the above statements is that \( Y \) tends to increase (or decrease) as \( X \) increases.

The assessment of the strength of the relationship is done using the magnitude of Pearson’s \( r \) coefficient in the form \( 100 \times r^2 \). We say the ‘Regression’ explains \( (100 \times r^2) \% \) of the variation in the dependent variable. The higher the value, the more confidence can be given to using the regression for prediction. However, care should be exercised in using the regression equation for prediction outside the range of the original data for \( X \) even if the strength of the relationship is high. For example, using the linear regression of income on age to predict income for age values larger than those in the original set could be misleading, as the income curve flattens out after a certain age.

**Example 16**

Consider the data points represented in the scatter diagram in Example 10 above (i.e. Figure 4). The data points do seem to have a linear trend and so it seems reasonable to fit a linear regression of \( Y \) on \( X \). The results are as below:

- The regression coefficient, \( m = 3.0 \); the intercept, \( c = 1.5 \). Therefore, the regression equation is:
  \[ Y = 3X + 1.5. \]

- The correlation coefficient \( r = 0.99 \) and so the regression equation explains \( (100 \times r^2) \% \) of the variation in the values of the variable \( Y \). This value shows the linear relationship to be very strong, and so using the equation for prediction can be done with confidence.

12.2.7.2 One or more variables ordinal/nominal

There are methods that can be used

A. if either the dependent or independent variable is ordinal or nominal; and

B. if the relationship is linear or non-linear

12.2.8 Analysis of three or more variables

For the analysis of three or more variables, the above methods can be used to study the bivariate conditional distribution of two of the variables for given values of the third. The easiest method for analysis involving more than two variables, say three variables, is to produce and analyse two-way tables of two of the variables for each value of the third variable. Several scenarios are possible:

- If the conditional distributions in each of the two-way tables are the same within each table and between the tables, the three variables are independent.

- If one or more of the conditional distributions of the two variables in each of the two-way tables are different, but this difference is the same for all of the tables, then the two variables are interdependent but they are jointly independent of the third.

- If one or more of the conditional distributions of the two variables in each of the two-way tables are different and the pattern of differences changes between the tables, then the three variables are interdependent.

When one variable is continuous, a two-way
table of the other two variables containing summary statistics of the continuous variable in the cells can be used. For example, a two-way table of occupation and SE with summary statistics of the income distribution as cell entries can be used to explain the relationship between income, occupation and SE. Differences in the summary statistics in some of the cells would point to the existence of a relationship.

Example 17
Table 9 below presents the conditional distributions of Occupation given SE and Sex in the US in 2007. These can be used as follows to analyse the relationship between sex, occupation and SE in the US in 2007.

The conditional distributions of occupations given SE in Table 9 are different both within and between the two two-way tables for males and females: for example, amongst male employees, legislators constitute only 14.5% as compared to 25.5% for male employers. Consequently, we can conclude that Occupation and SE are interdependent for the male employed population. Again, comparing the male distribution of occupations for employers with that for females, we see that whilst only 23.3% in the former are service workers, the corresponding value for the female distribution is 50.5%. Therefore, the conditional distributions are different both within each two-way table and between the two-way tables. We conclude that the three variables Occupation, SE and Sex are interdependent. We can go on to analyse the nature of the relationships using the conditional distributions.

Example 18
The relationship between Occupation, SE and year in the US is analysed as follows, using the

<table>
<thead>
<tr>
<th>TABLE 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONDITIONAL DISTRIBUTIONS OF OCCUPATION GIVEN SE AND SEX</td>
</tr>
<tr>
<td><strong>Males</strong></td>
</tr>
<tr>
<td><strong>Status in Employment</strong></td>
</tr>
<tr>
<td>Occupation</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Leg.</td>
</tr>
<tr>
<td>Profs</td>
</tr>
<tr>
<td>Clerks</td>
</tr>
<tr>
<td>Service W.</td>
</tr>
<tr>
<td>Sk. Agric.</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>Total</td>
</tr>
<tr>
<td><strong>Females</strong></td>
</tr>
<tr>
<td><strong>Status in Employment</strong></td>
</tr>
<tr>
<td>Occupation</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Leg.</td>
</tr>
<tr>
<td>Profs</td>
</tr>
<tr>
<td>Clerks</td>
</tr>
<tr>
<td>Service W.</td>
</tr>
<tr>
<td>Sk. Agric.</td>
</tr>
<tr>
<td>Others</td>
</tr>
<tr>
<td>Total</td>
</tr>
</tbody>
</table>

Source: LABORSTA, Web database, ILO
conditional distributions of Occupation given SE and Year in Table 10.

In each of the years, the conditional distributions of occupation given SE are different across values of SE. These differences are, however, roughly the same across the years, i.e. the distributions in 2007 are the same as those in 2006. Thus, we can conclude that whilst there is an interdependent relationship between occupation and SE, these variables are independent of the year. This is not surprising, as the years are too close together for any significant change to have occurred in the distributions of these structural variables.

Example 19
Table 11 presents the summary statistics for the conditional distributions of income for given values of industrial sector and sex.

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Employers</th>
<th>Employees</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leg.</td>
<td>22.6</td>
<td>13.8</td>
<td>4.4</td>
<td>14.4</td>
</tr>
<tr>
<td>Profs</td>
<td>15.6</td>
<td>20.6</td>
<td>3.0</td>
<td>20.2</td>
</tr>
<tr>
<td>Clerks</td>
<td>3.1</td>
<td>14.1</td>
<td>39.3</td>
<td>13.3</td>
</tr>
<tr>
<td>Service W.</td>
<td>33.6</td>
<td>27.9</td>
<td>28.9</td>
<td>28.2</td>
</tr>
<tr>
<td>Sk. Agric.</td>
<td>0.6</td>
<td>0.7</td>
<td>12.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Others</td>
<td>24.5</td>
<td>22.9</td>
<td>11.1</td>
<td>22.9</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Occupation</th>
<th>Employers</th>
<th>Employees</th>
<th>Others</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leg.</td>
<td>22.8</td>
<td>13.7</td>
<td>5.7</td>
<td>14.3</td>
</tr>
<tr>
<td>Profs</td>
<td>15.3</td>
<td>20.1</td>
<td>1.9</td>
<td>19.7</td>
</tr>
<tr>
<td>Clerks</td>
<td>3.3</td>
<td>14.3</td>
<td>35.8</td>
<td>13.4</td>
</tr>
<tr>
<td>Service W.</td>
<td>33.3</td>
<td>28.0</td>
<td>28.3</td>
<td>28.2</td>
</tr>
<tr>
<td>Sk. Agric.</td>
<td>0.6</td>
<td>0.7</td>
<td>15.1</td>
<td>0.7</td>
</tr>
<tr>
<td>Others</td>
<td>24.7</td>
<td>23.2</td>
<td>13.2</td>
<td>23.2</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Source: LABORSTA, Web database, ILO
### Table 11
**Summary Statistics for Conditional Distributions of Income by Industry and Sex**

<table>
<thead>
<tr>
<th>Sex</th>
<th>Summary statistics</th>
<th>Industrial sector</th>
<th>Agriculture</th>
<th>Industry</th>
<th>Services</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td></td>
<td>1,510.0</td>
<td>2,543.0</td>
<td>7,182.3</td>
</tr>
<tr>
<td></td>
<td>CV</td>
<td></td>
<td>1.32</td>
<td>1.41</td>
<td>1.92</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td></td>
<td>5.6</td>
<td>1.8</td>
<td>2.3</td>
</tr>
<tr>
<td></td>
<td>% distribution</td>
<td></td>
<td>43.4</td>
<td>32.1</td>
<td>24.5</td>
</tr>
<tr>
<td>Female</td>
<td>Mean</td>
<td></td>
<td>1,821.0</td>
<td>1,213.5</td>
<td>3,484.1</td>
</tr>
<tr>
<td></td>
<td>CV</td>
<td></td>
<td>2.91</td>
<td>1.72</td>
<td>1.12</td>
</tr>
<tr>
<td></td>
<td>Skewness</td>
<td></td>
<td>1.9</td>
<td>2.3</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>% distribution</td>
<td></td>
<td>72.1</td>
<td>10.5</td>
<td>17.4</td>
</tr>
</tbody>
</table>

### 13. Analysis of Supply and Demand of Labour

A convenient framework for the analysis of the labour market is the supply and demand of labour, schematically represented in Figure 5 below. In general, supply and demand is an economic model of price determination in a market. In a labour market, suppliers are individual persons who try to sell their labour at the highest price. On the other hand, demanders of labour are enterprises, which try to fill the jobs they need at the lowest price. The equilibrium price for a certain type of labour is the wage rate.

Labour supply at a given point of time comprises all currently employed persons and unemployed persons currently available for work and seeking work. Labour force surveys are generally recognised as a comprehensive means of data collection on the supply of labour. In a labour force survey, individuals are asked their current economic activity, in particular whether they are working or actively seeking work, in which case they are part of the labour supply, or whether they are doing neither of these (i.e., are engaged only in other activities that are non-economic) in which case they are not part of the supply of labour at that moment.

Labour demand, by contrast, is the sum of all occupied and vacant jobs that enterprises or employers require to conduct their economic activity. Occupied jobs are those currently filled, including those filled by employers or own-account workers themselves. Vacant jobs are unfilled jobs which the employers are actively taking steps to fill immediately or within a reasonable period of time. Establishment surveys in which employers are asked about their currently filled jobs and vacancies generally form a suitable source of data on labour demand.32

In such a system, the distinction between jobs and persons is important. A person may hold more than one job and, likewise, there may be jobs held by no one or held by more than one person. The symmetry between the labour supply and labour demand concepts is also instructive. While unemployment represents the unsatisfied supply of labour, job vacancies represent the unmet demand for labour.

In this Guide Book on data analysis, the focus is on the analysis of labour supply. It is, however, important to keep in mind the overall framework and the need to complement the analysis in certain cases with elements of labour de-

---

mand for a better understanding of the functioning of the labour market.

In what follows, examples are given on basic analysis of labour supply with numerical illustrations using data from selected African and other countries. The range of topics is wide, starting from the size and structure of the population, the labour force participation of men and women, the use of the employment-population ratio, the unemployment rate, the particular aspects of the youth and their transition from school to work, hours of work, the various forms of underemployment and labour slack, the classifications by branch of economic activity, occupation, SE and informal employment, as well as the analysis of earnings, low pay and the working poor.

