Lavoro: partecipazione, dinamica e valutazione di politiche

Misura, metodi, modelli

Progetto anno 2005

Inconsistencies in Reported Employment Characteristics among Employed Stayers: a Case-study with Two-wave Panels from the Italian Labour Force Survey, 1993-2003

F. Bassi[*], *A. Padoan*[**], *U. Trivellato*[*†]

* *Dipartimento di Scienze Statistiche, Università di Padova*
** *Ufficio di Statistica, Regione Veneto*
† *CESifo*

Working Paper n. 77, luglio 2008

La valutazione dell'impatto di interventi pubblici: metodi e studi di caso.

Cofinanziamento MIUR, anno 2005

Unità locali del progetto:

Dip. di Economia "S. Cognetti De Martiis", Univ. di Torino	(coord. B. Contini)
Dip. di Statistica "G.Parenti", Univ. Di Firenze	(coord. F. Mealli)
Dip. di Scienze Statistiche, Univ. di Padova	(coord. E. Rettore)
Dip. di Scienze Economiche, Univ.di Salerno	(coord. S. Destefanis)
Dip. di Politiche Pubbliche e Scelte Collettive, Univ. del Piemonte Orientale	(coord. D. Bondonio)

Dipartimento di Scienze Statistiche
via C. Battisti 241-243, 35121 Padova
www.valutazione2003.stat.unipd.it

Per la presente pubblicazione sono stati adempiuti gli obblighi previsti dalle norme per la consegna obbligatoria di esemplari degli stampati e delle pubblicazioni di cui alla legge del 2 febbraio 1939 n.374 e successive modificazioni, presso la Procura della Repubblica e la Prefettura di Padova.

1. Introduction[1]

In recent years labour markets in industrialised countries have shown quite a high degree of mobility. A large literature on the micro-dynamics of the labour market focuses on job-to-job flows (Steven and Haltiwanger, 1999; Fallick and Fleischman, 2004; Shimer, 2005, among many others). Only a few papers explore the kind of job changes that workers experienced. The complexity of job mobility demands to analyse also these changes while changing employer (Neal, 1999). Literature on job matching suggests that a significant fraction of workers who switches job changes also employment characteristics, mainly industry and occupation (Miller, 1984; McCall, 1990).

In Italy, information on job characteristics can be obtained from various sources. As a prominent example, WHIP (Work Histories Italian Panel), a panel built from a sample of micro-data from the administrative archives of the Italian social security agency (INPS), provides longitudinal information on work histories of employees in the private non-agricultural sector. Using data for the period 1985-1996 and a three-digit ATECO81 scheme, Leombruni and Quaranta (2002) show that 53% of job changes coincide with industry changes. Mobility among occupations, measured according to a 4-category classification[2], is lower: 7% of blue-collars and executives change occupation while changing job; the percentage rises to 17% for the white-collars. Updated estimates from the same authors (Leombruni and Quaranta, 2005) document the persistence of these patterns.

Another fundamental source for the analysis of short-tem dynamics and persistence in the Italian labour market is the Quarterly Labour Force Survey (QLFS). It has the distinctive advantage of referring to a sample of the resident non-institutional population. Thus, it collects information on job characteristics of (almost) all the employed. The survey is cross-sectional with a 2-2-2 rotating design, which allows one to get two-wave panels one quarter and one year apart (see, e.g., Trivellato, 1997).

A few studies document that job characteristics, in particular industry and occupation, collected in surveys are affected by measurement error (with reference to the QLFS, see, e.g., de Angelini and Giraldo, 2003). The effect of these errors is to exaggerate the occurrence of changes in such characteristics, at least when information is obtained at two points in time with independent interviews (Bound, Brown and Mathiowetz, 2001). Sala and Lynn (2006) compare estimates of change in industry and occupation obtained in two survey waves 17 months apart but with different interview techniques: traditional independent interviewing and dependent interviewing. They show that dependent interviewing results in lower levels of observed change and that this shrinkage represents a reduction in measurement error, since the effect is particularly pronounced among respondents who do not change job between waves. Other studies demonstrate that, in general, industry is reported more accurately than occupation and, not surprisingly, that the agreement rate between employees' and employers' reports classified according to a single-digit coding scheme is higher than that resulting when reports are categorized according to the more detailed three-digit classification (Mellow and Sider, 1983; Mathiowetz, 1992).

In this paper we deal with measurement error, and its potentially distorting role, in information on industry and occupation collected by the QLFS. We use yearly panel data for the period from April 1993 to April 2003. The focus of our analyses is on inconsistent information on

[1] Research for this paper was supported by grants TRIVPRIN03 and RETTPRIN05 from the Italian Ministry of Education, University and Research, for the years 2004-05 and 2006-07, under the PRIN Programme. Individual anonymised data from the Italian Labour Force Survey were kindly provided by Istat (the Italian statistical agency), under a research agreement with the Department of Statistics, University of Padova. An earlier version was presented at the European Conference on Quality in Official Statistics 2008, Rome, July 8-11, 2008. We are grateful to Guido Masarotto and Adriano Paggiaro for insightful suggestions on preliminary drafts, and to conference participants for useful comments. The usual disclaimer applies.

ADDRESS FOR CORRESPONDENCE: francesca.bassi@unipd.it

[2] The 4-category classification for occupation consists of Executives, White-collars, Blue-collars and Apprentices.

employment characteristics –industry and occupation – resulting from yearly transition matrices for workers who (reported they) where continuously employed over the year and did not change job.

First, we compute and comment upon some usual indicators of disagreement. We find neat evidence that there is sizable measurement error both in industry and occupation, and that industry is reported more accurately than occupation. Then, we expand our analysis in three directions: (i) we test whether the consistency of repeated information, provided by employment stayers, significantly increases when the number of categories is collapsed; (ii) we explore the pattern of inconsistencies among response categories using Goodman's (1968) quasi-independence model; (iii) we compare the appropriateness of alternative classifications jointly by occupation and industry.

