



**HAL**  
open science

# Comparability of Statistical Concepts in Labour Force Surveys

Giorgos Kouvatses, Gilbert Saporta, Antonis Spinakis

► **To cite this version:**

Giorgos Kouvatses, Gilbert Saporta, Antonis Spinakis. Comparability of Statistical Concepts in Labour Force Surveys. ETK & NTTS 2001, Eurostat, Jun 2001, Hersonissos, Greece. pp.907-910. hal-03059941

**HAL Id: hal-03059941**

**<https://cnam.hal.science/hal-03059941v1>**

Submitted on 13 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Comparability of Statistical Concepts in Labour Force Surveys

Giorgos KOUVATSEAS<sup>1</sup>, Gilbert SAPORTA<sup>2</sup> and Antonis SPINAKIS<sup>3</sup>

<sup>1,3</sup>QUANTOS SARL

191, rue Saint Honoré,

75001 Paris, FRANCE

e-mail: <sup>1</sup>gkouv@quantos-stat.com, <sup>3</sup>anspi@quantos-stat.com

<sup>2</sup>Conservatoire National des Arts et Metiers

292 rue Saint Martin,

75141 Paris Cedex03, FRANCE

e-mail: saporta@quantos-stat.com

**Abstract:** In the framework of an effort to examine whether there exist major differences in the definition of unemployment between national concepts and the one adopted by the International Labour Organization (ILO), as well as if labour force surveys (LFS) based on small samples can lead to valid results, a survey was conducted in France, Germany, Italy and the United Kingdom. The sample was chosen, in most cases, according to the methodology of sampling with quotas, taking under consideration various demographic features of the population, such as age, gender, professional status etc. and was approximately 1000 persons in each country. After controlling for several factors the analysis concluded that differences are present while the question of the sample size has to be further explored.

**Keywords:** Definition of unemployment, Labour Force Survey, small samples.

## 1. Introduction

At present Eurostat, in order to define the unemployment rates, conducts annually a Labor Force Survey (LFS) in the Member States of the European Union in accordance with Council Regulation (EC) No 577/98.

Eurostat, in order to conduct the LFS, processes the data collected from the countries based on a common questionnaire. However, national definitions are still in use, even though not used as the official EU figures, reflecting different conceptual paradigms. In the framework of the research regarding labor force methodology a survey was conducted in four countries of the European Union: France, Germany, Italy and the United Kingdom (UK) during the last week of June 2000 and the first week of July 2000. The scope of the survey was to explore the underlying conceptual differences that might exist between the national definitions and the definition adopted by Eurostat, in

order to identify the principles that designate the construction of national models for measuring unemployment.

Another methodological aspect of this study was the use of a relatively small data set of only 1000 persons, which compared to the sample sizes of the official surveys, which range from 15,000 to 150,000 persons, might seem inadequate. However, the sampling method mainly used in our study was quota sampling, which provides some methodological and economical advantages for samples of that size [1]. Between the two objectives exists strong interrelationship. Namely, poor sampling results in poor estimations and thus comparison of concepts is difficult if not misleading. Another objective of the study was to examine non-response, however the small size of the non-response subgroup allowed only for descriptive analysis, which due to the small length of the paper cannot be presented.

## **2. Methodological Considerations**

Italy and UK have adopted the ILO guidelines and thus do not have national definitions. Thus, their contribution in the analysis was to compare the application of the national definitions of Germany and France with ILO's, as well as to evaluate the precision of the small sample size.

BVA SA performed the sampling, using the quota method in France, Italy and UK, while simple random sampling was used in Germany. In France the sample size was 1008 individuals, 1010 in Italy, 1078 in Germany and 967 in UK, while the samples were weighted according to several demographic variables in order to increase precision. In a first phase the observations were classified in the three status categories (employed, unemployed, inactive) applying all definitions (national concepts and ILO) in all four countries. The ILO estimation was used for the evaluation of the small sample quota method, while the comparison of the definitions was performed with an analysis of variance procedure.

## **3. Results**

As it has been mentioned, the scope of the survey was to estimate the unemployment rates according to the national definition of each country and the definition adopted by Eurostat, which is formed following the ILO guidelines. The deviations of the estimations under Eurostat's definition from the ones published by Eurostat for the same period are theoretically able to reveal whether our estimation is valid or not. In June 2000 the unemployment rates published by Eurostat were 9.6% for France and 8.4% for Germany, while it was 10.7% in Italy for April 2000 and 5.5% in the UK for May 2000. The unemployment rate estimated by our study was in France 11.5% (95% Confidence Interval (CI): 8.55%-14.35%), in Germany 5.4% (95% C.I.: 3%-7.53%), in Italy 6.04% (95% C.I.: 3.29%-7.8%) and in the UK 8.9% (95% C.I.: 6.48%-11.43%). In all countries except France the estimation of the 95% CI for the unemployment rate, did not include the rate published by Eurostat.

