

CHOOSING AMONG ALTERNATIVE  
CLASSIFICATION CRITERIA TO MEASURE  
THE LABOUR FORCE STATE

---

*Erich Battistin*  
*Enrico Rettore*  
*Ugo Trivellato*

# Choosing Among Alternative Classification Criteria to Measure the Labour Force State\*

Erich Battistin

Department of Statistics, University of Padova  
and Institute for Fiscal Studies, London

Enrico Rettore

Department of Statistics, University of Padova

Ugo Trivellato

Department of Statistics, University of Padova and CESifo, Munich

September 26, 2005

## Abstract

Current labour force counting relies on general guidelines set by the International Labour Office (ILO) to classify individuals into three labour force states: employment, unemployment and inactivity. However, the resulting statistics are known to be sensitive to slight variations of operational definitions *prima facie* consistent with the general guidelines. In this paper two alternative classification criteria are considered: a ‘strict’ criterion followed by Eurostat, which results from a stringent interpretation of the ILO guidelines, and a ‘mild’ criterion followed by the Italian Statistical Office up to 1992. We first show that the labour force statistics resulting from the two classification criteria differ considerably. We then discuss the relative merits of the two criteria by comparing those individuals whose classification depends on the criterion adopted to individuals whose classification is common across criteria. Similarities are established with respect to characteristics known to be relevant to the labour force state to assess which benchmark group individuals whose state is questionable look like the most. An application is presented to samples of married women from the Italian Labour Force Survey from five survey occasions between 1984 and 2000. Results are neatly in favour of the ‘mild’ criterion and are rather robust to changes in the business cycle, the participation rate, local labour market conditions and the questionnaire design.

*Keywords:* ILO classification, Mixture Models, Unemployment

*JEL Classification:* J64, J22

---

\*This paper, previously circulated as “Measuring participation at work in the presence of fallible indicators of the labour force state” (first version June 2000), benefited from helpful comments by two anonymous referees and the editor, from discussion with Susan Fleck and Connie Sorrentino and from comments by audiences at ESPE 2002 and SMABS 2002. Financial support from MIUR to the project “Dynamics and inertia in the Italian labour market and policies evaluation (data-bases, measurement issues, substantive analyses)” is gratefully acknowledged. Address for correspondence: Enrico Rettore, Dipartimento di Scienze Statistiche, Università di Padova, Via Cesare Battisti 241, 35121 Padova, Italy. E-mail: enrico.rettore@stat.unipd.it.

## EXECUTIVE SUMMARY

Current labour force counting relies on general guidelines set by the International Labour Office to classify individuals into three labour force states: employment, unemployment and inactivity. However, the resulting statistics are known to be sensitive to slight variations of operational definitions *prima facie* consistent with the general guidelines. It follows that the operational criterion adopted does matter, changing the pattern of unemployment and participation rates over time and attenuating or emphasizing differences across regions. For this reason, the issue of measuring unemployment has been given considerable attention in several countries. For example, the Bureau of Labor Statistics in the United States adopted since the late 70s a set of alternative unemployment indicators, known as U1-U7. The topic has been reconsidered in the mid 90s, when a new set of alternative measures has been introduced. The same problem has been carefully dealt with also in the United Kingdom, with emphasis on the production of survey-based monthly rates.

In this paper two alternative classification criteria are considered: a ‘strict’ criterion followed by Eurostat, which results from a stringent interpretation of the guidelines set by the International Labour Office, and a ‘mild’ criterion followed by the Italian Statistical Office up to 1992. We first show that the labour force statistics resulting from the two classification criteria differ considerably. We then discuss the relative merits of the two criteria by comparing those individuals whose classification depends on the criterion adopted to individuals whose classification is common across criteria. Similarities are established with respect to characteristics known to be relevant to the labour force state to assess which benchmark group individuals whose state is questionable look like the most. An application is presented to samples of married women from the Italian Labour Force Survey from five survey occasions between 1984 and 2000.

Results are neatly in favour of the ‘mild’ criterion and are rather robust to changes in the business cycle, the participation rate, local labour market conditions and the questionnaire design. While admitting that our conclusions might not be robust to additional variables omitted from the criterion to establish similarities across groups, we believe that the set of variables considered have proven important enough in the literature on labour supply to make our results a challenge for the current practice in labour force classification.

# 1 Introduction

This paper deals with the empirical problems that arise from having clear-cut *conceptual* definitions of the labour force state which do not straightforwardly map into unique *operational* criteria to infer the labour force state of individuals from information available in a typical Labour Force Survey (LFS). A general method is developed to assess the merits of alternative classification rules to discriminate between unemployment and inactivity.

National statistical agencies classify individuals of a reference population into three labour force states - employment, unemployment and inactivity - following general guidelines set by the International Labour Office (ILO). However, moving from raw information collected in a LFS there is substantial room for alternative classification rules, all broadly consistent with ILO guidelines. It follows that the operational rules derived from these guidelines are, to a certain extent, conventional (see Hussmanns *et al.*, 1990).

If the resulting labour force statistics were robust to conceivable variations of the operational rules, there would be no issue at all. In fact, well documented evidence in the literature suggests that the head-count appreciably depends upon the operational rule adopted, for countries both in and outside Europe. For example, Sorrentino (2000) points out discrepancies in the labour force statistics arising from the criteria adopted by North-American and European statistical agencies, respectively, and shows that many of these discrepancies can be explained by differences in the classification of individuals at the boundary between unemployment and inactivity. Brandolini *et al.* (2004) show how participation and unemployment rates calculated from the European Community Household Panel, whose format closely resembles that of European LFSs, are sensitive to variations in the definition of unemployed individuals.

The issue of measuring unemployment has been given considerable attention in several countries. For example, the Bureau of Labor Statistics, following a suggestion by Shiskin

(1976), adopted since the late 70s a set of alternative unemployment indicators, known as U1-U7. The topic has been reconsidered in the mid-90s, when a new set of alternative measures has been introduced (see Bregger and Haugen, 1995). The same problem has been carefully dealt with also in the United Kingdom, with emphasis on the production of survey-based monthly rates (see Working Party on the Measurement of Unemployment in the UK, 1995, Steel, 1997, and Bartholomew, 1997).

A fair conclusion that can be drawn from the studies above is that measurement issues are far from ignorable in the analysis of labour force data, and that the classification of individuals at the boundary between labour force states is somehow problematic. This is particularly important for empirical applications, as the correct classification of individuals into distinct labour market states is not just an exercise in measurement. In addition to the problem for the head-count, it is well known that classification errors might severely affect the estimation of structural and causal parameters using micro-level data (see, for example, Hausman *et al.*, 1998, and Battistin and Sianesi, 2004).<sup>1</sup>

Throughout this paper we will move from the concepts and definitions recommended by the ILO to compile labour force statistics, according to which individuals of the working-age population are classified into *three* mutually exclusive states: *employment*, *unemployment* and *inactivity*. Based on elementary information collected in a typical LFS, the classification of a large number of individuals in the population turns out to be common (and unquestionable) across statistical agencies worldwide. This is the case for individuals reporting either (i) hours of work in the reference period, or (ii) no hours of work, very recent activity for seeking work and immediate availability for work, or (iii) no hours of work and no actual interest/availability for work. According to current classification rules (see, for example, Exhibit 1 in Sorrentino, 2000), these individuals

---

<sup>1</sup>See Bound *et al.* (2001) for a review of the potential sources of errors in survey information on labour market data.

are taken as *benchmark* groups and convincingly classified as employed, unemployed and inactive, respectively.

On the other hand, there is a non-negligible number of ‘grey’ individuals at the boundary between unemployment and inactivity whose state, as it results from the information collected in a typical LFS, depends on the classification criterion adopted. These are individuals who report to be looking for a job and immediately available to work, but whose search intensity, as measured by the time elapsed since the last search action, exceeds a certain threshold (typically the month preceding the interview). While the basic difficulties have been spelled out some thirty years ago by Shiskin (1976), it is worth noting that recent trends in developed countries’ economies have expanded the spectrum of these dubious situations (see for example Malinvaud, 1986). As far as Italy is concerned, the implications of a sizeable ‘underground economy’ should also deserve careful attention (see Schneider and Ernste, 2000, and Zizza, 2002).

In this paper we propose a fairly *general approach* to shed light on the merits of alternative labour force classification criteria to discriminate between unemployment and inactivity, and we discuss whether the current operational guidelines set by the ILO are appropriate for the measurement of unemployment.

Two alternative classification criteria will be considered. The first criterion, which is currently followed by Eurostat and recommended by ILO, places the boundary between unemployment and inactivity by considering whether the last search action occurred in the *four weeks* before the interview. According to this classification, the unemployed comprise all individuals who (i) during the reference period had no hour of work (nor they had an attachment to any job they are temporarily absent from), (ii) were immediately available to work and (iii) had actively looked for a job in the four weeks preceding the interview. The second criterion, which has been followed by the Italian Statistical Office

up to July 1992 before switching to the Eurostat criterion, shares condition (i) and (ii) with the previous criterion but only requires the individual to have actively looked for a job, regardless of how far in the past. As discussed in Sorrentino (2000) and Brandolini *et al.* (2004), and as we will show in the next section, the choice between these two alternative classification criteria can appreciably affect labour force statistics.

Discrepancies between the two classification criteria will be dealt with as follows. We will compare ‘grey’ individuals who are classified as unemployed or inactive depending on the criterion adopted to the three benchmark groups unquestioned by all statistical offices, and we will establish which benchmark group they look like the most. The comparison will take place with respect to individual characteristics (such as age, education and family composition) known, both on theoretical and empirical grounds, to be strongly correlated with the labour force state. By exploiting the assumption, common across the two classification criteria considered, that the three benchmark groups comprise only employed, unemployed and inactive, respectively, we will be able to identify the boundary between unemployment and inactivity.

Our empirical analysis exploits cross-section samples of married women from the Italian LFS for five selected quarters over the period 1984 to 2000, separately for Northern, Central and Southern Italy. Such a design will allow us to study the properties of the classification rules on a sensitive sub-population of individuals (married women) over a fairly diversified range of economic circumstances, the economic context varying with respect to labour market structures (the regional breakdown), business cycle (with five survey occasions, covering years of expansion, recession and slight recovery) and the level of female participation at work (it sharply increased over the time span considered in the analysis). As pointed out by Brandolini *et al.* (2004), the Italian LFS provides a unique source of data to study the boundary between unemployment and inactivity, as

information on the *timing* of the last search action is not collected in any other European LFS, nor in Canada or in the Current Population Survey in the United States.