As for the detail of variable classification for cross-section estimates (admittedly less demanding then estimates from two-wave panel data), Istat – the national statistical agency – provides the following indications. For occupation, a reliable classification is just a binary one: Employee and Self-employed. For industry, Istat asserts as dependable a classification into 12 categories. Based on the hierarchical Kappa coefficient, for each of the two variables – industry and occupation – we test if reducing the number of categories significantly increases the consistency of information reported in two interviews one year apart. Evidence from these analyses supports the first indication by Istat, but casts severe doubts on the second one. Significant results in terms of measurement error reduction are obtained for a 6 or 5-category classification of industry.

We further explore the patterns of inconsistencies among variables' categories by testing several specifications of Goodman's quasi-independence model. The quasi-independence model is almost always rejected

Finally, we consider and compare alternative 4-category classifications obtained by collapsing industry and occupation into a single variable. The standard classification labels respondents as Self-employed, Employee in agriculture, Employee in industrial sector, and Employee in services. An alternative, still 4-category classification was recently used by Trivellato *et al.* (2005) in their study of worker turnover, with the four categories given by Self-employed, Employee in agriculture, Employee in industrial sector and private services, Employee in Public Administration and social services. Interestingly enough, the latter classification turns out to be almost uniformly better than the former, standard one.

The paper proceeds as follows. Section 2 contains a brief description of the data and presents the methods and the design of the analyses. Section 3 reports the main results. Section 4 concludes.

2. **Data and methods**

2.1. Data

As anticipated, QLFS is a quarterly survey with a 2-2-2 rotating design[3]. It collects information about labour market participation on a sample of respondents from the resident non-institutional population. For all persons who declare themselves as employed or report they worked at least one hour during the reference week, the questionnaire includes a series of questions on employment characteristics: working hours, occupation, industry, profession, location of the firm, number of employees of the firm, type of contract, date since when working with the current employer or in the current activity.

[3] The description of the survey given here applies till 2003. In 2004 the survey has been substantially redesigned. The main change regards the timing of interview: while in the pre-2004 QLFS interviews took place just in one week – usually the second – of each quarter, with the new "Continuous" LFS interviews are spread out over all weeks of the quarter. Other notable changes regard the largely new questionnaire and the mode of interview: from paper and pencil to CAPI-CATI (Istat, 2004).

Occupation is asked by means of a closed-form question with 11 categories, 6 for employees (Manager, Executive, Clerk, Workman, Apprentice, Outworker) and 5 for self-employed (Entrepreneur, Professional, Own-account worker, Member of a producers' cooperative, Contributing family worker). Information on industry is collected with an open-ended question and coded by the interviewer according to the ATECO2002 classification. As already pointed out, Istat (2003) warns that these two variables may not be totally reliable, if used at their maximum degree of detail. It suggests to use the binary classification, Employee/Self-employed, for occupation and a 12-category classification, corresponding to the two-digit ATECO2002, for industry[4].

In this paper we use 10 two-wave yearly panels, from April to April, 1993-2003. Among workers who were continuously employed during the year – yearly employment stayers, we consider only those who did not change job, for a total of 263,584 sample units, around 25,000 for each panel. We adopted a fairly strict criterion for identifying these workers: they are those respondents who, in both one year apart interviews, were classified as employed and reported the same date when answering to the question: "When did you start working with the current employer or in the current self-employment?" [5].

Following de Angelini and Giraldo (2003), we consider inconsistencies in job characteristics reported one year apart as due to measurement errors affecting these variables, while we assume that no or negligible errors are made in reporting dates. We also assume that, among yearly employment stayers who did not change employer (or current self-employment), genuine levels of change in industry and/or occupation are likely to be very low (Sala and Lynn, 2006). Thus, operationally for those workers all observed change in industry and/or occupation is attributed to measurement error (Mathiowetz and McGonagle, 2000).

As for occupation, measurement error is likely to be due to the detailed classification offered to respondents; for industry, mainly to the nature of the open-ended question, which is used to collect information. Minor changes in the wording used by the respondents to describe their branch of activity or minor changes in the recoding of the information by the interviewer might lead to a different industry classification, when in fact no change has occurred.

2.2. Methods and design of the analyses

Our study consists of analysing inconsistencies in industry and occupation reported in two independent interviews one year apart by workers who were (sensibly assumed to be) continuously employed and did not change job. The analysis develops following several lines.

First, usual descriptive indicators to assess inconsistencies are computed and compared across the two variables and over time.

Transition matrices among job characteristics declared one year apart provide the basic information for quantifying inconsistencies. As an example, Tables 1 and 2 report, the transition matrices by industry and occupation, respectively, of the April 1993-April 1994 panel. The frequencies on the main diagonal of the matrices refer to consistent responses, while the frequencies outside the main diagonal point out to inconsistencies.

A simple indicator of disagreement is the percentage of frequencies outside the main diagonal (*P*).

[4] The 12 categories for industry are: Agriculture; Mining and raw material extraction; Manufacturing; Construction; Wholesale and retail trade; Accommodation and food services; Transportation and communication; Financial and real estate activities; Professional and support service activities; Public Administration, defence and compulsory social security; Education, health and other social services; Other public, social and personal service activities.

[5] The criterion is conservative: in accordance with the aim of our work, it is meant to reduce at the minimum the risk of including dubious cases. Thus, we decided to eliminate from the analysis (i) records with missing data in dates, and (ii) records with dates that differ in the two interviews, though they are consistent with the fact that the worker was continuously employed during the year and did not change job.

