

Estimation of world and regional unemployment

Farhad Mehran¹

Introduction

A statistical implication of the globalization of the world economy is the requirement for more precise, comparable data among countries. Another statistical implication of the globalization process is the requirement for global information on broad geographical regions and the world at large. One particular area of global concern is the evolution of world employment and unemployment.

Estimates of world and regional employment and unemployment serve a number of purposes. First, they provide a comprehensive view of the employment situation in the world and its major geographical regions. Its evolution through time may be analysed in conjunction with other economic and social variables to draw conclusions on future trends in various aspects of the global economy, and of regional and national labour markets. Another role of world and regional estimates is the provision of global benchmarks against which the economic and labour market performance of individual countries may be compared and assessed. This aspect should not be underestimated as countries are increasingly in competition for foreign investment and local production of global goods, and information on their relative performance plays a significant role in many trade and financial decisions. Still another aspect of world and regional estimates is their high visibility by the media and their effect on public policy. A single number alone, like “250 million working children”, is sufficient to draw the attention of the world to a major social problem and to invite everyone to do something about it.

Accordingly, a project has recently been launched by the ILO Bureau of Statistics to produce, on an annual basis, world and regional estimates of employment and unemployment, with breakdowns by sex and broad age groups. The project is carried out within the framework of the ILO programme on key indicators of the labour market (KILM) conducted by the ILO Department of Employment and Training in collaboration with the ILO Bureau of Statistics.² Two of the eighteen KILM indicators are the unemployment rate and the employment-population ratio. These two indicators are also included in the Minimum National

¹ Bureau of Statistics, International Labour Office. E-mail: mehran@ilo.org. The author would like to express his gratitude to Ms. Adriana Mata-Greenwood and Mr. Marc Copin for the translation of the summary of this paper into Spanish and French respectively.

² ILO, *Key Indicators of the Labour Market 1999*, Geneva, 1999.

Social Data Set (MNSDS) formulated by the Expert Group on the Statistical Implications of Recent Major United Nations Conferences, and endorsed by the UN Statistical Commission in 1997. These two indicators relate to concerns expressed, in particular, by the World Summit for Social Development (Copenhagen, March 1995) and the World Conference on Women (Beijing, September 1995) on the need for expansion of productive employment and reduction of unemployment.

The purpose of the present paper is to describe a possible approach for obtaining world and regional estimates of these indicators. The methodology is applied here on a trial basis for obtaining 1995 unemployment estimates with distinction by sex and broad age group (youth and adult). The data should be considered as preliminary and subject to revision. The results will be evaluated and further methodological work will be carried out in the next twelve months, before adopting a standard procedure for regular publication of world and regional estimates of employment and unemployment, starting, possibly, in the next issue of KILM.

World estimation

The methodology used to obtain world and regional estimates of unemployment is based on a top-down approach; first, the world estimates are obtained; and then these are broken down into regional components using appropriate statistical techniques. The world estimates are derived on the basis of a sample of available country data sets. A representative stratified sample of 33 countries with appropriate unemployment data is selected and the national data adjusted to the international standards,³ and weighted to transform sample estimates to world estimates.

The ILO database LABPROJ on *estimates and projections of the economically active population, 1950-2010*⁴ form the sampling frame for selecting the sample of countries. The database covers 178 countries and territories with a population of at least 200,000 in 1990 (164 countries when the European Union is counted as one entity). For each country or territory, and for each decennial year

³ ILO, *Resolution concerning statistics of the economically active population, employment, unemployment and underemployment*, 13th International Conference of Labour Statisticians, Geneva, 1982. (<http://www.ilo.org/public/english/120stat/res/ecacpop.htm>).

