Labour Force Data Analysis:
Guidelines with African Specificities



2012

AFRICAN DEVEVELOPMENT BANK

This report was prepared by the Statistical Capacity Building Division of the Statistics Department, in the Office of the Chief Economist at the African Development Bank. The findings reflect the opinions of the authors and not necessarily those of the African Development Bank or its Board of Directors. Every effort has been made to include inputs by labour practitioners from a number of national statistical offices who participated in various labour statistics workshops and seminars.

Statistical Capacity Building Division
Statistics Department
African Development Bank
Temporary Relocation Agency
BP 323, 1002 Tunis, Belvédère
Tunis, Tunisia
Tel.: (216) 71 10 36 54
Fax: (216) 71 10 37 43
E-mail: Statistics@afdb.org
Website: www.afdb.org

Copyright © 2012 African Development Bank

Design/layout by Phoenix Design Aid AS/Denmark. ISO 14001/ISO 900
Certified and approved CO_2 Company.
The Annual Report is printed on environmentally-friendly recycled paper with vegetable inks.
The printed matter is further recyclable.
ISO 14001 certified and EMAS-approved.

LABOUR FORCE DATA ANALYSIS:
Guidelines with African Specificities

TABLE OF CONTENTS

Preface IV
Acknowledgements V
List of figures VI
List of tables VII
Acronyms and abbreviations VIII

Part I: Measurement of the EAP 1

1. Introduction 2
2. The EAP 5
3. The current EAP (or labour force) 9
4. The employed population 11
5. The unemployed population 14
6. Population outside the labour force 16
7. Employment in the informal sector and informal employment 17
8. Working time measurement 20
9. Measures of underutilised labour 21
10. Measurement of long-term employment situation (usual activity) 25

Annex
a. Slides on hours concepts 31

Part II: Analysis of labour force data 35

11. Introduction 36
12. Descriptive analysis of survey variables 36
13. Analysis of supply and demand of labour 51
14. Size and composition of the population 52
15. Labour force participation of men and women 55
16. Employment-population ratio 60
17. Unemployment and its duration 63
18. Youth and school-to-work transition 67
19. Hours of work, underemployment and labour slack 71
20. Branch of economic activity and productivity 75
21. Occupational structure and segregation 78
22. SE and informal employment 81
23. Income from employment and earnings differential 83
24. Low pay and working poor 86

Part III: Quality assessment of labour force survey data 91

25. Coherence and comparability of survey estimates 92
26. Sampling errors and precision of estimates 97
27. Non-sampling errors and weights adjustment 102

PREFACE

Labour statistics are an important element in the measurement of economic growth and development and yet the data required to provide evidence-based decision making are scanty or non-existent. Within the implementation framework of its Statistical Capacity Building Program II, the African Development Bank (AfDB) in consultation with its regional member countries identified the lack of expertise in data analysis of labour force surveys as a major drawback in the production of labour force statistics in many countries. As a result, the data collected are not properly or fully analysed, even sometimes not analysed at all. This is due to a variety of reasons related to decisions and activities both before actual data collection and subsequent to it. Amongst these are questions relating to questionnaire design, field work, preparation of the collected data for analysis and the analysis itself. In an effort to address these issues, the AfDB decided to commission this guidebook so that statisticians in Africa will have a practical guide when analysing statistics on the labour force and users of these statistics will benefit from an introduction to methods for their analysis. The guidebook is in three parts, introducing the relevant concepts and definitions for labour force statistics, analysis of labour force data and quality assessment of labour force surveys.

PART I deals with the concepts and definitions of the economically active population (EAP); the current EAP (i.e. the labour force); employment and unemployment statistics; working time statistics; time-related underemployment and inadequate employment; labour under-utilisation; employment in the informal sector and informal employment; and the usual EAP. It also introduces the various uses and sources of these statistics. In addition, it reflects developments that have taken place since the publication of the manual of the International Labour Organization (ILO) on the EAP in 1990.[1]

PART II presents methods for analysing data from labour force surveys, from descriptive analysis used in the production of a survey report to analysis of labour force structures and trends that are particularly useful in analytical studies. The former describes approaches to understanding and using tables, diagrams and related statistics to write survey reports. The latter sections introduce key indicators of the labour market and their uses. Areas covered include the supply and demand of labour; the size and composition of the population; labour force participation of men and women; the employment-population ratio; unemployment and its duration; youth and school-to-work transition; hours of work; underemployment and labour slack; branches of economic activity and productivity; occupational structure and segregation; status in employment (SE) and informal employment; income from employment and earning differentials; and low pay and the working poor.

PART III discusses methods for assessing the quality of a labour force survey in terms of (a) the coherence and comparability of the statistics produced from it both internally and externally in comparison with other sources; and (b) the identification and estimation of both sampling and non-sampling errors associated with the survey.

This guidebook is directed at labour statistics practitioners in national statistics offices and policy analysts in development, economic and labour ministries, as well as those in central banks.



Mthuli Ncube
Chief Economist and Vice President

2 Hussmanns, R., Mehran, F., Verma, V. (1990): Surveys of economically active population, employment, unemployment and underemployment, ILO Geneva.

ACKNOWLEDGMENTS

This book was prepared by a team headed by Oliver J.M. Chinganya (Head of the Statistical Capacity Building Division, Statistics Department). Other core members of the team included Besa Muwele (Senior Statistician, Statistical Capacity Building Division, Statistics Department), and Abimbola Sylvester Young and Farhad Mehran (consultants), both of whom are former directors of the Bureau of Statistics of the International Labour Office.

The Guide Book was prepared under the close and direct supervision of Charles L. Lufumpa (Director, Statistics Department of the AfDB), under the overall guidance of Mthuli Ncube (Chief Economist and Vice President of the AfDB).

The preparation process benefited from the inputs of various labour practitioners in a number of national statistical offices who participated in various workshops and seminars.

LIST OF FIGURES

PART I

Figure 1. Labour statistics: Bridge between economic and social statistics
Figure 2. Labour force surveys: Advantages and limitations
Figure 3. Establishment surveys: Advantages and limitations
Figure 4. Administrative records: Advantages and limitations
Figure 5. Productive and non-productive activities
Figure 6. Labour force framework
Figure 7. International definition of employment
Figure 8. Usually active measurement framework

PART II

Figure 1. Employment by industry
Figure 2. Histogram and density curve of an income distribution (artificial data)
Figure 3. Conditional distributions of occupation given SE (US, 2006)
Figure 4. Scatter diagram of variables X and Y (artificial data)
Figure 5. Supply and demand of labour
Figure 6. Tanzania age pyramid 2010
Figure 7. Population age pyramids, 2010
Figure 8. Dependency ratios (2010)
Figure 9. Labour force participation rate (World, 2010)
Figure 10. Labour force participation rate, Tanzania (Zanzibar), 2008
Figure 11. Labour force participation rate, USA March 2006
Figure 12. Labour force participation rate of men and women by educational attainment
Figure 13. Labour force by educational attainment in four countries
Figure 14. Female labour force participation rate by marital status
Figure 15. Unemployment rate by age group
Figure 16. Youth versus national unemployment rate
Figure 17. Percentage number of unemployed persons by duration of unemployment
Figure 18. School leavers: academic year 2004-2005 (An illustration)
Figure 19. Youth school-to-work transition: An illustration with labour force survey data
Figure 20. Relative number of employed persons by hours usually worked per week at all jobs
Figure 21. Rate of time-related underemployment in seven countries
Figure 22. Correspondence of International Standard Classifications of All Economic Activities (ISIC Rev 3.1 and ISIC Rev 4)
Figure 23. Employment by branch of economic activity, South Africa (2008)
Figure 24. Labour productivity in selected branches of economic activity
Figure 25. SE in main job in two countries
Figure 26. Average monthly earnings by level of educational attainment
Figure 27. Average monthly earnings by age group
Figure 28. Definitions of working poor
Figure 29. Profile of working poor by age group (USA 2004)

PART III

Figure 1. Coherence of employment and output trends
Figure 2. Sampling distribution of unemployment estimates

LIST OF TABLES

PART II

Table 1. Employment by industry (South Africa, First Quarter 2008)
Table 2. Age distribution by labour force status (South Africa, First Quarter 2008)
Table 3. Summary statistics of age distribution by labour force status (South Africa, First Quarter 2008)
Table 4. Employment by industry
Table 5. Employed population by occupation and SE
Table 6. Conditional distributions of SE given occupation (US, 2006)
Table 7. Conditional distributions of occupation given SE (US, 2006)
Table 8. Summary statistics for conditional distributions of income by industrial sector
Table 9. Conditional distributions of occupation given SE and sex
Table 10. Conditional distributions of occupation given SE and year
Table 11. Summary statistics for conditional distributions of income by industry and sex
Table 12. Calculation of the core employment-population ratio
Table 13. Age at employment and at retirement
Table 14. Calculation of the unemployment rate
Table 15. Ratio of youth to national unemployment rate
Table 16. School-to-work transition matrix
Table 17. Short and long hours of work
Table 18. Alternative rates of labour slack in seven countries
Table 19. Occupational segregation among young men and women in three countries
Table 20. Education and occupation mismatch in terms of ISCO and ISCED
Table 21. Informal employment in three countries
Table 22. Average reported monthly earnings (USD 2007) in main job among men and women
Table 23. Estimated parameters of the earnings function
Table 24. Low earnings in seven countries

PART III

Table 1. Reconciliation of employment data: Labour force survey and National Accounts
Table 2. Reconciliation table: Number of employed persons in manufacturing establishments with 50 or more workers
Table 3. Reconciliation table: Unemployed persons seeking work at public or related employment offices
Table 4. Sampling errors of main labour force indicators: A numerical example
Table 5. Relative standard error by size of estimate

ACRONYMS AND ABBREVIATIONS

AfDB	African Development Bank
CV	Coefficient of Variation
EAP	Economically Active Population
ICSE	International Standard Classification of Status in Employment
ICLS	International Conference of Labour Statisticians
ILO	International Labour Organization
ILO WG	Working Group on the Advancement of Employment and Unemployment Statistics of the ILO
ISCED	International Standard Classification of Education
ISCO	International Standard Classification of Occupations
ISIC	International Standard Industrial Classification of All Economic Activities
KILM	Key Indicators of the Labour Market
LFPR	Labour Force Participation Rate
PSU	Primary Sampling Unit
SE	Status in Employment
SNA	System of National Accounts
UN	United Nations
UNSD	United Nations Statistics Division
US	United States of America
USD	United States Dollars

PART I

Measurement of the EAP

1. Introduction
2. The EAP
3. The current EAP (or labour force)
4. The employed population
5. The unemployed population
6. Population outside the labour force
7. Employment in the informal sector and informal employment
8. Working time measurement
9. Measures of underutilised labour
10. Measurement of long-term employment situation (usual activity)

Annex

a. Slides on hours concepts

1. INTRODUCTION

1.1 USES

Statistics on the EAP and its components, including indicators derived from them, have both economic and social uses (see Figure 1). They are used to describe and analyse the supply of labour (a) towards the formulation, implementation and monitoring of macro-economic and micro-economic policies and programmes and (b) in human resource development planning. These statistics provide measures of the supply side of the labour market such as the size and characteristics of the EAP, of labour input into production, of the structure and conditions of employment, and of the extent of utilisation of available labour resources (volume and persons). They include statistics on:

A. The EAP, including the size of its components (employment, unemployment and not economically active) and its socio-demographic characteristics, such as age, sex, geographical location, educational attainment, residency, and race (where applicable);

B. The characteristics of the employed population as measured through their various jobs, including their occupation, industry, SE, institutional sector, working conditions (such as employment income, working time, job security and safety, training, participation in decision making), as well as employment in the informal sector and informal employment;

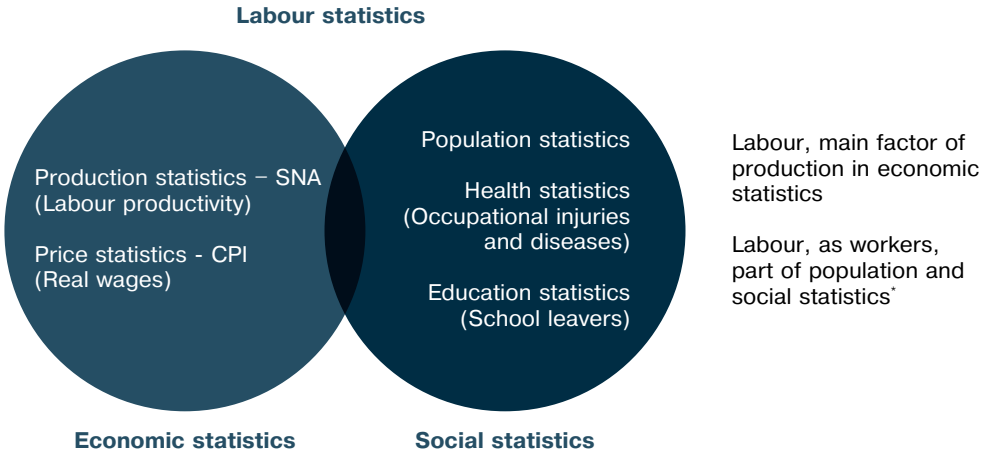
C. The characteristics of the unemployed population including duration of unemployment, job search methods, access to social benefits, and the characteristics of previous jobs they may have had; and

D. The characteristics of those not economically active, including their reasons for not being active and the extent of their labour market attachment, as well as characteristics of previous jobs they may have had.

NB: the above terms are as defined in various resolutions and guidelines of the International Conference of Labour Statisticians (ICLS).

The analysis of these statistics over time provides a basis for monitoring trends in the economy and the employment situation. For example, the unemployment rate is widely used

FIGURE 1
LABOUR STATISTICS: BRIDGE BETWEEN ECONOMIC AND SOCIAL STATISTICS



* Also: Household income statistics (Income from work main determinant of household income); Time-use statistics (Hours of work main element of time-use).

in developed economies to assess the current performance of the economy. These statistics along with labour demand statistics (such as filled jobs, vacancies, labour costs, compensation and training needs) and contextual statistics (such as economic growth, investment, population growth and skills training) are essential in the analysis of the labour market. Also, when used in conjunction with other relevant statistics, they are useful in understanding other economic phenomena and evaluating corresponding policies, such as the possible impact of a specific fiscal policy on employment.

Statistics on the EAP and its components play a key role in the design and evaluation of government socio-economic policies and programmes relating to welfare, migration, education and health, including income maintenance, poverty, social exclusion, job creation, vocational training and so on. The analysis of statistics on employment, income and other socio-economic variables provides information on the income-generating capacity of different economic activities, the extent to which there is decent work, the adequacy of employment income for welfare objectives, the existence of group inequalities in access to employment opportunities and so on.[2]

1.2 SOURCES

The best source for these statistics is labour force surveys. Other household surveys and population censuses could also be used. Statistics on employment are available from establishment/enterprise surveys and censuses as well as from administrative systems such as social security records, tax records, and public-sector payrolls. Unemployment statistics are obtainable from records of employment agencies on registered jobseekers and from those of social security/benefits schemes on recipients of unemployment benefits. These sources all have their relative merits[3] (see Figures 2, 3 and 4). They may be used together to obtain a comprehensive picture of the employment situation and to mutually improve the quality of their data. It should be noted that they could genuinely produce different estimates for the same phenomenon, for example, different estimates of employment from a labour force survey and from an establishment survey.

It is important to appreciate that statistics on the EAP and its components, on their own, cannot respond to questions on all the above issues. They can only do so in tandem with other relevant statistics. For example, in order to properly assess labour productivity, statistics are required on production as well as on labour

FIGURE 2
LABOUR FORCE SURVEYS: ADVANTAGES AND LIMITATIONS

ADVANTAGES

- Comprehensive coverage
- Independent measurement in line with international standards
- Joint data on employment, unemployment and economic inactivity
- Possibility of other data on labour slack, etc.

LIMITATIONS

- Relatively costly, especially if conducted with high frequency
- Subject to sampling and response errors
- Requires statistical infrastructure
- Limited geographic and other disaggregations

2 See United Nations & ILO (2009) for additional information on the uses of these statistics.

3 See Pember, R.J. & H. Djerma (2005): Development of labour statistics system, Bulletin of Labour Statistics, Vol. 1, ILO Geneva for full discussion of these merits.

FIGURE 3
ESTABLISHMENT SURVEYS: ADVANTAGES AND LIMITATIONS

ADVANTAGES

- More accurate
  information based on establishment records
- Detailed industry data
  (and occupation at 3-digit codes)
- Varied types of wage data
  (wage rate, earnings, labour cost)
- Varied types of hours data
  (hours paid for, normal hours of work, overtime hours)

LIMITATIONS

- Coverage often limited to formal sector
  establishments certain size
- Double-counting
  of multiple jobholders
- Non-response relatively high
- Auxiliary information limited
  to items available in establishment record

FIGURE 4
ADMINISTRATIVE RECORDS: ADVANTAGES AND LIMITATIONS

ADVANTAGES

- Lower cost
- Full count
- No reporting burden
- Timeliness
- Possibility of deep geographical and other disaggregation

LIMITATIONS

- Dependence on current legislation
- Scope and coverage problems
- Instability over time
- False/inadequate reporting
- Limited auxiliary variables
- Confidentiality-

input. Also, for analysis of the working poor, statistics are required on the employment and poverty situation of each person. Moreover, challenges relating to coverage, concepts, definitions, classifications, units, and so on could have a bearing on the combination or use of statistics from different sources. Producers and users of these statistics should therefore be aware of the issues raised in this sub-section.

1.3 INTERNATIONAL STANDARDS

As always, the production of statistics on a phenomenon of interest requires clarity and uniqueness in its concepts and definitions. In the case of the EAP, international standards on its concepts and definitions have been established through various resolutions and guidelines of the ICLS of the ILO. The main one is the *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment* adopted by the 13th ICLS in 1982 (hereinafter referred to as the 13th ICLS *Resolution*). It has been amplified by the *Guidelines on the implications of employment promotion schemes on the measurement of employment and unemployment* adopted by the 14th ICLS (1987); *Guidelines concerning the treatment in employment and unemployment statistics of persons on extended absences from work* adopted by the 16th ICLS (1998); and the *Resolution concerning the measurement of underemployment and inadequate employment situations* adopted by the 16th ICLS (1998). The 18th ICLS (2008) adopted a *Resolution* amending one of the key paragraphs of

the 1982 *Resolution* to bring it in line with the most recent version of the System of National Accounts (SNA). Other supporting resolutions have been adopted in relation to classification of occupations (18th ICLS 2008), to classification of SE (15th ICLS 1993), to income from employment (16th ICLS 1998) and working time (18th ICLS 2008). The 15th ICLS in 1993 adopted the *Resolution concerning statistics of employment in the informal sector* and the 16th ICLS in 1998 adopted the *Guidelines concerning a statistical definition of informal employment*.

In response to a resolution adopted by the 18th ICLS in 2008 and recommendations from the review of labour statistics commissioned by the UN Statistics Commission, the ILO established the Working Group on the Advancement of Employment and Unemployment Statistics to consider proposals for revising the 13th ICLS *Resolution* (the ILO WG). The ILO WG is well advanced in its work and has reached some conclusions on the contents of the new resolution which are also referenced in this part of the Guide Book.

1.4 PURPOSE AND MAIN REFERENCES

This part of the Guide Book explores the basic concepts and definitions used in the production of statistics on the EAP and its components based on the above international standards, the ILO WG proposals and current best practices in countries. It draws from various published sources, including the following:

A. Hussmanns, R. (2007): *Measurement of employment, unemployment and underemployment – Current international standards and issues in their application,* Bulletin of Labour Statistics, Vol. 1, pp. IX-XXXI, ILO Geneva.

B. Hussmanns, R., Mehran, F., Verma, V. (1990): *Surveys of economically active population, employment, unemployment and underemployment*, ILO Geneva.

C. ILO (2000): *Current International Recommendations on Labour Statistics*, ILO Geneva.

D. ILO Bureau of Statistics (2003): *International Training Compendium on Labour Statistics, Module 1, Statistics on employment, unemployment, underemployment: economically active population*, ILO Geneva.

E. UN (2008): *Principles and Recommendations for Population and Housing Censuses, Revision 2,* UN New York.

F. UN & ILO (2009): Handbook on Measuring the Economically Active Population and Related Characteristics in Population Censuses, UN New York.

All the resolutions and guidelines discussed below can be accessed at http://www.ilo.org/global/statistics-and-databases/standards-and-guidelines/lang--en/index.htm. Some are also available in ILO (2000).2.

2. THE EAP

The intention is to determine a definition of activity status that would allow the mutually exclusive and exhaustive categorisation of the population of interest into:

A. Those 'actively' supplying labour;

B. Those not 'actively' supplying labour, but willing and looking to do so; and

C. Those not doing either of the above.

2.1 DEFINITION

The current approach is that the most important reason for determining the activity status of a person is economic. Therefore, statistics on the EAP should be produced in such a way that they are consistent with economic statistics, in particular the SNA. This then requires the ac-

The economically active population comprises all persons of either sex who furnish the supply of labour for the production of goods or services within the production boundary, as defined by the latest version of the SNA, during a specified time-reference period. According to the SNA 2008, the relevant production of goods and services includes all production of goods, the production of market and non-market services, and the production for own final consumption of household services by employing paid domestic staff.

tivity concept of employment statistics to be the same as the activity concept of production statistics. Thus, the 13th ICLS *Resolution*, and its amendment in the 18th ICLS (2008), adopted the concept of economic activity as defined in the SNA as the basis for measuring the EAP. According to the 18th ICLS *Resolution*:[4]

Hence, a person is considered to be economically active if, and only if, that person contributes or is available and seeking to contribute to the production of goods and services that fall within the SNA production boundary. In any household survey, it is therefore crucial to be able to unambiguously identify which activities are economic in order to apply the above definition of EAP.

2.2 SNA PRODUCTION BOUNDARY[5]

The SNA 2008 defines production in general terms to be any activity performed under the control, management and responsibility of an institutional unit (financial corporation, non-financial corporation, government, non-profit institution serving households, and household) that uses labour and assets to transform inputs of goods and services into outputs of other goods and services. Excluded as a result are:

I. activities that are not under any direct human intervention such as the unmanaged growth of fish stocks in international waters;

II. basic human activities such as eating, sleeping and doing exercise that cannot be performed by a third party for an individual; and

III. activities not producing any outputs such as begging or stealing (unless the goods so acquired are resold).

A. Furthermore, the SNA production boundary, the basis on which all measurements relating to production are made in the national accounts, includes only some of the activities that are production in the above general terms. These are:

B. the production of all goods or services that are supplied to units other than their producers, or intended to be so supplied;

C. the production of all goods that are retained by their producers for their own final consumption or gross fixed capital formation;

D. the production for own final consumption of housing services by owner-occupiers of dwellings (involves no labour input and so is not of any relevance in the measurement of the EAP); and

E. the production for own final consumption of domestic or personal services produced by employing paid domestic staff.

These are the activities that are defined as economic activities in the SNA and used in the above definition of the EAP (see Figure 5).

4 18th ICLS (2008): Report of the Conference, p. 66, ILO Geneva.

5 Commission of the European Communities, International Monetary Fund, Organization for Economic Cooperation and Development, United Nations, World Bank (2008): System of National Accounts 2008, Brussels/Luxembourg, New York, Paris, Washington DC.

FIGURE 5
PRODUCTIVE AND NON-PRODUCTIVE ACTIVITIES

PRODUCTION IN GENERAL SENSE: ACTIVITIES LEADING TO						Non-Productive Activities
Production of Goods and Services (intended to be) supplied to other units		Production for own final use				
		All Goods[6]	Services			
Goods or services sold at economically significant prices, bartered or used for payments in kind by market producers (most of their products)	Goods or services provided free, or at prices not economically significant, by non-market producers (government units or non-profit institutions)	Producing agricultural or primary products; Processing agricultural products; Other kinds of processing; Construction or major renovation of own dwellings	Housing services produced by owner occupiers of dwellings (no labour input, however)	Domestic or personal services produced by paid domestic staff	Domestic or personal services provided unpaid by household members e.g. cleaning, repairing, preparation of meals, care work, training children, transporting of household members	Activities not under human intervention e.g. unmanaged growth of fish stocks in international waters; Basic human activities e.g. eating, sleeping; Activities without outputs e.g. begging, stealing
Within SNA Production Boundary					Beyond SNA Production Boundary	

Source: Adapted from ILO (2003), Figure 1b

The production in bullet point (a) above is subdivided into market production and non-market production. The former is carried out by establishments, most[7] of whose output is intended for sale at economically significant prices and typically includes corporations and unincorporated household enterprises. The latter, non-market production, is production by government units and non-profit institutions. Their output is either supplied free or at prices that are not economically significant. Activities relating to both these types of production typically involve some form of remuneration to those who participate in them, in the form of pay or profit. These activities may also be done on the basis of exchange of labour or barter of goods.

The production in bullet point (b), referred to as 'production for own final use', typically but not exclusively consists of:

B1 production of agricultural/primary products (e.g. milk, cereals, cotton, gathering of berries, forestry, wood-cutting, collection of firewood, hunting, fishing, mining salt, cutting peat, supply of water);

B2 processing of agricultural products (e.g. making butter, flour, wine, baskets, mats; preservation of meat, fish, fruits; curing of skins and making of leather; threshing of grain; etc.);

B3 other kinds of processing such as weaving cloth; making furniture, footwear, utensils, pottery; dress-making and tailoring; and

B4 construction, major renovation or extension of own dwellings, farm buildings, etc.[8]

The activities relating to this type of production are to be included as economic activities

6 Provided these are quantitatively important in relation to total supply in economy.

7 SNA 2008 considers enterprises for which only some (less than 50%) of their output is intended for sale at economically significant prices as producers for own final use. The ILO position, however, is also to include them as market producers.

provided the amount of the good produced is quantitatively important in relation to the total supply of that good in the country.[9]

Excluded from the SNA production boundary are the unpaid domestic and personal services produced by households for their own final consumption. These include services such as cleaning, decoration and minor maintenance of the dwelling occupied by the household; cleaning, servicing and repair of household durables or other goods; preparation and serving of meals; care, training and instruction of children; care of sick, infirm or old people; and transportation of household members or their goods. Whilst acknowledging that these activities are productive, in the above general terms, and have economic value, the SNA advances several reasons for excluding them from the definition of the SNA production boundary. Unlike corresponding goods produced in the same way, these services are consumed immediately on production and cannot be stored. This means there is no possibility to later decide to exchange them in the market. As they are also not produced for the market, no typical suitable market prices exist for use in valuing them. In addition, given the widespread performance of these activities by all old enough to do so and the definition of employment as engagement in production included within the SNA production boundary, the inclusion of these activities in this boundary would lead to virtually everyone being employed and so make nonsense of employment statistics.

However, although they are not included in the SNA production boundary, these activities should still be measured given their importance in gender analysis and in understanding the interaction between them and economic activities. For example, depending on the performance of the economy, households shift between out-sourcing care activities and doing them as unpaid domestic services. Their measurement is therefore necessary to better understand trends that may be observed in the measurement of the EAP over time or for different countries.