Cohen's (1960) Kappa coefficient, $K = (p_o - p_e)/(1 - p_e)$, where p_o is the observed proportion of agreement and p_e is the proportion of agreement expected under independence of answers, is a well established index of agreement. It is based on the comparison between observed counts on the main diagonal of the matrix with the corresponding expected cell counts under the model of independence. It ranges form –1 (total disagreement) to 1 (perfect agreement). The value 0 is obtained when agreement is totally due to chance[6] [7].

Table 1: *Transition matrix by industry, April 1993 to April 1994*

	1994											
1993	Agric.	Mining	Manuf.	Constr.	Wholes.	Accom.	Transp.	Finance	Profess.	P.A.	Education	Other
Agric.	1,624	4	26	8	17	4	2	1	3	30	7	4
Mining	1	270	32	22	14	0	2	4	2	8	1	1
Manuf.	16	28	5315	87	179	15	41	11	43	29	37	36
Constr.	12	12	90	1,721	28	0	16	6	27	17	12	21
Wholes.	24	6	171	55	3,648	40	31	6	22	19	28	34
Accom.	2	1	6	5	26	705	10	0	0	6	15	14
Transp.	11	2	47	22	31	4	1,306	10	4	44	11	21
Finance	2	5	17	8	21	2	10	777	12	12	11	10
Profess.	2	2	39	39	32	3	9	23	814	22	22	65
P.A.	23	11	37	22	25	10	48	13	12	2,098	164	27
Education	7	2	25	10	28	13	10	10	11	108	3,188	38
Other	9	5	25	14	39	19	25	8	36	40	58	938

Second, we want to ascertain if reducing the number of the categories in which the variables are coded significantly increases the consistency of information reported in two subsequent interviews. We test that by using the hierarchical Kappa coefficient (Cohen, 1968).

The hierarchical Kappa coefficient provides a framework for investigating if patterns of disagreement pertain primarily to interchanges among similar response categories, as opposed to substantively important misclassification. Partial credit is permitted for a certain type of

[6] Other indicators suggested by the literature, given a transition matrix with absolute frequencies, are (see, e.g., Hauser and Massagli, 1983): (i) The net difference rate $e_i = [(X_{.i} - X_{i.}) / X_{..}] \times 100$, where $X_{.i}$ is the ith marginal column sum, $X_{i.}$ is the ith marginal row sum and $X_{..}$ is the total count. This expression is simply the difference between interviews in the marginal proportions in the same categories. It ranges form –100 to 100; the value 0 is obtained when marginal proportions for a category are exactly the same. Its limitation is that there may appear to be significant differences in marginal proportions in several categories as a result of a smaller number of net classification differences. (ii) The index of inconsistency: $I_i = \{(X_{.i} + X_{i.} - 2X_{ii})/[X_{.i}(X_{..} - X_{i.}) + X_{i.}(X_{..} - X_{.i})]/X_{..}\} \times 100$, where X_{ii} is the ith diagonal entry. This is the ratio of observed discrepancies (off-diagonal counts) in a given category to those discrepancies expected under simple independence. It ranges from 0 – no inconsistencies – to 100 – complete randomness among answers. Its major defect is that the model of independence almost never fits with repeated measurements. We calculated the net difference rate and the index of inconsistencies on our transition matrices, and obtained results fully consistent with those revealed by indexes *P* and *K*, reported in Table 3.

[7] Cohen's Kappa is especially appropriate in the medical sciences, were studies are often designed to assess the agreement between different raters or different diagnostic instruments. If the two readings are from two different raters, *K* accounts for rater bias. On the contrary, if the two readings are from replicated measurements, an intraclass correlation coefficient may be appropriate, since we may assume no bias (Barnhart and Williamson, 2002). In our case, we aim at measuring agreement between responses obtained on the same sample in interviews one year apart: the interviewer usually is not the same; besides, also the respondent might change, since proxy respondents are allowed, and are indeed frequent (see, e.g., Gandolfo and Gennari, 2000, who state that, in the period from April 1998 to January 1999, the rate of proxy respondents in the QLFS was on average a little above 40%). For these reasons we do prefer Cohen's Kappa to the intraclass correlation coefficient.

disagreement, which implies assigning a set of weights to some specific matrix cells. For example, the weights can be chosen so that the associated Kappa measures indicate the increments in agreement which result by successively combining relevant categories of the response variable. In this case: $p_o = \sum_{i=1}^{I}\sum_{i'=1}^{I} w_{ii'} p_{ii'}$ and $p_e = \sum_{i=1}^{I}\sum_{i'=1}^{I} w_{ii'} p_{i.} p_{.i'}$, where $w_{ii'}$ is the weight assigned to the *ii'* cell, $p_{ii'}$ is the proportion observed in cell *ii'* and $p_{i.}p_{.i'}$ is the proportion expected under the model of independence (Koch *et al.*, 1977).

Table 2: *Transition matrix by occupation, April 1993 to April 1994*

	1994										
1993	Manager	Executive	Clerk	Workman	Appr.	Outw.	Entrep.	Prof.	Own-a	Coop.	Contr.
Manager	243	89	39	14	0	0	6	7	4	0	0
Executive	65	684	218	11	0	0	0	11	4	0	1
Clerk	55	284	6,790	465	0	2	13	31	44	8	12
Workman	15	16	557	7,882	25	11	8	6	138	22	35
Apprentice	0	0	11	79	98	1	0	1	2	0	1
Outworker	0	0	4	18	0	29	0	1	5	0	0
Entrepreneur	4	0	6	4	0	0	238	21	123	6	10
Professional	6	13	36	6	0	1	7	641	93	2	3
Own-account w	4	4	40	118	2	6	102	89	4,353	74	102
Coop's member	0	0	6	14	0	0	8	3	54	115	7
Contr. family w	0	3	36	36	5	1	12	7	117	14	767

The problem is formalised as a simple test of hypothesis. Let $\hat{K}_1$ and $\hat{K}_2$ be two hierarchical Kappa coefficients, estimated with two different sets of weights so that the second one implies a less disaggregated classification, the hypothesis test $H_o : \hat{K}_2 = \hat{K}_1$ *vs.* $H_1 : \hat{K}_2 > \hat{K}_1$ allows one to verify if aggregating categories improves significantly agreement (Landis and Koch, 1977). In our study, the weights $w_{ii'}$ are chosen so that they imply aggregation among categories identifying similar employment (industry or occupation) characteristics. The weights have value 1 for cells in which there is perfect agreement (on the main diagonal) and for cells outside the main diagonal linking similar categories – whose observed frequencies are considered as agreements, value 0 for all other cells.