14 SIZE AND COMPOSITION OF THE POPULATION

14.1 INTRODUCTION

The size and composition of the population is the starting point of the analysis of labour supply. Population constitutes the human capital of the nation and defines its potential labour supply. From an economic point of view, the working population is a factor of production and its aptitude and skill level contributes to the productivity of the national economy. From a social point of view, different categories of the population form social groups of particular concern and meeting their needs are major challenges faced by public institutions and society at large.

14.2 AGE PYRAMID

The current structure of the population and to some extent its past evolution and future trend can be examined with the help of the population age pyramid. It shows the size distribution of the age categories of the population for men and women, separately. The pyramid is constructed as back-to-back horizontal bar graphs. The left bar graph shows the age structure of men and the right bar graph that of women. The age structure is ordered from bottom to top, with lower age groups at the bottom and the higher age groups at the top. Because there are generally more young per-
sons than older people and about the same number of men and women, the diagram typically takes the form of a symmetric pyramid. Figures 6 and 7 show the age pyramid of Tanzania by single year for 2010, and those of the world population in 2010 and three other selected countries for comparison: Nigeria, Germany and Iran.

In contrast with the Tanzania age pyramid, the world age pyramid, top left portion of the display, has almost a round-belly shape, the sign of a stationary population pyramid, characterised by a combination of low fertility and low mortality. The younger population (15-29 years) is about the same size as the children population (0-14 years), 1.772 billion young people against 1.861 billion children. The adult population (30 years and over) is almost the same size as the younger population below 30 years of age.

The age pyramid of Nigeria has a closely pyramidal shape, very similar to that of Tanzania. Virtually every lower age group has a larger population than the next higher age group. This type of age pyramid reflects a population with a high birth rate, a high death rate and a short life expectancy, a typical pattern in a developing country. The youth population in Nigeria is 28.1% of total population, similar to the percentage in Tanzania (27.6%) and somewhat higher than the 25.7% world benchmark.

The age pyramid of Germany on the bottom left of the display is top heavy and is not really pyramidal. It has the shape for a typical ageing population in which the top part of the age pyramid is dominant and the bottom part is actually upside down, with a higher number of people in each age group than the next lower age group. The youth population in Germany is 17.2% of total population, significantly lower than the 25.7% world benchmark.

The age pyramid of Iran exhibits a particular shape, known as a youth bulge. The youth
Part II: Analysis of labour force data

The population 15-29 years old is abnormally large relative to the population in the lower age groups (0-14 years old). This youth bulge is the result of the extraordinarily high fertility experienced during the late 1970s and continuing in the 1980s, creating a large cohort of youth now 20-29 years old. The changing pattern of fertility and its sharp and steady fall in the last two decades can be observed in the narrow base of the age pyramid, below the youth bulge. The youth population in Iran is 33.9% of total population, significantly higher than the 25.7% world benchmark, reflecting the youth bulge.

14.3 Dependency Ratios

A useful summary measure to analyse the age structure of a population is the dependency ratio. It is a measure showing the number of dependents (children aged 0 to 14 and the older population aged 65 and over) to the core working age population (15-64 years):

\[
\text{Dependency ratio} = \frac{\text{Children(0 – 14 years) + Aged(65 years+)}}{\text{WorkingAge(15 – 64 years)}}
\]

The dependency ratio may be interpreted as the number of dependents that a worker on average must provide for in the society. The higher the ratio, the higher is the burden on those working. The dependency ratio may be decomposed into two parts, one showing:

\[
\text{Child dependency ratio} = \frac{\text{Children(0 – 14 years)}}{\text{WorkingAge(15 – 64 years)}}
\]

and the other:

\[
\text{Aged dependency ratio} = \frac{\text{Aged(65 years+)}}{\text{WorkingAge(15 – 64 years)}}
\]

Figure 8 compares the dependency ratio and its decomposition for the world population in 2010 and for the four countries mentioned earlier. It shows that the dependency ratios in Tanzania...
and Nigeria are above the world average, which is about 50%, i.e. for every dependent person there is on average two working age persons. In Tanzania and Nigeria, for every dependent person there is just a little more than one working age person.33

Figure 8 also shows that the majority of dependents are children (green bars), except in Germany where the majority of dependents are elderly people (yellow bars). It also shows that the lowest dependency ratio is in Iran (about 40%), reflecting the so-called “demographic window” defined as that period of time in a nation’s demographic evolution, lasting about 30-40 years, when the proportion of children under 15 years falls below 30% and the proportion of 65 years old is still below 15%.

15. LABOUR FORCE PARTICIPATION OF MEN AND WOMEN

15.1 LABOUR FORCE FRAMEWORK

The labour force or the EAP refers to all persons of either sex who furnish the supply of labour for the production of economic goods and services as defined by the UN systems of national accounts and balances during a specified time-reference period.34 The labour force is the sum of the employed and the unemployed. The population not economically active is generally classified according to the reason for inactivity.

The minimum age limit for measuring the labour force is not specified in the international standards, but it is recommended that the data should

---

Part II: Analysis of labour force data

National LFPRs and highlights of the data are published every two years as part of the ILO Key Indicators of the Labour Market (KILM).\textsuperscript{35} Annual estimates and projections spanning the period 1980 to 2020 are also available as part of the database of the ILO Department of Statistics.\textsuperscript{36}

In Figure 9, the global LFPR in 2010 by sex and age group is graphically presented for illustration.

Like most national rates, the world’s LFPR has an inverted-U shape, more pronounced for men than for women. The male curve is above the female curve, reflecting the higher LFPR of men at all age groups. For each sex, the curve increases at low ages as young people leave school and enter the labour market, reaches a peak in the age group 35-39 years for men and 40-49 years for women, before decreasing, slowly for women and more sharply for men, as people leave and retire from the labour market at older ages.

\textbf{15.2 LABOUR FORCE PARTICIPATION RATE (LFPR)}

The LFPR is an indicator of the level of labour market activity. It reflects the extent to which a country’s working age population is economically active. It is defined as the ratio of the labour force to the working age population expressed in percentage terms:

\[
\text{LFPR} = 100 \times \frac{\text{Labourforce}}{\text{WorkingAgePopulation}}
\]

The breakdown of the LFPR by sex and age group gives a profile of the distribution of the EAP within a country. As the international standards do not refer to a maximum age limit for the measurement of the labour force, in principle any person of working age (15 years and older) could be economically active.

\textbf{FIGURE 9}

\textbf{LABOUR FORCE PARTICIPATION RATE (WORLD, 2010)}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9}
\end{figure}

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|c|c|c|c|c|c|c|c|c|}
\hline
\hline
\textbf{Men} & 100 & 90 & 80 & 70 & 60 & 50 & 40 & 30 & 20 & 10 & 0 \\
\hline
\textbf{Women} & 100 & 90 & 80 & 70 & 60 & 50 & 40 & 30 & 20 & 10 & 0 \\
\hline
\end{tabular}
\caption{Labour Force Participation Rate (World, 2010)}
\end{table}

The shape of the LFPR by sex and age group varies somewhat among countries. As shown in Figure 10, in Tanzania (Zanzibar) the LFPRs both for men and women follow similar patterns as the world average, but the male and female differences are much less accentuated. The female LFPR is almost equal to the male rate for all age groups except at top ages, reflecting the dominance of agriculture in the economy and the limited coverage of social security for older people.

The curve also shows that most people, both men and women, are in the labour force after age 15, and remain so at all ages except for older women, who, by 60-64 years, are mostly out of the labour force. By contrast, for the world as a whole, the age at which most young people are in the labour force is for men between 15 and 19 years and for women between 20 and 24 years. Similarly, the world average age at which most people are out of the labour force is for men between 60 and 64 years and for women between 50 and 54 years.

The comparison with the US pattern is also instructive. As shown in Figure 11, the shape of the LFPR among men in the US follows the general inverted U-pattern, but that of women is more like an M-pattern, with two peaks, one reflecting the age at which women start leaving the labour market for reasons of marriage and child bearing (25-29 years) and the other when women (45-49 years) return to the labour market, albeit at a slightly lower rate, after children reach school age.

The US diagram also shows that the age at which most young people are in the labour force is between 20 and 24 years for both men and women. The age at which most older people are out of the labour force is between 60 and 64 years, also for both men and women.

15.3 LABOUR FORCE PARTICIPATION AT DIFFERENT LEVELS OF EDUCATION

The close relationship between educational achievement and employment opportunity is widely recognised in most countries. A typical
pattern is presented in Figure 12 below. It shows that the higher an individual’s educational attainment, the more likely the person would be in the labour force. It is instructive to note, however, that while labour force participation is higher among men than women at all levels of education, the gap almost vanishes at the tertiary level of education: women and men with

**FIGURE 11**
LABOUR FORCE PARTICIPATION RATE USA, MARCH 2006

**FIGURE 12**
LABOUR FORCE PARTICIPATION RATE OF MEN AND WOMEN BY EDUCATIONAL ATTAINMENT
university degrees are almost equally likely to participate in economic activity.

For more detailed analysis of the relationship between specific educational programmes and the labour market, further information would be needed on the nature of the programme (type of training and its duration); subsequent experience of participants (both those who completed the programme and those who dropped out); and their likely experience in the absence of the programme.

Another use of the data is to find out the skill levels of the labour force of the country. The larger the proportion of the labour force with secondary and tertiary (university) education, the higher the skill level of the labour force. Figure 13 presents the distribution of the labour force by educational attainment for four African countries: Botswana (2006) for age group 12 years old and over, Namibia (2004) for age group 15+ years, South Africa (2008) for age group 15–64 years and conscripts, and Tanzania (Zanzibar, 2006) for age group 15+ years. In all countries except Tanzania (Zanzibar), secondary education is the dominant level of educational attainment. The proportion of the labour force with tertiary education is highest in South Africa, followed by Namibia.


15.4 LABOUR FORCE PARTICIPATION OF WOMEN BY MARITAL STATUS

The relationship between female labour force participation and marital status is shown in Figure 14. In this example, the LFPR of single women is higher than that of married women (34% versus 24%). However, the highest LFPR is among divorced women, possibly due to their need to work in the absence of a partner. LFPR is lowest among widowed women, reflecting...
Part II: Analysis of labour force data


the age-effect as widowers tend to be older than women in other marital status categories, and older people tend to have lower LFPRs.