The 13th ICLS *Resolution*, including its amendment in the 18th ICLS, is in the process of being revised. One of the current proposals is to no longer consider the activities in (b) above as employment activities even though they are economic activities within the SNA production boundary. Persons engaged in these activities along with those relating to household services for own final consumption, who are excluded from the SNA production boundary, will be treated as unemployed if they are looking for work or as 'persons not in the labour force' otherwise. It should be noted that in those countries where the activities in (b) are carried out by a significantly sizeable group of persons, this change could have a great impact on their employment size and unemployment rate.

2.3 AGE LIMITS

The above definition of the EAP refers to the economic activity of 'all persons'. In reality, there are persons such as babies, infants and some young persons who are not in a position to perform any economic activity. There is therefore no need to ascertain their economic activity status when measuring the EAP. Moreover, for some older children and young persons, compulsory schooling and/or national labour legislation may prevent them from being involved in economic activities in any significant numbers. Thus, the cost and time wasted in identifying the very few with such activities may not be justifiable. The international guidelines therefore recommend that countries should specify a minimum age limit for the measurement of the EAP. Given the very different cultural, economic, legal and educational practices amongst countries, no universally applicable value was set in the guidelines. In practice, the values chosen vary between 14, 15 and 16 years of age depending on the educational system. Values such as 10 and even six

8 See Hussmanns et al (1990) for further examples.

9 The criterion adopted in the 13th ICLS Resolution is that the quantity should be an important contribution to the total consumption of the household.

years have also been used in some countries.[10] The population above the chosen minimum age is referred to as the **working age population.**

Although not recommended by the international standards and not supported by the increasing tendency to abolish retirement age, some countries do apply a maximum age limit for defining the working age population.

2.4 REFERENCE PERIODS

Measurement of the EAP based on the above definition also requires a specification of the period when the economic activity was done. The 13th ICLS *Resolution* distinguishes two such periods: a short reference period such as one day or one week and a long reference period such as a year. The EAP measured using the short reference period is referred to as the **currently active population (or the labour force)** whilst that based on the long reference period is called the **usually active population.**

3. THE CURRENT EAP (OR LABOUR FORCE)

3.1 DEFINITION

According to the 13th ICLS *Resolution*:

> The currently active population, or the **labour force,** comprises all persons above a specified minimum age who, during a specified brief period of one day or one week, fulfil the requirements for inclusion among the employed or the unemployed.

The definitions of employment and unemployment are treated later in Part I.

The measurement of the labour force thus requires the identification of persons who were employed and those who were unemployed, the rest of the population then being outside the labour force. The method by which this identification is carried out is referred to as the labour force framework.

3.2 LABOUR FORCE FRAMEWORK

The framework classifies persons in the population into three groups – employed population, unemployed population and population outside the labour force – based on their activities during the short reference period. It uses a set of priority rules to ensure that these groups are mutually exclusive and exhaustive. **The priority rules** are that, within the short reference period:

- The economic activity status of employment takes precedence over all other economic activity statuses; and
- For those who are not in employment, the activity status of unemployment takes precedence over the status of not being economically active.
- Thus, in classifying the population according to the labour force framework, the following steps are applied sequentially:
- First, identify persons who were employed amongst all those in the population above the minimum working age.
- Next, from amongst the rest of the working age population, identify those who were unemployed.
- Then the residual group in the population, including those below the minimum working age, constitute those outside the labour force (i.e. not currently active).

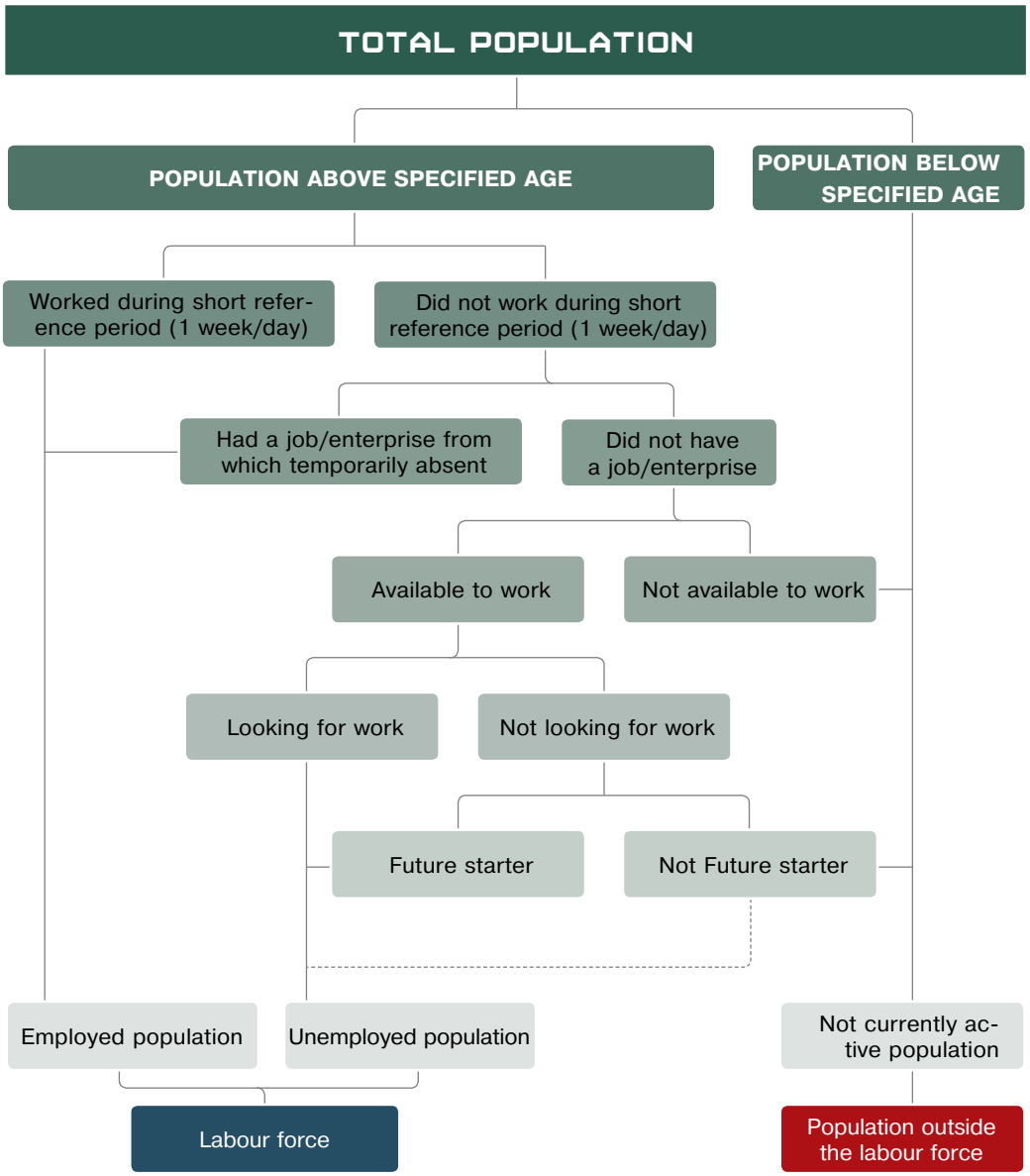
As a consequence, a person who was working and at the same time looking for other work is classified as employed whilst a student attending college and at the same time looking for work is classified as unemployed. A corollary of

10 See UN & ILO (2009), pp. 46–47 for further discussion on the choice of a minimum age.

the priority rules is that employment always takes precedence over other activities regardless of the amount of time devoted to it during the reference period, even if only one hour. As a result, unemployment is the state of total lack of work amongst those looking for and available to work.

Another key element in using the labour force framework is that the classification is based on the actual activity of the person during the specified short reference period. Thus, in general, persons are considered for inclusion in the labour force on the basis of whether they were engaged in economic activity or were looking for such an activity during the reference period. This is referred to as the activity principle. Some exceptions to the principle are allowed, however, such as the inclusion among the employed of persons who were temporarily away from work or the inclusion among the unemployed of persons waiting to start a specific

FIGURE 6
LABOUR FORCE FRAMEWORK



job they already have, at a date subsequent to the reference period and so not looking for a job during this period.

3.3 APPLICATION OF THE LABOUR FORCE FRAMEWORK

The labour force framework enables the measurement of the current situation in a country with respect to employment, unemployment and the employment characteristics of the population. The choice of one day or one week as the reference period has important consequences for the labour force status of individuals.[11] Most countries use the labour force as their preferred measure of the EAP, basing it on a one-week reference period. It is particularly suited to situations where the dominant type of employment is regular, full-time paid employment.[12]

4. THE EMPLOYED POPULATION

4.1 DEFINITION

An employed person is, broadly speaking, a person in the working age population who, during the short reference period, performed some economic activity or who would have performed such an activity had they not been temporarily absent from work. The employed population can be described as all persons in the working age population who, during a specified short reference period of one week or one day, met any one of the following conditions:

A. performed some work for wage, salary, profit or family gain, in cash or in kind;
B. were temporarily absent from a job in which they had already worked and to which they maintained a formal attachment; or
C. were temporarily absent from work in an enterprise they own (such as a business enterprise, a farm or a service undertaking), for some specific reason.

According to the 13th ICLS definition:

> … the employed comprise all persons above the age specified for measuring the economically active population who, during a specified short reference period of either one week or one day, were in the following categories:
>
> **Paid employment:**
>
> A. *at work:* persons who, during the reference period, performed some work for wage or salary, in cash or in kind;
> A. *with a job but not at work*: persons who, having already worked in their present job, were temporarily not at work during the reference period and had a formal attachment to their job;
>
> **Self-employment:**
>
> A. *at work:* persons who, during the reference period, performed some work for profit or family gain, in cash or in kind;
> A. *with an enterprise but not at work:* persons with an enterprise (which may be a business enterprise, a farm or a service undertaking) who were temporarily not at work during the reference period for any specific reason.
>
> … For operational reasons, the notion of "some work" may be interpreted as work for at least one hour.

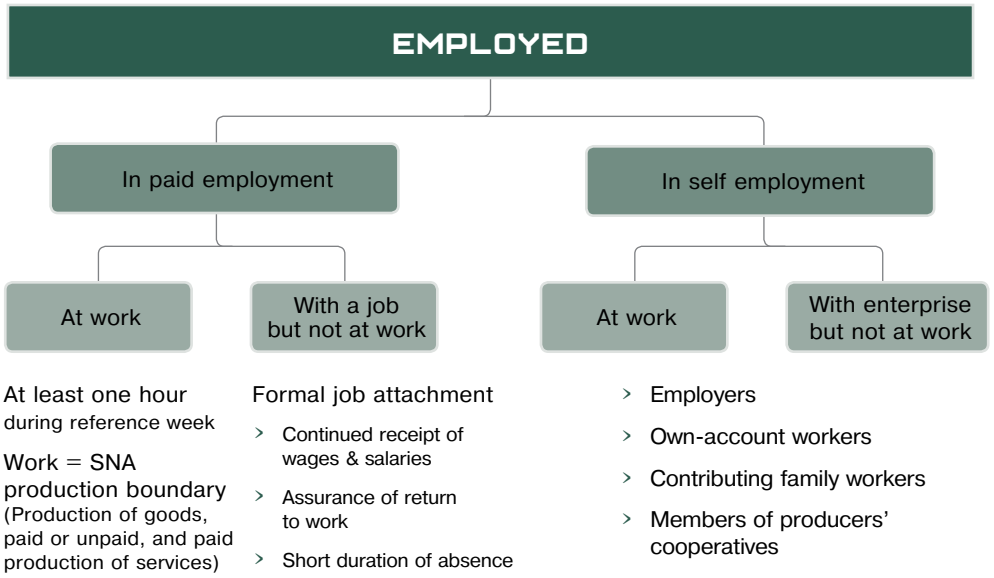
Thus, the international definition distinguishes between "paid employment (employees including apprentices and members of the armed forces) and "self-employment" (employers and own-account workers including members of producers' cooperatives, contributing family workers and producers of goods for own final use), providing separate criteria for their measurement (See Figure 7).

4.2 OPERATIONAL ISSUES

The concept of work in the definition is the same as the concept of economic activity in the SNA

11 See Hussmanns et al (1990) and UN & ILO (2009) for a full discussion of these consequences.
12 Hussmanns (2007)

FIGURE 7
INTERNATIONAL DEFINITION OF EMPLOYMENT



production boundary. Furthermore, the definition interprets the notion of "some work" as work for at least one hour during the reference period (the **one-hour criterion**). There are several inter-related reasons[13] for the use of this criterion in the definition of employment, including:

A. the priority rule of the labour force framework which gives precedence to any employment activity, even if only for one hour, over other economic activity statuses (hence unemployment is total lack of work);

B. the need to make the definition of employment as broad as possible to cover all types of employment situations such as short-term, part-time, casual and temporary work; and

C. ensuring that at the macro-level total labour input corresponds to total production in the economy.

Another operational issue arising out of the definition is the notion of **"temporary absence"** from a job or an enterprise. The 13th ICLS resolution identifies for inclusion as employed some specific groups of persons who were not at work during the reference period. These are persons who were temporarily absent from their work for reasons such as illness, injury, holiday, strike or lockout, educational or training leave, maternity or parental leave, temporary reduction in economic activity, temporary disorganisation or suspension of work due to bad weather, mechanical or electrical breakdown, shortage of raw materials or fuels, etc. or other temporary absences with or without leave.

Amongst the others who were not at work during the reference period, the international definition gives specific criteria for assessing absence from paid employment. These are based on the notion of "formal job attachment", which is to be determined in the light of national circumstances according to one or more of the following: (i) the continued receipt of wage or salary; (ii) an assurance of a return to work (with the same employer) following the end of the contingency or an agreement as to the date of return; and (iii)

13 See Hussmanns (2007) for a full discussion.

the elapsed duration of absence from the job, which may be that duration for which workers can receive compensation benefits without obligations to accept other jobs. There are, however, borderline situations in the treatment of certain categories of persons absent from work such as persons temporarily laid off, non-regular employees, seasonal employees, women on maternity leave, employees on unpaid leave initiated by the employer, employees on other extended leave such as parental leave and educational or training leave. Guidelines on their treatment were endorsed by the 16th ICLS[14] and discussed in ILO (2003) and Hussmanns (2007).

Due to the heterogeneity of working patterns amongst the self-employed, the international definition does not give specific guidelines for the treatment of persons temporarily absent from self-employment. In practice, however, the criterion used for inclusion as employed of persons temporarily absent from self-employment is the continued existence of their enterprise during their absence. The duration of absence, determined according to national circumstances, can be used to ascertain that the absence is temporary. For certain categories of self-employed workers, it can be deduced that their enterprise does not continue to exist in their absence. This is the case for casual own-account workers, like itinerant shoe-shiners or street vendors, when they are absent from work. It also applies to most seasonal employers, own-account workers and members of producers' cooperatives during the off-season. These otherwise self-employed persons should therefore not be classified as employed during the off-season. However, some enterprises do continue to exist in the off-season and their owners continue to do some work in them even during the off-season (e.g. farms which are maintained even during the off-season although the bulk of their activities are seasonal). In this case, the self-employed person not at work during the off-season could be classified as employed provided the duration of absence is within an acceptable limit. Contributing family workers, even though they participate in activities in a household enterprise, are not considered as owning or operating the enterprise. They cannot therefore be "with an enterprise but not at work". Thus, contributing family workers not at work during the reference period should not be included among the employed. Those in the above groups who are not included as employed should be classified as unemployed or not in the labour force according to their search activities and availability for work during the reference period.

4.3 SPECIAL GROUPS

The following groups of workers were given explicit attention in the 13th ICLS *Resolution* for inclusion among the employed: contributing family workers at work; persons engaged in the production of goods for own final use within their household; paid apprentices; working students, volunteer workers, homemakers, etc.; and members of the armed forces.

Contributing family workers are persons who work for family gain in an unincorporated market enterprise operated by a related person living in the same household.[15] Such persons are to be considered employed if at work in the same way as for other self-employed persons, irrespective of the number of hours worked during the reference period. Countries that choose for special reasons to set a minimum time criterion higher than one hour for the inclusion of contributing family workers among the employed should identify and separately classify those who worked less than the prescribed time.

Persons engaged in the production of goods for own final use by their household should be considered for inclusion as employed provided the production constitutes an important contribution to the total consumption of the household. The intention in the proviso is to exclude from consideration production that does not really have a significant impact on the subsistence of the household and which is done more as a leisure activity. This is similar to the macro-level exclusion of negligible production activities for

14 Guidelines concerning the treatment in employment and unemployment statistics of persons on extended absences from work, 16th ICLS, 1998, http://www.ilo.org/public/english/bureau/stat/download/guidelines/exleave.pdf

15 See Resolution concerning the International Classification of Status in Employment (ICSE), ILO (2000) pp. 20-23 for the full definition of contributing family workers and paid apprentices.

own final use from the national accounts. The implementation of the proviso in labour force surveys and population censuses is, however, problematic. Therefore, some countries use the volume of labour input into this production to assess its importance whilst others simply ignore it and include all of them as employed. As mentioned earlier, under the current proposals for revising the 13th ICLS *Resolution*, such persons will no longer be considered as employed.

Apprentices and trainees who receive pay in cash or kind should be considered in employment and treated in the same way as other persons in paid employment. The treatment of unpaid apprentices was not specified in the 13th ICLS *Resolution*, but it could be deduced on the basis of general principles. When at work they should be considered as employed if they fulfil the conditions for inclusion as contributing family workers. Otherwise, their inclusion as employed should depend on their contribution to the production of goods and services of the enterprise.

According to *Guidelines* agreed in the 14th ICLS,[16] participants in job training schemes are to be included as employed if the training takes place within the context of the enterprise and in connection with its production, irrespective of whether or not they were paid by the employer. If the training takes place outside the context of the enterprise or without connection to its production, the participants are considered as employed if they retain a formal job attachment to the enterprise in which they had formerly worked.[17]

Based on SNA 08 and the 18th ICLS, volunteers providing labour inputs for the production of goods should be included as employed, irrespective of whether the goods are intended as market outputs, non-market outputs or outputs for own final use. Volunteers contributing to the production of market or non-market services are also considered as employed. However, volunteers contributing to the production of services for own final consumption by other households are not to be considered as employed, since this production is outside the SNA production boundary.

In line with the priority rules of the labour force framework, students, homemakers, pensioners and those engaged mainly in non-economic activities during the reference period, but who were at the same time in paid employment or self-employment, should be considered as employed on the same basis as other categories of employed persons (and be identified separately, if possible).

Finally, all members of the armed forces are to be included as employed, according to the 13th ICLS *Resolution*. The statistics should include both regular and temporary members of the armed forces as well as persons performing civilian services as an alternative to compulsory military services.

5 THE UNEMPLOYED POPULATION

5.1 DEFINITION

According to the international standards adopted by the 13th ICLS:

> The unemployed population consists of those in the working age population who during the reference period were simultaneously:
>
> 1. "without work", i.e. not employed, where employment is as defined above;
> 2. "currently available to work", i.e. were available for employment during the reference period; and

16 14th ICLS (1987): Guidelines on the implications of employment promotion schemes on the measurement of employment and unemployment, Current International Recommendations on Labour Statistics, ILO (2000), pp. 86–87.

17 See Hussmanns (2007) for a full discussion.

3. "seeking work", i.e. had taken specific steps in a specified recent period to look for employment.

5.2 THE "WITHOUT WORK" CRITERION

The "without work" criterion follows from the priority rule of the labour force framework in which any employment takes precedence over other activity statuses. The criterion thus distinguishes clearly between employment and non-employment, also ensuring that the two categories are mutually exclusive. Therefore, an unemployed person has to be totally without work, i.e. did not work at all during the reference period (not even for one hour) and was not temporarily absent from work. The other two criteria serve to draw the distinction between the unemployed and those not economically active among the non-employed population.

5.3 THE "SEEKING WORK" CRITERION

The second criterion requires that specific action should have been taken in the recent past to look for work, in keeping with the activity principle of the labour force framework. The 13th ICLS *Resolution* gives as examples of specific formal steps to seek work: registration at an employment exchange; application to prospective employers; checking at work sites, farms, factory gates, market or other assembly places; placing or answering newspaper advertisements; looking for land, buildings, machinery or equipment to establish one's own enterprise; arranging for financial resources; applying for permits and licences, etc. It also suggests informal methods such as seeking the assistance of friends or relatives. With respect to registration at employment centres, the 14th ICLS further specified that this is allowable only if it is done for the purpose of finding a job and not for administrative purposes relating to the receipt of certain social benefits. Advertisements would now extend to include those on the internet and other e-media, but simply looking at them without responding to them is not considered as actively seeking work. The duration of the recent past for seeking work was not specified by the 13th ICLS *Resolution* but it is usually assumed that it could extend beyond the basic survey reference period of one day or one week. The trend is in fact to use the last month or the past four weeks as the search period.

5.4 THE "AVAILABLE FOR WORK" CRITERION

According to the international standards, to be considered unemployed a person should be available to take up employment during the reference period if such an opportunity arose. In practice, some countries extend the period to take into account that some persons seeking work are not expected to take it up immediately it is offered. They may need to make arrangements to cope with previous commitments (for example, childcare), may be sick on the start day or are required to start work on a particular date, e.g. 1st of the month. For example, the availability period for countries in the European Union is the reference week plus the subsequent two weeks. The availability criterion serves to exclude from unemployment persons who may be seeking work to start at a much later date, like students looking for work to start at the end of the school year, and persons who cannot take up work due to family responsibilities, illness, community responsibilities, and so on.

5.5 FUTURE STARTERS

The current standards give guidelines on the treatment of persons who are not employed but are not looking for work because they have already made arrangements to take up a job or start an enterprise at a date subsequent to the reference period (future starters). Such persons, if currently available for work, should be classified as unemployed even though they

did not look for work. This is one instance where an exception is made from the requirements that all three criteria should be applied simultaneously.

5.6 THE "RELAXED DEFINITION"

The current international standards recognise that in certain circumstances the "seeking work" criterion, as a means of identifying those searching for information on the labour market, may not be meaningful. They allow for the relaxation of the use of the criterion in situations:

> *"where the conventional means of seeking work are of limited relevance, where the labour market is largely unorganized or of limited scope, where labour absorption is at the time inadequate, or where the labour force is largely self-employed."*

This is likely to be the case in rural areas of developing countries where self-employment prevails and no institutions are properly developed to provide information on the availability of jobs. Moreover, most workers in these areas, as well as in agriculture, do not actively seek work at certain periods of the year since they already have a more or less complete knowledge of the employment opportunities that exist at various times of the year. Even in urban areas of these countries and in some industrialized countries, particular groups of workers may not actively seek work because they believe no work suitable to their skills is available in their areas. In these instances when the "seeking work" criterion is relaxed, the use of the "available for work" criterion becomes critical. Thus, the standards recommend that appropriate objective tests should be applied when assessing availability to work in these circumstances. The measure of unemployment under the relaxation provision is referred to as the **relaxed definition of unemployment.**

5.7 CURRENT PROPOSALS FOR REVISION

One of the proposals relating to the revision of the 13th ICLS resolution is to abandon the use of the relaxed definition of unemployment. In many countries, the production of multiple unemployment rates has caused confusion and loss of credibility in the statistics being produced as well as misunderstandings in public debates. Moreover, with the proposal to produce supplementary indicators of labour slack in the new resolution, the relaxation provision will no longer be necessary.

As mentioned earlier, another proposed revision is the exclusion of persons producing output exclusively for own final use from the employed population. Those amongst them who satisfy the above three unemployment criteria would now be included amongst the unemployed population. If their number is significant, this change could lead to a sizeable increase in the unemployed population and so create a break in the historical series of unemployment rates. Those who fail one or more of the three unemployment criteria would be placed in the population outside the labour force.

6 POPULATION OUTSIDE THE LABOUR FORCE

6.1 DEFINITION

According to the international standards adopted by the 13th ICLS:

> The population outside the labour force (or the population not currently active) comprises all persons in the population who, during the short reference period of one day or one week, were neither employed nor unemployed.

It should be noted that persons below the minimum working age specified for measurement of the EAP are included in this population.

6.2 CLASSIFICATIONS

The current international standards recommend that this population should be classified in terms of their reason for not being active during the specified short reference period, in order to assist governments in designing their human resources and development policies. The proposed groups consist of: (a) those attending educational institutions; (b) those performing household duties; (c) those who are retired or too old and are living on pension or capital income; and (d) those with other reasons. The 13th ICLS further proposed separating out those who were available for work but did not seek work during the reference period from the others in this population. The resolution also recommended the development of classifications reflecting the relative strength of attachment to the labour market of persons in this population.

The ILO WG is considering the following proposals for such a classification to be included in the new resolution:

A. Persons outside the labour force who looked for work (during the seeking work reference period) but were not available (during the availability reference period) to take up work;

B. Persons outside the labour force who wanted and were available to work (during the availability reference period) but did not seek work (during the seeking work reference period) because they were:

  I. Discouraged jobseekers: past failure to find suitable jobs, real or perceived lack of suitable jobs, real or perceived social barriers, real or perceived economic barriers, lack of infrastructure;

  II. Others: workers on lay-off without formal job attachment, workers in training, others;

C. Persons outside the labour force who wanted to work, but were neither seeking nor available for work during the respective periods for seeking and being available for work;

D. Persons outside the labour force who did not want to work.

7 EMPLOYMENT IN THE INFORMAL SECTOR AND INFORMAL EMPLOYMENT

7.1 USES

In many developing countries, especially countries in Africa, the informal sector has long played a significant role in employment creation, income generation and in the development debate. Increasingly, it is also becoming relevant in developed economies. Due to various factors, but especially in the wake of the global financial and economic crisis, employed persons are going into informal sector activities rather than into outright unemployment. Statistics are required on:

A. the number of persons employed in the informal sector, their socio-demographic and other characteristics as well as their conditions of work and employment;

B. the number and structural characteristics of the informal sector enterprises;

C. the production and incomes generated by these enterprises; and

D. other characteristics relating to the birth and death of these enterprises and their relationship with other production units.

7.2 INTERNATIONAL STANDARDS (EMPLOYMENT IN THE INFORMAL SECTOR)

Recognising the importance of the informal sector, the 15th ICLS adopted in 1993 the *Resolution concerning statistics of employment in the informal sector*.[18] According to the international standards:

> … the informal sector is regarded as a group of … household enterprises or, equivalently, unincorporated enterprises owned by households … and comprise (i) "informal own-account enterprises" … and (ii) … "enterprises of informal employers" …

where,

> … informal own-account enterprises may comprise, depending on national circumstances, either all own-account enterprises or only those which are not registered under specific forms of national legislation.
>
> … enterprises of informal employers may be defined, depending on national circumstances, in terms of one or more of the following criteria:
>
> I. size of the unit below a specified level of employment;
>
> II. non-registration of the enterprise or its employees.

Further, the *Resolution* specified that:

> The population employed in the informal sector comprises all persons who, during a given reference period, were employed … in at least one informal sector unit … irrespective of their status in employment and whether it is their main or secondary job.