Finally, the patterns of inconsistencies among response categories, at various levels of disaggregation, are explored by estimating log-linear models of quasi-independence.

Log-linear models can be usefully applied in order to detect inconsistencies in contingency tables (Hagenaars, 1990). In particular, the model of quasi-independence is used to evaluate if, when we leave the main diagonal cells aside, the remaining cells show particular systematic patterns of association or whether there is independence on this (truncated) table. In the model of quasi-independence the entries on the diagonal cells of a transition matrix are blocked, and the model of independence is specified for the off-diagonal cells (Goodman, 1968). The expression of the loglinear model of quasi-independence for an *IxI* table is the following:

$$\log F_{ij} = m + m_i + m_j + m_{ij},$$

where F_{ij} is the expected frequency in the generic cell of the two-way contingency table; m is the grand mean; m_i are row effects with $1 \le i \le I$; m_j are column effects with $1 \le j \le I$ and m_{ij} are interaction effects for diagonal cells, $m_{ij} = 0$ if $i \ne j$.

In our case, assuming quasi-independence implies that errors in reporting industry or occupation are independent in two interviews one year apart. Rejecting the model implies that inconsistencies do not occur randomly: rather, there are systematic patterns of associations among

response categories. After estimating the model, a close inspection of residual frequencies may give information on the sizes of associations.

3. Main results

3.1. Descriptive evidence

For our study we use 10 two-wave panels, obtained by linking information collected in two one-year apart occasions of the QLFS from April 1993 to April 2003. As clarified in Section 2.1, we extract those workers who were classified as continuously employed and did not change job. The total sample, over 10 years, consists of 263,584 workers.

Table 3 contains, for the 10 panels, the values of the descriptive statistics of inconsistency: the percentage of frequencies outside the main diagonal (*P*) and the Cohen's Kappa (*K*), with reference to industry classified with 12 categories, as recommended by Istat, and to occupation classified with 11 categories, as in the questionnaire.

It is worth noting that industry is reported with fewer inconsistencies than occupation, according to both indices[8]. Another interesting evidence is that there is no significant trend in the indices of inconsistencies: the effect of measurement error in the survey has been more or less constant over the decade.

Looking at the indices calculated for the various categories of the two variables – not reported here for the sake of space[9], it emerges that the most consistent categories of occupation are Clerk and Workman among the employees, and Professional and Own-account worker among the self-employed. As for industry, Agriculture, Mining and raw material extraction, Professional and support service activities and Other public, social and personal service activities are reported with least inconsistencies.

Table 3: *Measures of inconsistencies with reference to industry and occupation*

Panels	Industry		Occupation	
	P *	*K***	*P* *	*K***
93-94	11.8	0.8672	14.0	0.8132
94-95	10.6	0.8785	13.1	0.8255
95-96	10.9	0.8750	13.1	0.8265
96-97	9.5	0.8915	12.5	0.8349
97-98	9.7	0.8896	12.9	0.8297
98-99	10.6	0.8787	12.9	0.8301
99-00	10.9	0.8758	13.1	0.8276
00-01	10.9	0.8754	13.7	0.8195
01-02	10.3	0.8822	12.5	0.8346
02-03	9.7	0.8892	12.4	0.8355

* *P* is the percentage of frequencies outside the main diagonal.
** *K* is the Cohen's Kappa coefficient.

In order to interpret the values of Cohen's Kappa, two scales are mainly used in empirical studies. The scale proposed by Fleiss (1981) defines as marginal agreement values of the coefficient

[8] As already said, this result can be found in other studies. In our case study, though, it appears a bit puzzling, since occupation is reported answering to a closed-form question, while industry is asked by means of an open-ended question and answers are afterwards coded by the interviewer. Literature on measurement errors in surveys, specifically in reporting job characteristics, documents that inconsistencies over time are more likely when information is collected with open-ended questions (Mathiowetz and McGonagle, 2000).
[9] They are available from the authors on request.

which are lower than 0.40, as good agreement values between 0.41 and 0.75, as excellent agreement values over 0.75. Landis and Koch (1977) propose to consider as slight agreement values lower than 0.20, as fair agreement values between 0.21 and 0.40, as moderate agreement values between 0.41 and 0.60, as substantial agreement, values of the coefficient between 0.61 and 0.80. A value greater than 0.80 denotes almost perfect agreement.

As shown in Table 3, for both variables – industry and occupation – the values of Cohen's Kappa are pretty high, resulting in almost perfect agreement according to both scales. All coefficients, moreover, are statistically significant, *i.e.*, statistically different form zero. However, the peculiarity of our case should be taken into account. Indeed, in the absence of measurement error we do expect no inconsistencies between information reported in two subsequent interviews. A first evidence of non-negligible inconsistencies in our data is the percentage of frequencies outside the main diagonal of the matrices, around 10% for industry and around 12-14% for occupation. Besides, under the hypothesis of no inconsistencies one should expect a Kappa index equal, or very close, to 1; which does not seem to be the case[10]. Overall, we interpret these results as indicative of sizable measurement errors.

Therefore, we will proceed to verify if aggregating categories might improve agreement, that is to say significantly diminish the percentage of inconsistent information.