15.5 LABOUR FORCE PARTICIPATION OVER TIME

There is a widespread hypothesis that LFPR follows a U-shape pattern in the course of economic development, especially in the case of women. At low levels of development, agriculture is the dominant form of economic activity in which large numbers of men and women are engaged. The labour force participation is therefore high. Over time, economic activity shifts from home-based production to market-oriented activities in different sectors of the economy. Furthermore, increased mechanisation in agriculture reduces employment opportunities in that activity, leading to migration from rural areas to the cities in search for work or higher education, especially among young people. The result is that LFPR decreases over time at the lower levels of development before starting to increase at higher levels of development when industry and services start to become the dominant form of economy activity in country.37

This phenomenon may be examined by observing the national LFPR over a long period of time (if such a consistent time series is available) or by calculating the current LFPR by geographical area ordered according to level of development. The areas at the lowest levels of development should have higher LFPRs than those of areas at middle levels of development. Moreover, these areas should themselves also have a lower LFPR than those of areas at higher levels of development.

16. EMPLOYMENT-POPULATION RATIO

Aggregate employment generally increases with growing population. Therefore, the ratio of employment to the working age population is an important indicator of the ability of the economy to provide employment to its growing

A decline in the employment-population ratio is often regarded as an indicator of economic slowdown and a decline in total employment an indicator of an even more severe economic downturn.

The employment-population ratio is defined in percentage terms as:

\[
\text{Employment-population ratio} = \frac{\text{Number Employed}}{\text{Working Age Population}} \times 100
\]

The working age population is variously defined as the population 15 years old and over, the population 15 to 64 years old, or other more restrictive age intervals as proposed below. It may be calculated separately for men and women and by age group and other variables of interest, such as educational attainment and urban-rural place of residence. The ratio generally changes faster than the LFPR and slower than and in an opposite direction to the unemployment rate.

Two particular uses of the employment-population ratio are given below. One use is for monitoring the performance of the labour market over time by observing the direction of annual change in the employment-population ratio of the prime-age population (ages 25-54), called here the core employment-population ratio. When the economy is growing, this ratio should increase, or at least remain unchanged, reflecting a certain harmony between the growth of the population and employment. Because the proposed indicator is restricted to the working age population 25 to 54 years old, it is not affected by the increased schooling of young people or earlier and lengthier retirement among the elderly, two phenomena often observed in many countries. Also, because the calculation of the ratio does not require data on unemployment, the indicator should be less controversial than the unemployment rate.

Table 12 illustrates the use of the indicator, calculated based on data from the Labour Force Survey of South Africa:

The results show that in the period from 2007 to 2008, employment of prime-age people increased by 361,000, more than the increase in the size of the corresponding population. This net job gain is reflected in the slightly higher core employment-population ratio, which increased from 59.8% in 2007 to 60.7% in 2008.

Where quarterly or monthly labour force surveys are conducted, the proposed indicator if further restricted to the urban areas may also provide a reliable indicator of the performance of the urban labour market within the year. This is because seasonal variations in agriculture and changes in school attendance during the year should have only a limited effect on the urban ratio.

<table>
<thead>
<tr>
<th>South Africa (Labour Force Survey)</th>
<th>2007*</th>
<th>2008</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of employed persons aged 25–54 (‘000)</td>
<td>10,566</td>
<td>10,928</td>
<td>361</td>
</tr>
<tr>
<td>Population aged 25–54 (‘000)</td>
<td>17,684</td>
<td>17,998</td>
<td>314</td>
</tr>
<tr>
<td>Core employment–population ratio</td>
<td>59.8%</td>
<td>60.7%</td>
<td>1.0%</td>
</tr>
</tbody>
</table>

* September
Source: LABORSTA, Web database, ILO

Another use of the employment-population ratio concerns the two tails of the age distribution not covered by the proposed indicator, namely youth aged 15-24 and people aged 55 and over. The variations of the employment-population ratio for these two categories of persons depend on schooling behaviour and the retirement system of the country. In this regard, a useful indicator to monitor is the age at which most young people of that age are employed and another is the age at which most older people of that age have left employment (or perhaps preferably out of the labour force).

These indicators are best calculated on the basis of single-year age data, but can also be estimated using grouped data by linear interpolation as follows:

Age at which most youth of that age are employed =

\[ 17.5 + 5 \left( \frac{50\% - r_{15-19}}{r_{20-24} - r_{15-19}} \right) \]

rounded to the nearest complete year, where r15-19 and r20-24 are the employment-population ratio for young persons aged 15-19 and 20-24, respectively. A similar method may be used for estimating the age at which most older people aged 55 and above have left employment:

Age at which most people aged 55 and above have left employment

\[ = 57.5 + 5 \left( \frac{50\% - r_{55-59}}{r_{60-64} - r_{55-59}} \right) \text{ if } (r_{60-64} \geq 50\%) \]

\[ = 62.5 + 5 \left( \frac{50\% - r_{60-64}}{r_{65-64} - r_{60-64}} \right) \text{ if } (r_{60-64} < 50\%) \]

Table 13 illustrates the results obtained based on grouped data from the Labour Force Survey of South Africa (2007) for males and females separately. The results may be interpreted as follows: women tend to enter employment later than men (at 29 rather than 25) and tend also to leave employment earlier than men (at 54 rather than 61).

Graphically, these indicators correspond to the points of intersection of the employment-population curve (aged 15+) with the horizontal line of 50%.

Similar values may be obtained using the curve of the LFPR instead of the curve of the employment-population ratio. This gives corresponding estimates for entry age in the labour force (men = 22; women = 23) and exit age from the labour force (men = 62; women = 55). The accuracy of the numerical results may be improved using single-year data rather than grouped data. Also, based on appropriate data, other calculations may be done to derive estimates of related concepts such as average age at first job or average age at first entry into the labour force and, similarly, average age at last job or average age at last exit from the labour force.
17. UNEMPLOYMENT AND ITS DURATION

17.1 UNEMPLOYMENT RATE

The unemployment rate is the most commonly used indicator of the labour market. It is defined as the percentage of persons in the labour force who are unemployed:

Unemployment rate = \[
\frac{\text{Number Unemployed}}{\text{Labour Force}} \times 100
\]

The unemployment rate is a measure of imbalance in the labour market representing the extent of unutilised labour supply of the country. It is also sometimes used in a general sense as an indicator of the health of the economy, not just the labour market. Unemployment rates for specific categories of the labour force, such as men, women, youth, adults, geographic regions, or specific (past) occupations and branches of economic activity, shed light on the groups of workers and sectors of the economy or regions most affected by unemployment.

Table 14 illustrates the calculation of the unemployment rate using data from the Labour Force Survey of South Africa. The results suggest a very slight decline in the unemployment rate in South Africa from 23.0% in 2007 to 22.9% in 2008, in line with the slight improvement noted earlier on the core employment-population ratio (59.8% in 2007 versus 60.7% in 2008).

Unemployment statistics have created a broad range of controversies in many countries, both developing as well as developed countries. The underlying definition of unemployment has been widely criticised, especially its reliance on the “one-hour criterion” of its companion employment definition, which leads to excluding from the classification persons working a few hours during the week who otherwise satisfy the other criteria of the unemployed. Many other issues have been raised and sometimes vehemently, such as those concerning the borderline between employment and unemployment (e.g. young jobseekers working on community work programmes) and between unemployment and inactivity (e.g. so-called discouraged workers) or concerning the confusion between unemployment data based on surveys and registered jobseekers data based on administrative records.

Facing these challenges, some countries have introduced alternative measures of the unemployment rate, providing analysts and the public with a wider range of data for assessing the conditions of the labour market. An example is the set of six “alternative measures of labour underutilization” U1-U6 regularly published by the US Bureau of Labour Statistics:

### Table 14

**Calculations of the Unemployment Rate**

<table>
<thead>
<tr>
<th>South Africa (Labour Force Survey)¹</th>
<th>2007²</th>
<th>2008</th>
<th>Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of unemployed persons ('000)</td>
<td>3,945</td>
<td>4,075</td>
<td>130</td>
</tr>
<tr>
<td>Labour force ('000)</td>
<td>17,178</td>
<td>17,788</td>
<td>610</td>
</tr>
<tr>
<td>Unemployment rate (%)</td>
<td>23.0%</td>
<td>22.9%</td>
<td>-0.1%</td>
</tr>
</tbody>
</table>

¹ Data cover population 15+ years old in 2007 and 15-64 years old in 2008.
² September

---

Part II: Analysis of labour force data

Unemployment generally disproportionately affects the youth labour force. The result is that the unemployment rate of young people is almost invariably higher than that of adults. The following graph shows the unemployment rate by age group in South Africa in September 2007. The rate generally decreases with age, the highest rate being 57% among those aged 15-19, followed by 45% among the 20-24 age bracket, and so on.

Analysis of the youth unemployment rate in a broad range of countries indicates that, in all countries, the youth unemployment rate is higher than the national rate, in most cases more than twice the national rate. A categorisation of the relative seriousness of youth unemployment is shown in Figure 16 below, where the vertical axis measures the youth unemployment rate (ages 15-24) and the horizontal axis the national rate (ages 15+). The light blue area refers to the expected situation where the youth unemployment rate is more than the national rate and the dark blue area shows the seriousness of the situation when the youth unemployment rate is more than twice the national rate.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

17.2 Youth Unemployment

Unemployment generally disproportionately affects the youth labour force. The result is that the unemployment rate of young people is almost invariably higher than that of adults. The following graph shows the unemployment rate by age group in South Africa in September 2007. The rate generally decreases with age, the highest rate being 57% among those aged 15-19, followed by 45% among the 20-24 age bracket, and so on.

Analysis of the youth unemployment rate in a broad range of countries indicates that, in all countries, the youth unemployment rate is higher than the national rate, in most cases more than twice the national rate. A categorisation of the relative seriousness of youth unemployment is shown in Figure 16 below, where the vertical axis measures the youth unemployment rate (ages 15-24) and the horizontal axis the national rate (ages 15+). The light blue area refers to the expected situation where the youth unemployment rate is more than the national rate and the dark blue area shows the seriousness of the situation when the youth unemployment rate is more than twice the national rate.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from $U_1=5.0\%$ to $U_6=15.3\%$ with the official unemployment at $U_3=8.5\%$. Later in this Part, some elements of these alternative measures will be examined in more detail.