⁴ LABPROJ data available in the form of four ASCII text files in two diskettes (*Economically Active Population 1950-2010, Fourth Edition, Diskettes User's Guide*, International Labour Office, Geneva, 1997).

between 1950 and 2010 and for 1995, there are data on total population, activity rate and labour force disaggregated by sex and five-year age group (from 10 to 64 years and 65 years and over). Estimates of the distribution of the labour force by sex and major sector of the economy, i.e., agriculture, industry and services, are also available for the period 1950-1990, with separate data on manufacturing for 1980 and 1990.

The 1995 data of LABPROJ was used to stratify the 178 countries and territories according to six regional groupings and three size-categories of the labour force. The regional groupings are those used in the presentation of the KILM indicators. They combine geographical positions and developmental rankings as follows: KILM Region-1 Developed/Industrialised countries (the European Union and 10 other countries); KILM Region-2 Transition countries (27 countries); KILM Region-3 Asia and the Pacific (28 countries); KILM Region-4 Latin America and the Caribbean (31 countries); KILM Region-5 Sub-Saharan Africa (44 countries); and KILM Region-6 Middle East and North Africa (23 countries).

The three size-categories include “large” countries, i.e., the largest countries in LABORSTA that make up together just more than two-thirds of the world labour force in 1995. Nine countries and the European Union qualify, to which we have added the next largest country, Nigeria, to include one African country in the stratum. All eleven “large” countries are included in the sample without sampling.

The “middle-size” countries are the next largest countries forming two-thirds of the remaining world labour force in 1995. There are thirty “middle-size” countries as defined here, to which we have added four in order to cover together with the “large” countries just more than 90% of the 1995 world labour force, a convenient round figure. A random sample of eleven of them is then selected with probability one-third.

The “small” countries form the third stratum. There are 119 “small” countries in LABPROJ, making up together less than 10% of the world labour force in 1995. They are stratified by regions and eleven are selected with probability of selection equal to one-tenth.

The thirty-three countries forming the sample are listed in the following table. The stratified sampling approach is found to be convenient as it allows the concentration of effort on a limited number of countries that have non-negligible influence on the world and regional totals, yet maintaining representation of the remaining countries.

**Table 1. Sample Countries for
World and Regional Estimation of Unemployment**

<u>KILM REGIONS</u>	<u>LARGE</u>	<u>MIDDLE-SIZE</u>	<u>SMALL</u>	<u>TOTAL</u>
1. DEVELOPED COUNTRIES	European Union United States Japan	Turkey	New Zealand	5 (11)
2. TRANSITION ECONOMIES	Russian Federation	Ukraine Poland	Lithuania Slovenia	5 (27)
3. ASIA AND PACIFIC	China India Indonesia Bangladesh Pakistan	Viet Nam Thailand Malaysia	Singapore	9 (28)
4. LATIN AMERICA AND THE CARIBBEAN	Brazil	Mexico Colombia	Bolivia Jamaica	5 (31)
5. SUB-SAHARA AFRICA	Nigeria	Uganda	Côte d'Ivoire Zambia Mauritius	5 (44)
6. MIDDLE EAST NORTH AFRICA	-	Egypt Morocco	Syria Palestine (Gaza)	4 (23)
<u>TOTAL</u>	11 (11)	11 (34)	11 (119)	33 (164)
LABOUR FORCE (%)	(67%)	(23%)	(10%)	(100%)

Notes: Totals refer to number of sample countries in a particular category. Figures in parentheses refer to the corresponding number of countries in the LABPROJ database. "Labour force" refers to the percentage of world labour force in the particular category according to LABPROJ.

Regional estimation

As the number of sample countries in each region is small, sufficiently

accurate estimates of regional employment and unemployment may not be obtained by direct methods. A preferred approach, adopted here, is to disaggregate the world estimates into their regional components in much the same way as national statistical offices derive small area statistics by combining survey estimates with correlated data. Below, the main ideas are highlighted.

Because the designated “large” countries are in the sample with certainty, the small area estimation methodology is applied only to the middle-size” and “small” countries within each region.