Thus, employment in the informal sector is determined by the **characteristics of the enterprise in which a person is employed.** According to the above, the identification of the enterprise as an informal sector enterprise is based on its size or the registration of the enterprise or that of its employees, depending on the SE of the employed person. The Delhi Group on Informal Sector Statistics recommends that, for international reporting, a limit of 'less than five' should be used as the size cut-off.[19]

In practice some other additional criteria are used, such as the keeping of separate accounts for the enterprise and the type of workplace of the enterprise. The former comes from the conceptual definition of the informal sector and the latter from the fact that informal sector enterprises tend to operate from or close to the entrepreneur's home, from a market stall, or from some temporary location.

Operationally, persons employed in the informal sector can be identified as follows:

A. An employee or contributing family worker is employed in the informal sector if:

   I. the number of persons working in the enterprise where that person works is less than some specified cut-off (e.g. five); and

   II. either the enterprise is not formally registered (for purposes of tax, social security or some form of trade licence) or it operates in certain types of informal places such as at or near the home, in the street or in some moveable location.

B. An employer, own-account worker or member of a producers' cooperative is employed in the informal sector if the enterprise they own is not formally registered (for purposes of tax, social security or some form of trade licence) or does not have a bank account separate from the personal account of the owner.

18 Current International Recommendations on Labour Statistics, p. 32, ILO Geneva 2000.

19 See http://mospi.nic.in/report_3.htm

C. Persons employed in government and para-statal bodies, cooperatives, incorporated enterprises and quasi-corporate enterprises are excluded.

D. Also excluded are persons employed in agricultural activities, for practical reasons,[20] and those employed in households as domestic workers.

It should be noted that a person may be employed in the informal sector through their main, secondary or other jobs. For analysis linking employment in the informal sector with the characteristics of jobs, a person may be counted as many times as that person has a job in the informal sector.

7.3 INTERNATIONAL STANDARDS (INFORMAL EMPLOYMENT)

The 17th ICLS, 2003, adopted the *Guidelines concerning a statistical definition of informal employment.*[21] In them, **informal employment is determined by the characteristics of the job** the person does, unlike the previous concept of employment in the informal sector, which was based on the characteristics of the enterprise. The concept of informal employment was introduced to complement that of employment in the informal sector and in response to requests from the Delhi City Group on Informal Sector and the International Labour Conference.

According to the international guidelines:

1. Informal employment comprises the total number of informal jobs as defined ... below, whether carried out in formal sector enterprises, informal sector enterprises, or households, during a given reference period.
2. ... informal employment includes the following jobs:
   - A. Own-account workers and employers employed in their own informal sector enterprises;
   - B. Contributing family workers irrespective of whether they work in formal or informal sector enterprises;
   - C. Members of informal producers' cooperative;
   - D. Employees holding informal jobs in formal sector enterprises, informal sector enterprises, or as paid domestic workers employed by households. Employees are considered to have informal jobs if their employment relationship is, in law or in practice, 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.);
   - E. Own-account workers engaged in production of goods exclusively for own final use by their household.

Operationally, persons employed in informal jobs are identified as follows:

E. An employee is in informal employment with respect to that person's job (main, secondary or other job) if in that job the person does not have a written contract, is not entitled to paid (annual, sick or maternity) leave or does not have social security protection. Note that in this case employees in agricultural activities or those employed in private households as domestic workers are included, if their jobs have these characteristics.

F. A contributing family worker is always in informal employment.

G. An employer, own-account worker or member of a producers' cooperative is in informal employment if the enterprise they own is an informal sector enterprise.

18 Current International Recommendations on Labour Statistics, p. 32, ILO Geneva 2000.

19 See http://mospi.nic.in/report_3.htm

20 Some countries do however include persons engaged in agricultural activities if they satisfy the conditions in a) or b).

21 See http://www.ilo.org/public/english/bureau/stat/download/guidelines/defempl.pdf

8 WORKING TIME MEASUREMENT

8.1 USES

Statistics on working time (i.e. 'hours of work') are useful in the analysis of both economic and social issues. They serve to:

- derive a complete assessment of labour input into production;
- analyse employment statistics in depth, given that employment is defined as work even for just one hour in a week;
- measure time-related underemployment, an important component of labour underutilisation as well as labour productivity, wage rates, rates of occupational injuries, etc.;
- compute indicators on time-related overemployment and work–family balance, key indicators in the measurement of decent work; and
- facilitate negotiations in social dialogue relating to collective bargaining, etc.

8.2 INTERNATIONAL STANDARDS

The current international standards on the measurement of working time are the *Resolution* concerning the measurement of working time adopted by the 18th ICLS in 2008. It recognises seven concepts of working time, each serving a specific objective. They include **hours actually worked, hours paid for, normal hours of work, contractual hours of work, hours usually worked, overtime hours of work and absence from work hours.**

8.3 HOURS ACTUALLY WORKED

According to the 18th ICLS *Resolution*:

> *Hours actually worked* is 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 further defines it as follows:

> *Hours actually worked* includes …
>
> **A.** "Direct hours" … spent carrying out the tasks and duties of a job. …
>
> **B.** "Related hours" … spent maintaining, facilitating or enhancing productive activities …
>
> **C.** "Down time" … time when a person in a job cannot work due to machinery or process breakdown, accident, … but continues to be available for work. …
>
> **D.** "Resting time" is time spent in short periods of rest, relief or refreshment … according to established norms and/or national circumstances.

Related hours refer to time spent cleaning, repairing, preparing etc. the work location itself, as well as changing time. The category also includes time spent transporting goods, waiting time, time on-call, on work-related travel and on training activities required by the job. Resting time covers tea, coffee or prayer breaks. Hours actually worked excludes time not worked, such as in annual leave, public holidays, sick leave and other absences, commuting time between work and home, longer breaks such as meal breaks and other educational activities not required by the job.

8.4 HOURS USUALLY WORKED

The international standards define:

> *Hours usually worked* is the typical value of hours actually worked in a job per short reference period such as one week, over a long observation period of a month, quarter, season or year.

It can be calculated, for example, as the most frequent number of hours that a person actually worked per week during the past month or, again, as the mode of the annual distribution of weekly hours actually worked. In both instances, it is assumed the mode exists and the data is available. In most cases, however, it is determined on the basis of self-reporting of the typical value by the respondent.

8.5 OTHER HOURS CONCEPTS

The slides in the Annex at the end of Part I present the definitions of the other hours concepts, as well as the above definitions and an example illustrating all of them.

8.6 SOURCES

Labour force surveys and population censuses can be used to collect data on hours actually worked and hours usually worked. The former source is particularly suited for this and may also produce data on hours paid for and normal or contractual hours. Establishment-based surveys are well suited to collecting data on these latter and the other concepts. Administrative registers are useful for providing information on contractual hours, hours paid for, paid absence from work hours and normal hours of work.

9 MEASURES OF UNDERUTILISED LABOUR

9.1 INTRODUCTION

The unemployment rate, as a measure of total lack of work, is widely used by the media, politicians, policy-makers and so on as a barometer for the performance of the economy. It assesses the inability of the economy to provide jobs for those who do not have jobs and are looking for jobs. The indicator does not, however, reflect the economy's inability to fully and adequately utilise the available labour supply in the sense that:

I. Employed persons may experience
   - Partial lack of work;
   - Under-utilisation of their skills;
   - Low productivity in their employment; and
   - Low returns from their work.

II. Persons outside the labour force may be in a position to supply their labour for use in the economy under the right circumstances.

In situations where

A. unemployment insurance, social security schemes or other public relief schemes are limited or non-existent;

B. self-employment is prevalent; or

C. agricultural activities predominate;

such as in many developing countries, very few persons can afford to be unemployed, in the sense of a total lack of work. The majority have to engage in some economic activity, often in the informal sector or in informal jobs, however slight or inadequate they may be. The unemployment rate is then consistently so low that it loses its usefulness as a barometer of the performance of the economy.

Also, due to volatility in financial and other markets, labour market flexibility, job contraction, changes in the employment situation and the rise of various forms of non-standard employment, the unemployment rate is no longer sufficient to describe and analyse the functioning of the labour market, even in industrialized countries. Many employed persons now have to resort to jobs in which they make less use of their skills, work less or receive less income, especially in self-employment, rather than going into unemployment.

Amongst persons outside the labour force who are available to work are: (a) discouraged workers; (b) those that have stopped looking for work because they believe no work suitable to their qualifications or circumstances exist; and (c) those who could work if personal issues such as child care were resolved. These have varying degrees of marginal at-

tachment to the labour market and form part of the potential labour force. Their size and characteristics are also important in assessing the economy's ability to utilise the available labour supply.

The unemployment rate therefore does not completely portray the challenges in the employment situation and in the labour market of a country. The need to complement it with other indicators of labour underutilisation to get a complete picture of the labour market has long been recognised. International standards have so far only provided guidance for measuring the underutilisation of labour supply amongst the employed, referred to as 'underemployment'. Proposals being considered by the ILO WG for the revision of the 13th ICLS *Resolution* would extend this guidance to underutilisation of labour supply in the working age population.

9.2 UNDEREMPLOYMENT

Underemployment, in its wide sense, relates to the underutilisation of the productive capacity of the employed population. As already stated, this could be in the form of partial lack of work, low income, low productivity, inadequate use of skills, etc. of the employed population. However, the current international standards for the measurement of underemployment[22] are restricted only to the measurement of the partial lack of work of employed persons, referred to as **time-related underemployment.** The concept of underemployment adopted by the 16th ICLS reflects the situation in which persons who worked or were temporarily absent from work were available and willing to work additional hours, thus complementing the statistics on employment and unemployment.

With respect to the other aspects of underutilisation of the capacity of the employed population, the 16th ICLS recognised their importance to complement the measurement of time-related underemployment. Whilst identifying possible indicators that could be used in their measurement and providing some guidelines for doing so, the *Resolution* recommended that further development was required on their statistical definitions and methods. As a result, no international standards were adopted on their measurements. These other types of underutilisation of the productive capacity of the employed population are referred to as **inadequate employment situations.** They describe situations in the workplace which reduce the capacities and well-being of workers, as compared to an alternative employment situation they are willing and able to engage in.

9.2.1 Time-related underemployment

According to the international standards:

> Persons in time-related underemployment comprise all persons in employment (as defined above) who, during the reference period used to define employment, satisfied simultaneously the following three criteria:
>
> **A.** Were willing to work additional hours;
>
> **B.** Were available to work additional hours;
>
> **C.** Had worked less than a threshold relating to working time.

The definition applies to all employed persons, including the self-employed and those temporarily absent from work. It should be noted that the reference is to work on all jobs by an employed person during the reference period.

In the *Resolution*, the notion of **willingness to work additional hours** translates into wanting to increase hours of work by:

A. having another job (or jobs) in addition to their current job (or jobs);
B. replacing any of their current jobs with another job (or jobs) with increased hours of work;
C. increasing their hours of work in any of their current jobs; or
D. a combination of the above options.

22 16th ICLS (1998): Resolution concerning the measurement of underemployment and inadequate employment situations, Current International Recommendations on Labour Statistics, ILO Geneva, 2000, pp. 29–31.

There is no requirement for the employed person to have actively sought these additional hours, but it is recommended that those who did so should be separately identified for analytical purposes.

The **availability** criterion requires the employed person to be ready (i.e. have the time to do so), within a specified subsequent period, to work additional hours if given the opportunity to do so. The choice of a specified subsequent period is according to national circumstances, but it should comprise the usual period required for someone to change jobs. The opportunity to work should be interpreted within the context of how the worker wishes to work additional hours. For example, if the desire is to do more hours within a current job, the opportunity must be with regard to the possibility of doing so. It should be noted that, in practice, respondents have experienced some difficulties in differentiating between "willingness" and "availability." Some countries have therefore resorted to combining them in one question on "willingness and availability".

The third criterion, **working less than a threshold**, relates to an employed person whose "hours actually worked" in all jobs during the reference period is less than a threshold, determined according to national circumstances. The definition of "hours actually worked" which is applied should be that in the current international guidelines, which at present are those adopted by the 18th ICLS.[23] Suggestions in the *Resolution* for choice of the threshold are: (i) the boundary between full-time and part-time employment; (ii) the norm established by national legislation, collective agreements, etc.; or (iii) the median or average values. The purpose of the threshold is to exclude from the time-related underemployed population those employed persons who already can be considered to have attained full-time employment in terms of the duration of their work, even if they were willing and available to work additional hours. They cannot be considered as suffering from a partial lack of work and their underemployment may be due to other factors such as low hourly income. The 16th ICLS *Resolution* nevertheless recommended that data should be collected on all workers who were willing and available to work additional hours, irrespective of the hours they actually worked during the reference period (see Hussmanns (2007) and ILO (2003) for further discussion of these issues).

The use of "hours actually worked" during the short reference period instead of "hours usually worked" is in contention by some experts. The former choice includes both persons in involuntary part-time work (persons who usually work part-time and are willing and available to work additional hours)[24] and those who, during the short reference period, had to work less than their normal hours as an exceptional situation (e.g. those temporarily absent from work). The use of usual hours leads to limiting the underemployed to the first group, involuntary part-time workers. The European Union has decided to base their definition of underemployment on usual hours, specifying that the population of interest to them are the underemployed part-time workers. Usual hours, however, pose data-collection challenges associated with its reliable and accurate measurement in household surveys. The 16th ICLS *Resolution* opted for actual hours but recommended identifying these two groups separately.

9.2.2 Inadequate employment situations

As mentioned earlier, the 16th ICLS *Resolution* did not adopt definite standards for use in measuring inadequate employment situations. However, it did propose for consideration by countries the following definition:

> Persons in inadequate employment situations comprise employed persons who, during the reference period, wanted to change their current work situation, or (particularly for the self-employed) to make changes to their work activities and/or environment, for any of a set of reasons, chosen according to national circumstances.

23 18th ICLS (2008): Resolution concerning the measurement of working time, See http: http://www.ilo.org/public/english/bureau/stat/download/guidelines/

24 In fact, not all involuntary part-time workers may be included as some could have actually worked more hours than the threshold during the reference week due to special circumstances.

Examples of reasons given in the 16th ICLS *Resolution* include: inadequate use and mismatch of occupational skills; inadequate income in the current job(s); excessive hours of work; precariousness of the job(s); inadequate tools, equipment or training for the assigned tasks; inadequate social services; difficulties in travelling to work; variable, arbitrary or inconvenient work schedules; recurring work stoppages due to delivery failures of raw materials or energy; prolonged non-payment of wages; and long overdue payment from customers. These reasons are neither mutually exclusive nor exhaustive.

Workers may want to change their current work situation during the reference period in one or both of the following ways:

A. replace any of their current jobs for another one; and/or

B. carry out changes in any of their current jobs (e.g. by reorganising the work; upgrading tools, machinery or equipment; diversifying the type of products produced; developing innovative marketing strategies; etc.)

The *Resolution* further suggested two additional criteria that may be used in identifying employed persons in inadequate employment situations:

> … Workers availability to change their current situation, as well as their active job search, as understood in the definition of time-related underemployment, may also be applied.

The use of these additional criteria brings some objectivity into the measurement, as the desire for changing current work situations on its own could lead to rather subjective results. The availability criterion should thus be understood in terms of 'being in a position to do so', either in terms of time or of current capacity. Most countries collecting data on inadequate employment situations in fact apply them.

The 16th ICLS identified three particular types of inadequate employment situations that countries may wish to consider:

A. **Skill-related inadequate employment:** persons who were willing or seeking to change their current work situation during the reference period in order to use their current occupational skills more fully and were available to do so.

B. **Income-related inadequate employment**: persons who were willing or seeking to change their current work situation to increase their income by increasing the levels of work organisation or productivity, by improving tools and equipment, training or infrastructure, and were available to do so.

C. **Inadequate employment related to excessive hours:** persons who were willing or seeking to change their current work situation to work fewer hours, either in the same job(s) or in another job, with a corresponding reduction of income.

In applying (b), it may be useful to set an income threshold, based on national circumstances, above which a person does not qualify to be included. Similarly, in applying (c), an hour's threshold below which persons do not qualify for inclusion could also be set.

The identification of employed persons as belonging to one or more of the above groups is based on their current capacities and desire to change their current work situation for one or more of the above reasons, as expressed by the employed persons themselves. This micro labour supply approach introduces some subjectivity into the measures of inadequate employment.

An alternative approach, being considered by the ILO WG for the indicators of labour underutilisation in the new resolution, is the macroeconomy level approach of identifying persons in these inadequate employment situations through a more 'objective' basis. For example,

skills-related inadequate employment would be determined by identifying all employed persons who are in jobs for which they are over-skilled. The method proposed is to compare their skill level, as measured by their educational attainment level, with the skill level required for the job, as measured by its level in the occupational classification. Similarly, the income-related inadequate employment situation will be determined by comparing the person's earnings with a threshold, such as one-third of the median income value (i.e. low pay criterion). The excessive hours situation can be objectively determined by comparing the hours actually worked with some appropriate cut-off, similar to the measurement of time-related underemployment. In all instances, there will be no need to know whether the employed persons want to change their work situation, were available to or were seeking to do so.

9.3 LABOUR SLACK AND OTHER FORMS OF LABOUR UNDERUTILISATION

As already mentioned, the ILO WG is developing proposals for indicators of labour underutilisation that go beyond the labour force to the working age population. One such is a measure of labour slack which would encompass the persons in time-related underemployment, those in unemployment and those persons outside the labour force who:

A. were seeking work, but not available for work within the availability reference period;

B. were available to work but not seeking work during the seeking work reference period because they were discouraged, as described above in section 6.2; and

C. possibly some others yet to be determined.

EUROSTAT has decided to produce measures of labour underutilisation on the above basis using only the criteria in a) and b) for those outside the labour force.

Development work is continuing in the ILO to come up with a proposal for a resolution to be adopted by the 19th ICLS that would include the above indicator of labour slack (insufficient volume of work). These proposals would also include the indicators on skill underutilisation and low pay discussed above in section 9.2, as measures of other forms of labour underutilisation. There may be overlaps, in the sense of an employed person belonging to two or more of these different types of labour underutilization, and these would need to be resolved if the intention is to have a single indicator of labour underutilisation to complement the unemployment rate. Another challenge would be which denominator to use when computing the indicator(s), with the choices being between the size of the labour force, that of the working age population or some number in between.

10. MEASUREMENT OF LONG-TERM EMPLOYMENT SITUATION (USUAL ACTIVITY)

10.1 INTRODUCTION

The labour force measurement framework produces statistics on the employment situation of individuals and that in the economy during a short reference period, such as one week. These statistics thus present a snapshot of the employment situation at a 'current' point in time, i.e. the reference week. Almost certainly, a one-off measurement of this type will be inadequate for the description and analysis of the long-term employment situation of a substantial number of individuals and that in the economy. Admittedly, for many persons, their employment situation remains relatively stable over such long periods and so the picture obtained from such a one-off survey should persist throughout the longer period. There are, however, a substantial number of persons whose employment situation may change over such a long period. Their labour force status and/or the characteristics of their job(s) could have changed during this period,

which when aggregated over all individuals could also apply to the economy. In economies with significant seasonal sectors such as agriculture or tourism, these changes will occur as a result of the seasonality in production. Even in economies where seasonality is not prominent, changes occur for a variety of other reasons such as financial or economic crisis, labour market flexibility, demographic factors, job contraction and the rise of various forms of non-standard employment. This is evidenced, for example, through the statistically significant changes in monthly or quarterly unemployment rates in those countries where they are measured as often.

Statistics on the long-term employment situation in the economy, when summarised over the long period, are useful for economic analysis, projections, manpower and development planning. They are particularly valuable as a basis for national account estimates of labour input into production and for cross analysis with other national account aggregates. Statistics on the long-term employment situation of an individual, when summarised over the long period, are important for social analysis in the areas of poverty and social exclusion, and for formulating policies in social areas such as education and training. They fit in with other social statistics determined on an annual basis such as household income and school enrolment. Statistics on the long-term employment situation of individuals and that in the economy are also key in the analysis of seasonal patterns in employment statistics, flows in and out of labour force categories, duration of stay in each category, etc.

10.2 SOURCES

Statistics on the long-term employment situation in the economy can be obtained through frequent labour force surveys over the long period. For example, these can be quarterly, monthly or continuous surveys. The data collected can be used to study changes across the different periods, such as changes in unemployment rates, or summarised (aggregated or averaged) across the different periods to produce annual estimates such as annual employment. These statistics can also come from spreading the labour force survey sample over a one-year period. In some instances, these sub-samples are nationally representative and so capture significant seasonal patterns of activities or substantial labour force movements during the year. When frequent labour force surveys also use sample rotation schemes, they generate data on the long-term employment situation at the individual level. This is the most reliable way of producing statistics for analysing flows in the labour market, seasonality and duration of stays in different labour force categories.[25] When summarised over the long period for each individual, statistics on the employment situation of that individual can then be analysed along with other personal characteristics of that individual and those of the person's household. Statistics on the employment situation in the economy are then obtained by summarising the data over the long period and over individuals.

A growing number of countries, especially those with developed economies, are implementing frequent labour force surveys, on a monthly or continuous basis, with a rotational sample. The challenges they face are analytical, in the sense of deciding which form of summarisation to do and, if studying changes, determining the extent to which observed changes are statistically significant and not just noise. Others amongst them implement frequent labour force surveys, but with different samples. These are used to summarise the employment situation or study changes in it at the level of the economy. Very few developing countries, even less so in Africa, are in a position to implement frequent labour force surveys with sample rotation. Some carry out labour force surveys with the sample spread over the year, which can then be used to produce statistics on the employment situation at the level of the economy.

The majority of countries in Africa carry out either regular but infrequent labour force surveys, such as every three or five years, or ad hoc labour force surveys. In settings such as these, where it is not possible to implement frequent labour force

25 Countries that have a reliable register-based statistical system may also be able to produce these statistics from their registers.

surveys or spread the labour force survey sample over time, some other method is necessary to generate statistics on the long-term employment situation of individuals and that in the economy. The alternative method is to generate the data required through retrospective measurement over the long period.

10.3 USES OF THE STATISTICS

Data generated from surveys or from retrospective measurement on the employment situation of individuals during each time period (e.g. each month) over a long reference period can be used to study the changes that occurred in their labour force status, if any. Alternatively, these data can be summarised:

C. for each individual across these time periods to determine and analyse an annual 'picture' of the person's employment situation; and/or

D. across individuals for the annual employment situation in the economy.

In a), it may be useful to be able to classify a person over the long period as:

I. 'somehow economically active' or 'somehow not economically active', including amongst the former those who are 'somehow employed' and those who are 'somehow unemployed';

II. 'mostly employed', 'mostly unemployed' or 'mostly not economically active'; or

III. 'ever employed' or 'never employed'; and so on.

In b), one can, for example, compute:

I. average annual employment by averaging the number of employed persons each time period over the long reference period; and

II. total labour input into production by summing the number of hours worked each time period across individuals and time periods; and so on.

10.4 MEASUREMENT FRAMEWORKS

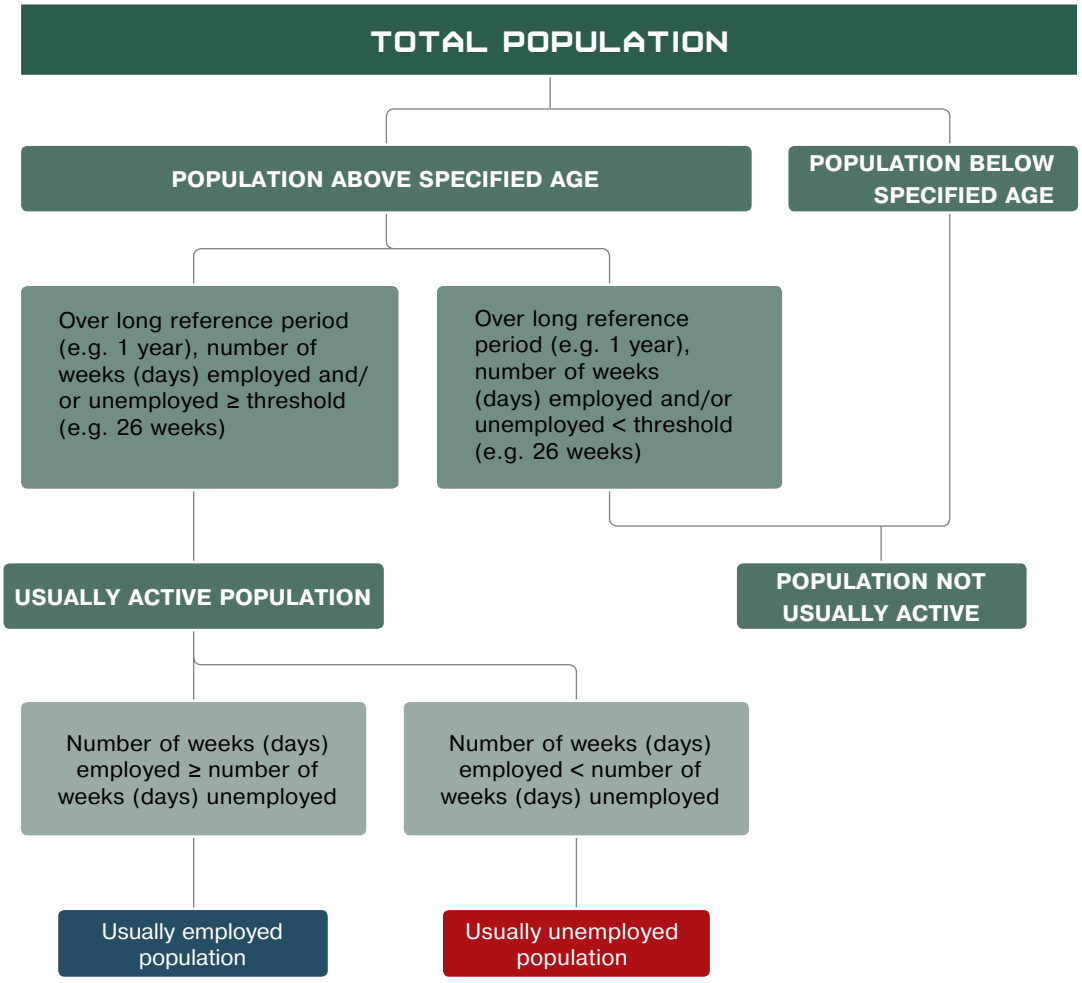
In recognition of the need to produce some information about the long-term employment situation in the absence of regular frequent labour force surveys, the 13th ICLS *Resolution* introduced the **'usually active measurement framework'** (see Figure 8). According to these international standards:

> The usually active population comprises all persons above a specified age whose main activity status as determined in terms of number of weeks or days during a long specified period (such as the preceding 12 months or the preceding calendar year) was employed or unemployed.