NOTE

Persons marginally attached to the labour force are those who currently are neither working nor looking for work but indicate that they want and are available for a job and have looked for work sometime in the past 12 months. Discouraged workers, a subset of the marginally attached, have given a job-market related reason for not currently looking for work. Persons employed part time for economic reasons are those who want and are available for full-time work but have had to settle for a part-time schedule.

FIGURE 15
UNEMPLOYMENT RATE BY AGE GROUP
(SOUTH AFRICA SEPTEMBER 2007)

FIGURE 16
YOUTH VERSUS NATIONAL UNEMPLOYMENT RATE
Part II: Analysis of labour force data

The international standards on unemployment statistics do not define how long a spell of unemployment should be for it to be considered as long-term unemployment. In practice, however, many countries define long-term unemployment as unemployment of more than one year. The following diagram shows the distribution of the unemployed in Tanzania, Zanzibar (2006) by length of unemployment with distinction for men and women. The data show that more than 80% of the unemployed, among men as well as among women, report an unemployment duration of two years or more.

These results indicate not only the gravity of unemployment in Tanzania, Zanzibar, but also the difficulty of measuring and interpreting data on the duration of unemployment in some areas.

The difference between the estimated age of entry in the labour force and the estimated age of entry into employment calculated earlier for South Africa (September 2007) suggests a spell of unemployment (which may include periods of inactivity) of about three years before employment for young men (3 = ages 25-22) and six years for young women (6 = ages 29-23).

### TABLE 15
RATIO OF YOUTH TO NATIONAL UNEMPLOYMENT RATE

<table>
<thead>
<tr>
<th>South Africa (Labour force survey)</th>
<th>2007¹</th>
<th>2008</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youth unemployment rate (15-24)</td>
<td>53.1%</td>
<td>54.4%</td>
</tr>
<tr>
<td>National unemployment rate (15+)</td>
<td>23.0%</td>
<td>22.9%²</td>
</tr>
<tr>
<td>Ratio of youth to national unemployment rate</td>
<td>2.3</td>
<td>2.4</td>
</tr>
</tbody>
</table>

¹ September
² Data for 2008 refer to persons 15-64 years old.

The youth unemployment rate is compared to the national unemployment rate using data from the Labour Force Survey of South Africa in September 2007 and 2008.

The results show that the ratio of the youth to national unemployment rate is more than two, falling in the plain blue area category of high youth unemployment rate both in September 2007 and 2008. It is instructive to note that, while the national unemployment rate changed little or slightly decreased between September 2007 and 2008, the youth unemployment rate increased considerably during that period.

### 17.3 LONG-TERM UNEMPLOYMENT

Duration of unemployment is the length of time that an unemployed person has been without work, available for work and actively seeking work. In practice, what is measured in a survey is the duration of unemployment up to the time of the survey. The completed spell of unemployment runs through the survey and beyond and is not directly measurable by conventional labour force surveys.⁴²

The international standards on unemployment statistics do not define how long a spell of unemployment should be for it to be considered as long-term unemployment. In practice, however, many countries define long-term unemployment as unemployment of more than one year. The following diagram shows the distribution of the unemployed in Tanzania, Zanzibar (2006) by length of unemployment with distinction for men and women. The data show that more than 80% of the unemployed, among men as well as among women, report an unemployment duration of two years or more.

These results indicate not only the gravity of unemployment in Tanzania, Zanzibar, but also the difficulty of measuring and interpreting data on the duration of unemployment in some areas.

The difference between the estimated age of entry in the labour force and the estimated age of entry into employment calculated earlier for South Africa (September 2007) suggests a spell of unemployment (which may include periods of inactivity) of about three years before employment for young men (3 = ages 25-22) and six years for young women (6 = ages 29-23).

---
18. YOUTH AND SCHOOL-TO-WORK TRANSITION

18.1 INTRODUCTION

The youth population is economically the most vital and socially the most vocal group in most countries. The quality and attributes of the youth workforce is largely determined by the education system and the transition from school to work constitutes an important policy concern in all countries. The interaction between the education system and the labour market is in fact a two-way relationship. The education system provides the labour market with the supply of educated workforce for the national economy, while the labour market through the wage structure of occupations and other labour market variables transmits signals on the types of qualifications and specialisations expected from the education system.

18.2 YOUTH POPULATION

The ILO resolution concerning youth employment, adopted at the International Labour Conference (86th ILC 1998), defines in its preamble the youth population as persons 15 to 24 years old.43 While the international standards should be regarded as the norm, it may be appropriate in countries where entry into the labour market for many people is at a later age to extend the definition to include all persons in the age category 15 to 29 years.

18.3 SCHOOL LEAVERS

An important indicator of the relationship between the education system and the labour market is the number of school leavers in a given year. The school leavers may be defined as those young persons who have just left the education system permanently or for a period of time whether after graduating from a specific level of education or dropping out without...
Part II: Analysis of labour force data

18.4 school-to-work transition

An instructive way of examining the transition from school to work of young people based on data generally available from labour force surveys is given below for a particular country. The graph plots the percentage of youth not in school (not in school, not employed) against the percentage at work (employed) by single years of age.

It can be observed that before the age of 17 virtually all young people are at school, with very few

Completing the level of education in which they were. It is thus convenient to subdivide school leavers into graduates and dropouts. The graduates are in a sense the successful school leavers and the dropouts the unsuccessful. The classification by educational attainment determines the level of education from which the graduates have graduated or the dropouts have dropped out. Data on school-leavers in a small island country in the Middle East are given below for illustration purposes:

It should be mentioned that not all school leavers enter the labour market. Some leave the education system with no intention to enter the labour market, such as young women who get married and become homemakers or young men and women who get engaged in voluntary work in their community or abroad. There are also those who delay their entrance into the labour market with a period of economic inactivity for personal or family reasons. Also, the interaction between the education system and the labour market is not limited only to school leavers. There are those who combine education and employment during their schooling. Also, at any given time, there are those adults with work experience who re-enter the labour market after a period of training and education. Similarly, there are leavers from the labour market who re-enter the education system to get more training and education after a period of unemployment or employment.
working. From around 17, the relative number of young people in school diminishes and the number employed increases with age as expected. However, there is a gap between schooling and employment. The gap reflects those youth not in school or employed. Some of these young people are looking for work (corresponding to the dash line labelled ‘unemployed’) and some are inactive. The graph shows that after the age of 17 the gap gets wider until the age of 20 when it remains more or less constant, indicating that in this country there is a significant group of young people who do not move from school to work upon leaving the education system. An interesting question is to determine whether these inactive young people have particular characteristics that distinguish them from others with a normal transition from school to work.

For countries where a large part of young people even at age 15 are not in school and already working, the starting age of the school-to-work transition diagram should be set at a lower value, say, at 10 years old.

The surface of the area formed between the “at school” line and the “employed” line standardised for the length of the age span can be regarded as an index of the transition gap between school and work. A small value of the index indicates that the transition between school and work is smooth and young people go directly from school to work without any spell of unemployment or inactivity. By contrast, a large value indicates the degree of difficulty or delay in making the transition from school to work in that country.

Transition gap = Standardised area between “at school” line and “employed” line

\[ \frac{1}{15} \sum_{i=15}^{29} (100 - S_i - E_i) - (100 - S_o - E_o) = 16.4\% \]

where \( E_i \) and \( S_i \) are the percentage of young people i years old who are “employed” and “at school, not employed” respectively, and \( E_o \) and \( S_o \) are the average of the corresponding end-point values, \( E_o = (E_{15} + E_{29})/2 \) and \( S_o = (S_{15} + S_{29})/2 \). The number in the denominator, 15, is the length of the age span from ages 15 to 29. The transition index ranges from 0 to 100. It is 0 when all young people are either at school, or in employment.
is 100 in the extreme situation where no young person is in school and none is employed.

The transition gap may be decomposed into two components, one reflecting the ease or difficulty of the transition (unemployment) and the other reflecting delay in transition or non-transition (inactivity):

$$\text{Transition gap} = \text{Unemployment} + \text{Inactivity} = 8.3\% + 8.1\%$$

where

Unemployment

$$= \frac{1}{15} \times \sum_{i=15}^{29} (100 - U_i) - (100 - U_o)$$

and

Inactivity

$$= \frac{1}{15} \times \sum_{i=15}^{29} (100 - S_i - E_i - U_i) - (100 - S_o - E_o - U_o)$$

The transition gap and its components may also be measured on the basis of data classified by age group instead of single age provided care is taken to correspondingly adjust the formulas. The transition gap may also be measured for young men and women separately, as well as other population breakdowns of interest, such as by geographical area, urban–rural, marital status or educational attainment.

18.5 FURTHER ANALYSIS

The transition of young people from school to work may be studied more closely on the basis of specialised surveys linked to the labour force survey or conducted independently, such as the ILO School-to-Work Transition Survey designed to collect in-depth information concerning the labour market situation of young men and women and, in particular, to quantify the relative ease or difficulty of labour market entry of young people as they exit school.\textsuperscript{46}

Detailed analysis may also be made with labour force survey data directly using retrospective questions such as:

- **Q1.** At what age did you start your first regular job or business? (After you first left full-time education)
  First job means a job that lasted at least six months, unless terminated by period of labour market activity (i.e. period of unemployment or another job). Casual work, part-time jobs or vacation jobs while in school or university should not be considered as first job in the present context. With follow-up questions:

- **Q2.** Before you started your first regular job or business, roughly how many months or years have you spent as unemployed?

- **Q3.** Before you started your first regular job or business, roughly how many months or years have you spent at casual work, or in part-time or vacation jobs, either as an employee or as a self-employed?

Labour force surveys with rotation sampling schemes\textsuperscript{47} can also be used to construct appropriate labour force flow matrices and study the school-to-work transitions of young people. An example is given below in Table 16, showing the transition matrix for young people 15 to 29 years old, describing the joint school and labour

---

\textsuperscript{46} Elder, Sara, ILO School-to-work transition survey: A methodological guide, Youth Employment Programme, (Geneva; ILO, 2009). The ILO School-to-work transition survey is in fact a combination of two surveys: one addressing young people and the other employers.