To preview our conclusions, given the set of individual characteristics the comparison is based on, empirical evidence is provided against Eurostat classification rules and in favour of the criterion previously adopted by the Italian Statistical Office. In particular, we find poor evidence to support the practice of classifying as inactive those individuals with no hours of work, immediately available for work and looking for a job but with no recent active steps for seeking work. Rather, according to our results, most of these ‘loose’ job seekers look quite similar to the (benchmark) unemployed. As we shall show, moving these individuals from inactivity to unemployment results in a non-negligible increase of participation and unemployment rates. Our results, though based on similarities with respect to individual characteristics which are typically observable in a LFS, are robust to variations of the business cycle, local labour market conditions, level of participation to the labour market as well as design of the survey questionnaire. For this reason, we believe that they provide evidence deserving careful attention.<sup>2</sup>

The remainder of the paper is organized as follows. Section 2 discusses to what extent alternative classification rules, all broadly consistent with ILO guidelines, may lead to significant differences in labour force statistics. An illustration is provided using data from the Italian LFS. Section 3 presents the approach that we take to discuss the classification of individuals at the boundary between labour force states. Section 4 presents the data used for our empirical application, and Section 5 describes the results.

---

<sup>2</sup>This paper focuses on the case where cross-sectional micro-data from a LFS are available to the researcher. As we will discuss further below, an extension of the analysis exploiting the LFS panel along the lines suggested by Jones and Riddell (1999) would be desirable. Indeed, the Italian Statistical Office has just released public use two-wave panel samples, obtained by exploiting the rotating sample scheme of the Italian LFS (see Brandolini *et al.*, 2004). However, it is hard to imagine that it will be ever feasible to gather longitudinal data for the last two decades of the 20th century. It follows that most of the analysis that we do in this paper to check the robustness of results over a wide time range would be lost. Arguably, also for other European countries it would be difficult to get LFS panel datasets extending backwards to the 80s.



A sensitivity analysis to violations of the assumptions our approach rests on is presented in Section 6, while Section 7 concludes.

## 2 The problem

### 2.1 Evidence from the motivating case-study

Current statistics from LFSs rely on conventional definitions to count employed and unemployed individuals. To improve international comparability of labour force indicators, ILO provides national statistical offices with recommendations on the definition and measurement of labour force participation. Over the years, these guidelines have become the standard for many countries. Consequently, operational rules adopted by LFSs are now broadly similar in outline and spirit.

According to the general ILO guidelines (see International Labor Office, 1983), a subject above a specified age (usually 14 or 15) is classified as (1) employed, if during the *reference period* s/he worked at least *a bit* (or was not at work for any reason, but is bound to get back to a job s/he has an attachment to); (2) unemployed, if (i) during the *reference period* s/he did not work at all, (ii) s/he is looking for a job and *recently* took *specific steps* for seeking work, and (iii) s/he is *immediately available* to work; (3) inactive, i.e. out of the labour force, otherwise.

Clearly, there is room for alternative operational definitions of the labour force state, depending on how the terms *reference period*, *a bit*, *recently*, *specific steps*, *immediately available* are translated into clear-cut rules for classification. All countries agree that even a single hour of work during the reference period (set to be the week prior to the interview) suffices to classify individuals as employed. Moreover, there is a general consensus that unemployed individuals should be available for work and actively seeking work.

However, the two latter conditions have been implemented differently across countries. As the boundary between unemployment and inactivity is determined with respect to the timing of the last search for work (as seen from the interview time), differences in the implementation of this concept may affect the comparability of international labour force statistics. Examples of these limitations are discussed in Sorrentino (2000), where the interpretation of the ILO guidelines across different countries in North-America and in Europe is reviewed.

It is also worth pointing out that economic theory does not help to solve the classification problem, as the economic concepts of ‘employment’ and ‘unemployment’ are typically formulated in fairly general, sometime different terms (this is especially the case for unemployment, frictional vs. keynesian, say). Crucial to our purposes is the fact that the economic definitions of the labour force state do not provide any mean to discriminate among competing operational labour force classifications (see, for example, Killingsworth and Heckman, 1986).

The empirical question then arises of whether the study of unemployment and participation rates, and more generally the analysis of labour market outcomes, is sensitive to how labour force states are conventionally defined. This paper contributes to this discussion by considering two alternative classification rules relevant to the problem. For the sake of brevity, they will be referred to as Eurostat criterion (EC) and Istat criterion (IC), respectively. These two criteria differ in the way the ILO guideline *recently took specific steps for seeking work* is made operational to discriminate between unemployment and inactivity.

The EC results from a strict interpretation of the condition of being actively seeking work and it is currently the criterion followed by Eurostat (see Eurostat, 1997). For an individual to be classified as unemployed active steps must have been taken within the

Table 1: Participation and unemployment rates for married women in Italy using Eurostat (EC) and Istat (IC) classification criteria

	Participation		Unemployment	
	IC	EC	IC	EC
<b>1984</b>				
Northern	43.92	42.29	8.30	4.75
Central	39.36	37.77	7.03	3.12
Southern	31.14	29.14	13.47	7.55
Countrywide	40.48	38.78	9.16	5.06
<b>1990</b>				
Northern	47.35	45.97	5.93	3.10
Central	48.09	45.42	10.32	5.07
Southern	34.85	31.32	21.35	12.48
Countrywide	43.28	40.89	11.20	6.01
<b>1993</b>				
Northern	51.28	49.37	6.24	2.61
Central	48.90	46.18	10.51	5.24
Southern	36.30	31.73	23.44	12.42
Countrywide	45.87	42.91	11.68	5.60
<b>1995</b>				
Northern	52.90	51.12	6.17	2.91
Central	50.83	47.24	11.86	5.17
Southern	37.55	32.12	26.69	14.30
Countrywide	47.40	44.02	12.85	6.17
<b>2000</b>				
Northern	58.06	55.77	6.68	2.85
Central	55.12	51.58	11.03	4.92
Southern	40.26	33.83	29.36	15.95
Countrywide	51.61	47.70	16.46	6.37

four weeks prior to the interview. On the contrary, the IC criterion refers to the definition that was followed by the Italian Statistical Office (Istat) up to the second quarter of 1992 (after then, it was replaced by the current EC criterion). It only requires that active steps for seeking work have been taken, regardless of how far in the past.

One might argue that for any practical purpose alternative classification rules lead to consistent results; unfortunately, this is not the case. Table 1 presents participation and unemployment rates from the Italian LFS for married women, over time (1984, 1990, 1993, 1995 and 2000 - second quarter) and by region (to control for area effects).

Weighted rates are reported using the EC and the IC. To summarise results, we regressed unemployment and participation rates on a quadratic polynomial in time, area dummies, a dummy for the classification rule and interactions of time and area dummies with the rule being used. Results are based on 30 observations, separately for unemployment and participation rates, obtained from 5 different time periods (1984, 1990, 1993, 1995 and 2000), 3 regions (Northern, Central and Southern Italy) and 2 classification methods (EC and IC).

The main findings can be summarised as follows. Both unemployment and participation rates are higher using the IC, and differences between the two classification criteria increase over time (see also Figure 1 and Figure 2 below). Regression results point to a significant effect of time, for unemployment rates, and of time and region, for participation rates. More importantly, the interaction effects between the rule being used and regional dummies are statistically significant in both regressions. It follows that the operational criterion adopted does matter, changing the pattern of unemployment and participation rates over time and attenuating or emphasizing differences across regions. Accordingly, the classification of those individuals at the boundary between unemployment and inactivity appears to be a crucial problem. Along the same lines, Jones and Riddell (1999) and Sorrentino (2000) document policy relevant differences in Canadian and American unemployment rates arising from varying the boundary between unemployment and inactivity.<sup>3</sup>

Given the results presented in Table 1, one might wonder whether such evidence

---

<sup>3</sup>Even leaving aside the potential problem for the head-count, critical problems arise for the estimation of gross flows and for the structural modelling of labour supply and unemployment. Classification errors in the observed state generally induce substantial bias in the estimation of gross flows, thus leading to erroneous conclusions about labour market dynamics (Bassi *et al.*, 2000). Besides, at the micro level classification error of a dependent variable, such as the labour force state, might severely affect the estimation of structural parameters (see for example Hausman *et al.*, 1998, and Battistin and Sianesi, 2004). Rettore and Trivellato (1993) show that the estimates of a simple model of labour supply with unemployment based on the 1984 wave of the Italian LFS are quite sensitive to the labour force state definition; the topic is further elaborated in Rettore and Trivellato (1998).

can be used to study the merits of alternative labour force counting criteria. Section 3 will address this issue by exploiting information on individuals whose classification the EC and the IC agree on, to investigate the labour force state of individuals whose classification depends on the operational rule followed. The remainder of this section presents a general formulation of the problem and introduces the notation that we will use throughout the analysis.

## 2.2 General set-up

Let  $T$  be the *actual* labour force state and let  $R$  be a categorical index summarizing the basic information on the labour force condition available in a LFS. Categories of  $R$  typically summarize individuals' activity in the reference week (work/no work) and their attachment to the labour market as it results from self-reported information on the timing of their job-search and on their availability to work.

Throughout the analysis it will be assumed that three mutually exclusive labour force states exist: *employment* ( $E$ ), *unemployment* ( $U$ ) and *inactivity* ( $OLF$ ). Any classification rule defines a correspondence between categories of  $R$  and labour force states. Individuals are classified into one of the three labour force states by grouping the categories of  $R$  according to that rule.

The top panel of Table 2 presents the categories of  $R$  available from the Italian LFS questionnaire relevant to EC and IC. Definitions of the mutually exclusive categories of  $R$  are given by column. Category  $W$  identifies those individuals who either report *at least one* hour of work in the reference week or have a formal attachment to a job from which they are temporary absent for any reason (temporary lay-off included). Categories  $S1$  through  $S4$  refer to individuals not at work and *actively seeking work*, and differ according to the timing of the last specific step to seek work. As pointed out earlier in the paper, the collection of information on the timing of search is a unique characteristics of the

Table 2: Definition of  $R$  and sample size

	Definitions	Working			Actively searching			Not searching		
		$W$	$S1$	$S2$	$S3$	$S4$	$NS1$	$NS2$		
	At least one hour of work in the reference week	X								
	No hours of work in the reference week		X	X	X	X	X	X		
	Looking for a job and immediately available for work		X	X	X	X				
	Last search undertaken during the last four weeks		X							
	Last search undertaken from one to six months ago			X						
	Last search undertaken more than six months ago				X					
	No search step undertaken yet (before 1992)					X				
	Search activity in the future (after 1992)						X			
	Not looking for a job because discouraged								X	
	Not looking for a job because not willing to work									X
<b>1984</b>	North	17,249	853	466	55	188	-	365	22,028	41,204
	Center	4,346	161	110	28	49	-	171	6,443	11,308
	South	3,904	326	207	37	45	-	332	8,987	13,838
<b>1990</b>	North	25,499	1,340	783	120	282	-	868	37,458	66,350
	Center	18,192	586	373	62	141	-	305	19,684	39,343
	South	5,481	301	204	35	72	-	128	5,764	11,985
<b>1993</b>	North	5,961	948	695	63	111	-	421	13,757	21,956
	Center	29,634	1,835	1,272	160	324	-	854	39,205	73,284
	South	7,314	198	185	62	-	27	70	7,640	15,496
<b>1995</b>	North	3,845	216	171	76	-	8	59	4,301	8,676
	Center	3,598	512	415	177	-	24	144	7,621	12,491
	South	14,757	926	771	315	-	59	273	19,592	36,663
<b>2000</b>	North	7,541	235	167	68	-	36	90	7,368	15,505
	Center	3,927	243	242	104	-	7	106	4,098	8,727
	South	3,520	578	518	168	-	22	232	7,170	12,208
		14,988	1,056	927	340	-	65	428	18,636	36,440
	North	7,813	232	168	95	-	45	85	6,304	14,742
	Center	3,950	221	191	96	-	5	95	3,541	8,099
	South	3,450	677	540	223	-	15	336	6,583	11,824
		15,213	1,130	899	414	-	65	516	16,428	34,665