3.2. A strategy for testing a sequence of less disaggregated classifications

As already explained, the procedure based on the hierarchical Kappa statistics is appropriate for assessing the pattern of agreement among two or more classifications of some categorical response variable. A sequence of hierarchical Kappa statistics refers to progressively less stringent, usually nested, definitions of agreement. The values of the coefficient obtained yield larger values for corresponding broader views of agreement. Since the Kappa statistics have an approximate multivariate normal distribution for large samples, it is possible to test the significance of successive differences, by means of the Wald statistics.

Hierarchical Kappas are formulated using sets of criterion weights. In our application the first set of weights defines the agreement as the occurrence of the same response category in both interviews. The other sets of weights correspond to more aggregated classifications, which consider as agreement also the occurrence, in two consecutive interviews, of responses which are different but belong to similar categories. Similar categories are treated as equivalent, obtaining a less stringent definition of consistency.

Table 4: *Hierarchical Kappa coefficients and Wald test: industry*

Panels	Kappa coefficients				Wald test		
	12 categories	6 categories	5 categories	3 categories	6 *vs.* 12	5 *vs.* 6	3 *vs.* 5
93-94	0.8672	0.8833	0.8940	0.9020	217.75***	96.26***	26.59***
94-95	0.8785	0.8939	0.9037	0.9113	174.26***	71.25***	21.41**
95-96	0.8750	0.8899	0.8989	0.9005	193.51***	71.79***	1.17
96-97	0.8915	0.9044	0.9104	0.9159	165.64***	38.77***	14.39
97-98	0.8896	0.8982	0.9044	0.9082	81.08***	34.85***	6.19
98-99	0.8787	0.8894	0.8964	0.9008	107.82***	40.62***	7.65
99-00	0.8758	0.8880	0.8931	0.8944	132.95***	21.90**	0.65
00-01	0.8754	0.8865	0.8883	0.8903	109.25***	2.91	1.39
01-02	0.8822	0.8910	0.8926	0.8943	80.50***	2.47	1.08
02-03	0.8892	0.9008	0.9046	0.9044	131.86***	14.05	0.02

[10] Note that it would be quite complicated to build a test-statistic to ascertain the hypothesis that the empirical Kappa coefficients were significantly different from 1, since it implies testing a parameter value at the boundary of the parameter's space.

* Significant at a=0,1, ** significant at a=0,05, *** significant at a=0,01

For industry, the first set of weights is given by the 12-category classification recommended by Istat. The second set of weights results into a 6-category classification consisting of Agriculture, Manufacturing and mining (obtained aggregating Mining and raw material extraction and Manufacturing), Construction, Wholesale and retail trade, Services (obtained aggregating Accommodation and food services, Transportation and communication, Financial and real estate activities, Professional and support service activities) and Public Administration (obtained aggregating Public Administration, defence and compulsory social security, Education, health and other social services, Other public, social and personal service activities). The third set of weights corresponds to a 5-category classification derived from the previous one by aggregating Services and Public Administration into the category Other Activities. The last set of weights implies the usual 3-category classification: Agriculture, Industrial sector and Services (for a summarizing scheme, see Table A1 in Appendix).

Table 4 contains the values of the Kappa coefficients calculated using the four sets of weights/classifications on the 10 panels, and the results of the Wald tests performed on the differences between each hierarchical Kappa coefficient and the one corresponding to a more aggregated classification[11]. Switching from 12 to 6 categories significantly improves agreement among responses in all 10 panels; reducing further categories to 5, significantly improves agreement in 7 out of 10 panels; no significant increase is obtained when reducing answers to the usual 3-category classification. From these results it appears that the two classifications that minimise inconsistencies in information on industry collected in the QLFS are those with 6 or 5 categories. This evidence contradicts the recommendation by Istat to use the 12-category classification.

For occupation, the first set of weights consists of the 11-category classification used in the questionnaire. The second set of weights results into a 6-category classification consisting of White-collar (obtained aggregating Manager, Executive and Clerk), Blue-collar (obtained aggregating Workman and Apprentice), Outworker, Self-employed (obtained aggregating Entrepreneur, Professional and Own-account worker), Member of a producers' cooperative, Contributing family worker. The last set of weights corresponds to the binary classification recommended by Istat: Employee and Self-employed (for a summarizing scheme, see Table A2 in Appendix).

Table 5: *Hierarchical Kappa coefficients and Wald test: occupation*

Panels	Kappa coefficients			Wald test	
	11 categories	6 categories	2 categories	6 *vs.* 11	2 *vs.* 6
93-94	0.8132	0.8709	0.9317	1,035.71***	621.99***
94-95	0.8255	0.8803	0.9371	816.87***	486.58***
95-96	0.8265	0.8804	0.9361	931.28***	560.05***
96-97	0.8349	0.8904	0.9402	965.87***	487.31***
97-98	0.8297	0.8850	0.9406	927.84***	552.70***
98-99	0.8301	0.8863	0.9398	922.88***	512.05***
99-00	0.8276	0.8816	0.9388	874.81***	565.95***
00-01	0.8195	0.8764	0.9317	893.42***	481.62***
01-02	0.8346	0.8911	0.9413	922.73***	466.68***
02-03	0.8355	0.8894	0.9432	893.83***	527.38***

*** Significant at a=0,01.

Table 5 parallels Tables 4, and contains the values of the Kappa coefficients calculated using the three sets of weights on the 10 panels, as well as the results of the tests performed on the differences between the hierarchical Kappa coefficients. Switching from 11 to 6 categories

[11] Since the Kappa coefficients have an approximated multivariate normal distribution for large samples, chi-square tests for linear hypotheses about them can be carried out with Wald statistics.

significantly improves agreement among responses in all panels[12]; reducing categories to 2 further increases agreement. The recommendation by Istat to classify occupation with a dichotomous variable, Employee/Self-employed, is neatly confirmed.