\textsuperscript{47} In rotation sample designs, a fraction of sample households and persons remain in the sample at different points in time and provide the basis to derive flow data by matching records from successive survey rounds.
force status of the individual: E = Employed; U = Unemployed; S = Inactive, at school; O = Inactive, not at school.

The first corner cell \((E_{t-1}^e, E_t^e)\) indicates that there were 38,000 young people who were employed in both April and October 2009. The next cell in the same row \((E_{t-1}^e, U_t^u)\) indicates that 2,200 who were employed in April 2009 lost their job and became unemployed in October 2009, and so on. The cells in the third row provide data on the school leavers: 800 young people at school in April 2009 found a job in October 2009; 1,900 became unemployed; 61,100 remained at school; and 1,700 became inactive.

The transition matrix can be analysed under certain assumptions to derive a variety of analytically useful concepts such as:

- estimates of the number of school leavers;
- estimates of the proportion of school leavers entering the labour market;
- likelihood of obtaining employment without any spell of unemployment;
- expected duration of employment before first spell of unemployment; and
- expected duration of unemployment of school leavers entering the labour market.

The analysis could also be carried out for young men and women separately, with further breakdowns by age group, educational attainment, and other characteristics of interest.

### 19. Hours of Work, Underemployment and Labour Slack

#### 19.1 Hours of Work

Because the international definition of employment is expansive, covering all lengths of work including just one hour during a week, it is important that employment is analysed in conjunction with data on hours of work in order to distinguish the various intensities of employment. Data on hours of work are also necessary to calculate time-related underemployment as well as average wages per hour so that the resulting wage data are comparable across different categories of workers.

The recent international standards on the measurement of working time recognise several concepts of hours of work serving different purposes, including contractual hours.
of work, normal hours of work, hours usually worked, hours actually worked, and hours paid for.\textsuperscript{48} Data on hours usually worked and hours actually worked at the main job are generally collected through labour force surveys. Sometimes, these surveys also collect the data for all jobs, including subsidiary jobs. Data on contractual hours and hours paid for are generally collected through establishment surveys and administrative registers. These sources also sometimes provide data on normal hours of work and hours actually worked.

Hours actually worked are the time spent in a job for the performance of activities that contribute to the production of goods and/or services during a specified reference period. It includes the direct hours that the person is engaged in the activities, as well as related hours such as waiting time, time on call, and resting time, coffee break, prayer, etc. It excludes annual leave, public holidays, sick leave and other leaves, as well as commuting time between work and home, longer breaks such as meal breaks and educational activities, even if authorised by the employer.

Hours usually worked are the hours actually worked in a job during a typical week (or in general any specific reference period). It can be calculated, for example, as the most frequent number of hours that a person actually worked per week during the past month.

Figure 20 shows a typical distribution of employed persons according to hours usually worked per week at all jobs. It can be observed that most employed persons usually work between 15 and 49 hours per week at their jobs (61%). The number of persons working less than 15 hours per week is relatively small (12%), whereas the number of persons working long hours of more than 50 hours per week is relatively large (26%).

Many persons working short hours (less than 15 hours per week) are women and young people, in rural areas, working as unpaid family work-

\textsuperscript{48} ILO, Resolution concerning the measurement of working time, 18th ICLS, Geneva, 24 November – 5 December 2008.

\textbf{FIGURE 20}

RELATIVE NUMBER OF EMPLOYED PERSONS BY HOURS USUALLY WORKED PER WEEK AT ALL JOBS

\begin{center}
\begin{tikzpicture}[scale=0.8]
\begin{axis}[
    ybar,\n    bar width=15pt,\n    symbolic x coords={<15 hours, 15-39 hours, 40-49 hours, 50+ hours},\n    xtick=data,\n    ytick={0,5,...,35},\n    ylabel=\text{Relative Number of Employed Persons},\n    xlabel=\text{Hours Usually Worked Per Week At All Jobs},\n    ymax=37,\n    ymin=0,\n    %axis on top,\n    %axis line style={line width=1pt,draw=black},\n    %grid=major,\n\end{axis}
\end{tikzpicture}
\end{center}
The international standards on this topic are limited to the measurement of time-related underemployment, referring to situations where the hours of work of an employed person are insufficient in relation to an alternative employment situation in which the person is willing and available to engage.\textsuperscript{50}

Rates of time-related underemployment for seven countries are presented in Figure 21.\textsuperscript{51} The rates are calculated as percentage of the total number of persons employed and using the threshold of 35 hours of work during the week. The highest rate is for Tanzania in 2005-6.

19.2 UNDEREMPLOYMENT

Underemployment reflects underutilisation of the productive capacity of the employed population. The international standards on this topic are limited to the measurement of time-related underemployment, referring to situations where the hours of work of an employed person are insufficient in relation to an alternative employment situation in which the person is willing and available to engage.\textsuperscript{50}

The data indicate that the incidence of long hours of work (more than 50 hours per week) is relatively higher among men, in urban areas, and among private sector employees and self-employed persons (see Table 17).

\begin{table}[h]
\centering
\begin{tabular}{|l|c|c|}
\hline
 & Short hours of work (<15 hours per week) & Long hours of work (50+ hours per week) \\
\hline
Total & 4.8\% & 42.6\% \\
\hline
Men & 3.5\% & 48.0\% \\
\hline
Women & 11.5\% & 15.6\% \\
\hline
Urban & 3.8\% & 45.7\% \\
\hline
Rural & 6.9\% & 36.4\% \\
\hline
Public-sector employees & 2.0\% & 48.5\% \\
\hline
Private-sector employees & 4.5\% & 26.9\% \\
\hline
Self-employed & 6.6\% & 49.4\% \\
\hline
Contributing family workers & 8.7\% & 23.7\% \\
\hline
\end{tabular}
\caption{SHORT AND LONG HOURS OF WORK}
\end{table}

19.3 LABOUR SLACK

Labour slack reflects the total insufficiency of the volume of work and constitutes a form of labour underutilisation. In addition to unemployment and time-related underemployment, it also includes discouraged workers and other persons.


\textsuperscript{50} ILO, Resolution concerning the measurement of underemployment and inadequate employment situations adopted by the 16th ICLS, Geneva, October 1998.

### FIGURE 21
RATE OF TIME-RELATED UNDEREMPLOYMENT IN SEVEN COUNTRIES

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanzania</td>
<td>16.7%</td>
<td>9.5%</td>
<td>7.2%</td>
<td>6.5%</td>
<td>4.3%</td>
<td>2.8%</td>
</tr>
<tr>
<td>Philippines (2003 Q4)</td>
<td>9.5%</td>
<td>7.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moldova (2007)</td>
<td>6.5%</td>
<td>4.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td>17.8%</td>
<td>12.7%</td>
<td>15.3%</td>
<td>16.5%</td>
<td>8.2%</td>
<td>7.5%</td>
</tr>
<tr>
<td>Turkey (2007)</td>
<td>16.5%</td>
<td>12.3%</td>
<td>15.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Philippines (2006)</td>
<td>19.3%</td>
<td>12.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bosnia and Herzegovina</td>
<td>44.8%</td>
<td>42.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moldova (2007)</td>
<td>13.2%</td>
<td>7.3%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Panama (2007 Oct)</td>
<td>8.2%</td>
<td>5.8%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mexico (2007 Q2)</td>
<td>7.5%</td>
<td>3.7%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 18
ALTERNATIVE RATES OF LABOUR SLACK IN SEVEN COUNTRIES

<table>
<thead>
<tr>
<th>Alternative rates of labour slack</th>
<th>U3 Unemployment rate</th>
<th>U4 Unemployed plus discouraged workers</th>
<th>U5 Unemployed, plus discouraged workers, plus other marginally attached workers</th>
<th>U6 Unemployed, plus all marginally attached workers, plus time-related underemployed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bosnia Herzegovina (2006)</td>
<td>32.0%</td>
<td>37.4%</td>
<td>42.8%</td>
<td>44.8%</td>
</tr>
<tr>
<td>Philippines (2003 Q4)</td>
<td>5.8%</td>
<td>7.9%</td>
<td>12.7%</td>
<td>19.3%</td>
</tr>
<tr>
<td>Tanzania (2005–2006)</td>
<td>3.4%</td>
<td>3.9%</td>
<td>4.7%</td>
<td>17.8%</td>
</tr>
<tr>
<td>Turkey (2007)</td>
<td>9.9%</td>
<td>12.3%</td>
<td>15.3%</td>
<td>16.5%</td>
</tr>
<tr>
<td>Moldova (2007)</td>
<td>5.1%</td>
<td>6.9%</td>
<td>7.3%</td>
<td>13.2%</td>
</tr>
<tr>
<td>Panama (2007 Oct)</td>
<td>5.7%</td>
<td>5.8%</td>
<td>5.8%</td>
<td>8.2%</td>
</tr>
<tr>
<td>Mexico (2007 Q2)</td>
<td>3.4%</td>
<td>3.5%</td>
<td>3.7%</td>
<td>7.5%</td>
</tr>
</tbody>
</table>
with some degree of labour force attachment. Though not classified as unemployed because they do not meet the three criteria of the definition of unemployment, these persons are in situations close to unemployment, in the sense that they are more likely to become unemployed in the next time period than become employed.

No international standard definition of discouraged workers exists, but they could be defined as persons not economically active (i.e. not categorised as employed or unemployed in the sense of the labour force framework), currently available for work and seeking work during the past six months, but not actively looking for work during the last four weeks because of their discouragement from past failure in finding work.

Others with some degree of labour force attachment include persons not in the labour force who want to work and are currently available for work but not actively seeking work for involuntary or economic reasons such as illness, bad weather, etc.

Table 18 shows the values of labour slack calculated for the seven countries in Figure 21, presented in the form U3 to U6 in line with definitions of the US Bureau of Labour Statistics described in section 7.1.