Italian LFS, which allows us to distinguish different groups of individuals amongst those actively searching for an occupation. The *NS1* category refers to (or at least includes) the so-called ‘discouraged’ workers, that is those individuals not at work and *not looking* for a job because either (i) they have been unsuccessfully searching in the past or (ii) they believe not to be skilled enough or (iii) they believe employers consider them too young or too old (see OECD, 1987, for a discussion of the criteria to identify discouraged workers). Finally, the *NS2* category consists of the *unattached*, that is individuals not at work, *not looking* for a job and definitely not willing to work.<sup>4</sup>

With this notation, the classification rules implied by the EC and the IC can be straightforwardly summarised as follows. According to the EC, sample information on *R* is used to identify the actual state by means of the following rule:

$$\begin{aligned} T = E & \iff R = W, \\ T = U & \iff R = S1, \\ T = OLF & \text{ otherwise.} \end{aligned} \tag{1}$$

Thus, individuals out of the labour force are characterised for not seeking work or for having conducted their last search more than one month before the interview.

On the other hand, the IC can be formulated as follows:

$$\begin{aligned} T = E & \iff R = W, \\ T = U & \iff R = S1, S2, S3, S4 \\ T = OLF & \text{ otherwise,} \end{aligned} \tag{2}$$

so that whether or not an individual reports any search for work (regardless of its timing) determines her inclusion among the unemployed or the inactive, respectively.

Clearly, the EC and the IC agree on individuals reporting *W*, *S1*, and *NS2*, which are classified as *E*, *U* and *OLF*, respectively. The two criteria also agree on the classification of the *NS1* category into *OLF*. What is questioned instead, and where the two criteria

---

<sup>4</sup>Since the questionnaire of the Italian LFS closely follows the standards set by Eurostat, it is worth noting that the distinction between individuals working (*W*), actively seeking an occupation with recent (*S1*) or less recent (the group resulting from the union of our *S2*, *S3* and *S4*) steps taken, discouraged (*NS1*) and unattached (*NS2*) could be derived for any other European and North-American LFS. For example, Brandolini *et al.* (2004) use the European Community Household Panel, which is a longitudinal survey coordinated by Eurostat, to look at these groups of individuals.

differ, is how individuals presenting any of the remaining categories of  $R$  ( $S2$ ,  $S3$  or  $S4$ ) are classified, that is where the boundary is set between unemployment and inactivity. Both classification criteria require availability and active seek for work to be classified as unemployed, but the requirement of active job-search is interpreted in different ways.

In what follows, information for the three benchmark groups  $W$ ,  $S1$  and  $NS2$  will be exploited to shed light on the relative merits of the classification rules implied by the EC and the IC for the remaining individuals in the population. The logic of what we do closely resembles previous work by Flinn and Heckman (1983), Jones and Ridell (1999) and Brandolini *et al.* (2004). As the approach taken in this paper, the above mentioned research looks for behavioural similarities between some benchmark groups whose labour market state is known on the one hand, and groups whose labour market state is unclear on the other. However, while in these papers similarities are established with respect to transition rates towards the benchmark states  $W$  and  $NS2$  (i.e. towards employment and inactivity), we instead look for similarities with respect to a set of characteristics relevant to the labour force state. Advantages and disadvantages of the two approaches to discriminate among alternative classification rules will be discussed in Section 6.2 below.

### **3 A model with fallible indicators of the labour force state**

#### **3.1 Model specification**

Our analysis develops by relaxing the deterministic relationship between the index  $R$  and the labour force state  $T$  postulated by both the EC and the IC. Information will be exploited on individual characteristics collected in a typical LFS, known to matter for labour force state membership both on theoretical and empirical grounds. We will use such characteristics to assess how the categories of  $R$  relate to the labour force state  $T$



and, eventually, to choose between the classification criteria (1) and (2).

Let  $x$  be a set of individual characteristics relevant to the probability of membership in each labour force state, and let  $f(x)$  be their distribution. If  $x$  affects the probability of membership in each labour force state, then it must be that  $f(x|A) = \frac{Pr(A|x)f(x)}{Pr(A)}$ ,  $A = W, S1, NS2$ , varies with  $A$ , since subjects presenting  $W$ ,  $S1$  and  $NS2$  are taken as out of question employed, unemployed and inactive, respectively. These distributions will play as our benchmark in the analysis. The validity of alternative classification criteria can be assessed by looking at the distribution of  $x$  for those individuals in the remaining categories of  $R$ , to check which benchmark distribution they look like the most.

Formally, if the EC were right the following equalities would hold at least approximately:

$$f(x|R) = f(x|NS2), \quad R = S2, S3, S4, NS1.$$

In other words, if individuals who report  $S2$ ,  $S3$ ,  $S4$  and  $NS1$  were truly inactive, they should look like benchmark *OLF* individuals with respect to  $x$ . Analogously, if the IC were right, the following equalities:

$$\begin{aligned} f(x|R) &= f(x|S1), & R = S2, S3, S4 \\ f(x|NS1) &= f(x|NS2), \end{aligned}$$

would approximately be verified. It is worth noting that the  $NS1$  group (which, loosely speaking, consists of ‘discouraged workers’) can also be investigated. According to common practice, individuals presenting this category are classified as inactive since they miss the ‘actively seeking work’ condition.

There is a third alternative, however, which is somewhere in between the IC and the EC. Categories  $S2$ ,  $S3$ ,  $S4$  and  $NS1$  might be a *mixture* of unemployed and inactive individuals. If this were the case, the conditional distributions  $f(x|R)$ ,  $R = S2, S3, S4, NS1$ ,

would approximately equal a *weighted sum* of the three benchmark distributions by a proper choice of the weights.

In what follows we will test for this possibility by seeking for a weighted mean of the three benchmark distributions able to provide a reasonable approximation to  $f(x|R)$ ,  $R = S2, S3, S4, NS1$ . Formally, the problem may be formulated in mixture terms by writing:

$$f(x|R) = \sum_{T \in \{E, U, OLF\}} p(T|R) f(x|T), \quad R = S2, \dots, NS1, \quad (3)$$

where the components of the mixture are known due to the maintained assumptions:

$$\begin{aligned} f(x|E) &= f(x|W), \\ f(x|U) &= f(x|S1), \\ f(x|OLF) &= f(x|NS2). \end{aligned} \quad (4)$$

The next section discusses the estimation strategy to identify the mixture weights  $p(T|R)$  in (3).

### 3.2 Estimation issues

The crucial restriction equation (3) rests on is that the observed responses of the manifest variables  $x$  are independent of  $R$  once the labour force state is accounted for. More formally, (3) follows once the restriction

$$f(x|T, R) = f(x|T)$$

is imposed. According to this formulation, in our problem the association between  $x$  and  $R$  arises because of their joint dependence on the labour force state  $T$ , with  $T = E, U, OLF$ . The distributions  $f(x|T)$  in (3) are assumed known *a priori*, since they correspond to the distributions of  $x$  for the categories  $R = W, S1, NS2$ , respectively. As

a consequence model (3) can be interpreted as a mixture model with known components where the weights  $p(T|R)$  are unknown.<sup>5</sup>

Let:

$$\mathfrak{S}_A = \left\{ f(x|A), A \in \tilde{A} \right\}$$

denote the family of conditional distribution functions of the variable  $x$  indexed by a point  $A$  in a discrete set  $\tilde{A}$ . The relationship in (3) states that each member of the family  $\mathfrak{S}_R$  belongs to the three dimensional convex hull generated by the family  $\mathfrak{S}_T$ . The relationship between the distributions in  $\mathfrak{S}_R$  and the distributions in  $\mathfrak{S}_T$  is established once the mixing weights  $p(T|R)$ , i.e. the probability of being in state  $T$  conditional on reporting  $R$ , are specified. Such weights summarise the properties of the measurement instrument.

Two logically different types of restrictions can be imposed on the mixture weights. The first set of restrictions follows directly from the classification rules on which EC and IC *agree*. The identification of the mixture components in (3) is driven by these restrictions, that can be summarized as follows:

$$\begin{aligned} R = W &\implies T = E, \\ R = S1 &\implies T = U, \\ R = NS2 &\implies T = OLF. \end{aligned} \tag{5}$$

The foregoing relationships impose restrictions on the mixture weights, since they imply that:

$$p(E|W) = p(U|S1) = p(OLF|NS2) = 1.$$

---

<sup>5</sup>The formulation of the problem features some loose similarities with a Latent Class Analysis (LCA; see Goodman, 1974, and Hagenaars, 1990). LCA assumes that a set of latent (unobservable) classes exists such that, conditional on latent class membership, the manifest variables are mutually independent of each other. However, our problem departs from the traditional LCA set-up since the latent distributions  $f(x|T)$  are assumed known *a priori*.

As a result, the mixture components turn out to be identified as distinct members of the family  $\mathfrak{S}_R$ , so that  $\mathfrak{S}_T \subset \mathfrak{S}_R$  and (4) is also satisfied.

The second set of restrictions on the mixture weights refers to those categories of  $R$  on which EC and IC *disagree*. For example, by stating that individuals presenting  $R = S2, S3, S4$  belong to *OLF*, EC imposes the following additional restrictions on the mixture weights:

$$p(U|S2) = p(U|S3) = p(U|S4) = 0,$$

$$p(E|S2) = p(E|S3) = p(E|S4) = 0.$$

A similar set of restrictions results by applying IC.

It is worth noting that, given the first set of restrictions, the second set consists of *over-identifying* restrictions, namely restrictions that can be tested. Otherwise stated, by relying upon (5), the weights associated to  $R = S2, \dots, NS1$  can be estimated without additional restrictions. The major implication is that any *a priori* restriction on individuals reporting  $R = S2, \dots, NS1$  can be tested against the data.