The above definition introduces the notion of 'activity status', which is determined over a short reference period such as a week or a day in terms of being 'economically active (employed or unemployed)' or 'not economically active'. This could be done on the same basis as the labour force framework. The 'main activity status' is then determined based on a summation of the various activity statuses of the individual during the preceding 52 weeks or 365 days of the 12-month reference period. The use of the main activity status in this framework differs from the labour force framework priority criterion, which accords priority to any economic activity over activities that are not economic. Although not explicitly included in the standards, the procedure mostly followed by those countries using the usual activity framework is to interpret 'main activity status' in terms of the **majority criterion.** That is, to classify a person as usually active if the number of weeks of being in the 'active status' is not smaller than the number of weeks of being in the 'not active' status. Another possibility discussed in Hussmanns et al (1990) is to classify a person as usually active if the number of weeks of being active is not less than some cut-off point such as the length of the tourist season.

As for the labour force framework, the classification of individuals could differ depending on whether it is based on weeks or on days as

FIGURE 8
USUALLY ACTIVE MEASUREMENT FRAMEWORK



Source: UN & ILO (2010), Figure IV, p. 85.

the unit of measurement. When employment is mostly of a regular and continuing nature and hence a week of employment usually means a week of full-time employment or of employment for the major part of the working time, it is recommended to base the main activity status on weeks. Otherwise, days are better. However, most countries using this framework base it on weeks.

The usually active population can be further subdivided into the usually employed population and the usually unemployed population according to the main activity during the active period, i.e. whether they were mainly active employed or mainly active unemployed. Thus, a usually active person is **usually employed** if during the active period the number of weeks of employment is not less than the number of weeks of unemployment. The usually active person is **usually unemployed** if the contrary holds. This is another difference between the labour force framework and the usually active framework. In the former, persons are first classified as employed or unemployed and then the two sub-populations are merged to get the labour force. In the latter, it is the reverse.

According to international standards,

> The **'population not usually active'** comprises all persons whose main activity status during the longer specified period was neither employed nor unemployed.

It is further recommended that this population is classified into the following four groups: *students; homemakers; pension or capital income recipients;* and *others*.

There are some other measurement frameworks for long-term employment situations, although these are not included in the international standards. One such is the **predominant activity framework.** In this, a person is classified according to the number of weeks (or days) spent in each of the three activity states (employed, unemployed or not active) during the long reference period of 12 calendar months or one calendar year. The classification is done using the majority criterion. Thus, the person is classified as mainly employed, mainly unemployed or mainly not active depending on which of these three states the person spent the longest time during the 12 months or one calendar year. The person is then classified as mainly active if they were either mainly employed or mainly unemployed, in a similar way to the labour force framework.[26] Another framework discussed in Hussmanns et al (1990) is **population economically active at some point during the year,** which refers to all persons above the specified minimum age who experienced at least one week (or one day) of employment or unemployment in the course of the reference year.

10.5 EXAMPLE

A person during the course of the reference year was employed for 12 weeks, unemployed for 17 weeks and not economically active during the remaining 23 weeks.

Usually active framework: The person is first classified as usually active since the total number of weeks of being active (employed or unemployed) is 29 weeks, which amounts to more than the 23 weeks of not being active. Since the person is usually active, they can then be classified as usually unemployed, as the number of weeks of unemployment during the active period of 29 weeks is 17 weeks, which is larger than the number of weeks of employment (12 weeks). Thus, the person is **usually active, unemployed.**

Predominant activity framework: Of the three states of employment, unemployment and not being active, the person spent the largest amount of time (23 weeks) in being not active. So the person is classified as **mainly not active.**

The person was **economically active at some point during the year** since they had 29 weeks of employment/unemployment.

10.6 DATA COLLECTION

In theory, the data needed for any of the above measurement frameworks can be generated through frequent labour force surveys based on rotational samples without the use of retrospective questioning. In practice, the data is generated from the other types of labour force surveys using retrospective measurement. When retrospective measurement is based on the recall memory of individuals, experience has shown that it could be subject to substantial recall errors, especially for those individuals with unstable employment situations. This poses a major measurement challenge which has been addressed using various methods, but never completely satisfactorily.

For the usual activity framework, the following methods are used:

A. Directly asking respondents about their main activity status during the long reference period – simple, but most unreliable as subject to misunderstanding of the concept of main activity status as well as recall errors;

B. Asking respondents about their number of weeks (or days) of employment or of unemployment during the long reference period (referred to as **whole year recall**) – unreliable due to recall errors of the exact numbers over such a long period; and

C. Asking respondents the number of half-months, weeks or days of employment or

of unemployment, for each month of the reference year (**month-by-month recall**) – less recall errors but more complex questionnaire.

In a), persons are classified directly from their responses. In b) and c), responses have to be further processed to classify the person. With c), there is also the possibility of producing statistics on flows and on seasonal variations.

For the predominant activity framework, a direct question on which activity state (employed, unemployed or not active) was predominant during the reference period is usual. Although still subject to recall errors, the risk of this should be less than for the usual activity framework. Countries collecting statistics on long-term employment situations through population censuses tend to favour this approach.

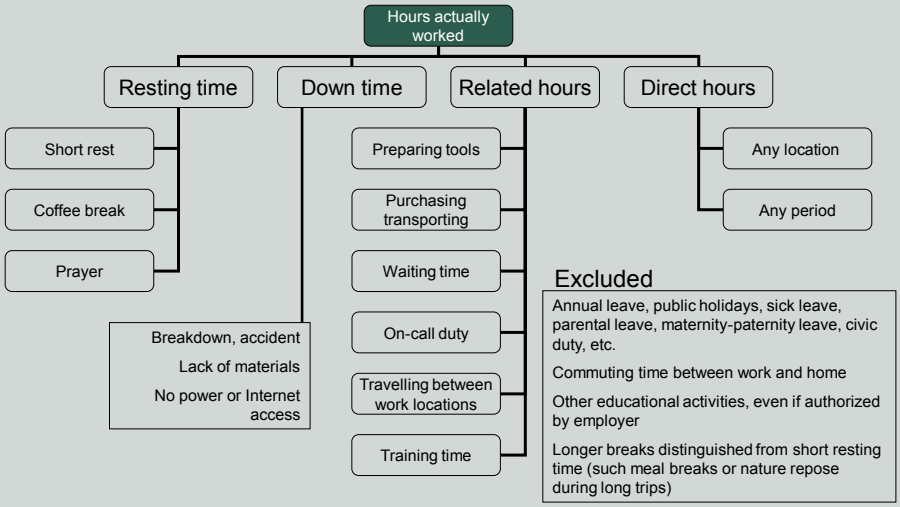
10.7 FUTURE PROPOSALS

Due to the limited quality of the data from retrospective questioning and the fact that a growing number of countries are now implementing frequent labour force surveys, the ILO WG is examining a proposal to exclude these long-term frameworks from the new standards. However, as argued above, there are still many African countries that do not do such surveys and yet need some information on their long-term employment situation. They would still need some guidance on how best to do so, even if with a proviso that the statistics should only be considered as indicative due to the data collection challenges.

26 See, UN & ILO, 2009, p. 83.

ANNEX: SLIDES ON HOURS CONCEPT

HOURS ACTUALLY WORKED



HOURS PAID FOR

- In paid employment jobs

  Time for which payment received by employer (at normal or premium rates, in cash or in kind) regardless of whether the hours actually worked or not

  Includes: paid annual leave, paid public holidays, paid sick leave

  Excludes: Unpaid overtime, unpaid educational leave, maternity leave unpaid or paid through social security systems

- In self-employment jobs

  Hours paid on the basis of time units

  Hours paid for equivalent to hours actually worked

ANNEX: SLIDES ON HOURS CONCEPT

NORMAL HOURS OF WORK

- Hours of work fixed in paid employment jobs by or in pursuance of
  - laws and regulations
  - collective agreements
  - arbitral awards
- Also applies to self-employment jobs
  - when the hours fixed for all jobs in a specified industry or occupation (such as for drivers to ensure public safety)

CONTRACTUAL HOURS OF WORK

Time expected to be performed

- according to contract for a paid-employment job
- for provision of goods/services in a self-employment job (or volunteer job)

Comparison with normal hours of work

- may be equivalent to or established in conformity with prevailing normal hours of work
- may be above normal hours (e.g., contractual overtime hours)

HOURS USUALLY WORKED

Definition

Hours actually worked in an activity during a typical week

Calculation

Modal value of hours actually worked per week over a long period

OVERTIME HOURS OF WORK

- Applies to all jobs
  - Hours stipulated as overtime in contract
  - Hours actually worked in excess of hours usually worked if no contract
- Excludes
  - Overtime as result of rotation
- Overtime hours of work
  - Paid
  - Unpaid

ABSENCE FROM WORK HOURS

- Applies to all jobs
  - Contractual hours of work not actually worked if contractual hours exist
  - Hours usually worked but not actually worked when contractual hours do not exit
- Excludes
  - time not worked as result of established work arrangements (such as flexi-time off or shift work)
- Absence from work hours
  - Paid or unpaid
  - Initiated by worker or by employer

HOURS OF WORK: NUMERICAL EXAMPLE

Reference week

Contractual hours of work							Σ
M	T	W	T	F	S	S	
9:00	9:00	9:00	9:00	9:00	-	-	
17:00	17:00	17:00	17:00	17:00	-	-	
8	8	8	8	8	0	0	40

Hours actually worked							Σ
M	T	W	T	F	S	S	
9:00	9:00	9:00	Leave		9:00	-	
18:00	18:00	17:00	Leave		13:00	-	
9	9	8	0	0	4	0	30

Hours usually worked							Σ
M	T	W	T	F	S	S	
9:00	9:00	9:00	9:00	9:00	9:00	-	
17:00	17:00	17:00	17:00	17:00	13:00	-	
8	8	8	8	8	4	0	44

Hours paid for							Σ
M	T	W	T	F	S	S	
9:00	9:00	9:00	Leave		9:00	-	
18:00	18:00	17:00	Leave		13:00	-	
9	9	8	8	8	0	0	42

2 hrs paid overtime

16 hrs absence from work

Saturday work not paid: 4 hrs unpaid overtime

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,[27] 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 1
EMPLOYMENT BY INDUSTRY (SOUTH AFRICA, FIRST QUARTER 2008)

	Jan–Mar (2008)	
Industry	Number ('000)	% share
Agriculture	799	5.9
Mining	333	2.4
Manufacturing	1,988	14.6
Utilities	95	0.7
Construction	1,112	8.2
Trade	3,156	23.2
Transport	747	5.5
Finance	1,667	12.2
Community and social services	2,564	18.8
Private households	1,163	8.5
Total	**13,623**	**100.0**

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.

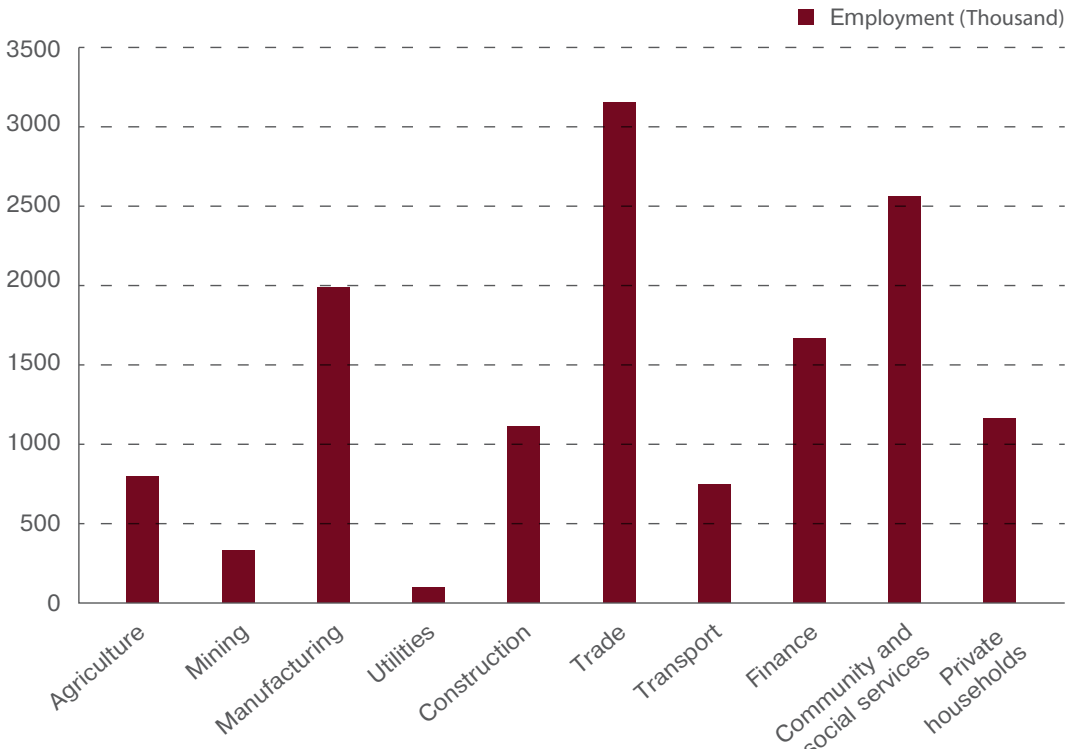
12.1.2 Single variable: Using summary statistics

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

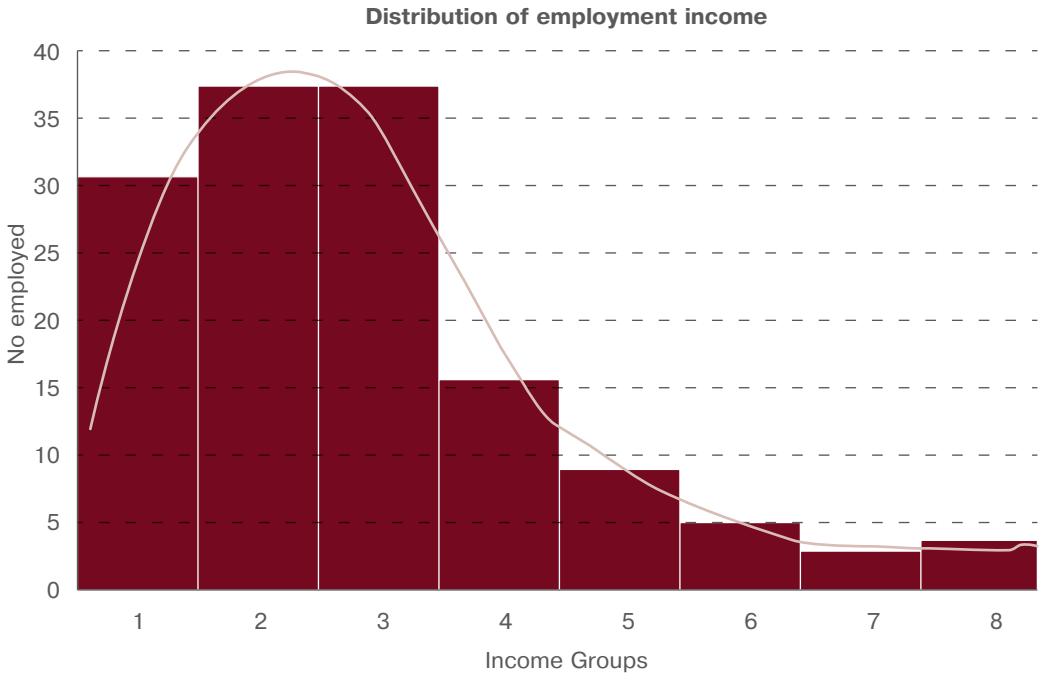
EMPLOYMENT BY INDUSTRY (1ST QTR 2008)



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

FIGURE 2

HISTOGRAM AND DENSITY CURVE OF AN INCOME DISTRIBUTION (ARTIFICIAL DATA)



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

› 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

D. Data on annual unemployment rates of a selection of countries in 2002[28] yield the following summary statistics of the distribution of annual unemployment rates across countries: Mean = 9.5%; SD = 6.7%; 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).

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

Mean = 3.99 days; SD= 1.56 days; 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

28 Source: LABORSTA, ILO statistical database.

29 Source: Based on Labour force survey, Switzerland, 2004.

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 2

AGE DISTRIBUTION BY LABOUR FORCE STATUS (SOUTH AFRICA, FIRST QUARTER 2008)

	Population (in thousands)					
Age group	Employed		Unemployed		Outside labour force	
(years)	Number	%	Number	%	Number	%
15–24	1,657	12.2	1,418	33.8	6,818	53.3
25–34	4,645	34.1	1,671	39.9	1,953	15.3
35–44	3,604	26.5	700	16.7	1,186	9.3
45–54	2,589	19.0	308	7.3	1,268	9.9
55–64	1,128	8.3	95	2.3	1,569	12.3
Total	**13,623**	**100**	**4,192**	**100**	**12,794**	**100**

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

TABLE 3

SUMMARY STATISTICS OF THE AGE DISTRIBUTIONS BY LABOUR FORCE STATUS (SOUTH AFRICA FIRST QUARTER 2008)

	Employed population	Unemployed population	Population outside labour force
Mean age	37.7	30.4	31.3
St. Dev.	11.4	10.0	14.5
CV	0.30	0.33	0.46
Skewness	-0.3	-0.9	-0.9

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

	Jan–Mar (2008)		Apr–Jun (2008)		Qtr to Qtr
Industry	Number ('000)	% share	Number ('000)	% share	Change ('000)
Agriculture	799	5.9	790	5.8	-9
Mining	333	2.4	346	2.5	13
Manufacturing	1,988	14.6	1,968	14.3	-20
Utilities	95	0.7	97	0.7	2
Construction	1,112	8.2	1,138	8.3	26
Trade	3,156	23.2	3,105	22.6	-51
Transport	747	5.5	774	5.6	27
Finance	1,667	12.2	1,687	12.3	20
Community social services	2,564	18.8	2,635	19.2	and 71
Private households	1,163	8.5	1,185	8.6	22
Total	**13,623**	**100.0**	**13,729**	**100.0**	**106**

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

TABLE 5

EMPLOYED POPULATION BY OCCUPATION AND SE ('000S) (US, 2006)

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

Source: LABORSTA, Web database, ILO

TABLE 6

CONDITIONAL DISTRIBUTIONS OF SE GIVEN OCCUPATION (US, 2006)

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

Source: LABORSTA, Web database, ILO

TABLE 7

CONDITIONAL DISTRIBUTIONS OF OCCUPATION GIVEN SE (US, 2006)

		SE			
Occupation		Employers & Own-account workers	Employees	Unpaid family workers	All workers
Legislators, senior officials & managers		22.8	13.7	5.7	14.4
Profs., Technicians & Assoc. profs.		15.3	20.1	1.9	19.8
Clerks		3.3	14.3	35.8	13.5
Service workers & shop and market sales workers		33.3	28.0	28.3	28.4
Skilled agricultural and fishery workers		0.6	0.7	15.1	0.7
Craft & related trade workers; Plant & machine operators		24.7	23.2	13.2	23.3
Total	**%**	**100.0**	**100.0**	**100.0**	**100.0**
	Numbers	**10,879**	**139,793**	**106**	**150,778**

Source: LABORSTA, Web database, ILO

ticular 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)

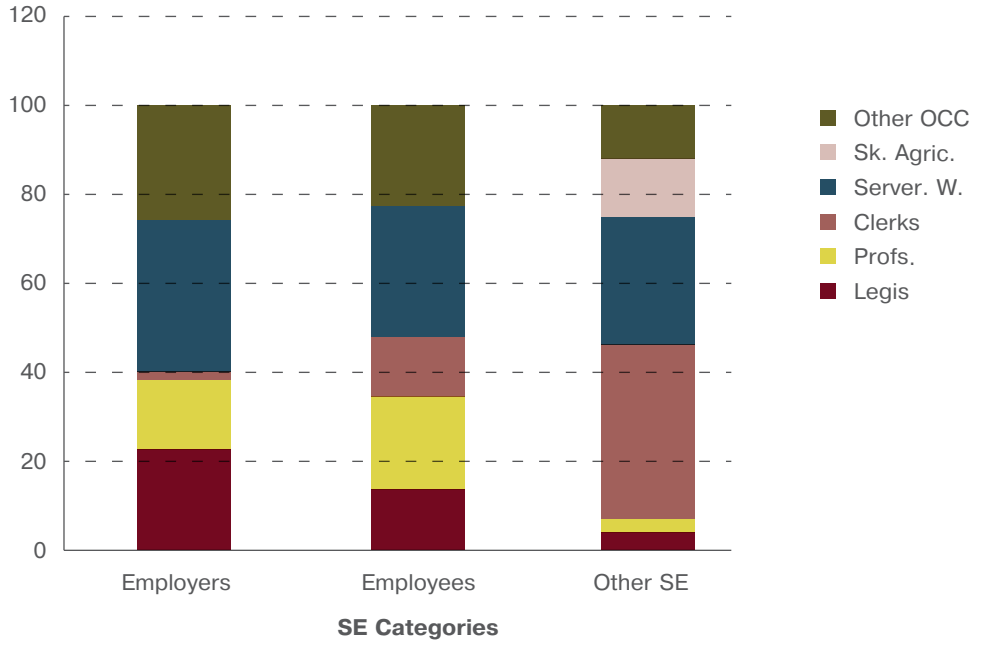
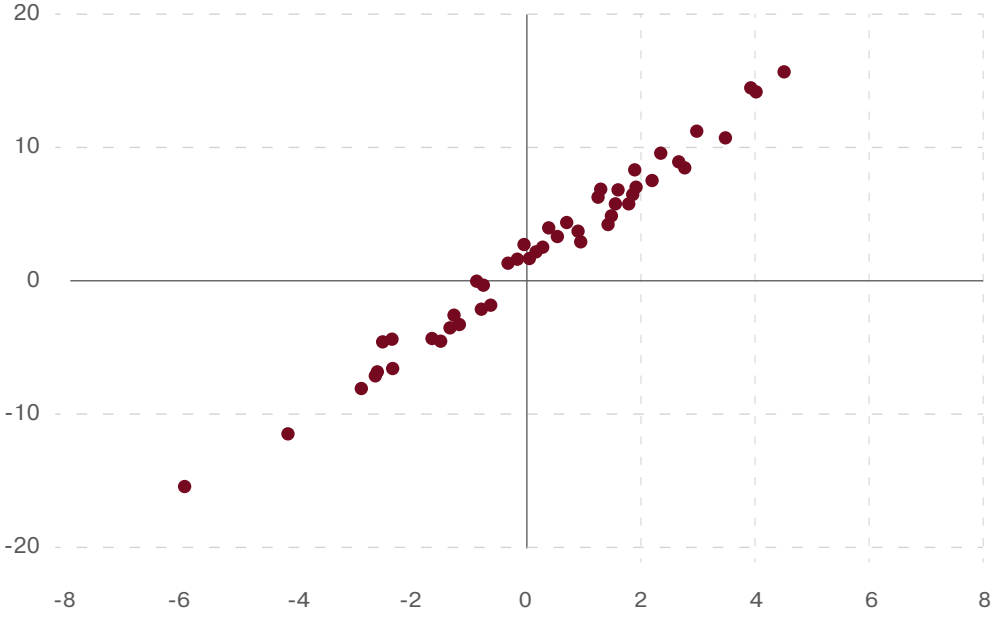


FIGURE 4

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 (& medians) 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 8
SUMMARY STATISTICS FOR CONDITIONAL DISTRIBUTIONS OF INCOME BY INDUSTRIAL SECTOR

	Agric	Industry	Services	Others
Mean	1,702.0	1,378.1	5,147.2	21.7
Median	301.2	355.7	1,596.7	2.0
Standard Deviation	4,150.8	2,102.7	7,543.4	65.7
CV	2.44	1.53	1.47	3.03
Skewness	4.4	2.1	2.1	4.7
Minimum	0.3	10.5	69.2	0.0
Maximum	26,369.3	8,070.0	31,798.0	410.0
1st Quartile	69.2	97.8	518.9	0.1
3rd Quartile	680.7	1,411.5	5,909.8	7.7

age and sex, with income as the endogenous variable.

C. 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 the 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 (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:

- Value 0 means no linear relationship whatsoever;
- Value 1 means perfect linear relationship (i.e. an exact mathematical relationship, e.g. Circumference of circle = 2 times diameter); and

› The closer the magnitude is to 1, the stronger the relationship; the closer to 0, the weaker the relationship.

The sign shows if the linear relationship is positive/direct (+) or negative/inverse (-). In the former, the values of the two variables tend to increase (or decrease) together. In the latter, the values tend to go in opposite directions (i.e. one tends to increase as the other decreases or vice versa).

Example 14

A. Values of 'Income' and 'Age' were obtained from employed persons in a survey. To investigate the relationship between the two variables, the correlation coefficient between them was computed. The value obtained is r = .65. This implies that Income and Age are directly related (positive value of r), i.e. as Age increases Income levels also tend to increase. The magnitude of 0.65 is fairly strong, indicating a close linear relationship between income and age.[30]

B. The correlation coefficient between the 'Height above sea level' and the 'Temperature' of selected sites in a country was computed as -0.97. The negative sign indicates that 'Height above sea level' and 'Temperature' are negatively related, so as height above sea level increases, temperature tends to drop. The very high magnitude of 0.97 implies a very strong, almost exact, linear relationship between these two variables.

Note that a value of 0 for the correlation coefficient does not necessarily mean there is no relationship between the variables. Such a value could be obtained for two variables with perfect curvilinear relationship, e.g. an upturned U when the scatter diagram is plotted. It only signifies that there is no linear relationship.

12.2.6.2 Both variables nominal/ordinal

When both variables are ordinal or nominal (e.g. educational attainment, occupation, SE), 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 version 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 * r2. We say the 'Regression' explains (100 * r2) % 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 = 3 X + 1.5.

- The correlation coefficient r = 0.99 and so the regression equation explains (100 * r2) = 98% 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-liner

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

31 If m = 0, knowing X does not give any information about Y. So Y does not depend linearly on X. Note there could still be another form for the dependence of Y on X even if m is 0. For example, the relationship could be curvilinear.

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 9

CONDITIONAL DISTRIBUTIONS OF OCCUPATION GIVEN SE AND SEX

Males

	Status in Employment			
Occupation	Employers	Employees	Others	Total
Leg.	25.5	14.5	2.4	15.3
Profs	13.6	16.8	7.1	16.4
Clerks	0.8	6.7	9.5	6.2
Service W.	23.6	24.2	38.1	24.1
Sk. Agric.	0.8	1.0	19.0	1.0
Others	35.7	36.8	23.8	36.5
Total	100.0	100.0	100.0	100.0

Females

	Status in Employment			
Occupation	Employers	Employees	Others	Total
Leg.	17.7	13.1	5.4	13.3
Profs	19.1	24.9	1.1	24.5
Clerks	6.9	22.4	52.7	21.5
Service W.	50.5	32.1	24.7	33.0
Sk. Agric.	0.2	0.3	9.7	0.3
Others	5.5	7.2	6.5	7.1
Total	100.0	100.0	100.0	100.0

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.

Whilst the Services sector has the highest mean income for both sexes, the second highest for males is in Industry whilst that of females is Agriculture. The pattern of dispersion for the income distributions is also different between males and females and between the industrial sectors. The same holds for skewness, with the income distribution for females in Industry the most positively skewed distribution whilst for males it is in Agriculture. The conditional distributions of industrial sector for given sex values are also different. We can conclude that there is a relationship between income, industrial sector and sex, with income dependent on both sex and industrial sector. There is also a relationship between industrial sector and sex, with most female employment tending to predominate in agriculture compared to the other sectors.