According to these figures, Tanzania (2005-2006) has the lowest unemployment rate at 3.4% but the third largest labour slack rate (U6) at 17.8%. The larger labour slack rate reflects not only the unemployment situation of the country, but also other aspects of labour underutilisation such as time-related underemployment, discouragement, and other marginal attachment to the labour force. This example clearly shows the importance of providing a range of alternative summary statistics to better reflect the state of the labour market, as well as the advantages to not focusing only on a single indicator in the form of the standard unemployment rate.

20 **BRANCH OF ECONOMIC ACTIVITY AND PRODUCTIVITY**

It is often argued that in the course of economic development labour flows from agriculture and other labour-intensive primary activities to industry and then to services, along the process of migration of workers from rural to urban areas.\(^5^2\) Data on employment by branch of economic activity allow the monitoring of this development and understanding its causes.

### 20.1 BRANCH OF ECONOMIC ACTIVITY

Branch of economic activity refers to the activity of the establishment in which an employed person worked during the time-reference period. It describes the activity of the establishment, not the type of work that the individual does when working for that establishment. For example, a person may work as a security guard in a department store, an accountant at a hotel or bus driver who drives passengers to the aircraft at an airport. An establishment may be a farm, a mine, a factory, a workshop, a store, an office or a similar type of economic unit. It is important to distinguish enterprises from establishments. “Enterprise” is a broader concept than “establishment”. An enterprise is a legal entity (or group of legal entities) and may have a number of establishments with different economic activities and different locations.

For classifying economic activities in surveys, most countries are currently using the International Standard Industrial Classification of All Economic Activities, ISIC Rev 3.1, although increasingly the newer version ISIC Rev 4 is also being used. ISIC Rev 3.1 classifies economic activities into 17 broad categories as follows.\(^5^3\) The cor-

---


Part II: Analysis of labour force data

be analysed in conjunction with production data from the SNA to derive preliminary estimates of labour productivity in different branches of economic activity. Labour productivity is generally calculated as the ratio of an index of output to an index of labour input:

$$\frac{\text{Quantity index of gross output}}{\text{Quantity index of labour input}}$$

In practice, the indicator of primary interest is the change in labour productivity where growth in output (or value-added) is compared at constant prices with growth of labour input. This indicator may expressed by

$$\Delta LP = \Delta Q - \Delta L,$$

where $Q$ and $L$ represent, respectively, output and labour and $\Delta$ refers to their arithmetic percentage change with respect to time. Productivity changes

### 20.2 LABOUR PRODUCTIVITY

Employment data from labour force surveys may be analysed in conjunction with production data from the SNA to derive preliminary estimates of labour productivity in different branches of economic activity. Labour productivity is generally calculated as the ratio of an index of output to an index of labour input.

**Figure 23** based on data from the Labour Force Survey of South Africa (2008), shows the division of employed population (aged 15-64) into three broad categories of economic activity: Agriculture, Industry and Services. The division of employment by sector in South Africa is compared with the world average for the same year. The results show that the share of the employment in industry in South Africa is about the same as the world average (26% versus 25%). The share in services is, however, significantly higher (69% versus 44%) and correspondingly the share in agriculture is significantly lower (6% versus 31%).

**Figure 22**

CORRESPONDENCE OF INTERNATIONAL STANDARD CLASSIFICATIONS OF ALL ECONOMIC ACTIVITIES (ISIC REV 3.1 AND ISIC REV 4)

<table>
<thead>
<tr>
<th>ISIC Rev 3.1</th>
<th>ISIC Rev 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Agriculture, hunting and forestry</td>
<td>A. Agriculture, forestry and fishing</td>
</tr>
<tr>
<td>B. Fishing</td>
<td></td>
</tr>
<tr>
<td>C. Mining and quarrying</td>
<td>B. Mining and quarrying</td>
</tr>
<tr>
<td>D. Manufacturing</td>
<td>C. Manufacturing</td>
</tr>
<tr>
<td>E. Electricity, gas and water supply</td>
<td>D. Electricity, gas, steam and air conditioning supply</td>
</tr>
<tr>
<td>F. Construction</td>
<td>E. Water supply; sewerage, waste management and remediation activities</td>
</tr>
<tr>
<td></td>
<td>F. Construction</td>
</tr>
<tr>
<td>G. Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods</td>
<td>G. Wholesale and retail trade; repair of motor vehicles and motorcycles</td>
</tr>
<tr>
<td>H. Hotels and restaurants</td>
<td>H. Transportation and storage</td>
</tr>
<tr>
<td>I. Transport, storage and communications</td>
<td>I. Accommodation and food service activities</td>
</tr>
<tr>
<td>J. Financial intermediation</td>
<td>J. Information and communication</td>
</tr>
<tr>
<td>K. Real estate, renting and business activities</td>
<td>K. Financial and insurance activities</td>
</tr>
<tr>
<td>L. Public administration and defence; compulsory social security</td>
<td>L. Real estate activities</td>
</tr>
<tr>
<td>M. Education</td>
<td>M. Professional, scientific and technical activities</td>
</tr>
<tr>
<td>N. Health and social work</td>
<td>N. Administrative and support service activities</td>
</tr>
<tr>
<td>O. Other community, social and personal service activities</td>
<td>O. Public administration and defence; compulsory social security</td>
</tr>
<tr>
<td>P. Activities of private households as employers and undifferentiated production activities of private households</td>
<td>P. Education</td>
</tr>
<tr>
<td>Q. Extraterritorial organizations and bodies</td>
<td>Q. Human health and social work activities</td>
</tr>
<tr>
<td></td>
<td>R. Arts, entertainment and recreation</td>
</tr>
<tr>
<td></td>
<td>S. Other service activities</td>
</tr>
<tr>
<td></td>
<td>T. Activities of households as employers; undifferentiated goods and services-producing activities of households for own use</td>
</tr>
<tr>
<td></td>
<td>U. Activities of extraterritorial organizations and bodies</td>
</tr>
</tbody>
</table>
may be calculated on the basis of the simple concept of head count, output per worker, or on the basis of hours of work or output per hours.

Figure 24 shows annual changes in labour productivity, defined as output per worker, in nine branches of economic activity (ISIC Rev. 3.1) for illustration purposes. According to this example, labour productivity increased in only two branches of economic activity (manufacturing; and hotels and restaurants). In wholesale and retail trade, productivity change was nil. In all the others, the change in the level of output was less than the change in the level of employment.
21 OCCUPATIONAL STRUCTURE AND SEGREGATION

21.1 INTRODUCTION

An important use of occupation data is the study of the relationship between occupations and skill levels. The relationship between occupations and the level and field of education help to develop appropriate career development and labour market policies for the youth in particular. Occupation data also serve a broad variety of other analytical purposes, including the measurement of occupational segregation between men and women, the share of women in managerial positions, and the degree of mismatch between occupation and educational level of the employed population.

21.2 INTERNATIONAL STANDARD CLASSIFICATION OF OCCUPATIONS (ISCO)

Occupation refers to the kind of work usually done by a person employed (or the kind of work done previously or wanted if the person is unemployed), irrespective of the branch of economic activity or the SE of the person. An occupational classification system puts together occupations of similar tasks and duties or in terms of the similarity of skills required to fulfil the tasks and duties of the job. The new International Classification of Occupations (ISCO-08) classifies occupations in 10 major occupational groups subdivided into 43 sub-major groups, 130 minor groups and 436 unit groups comprising in most cases a detailed number of occupations. ISCO-08 uses the International Standard Classification of Education to define four broad categories of skill levels: 1st level – Primary education or first stage of basic education; 2nd level – Lower secondary education; and 3rd level – Upper secondary education; and 4th level – Higher education.

21.3 OCCUPATIONAL SEGREGATION

A commonly used index of occupational segregation is given by

$$D = \frac{1}{2} \sum \left| \frac{n_{AI}}{n_A} - \frac{n_{BI}}{n_B} \right|$$

where $n_{AI}$ and $n_{BI}$ are, respectively, the number of men and women in occupational category I and $n_A$ and $n_B$ are, respectively, the total number of men and women in all occupational categories. The value of the segregation index $D$ ranges from 0 to 1, 0 indicating no segregation and 1 indicating complete segregation. The index may be interpreted as the fraction of persons that need to change occupations to achieve zero segregation.

The occupational sex segregation index is one of the ILO decent work indicators. It is a commonly used proxy indicator for equality of opportunity in employment and occupation. The index measures the extent to which labour markets are separated into “male” and “female” occupations, e.g. the percentage of female (or male) non-agricultural employment in a female-dominated (or male-dominated) occupation, or to the total non-agricultural employment in a gender-dominated occupation. The indicator reflects direct and indirect discrimination in access to employment opportunities prior to and outside the labour market (i.e. in education and training, perceived suitability of jobs to female roles), at entry and within the labour market (i.e. recruitment, on-the-job training opportunities, promotion, job change during upgrading). Crowding of women in occupations with

low wages explains a big part of differences in earnings between men and women. The indicator can also reflect differences in occupational preferences between genders.

The following table shows the value of the segregation index for the youth population in Iran and compares it with corresponding figures for selected countries.

According to the figures in the table, the index of occupational sex segregation in Iran was 0.6351 in 2006 when measured at four-digit occupation classification and 0.6138 when measured at three-digit level. Note that the value of the index is lower with smaller number of occupational categories, and higher with larger number of categories. The segregation index for Sweden and for the United States are calculated at three-digit occupation classification (based on labour force survey results for young people less than 35 years old) and the resulting values are 0.5298 and 0.5078, respectively.

A simple method that uses level of educational attainment and one-digit occupation data generally available from labour force surveys defines a mismatch when the educational attainment of the worker is higher than the educational level required by the main current job. The following matrix shows the measurement of mismatch in terms of ISCO and International Standard Classification of Education (ISCED), where the grey areas represent mismatch:

According to these data, about 14% of the employed population is working, in their main jobs, in occupations with skill requirements below their educational attainment. The percentage is about the same among men and women. However, the breakdown is different for men and women. The bulk of women with an occupation–education mismatch have high educational attainment (88%) while among men the breakdown of occupational mismatch is about even between those with secondary education (42%) and those with tertiary education (58%).