Table 6: *Goodness-of-fit statistics for the quasi-independence model: industry*

Panels	Number of classes	X^2*	L^2*	Degrees of freedom	BIC index
93-94	12	1,310.95	1,180.35	109	74.87
	6	651.81	596.52	89	-306.12
	5	350.64	308.96	65	-350.27
	3	202.72	190.25	47	-286.42
94-95	12	1,117.23	993.54	109	-93.73
	6	557.06	514.57	89	-373.20
	5	234.53	219.63	65	-428.75
	3	111.99	108.30	47	-360.52
95-96	12	1,073.58	965.40	109	-138.19
	6	520.16	501.48	89	-399.61
	5	253.15	240.34	65	-417.77
	3	145.81	145.19	47	-330.67
96-97	12	1,038.38	896.49	109	-206.29
	6	411.06	393.80	89	-506.63
	5	220.96	212.57	65	-445.05
	3	117.53	119.77	47	-355.74
97-98	12	776.94	692.80	109	-407.74
	6	359.91	365.29	89	-533.31
	5	240.31	229.31	65	-426.97
	3	111.58	111.95	47	-362.59
98-99	12	990.77	887.54	109	-210.85
	6	511.34	489.06	89	-407.79
	5	277.61	273.25	65	-381.76
	3	138.15	142.21	47	-331.40
99-00	12	1,087.13	997.56	109	-102.56
	6	542.33	524.01	89	-374.25
	5	284.80	273.11	65	-382.92
	3	183.87	174.50	47	-299.86
00-01	12	976.55	861.55	109	-234.43
	6	360.33	351.77	89	-543.10
	5	204.88	197.68	65	-455.88
	3	125.39	131.36	47	-351.22
01-02	12	949.90	841.43	109	-258.07
	6	470.26	426.90	89	-470.86
	5	233.38	219.82	65	-435.85
	3	120.53	121.45	47	-352.65
02-03	12	952.57	873.38	109	-228.43
	6	434.87	428.31	89	-471.33
	5	244.36	235.23	65	-421.81
	3	113.92	115.33	47	-359.76

* All p-values are lower than 0.0001.

[12] In section 2.1 we motivated the assumption that among continuously employed workers who did not change job, genuine levels of change in occupation (and/or industry) are likely to be very low. In any case, with the 6-category classification true transitions among occupations for those workers are definitely implausible.

3.3. Does the quasi-independence model hold?

In the case of industry, the log-linear model of quasi-independence has been estimated so to reproduce the 12, 6, 5 and 3-cateogry classifications. Table 6 reports the results of model fitting (Pearson X^2 and log-likelihood ratio L^2 statistics and associated *p*-values and BIC index) for the 10 panels. The hypothesis of quasi-independence is always rejected, for all panels and all levels of aggregation, indicating that one-year apart responses show non-random association even at the maximum level of aggregation (3 categories).

Although the model of quasi-independence does not fit, the BIC index gives some interesting evidence. It reaches its minimum value in correspondence with the 6-category classification in 7 out of 10 panels, indicating that this is the level of aggregation with the best fit to the data; in correspondence with the 5-category classification in the remaining three panels, confirming the conclusions reached applying the hierarchical Kappa procedure.

Estimated residuals of a quasi-independence model measure association exceeding that expected under random behaviour. The greater their value, the higher the difference between observed association and that expected under the null hypothesis of randomness. Specifically, positive estimated residuals indicate that the model underestimates association between the two categories involved; negative estimated residuals indicate overestimation. Inspecting estimated residuals of the quasi-independence model implying 3 categories, *i.e.*, testing quasi-independence among inconsistencies which are due to response categories aggregated in different sectors (Agricultural, Industrial sector, Services) in the two occasions, we note that it underestimates association between the following couples of categories: Agriculture and Public Administration, defence and compulsory social security; Manufacturing and Wholesale and retail trade; Construction and Professional and support service activities. On the contrary, the model overestimates association between Manufacturing and Public Administration, defence and compulsory social security; Construction and Wholesale and retail trade; Professional and support service activities and Agriculture. It is hard to attribute these results only to inconsistencies in responses given one year apart, while it appears more convincing to ascribe them also to non-random measurement error affecting responses in each wave of the survey.

In the case of occupation, the log-linear model of quasi-independence has been estimated so to reproduce the 11, 6 and 2-cateogry classifications. Table 7 parallels table 6, and reports the results of model fitting for the classifications of occupation for the 10 panels. Also for this variable the hypothesis of quasi-independence is always rejected, for all panels and all levels of aggregation, pointing out that one year apart responses show non-random association even at the maximum level of aggregation (2 categories).

The BIC index reaches its minimum value always in correspondence with the 2-category classification. Such evidence indicates that this is the level of aggregation with the best fit to the data, and confirms again the conclusions reached applying the hierarchical Kappa procedure.

Inspecting estimated residuals of the quasi-independence model implying 2 categories, *i.e.*, testing quasi-independence among inconsistencies which are due to response categories classified as self-employed in one occasion and as employee in the other occasion, we note that it underestimates the association between the following couples of categories: Manager and Entrepreneur; Executive and Professional; Clerk and Professional. In this case patterns of association appear more reasonable and easier to understand than in the case of industry. The result shows, in fact, that residual association tends to concentrate among the highest classes of both employees and self-employed. Usually high positions, even if as employee, enjoy more flexible working conditions so that they might be confused with self-employment. Another sensible explanation could be that in Italy, as in other European countries, in recent years the standard dichotomising of workers into Employee/Self-employed and Employee has become too rigid, and is unable to cope with the growth of non-standard forms of employment (see, for example, Burchell, Deakin and Honey, 1999).