TABLE 10

CONDITIONAL DISTRIBUTIONS OF OCCUPATION GIVEN SE AND YEAR

2007

	Status in Employment			
Occupation	Employers	Employees	Others	Total
Leg.	22.6	13.8	4.4	14.4
Profs	15.6	20.6	3.0	20.2
Clerks	3.1	14.1	39.3	13.3
Service W.	33.6	27.9	28.9	28.2
Sk. Agric.	0.6	0.7	12.6	0.7
Others	24.5	22.9	11.1	22.9
Total	100.0	100.0	100.0	100.0

2006

	Status in Employment			
Occupation	Employers	Employees	Others	Total
Leg.	22.8	13.7	5.7	14.3
Profs	15.3	20.1	1.9	19.7
Clerks	3.3	14.3	35.8	13.4
Service W.	33.3	28.0	28.3	28.2
Sk. Agric.	0.6	0.7	15.1	0.7
Others	24.7	23.2	13.2	23.2
Total	100.0	100.0	100.0	100.0

Source: LABORSTA, Web database, ILO

TABLE 11

SUMMARY STATISTICS FOR CONDITIONAL DISTRIBUTIONS OF INCOME BY INDUSTRY AND SEX

		Industrial sector		
Sex	**Summary statistics**	**Agriculture**	**Industry**	**Services**
Male	Mean	1,510.0	2,543.0	7,182.3
	CV	1.32	1.41	1.92
	Skewness	5.6	1.8	2.3
	% distribution	43.4	32.1	24.5
Female	Mean	1,821.0	1,213.5	3,484.1
	CV	2.91	1.72	1.12
	Skewness	1.9	2.3	1.8
	% distribution	72.1	10.5	17.4

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.

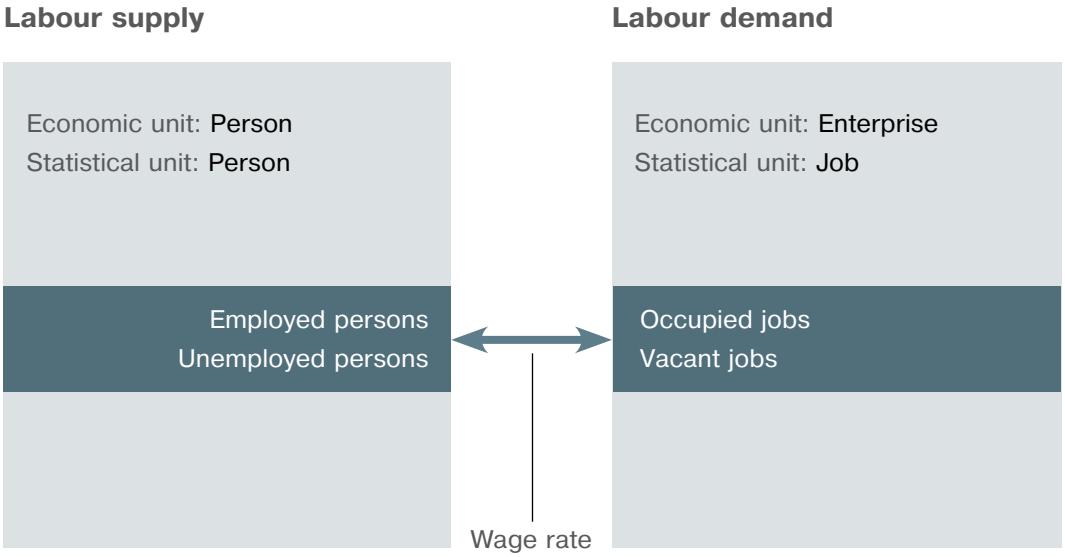
Labour demand, by contrast, is the sum of all occupied and vacant jobs that enterprises or employers require to conduct their economic activity. Occupied jobs are those currently filled, including those filled by employers or own-account workers themselves. Vacant jobs are unfilled jobs which the employers 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-

32 Hoffmann, Eivind, "Measuring the demand for labour," ILO Bulletin of Labour Statistics, Geneva 1992-1.

FIGURE 5
SUPPLY AND DEMAND OF LABOUR



Data source: Labour force survey

Data source: Establishment survey

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

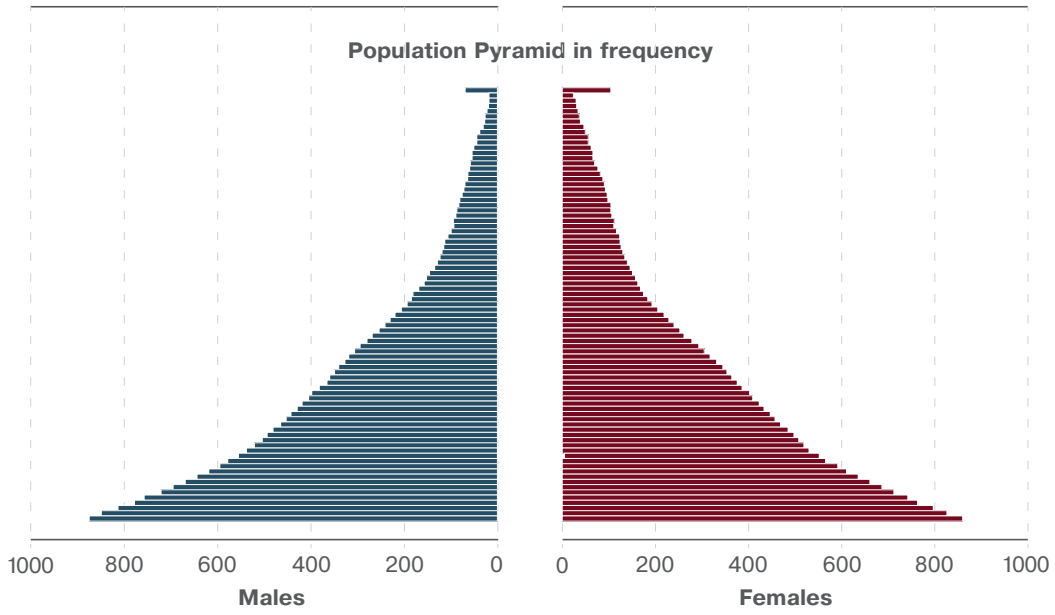
The size and composition of the population is the starting point of the analysis of labour supply. Population constitutes the human capital of the nation and defines its potential labour supply. From an economic point of view, the working population is a factor of production and its aptitude and skill 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-

FIGURE 6

TANZANIA AGE PYRAMID 2010



Source: United Nations, Department of Economic and Social Affairs, Population Division (2009). World Population Prospects: The 2008 Revisin, CD-ROM Edition.

Pyramid: Function "pyramid" of the contributed package epicalc in R (www.cran.r-project.org). For creating a population puramid online (www.projectos.org/statitisticum).

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.

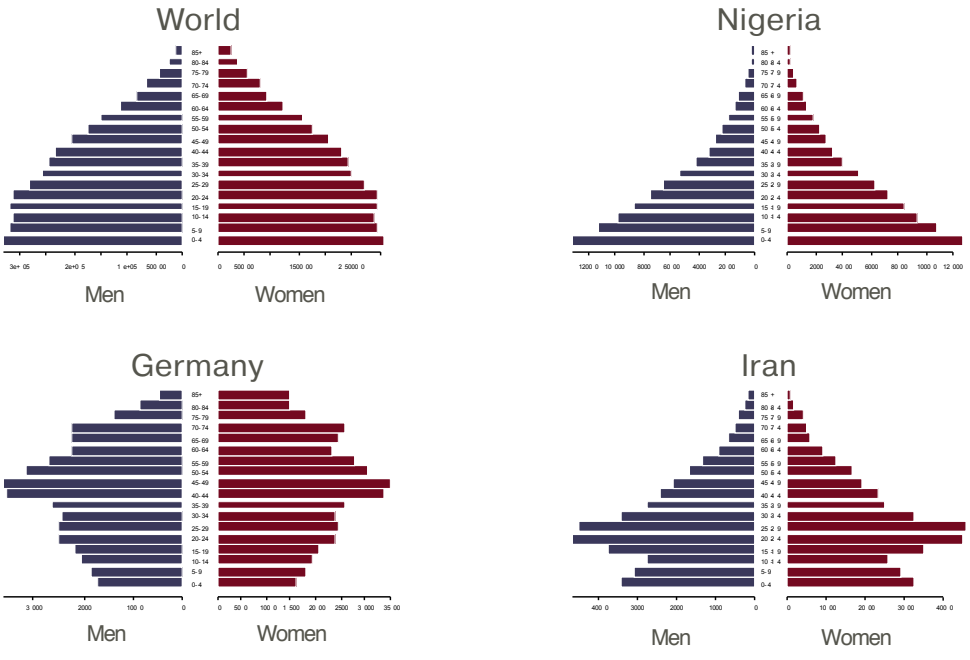
The age pyramid of Nigeria has a closely pyramidal shape, very similar to that of Tanzania. Virtually every lower age group has a larger population than the next higher age group. This type of age pyramid reflects a population with a high birth rate, a high death rate and a short life expectancy, a typical 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

FIGURE 7

POPULATION AGE PYRAMIDS, 2010



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):

Dependency ratio =

$$\frac{Children(0-14\,years)+Aged(65\,years+)}{WorkingAge(15-64\,years)}$$

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:

Child dependency ratio =

$$\frac{Children(0-14\,years)}{WorkingAge(15-64\,years)}$$

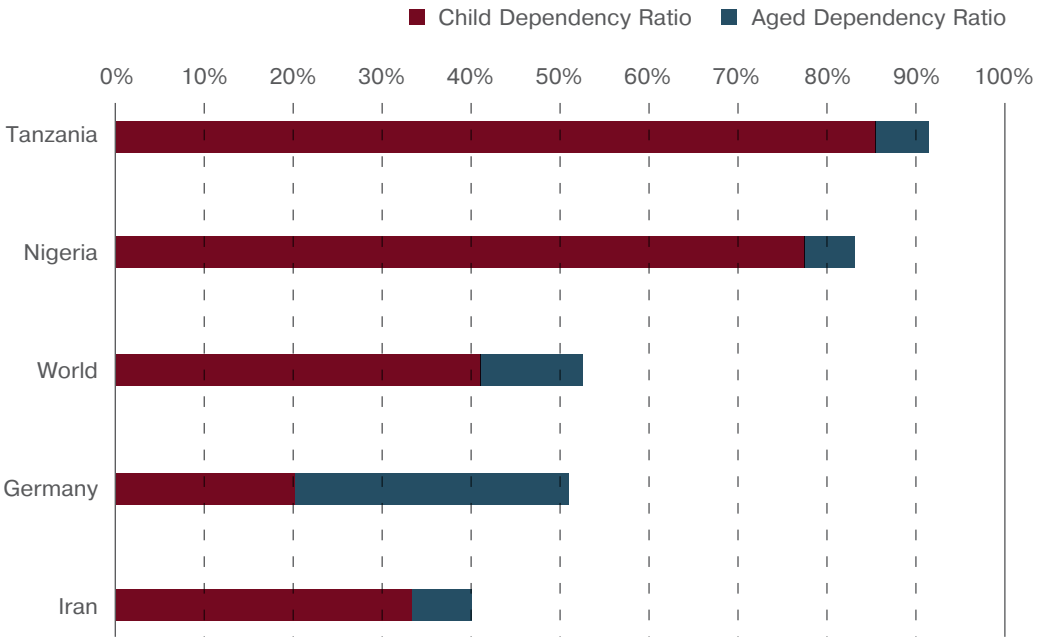
and the other:

Aged dependency ratio =

$$\frac{Aged(65\,years+)}{WorkingAge(15-64\,years)}$$

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

FIGURE 8
DEPENDENCY RATIOS (2010)



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

33 UN Population Division, World Population in 2300. ST/ESA/SER.A/236, New York, 2004.

34 ILO, Resolution concerning statistics of the economically active population, employment, unemployment and underemployment, 13th ICLS, Geneva, October 1982, http://www.ilo.org/. For an authoritative description and update see Ralf Hussmanns, "Measurement of employment, unemployment and underemployment – Current international standards and issues in their application" (with summaries in French and Spanish), ILO Bulletin of Labour Statistics, Geneva, 2007-1, pp. IX–XXX1.

at least distinguish between persons below 15 years of age and those 15 years and over. 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.

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:

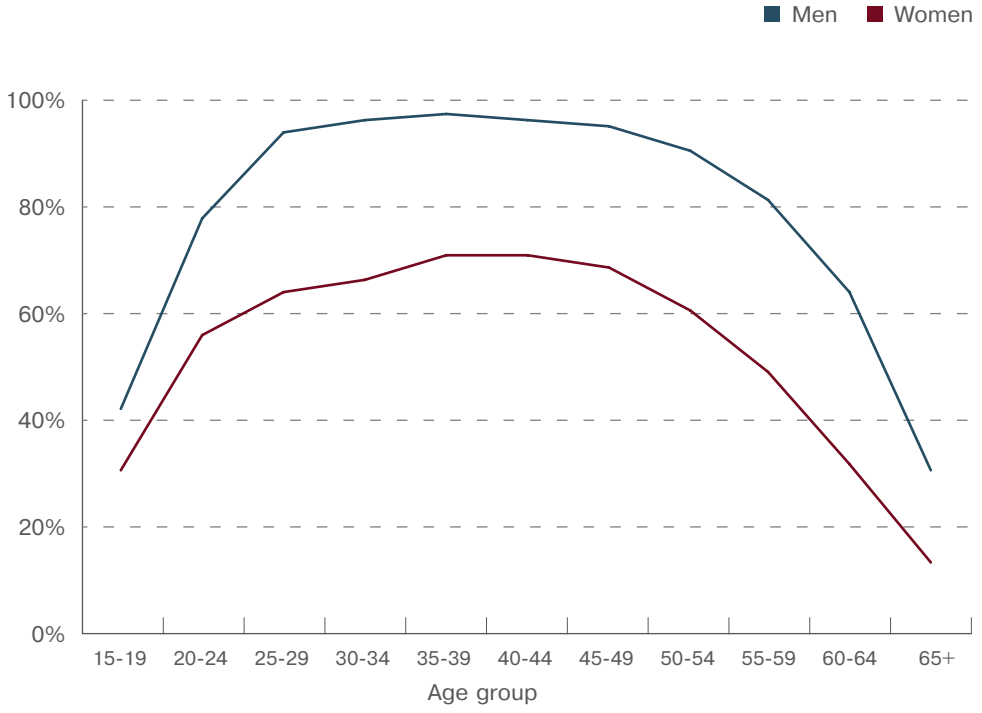
$$\text{LFPR} = 100 \times \frac{Labourforce}{WorkingAgePopulation}$$

The breakdown of the LFPR by sex and age group gives a profile of the distribution of the EAP within a country. National LFPRs and highlights of the data are published every two years as part of the ILO Key Indicators of the Labour Market (KILM 1).[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.[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.

FIGURE 9
LABOUR FORCE PARTICIPATION RATE (WORLD, 2010)



35 ILO, Key Indicators of the Labour Market 7th edition, Geneva, 2011, KILM 9.

36 ILO, Economically Active Population Estimates and Projections 1980–2020. Version 5, revision 2009. http://laboursta.ilo.org/.

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 reason 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.

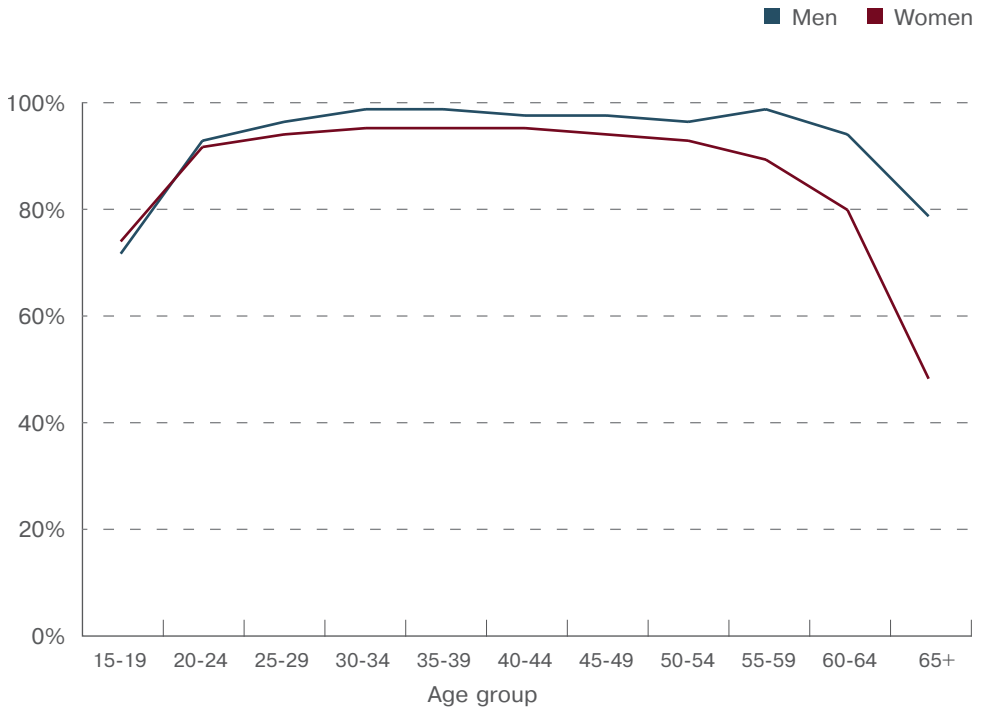
The US diagram also shows that the age at which most young people are in the labour force is between 20 and 24 years for both men and women. The age at which most older people are out of the labour force is between 60 and 64 years, also for both men and women.

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

FIGURE 10

LABOUR FORCE PARTICIPATION RATE, TANZANIA (ZANZIBAR), 2008



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

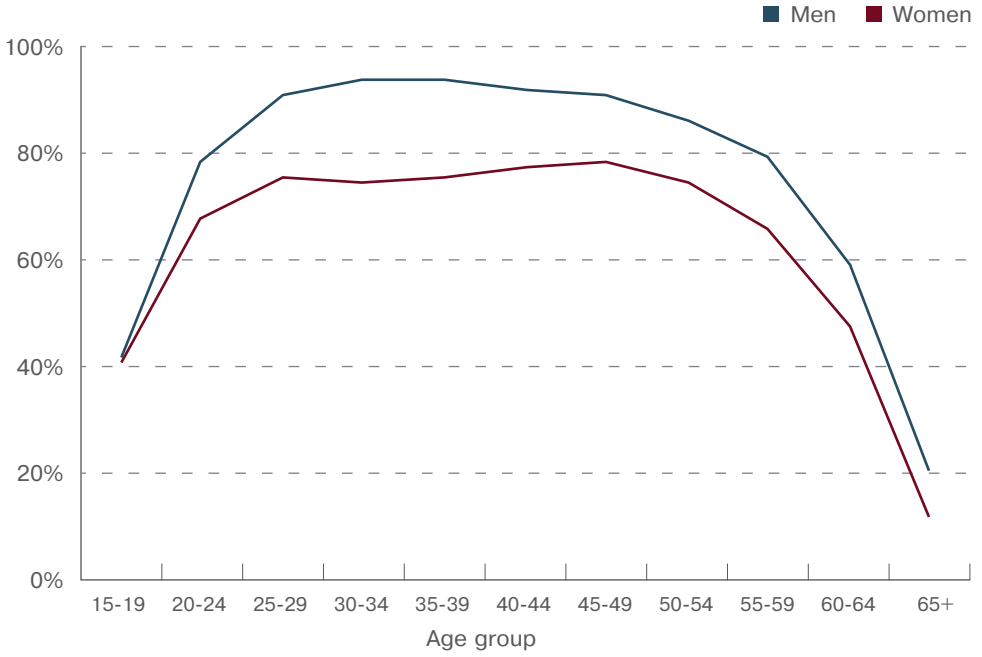
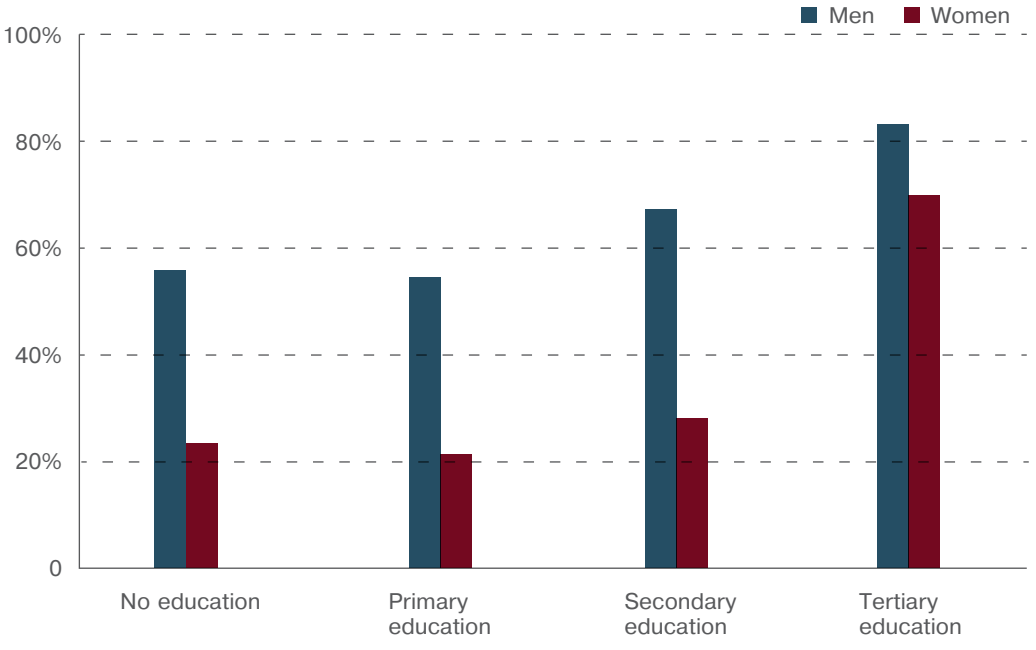


FIGURE 12

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.

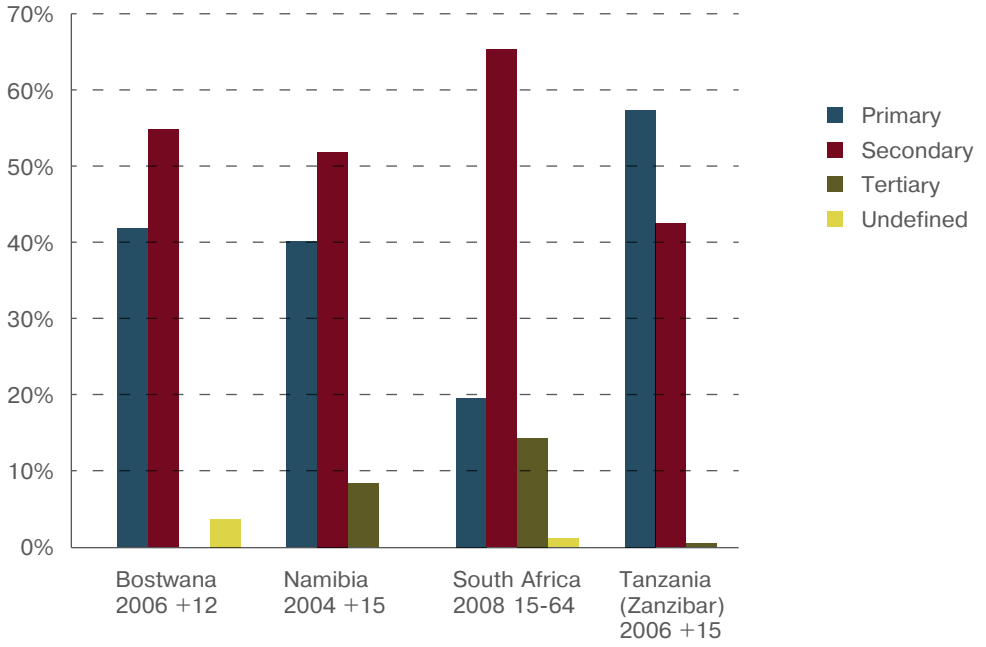
Another use of the data is to find out the skill levels of the labour force of the country. The larger the proportion of the labour force with secondary and tertiary (university) education, the higher the skill level of the labour force. Figure 13 presents the distribution of the labour force by educational 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.

Laborsta for Botswana, Namibia and South Africa. For Tanzania (Zanzibar): 2006 Integrated Labour Force Survey, General Report, Volume II, Labour Commission and Office of Chief Government Statistician, Zanzibar, June 2008, pp. 13–15.

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

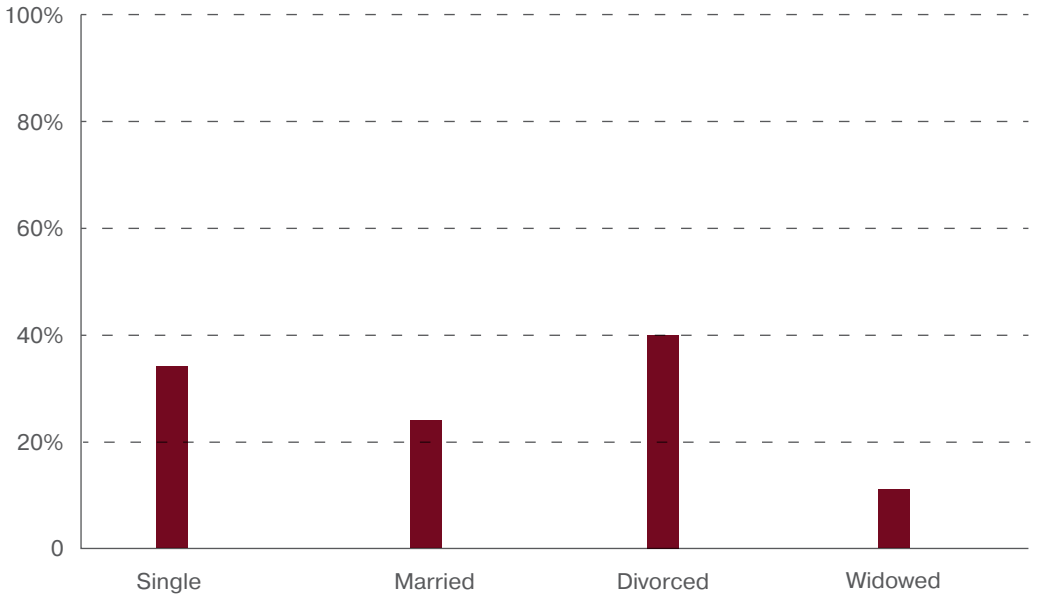
FIGURE 13
LABOUR FORCE BY EDUCATIONAL ATTAINMENT IN FOUR COUNTRIES



Source: ILO database Laborsta for Botswana, Namibia and South Africa. For Tanzania (Zanzibar): 2006 Integrated Labour Force Survey, General Report, Volume II, Labour Commission and Office of Chief Government Statistician, Zanzibar, June 2008, pp. 13–15.