---

**TABLE 19**

**OCCUPATIONAL SEGREGATION AMONG YOUNG MEN AND WOMEN IN THREE COUNTRIES**

<table>
<thead>
<tr>
<th></th>
<th>Year</th>
<th>Occupation classification</th>
<th>Occupational categories</th>
<th>Segregation index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Iran 2006</td>
<td>4-digit</td>
<td>325</td>
<td>0.6351</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-digit</td>
<td>111</td>
<td>0.6138</td>
<td></td>
</tr>
<tr>
<td>Sweden 2000</td>
<td>3-digit</td>
<td>125</td>
<td>0.5297</td>
<td></td>
</tr>
<tr>
<td>United States</td>
<td>3-digit</td>
<td>491</td>
<td>0.5078</td>
<td></td>
</tr>
</tbody>
</table>

Sources:
- Iran data from 2% population census file; Sweden and US data from labour force surveys reported in ILO SEGREGAT database, http://www.ilo.org/stat.
Notes:
- Iran age group: 15–24 years old; Sweden and United States age group: less than 35 years old.

---

21.4 OCCUPATION AND EDUCATION MISMATCH

The incompatibility between education and occupation of workers refers to the situation where the educational attainment of the worker is above the skill requirement of his or her job. In a sense, this means that the return on investment in education and training is below optimum and somewhat wasted. Different indicators have been used to measure education and occupation mismatch.\(^58\)

A simple method that uses level of educational attainment and one-digit occupation data generally available from labour force surveys defines a mismatch when the educational attainment of the worker is higher than the educational level required by the main current job. The following matrix shows the measurement of mismatch in terms of ISCO and International Standard Classification of Education (ISCED), where the grey areas represent mismatch:

According to these data, about 14% of the employed population is working, in their main jobs, in occupations with skill requirements below their educational attainment. The percentage is about the same among men and women. However, the breakdown is different for men and women. The bulk of women with an occupation–education mismatch have high educational attainment (88%) while among men the breakdown of occupational mismatch is about even between those with secondary education (42%) and those with tertiary education (58%).

---

However, the breakdown is different for men and women. The bulk of women with an occupation-education mismatch have high educational attainment (88%) while among men the breakdown of occupational mismatch is about even between those with secondary education (42%) and those with tertiary education (58%).

An application of this scheme to data from a labour force survey is given below.

According to these data, about 14% of the employed population is working, in their main jobs, in occupations with skill requirements below their educational attainment. The percentage is about the same among men and women.

<table>
<thead>
<tr>
<th>ISCO-08 Major Groups</th>
<th>ISCED-97 Educational attainment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 1 2 3 4 5 6</td>
</tr>
<tr>
<td></td>
<td>1st 2nd 3rd 4th</td>
</tr>
<tr>
<td>1 Managers</td>
<td></td>
</tr>
<tr>
<td>2 Professionals</td>
<td></td>
</tr>
<tr>
<td>3 Technicians and associate professionals</td>
<td></td>
</tr>
<tr>
<td>4 Clerical support workers</td>
<td></td>
</tr>
<tr>
<td>5 Service and sales workers</td>
<td></td>
</tr>
<tr>
<td>6 Skilled agricultural, forestry, fishery workers</td>
<td></td>
</tr>
<tr>
<td>7 Craft and related trades workers</td>
<td></td>
</tr>
<tr>
<td>8 Plant and machine operators and assemblers</td>
<td></td>
</tr>
<tr>
<td>9 Elementary occupations</td>
<td></td>
</tr>
<tr>
<td>0 Armed forces occupations</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal employment (share in non-agricultural employment)</td>
<td>33%</td>
<td>67%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Informal sector employment</td>
<td>54%</td>
<td>81%</td>
<td>86%</td>
</tr>
<tr>
<td>Informal employment outside the informal sector</td>
<td>46%</td>
<td>19%</td>
<td>14%</td>
</tr>
</tbody>
</table>
22. SE AND INFORMAL EMPLOYMENT

SE classifies jobs held by persons at a given point of time with respect to the type of explicit or implicit contract of employment of the person with other persons or organisations. It may refer to the current job or jobs of an employed person or to the last job of an unemployed person who had past work experience.

22.1 EMPLOYED POPULATION BY SE

The International Standard Classification of Status in Employment (ICSE-1993) identifies five categories of persons with respect to their SE:

› **Employees**: Persons working in “paid employment jobs”, i.e. holding an explicit (written or oral) or implicit employment contract with remuneration not directly dependent upon the revenue of the unit for which they work. Remuneration could be in the form of wages or salaries, commission from sales, piece-rates, bonuses, or in-kind payments such as food, housing or training.

› **Employers**: Persons working on own-account or with one or a few partners in “self-employment jobs”, i.e. (a) remuneration is directly dependent on the profits (or potential for profits) derived from the goods and services produced or for own consumption and (b) engaging one or more “employees”, on a continuous basis.

› **Own-account workers**: Persons working on own-account or with one or a few partners in a “self-employment job”, not engaging any “employees”, on a continuous basis.

› **Contributing family workers**: Persons working in a market-oriented establishment operated by a household member, who cannot be regarded as a partner, in a “self-employment job”, not engaging any “employee” on a continuous basis.

› **Members of producers’ cooperatives**: Persons working in a cooperative producing goods and services, in a “self-employment job”, not engaging any “employee” on a continuous basis.

At early ages, in many countries, young people are working as contributing family workers, often combining schooling and work at the family farm or at the household enterprise. At later ages, the relative number of contributing family workers decreases and the relative number of employees increases as young people leave the schooling system and enter the labour market for the first time, often taking an employee job.

The following diagram shows the distribution of the employed population by SE in the main job for two countries, South Africa (2008) and Tanzania (2006). In South Africa, the bulk of the employed are employees (84%) while in Tanzania the bulk are own-account workers (76%). The share of employers is significantly higher in South Africa (5.5%) than in Tanzania (2%), while the relative number of contributing family workers is considerably smaller in South Africa (1%) than in Tanzania (11%).

22.2 INFORMAL EMPLOYMENT

Many persons who cannot find a decent job with sufficient income try to earn a living by working in the informal sector or accepting employment on an informal basis, for example as apprentice or temporary part-time employee, a casual worker without contract, or a home-based producer, street vendor, waste collector, or domestic worker in another household.

An analytically useful concept that encompasses most of these activities is *informal employment*. In line with the international statistical standards on the topic, informal employment is defined to include:

› **Employment in the informal sector (except rare employees in that sector who may have**


In Table 21, the share of informal employment in total non-agricultural employment is reported for three selected countries, South Africa (2010), Uganda (2010) and Mali (2004). In these countries, the informal sector is defined in terms of the informality of the employment relationship. The employment relationship is informal if, in law or in practice, it is not subject to national labour legislation, income taxation, social protection or entitlement to certain employment benefits (advance notice of dismissal, severance pay, paid annual or sick leave, etc.).

According to these data, informal employment is highest in Mali (82%) followed by Uganda (67%) and South Africa (33%). In the first two countries, the bulk of informal employment is made of employment in the informal sector (86% in Mali and 81% in Uganda), while in South Africa informal employment is divided almost equally between employment in the informal sector (54%) and informal employment outside the informal sector (48%).
### TABLE 21
INFORMAL EMPLOYMENT IN THREE COUNTRIES

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Informal employment (share in non-agricultural employment)</td>
<td>33%</td>
<td>67%</td>
<td>82%</td>
</tr>
<tr>
<td></td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Informal sector employment</td>
<td>54%</td>
<td>81%</td>
<td>86%</td>
</tr>
<tr>
<td>Informal employment outside the informal sector</td>
<td>46%</td>
<td>19%</td>
<td>14%</td>
</tr>
</tbody>
</table>

### 23. INCOME FROM EMPLOYMENT AND EARNINGS DIFFERENTIAL

#### 23.1 INTRODUCTION

Increasingly, labour force surveys collect data on income from employment, some limited to earnings of employees from main job, some covering also the earnings of self-employed and income from secondary jobs. Such surveys permit the analysis of earnings differential among different categories of workers and different branches of economic activity. They also permit the analysis of the relationship between economic hardship and the activity status of individuals, in particular the identification of low pay workers and in certain cases the working poor.

- Income related to paid employment includes direct wages and salaries in cash for time worked and work done, remuneration for time not worked, cash bonuses and gratuities, and remuneration in kind and services, profit-related pay and employment-related social security benefits.
- Income related to self-employment is the profit or share of profit generated by the self-employment activity. It can be calculated as the difference between the value of the gross output of the activity and the operating expenses. Income from self-employment includes remuneration received by owner-managers of corporations and quasi-corporations, where relevant. It also includes employment-related social security benefits received by self-employed persons.

#### 23.2 INCOME FROM EMPLOYMENT

Income from employment consists of the payments, in cash, in kind or in services, which are received by individuals, for themselves or in respect of their family members, as a result of their current or former involvement in paid or self-employment jobs. Income from employment excludes income derived from other sources not related to employment such as property, social assistance, transfers, etc. For definitional and measurement purposes, the concept of income related to paid employment is distinguished from the concept of income related to self-employment.

---

62 The term “income from employment” is used here in the same sense as the more exact term “employment-related income” adopted by the 16th ICLS, Resolution concerning the measurement of employment-related income (October 1998).
23.3 DETERMINANTS OF EARNINGS

The main elements determining the level of earnings of a worker are educational attainment and work experience. The higher the educational attainment of a worker, the higher one would expect his or her earnings to be. Similarly, the longer the work experience of the worker, the higher one would expect his or her earnings would be.

Figure 26 shows the average earnings of workers by level of educational attainment. It clearly shows that the average monthly earnings of workers increase with level of educational attainment. The largest increase occurs when passing from diploma to university degree (bachelor’s and above). The average monthly earnings of workers with a diploma are 518 USD while those with a bachelor’s degree and above are 766 USD.

Figure 27 shows the average monthly earnings of the workers by age group, a substitute for work experience. It can be observed that average monthly earnings steadily increase with age and peak at ages 50-54 before starting a generally downward trend. This parabolic pattern of earnings is consistent with the general observation that, after a certain point, productivity decreases with age.

23.4 EARNINGS DIFFERENTIALS

Major studies on the work of women and men show that women are more likely than men to work in the informal economy and that the average earnings of women in the informal economy are lower than those of men in the informal economy. In general, many data sets, like the one in Table 22 below, show that the average earnings of women are lower than that of men across occupations.