A sufficient condition for the identifiability of the weights in (3) is that the set of mixture components:

$$\mathfrak{S}_T = \{f(x|W), f(x|S1), f(x|NS2)\}$$

is linearly independent, i.e. that none of them can be written as a linear combination of the remaining ones. The likelihood equations for  $p(T|R)$ ,  $R = S2, \dots, NS1$ , can be straightforwardly derived from the relationship in (3). The EM algorithm is particularly useful to obtain the maximum likelihood estimates of the mixing weights in this case (Everitt and Hand, 1981, and Maritz and Lwin, 1989). Starting with initial values  $p(T|R)^{(0)}$ , new values  $p(T|R)^{(1)}$  are obtained by iteration according to the following

equations:<sup>6</sup>

$$\begin{aligned}
 p(T|R, x)^{(1)} &= \frac{f(x|T)p(T|R)^{(0)}}{\sum_s f(x|s)p(s|R)^{(0)}}, \\
 p(T|R)^{(1)} &= \sum_x f(x|R)p(T|R, x)^{(1)}.
 \end{aligned}$$

Once the mixture weights  $p(T|R)$  have been estimated, the probability of membership in the three labour force states implied by the model can be expressed as:

$$p(T) = \sum_R p(T|R)p(R), \quad T = E, U, OLF. \quad (6)$$

### 3.3 Specification testing

In this section we discuss how we test for the correct specification of the mixture model (3)-(4). Two alternative indicators of the goodness of fit will be considered: the log likelihood test and the Schwarz (1978) statistic.

The problem we deal with is whether, by properly weighting the benchmark distributions  $f(x|W)$ ,  $f(x|S1)$  and  $f(x|NS2)$ , we succeed in approximating the four distributions  $f(x|R)$ ,  $R = S2, S3, S4, NS1$ . A rejection of the model should be taken as evidence that the three states, as defined by the maintained operational criteria (5), are not enough to fully account for what happens in the labour market. In turn, this could either imply that (i) the three benchmark distributions resulting from the maintained operational criteria (i.e. those on which the EC and the IC agree) do not correspond to  $f(x|E)$ ,  $f(x|U)$  and  $f(x|OLF)$  or that (ii) the number of labour force states is larger than three.

A test for the restrictions imposed by (3)-(4) can be derived by comparing the estimate of  $f(x|R)$  obtained under the mixture model to the estimate obtained by taking its empirical counterpart, thus assessing the model fit by means of a *likelihood ratio test*. To fix ideas, let  $x$  be a  $k$ -dimensional multinomial distribution resulting from discretizing

---

<sup>6</sup>In our analysis, convergence was achieved after few iterations and results appeared to be robust with respect to the choice of initial values.

the variables relevant to the labor force state (the set of variables used in the analysis will be described in Section 4). Under the null hypothesis of correct specification, the probabilities  $f(x|R)$ ,  $R = S2, \dots, NS1$ , are equal to a weighted mean of  $f(x|W)$ ,  $f(x|S1)$  and  $f(x|NS2)$ , with two weights to be estimated (since there is an obvious adding-to-one restriction). Under the alternative hypothesis there are  $k - 1$  parameters to be estimated, leading to  $k - 3$  degrees of freedom for the likelihood ratio test.

To sensibly compare the goodness of fit at different levels of model parsimony, a *penalized version of the log likelihood* is also considered, the penalty term depending on both the dimension of the parameter and the sample size (see Schwarz, 1978, and Kass and Raftery, 1995). The criterion suggested is derived as the large-sample limit of a bayesian procedure under a special but fairly general class of priors (see Schwarz, 1978). The rule derived within such a framework consists of choosing the model so that

$$\ell_\tau - 0.5\tau \log n$$

is maximized, where  $\ell_\tau$  is the log-likelihood for the model whose dimension is  $\tau$  and  $n$  is the sample size. Under the null hypothesis that model (3)-(4) is correctly specified the dimension is  $\tau = 2$  (i.e. the number of weights to be estimated), while under the alternative hypothesis  $\tau = k - 1$ . It follows that the mixture model is not rejected if and only if the inequality

$$\ell_2 - \ell_{k-1} + 0.5(k - 3) \log n > 0$$

holds. Accordingly, the usual likelihood ratio criterion is corrected by a term reflecting the different degree of parsimony of the two competing models and the penalty increases with the sample size.

## 4 The data

Our plan for the empirical analysis extends to micro-data from the Italian LFS on a sample of married women aged no more than 60 whose husband is no more than 65 years old, on five survey occasions - 1984, 1990, 1993, 1995 and 2000, always second quarter - and separately for Northern, Central and Southern Italy. For the period covered by our empirical analysis the lower age limit to enter the labour force in Italy was set at 14 up to 1992 and at 15 since then.<sup>7</sup>

The five sample years have been selected to reflect the variability in the business cycle, with 1984 and 1990 years of expansion, 1993 a year of recession, 1995 a year of slight recovery from the recession and 2000 a year of moderate economic growth and sharp employment growth (see Altissimo *et al.*, 2000). The regional breakdown is intended to capture structural differences in the Italian labour market and in the overall economy, with the relatively well-developed Northern area contrasted with the much less developed South (see Table 1). In addition, the countrywide EC participation rate of married women grew from 38.78 in 1984 to 47.70 in 2000, allowing to check whether changes in the composition of the pool of participants affect the results.<sup>8</sup>

We look at married women because they represent a sub-set of the population whose labour supply is particularly sensitive to individual characteristics as well as to labour demand conditions (see Killingsworth and Heckman, 1986). The individual characteristics available from the Italian LFS that we consider are the following: woman's age and education, husband's age and education, number of children and their age. The  $f(x|R)$ 's

---

<sup>7</sup>It is worth pointing out that by focusing on married women we consider a large fraction of the entire female population aged below 60. This fraction ranges between 80 and 85 percent on average for the *W*, *NS1* and *NS2* groups, and it is well above 90 percent for the remaining categories (almost 100 percent for Central and Southern Italy).

<sup>8</sup>One reason to check the robustness of our results with respect to the business cycle and the regional labour market conditions is that the four-week requirement may be excessively rigid for discriminating between individuals searching for an occupation and individuals out of the labour force, as the timing of search can be endogenously determined by the conditions of the overall economy.

coincide with the multinomial distributions obtained by discretizing such variables into  $k = 41$  cells of reasonable sample size (see Table A.4).

Tables A.1, A.2 and A.3 present descriptive statistics for the variables above. Results from a multinomial regression of the categorical variable taking values  $W$ ,  $S1$  and  $NS2$  on the individual characteristics  $x$  as well as on year and regional dummies point to a strong relevance of the explanatory variables for the labour force state. Further empirical evidence on the relevance of these characteristics to the female labour force state is discussed in Rettore and Trivellato (1993). As the distribution of the variables  $x$  varies across the three benchmark groups, it follows that the condition required for the identification of weights in (3) is met.

The bottom panel of Table 2 presents sample size by year, region and categories of  $R$ . The main group of non-working individuals is  $NS2$  (around 89 percent on average), followed by  $S1$  (4 percent on average). The remaining groups account for a much smaller proportion of individuals, and their size shrinks considerably as the number of months since the last search increases.<sup>9</sup>

## 5 Results

### 5.1 Goodness-of-fit

The empirical analysis is carried out separately by year and geographic area (North, Center, South), allowing each distribution  $f(x|R)$ ,  $R = S2, S3, S4, NS1$ , to be a weighted mean of the three benchmark distributions. In particular, we allow for  $f(x|W)$  to enter this weighted mean: otherwise stated, we allow for similarities with respect to the

---

<sup>9</sup>Changes in the questionnaire and survey operations took place over the period covered by this analysis, due to an overall revision of the Italian LFS which took place when Istat moved to the EC criterion (see Casavola and Sestito, 1994, and Trivellato, 1997). Because of these changes, the definition of the residual category  $S4$  amongst actively searching individuals slightly changes after October 1992. While before October 1992 subjects presenting  $S4$  are those who report that ‘no search step has been undertaken at the moment of the interview’, after then they are those reporting they plan to search in the future among those not seeking work at the interview time. Also, it can be noticed the sharp drop in total sample size after 1992.



Table 3: Estimation results

	North				Center				South			
	S2	S3	S4	NS1	S2	S3	S4	NS1	S2	S3	S4	NS1
<b>1984</b>												
Employment	0.1487	0.1647	0.1394	0.0277	0.1460	-	0.2686	0.0000	0.0763	0.2425	0.2375	0.0001
Unemployment	0.8513	0.7362	0.8606	0.0000	0.8540	-	0.6370	0.0000	0.9237	0.7575	0.7134	0.0000
Out of the labour force	0.0000	0.0992	0.0000	0.9723	0.0000	-	0.0944	1.0000	0.0000	0.0000	0.0491	0.9999
sample size	466	55	188	365	110	28	49	171	207	37	45	332
<b>1990</b>												
Employment	0.1577	0.0060	0.1681	0.0000	0.1718	0.4810	0.6466	0.1055	0.0000	0.0283	0.1393	0.0000
Unemployment	0.7861	0.8590	0.8286	0.0000	0.8282	0.5190	0.2464	0.0160	0.9999	0.9717	0.7384	0.0368
Out of the labour force	0.0562	0.1350	0.0033	1.0000	0.0000	0.0000	0.1070	0.8786	0.0001	0.0000	0.1223	0.9632
sample size	373	62	141	305	204	35	72	128	695	63	111	421
<b>1993</b>												
Employment	0.4890	0.4527	-	0.0039	0.2016	0.1130	-	0.0000	0.1188	0.0788	-	0.0862
Unemployment	0.5110	0.5473	-	0.1063	0.7323	0.8126	-	0.1357	0.7694	0.9212	-	0.0299
Out of the labour force	0.0000	0.0000	-	0.8898	0.0662	0.0745	-	0.8643	0.1117	0.0000	-	0.8838
sample size	185	62	27	70	171	76	8	59	415	177	24	144
<b>1995</b>												
Employment	0.5621	0.1253	0.7892	0.0488	0.3832	0.1651	-	0.0000	0.0904	0.1706	-	0.0000
Unemployment	0.4379	0.5488	0.2108	0.0024	0.6167	0.8087	-	0.0000	0.9095	0.8294	-	0.0000
Out of the labour force	0.0000	0.3258	0.0000	0.9489	0.0001	0.0261	-	1.0000	0.0001	0.0000	-	1.0000
sample size	167	68	36	90	242	104	7	106	518	168	22	232
<b>2000</b>												
Employment	0.3738	0.3210	0.1293	0.0748	0.1753	0.1760	-	0.0000	0.0503	0.0712	-	0.0000
Unemployment	0.6261	0.4187	0.2474	0.0000	0.6197	0.6747	-	0.1490	0.8901	0.8513	-	0.1427
Out of the labour force	0.0000	0.2603	0.6233	0.9252	0.2050	0.1493	-	0.8510	0.0596	0.0775	-	0.8573
sample size	181	96	47	87	191	96	5	95	540	223	15	336