FIGURE 14

FEMAILE LABOUR FORCE PARTICIPATION RATE MY MARITAL STATUS



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

37 Standing, G, Labour Force Participation and Development, International Labour Office, Geneva, 1978.

population. 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[38] is defined in percentage terms as:

$$\text{Employment-population ratio} = \frac{Number Employed}{Working Age Population}$$

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 12
CALCULATION OF THE CORE EMPLOYMENT–POPULATION RATIO

South Africa (Labour Force Survey)	2007*	2008	Change
Number of employed persons aged 25–54 ('000)	10,566	10,928	361
Population aged 25–54 ('000)	17,684	17,998	314
Core employment–population ratio	**59.8%**	**60.7%**	**1.0%**

* September
Source: LABORSTA, Web database, ILO

38 ILO, Key Indicators of the Labour Market 7th edition, Geneva, 2011, KILM 2.

TABLE 13
AGE AT EMPLOYMENT AND AT RETIREMENT

South Africa (Labour Force Survey – Sep 2007)	Male	Women
Age at which most youth of that age are employed	25	29
Age at which most older people of that age have left employment	61	54

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 =

$$17.5 + 5\frac{(50\% - r_{15-19})}{(r_{20-24} - r_{15-19})}$$

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

$$= 57.5 + 5\frac{(50\% - r_{55-59})}{(r_{60-64} - r_{55-59})} \; if\,(r_{60-64} \geq 50\%)$$

$$= 62.5 + 5\frac{(50\% - r_{60-64})}{(r_{65+} - r_{60-64})} \; if\,(r_{60-64} < 50\%)$$

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 =

$$100 \times \frac{Number\,Unemployed}{Labour\,Force}$$

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.[39] 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:[40]

TABLE 14
CALCULATIONS OF THE UNEMPLOYMENT RATE

South Africa (Labour Force Survey)[1]	2007[2]	2008	Change
Number of unemployed persons ('000)	3,945	4,075	130
Labour force ('000)	17,178	17,788	610
Unemployment rate (%)	**23.0%**	**22.9%**	**-0.1%**

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

39 *ILO, World Labour Report 1995. Chapter 1: "Controversies in labour Statistics," Geneva, 1995, pp. 11–30.*
40 *US Bureau of Labor Statistics, Economic News Release, Alternative measures of labor underutilization,* http://www.bls.gov/news.release/empsit.t15.htm.

U1. Persons unemployed 15 weeks or longer, as a percent of the civilian labour force

U2. Job losers and persons who completed temporary jobs, as a percent of the civilian labour force

U3. Total unemployed, as a percent of the civilian labour force (the official unemployment rate)

U4. Total unemployed plus discouraged workers, as a percent of the civilian labour force plus discouraged workers

U5. Total unemployed, plus discouraged workers, plus all other marginally attached workers as a percent of the civilian labour force plus all marginally attached workers

U6. Total unemployed, plus all marginally attached workers, plus total employed part time for economic reasons, as a percent of the civilian labour force plus all marginally attached workers

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.

The resulting values of these alternative rates may have wide variations. For example, in October 2011, it ranged from U1=5.0% to U6=15.3% with the official unemployment at U3=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.[41] 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 na-

41 ILO, Key Indicators of the Labour Market 1999. Table 8b. ILO-comparable unemployment rates, Geneva, 1999, pp. 220–230.

FIGURE 15

UNEMPLOYMENT RATE BY AGE GROUP (SOUTH AFRICA SEPTEMBER 2007)

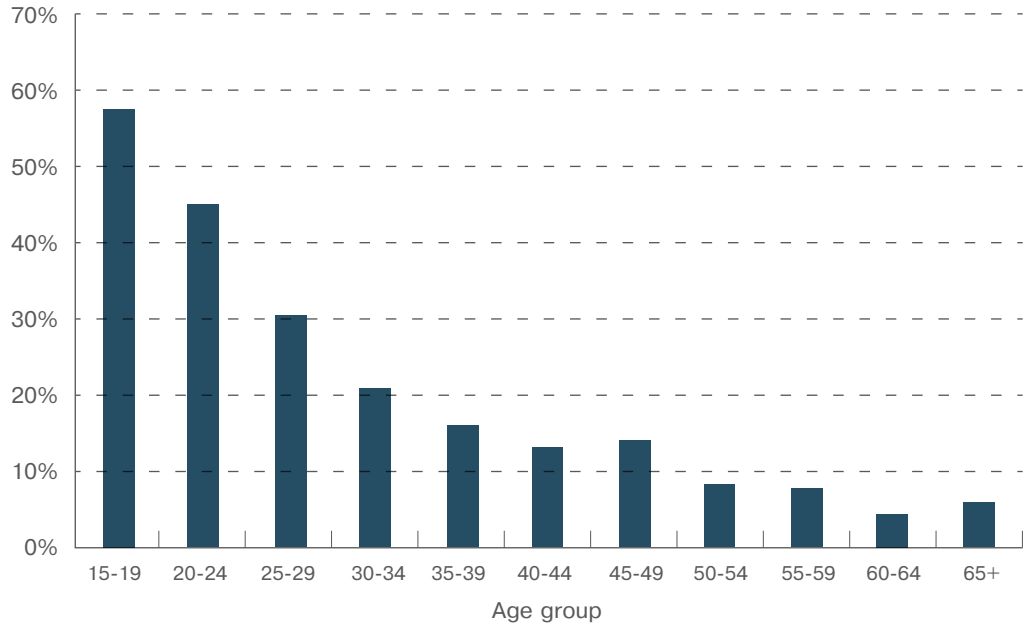


FIGURE 16

YOUTH VERSUS NATIONAL UNEMPLOYMENT RATE

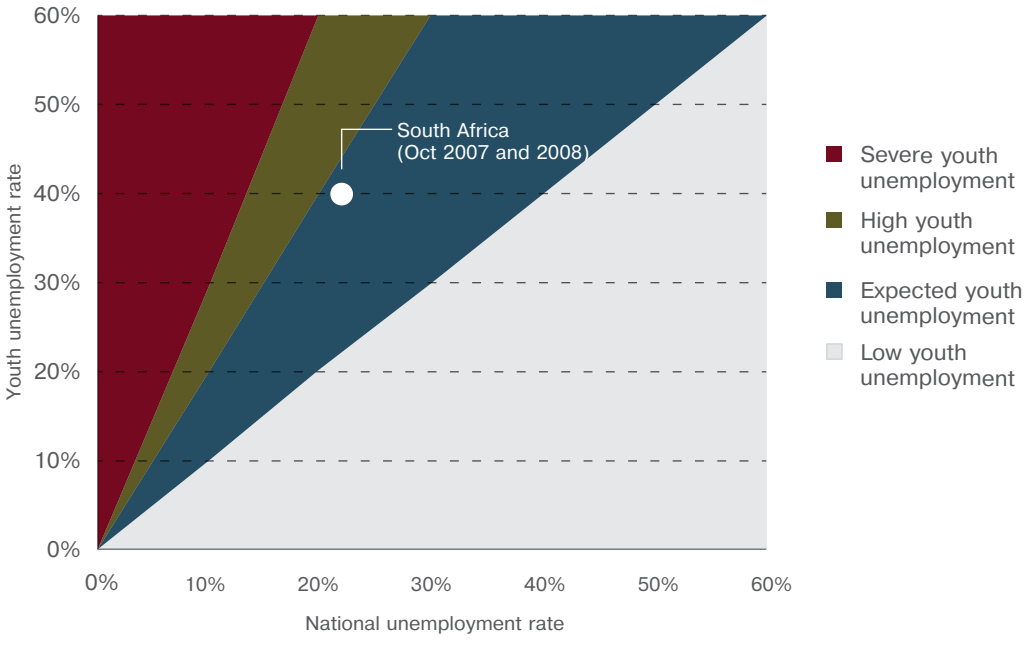


TABLE 15
RATIO OF YOUTH TO NATIONAL UNEMPLOYMENT RATE

South Africa (Labour force survey)	2007[1]	2008
Youth unemployment rate (15-24)	53.1%	54.4%
National unemployment rate (15+)	23.0%	22.9%[2]
Ratio of youth to national unemployment rate	2.3	2.4

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

tional rate, but less than twice the national rate; the plain blue area refers to the more severe situation where the youth rate is more than twice but less than three times the national rate; and finally the dark blue area refers to the extreme situation where the youth rate is more than three times the national rate.

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.

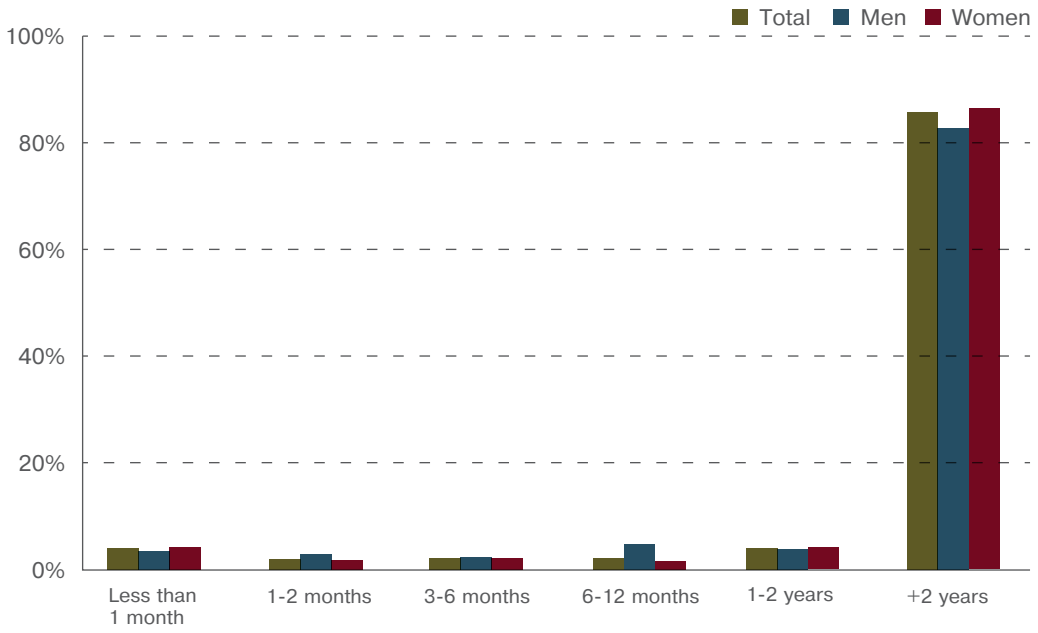
These results indicate not only the gravity of unemployment in Tanzania, Zanzibar, but also the difficulty of measuring and interpreting data on the duration of unemployment in some areas.

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).

42 Kiefer, Nicholas, M & Lundberg, Shelly J & Neumann, George R, "How Long Is a Spell of Unemployment? Illusions and Biases in the Use of CPS Data," Journal of Business & Economic Statistics, American Statistical Association, vol. 3(2), April 1985, pp. 118–128.

FIGURE 17

PERCENTAGE NUMBER OF UNEMPLOYED PERSONS BY DURATION OF UNEMPLOYMENT



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

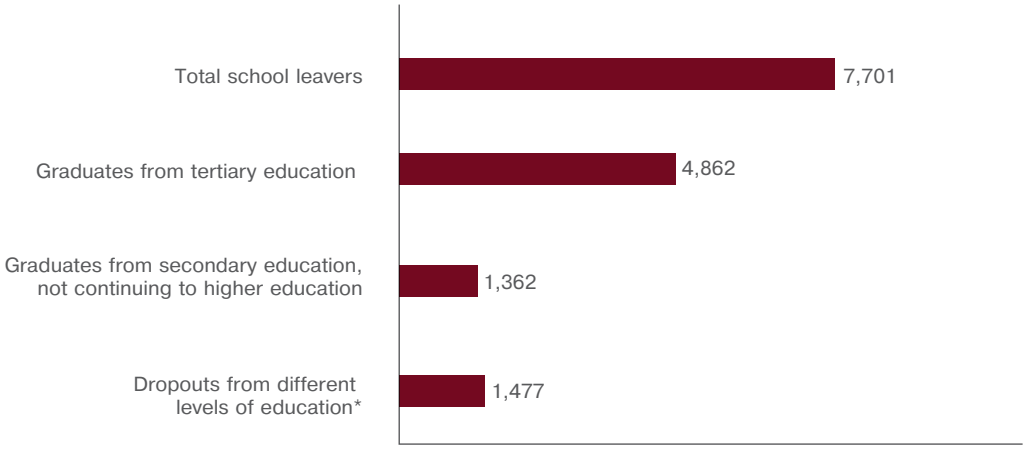
The ILO resolution concerning youth employment, adopted at the International Labour Conference (86th ILC 1998), defines in its preamble the youth population as persons 15 to 24 years old.[43] While the international standards should be regarded as the norm, it may be appropriate in countries where entry 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

FIGURE 18

SCHOOL LEAVERS: ACADEMIC YEAR 2004-2005 AN ILLUSTRATION



* Dropouts from tertiary education (506); from upper secondary education (596); from lower secondary education (253); and from primary education (122).

completing the level of education in which they were.[44] 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.[45] 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.

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

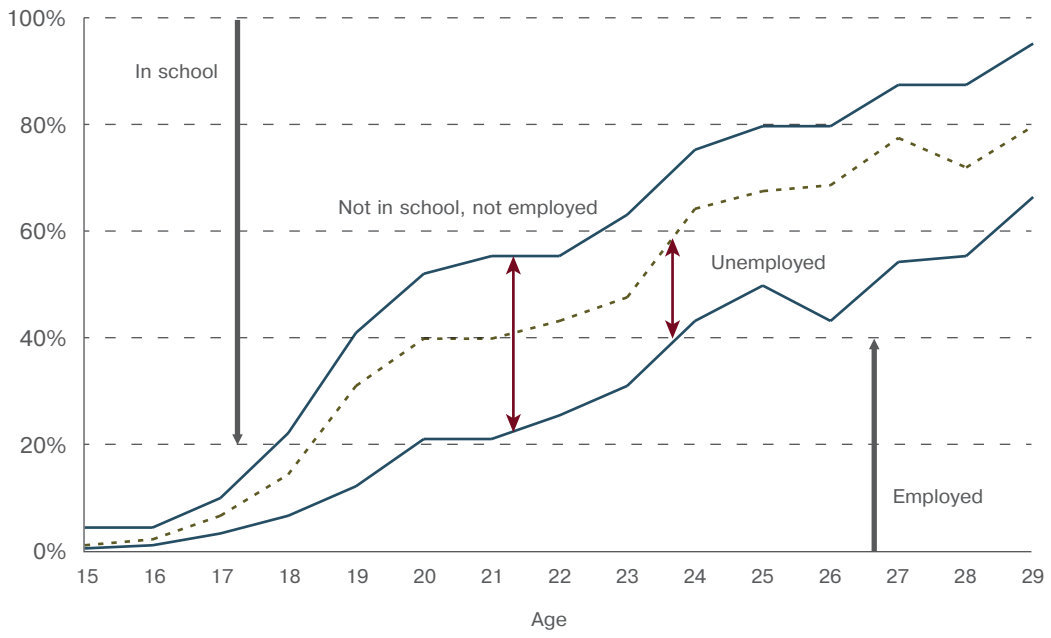
43 ILO Resolution concerning youth employment, 86th International Labour Conference, 1998

44 For indicators on graduates and dropouts see UNESCO, Education Indicators, Technical guidelines, The Institute for Statistics, November 2009.

45 UNESCO, International Standard Classification of Education 1997, Institute for Statistics, May 2006, Re-edition, www.uis.unesco.org.

FIGURE 19

AN ILLUSTRATION WITH LABOUR FORCE SURVEY DATA



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

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

where Ei and Si are the percentage of young people i years old who are "employed" and "at school, not employed" respectively, and Eo and So are the average of the corresponding end-point values, Eo = (E15+E29)/2 and So=(S15+S29)/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. It

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

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

and

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

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.[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[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

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.

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.

TABLE 16
SCHOOL-TO-WORK TRANSITION MATRIX

School-to-work transition matrix (15-29 years old)			Ocotber 2009				
			Et	Ut	St	Ot	Total
April 2009	Employed	E_{t-1}	38000	2200	800	1900	**42900**
	Unemployed	U_{t-1}	1900	20100	0	300	**22300**
	Inactive, at school	S_{t-1}	800	1900	61100	1700	**65500**
	Inactive, not at school	O_{t-1}	1200	500	5600	14100	**21400**
	Total		**41900**	**24700**	**67500**	**18000**	**152100**

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_t) 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},U_t) 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, UNDER-EMPLOYMENT 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.[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.

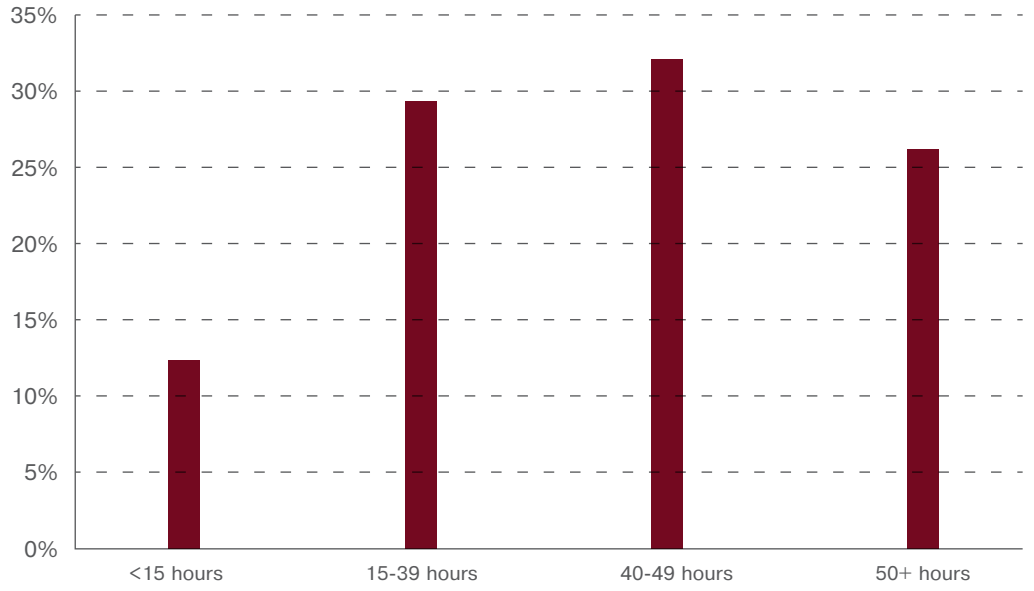
Hours usually worked are the hours actually worked in a job during a typical week (or in general any specific reference period). It can be calculated, for example, as the most frequent number of hours that a person actually worked per week during the past month.

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-

FIGURE 20

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



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

TABLE 17
SHORT AND LONG HOURS OF WORK

	Short hours of work (<15 hours per week)	Long hours of work (50+ hours per week)
Total	4.8%	42.6%
Men	3.5%	48.0%
Women	11.5%	15.6%
Urban	3.8%	45.7%
Rural	6.9%	36.4%
Public-sector employees	2.0%	48.5%
Private-sector employees	4.5%	26.9%
Self-employed	6.6%	49.4%
Contributing family workers	8.7%	23.7%

ers in agriculture during the off-season. To the extent that short hours of work are voluntary or for non-economic reasons, it should not be regarded as underemployment. At the other extreme, long hours of work or excessive hours of work as termed in the framework of decent work indicators are considered a threat to physical and mental health, interfering with the balance between work and family life, reducing productivity and often signalling an inadequate hourly pay.[49]

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).

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.[50]

Rates of time-related underemployment for seven countries are presented in Figure 21.[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.

19.3 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

49 ILO, Measuring Decent Work, Discussion Paper for Tripartite Meeting of Experts on the Measurement of Decent Work, Geneva, 8-10 September 2008.

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

51 ILO, Beyond Unemployment: Measurement of Other Forms of Labour Underutilization, Room Document 13 (Annex), 18th ICLS, Geneva, 24 November to 5 December 2008.

FIGURE 21

RATE OF TIME-RELATED UNDEREMPLOYMENT IN SEVEN COUNTRIES

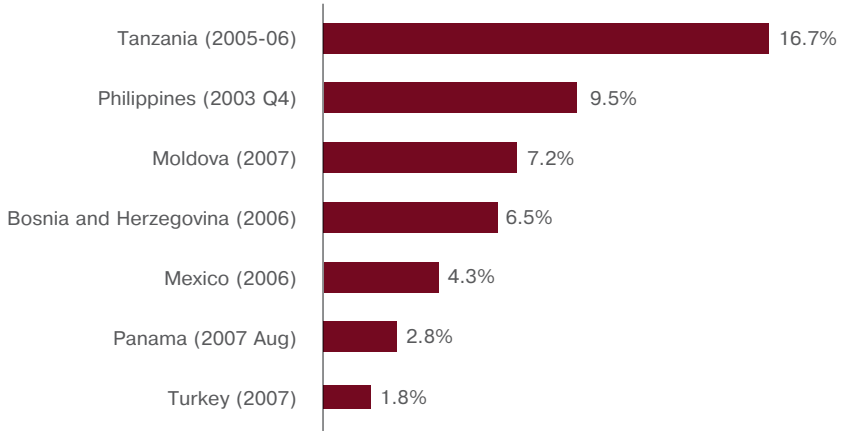


TABLE 18

ALTERNATIVE RATES OF LABOUR SLACK IN SEVEN COUNTRIES

Alternative rates of labour slack	U3 Unemployment rate	U4 Unemployed plus discouraged workers	U5 Unemployed, plus discouraged workers, plus other marginally attached workers	U6 Unemployed, plus all marginally attached workers, plus time-related underemployed
Bosnia Herzegovina (2006)	32.0%	37.4%	42.8%	44.8%
Philippines (2003 Q4)	5.8%	7.9%	12.7%	19.3%
Tanzania (2005–2006)	3.4%	3.9%	4.7%	17.8%
Turkey (2007)	9.9%	12.3%	15.3%	16.5%
Moldova (2007)	5.1%	6.9%	7.3%	13.2%
Panama (2007 Oct)	5.7%	5.8%	5.8%	8.2%
Mexico (2007 Q2)	3.4%	3.5%	3.7%	7.5%

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.[52] 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.[53] The cor-

52 ILO, Key Indicators of the Labour Market. 7th edition, KILM 4, Geneva, 2011.

53 United Nations, International Standard Industrial Classification of All Economic Activities, ISIC-88, Rev. 3,1 and Rev. 4 and correspondence tables http://unstats.un.org/unsd (click Methods & Classifications).

respondence between ISIC Rev 3.1 and ISIC Rev 4 at the aggregate level is given in Figure 22 below.

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%).

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:

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

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 indictor may expressed by

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

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

FIGURE 22
CORRESPONDENCE OF INTERNATIONAL STANDARD CLASSIFICATIONS OF ALL ECONOMIC ACTIVITIES (ISIC REV 3.1 AND ISIC REV 4)

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

FIGURE 23

EMPLOYMENT BY MAJOR BRANCH OF ECONOMIC ACTIVITY, SOUTH AFRICA (2008)

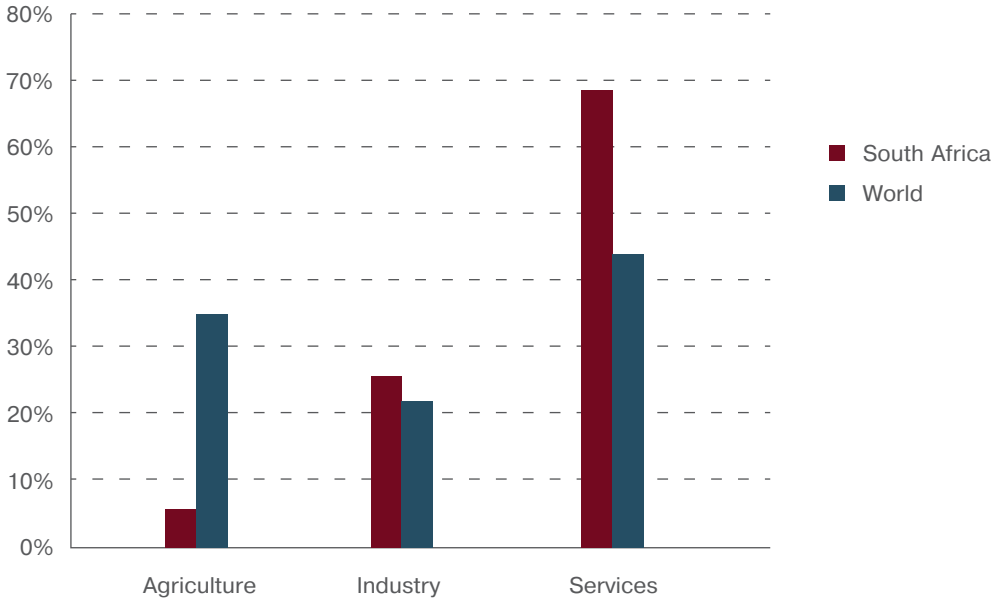
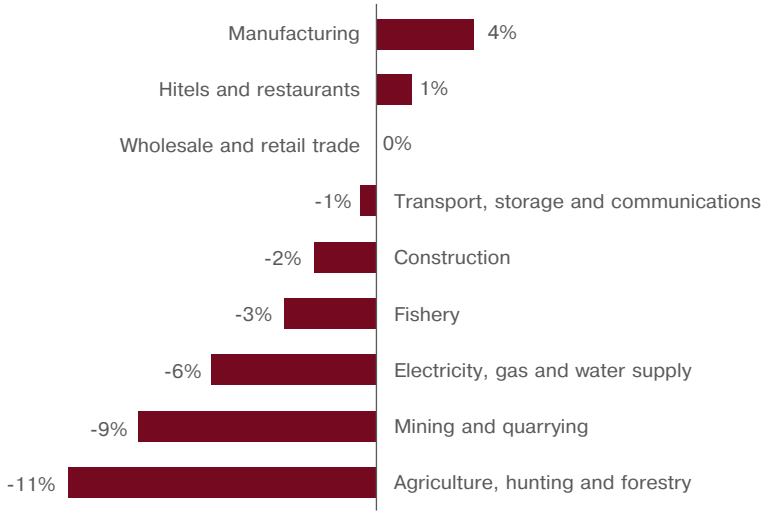


FIGURE 24

LABOUR PRODUCTIVITY IN SELECTED BRANCHES OF ECONOMIC ACTIVITY



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.[54] ISCO-08 uses the International Standard Classification of Education[55] to define four broad categories of skill levels: 1st level – Primary education or first stage of basic education; 2nd level – Lower secondary or second stage of basic education; 3rd level – Upper secondary education; and 4th level – Higher education.