According to these results, the largest earnings differentials in absolute terms are among professionals (196 USD) and the lowest among craft- and trade-related workers (4 USD). In

---

**FIGURE 26**

**AVERAGE MONTHLY EARNINGS BY LEVEL OF EDUCATIONAL ATTAINMENT**

---

Figure 27
AVERAGE MONTHLY EARNINGS BY AGE GROUP

Table 22
AVERAGE REPORTED MONTHLY EARNINGS IN MAIN JOB (USD 2007) AMONG MEN AND WOMEN

<table>
<thead>
<tr>
<th>Branch of Economic activity</th>
<th>Men</th>
<th>Women</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>492</td>
<td>455</td>
<td>37</td>
</tr>
<tr>
<td>Managers</td>
<td>906</td>
<td>840</td>
<td>67</td>
</tr>
<tr>
<td>Professionals</td>
<td>760</td>
<td>563</td>
<td>196</td>
</tr>
<tr>
<td>Technicians and Associate professionals</td>
<td>535</td>
<td>406</td>
<td>129</td>
</tr>
<tr>
<td>Clerical support workers</td>
<td>463</td>
<td>394</td>
<td>69</td>
</tr>
<tr>
<td>Service and sales workers</td>
<td>361</td>
<td>256</td>
<td>105</td>
</tr>
<tr>
<td>Skilled Agricultural, Forestry and Fishery workers</td>
<td>397</td>
<td>273</td>
<td>124</td>
</tr>
<tr>
<td>Craft and related trades workers</td>
<td>330</td>
<td>326</td>
<td>4</td>
</tr>
<tr>
<td>Plant and machine operators and assemblers</td>
<td>299</td>
<td>193</td>
<td>106</td>
</tr>
<tr>
<td>Elementary occupations</td>
<td>349</td>
<td>268</td>
<td>80</td>
</tr>
<tr>
<td>unspecified</td>
<td>474</td>
<td>867</td>
<td>(383)</td>
</tr>
</tbody>
</table>
relative terms, the largest earnings differentials are among the plant and machine operators (1.55=299/193) and assemblers and the lowest again among craft- and trade-related workers (1.01=330/326).

The earnings differentials among men and women may be corrected for differences in levels of educational attainment and work experiences using the Mincer model.64 The Mincer earnings function is specified here by

\[ w = \beta_0 + \beta_1 \text{sex} + \beta_2 \text{age} + \beta_3 \text{age}^2 + \beta_4 \text{educ} + \epsilon \]

where \( w \) is the logarithm of earnings, \( \text{sex} \) is a variable with value 1 for men and -1 for women, \( \text{age} \) represents the age variable, \( \text{educ} \) the educational attainment (1 for illiterate to 5 for bachelor’s and above) and \( \epsilon \) represents a residual variable with conditional expected value equal to zero. The application of the Mincer function to the data above gives the following estimates, all highly significant:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter</th>
<th>Estimate</th>
<th>Standard error</th>
<th>t-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>( \beta_0 )</td>
<td>3.742</td>
<td>6.483 e-02</td>
<td>41.014 ***</td>
</tr>
<tr>
<td>Sex</td>
<td>( \beta_1 )</td>
<td>0.125</td>
<td>8.060 e-03</td>
<td>11.841 ***</td>
</tr>
<tr>
<td>Age</td>
<td>( \beta_2 )</td>
<td>0.050</td>
<td>3.397 e-03</td>
<td>31.872 ***</td>
</tr>
<tr>
<td>Age^2</td>
<td>( \beta_3 )</td>
<td>-0.0003</td>
<td>4.238 e-05</td>
<td>-9.362 ***</td>
</tr>
<tr>
<td>Educ</td>
<td>( \beta_4 )</td>
<td>0.287</td>
<td>6.057 e-03</td>
<td>33.032 ***</td>
</tr>
</tbody>
</table>

The highly significant coefficient of the sex variable (\( \beta_1=0.125 \)) indicates that, even after adjusting for differences in length of work experience (age and \( \text{age}^2 \)) and level of educational attainment (\( \text{educ} \)), there is a statistically significant difference between the earnings of men and women, with men receiving on average higher earnings than women (because \( \beta_1 > 0 \)).

24. LOW PAY AND WORKING POOR

24.1 LOW EARNINGS

The size distribution of earnings is an indicator of income inequality in the country. For many people, the most important characteristic of work is pay and the principle of an ‘adequate living wage’ is mentioned in the Preamble to the ILO Constitution.

In the latest version of the ILO framework of decent work indicators, a low pay rate is defined as the percentage of the employed population whose average hourly earnings are below two-thirds of the median of the distribution or an absolute minimum, whichever is greater. Formulating the indicator in terms of a percentage of the median makes it independent of the national currencies and facilitates international comparison. The choice of two-thirds, recommended by the Working Group on Decent Work Indicators at the

---

64 For a critical review see Heckman, James J., Lance J. Lochner, and Petra E. Todd, “Fifty Years of Mincer Earnings Regressions,” First draft June 1998, Revised 19 March 2003.”
87th ICLS has the virtue of simplicity and wide applicability, including in countries that have either not adopted minimum wage legislation or which have set the statutory minimum wage far below the prevailing market wage.

In the analysis of low pay, it is important to distinguish between different categories of low earners. The core is the full-time employed with low monthly earnings. For full-time employed, defined as all employed persons who during the reference week, low earnings involve:

- usually working 40–48 hours per week at all jobs; and
- the total monthly earnings at all jobs being less than a specified threshold. The threshold is defined as two-thirds of the median monthly earnings of all employed persons during the reference week with 40 to 48 of usual work per week at all jobs, young and adults combined. The threshold of low earnings is defined in terms of the distribution of monthly earnings of full-time workers in order to accommodate different national norms regarding minimum wage and differences in hours of work among full-time workers in each country. Since the median of the distribution determines the relative position of a typical worker in that distribution, it is plausible that national norms regarding minimum pay are indeed tied to that median.66

For less than full-time employed persons, low earnings may be defined as all employed persons during the reference week:

- usually worked less than 40 hours per week at all jobs; and
- Received hourly earnings less than a specified threshold consistent with the monthly threshold of full-time workers specified earlier. It can be calculated as the ratio of the threshold of low monthly earnings of full-time workers to the average hours of work per week of full-time workers*52/12.

For employed persons working more than the typical number of hours for full-time work, low earnings may be defined as all employed persons who during the week:

- usually worked more than 48 hours per week at all jobs; and
- received monthly earnings at all jobs which were less than the threshold of low monthly earnings of full-time workers or whose hourly earnings at all jobs were less than the threshold of low hourly earnings of less than full-time workers determined earlier.

Table 24 is based on the data for seven countries reported in the ILO document on the measurement of labour underutilisation mentioned earlier.67

The data show that low earners are substantial in most countries. Tanzania (2005–6) has the highest percentage of employed persons with low earnings (30.4%) followed by Panama (20.5%) and the Philippines (20.5%). It is instructive to note that the principal component of low earnings concerns persons working long hours (more than 48 hours per week). This means that, despite long hours of work, the earnings of a large portion of people remain below the threshold.

24.2 WORKING POOR

Labour force surveys that in addition to income from employment or earnings also collect data on total household income can be used to measure and analyse the working poor. The working poor are broadly defined as working persons who are unable to earn enough to maintain the welfare of themselves and their families. More specifically, the working poor are persons who are working

---

Table 24
LOW EARNINGS IN SEVEN COUNTRIES

<table>
<thead>
<tr>
<th>Low Earnings</th>
<th>Total</th>
<th>Full time employed (40-48 Hours per week)</th>
<th>Less than full time employed (&lt;40 Hours per week)</th>
<th>Overly employed (&gt;48 hours per week)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanzania (2005-2006)</td>
<td>30.4%</td>
<td>6.1%</td>
<td>2.9%</td>
<td>21.4%</td>
</tr>
<tr>
<td>Panama (2007 Aug)</td>
<td>20.5%</td>
<td>7.1%</td>
<td>7.9%</td>
<td>5.4%</td>
</tr>
<tr>
<td>Phillipines (2003 Q4)</td>
<td>20.5%</td>
<td>8.1%</td>
<td>1.0%</td>
<td>11.4%</td>
</tr>
<tr>
<td>Mexico (2007 Q2)</td>
<td>11.3%</td>
<td>2.6%</td>
<td>3.2%</td>
<td>5.4%</td>
</tr>
<tr>
<td>Moldova (2007)</td>
<td>7.2%</td>
<td>5.0%</td>
<td>0.8%</td>
<td>1.3%</td>
</tr>
<tr>
<td>Bosnia &amp; Herzegovina (2006)</td>
<td>6.8%</td>
<td>4.2%</td>
<td>0.0%</td>
<td>2.5%</td>
</tr>
<tr>
<td>Turkey (2007)</td>
<td>6.8%</td>
<td>1.8%</td>
<td>0.8%</td>
<td>4.2%</td>
</tr>
</tbody>
</table>

Figure 28
DEFINITIONS OF WORKING POOR

<table>
<thead>
<tr>
<th>BROAD DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>› Working persons who are unable to earn enough to maintain the welfare of themselves and their family</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SPECIFIC DEFINITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>› Working persons in households with income below the poverty line</td>
</tr>
</tbody>
</table>

Household = All persons, related family members and all unrelated persons, who occupy a housing unit and have no other usual address.

and live in households with income below the poverty line, as shown in Figure 28.

The US collects data on the working poor based on a survey conducted as part of an annual survey, called the Annual Social and Economic Supplement to the Current Population Survey (i.e. the US monthly labour force survey). The definition of the working poor distinguishes between two categories of persons: (a) Working persons living as unrelated individuals with income below the poverty level; and (b) Working persons living in families with total income below the poverty level.

Family is defined as a group of two or more persons residing together who are related by birth, marriage, or adoption. The threshold of poverty is defined on the basis of the different consumption requirements of individuals and families and factors such as family size and number of children younger than 18. A working

person is a person who has been in the labour force during the year, in one of the following categories:

- Did not work during the year
- Worked during the year
- Usual full-time worker
- Usual part-time worker
  - Involuntary
  - Voluntary

The resulting data presented above indicate that young workers are more likely to be among the working poor than adult workers. Among young people aged 16 to 19 who are working, about 10.2% are poor. The percentage is even higher (11.7%) among those who are 20 to 24.