Table 4: Goodness of fit

	<i>North</i>				<i>Center</i>				<i>South</i>			
	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>
<b>1984</b>												
p-value $\chi^2$	0.0450	0.1010	0.0120	0.0020	0.0210	-	0.0720	0.0040	0.0200	0.0010	0.7420	0.0000
Schwarz	88.9930	51.3645	66.7967	21.3632	65.7523	-	51.0711	48.1997	67.9159	40.1597	56.4136	62.9514
<b>1990</b>												
p-value $\chi^2$	0.2170	0.0510	0.0280	0.0040	0.0010	0.0090	0.0000	0.0050	0.0010	0.0780	0.4490	0.0010
Schwarz	87.9787	53.4391	65.5384	52.4299	64.0921	43.1144	42.8575	56.5600	90.7684	50.9024	67.3242	72.9154
<b>1993</b>												
p-value $\chi^2$	0.0280	0.0180	-	0.1920	0.0020	0.0010	-	0.5610	0.0000	0.0840	-	0.0380
Schwarz	69.0126	52.5186	-	62.0177	60.0354	49.8326	-	57.8071	85.8691	70.6865	-	67.2565
<b>1995</b>												
p-value $\chi^2$	0.1400	0.1310	0.0530	0.0720	0.0100	0.3200	-	0.3610	0.0030	0.0240	-	0.0090
Schwarz	74.4889	56.3545	44.9364	52.9531	70.7131	64.7685	-	64.4423	86.9259	68.5848	-	73.4518
<b>2000</b>												
p-value $\chi^2$	0.0140	0.0120	0.4910	0.2770	0.0050	0.0050	-	0.1240	0.0020	0.0230	-	0.0000
Schwarz	63.5744	50.1314	55.5625	58.7673	59.4030	47.9356	-	61.0158	75.0420	65.7738	-	65.4566

observable characteristics  $x$  between individuals presenting  $R = S2, S3, S4, NS1$  and individuals presenting  $R = W$ .

Table 3 presents estimated weights from model (3) separately by year and region. Results are not reported when the sample size is smaller than 30, thus excluding 8 out of 60 cells defined by categories  $R$ , year and region. Table 4 reports the  $p$ -value associated to the likelihood ratio test of the constrained against the unconstrained model (that is, the  $p$ -value of the  $\chi^2$  statistic) and the Schwarz statistic (positive values are not against the constrained model). Bootstrap  $p$ -values are reported as the result of 1,000 simulations under the null hypothesis - namely, by assuming that model (3) is correctly specified - and using the mixture weights estimated from the actual sample. The likelihood-ratio statistic is evaluated on each pseudo-sample and its distribution under the null hypothesis is calculated.

Although the overall picture suggests a fairly good fit, results vary appreciably depending on the categories of  $R$ , on the time period and on the criterion used.<sup>10</sup> According to the likelihood ratio test, the model is often not rejected for  $f(x|S3)$  both before and after 1992, and it also provides a fairly good picture for  $f(x|S4)$ . Instead, the model is most times rejected for  $f(x|S2)$  (with the exception of Northern Italy) and  $f(x|NS1)$  (though the overall picture looks better after 1992). Overall, the model fits the data better after 1992. Results for the Schwarz statistic are much more encouraging, as they suggest that, once the parsimony of the competing specifications is accounted for, the mixture model is *never* rejected. Overall, results in this section, consistently with the conventional wisdom, point to the existence of three labour force states.

---

<sup>10</sup>The same model estimated separately for self- and proxy-respondents (see Blair *et al.*, 1991) leads to similar conclusions. For this reason, results are presented for the overall sample. It is worth stressing again that, because of the change in the questionnaire documented in Section 4, the definitions of category  $S4$  before and after 1992 are not fully comparable.

## 5.2 Mixture weights

We now turn to the main aim of our exercise: which benchmark distribution do the  $f(x|R)$ 's,  $R = S2, S3, S4, NS1$ , look like the most? Though the mixture model is found not fully satisfactory for all categories of  $R$  and all combinations of time and geographic areas, it is nonetheless an interesting exercise to use mixture weights as a tool to classify into the usual labour force states individuals in the uncertain categories of  $R$ .

The main results about the merits of the two classification criteria considered in this paper can be summarized as follows. For individuals presenting  $R = NS1$ , the current practice of classifying them as inactive is *not* called into question by our test. In fact, this group appears to comprise only (or, with few exceptions, mainly) inactive individuals. As for the groups  $R = S2, S3, S4$ , our results suggest that the practise followed by the IC seems more appropriate: with very few exceptions, the bulk of these groups consists of unemployed. Finally, it is also worth noting that a non-negligible fraction of individuals presenting  $R = S2, S3, S4$  turns out to be similar to the employed, particularly in Northern and Central Italy.

More specifically, conditional on (i) not being at work, (ii) looking for a job and (iii) being immediately available for work:

- $R = S2$  group: there is *no* evidence in our data supporting the Eurostat practice of classifying these individuals as inactive (the only relevant exceptions are South 1993 and Center 2000). These individuals are mostly identical to individuals in  $U$ , and their estimated probability of being in unemployment ranges from 70 percent to 100 percent (with few minor exceptions featuring slightly smaller estimates).
- $R = S3$  group: the evidence for this group is less clear-cut, although the estimated proportion of inactive individuals is well below 10 percent on average (notable

exceptions are North 1995 and North 2000).

- $R = S4$  group: results for this group vary over time and across areas. For Northern Italy, the probability of being inactive is different from zero only in 2000. For Central and Southern Italy, the same probability is around 10 percent on average in 1984 and 1990 (results for 1993 and 1995 are not reported due to the very small sample size)

Finally, we also have that:

- $R = NS1$  group: Individuals in  $NS1$  are definitely close to  $OLF$ . The current practice, common to both classification criteria, is not rejected by our results. With minor exceptions, the so-called ‘discouraged workers’ look inactive.

These results closely resemble those obtained from previous research in the literature, in which the validity of alternative labour force classifications for individuals at the boundary between unemployment and inactivity was established by looking at their transition rates towards the benchmark states. The intuition behind this approach to the problem, pioneered by Flinn and Heckman (1983), is that if transition probabilities from two or more states towards  $W$ ,  $S1$  and  $NS2$  are statistically equivalent, those states cannot be regarded as behaviourally different. Resting on this intuition, Jones and Riddell (1999) and Brandolini *et al.* (2004) provide an empirical assessment of the appropriate definitions of unemployment and inactivity for the dubious categories discussed in this paper (the only exception being the  $S4$  category, which is not considered by either of these studies).

Using longitudinal data from the 2000 Italian LFS, Brandolini *et al.* (2004; see Table 6) find that the groups  $S2$  and  $S3$  are always behaviourally different from non-searchers, namely from  $NS1$  and  $NS2$ . Their sample is selected without controlling for

the marital status and including women aged below 65, and pooled results are presented for Northern and Central Italy. Bearing this in mind, Brandolini *et al.* (2004; see Table 7) also find that, with the exception of women aged 15-34 and living in North-Central Italy, the category *S2* behaves like our benchmark unemployed. A similar result applies to *S3*, but only for women aged 35-64 in Southern Italy, while the similarity of *S3* with unemployment is rejected for younger women in the same region and for all women in the rest of Italy. Based on these results, Brandolini *et al.* (2004) conclude that there might be a fourth labour market state in between unemployment and inactivity.

As for discouraged workers, Jones and Riddell's (1999, see Table 1) group labeled 'M(D)', whose definition corresponds to *NS1* in this paper, displays a behaviour closer to the benchmark inactive than to the benchmark unemployed. However, although their analysis is not directly conducted on the 'M(D)' group and does not control for gender, the equivalence of discouraged workers and inactive is apparently rejected in their data. They again conclude for the existence of four distinct labour force states, as in Brandolini *et al.* (2004).

It is worth noting, however, that the approach taken in this paper provides a simple way to reconcile the contradictory evidence on the number of states coming from our analysis and the analysis done by previous studies. It could indeed be the case that the fourth group of individuals found by Jones and Riddell (1999) and Brandolini *et al.* (2004) displays a pattern of transition rates different from those of the three benchmark groups because it comprises three distinct unobservable sub-groups of individuals from the usual three states. Note that, if this were the case, the fourth group would not represent a real state, but just a mixture of the usual three states.

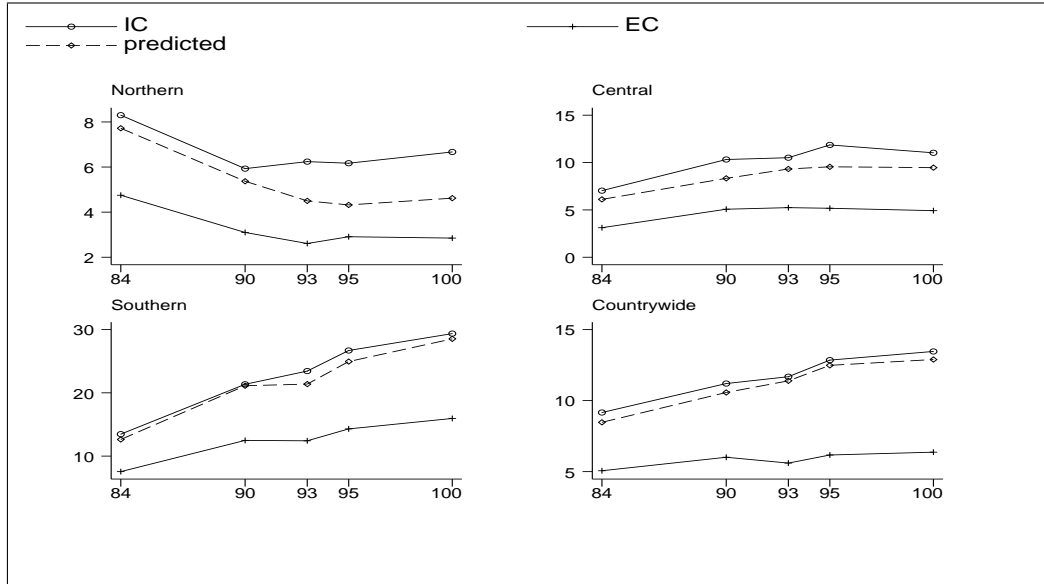


Figure 1: Married women unemployment rates: Eurostat criterion (EC), Istat criterion (IC) and implied by the estimated model