21.3 OCCUPATIONAL SEGREGATION

A commonly used index of occupational segregation is given by[56]

$$D = \frac{1}{2}\sum_i \left| \frac{n_{Ai}}{n_A} - \frac{n_{Bi}}{n_B} \right|$$

where n_{Ai} and n_{Bi} are, respectively, the number of men and women in occupational category I and nA and nB 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.[57] 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

54 ILO, ISCO-08, http://www.ilo.org/public/english/bureau/stat/download/res/futisco.pdf.

55 UNESCO, International Standard Classification of Education ISCED 1976, Paris. http://www.unesco.org/education/information/nfsunesco/doc/isced_1997.htm

56 James, D. R. and Taeuber, K. E. (1985) "Measures of segregation". In Tuma, N. (ed.) Sociological Methodology, Jossey-Bass, San Francisco, pp.1–32. See also Watts, M. (1998) "Occupational gender segregation: Index measurement and econometric modeling", Demography 35(4), pp. 489-496.

57 ILO (2008), Measuring Decent Work, Discussion Paper for Tripartite Meeting of Experts on the Measurement of Decent Work, International Labour Office, Geneva, 8-10 September 2008.

TABLE 19

OCCUPATIONAL SEGREGATION AMONG YOUNG MEN AND WOMEN IN THREE COUNTRIES

	Year	Occupation classification	Occupational categories	Segregation index
Iran	2006	4-digit	325	0.6351
	-	3-digit	111	0.6138
Sweden	2000	3-digit	125	0.5297
United States	2000	3-digit	491	0.5078

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.

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.

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]

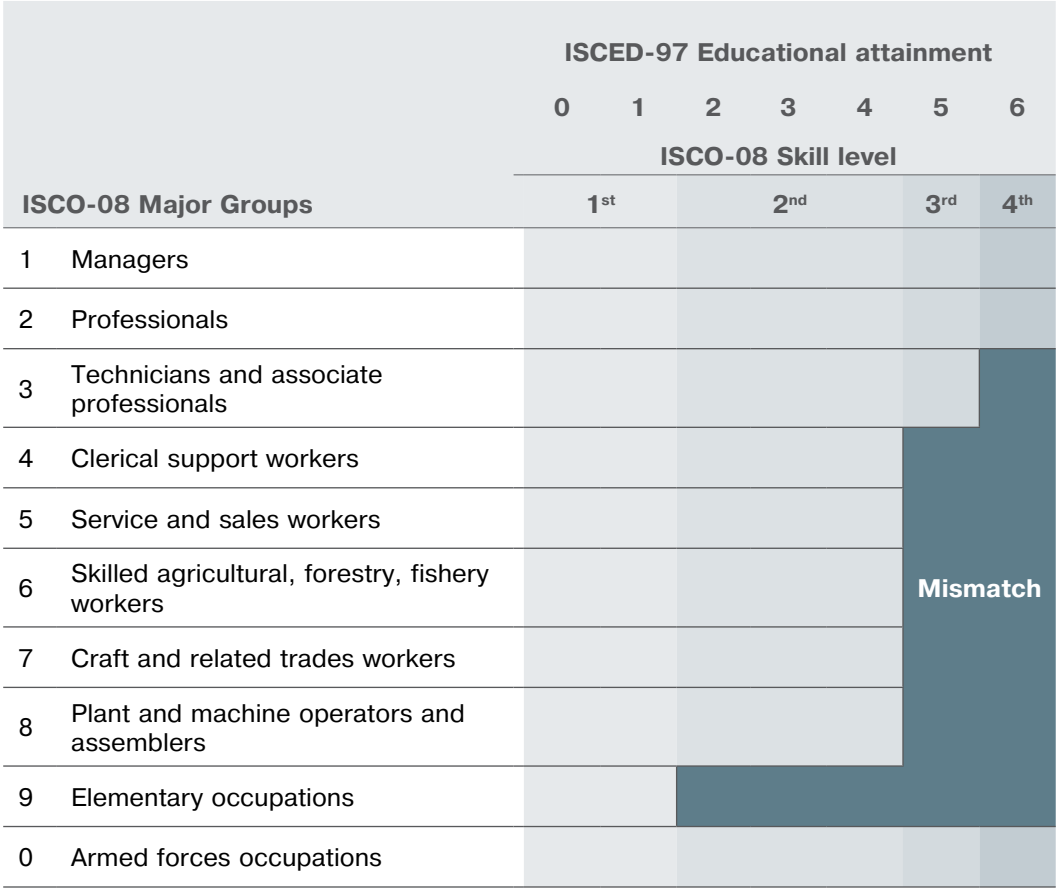
A simple method that uses level of educational attainment and one-digit occupation data generally available from labour force surveys defines a mismatch when the educational attainment of the worker is higher than the educational level required by the main current job. The following matrix shows the 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%).

58 Sullivan, Teresa, A., and Hauser, Philip, M., "The Labor Utilization Framework: Assumptions, Data, and Policy Implications," in Counting the Labor Force, National Commission on Employment and Unemployment Statistics, Appendix Vol. I, Concepts and Data Needs, US Government Printing Office, Washington DC, 1977.

TABLE 20

OCCUPATIONAL SEGREGATION AMONG YOUNG MEN AND WOMEN IN THREE COUNTRIES



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. 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%).

Informal employment	South Africa (2010)	Uganda (2010)	Mali (2004)
Informal employment (share in non-agricultural employment)	33%	67%	82%
	100%	100%	100%
Informal sector employment	54%	81%	86%
Informal employment outside the informal sector	46%	19%	14%

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:[59]

- 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,[60] informal employment is defined to include:

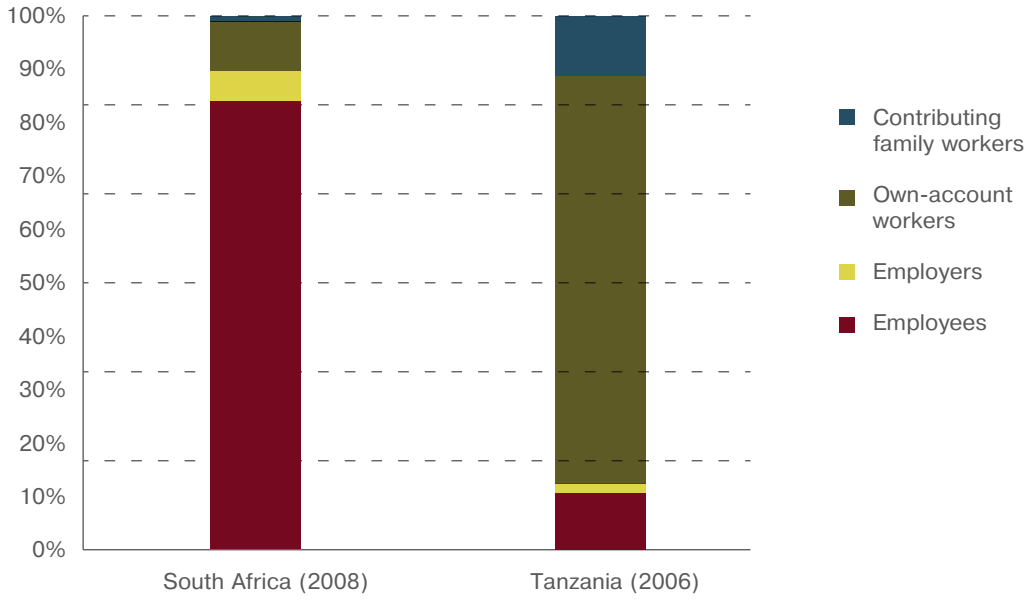
- Employment in the informal sector (except rare employees in that sector who may have

59 ILO, International Classification of Status in Employment, ICSE-93, 15th ICLS, Geneva, http://laborsta.ilo.org.

60 ILO, Resolution on the measurement of employment in the informal sector, adopted by the 15th ICLS, Geneva, 1993, and Guidelines concerning a statistical definition of informal employment, adopted by the 17th ICLS, Geneva, 2003.

FIGURE 25

EMPLOYMENT BY MAJOR BRANCH OF ECONOMIC ACTIVITY, SOUTH AFRICA (2008)



formal employment). The informal sector consists of informal own-account enterprises and enterprises of informal employers. For operational purposes, informal own-account enterprises are defined as either all own-account enterprises or only those that are not registered under specific forms of national legislation. Enterprises of informal employers are defined in terms of one or more of the following criteria: (i) size of the unit below a specified level of employment; and (ii) non-registration of the enterprise or its employees.

› Informal employment outside the informal sector. Informal employment outside 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.).

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).[61]

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%).

61 ILO Key Indicators of the Labour Market Seventh edition, Geneva 2011, Table 8.

TABLE 21
INFORMAL EMPLOYMENT IN THREE COUNTRIES

Informal employment	South Africa (2010)	Uganda (2010)	Mali (2004)
Informal employment (share in non-agricultural employment)	33%	67%	82%
	100%	100%	100%
Informal sector employment	54%	81%	86%
Informal employment outside the informal sector	46%	19%	14%

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.

23.2 INCOME FROM EMPLOYMENT[62]

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.

- 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.

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.

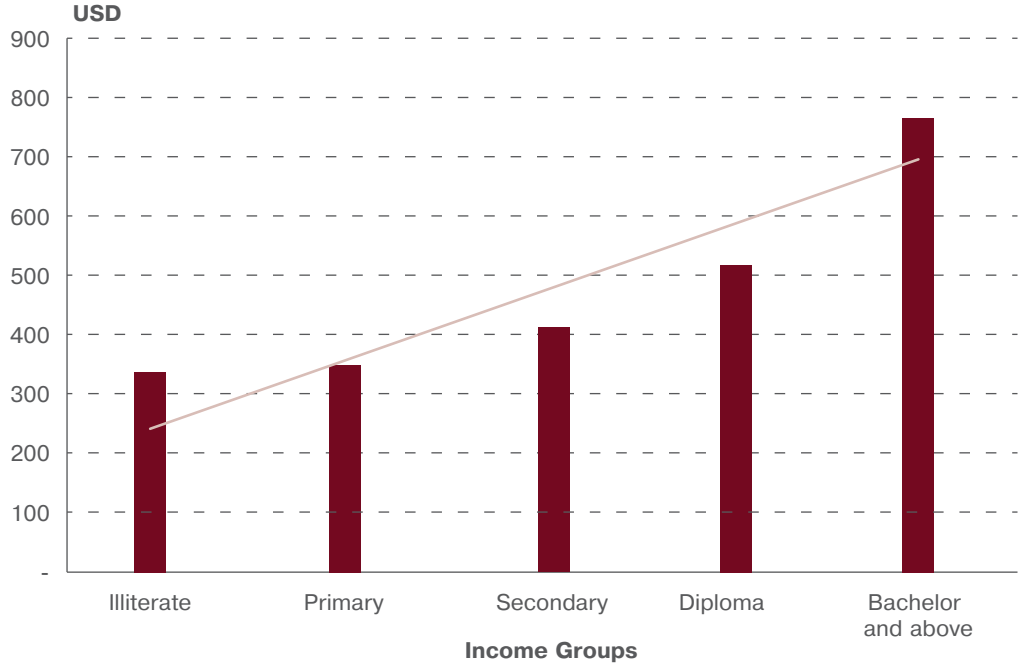
23.4 EARNINGS DIFFERENTIALS

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.[63] 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



63 Chen, Martha, Joann Vanek, Francie Lund, James Heintz with Renana Jhabvala and Christine Bonner, Progress of the World's Women 2005, Women, Work & Poverty, UNIFEM, New York, 2005, pp. 46–50.

FIGURE 27

AVERAGE MONTHLY EARNINGS BY AGE GROUP

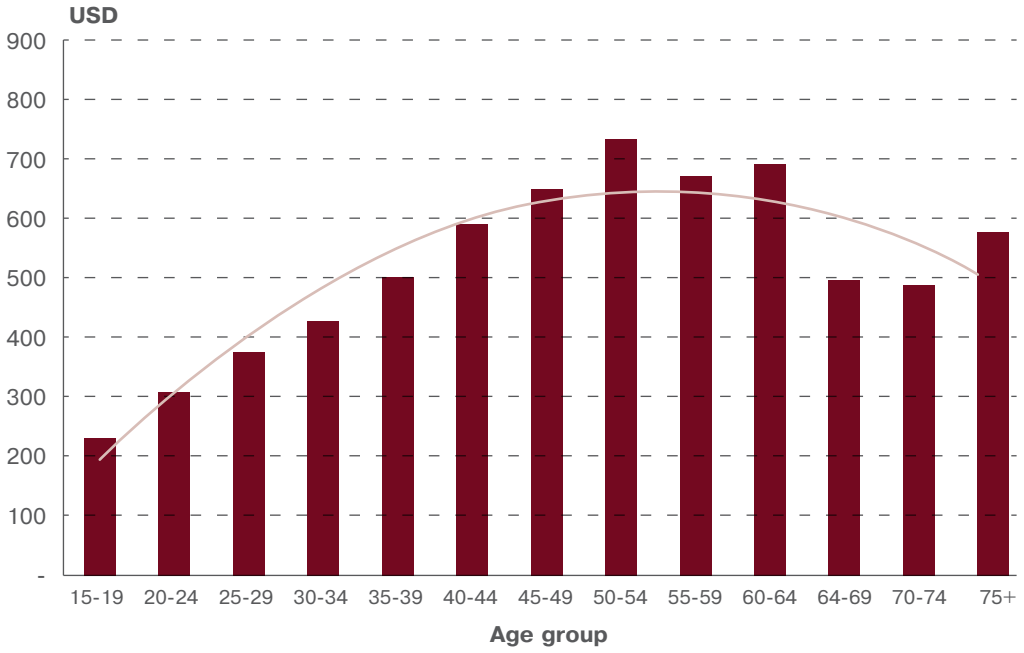


TABLE 22

AVERAGE REPORATED MONTHLY EARNINGS IN MAIN JOB (USD 2007) AMONG MEN AND WOMEN

Branch of Economic activity	Men	Women	Difference
Total	492	455	37
Managers	906	840	67
Professionals	760	563	196
Technicians and Associate professionals	535	406	129
Clerical support workers	463	394	69
Service and sales workers	361	256	105
Skilled Agricultural, Forestry and Fishery workers	397	273	124
Craft and related trades workers	330	326	4
Plant and machine operators and assemblers	299	193	106
Elementary occupations	349	268	80
unspecified	474	867	(383)

TABLE 23
ESTIMATED PARAMETERS OF THE EARNINGS FUNCTION

Variable	Parameter	Estimate	Standard error	t-value
Intercept	β_o	3.742	6.483 e-02	41.014 ***
Sex	β_1	0.125	8.060 e-03	11.841 ***
Age	β_2	0.050	3.397 e-03	31.872 ***
Age^2	β_3	-0.0003	4.238 e-05	-9.362 ***
Educ	β_4	0.287	6.057 e-03	33.032 ***

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

$$w = \beta_o + \beta_1 \, sex + \beta_2 \, age + \beta_3 \, age_2 + \beta_4 \, educ + \varepsilon$$

where w is the logarithm of earnings, sex is a variable with value 1 for men and -1 for women, age represents the age variable, educ the educational attainment (1 for illiterate to 5 for bachelor's and above) and ε 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:

The highly significant coefficient of the sex variable (β_1=0.125) indicates that, even after adjusting for differences in length of work experience (age and age_2) and level of educational attainment (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.´

18th ICLS,[65] 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 who 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

65 See, for example, Tripartite Meeting of Experts on the Measurement of Decent Work, ILO, Geneva, 8–10 September 2008, Chairperson's report, paragraph 51.

66 Bescond, D., A. Châtaignier and F. Mehran, "Seven indicators to measure decent work: An international comparison," International Labour Review, ILO, Geneva, 2003/2, pp. 179-211.

67 ILO, Beyond Unemployment: Measurement of Other Forms of Labour Underutilization, Room Document 13, 18th ICLS, Geneva, 24 November – 5 December 2008.

TABLE 24
LOW EARNINGS IN SEVEN COUNTRIES

Low Earnings	Total	Full time employed (40-48 Hours per week)	Less than full time employed (<40 Hours per week)	Overly employed (>48 hours per week)		
				Total	Low Monthly Earnings	Low Hourly Earnings
Tanzania (2005-2006)	30.4%	6.1%	2.9%	21.4%	14.8%	6.6%
Panama (2007 Aug)	20.5%	7.1%	7.9%	5.4%	3.3%	2.1%
Phillipines (2003 Q4)	20.5%	8.1%	1.0%	11.4%	8.8%	2.6%
Mexico (2007 Q2)	11.3%	2.6%	3.2%	5.4%	2.3%	3.1%
Moldova (2007)	7.2%	5.0%	0.8%	1.3%	0.6%	0.7%
Bosnia & Herzegovina (2006)	6.8%	4.2%	0.0%	2.5%	0.9%	1.7%
Turkey (2007)	6.8%	1.8%	0.8%	4.2%	4.2%	-

FIGURE 28
DEFINITIONS OF WORKING POOR

BROAD DEFINITION

› Working persons who are unable to earn enough to maintain the welfare of themselves and their family

SPECIFIC DEFINITION

› Working persons in households with income below the poverty line

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

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

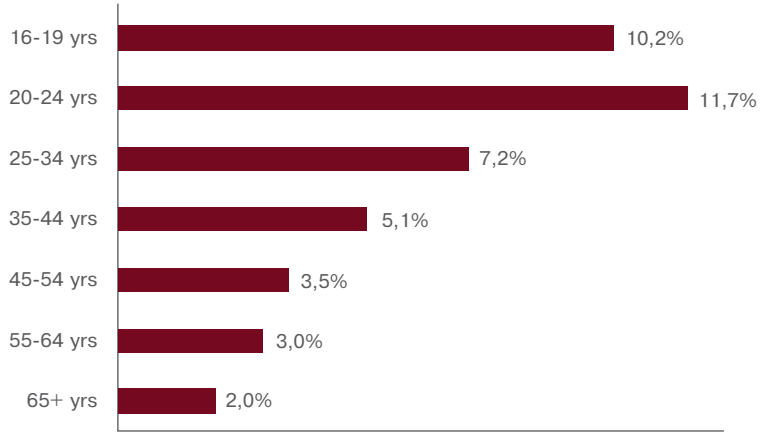
The US collects data on the working poor based on a survey conducted as part of an annual survey, called the Annual Social and Economic Supplement to the Current Population Survey (i.e. the US monthly labour force survey).[68] The definition of the working poor distinguishes between two categories of 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

68 U.S. Department of Labor, U.S. Bureau of Labor Statistics, A Profile of the Working Poor, 2004, Report 994, May 2006.

FIGURE 29

PROFILE OF WORKING POOR BY AGE GROUP (USA 2004)



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.

PART III:

Quality assessment of labour force survey data

25. Coherence and comparability of survey estimates
26. Sampling errors and precision of estimates
27. Non-sampling errors and weights adjustment

25. COHERENCE AND COMPARABILITY OF SURVEY ESTIMATES

25.1 INTRODUCTION

The quality of the results of the labour force survey may be assessed by comparing the survey data with data from other sources deemed more accurate. Lack of coherence of data from different sources is a problem because users of data are puzzled when faced with different figures referring to identical or similar concepts. From the point of view of data producers, reconciliation of the survey data with other sources can help to identify measurement errors and to take actions to reduce them.

Here, three examples are given for reconciling labour force survey data with national accounts, industrial surveys of large establishments, and administrative data on registered jobseekers.

25.2 COHERENCE WITH NATIONAL ACCOUNTS

The labour force framework is designed to maintain consistency between the employment statistics obtained from labour force surveys and the production statistics compiled in national accounts. As mentioned in Part I of this Guide Book, the 13th ICLS Resolution (1982), and its amendment in the 18th ICLS Resolution (2008), stipulates that "The economically active population comprises all persons of either sex who furnish the supply of labour for the production of goods or services within the production boundary, as defined by the latest version of the SNA, during a specified time-reference period."[69]

One therefore expects that the trends of employment and production follow the same pattern. If the labour force survey shows an increase in employment, one would expect to see also an increase in production in the national accounts. Similarly, a decline in the level of employment according to the labour force survey should be reflected in a decline in the level of production in the national accounts.

The following numerical example illustrates the coherence between employment and output trends for a particular country over the period 2006–2009. The diagram plots the annual growth rate of employment derived from labour force surveys over time and the corresponding annual growth rate of gross output measured in constant prices from national accounts.

FIGURE 1

PROFILE OF WORKING POOR BY AGE GROUP (USA 2004)



69 ILO, Report of the Conference, 18th ICLS, Geneva, 2008, pp.66.

It can be observed that the two curves closely follow each other, indicating the coherence between the survey employment data and the national accounts output data. The difference between the heights of the curves measures the change in labour productivity over the period.

Where estimates of national employment are also produced as part of national accounts, they can be compared with the employment data obtained from labour force surveys. In such cases, attempts may be made to reconcile the two sets of data to establish their degree of coherence or incoherence. Labour force survey estimates are generally lower than the national accounts estimate due to the more limited coverage of the survey.

The elements for reconciling the employment data from labour force surveys and the corresponding data from national accounts are presented in the following reconciliation table.[70]

25.2 ESTABLISHMENT SURVEYS AS BENCHMARK

Where national surveys of establishments collecting employment data with relative precision exist, the results may be used as benchmark for evaluating the quality of the corresponding employment data obtained from labour force surveys. For such comparisons to be valid, the timing of the two types of surveys should coincide and there should be sufficient information for reconciling the differences in the measurement units of employment and the scope of the statistics in terms of economic activity and size of establishment.

There are a large variety of establishment surveys, each designed to obtain specific information such as production, export, employment and average earnings, occupational employment and wage structure, labour cost, job vacancies, hiring and firing practices, skill level, business sentiment, future employment prospects, informal sector, child labour, etc. Of particular interest in the present context are national establishment surveys on employment and earnings in which data are collected, among other things, on number of persons engaged, number of paid employees or other similar units.

In many countries, the relevant establishment surveys focus on specific branches of economic activity, such as manufacturing industries, and cover establishments of certain sizes, for exam-

TABLE 1
RECONCILIATION OF EMPLOYMENT DATA: LABOUR FORCE SURVEY AND NATIONAL ACCOUNTS

Labour force survey	National accounts
LFS employed persons	Domestic employment
+ Conscripts	+ Residents working outside
+ Employed living in collective households	- Non-residents working inside
+ Unpaid trainees and apprentices	
+ Farmers producing wholly for own consumption	
Total employed persons	**National employment**
+/- Unaccounted difference	+ Hidden/unrecorded employment
Reconciled estimate	Reconciled estimate

70 European Commission, Task Force on the Quality of Labour Force Survey, Final report, Eurostat Methodologies and Working Papers, 2009 edition.

ple, large establishments with 50 or more workers. For comparison purposes, it is therefore necessary to use the corresponding employment data from the labour force survey. Such correspondence is generally possible on the basis of questions on branch of economic activity and size of establishment in terms of employment available in many labour force surveys. Regarding industry coverage, care should be taken to ensure proper correspondence especially in relation to the inclusion or exclusion of public-sector establishments. Regarding size of establishments, if the answer category in the labour force survey does not exactly correspond to the size coverage of the establishment survey, the employment data of the labour force survey should be interpolated to obtain the corresponding estimate.

Another issue concerns the unit of measurement. In establishment surveys, the unit of measurement is generally a job while in labour force surveys the unit of measurement is a person. A person may have more than one job and therefore the correspondence between the employment data from establishment surveys and labour force surveys should take multiple jobholding into account. This can generally be satisfactorily accomplished with labour force surveys covering data on secondary jobs. If such data are not available, everything else being equal, one should expect that the employment data from the establishment survey be higher than the corresponding data from the labour force survey.

A numerical example is presented below in Table 2, illustrating a quality assessment of employment data of a labour force survey using a national establishment survey as benchmark. The establishment survey refers to an annual survey designed to measure the value of output and valued added of large manufacturing establishments with 50 workers or more on the basis of complete enumeration. The survey covers all establishments with private, public and cooperative types of ownership engaged in the production of manufacturing goods (ISIC Rev 4 – Code D). The survey collects data on the average number of persons engaged during the past calendar year, including full-time and part-time paid employees and apprentices, employers and active partners, and unpaid family workers. According to the survey results, this figure for 2006 was 804,805.

The total employment for the same year according to the labour force survey was 21,093,000. The number of persons employed in manufacturing in their main job was 3,834,000, of whom the number employed in large establishments with 50 workers and more was 775,500.

As expected, the establishment survey estimate of employment (804,805) is higher than the cor-

TABLE 2
RECONCILIATION TABLE: NUMBER OF EMPLOYED PERSONS IN MANUFACTURING ESTABLISHMENTS WITH 50 OR MORE WORKERS

Labour force survey	Establishment survey
Main job: **775,500** Secondary job: **17,300**	804,805
Total: **792,800** Unaccounted difference: **12,000**	804,805
Reconciled estimate: **804,800**	804,805

responding estimate from the labour force survey (775,500). The difference may partly be due to persons with secondary jobs in large manufacturing establishments, estimated from the labour force survey to be 17,300. Taking this factor into account, the results can be presented in the following reconciliation table, which shows that the unaccounted difference between the estimate of employment in large manufacturing establishments and the corresponding estimate from the labour force survey is 12,000, representing an error rate of about 1.5% for the labour force estimate (if the establishment survey estimate is considered as more accurate).

Similar values of the percentage unaccounted difference were calculated for the period 2007–2009. The results ranged from 1.2% to 1.6%, providing some confidence that the labour force survey measures employment in large manufacturing establishments with a relatively low margin of error. In some cases, the comparison and reconciliation process may also be carried out in more detail, for example by sex, SE and subdivisions of manufacturing activities. Depending on the availability of data, it may also be carried out for other variables common in the two surveys, such as average hours of work and average earnings.

In some cases, the comparison and reconciliation process may also be carried out in more detail, for example by sex, SE and subdivisions of manufacturing activities. Depending on the availability of data, it may also be carried out for other variables common in the two surveys, such as average hours of work and average earnings.

25.3 ADMINISTRATIVE RECORDS OF REGISTERED JOBSEEKERS AS BENCHMARK

Many countries compile labour-related administrative data from a variety of sources, including social security organisations, public and private employment offices, unemployment insurance schemes, civil service administrations, etc.[71] Because they are by-products of administrative procedures, the statistics generally involve little extra cost and can be made available at frequent intervals and fine population categories and geographical areas. Data derived from administrative records, however, often suffer from a number of shortcomings, with respect to coverage, concepts and availability of auxiliary variables in particular.

As shown here, under certain conditions labour-related administrative data and in particular data on registered jobseekers can nevertheless be useful among other things for comparison with unemployment data obtained from labour force surveys. Reconciliation of the unemployment statistics generally requires data adjustments in both sources.

The measurement of unemployment in labour force surveys is generally based on the international definition, which classifies as unemployed all persons, above the age specified for measuring the EAP, who satisfy three conditions[72]: "without work during the reference period"; "currently available for work"; and "seeking work during a specified recent period", generally the last four weeks in survey applications.

Seeking work means taking active steps to look for work and goes beyond registration at public or private employment exchange (for the purpose of obtaining a job offer) and also includes other means of looking for work, such as direct application to employers, placing or answering advertisements, seeking assistance from friends or relatives, and looking for resources and permits to establish one's own enterprise. For comparison with administrative data on registered jobseekers, the labour force survey data should therefore be restricted only to those persons who reported having looked for work by registering with the employment offices consistent with the coverage of the administrative data.