To fix ideas let us consider a particular region, say, Latin America and the Caribbean. There are 31 countries in this region one of which, Brazil, is a large country. Out of the 30 middle-size and small countries in this region, four are in the sample, Mexico, Columbia, Bolivia, and Jamaica. The challenge, therefore, is to use the information for these four sample countries to obtain regional estimates for all the middle-size and all the small countries in the region. To do so, the small area estimation technique borrows “strength” from information on other variables and on other regions.

Let “i” represent a particular region, and suppose the objective is to obtain the estimate of total unemployment for a particular year and a particular sex-age category in the middle-size or small countries of the region “i”. The corresponding regional labour force statistics are considered given and obtained from the LABPROJ database mentioned earlier. Let X_i denote this number. The required regional unemployment, denoted by Y_i is related to X_i by the simple relationship

$$Y_i = u_i X_i$$

where u_i is the corresponding unknown regional unemployment rate.

A direct estimate of the regional unemployment rate u_i can be obtained by calculating the unemployment rate of the appropriate sex-age category of the sample countries in region “i”. This estimate may be expressed as

$$\hat{u}_i = y_i / x_i$$

where y_i denotes the total unemployment of the particular sex-age category in the sample countries of region “i” and x_i the corresponding labour force.

As another possible estimate of u_i , one could consider the use of the same unemployment rate but calculated for all sample countries irrespective of the region. This estimate may be expressed by

$$\hat{u} = y / x$$

where y denotes the total unemployment of the particular sex-age category in all sample countries, irrespective of the region, and x denotes the corresponding labour force.

In comparing the two estimates, it should be noted that the advantage of one is the drawback of the other. The first estimate, \hat{u}_i , obtained directly from the sample, is essentially unbiased if the sample on which it is based is representative of the countries of the region. The precision of the estimate is, however, low because the data used for its calculation are based on a limited number of countries. On the other hand, the second estimate, \hat{u} , should have a higher precision than \hat{u}_i , because it is based on a significantly larger number of countries, but in contrast with \hat{u}_i , it is biased because it does not refer to the particular region in question.

The synthetic methods of small area estimation combines the two estimates by appropriate procedures to produce various weighted estimates, two of which are given below. The first, generally referred to as the regression estimate, may be expressed as the weighted estimate,

$$\hat{u}_i^{\text{reg}} = w_i \hat{u}_i + (1 - w_i) \hat{u}$$

where the weight w_i is the ratio of the sample estimate of the correlated variable, the regional labour force, to its known value. In terms of the notation introduced earlier, w_i is defined by $w_i = (x_i / f_i) / X_i$, where $f_i = n_i / N_i$, is the sampling fraction, with n_i equal to the number of sample countries in region i and N_i the total number of countries in that region. Thus, the closer the sample estimate of regional labour force is to its known value, the higher the weight of the sample estimate of the unknown regional unemployment rate. The procedure is reasonable because it corrects itself if the sample estimate is found to be inadequate where known figures are available.

An improved synthetic estimate is given by

$$\hat{u}_i^{\text{sd}} = w_i^* \hat{u}_i + (1 - w_i^*) \hat{u}$$

where the weights are sample-size dependent, defined by $w_i^* = (x_i / f_i^*) / X_i$, with $f_i^* = \max(f_i, f)$ and $f = n/N$, n being the number of sample countries irrespective of regions, and N the total number of countries making up the world. This weighting scheme is shown to be close to the optimum weights which would minimize the average error of the estimate taking account of both bias and

precision.⁵

The regional estimates of unemployment produced in the first release of the KILM is based on the sample-size dependent estimation method described above. The results are to be evaluated against estimates obtained from more complex procedures, in particular, model-based methods (empirical Bayes and hierarchical Bayes estimation) which treat small area estimations like missing value problems.