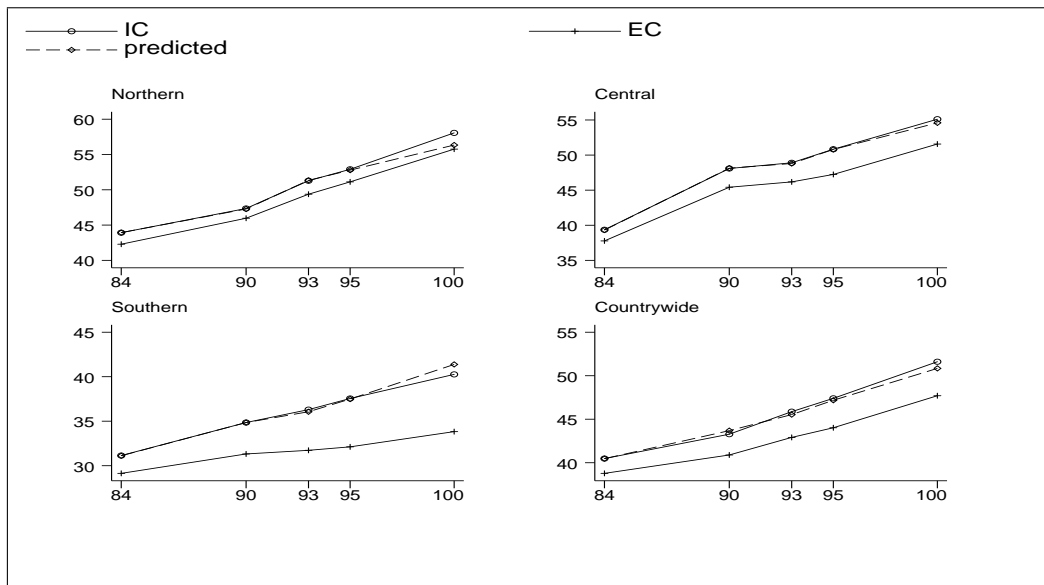


Figure 2: Married women participation rates: Eurostat criterion (EC), Istat criterion (IC) and implied by the estimated model

### 5.3 Unemployment and participation rates

Since the unemployment rate is defined as the proportion of unemployed individuals out of the total labour force:

$$\frac{Pr(U)}{Pr(U) + Pr(E)},$$

the estimated value of this indicator as implied by the model can be derived using the relationship in (6). Figure 1 presents the unemployment rate for 1984, 1990, 1993, 1995 and 2000 by region as it results from EC and IC calculations (that is, the numbers reported in Table 1) and from model (3). Participation rates are presented in Figure 2. As for the unemployment rates, those implied by the model are very close to those derived according to the IC, uniformly over time and across areas (the only exception is North after the 1992 change of the survey questionnaire, in which case the model based rates are approximately in between the EC and IC ones). As for the model-based participation rates, they are always nearly equal to those derived from the IC (the only exception is North 2000).

## 6 Assessing the validity of the model

### 6.1 Robustness to violations of the identifying restrictions

Estimating the proportion of employed, unemployed and inactive individuals among those presenting  $R = S2, S3, S4, NS1$  crucially relies on assumption (4), which yields identification of the distributions of  $x$  conditional on employment, unemployment and inactivity, respectively. In this section we study the robustness of our findings to violations of this assumption by considering *two* types of sensitivity analysis.

First, we study the effects of contaminating the three benchmark groups  $W, S1, NS2$  with non-employed, non-unemployed and active individuals, respectively. In other words, we relax the assumption that the three benchmark distributions comprise only



Table 5: Estimates of the mixture weights by pooling samples over the years

	<i>North</i>				<i>Center</i>				<i>South</i>			
	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>
Employment	0.0886	0.1862	0.0608	0.0000	0.0403	0.1004	0.1962	0.0000	0.0242	0.1196	0.0507	0.0000
Unemployment	0.9114	0.6727	0.9392	0.0000	0.9597	0.8989	0.7471	0.0000	0.9758	0.8803	0.8352	0.0268
Out of the labour force	0.0000	0.1410	0.0000	1.0000	0.0000	0.0007	0.0567	1.0000	0.0000	0.0000	0.1141	0.9732

Table 6: Estimation results for individuals working less than 20 hours per week

<b>1984</b>	<i>North</i>	<i>Center</i>	<i>South</i>
Employment	0.9048	1.0000	1.0000
Unemployment	0.0000	0.0000	0.0000
Out of the labour force	0.0952	0.0000	0.0000
sample size	1,781	412	486
<b>1990</b>	<i>North</i>	<i>Center</i>	<i>South</i>
Employment	0.9999	1.0000	0.9915
Unemployment	0.0000	0.0000	0.0085
Out of the labour force	0.0001	0.0000	0.0000
sample size	2,324	649	943
<b>1993</b>	<i>North</i>	<i>Center</i>	<i>South</i>
Employment	0.8862	1.0000	1.0000
Unemployment	0.1137	0.0000	0.0000
Out of the labour force	0.0002	0.0000	0.0000
sample size	1,123	537	520
<b>1995</b>	<i>North</i>	<i>Center</i>	<i>South</i>
Employment	0.8950	1.0000	1.0000
Unemployment	0.0443	0.0000	0.0000
Out of the labour force	0.0607	0.0000	0.0000
sample size	1,165	544	527
<b>2000</b>	<i>North</i>	<i>Center</i>	<i>South</i>
Employment	0.7474	0.9659	0.9870
Unemployment	0.2049	0.0341	0.0000
Out of the labour force	0.0476	0.0000	0.0130
sample size	1,396	611	620

employed, unemployed and inactive individuals, and we check the implications for our analysis. Second, we allow the hard core of the category  $W$  to be less homogeneous than we have maintained so far by splitting working individuals into two groups depending on the number of hours worked (below and above 20 per week). Individuals working ‘part-time’ are treated as a dubious category, and we apply the same procedure described in Section 3.1 after considering one additional group. Our results survive these checks.

As for the first type of sensitivity analysis, the impact of (4) failing to hold can be easily characterised in our setting by means of the following relationship between the distributions of  $x$  for the benchmark groups  $W$ ,  $S1$  and  $NS2$  and the distributions of  $x$

for the three labour force states  $E$ ,  $U$  and  $OLF$ :

$$[f(x|W), f(x|S1), f(x|NS2)] = [f(x|E), f(x|U), f(x|OLF)] \mathcal{A}, \quad (7)$$

where  $\mathcal{A}$  is the  $3 \times 3$  matrix whose columns are the  $3 \times 1$  vectors of probabilities  $[p(E|R), p(U|R), p(OLF|R)]'$ ,  $R = W, S1, NS2$ . It follows that, after writing (3) in matrix notation:

$$f(x|R) = [f(x|W), f(x|S1), f(x|NS2)] \begin{bmatrix} p(W|R) \\ p(S1|R) \\ p(NS2|R) \end{bmatrix}, \quad R = S2, \dots, NS1$$

and substituting (7) into the last expression, the weights obtained through our identification strategy are related to the correct weights according to the following identity:

$$\begin{bmatrix} p(E|R) \\ p(U|R) \\ p(OLF|R) \end{bmatrix} = \mathcal{A} \begin{bmatrix} p(W|R) \\ p(S1|R) \\ p(NS2|R) \end{bmatrix} \quad R = S2, \dots, NS1. \quad (8)$$

The last expression clarifies that the restrictions in (4) set  $\mathcal{A}$  to the identity matrix  $\mathcal{I}_3$ , as we have  $p(E|W) = 1$ ,  $p(U|S1) = 1$  and  $p(OLF|NS2) = 1$ . When such restrictions are verified, the mixture weights identified by our exercise are in fact the correct ones. If  $\mathcal{A} \neq \mathcal{I}_3$ , they are affected by the contamination of the three benchmark groups.

We investigate the sensitivity of our results to the presence of contamination by allowing for non-zero values off the diagonal of  $\mathcal{A}$ , and then recovering the true weights using (8). We will focus on the two main results from our analysis, namely that (i) there are no inactive individuals in the  $S2$  group, and that (ii) the  $NS1$  group nearly comprises only inactive individuals, and we will study their sensitivity to different levels of contamination. Table 5 summarises these results by presenting estimates of weights obtained by pooling our samples over time while preserving the geographical breakdown.

After noting that, according to our estimation results,  $p(W|S2) \simeq 0$  and  $p(NS2|S2) \simeq 0$  and that (8) implies:

$$p(OLF|S2) = [p(OLF|W), p(OLF|S1), p(OLF|NS2)] \begin{bmatrix} p(W|S2) \\ p(S1|S2) \\ p(NS2|S2) \end{bmatrix},$$

it follows from the last expression that:

$$p(OLF|S2) \simeq p(OLF|S1).$$

In words, our estimation results imply that the true fraction of inactive individuals in the  $S2$  group depends on the value of just one of the probabilities off the main diagonal of the matrix  $\mathcal{A}$ . It therefore follows that our first result, namely that  $p(OLF|S2) = 0$ , would be severely biased only if the  $S1$  group - people reporting no work in the reference period, immediate availability and active steps taken in the thirty days prior to the interview - comprised a large fraction of inactives.

By applying the same line of reasoning and by noting that our estimation results imply  $p(W|NS1) \simeq 0$  and  $p(S1|NS1) \simeq 0$ , it also follows that:

$$p(OLF|NS1) \simeq p(OLF|NS2).$$

Accordingly, our second result, namely that  $p(NS1|NS2) = 1$ , would be severely biased only if the  $NS2$  group - people reporting no work in the reference period and no availability/interest for work regardless of the work arrangement - included a large fraction of active.

As for the  $S3$  and  $S4$  groups, the relationship in (8) leads to less clearcut implications: the  $p(S1|R)$  probabilities,  $R = S3, S4$ , are in fact very large across geographical areas (varying from 0.67 for  $S3$  in the North to 0.94 for  $S4$  in the North) but not equal to one as for the  $S2$  group. It follows that the bias in the estimated probabilities  $p(OLF|S3)$  and  $p(OLF|S4)$  resulting from violations of the identifying restrictions is driven by, but not equal to,  $p(OLF|S1)p(S1|R)$ ,  $R = S3, S4$ , which - as in the case of the  $S2$  group - is large if the  $S1$  group comprises a large fraction of inactive.

As a further check of the robustness of our results, we relax assumption (4) by splitting the category  $W$  into two sub-categories depending on the number of hours worked

per week. In particular, we consider the group of individuals working full-time (i.e. more than 20 hours) *vis-a-vis* the group of part-timers (i.e. working less than 20 hours per week), and we use the former group as the benchmark to identify the  $x$  distribution for the employed.<sup>11</sup> The estimation procedure described in Section 3.1 is carried out by including the part-timers among the uncertain categories, thus estimating their probability of looking like full-timers, unemployed and inactive individuals. Results for part-timers are in Table 6, where the estimated mixture weights are reported separately for the five survey occasions and the three geographic areas. With minor exceptions for Northern Italy, the part-timers look very much the same as full-timers with respect to the characteristics we include in  $x$ . As a result, excluding them from the  $W$  benchmark group does not make the difference for the estimation of the mixture weights for the categories  $S2$  to  $NS1$ .