71 Pember, Bob, Labour statistics based on administrative records: Guidelines on compilation and presentation, ILO East Asia Multidisciplinary Advisory Team (ILO/EASTMAT), ILO Regional Office for Asia and the Pacific, Bangkok, 1997.

72 The international definition of unemployment provides an exception for persons without work who are making arrangements to start work at a date subsequent to the reference period (future starts), who are classified as "unemployed" irrespective of their job-search activity, provided they were without work during the reference period and currently available for work.

The administrative data on registered jobseekers may also need to be adjusted for comparability with the survey unemployment data. Some administrative data include jobseekers already employed, registering with the employment office in search of a different job. This category of jobseekers should be excluded from the administrative count for comparison with the survey data. Also, in principle, for comparison purposes, only those jobseekers in the administrative data should be retained who have registered with the employment offices in the last four weeks or have made contacts in the last four weeks for follow-up on earlier registrations.

Table 3 illustrates the reconciliation of survey unemployment data with administrative data on registered jobseekers. Total unemployment according to a national labour force survey was 2,486,200 in 2006. The majority of the unemployed (1,537,600) were looking for work through direct contact with employers or through friends and relatives and other non-official means. By contrast, 948,600 reported looking for work through registration at public or related employment offices.

The corresponding administrative figure (903,514) is obtained after adjusting the data on registered jobseekers (1,057,086) for those already employed but seeking alternative or additional employment (63,054) and those who have not been active in the last four weeks (90,518).

Thus, the difference between the survey estimate of jobseekers at public or related employment offices (948,600) and the corresponding adjusted administrative data (903,514) is about 45,100, indicating a percentage difference of about 5%. Similar values of the percentage un-

TABLE 3

RECONCILIATION TABLE: UNEMPLOYED PERSONS SEEKING WORK AT PUBLIC OR RELATED EMPLOYMENT OFFICES

Labour force survey		Administrative records	
Total unemployed persons	**2,486,200**	Total registered jobseekers	**1,057,086**
Reporting other means of job search only	**-1,537,600**	Registered jobseekers already employed	**-63,054**
		Registered jobseekers, no active job search in last four weeks	**-90,518**
Reporting job search at public or related employment offices	**948,600**	Registered jobseekers within scope	**903,514**
Unaccounted difference	**-45,100**		
Reconciled estimate	**903,500**	Reconciled estimate	**903,514**

accounted difference can be calculated for other years, different geographical regions and for men and women separately.

25.4 COMPARABILITY OVER TIME

In addition to reconciliation of survey data with external sources, the comparability of survey data over time should be examined as lack of coherence of the data over time has important impacts on the credibility of the results perceived by the public. Issues concerning comparability over time arise whenever changes in concepts, organisation or methodology are introduced in the labour force survey programme. Their impact on the results is often difficult to isolate and the resulting breaks in the time series hamper the correct assessment of the evolution of the labour market.

26. SAMPLING ERRORS AND PRECISION OF ESTIMATES

26.1 SAMPLING ERRORS

Like in all sample surveys, the results of labour force surveys are subject to sampling errors. Sampling errors arise due to the fact that the survey does not cover all elements of the population but rather only a selected portion. The sampling error of an estimate is based on the difference between the estimate and the value that would have been obtained on the basis of a complete count of the population under otherwise identical conditions.

Although survey data are subject to errors from a variety of sources, knowing about the magnitude of sampling errors is of fundamental importance. Information on sampling errors is crucial for interpreting the survey results. It allows decisions to be made on the precision of the estimates and on the degree of confidence that may be attached to them. In the same vein, it allows decision on the degree of detail with which the survey data may be meaningfully tabulated and analysed. Information on sampling errors is also crucial in determining whether the survey estimates of change over time or the estimates of differences between two or more population subgroups are statistically significant and not just produced by chance.

Information on sampling errors is also crucial for sample design and evaluation. While cost and other practical considerations affect the sample design of a survey, rational decisions on the choice of sample size, sample allocation among strata, clustering and estimation procedures can only be made on the basis of detailed knowledge of their effect on the magnitude of sampling errors in the resulting statistics obtained from the survey.

In principle, sampling errors may be decomposed into two components: (i) sampling bias; and (ii) sampling variance. Sampling bias reflects the systematic errors that may occur due to the failures of the sample design, such as, for example, certain elements of the population receiving zero probability of selection. The sampling variance, on the other hand, reflects the uncertainty associated with a sample estimate due to the particular sample used for its calculation, among all possible other samples that could have been selected from the frame with the same sampling design.

As an illustration, consider a country with a population of 2 million. A simple random sample of 10,000 persons is drawn from this population: 7,000 are found to be in the labour force and among them 500 unemployed. The estimates of the total number of persons in the labour force and the total number of unemployed persons in the country are accordingly 1,400,000 and 100,000 respectively.

Clearly, a different sample may produce different sample counts and therefore different estimates of the national labour force and unemployment. For example, after drawing a different sample of the same size, we found 8,000 persons in the labour force and 400

among them unemployed. With the new sample results, the survey estimate of total labour force is 1,600,000 and of total unemployment 80,000.

Suppose we take all possible samples from the population in exactly the same way (there are essentially an infinite number of possible such samples) and for each sample we compute the estimate of total unemployment in the country. We would have the same number of estimates as the number of possible samples. These estimates can be used to construct a single-variable distribution, called the sampling distribution of the unemployment estimator. It is the **distribution** of the estimates of unemployment that would result from taking all possible samples from the population.

The resulting sampling distribution may look like the bell-shape density distribution shown below:

The sampling bias of an estimator is the difference between the mean of its sampling distribution and the true value of the variable being estimated. Thus, in the above example, if the true total number of unemployed person is 85,000, the sampling bias of the estimator would be:

Sampling bias = 80,000 – 85,000 = -5,000

In relative terms, the sampling bias is about -6%. The sampling variance of the estimate of the number of unemployed in the country is the variance of the sampling distribution. The square root of the sampling variance is the standard deviation, which, in this example, is as follows:

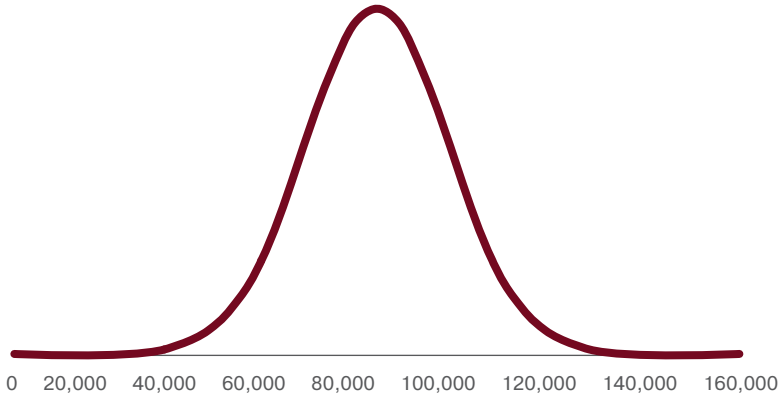
Standard deviation = 14,000.

A remarkable feature of probability sampling is that the shape of the sampling distribution and in particular its sampling variance or equivalently its standard deviation can be estimated using a single sample, not requiring repeated sampling of the population as suggested in the above illustration.

26.2 CALCULATION OF THE SAMPLING VARIANCE

The calculation of the sampling variance of survey estimates for complex multi-stage designs is generally based on the following principle: the variance contributed by the later stages of sampling is, under broad conditions, reflected in the observed variation among the sample results for first-stage units. Thus, the sampling variance of a variety of statistics, such as totals, means, ratios, proportions, and their differences, can be obtained on the basis of totals calculated for primary sampling units (PSUs).[73]

FIGURE 2
SAMPLING DISTRIBUTION OF UNEMPLOYMENT ESTIMATES



73 Verma, Vijay, Sampling Methods, Manual for Statistical Trainers Number 2, Statistical Institute for Asia and the Pacific (SIAP), Tokyo, Revised 2002.

The methodology is illustrated below in the context of labour force surveys. Consider a national labour force survey with estimated labour force (x) and unemployment (y). A typical design of the survey is a stratified two-stage sample with enumeration areas from the latest population census as primary sampling units (PSUs) and households or dwellings (or clusters of households or dwellings) as secondary sampling units (SSUs).

Let m^h be the number of sample PSUs in stratum h selected from a total of Mh sample PSUs in stratum h. The sample estimates of total unemployment and total labour force are expressed respectively as:

$$y = \sum_h \sum_i y_{hi}$$

$$x = \sum_h \sum_i x_{hi}$$

where y_{hi} and x_{hi} are the corresponding sum of sample results for sample PSU i,

$$y_{hi} = \sum_i w_{hij} y_{hij} \quad \text{and} \quad x_{hi} = \sum_i w_{hij} x_{hij}$$

where w_{hij} is the sampling weight associated to the sample person j with labour force status y_{hij} and unemployment status x_{hij} .

Under the principle stated above, the sampling variance of the estimates y and x can be calculated by:

$$\text{var}(y) = \sum_h (1 - f_h) \frac{m_h}{m_h - 1} \sum_i (y_{hi} - \frac{y_h}{m_h})^2$$

$$\text{var}(x) = \sum_h (1 - f_h) \frac{m_h}{m_h - 1} \sum_i (x_{hi} - \frac{x_h}{m_h})^2$$

where $f_h = \frac{m_h}{M_h}$ is the sampling fraction and $y_h = \sum_i y$ and $x_h = \sum_i x_{hi}$.

The sample principle applies for the calculation of the sampling variance of means, proportions, percentages, and ratios where both the numerator and denominator are sample estimates, such as for the calculation of the sampling variance of the unemployment rate:

$$r = \frac{y}{x}$$

The sampling variance is derived by Taylor linearization of the statistic. This gives

$$\text{var}(z) = \sum_h (1 - f_h) \frac{m_h}{m_h - 1} \sum_i (z_{hi} - \frac{z_h}{m_h})^2$$

where $z_{hi} = \frac{1}{x}(y_{hi} - r x_{hi})$

Taylor linearization can be applied for the calculation of the sampling variance of more complex statistics, such as differences of ratios, ratio of ratios, regression coefficients, etc. For example, in the case of differences of ratios:

$$r - r' = \frac{y}{x} - \frac{y'}{x'}$$

The sampling variance can be calculated by:

$$\text{var}(r - r') = \sum_h (1 - f_h) \frac{m_h}{m_h - 1} \sum_i (z_{hi} - \frac{z_h}{m_h})^2$$

where $z_{hi} = \frac{1}{x}(y_{hi} - r x_{hi}) - \frac{1}{x'}(y'_{hi} - r' x'_{hi})$.

The underlying assumptions of the methodology of variance estimation described here are that (a) at least two PSUs are drawn from each stratum and (b) the sample PSUs are selected independently of each other, with random choice and with replacement. These conditions are not always exactly satisfied, but they may sufficiently well approximate many situations in practice.

When at least two PSUs per stratum are not available, "pairing" or "collapsing" of similar strata to define new strata can be implemented, so that each new stratum contains at least two selections which are then assumed to be independent. When the PSUs are too small, variable or otherwise inappropriate to be used directly for variance estimation, more suitable units are defined, for example by random grouping of units within strata and linking or combining of units across strata.

TABLE 4

RECONCILIATION TABLE: UNEMPLOYED PERSONS SEEKING WORK AT PUBLIC OR RELATED EMPLOYMENT OFFICES

	Survey estimate	Standard error	Relative standard error	Confidence interval	
				Lower bound	Upper bound
Population 15+ years	4,092		0.2%		
Labour force	2,047	61	0.3%	2,035	2,059
- Employed	1,963	61	0.3%	1,951	1,975
- Unemployed	84	1.8	2.1%	81	88
			(%)		
LFPR	66.2%	0.17%-pt	0.3%	65.9%	66.5%
Unemployment rate	4.1%	0.09%-pt	2.2%	3.9%	4.3%
- Men	4.5%	0.12%-pt	2.5%	4.3%	4.7%
- Women	3.9%	0.11%-pt	2.8%	3.7%	4.1%

Finally, it should be mentioned that there are other methods of variance estimation for complex designs. Some of these alternative methods are based on comparison among replications of the full sample, such as jack-knife repeated replications, balanced repeated replications and bootstrapping. A major feature of these procedures is that, under general conditions for their application, the same and relatively simple variance estimation formula holds for statistics of any complexity.

26.3 PRECISION OF ESTIMATES: A NUMERICAL EXAMPLE

The following table should help to illustrate the use and interpretation of different measures of sampling errors in a labour force survey. The table shows the estimates of the main labour force indicators and their standard errors calculated using the methodology described in the preceding section.

One use of the standard error is to assess the level of precision of survey estimates. A low relative standard error indicates a high precision of the estimate. In general, the lower the relative standard error of an estimate, the higher the precision of the estimate. The relative standard error of an estimate is the ratio of the standard error to the size of the estimate.

Accordingly, it can be deduced from Table 1 earlier in this part of the Guide Book that the working age population (persons 15 years old and over) is more precisely estimated than the labour force. The relative standard error of the estimate of the working age population is 0.2%, while the relative standard error of the estimate of the labour force is 0.3%. Similarly, the labour force is more or less as precisely estimated as employment (relative standard error equal to 0.3%), which is substantially more precisely estimated than unemployment (relative standard error 2.1%). These results reflect the fact that the estimates based on larger effective sample size are generally more precise than estimates based on smaller effective sample size.

The table also shows that ratios such as the LFPR and the unemployment rate are about as precisely estimated as the corresponding aggregate levels of labour force and unemployment.

Another use of the standard error is for the calculation of confidence intervals. Under certain

broad assumptions, it can be stated that the true value of the variable of interest lies in between the survey estimate and a multiple of the standard error, with a certain degree of probability. In general, if y represents the survey estimate of a variable of interest, the true value of the variable represented by, say, θ lies with $(1-\alpha)\%$ confidence in the following interval:

$$y - z_{(\alpha/2)}\sigma \leq \theta \leq y + z_{(\alpha/2)}\sigma$$

where σ is the standard error of the estimate and $z_{(\alpha/2)}$ is the value of the standard normal distribution corresponding to the $(1-\alpha)$ % confidence probability. For a 95% confidence probability, $\alpha = 5\%$ and $z_{(\alpha/2)}$ is approximately equal to 1.96.

With respect to the results shown in Table 1, it can be stated, for example, that the true value of the total number of unemployed is within the following interval,

$$84{,}000 - 1.96 \times 1{,}800 \leq \theta \leq 84{,}000 + 1.96 \times 1{,}800$$
$$81{,}000 \leq \theta \leq 88{,}000$$

Similarly, it can be calculated that the unemployment rate lies with 95% confidence within the following interval:

$$4.1\% - 1.96 \times 0.09\% \leq \theta \leq 4.1\% + 1.96 \times 0.09\%$$
$$3.9\% \leq \theta \leq 4.3\%$$

A further use of the standard error is for determining the statistical significance of differences in survey estimates. Survey estimates of two variables of interest differ significantly from each other if zero is not in the confidence interval of their difference. An approximate significance test of the difference between two estimates may be done by simply comparing the confidence intervals of the estimates and checking their overlap. If they do not overlap, the two estimates may be said to be significantly different. Accordingly, from Table 1, it can be concluded the unemployment rate of men (estimated at 4.5% with confidence interval 4.3%–4.7%) is significantly different than women's unemployment rate (estimated at 3.9% with confidence interval 3.7%–4.1%).

As it is not practical to compute and report sampling variances for all the published statistics of the labour force survey, certain countries give general variance estimates using the approximate relationship between the variance of an estimate and its size (expressed by $\text{var}(y)/y = a + by$). Others provide approximate values of the relative standard error for different sizes of the estimate, as in Table 5 below containing data from the Australian labour force survey:

TABLE 5
RELATIVE STANDARD ERROR BY SIZE OF ESTIMATE

Size of estimate (persons in the category)	Relative standard error (% of the estimate)
5,000,000	0.14
2,000,000	0.31
1,000,000	0.54
500,000	0.83
200,000	1.37
100,000	1.97
50,000	2.80
20,000	4.45
10,000	6.29
5,000	8.91
2,000	14.09
1,000	19.93
500	28.19

27. NON-SAMPLING ERRORS AND WEIGHTS ADJUSTMENT

27.1 INTRODUCTION

In addition to sampling errors, survey data are subject to different types of non-sampling errors (coverage errors, non-response errors, response errors, and other errors such as editing, coding and processing errors).[74] In many situations, the non-sampling errors may have considerable impact on the quality of the survey results. Analysts should therefore pay careful attention to non-sampling errors when using the survey data. In particular, they should examine the magnitude of these errors, and verify to what extent some of the errors have been corrected through the adjustment of the sampling weights when the sample data are extrapolated to population aggregates. In this section, the main types of non-sampling errors are described and the weighting schemes used for adjusting some of their impacts are briefly reviewed.

27.2 COVERAGE ERRORS

Probability sampling requires each element in the target population to have a known non-zero probability of being selected in the sample. This condition is violated if the target population is not fully represented in the sample frame and if the sample selection of units from the frame is not according to the procedures specified in the sample design. The violation of these conditions generates coverage errors.

Coverage errors may occur explicitly from the use of imperfect frames (under-coverage, over-coverage, or duplication of units) or implicitly due to confusion in the boundaries of units and in rules of association between units of different types. In multi-stage household surveys such as most national labour force surveys, coverage errors may arise at any of the stages of selection of sample units: area units; households or dwellings; and persons.

In labour force surveys based on recently constructed sampling frames, one can expect the lists of enumeration areas to be complete and up to date. However, serious coverage problems can arise when the sampling frame is out of date or when maps of the area units are inexistent or poorly drawn. Newly settled areas, especially in the outskirt of cities, are often not properly represented in out-of date area frames. Also, inadequate maps and descriptions make the identification and delineation of the selected sample areas by survey workers extremely difficult in practice.

At the next stage of sampling, when households or dwellings are to be drawn from the sample areas, coverage errors may again occur. Lists of dwellings or housing units are generally less durable than area frames. Thus, the coverage errors mentioned in connection with area units are also present and generally more so in the case of households or dwellings. In particular, if the time lag between listing and interviewing is not sufficiently short, the risk of coverage errors is high. To minimise such errors, it is necessary to freshly list the households or dwellings in the selected sample areas or to draw supplementary samples based on information from housing permits or similar sources. Other methods of minimising coverage errors at this stage of sampling is to select the households or dwellings in small contiguous groups or clusters rather than singly.

Coverage errors may also occur at the stage of selection of individual persons in the sample household. Non-coverage can arise due to a failure to identify some eligible persons, such as lodgers, domestic workers or other non-family members of the household. It can also happen due to incorrect information on personal characteristics needed for sample selection, for example if the age a person is incorrectly recorded

74 Hussmanns, Ralf, Farhad Mehran, and Vijay Verma, ILO Manual on Concepts and Methods: Surveys of Economically Active Population, Employment, Unemployment and Underemployment (Part II).

as below the age set for measuring labour force characteristics (under-coverage error) or, alternatively, the age is incorrectly recorded as above the threshold age (over-coverage error).

27.3 NON-RESPONSE ERRORS

Non-response occurs due to failure to obtain the required information from the units selected in the sample (unit non-response) or to failure to obtain some items of information for the selected unit (item non-response). The distinction between unit non-response and item non-response need not be rigid in all circumstances. For example, in labour force surveys with a sample rotation scheme, some units may remain in the sample over a number of rounds and it may be possible to impute some information (such as age, sex, etc.) from an earlier round to a later round where non-response has occurred. This procedure will have the effect of changing some cases of unit non-response to that of item non-response.

Unit non-response may occur due to an incorrect address for the sample household, inaccessibility of certain dwellings, a refusal by the sample household to be interviewed, because no one was at home when the interviewer contacted the household, or if information on a particular person in the sample household could not be obtained for some reason. Vacant or demolished dwellings, non-existent or out-of-scope addresses, such as finding an enterprise or workshop instead of a household dwelling, are not to be considered as unit non-response.

In some countries, in the presence of unit non-response, the designated sample households that cannot be interviewed are substituted with other households in the vicinity. This practice is not recommended as the disadvantages far outweigh the benefits.[75] Substitution introduces extra bias into the estimates because responding households often have characteristics that differ from those of non-responding households. Also, substitution generates practical problems in fieldwork, including the use of time and risk of error in attempting to substitute and the difficulty of checking and monitoring the process.

Non-response can generally be measured directly from records internal to the survey. The overall non-response rate taking into account the three possible levels of non-response (dwellings, households and persons) may be calculated as 1 minus the response rate expressed by:

$$Non-response-rate = 1 - \frac{m'}{m} \times \frac{n'}{n} \times \frac{p'}{p}$$

where m = number of dwellings selected in the sample

m′ = number of dwellings successfully contacted

n = total number of sample households in contacted dwellings

n′ = number of successfully interviewed households

p = number of eligible persons listed in the interviewed households

p′ = number of eligible persons successfully interviewed.

The impact of non-response on survey results is, in principle, similar to that of under-coverage. In estimating population totals, the effect will depend on the relative value of missing units. In estimating averages, percentages, ratios or proportions, the effect will depend on the extent to which non-responding units differ in characteristics from the responding units.

Corrections for non-response errors can be made by inflating the survey estimates by the inverse of the response rate (1 minus the non-response rate defined above). For this purpose,

75 Vehovar, Vasja, "Field Substitution and Unit Non-Response," Journal of Official Statistics, Vol. 15, No. 2, 1999, pp. 335-350.

the sample is divided into a number of classes expected to be homogenous with respect to the survey variables. Then, the correction for non-response is applied separately for each class, i.e. the non-respondents in each class are assumed to be similar to the respondents in that class. The classes may be defined in terms of small geographical areas or simple socio-demographic characteristics of individuals.

27.4 RESPONSE ERRORS

Response errors refer to errors originating at the data collection stage.[76] In relation to an individual respondent, response error refers to the deviation between the response obtained from the survey and the actual or true value for the individual. They can occur because the respondent may be unwilling to divulge certain information (intentional error). They may also occur because the respondent does not know the answer to the question asked or does not fully understand the meaning of the question (unintentional error). They can also occur by memory lapses, such as by forgetting to report an event (omission) or incorrectly reporting its timing (telescoping).

Response errors may occur not only because of errors made by the respondent but also because of errors made by the interviewer or by the instrument used for measurement. Interviewers may introduce errors through haste and misrecording of responses, misunderstanding of the survey concepts and procedures, preconceptions and subjective biases. Sometimes, the questionnaire itself may be faulty, with wrong wording and an incorrect skipping pattern, causing errors to be made by interviewers and respondents.

An additional source of uncertainty arises from the fact that, in many labour force surveys, for reasons of economy, a certain degree of response by proxy has to be tolerated.

The measurement of response errors is often one of the most difficult parts of quality assessment of survey data.[77] The measurement framework generally used is based on the following decompositions:

Response errors = Response bias + Response variance

Response variance = Simple response variance + Correlated response variance[78]

Response bias is the effect of systematic errors that occur as a consequence of the general conditions under which the survey response is obtained. In a continuing labour force survey with more or less standardised design and procedures, essentially similar response biases may affect the survey results from all rounds. Consequently, such biases may go undetected unless the results are validated through external comparisons. Response biases are generally measured by re-interviews using similar but improved procedures (re-interviews with reconciliation) or through comparison of the survey results with data from external sources known to be more accurate.

Response variance is the result of random factors that affect individual responses. Surveys repeated under the same procedures and identical general conditions would not necessarily give the same results. Results from one repetition to another would differ even with the same sample. Response variance is a measure of this variability. Two components of response variance may be distinguished: simple response variance and correlated response variance.

Simple response variance is the measure of the variability of responses on individual items of the survey. The measurement of simple response variance requires independent repetitions of data collection under identical conditions, although, in practice, it is difficult to satisfy the conditions exactly. In labour force

76 Errors arising at subsequent stages of data processing such as data entry, coding, and editing are not treated here.

77 Biemer, Paul P. and Gösta Forsman, "On the Quality of Reinterview Data With Application to the Current Population Survey," Journal of the American Statistical Association, December 1992, Vol. 87, No. 420, pp. 915–923.

78 Expressed in more precise terms, the response variance is the weighted sum of the simple response variance and the correlated response variance.

surveys, re-interviews under similar conditions within a short time after the main survey are often used to measure the simple response variance of the main labour force indicators (re-interviews without reconciliation).

Correlated response variance arises from the variability of the work of interviewers. It indicates the lack of uniformity and standardisation in interviewers' work. The measurement of correlated response variance is generally conducted through re-interviews under experimental design conditions. The basic requirement is a degree of randomisation in the allocation of sample units to individual interviewers. The same approach applies also to measuring the correlated effects of supervisors, editors, coders, etc.

27.5 WEIGHTS ADJUSTMENT

Some of the non-sampling errors may be corrected to a certain extent by appropriate adjustments of the sampling weights used for extrapolating the sample results to population aggregates.

In general, the overall weight attached to each sample unit is formed by three components: the weights reflecting the sample design; the adjustment factors for non-response; and the final factors to calibrate the results to population controls obtained from external sources.

In a two-stage stratified sampling design, the sampling weight for sample households and individuals in sample PSU i in stratum h is given by the inverse of the probability of selection of the households and individuals in that PSU and stratum:

$$w_{hi} = \frac{1}{p_{hi}}$$

where p_{hi} is the probability of selection of the sample households and individuals in PSU i of stratum h.

The design weight for each sample PSU may be adjusted for non-response households or individuals by inflating the sampling weights by the inverse of the response rate:

$$w'_{hi} = w_{hi} \times \frac{1}{R_{hi}}$$

where is the response rate in PSU i of stratum h. It can be derived by the proper combination of the non-response rates defined in the previous section.

The adjusted weights may in turn be further adjusted to population totals obtained from population censuses and demographic analysis, if these are considered to provide more accurate results. This process of weight adjustment is called calibration and involves the calculation of the minimum required adjustment factor to be applied to the weights so that the resulting survey estimates agree exactly to the known population aggregates.[79]

Mathematically, the adjustment procedure may be described in matrix notation as follows:

$$w''_{hi} = w'_{hi} \times (1 + x'_{hi} \times \lambda)$$

where w'_{hi} is the adjusted weight obtained from the previous round of the adjustment process, $x'_{h,i}$ is the vector of auxiliary variables for which the aggregate totals are known (for example, male and female population), and λ is adjustment factor given by:

$$\lambda = T^{-1} \times (t_x - \hat{U}_x)$$

$$T = \sum_{hi} w'_{hi} \times x_{hi} x'_{hi}$$

where t_x is the vector of known population aggregates and $\hat{U}_x$ the vector of corresponding survey estimates using the adjusted weights obtained from the previous round of the adjustment process.

79 Särndal, Carl-Erik, and Jean-Claude Deville, "Calibration Estimators in Survey Sampling," Journal of the American Statistical Association, June 1992, Vol. 87, No. 418, pp. 376–382.



Statistical Capacity Building Division
Statistics Department
Chief Economist Complex
African Development Bank
BP. 323, 1002 Belvedere, Tunis
Tel: (216) 7110 2343
Fax: (216) 7110 3743
Internet: http://www.afdb.org
E-mail: statistics@afdb.org