Table 7: *Goodness-of-fit statistics for the quasi-independence model: occupation*

Panels	Number of classes	X^2*	L^2*	Degrees of freedom	BIC index
93-94	11	3,941.76	3,452.42	89	2,549.77
	6	1,716.59	1,300.78	75	540.12
	2	252.24	240.00	39	-155.53
94-95	11	3,248.49	2,765.17	89	1,877.40
	6	1,332.34	981.83	75	233.71
	2	180.09	177.74	39	-211.28
95-96	11	3,816.64	3,182.43	89	2,281.33
	6	1,532.71	1,143.73	75	384.38
	2	210.07	207.18	39	-187.68
96-97	11	3,496.75	3,051.29	89	2,150.86
	6	1,346.41	1,047.90	75	289.11
	2	202.00	209.37	39	-185.20
97-98	11	3,502.93	3,035.06	89	2,136.46
	6	1,229.64	947.31	75	190.06
	2	129.44	129.67	39	-264.10
98-99	11	3,525.77	3,077.31	89	2,180.46
	6	1,433.62	1,047.19	75	291.42
	2	201.73	218.96	39	-174.04
99-00	11	3,469.02	3,103.43	89	2,205.16
	6	1,496.51	1,134.56	75	377.59
	2	198.24	208.87	39	-184.75
00-01	11	3,509.98	2,986.26	89	2,091.38
	6	1,280.38	958.92	75	204.81
	2	172.38	180.34	39	-211.80
01-02	11	3,454.76	3,029.94	89	2,132.18
	6	1,168.10	890.99	75	134.45
	2	169.99	175.73	39	-217.67
02-03	11	3,439.02	3,047.52	89	2,147.87
	6	1,324.17	990.89	75	25.55
	2	186.34	161.28	39	-232.94

* All p-values are lower than 0.0001.

3.4. Testing a set of different classifications jointly by occupation and industry

It might be of some interest to evaluate what happens in terms of inconsistencies of one year apart responses, when we specify joint classifications by occupation and industry, starting from one with 13 classes: Self-employed and the employees split up by the 12-category industry.

Table 8 presents the values of the hierarchical Kappa coefficients calculated using four sets of weights (that corresponding to the 13-category classification and those corresponding to a 7, 6 and 4-category classifications, respectively, obtained still considering Self-employed workers in one class and aggregating employees according to industry in 6, 5 and 3 categories, according to the strategy for collapsing classes previously used for the variable industry; see Table A.3 in the Appendix for a summarizing scheme) on the 10 panels, and the results of the tests performed on their differences. Switching from 13 to 7 categories significantly improves agreement among responses in all panels; reducing categories to 6 and then to 4 further increases significantly agreement.

Finally, we compare this 4-category classification (Self-employed, Employee in agriculture, Employee in the industrial sector, and Employee in services) with an alternative still 4-category

classification recently introduced, on heuristic grounds, by Trivellato *et al*. (2005) in their study of worker turnover, with the four categories given by Self-employed, Employee in agriculture, Employee in industrial sector and private services, Employee in public administration and social services (see again Table A.3 in Appendix for a summarizing scheme).

Table 8: *Hierarchical Kappa coefficients and Wald test – joint classification*

Panels	Kappa coefficients				Wald test		
	13 categories	7 categories	6 categories	4 categories	7 *vs*. 13	7 *vs*. 6	6 *vs*. 4
93-94	0.8713	0.8874	0.8972	0.9066	242.54***	108.54***	88.87***
94-95	0.8866	0.9020	0.9116	0.9205	198.08***	93.26***	73.29***
95-96	0.8822	0.8977	0.9058	0.9123	231.98***	84.40***	50.84***
96-97	0.8961	0.9096	0.9148	0.9213	198.07***	43.66***	53.38***
97-98	0.8957	0.9046	0.9121	0.9189	103.14***	69.91***	54.33***
98-99	0.8881	0.9003	0.9075	0.9143	156.82***	63.94***	51.69***
99-00	0.8834	0.8969	0.9036	0.9116	183.26***	55.30***	63.87***
00-01	0.8814	0.8942	0.8989	0.9058	160.76***	28.69	45.10***
01-02	0.8905	0.9014	0.9061	0.9125	137.23***	31.17***	43.51***
02-03	0.8945	0.9074	0.9132	0.9181	181.95***	14.05***	29.19***

* Significant at a=0,1, ** significant at a=0,05, *** significant at a=0,01.

As it is shown in Table 9, also in this case switching form 13 to 4 categories significantly increases agreement among responses in all panels. Besides, this alternative 4-category classification has a higher (and statistically significant) level of agreement in 9 (8) out of 10 panels. The overall $?^2$, with 10 degrees of freedom, is equal to 145.05, with a *p*-value close to zero, and definitely confirms that the latter 4-category classification is superior to the former one. It suggests that for respondents is easier to distinguish between private and social services, than between the bunch of all services (public, social and private) and the industrial sector.

Table 9: *Hierarchical Kappa coefficients and Wald test: alternative classifications*

Panels	Kappa coefficients		Wald test
	13 categories	4 categories	
93-94	0.8713	0.9083	492.59***
94-95	0.8866	0.9195	370.86***
95-96	0.8822	0.9170	466.25***
96-97	0.8961	0.9263	393.49***
97-98	0.8957	0.9224	315.62***
98-99	0.8881	0.9221	439.54***
99-00	0.8834	0.9178	438.78***
00-01	0.8814	0.9162	422.11***
01-02	0.8905	0.9234	419.25***
02-03	0.8945	0.9255	390.21***

*** Significant at a=0.01.

4. Conclusions

The focus of the paper is on inconsistencies in job characteristics reported in one-year apart independent waves of the QLFS by workers continuously employed and who did not change job. Transition matrices by occupation (collected with 11 categories) and industry (recoded in 12 categories) show a significant percentage of frequencies outside the main diagonal.