## 6.2 A cautionary assessment

The results presented above point to similarities between the  $S2, S3, S4, NS1$  groups on the one hand and the  $W, S1, NS2$  groups on the other, which have been established with respect to a set of individual characteristics  $x$ . It is based on these similarities that the mixture model yields an estimate of the proportion of individuals to be classified as  $T = E, U, OLF$  conditional on  $R = S2, S3, S4, NS1$ .

In principle, one might argue that there are other individual characteristics we are not accounting for,  $u$  say, relevant to labour force state membership. Clear-cut similarities with respect to  $x$  do not imply that individuals are similar also with respect to unobserved characteristics  $u$ . Otherwise stated, the pattern of the mixture weights could have been different had we had available the joint distributions  $f(x, u|R)$ . While admitting that

---

<sup>11</sup>The distribution of weekly hours worked has somehow changed over time for the sample considered in our analysis, with the proportion of married women working up to 20 hours per week increasing from 6.46 percent in 1984 to 9.85 percent in 2000. We decided to use the 20 hours threshold because lower thresholds such as 12 or 15 hours would result in small sub-sets of workers.

from a theoretical point of view there might be room for improvement, we believe that, in the light of the literature on labour supply, the  $x$  variables we consider are rich enough to make our results at least an evidence deserving careful consideration. Moreover, results from the literature are roughly in line with those that we presented here. In particular, Jones and Riddell (1999) and Brandolini *et al.* (2004) find that their dubious groups analogous to our  $S2$  and  $S3$  groups are definitely behaviourally distinct from the benchmark inactive group.

It is also worth pointing out again that a non-negligible fraction of individuals presenting  $R = S2, S3, S4$  shares similarities with the  $W$  group. Should we count them as employed? This is clearly a tempting interpretation, as it points to the existence of ‘underground workers’: some individuals could purposely lack to mention positive hours of work in the reference week and conceal in the ‘looking for a job’ groups, presumably without reporting recent steps for seeking work (thus precisely either in  $S2$  or in  $S3$  or in  $S4$ ). At least two additional considerations are worth making in this respect. First, it might well be that  $S2, S3, S4$  individuals differ from the  $W$  group along some unobservable dimension  $u$ , so that the general comment made above applies. Second, such an evidence might also point to the existence of individuals who are very close to people at work, both in term of employability and availability, but still queuing for a job. Note that if we were to classify these individuals as unemployed, the unemployment rate implied by the model would turn out even closer to the IC rates reinforcing the evidence in Figure 1.

## 7 Concluding remarks

In this paper we have addressed the problem of inferring the labour force state from the elementary information collected by the Labour Force Surveys. Following the ILO

guidelines, national statistical offices define the labour force state of individuals from information on the activity (hours of work) and the timing of the search steps undertaken during conventionally defined reference periods. We have discussed how classification errors may arise in practise because the conceptual definitions of the labour force state given by the ILO do not straightforwardly map into unique operational criteria, so that the classification of individuals into employment, unemployment and inactivity depends on the operational rule adopted. Previous research in the literature has shown that labour force statistics are in general sensitive to changes in the operational rules, all broadly consistent with the ILO guidelines. We have provided additional evidence about the implications of such a problem by exploiting information on married women from five waves of the Italian LFS between 1984 and 2000.

To shed light on the merits of different classification criteria we have focused on two alternative classification rules resulting from a strict and a less stringent interpretation of the condition of being actively seeking work (the Eurostat EC criterion and the Istat IC criterion, respectively). Evidence on the merits of the two competing criteria has been assessed in two steps. First, we have identified some benchmark groups of individuals whose labour force state is agreed upon by both the EC and the IC, and we have considered a set of variables  $x$  known to matter for the labour force state of married women. Second, we have focused on those individuals whose classification depends on the operational rule being followed, and we have established which benchmark group they look like the most with respect to the  $x$  variables.

Our main result is that the operational rules followed by Eurostat do not fit the evidence provided by our sample. We find that individuals not at work, reporting to seek work and to be immediately available for work but with no recent steps undertaken are similar (at least with respect to  $x$ 's we consider) to individuals who are unquestionably

in the labour force. The same individuals are instead currently classified as inactive by Eurostat. This result is robust to changes in the business cycle, to geographical area effects, to different levels of married women participation in the labour market as well as to changes in the survey questionnaire. While admitting that our conclusions might not be robust to additional variables omitted from the criterion to establish similarities across groups, we believe that the  $x$ 's considered here have proven important enough in the literature on labour supply to make our results a challenge for the current practice in labour force classification.



## References

- [1] Altissimo, F., Marchetti, D.J. and Oneto, G.P. (2000), “The Italian Business Cycle: Coincident and Leading Indicators and Some Stylized Facts”, *Temi di Discussione* No. 377, Roma: Banca d’Italia.
- [2] Bartholomew, D. (1997), “Editorial: The Measurement of Unemployment in the UK: the Position at June 1997”, *Journal of the Royal Statistical Society A*, 160, 3, 385-388.
- [3] Bassi F., Hagenaaers, J.A., Croon, M.A. and Vermunt J.K. (2000), “Estimating true changes when categorical panel data are affected by uncorrelated and correlated errors. An application to unemployment data”, *Sociological Methods and Research*, 29, 230-268.
- [4] Battistin, E. and Sianesi, B. (2004), “Misreported schooling and the returns to education: evidence from the UK”, mimeo, London: Institute for Fiscal Studies.
- [5] Blair, J., Menon, G. and Bickart, B. (1991), “Measurement Effects in Self vs. Proxy Response to Survey Questions: An Information-Processing Perspective”, in Biemer et. al. (eds), *Measurement Errors in Surveys*, New York: Wiley.
- [6] Bound, J., Brown, C. and Mathiowetz, N. (2001), “Measurement error in survey data”, in J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics. Vol. 5*, Amsterdam: North-Holland, 3705-3843.
- [7] Brandolini, A., Cipollone, P. and Viviano, E. (2004), “Does the ILO definition capture all unemployment”, *Temi di Discussione* No. 529, Roma: Banca d’Italia.
- [8] Bregger, J.E., and Haugen, S.E. (1995), “BLS introduces new range of alternative unemployment measures”, *Monthly Labor Review*, 118, 10, 19-26.

- [9] Casavola, P. and Sestito, P. (1994), “L’indagine Istat sulle forze di lavoro”, *Lavoro e Relazioni Industriali*, 1, 179-195.
- [10] Eurostat (1997), *Labour Force Surveys. Methods and definitions*, Luxembourg.
- [11] Everitt, B.S. and Hand, D.J. (1981), *Finite mixtures distributions*, London: Chapman and Hall.
- [12] Flinn, C. J. and Heckman, J.J. (1983), “Are unemployment and out of the labor force behaviorally distinct labor force states?”, *Journal of Labor Economics*, Vol. 1, No. 1, 28-42.
- [13] Goodman, L.A. (1974), “Exploratory latent structure analysis using both identifiable and unidentifiable models”, *Biometrika*, 61, 215-231.
- [14] Hagenaars, J.A. (1990), *Categorical Longitudinal Data: Log-linear, Panel, Trend and Cohort Analysis*, Newbury Park: Sage.
- [15] Hausman, J.A., Abrevaya, J. and Scott Marton, F.M. (1998), “Misclassifications of a dependent variable in a qualitative response setting”, *Journal of Econometrics*, 87, 239-287.
- [16] Hussmanns, R., Merhan, F. and Verma, S.M. (1990), *Surveys of economically active population, employment, unemployment and underemployment: an ILO manual on concepts and methods*, Geneva: International Labour Office.
- [17] ILO (1983), ”Resolution Concerning Statistics of the Economically Active Population, Employment, Unemployment and Underemployment, *Bulletin of Labor Statistics*, No.13, IX-XV.
- [18] Jones, S.R.G. and Riddell, W.C. (1999), ”The measurement of unemployment: an empirical approach”, *Econometrica*, 67, 147-162.

- [19] Kass, R.E. and Raftery, A.E. (1995), "Bayes factors", *Journal of the American Statistical Association*, 90, 773-795.
- [20] Killingsworth, M. and Heckman, J.J. (1986), "Female labour supply: a survey", in Ashenfelter, O. and Layard, R. (eds), *Handbook of Labor Economics*, Amsterdam: North-Holland, 103-204.
- [21] Malinvaud, E. (1986), *Sur les statistiques de l'emploi et du chômage*, Paris: La Documentation Française.
- [22] Maritz, J.S. and Lwin, T. (1989), *Empirical Bayes methods*, London: Chapman and Hall.
- [23] OECD (1987), "On the Margin of the Labour Force: An Analysis of Discouraged Workers and Other Non-Participants", *Employment Outlook*, Paris: OECD.
- [24] Rettore, E. and Trivellato, U. (1993), "A Double-Hurdle labour supply model with fallible indicators of labour force state", *Statistica*, 3, 341-367.
- [25] Rettore, E. and Trivellato, U. (1998), "La misura della disoccupazione e la modellazione dell'offerta di lavoro: definizioni a priori e stime dipendenti da modelli a confronto", in E. Giovannini (a cura di), *La misurazione delle variabili economiche e i suoi riflessi sulla modellistica econometrica*, Annali di Statistica, Serie X, vol. 15, Roma: Istat, 127-146.
- [26] Schneider, F. and Ernste, D. H. (2000), "Shadow economies: size, causes and consequences", *Journal of Economic Literature*, 38, 77-114.
- [27] Schwarz, G. (1978), "Estimating the dimension of a model", *The Annals of Statistics*, 6, 461-464.

- [28] Shiskin, J. (1976), "Employment and unemployment: the doughnut or the hole?", *Monthly Labour Review*, 99, 2, 3-10.
- [29] Sorrentino, C. (2000), "International unemployment rates: how comparable are they?", *Monthly Labor Review*, 123, 6, 3-20.
- [30] Steel, D. (1997), "Producing Monthly Estimates of Unemployment and Employment According to the International Labour Office Definition", *Journal of the Royal Statistical Society, Series A*, 160, 1, 5-46.
- [31] Trivellato, U. (1997), "Le misure della partecipazione al lavoro nel quadro comunitario", in L. Frey (a cura di), *Le informazioni sul lavoro in Italia: significato e limiti delle informazioni provenienti dal lato delle famiglie*, Quaderni di Economia del Lavoro, No.59, Milano: Franco Angeli, 9-34.
- [32] Working Party on the Measurement of Unemployment in the UK (1995), "The Measurement of Unemployment in the UK (with discussion)", *Journal of the Royal Statistical Society, Series A*, 158, 3, 363-417.
- [33] Zizza, R. (2002), "Metodologie di stima dell'economia sommersa: un'applicazione al caso italiano", Temi di Discussione No. 463, Roma: Banca d'Italia.