Results

The results of an application of the above procedure to the 1995 KILM data are presented in the following table.⁶ The data show preliminary world and regional estimates of unemployment by sex and by broad age groups (youth, 15-24 years old; adult, 25 years and over). The regional details are limited to the developed countries (KILM Region-1), transition economies (KILM Region-2), and Latin America and the Caribbean countries (KILM Region-4). The estimates for the three other KILM regions could not be presented separately in this first edition, due to the fragility of the underlying data. Further work is being carried out by the ILO Bureau of Statistics to improve the base data as well as the methodology to derive the world and regional estimates.

According to the results shown in Table 2, the total number of unemployed in the world in 1995 was 137 million persons, corresponding to a world unemployment rate of 5.2%. Half of the world unemployment is estimated to have occurred in Africa, Asia and the Pacific and the Middle East, essentially because

Table 2. World and Regional Unemployment 1995
Preliminary estimates in 000's

Total Male	Female	Youth	Adult
------------	--------	-------	-------

⁵ Ghosh, M., and J.N.K. Rao, "Small area estimation: an appraisal," *Statistical Science*, Volume 9, Number 1, February 1994, pp. 55-76.

⁶ In a number of cases where data were not available for 1995, estimates from neighbouring years or non-KILM data sources were used: India (1993-94), Indonesia (1996); Malaysia (1992); Syrian Arab Republic (1991); and Gaza (1996). In the case of Bangladesh, China, Côte d'Ivoire, Nigeria, Uganda, and Zambia, as no suitable data could be found, in each case the strata regional estimates were used.

World	137,000	75,000	62,000	56,000	81,000
1. Developed countries	29,000	15,000	14,000	9,000	20,000
2. Transition economies	17,000	9,000	8,000	6,000	11,000
4. Latin American and the Caribbean	12,000	7,000	5,000	6,000	6,000
5. Other countries	70,000	38,000	32,000	30,000	40,000

Rates in %

World	5.2	4.7	5.9	8.9	4.0
1. Developed countries	6.5	6.0	7.3	12.4	5.4
2. Transition economies	8.2	8.2	8.2	16.7	6.6
4. Latin America and the Caribbean	6.2	5.3	8.0	12.0	4.1
5. Other countries	3.8	3.5	4.4	6.4	3.0

Source: ILO Bureau of Statistics, 1999. Composite estimation based on a sample of 33 national data sets.

of the large labour force in these regions. When adjusted for the size of the labour force, it is found that the average unemployment rate in these regions (3.8%) was in fact lower than the other regions. The highest regional unemployment rate for 1995 is recorded for transition economies (8.2%), probably reflecting the deep transformation that these economies were undergoing since 1991, with the breakdown of the Former Soviet Union.

The estimates in Table 2 also show that although the number of male unemployed was higher than the female number, in relative terms the burden of unemployment was heavier among women than men: the female global unemployment rate is estimated at 5.9% for 1995, against 4.7% for males.

A similar statement may be made concerning the youth as opposed to the adult population. The table shows that the incidence of unemployment among youths (8.9%) was more than twice the incidence of that of adults (4.0%). That the world youth unemployment rate should be higher than the corresponding adult rate is not surprising. It simply reflects the fact that the risk of unemployment is higher at labour market entry points than at any other points. But that the youth rate is more than twice the adult rate is revealing and points to the importance that youth unemployment should receive in national labour market policies.

The imbalance incidence of unemployment towards females and youths

appears to be a constant in all regions. Indeed, the 1995 female unemployment rate was higher than the male rate in all regions except in the transition economies where the two rates were about the same, at the relatively high value of 8.2%. The youth unemployment rate was, however, higher than the adult rate in all regions without exception. The widest gap recorded is for Latin America and the Caribbean where the youth unemployment rate (16.7%) was close to three times as high as the adult rate (6.6%) in 1995.

Although the global and regional figures presented here are based on data from a sample of countries and shown only for illustration purposes, they reveal the basic characteristics of unemployment at the world level, and constitute the start of a process to provide regular information on the evolution of employment and unemployment for large geographical areas in the world.