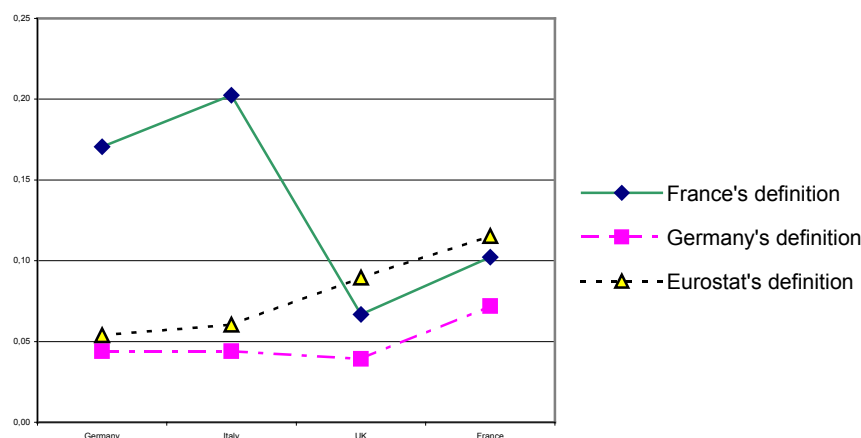
These deviations cast doubt on the reliability of derived estimations. Therefore we cannot assume that small samples have comparable efficiency. However, the design of

the study cannot shed light to whether the inefficiency is due to the size of the sample or to other factors like the selected quota variables. Nevertheless, we do not expect that the comparison of the definitions will be affected, as in any case the lack of precision should be present equivocally in all conceptual approaches.

### 3.1 Analysis of Variance

The analysis of variance was conducted, with primary aim to examine two factors: the definitions and the countries. Two additional factors, as confounders, were examined, age (<45 vs. ≥45 years) and gender, since they participate in all four weighting methods. Finally, in each analysis of variance model, the data were weighted by the inverse of the standard error of the rate estimate.

Figure 1 implies the presence of interaction between the countries and the three definitions. Namely, the difference between the mean rates of unemployment according to the definitions of France on one hand and Eurostat's or Germany's on the other is much greater in Germany and Italy than in UK or France. However, the results yielded a marginal non-significant (at the 5% confidence level) difference among the definitions (p-value=0.075).



**Figure 1** Pattern of mean unemployment rate for the data.

When the analysis was repeated taking into consideration age, the main effect of country was not significant (p=0.201). On the contrary the effect of the definition and the effect of age were significant (p<0.001 and p=0.001 respectively). In addition to that, interaction of country with definition had reached significance (p=0.015), which implies that the application of the national concept of France in Germany and Italy produces highly different results to the ILO definition as compared to the cases of UK and France where the estimated rates of unemployment are close.

The resulting model, when gender was considered as well, showed that only the effects of definition and the interaction of country with definition were significant even though the p-value of the interaction was not so strong (p=0.002 and 0.032 respectively). The inspection of the parameters of the model provides some useful information for the

interpretation of Figure 1. The same considerations for interaction, as in the previous paragraph, apply here as well.

An attempt to explain these results revealed that the source of disagreement is probably located in the identification of the inactive individuals. For example, in Germany the French definition classifies 93 persons as unemployed while the Eurostat's definition considers 77 of them as inactive. This relation is not the same the other way around. In Italy the French definition identifies 102 unemployed persons but the majority of them, 97, are inactive according to the Eurostat's definition. Again the effect is not the same in the other direction. In the other countries differences in the numbers of unemployed and inactive persons are not so profound.

#### **4. Conclusions**

One of the secondary objectives of the study was to estimate unemployment rates using the quota method with a sample size of 1000 individuals (approximately). As it is discussed in [1], samples sizes of 1000 are candidates for non-probabilistic sampling as they introduce lower sampling error than probabilistic alternatives. Nevertheless, this gain is lost if important quota variables are excluded and as a result this can lead to strong biases. However, in this study we have considered major socio-economic characteristics. Finally, there seems to be no way to control for local factors except from using strict experimental design, where two or more field strategies would be compared. Overall the sources of error could be attributed to three factors: a) sample size, b) the choice of the quota variables and c) local characteristics that require differential field strategy.

Regarding the comparison of the definitions the analysis has resulted in two major findings: a) The national definition of Germany follows the trend of the Eurostat's results and is always below, though the difference is not statistically significant and b) The national definition of France is not applicable in Germany and Italy while it provides almost identical estimations with the national definition and with Eurostat's, in UK and France. Overall, future studies have to be performed in order to further explore the questions that have been raised from this project.

#### **Acknowledgements**

This project was funded by the European Union under the direction of Eurostat in SUP. COM 1998, 98/S 119-76952/EN LOT 17, which was awarded to BVA SA. BVA SA prepared the questionnaire, performed the sampling and the weighting.

#### **References**

- [1] Ardilly P. *Les Techniques de Sondage*, Editions Technip, Paris, (1994).
- [2] Cochran W. G. "Sampling Techniques", 3rd edition, John Wiley and Sons, New York, (1977).
- [3] Kleinbaum D. G., Kupper L. L., Muller K. E., Nizam A. "Applied regression analysis and other multivariate methods", 3rd edition, International Thompson Publishing, Pacific Grove, (1998).
- [4] Labour Force Survey-Results, 1998 Luxembourg: Office for Official Publications of the European Communities, 1999, Theme3: Population and social conditions.
- [5] Labour Force Survey-Methods and Definitions, 1998 Luxembourg: Office for Official Publications of the European Communities, 1999, Theme3: Population and social conditions.