Table A.1: Descriptive statistics for the 1984 and 1990 samples

	Working			Actively searching			Not searching			Working			Actively searching			Not searching							
	W	S1	S2	S3	S4	NS1	NS2	NS2	W	S1	S2	S3	S4	NS1	NS2	NS2	W	S1	S2	S3	S4	NS1	NS2
<b>North 1984</b>																							
W's age	37.44	34.78	33.95	36.93	33.50	46.32	43.51		38.64	34.32	35.27	36.12	34.60	47.53	44.70		38.64	34.32	35.27	36.12	34.60	47.53	44.70
H's age	40.89	38.43	37.59	41.15	37.32	49.71	47.05		42.05	38.11	39.05	40.14	37.63	50.93	47.96		42.05	38.11	39.05	40.14	37.63	50.93	47.96
W's education	8.28	7.21	7.49	7.07	7.69	5.95	6.41		9.20	8.14	8.23	7.98	8.25	6.44	7.19		9.20	8.14	8.23	7.98	8.25	6.44	7.19
H's education	8.21	7.47	7.61	7.06	7.80	6.56	7.24		8.97	8.04	8.23	7.91	8.50	6.96	7.84		8.97	8.04	8.23	7.91	8.50	6.96	7.84
Nr. Children	1.32	1.34	1.35	1.50	1.20	1.07	1.54		1.34	1.22	1.30	1.55	1.33	1.14	1.51		1.34	1.22	1.30	1.55	1.33	1.14	1.51
<b>Center 1984</b>																							
W's age	39.29	33.86	32.66	36.32	34.87	44.92	41.43		39.42	33.96	33.11	37.51	35.26	43.52	43.27		39.42	33.96	33.11	37.51	35.26	43.52	43.27
H's age	42.74	37.67	36.21	40.71	38.43	48.09	45.58		43.24	38.28	37.21	41.11	38.77	48.07	47.35		43.24	38.28	37.21	41.11	38.77	48.07	47.35
W's education	8.45	8.26	8.79	9.75	8.04	6.16	6.68		9.53	8.97	9.93	9.34	9.53	7.24	7.37		9.53	8.97	9.93	9.34	9.53	7.24	7.37
H's education	8.56	7.89	8.50	9.18	7.51	6.82	7.52		9.19	8.67	9.61	8.23	9.06	8.10	7.93		9.19	8.67	9.61	8.23	9.06	8.10	7.93
Nr. Children	1.43	1.46	1.32	1.54	1.45	1.11	1.57		1.49	1.38	1.39	1.57	1.29	1.40	1.51		1.49	1.38	1.39	1.57	1.29	1.40	1.51
<b>South 1984</b>																							
W's age	38.91	32.86	32.01	32.31	33.77	41.19	39.66		39.31	32.74	32.77	33.48	34.78	41.32	39.80		39.31	32.74	32.77	33.48	34.78	41.32	39.80
H's age	42.74	37.24	36.45	36.72	37.93	44.78	43.93		43.24	37.00	37.18	38.00	37.99	45.59	43.96		43.24	37.00	37.18	38.00	37.99	45.59	43.96
W's education	8.88	8.45	8.77	10.06	8.34	5.94	6.29		9.65	8.68	8.96	10.19	8.75	6.63	7.02		9.65	8.68	8.96	10.19	8.75	6.63	7.02
H's education	8.87	7.89	8.31	10.61	7.91	6.62	6.93		9.49	8.54	8.51	8.97	8.79	7.41	7.54		9.49	8.54	8.51	8.97	8.79	7.41	7.54
Nr. Children	1.88	1.77	1.56	1.89	1.73	1.64	2.10		1.79	1.62	1.63	1.76	1.72	1.55	1.98		1.79	1.62	1.63	1.76	1.72	1.55	1.98

Education is measured as years of schooling.

Table A.2: Descriptive statistics for the 1993 and 1995 samples

	Working			Actively searching			Not searching			Working			Actively searching			Not searching								
	W	S1	S2	S3	S4	NS1	NS2	NS2	W	S1	S2	S3	S4	NS1	NS2	NS2	W	S1	S2	S3	S4	NS1	NS2	
<b>North 1993</b>																								
W's age	39.42	36.39	37.00	36.18	38.33	43.89	43.12		39.26	36.54	37.16	37.68	36.59	45.19	43.89		39.26	36.54	37.16	37.68	36.59	45.19	43.89	
H's age	42.70	39.90	40.60	39.90	40.89	48.36	46.60		42.51	40.24	40.46	42.32	40.06	48.84	47.38		42.51	40.24	40.46	42.32	40.06	48.84	47.38	
W's education	9.86	8.54	8.81	9.18	8.74	7.28	7.61		10.10	9.29	8.96	8.72	9.44	7.03	8.05		10.10	9.29	8.96	8.72	9.44	7.03	8.05	
H's education	9.56	8.84	8.61	8.58	9.07	7.64	8.25		9.93	9.26	9.51	9.05	9.50	7.54	8.87		9.93	9.26	9.51	9.05	9.50	7.54	8.87	
Nr. Children	1.26	1.29	1.26	1.47	1.04	1.30	1.53		1.22	1.26	1.30	1.53	0.84	1.29	1.45		1.22	1.26	1.30	1.53	0.84	1.29	1.45	
<b>Center 1993</b>																								
W's age	40.08	34.33	34.86	33.46	36.13	43.68	41.18		41.20	35.30	36.03	36.48	36.14	44.73	41.55		41.20	35.30	36.03	36.48	36.14	44.73	41.55	
H's age	43.39	38.57	38.70	36.94	42.13	47.40	45.19		44.90	38.88	39.93	40.51	39.57	48.88	45.74		44.90	38.88	39.93	40.51	39.57	48.88	45.74	
W's education	10.29	9.76	9.54	9.97	8.50	6.81	7.75		10.26	9.58	9.54	9.11	11.43	7.40	8.18		10.26	9.58	9.54	9.11	11.43	7.40	8.18	
H's education	10.41	9.54	9.41	9.34	7.13	7.04	8.45		10.24	9.68	9.57	9.32	11.86	7.61	8.81		10.24	9.68	9.57	9.32	11.86	7.61	8.81	
Nr. Children	1.43	1.39	1.38	1.54	1.88	1.51	1.57		1.44	1.42	1.44	1.46	0.71	1.56	1.53		1.44	1.42	1.44	1.46	0.71	1.56	1.53	
<b>South 1993</b>																								
W's age	39.66	33.56	34.10	32.15	35.88	39.72	38.65		40.34	34.16	34.79	35.40	37.60	40.11	38.94		40.34	34.16	34.79	35.40	37.60	40.11	38.94	
H's age	43.34	37.72	38.06	37.05	39.76	44.03	42.83		44.53	38.27	38.91	39.87	42.75	44.19	43.19		44.53	38.27	38.91	39.87	42.75	44.19	43.19	
W's education	10.95	8.77	8.75	9.62	7.20	7.15	7.46		11.01	9.18	9.22	9.45	7.10	7.19	7.69		11.01	9.18	9.22	9.45	7.10	7.19	7.69	
H's education	10.50	8.66	8.72	9.26	6.76	7.75	8.14		10.50	8.95	9.03	9.22	7.75	7.82	8.25		10.50	8.95	9.03	9.22	7.75	7.82	8.25	
Nr. Children	1.72	1.74	1.70	1.68	2.12	1.67	1.83		1.79	1.73	1.71	1.92	2.05	1.62	1.88		1.79	1.73	1.71	1.92	2.05	1.62	1.88	

Education is measured as years of schooling.

Table A.3: Descriptive statistics for the 2000 sample

	<i>Working</i>			<i>Actively searching</i>			<i>Not searching</i>		
	<i>W</i>	<i>S1</i>	<i>S2</i>	<i>S3</i>	<i>S4</i>	<i>NS1</i>	<i>NS2</i>	<i>NS3</i>	<i>NS4</i>
<b>North 2000</b>									
W's age	40.44	38.22	38.51	40.20	43.20	47.58	47.66	47.58	47.66
H's age	43.67	41.63	42.32	43.99	46.20	51.91	51.06	51.91	51.06
W's education	10.22	8.93	8.42	8.42	8.04	7.28	7.66	7.28	7.66
H's education	9.99	9.23	8.71	8.15	7.76	8.40	8.43	8.40	8.43
Nr. Children	1.16	1.26	1.30	1.43	1.22	0.95	1.21	0.95	1.21
<b>Center 2000</b>									
W's age	42.00	37.64	37.38	37.45	34.80	44.60	46.19	44.60	46.19
H's age	45.42	41.23	41.10	41.29	37.20	48.59	49.96	48.59	49.96
W's education	10.97	9.73	9.52	9.38	13.00	7.68	7.89	7.68	7.89
H's education	10.63	9.49	9.48	9.41	13.00	8.31	8.60	8.31	8.60
Nr. Children	1.30	1.36	1.40	1.40	0.80	1.22	1.32	1.22	1.32
<b>South 2000</b>									
W's age	42.04	37.19	36.51	37.79	42.93	42.06	43.01	42.06	43.01
H's age	45.61	41.56	40.66	41.49	46.13	46.29	47.18	46.29	47.18
W's education	11.50	8.98	8.83	8.70	8.13	7.16	7.39	7.16	7.39
H's education	10.96	8.49	8.67	8.80	8.47	7.29	8.10	7.29	8.10
Nr. Children	1.52	1.51	1.51	1.62	1.60	1.43	1.53	1.43	1.53

Education is measured as years of schooling.

Table A.4: Definition of cells

Number of children	Age of the youngest child	Education		Age of the wife
		Wife	Husband	
0		Low	Low	16-40
0		Low	Low	41-60
0		Low	High	
0		High	Low	
0		High	High	16-30
0		High	High	31-60
1	6-	Low	Low	16-30
1	6-	Low	Low	31-60
1	6-	Low	High	16-30
1	6-	Low	High	31-60
1	6-	High	Low	16-30
1	6-	High	Low	31-60
1	6-	High	High	16-30
1	6-	High	High	31-60
1	7-19	Low	Low	16-40
1	7-19	Low	Low	41-60
1	7-19	Low	High	
1	7-19	High	Low	
1	7-19	High	High	16-40
1	7-19	High	High	41-60
1	20+	Low	Low	
1	20+	other		
2+	6-	Low	Low	16-30
2+	6-	Low	Low	31-60
2+	6-	Low	High	
2+	6-	High	Low	
2+	6-	High	High	16-30
2+	6-	High	High	31-60
2+	7-19	Low	Low	16-30
2+	7-19	Low	Low	31-40
2+	7-19	Low	Low	41-60
2+	7-19	Low	High	16-40
2+	7-19	Low	High	41-60
2+	7-19	High	Low	16-30
2+	7-19	High	Low	31-40
2+	7-19	High	Low	41-60
2+	7-19	High	High	16-30
2+	7-19	High	High	31-40
2+	7-19	High	High	41-60
2+	20+	Low	Low	
2+	20+	other		

High education: 8+ years

Low education: 8- years