Aggregating categories improves agreement, as the application of the hierarchical Kappa procedure clearly demonstrates. For occupation the “best” level of aggregation results the binary one: Employee/Self-employed. For industry two classifications minimise inconsistencies: with 5

classes (Agriculture, Manufacturing and mining, Construction, Wholesale and trade, Other activities) and 6 classes (with a split of Other activities in Services and Public Administration), respectively. In the case of a joint classification of occupation and industry, the "best" level of aggregation is given by a 4-category classification recently advocated by Trivellato *et. al.* (2005), which distinguishes Self-employed, Employee in agriculture, Employee in industrial sector and private services, Employee in public administration and social services.

Inspection of estimated residuals of the log-linear model of quasi-independence – which does not fit the data even at the maximum level of aggregation – suggests that even cross-section information is affected by non-random measurement error, since not all residual association may be explained by inconsistencies among responses perceived as similar by respondents.

Abundant literature (Mathiowetz and McGonagle, 2000; Sala and Lynn, 2006; Lynn, Jäckle, Jenkins and Sala, 2006,among others) shows that dependent interviewing results in lower levels of observed change in job characteristics collected in two subsequent interviews, compared with change observed with independent interviewing. Besides, it documents that such reduction in observed change coincides with reduction of measurement error, since it is particularly pronounced among workers who do not change job.

The partly disappointing results of our analyses on 1993-2003 two-wave panel data from the QLFS, coupled with evidence from the literature on dependent interviewing, point out to the importance of innovations with the new "Continuous" LFS, in operation in Italy from 2004. It introduced a definitely better questionnaire and CAPI-CAPI mode of interview from the start, and is progressively extending dependent interviewing to several sections of the questionnaire, the one on job characteristics included.

References

de Angelini A., A. Giraldo (2003), "La mobilità dei lavoratori nel Veneto. Confronto fra misure su dati RTFL e su dati NETLABOR", Progetto di ricerca PRIN "Dinamiche e persistenze nel mercato del lavoro italiano ed effetti di politiche", *Working Paper n. 61*, Dipartimento di Scienze Statistiche, Padova.

Barnhart H.X., J.M. Williamson (2002), "Weighted lest-squares approach for comparing correlated Kappa", *Biometrics*, 58, 1012-1019.

Bound J., C. Brown, N. Mathiowetz (2001), "Measurement error in survey data", in *Handbook of Econometrics. Volume 5*, eds. J.J. Heckman and E. Leamer, Elsevier Science, Amsterdam.

Burchell B., S. Deakin, S. Honey (1999), "The Employment Status of Individuals in Non-standard Employment", *EMAR Publications No. 6*, Department of Trade and Industry, London.

Cohen J. (1960), "A coefficient of agreement for nominal tables", *Educational and Psychological Measurement*, 20, 37-46.

Cohen J. (1968), "Weighted kappa: Nominal scale agreement with provision for scale disagreement or partial credit", *Psychological Bulletin*, 70, 213-220.

Fallick B., C.A. Fleischman (2004), "Employer-to-employer Flows in the U.S. Labour Market: The Complete Picture of Gross Worker Flows", *Finance and Economics Discussion Series 2004-34*, Board of Governors of the Federal Reserve System (U.S.), Washington D.C..

Fleiss J.L. (1981), *Statistical Methods for Rates and Proportions*, Wiley, New York.

Gandolfo M., P. Gennari (2000), "Il lavoro sommerso, è possibile rilevarlo?", in Società Italiana di Statistica. Atti della XL Riunione Scientifica. Firenze 26-28 Aprile 2000, Dipartimento di Scienze Statistiche, Firenze.

Goodman L.A. (1968), "The analysis of cross-classified data: independence, quasi-independence, and interaction in contingency tables", *Journal of the American Statistical Association*, 63, 1019-1131.

Hagenaars J.A. (1990), *Categorical Longitudinal Data*, Sage, Newbury Park.

Hauser R.M., M.P. Massagli (1983), "Some models of agreement and disagreement in repeated measurements of occupation", *Demography*, 20, 449-460.

Koch G.G., J.R. Landis, J.R., Freeman, J.L., D.H. Freeman, R.G. Lehnen (1977), "A general methodology for the analysis of experiments with repeated measurements of categorical data", *Biometrics*, 33, 133-158.

Istat (2003), "Rilevazione trimestrale sulle forze di lavoro. Tracciato record del file standard longitudinale (Aprile 2003-Aprile 2002)", Roma (mimeo).

Istat (2004), "La nuova rilevazione sulle forze di lavoro - Contenuti, metodologie, organizzazione", Roma.

Landis J.R., G.G. Koch (1977), "The measurement of observer agreement for categorical data", *Biometrics*, 33, 159-174.

Leombruni R., R. Quaranta (2002), "Mobilità dei lavoratori in Italia, 1985-1996", in *Osservatorio sulla mobilità del lavoro in Italia*, a cura di B. Contini, Il Mulino, Bologna.

Leombruni R., R. Quaranta (2005), "Eppure si muoveva già. Una breve storia della mobilità dei lavoratori in Italia", in *Eppur si muove. Dinamiche e persistenze nel mercato del lavoro italiano*, a cura di B. Contini and U. Trivellato, Il Mulino, Bologna.

Lynn P., A. Jäckle, A., S.P. Jenkins, E. Sala (2006), "The effects of dependent interviewing on responses to question on income sources", *Journal of Official Statistics*, 22, 357-384.

Mathiowetz N. (1992), "Errors in reports of occupation", *Public Opinion Quarterly*, 56, 352-355.

Mathiowetz N., A. McGonagle (2000), "An assessment of the current state of dependent interviewing in household surveys", *Journal of Official Statistics*, 16, 401-418.

Mellow W., H. Sider (1983), "Accuracy of response in labor market surveys: Evidence and implications", *Journal of Labor Economics*, 1, 331-344.

McCall B.P. (1990), "Occupational matching: A test of sorts", *Journal of Political Economy*, 89, 45-69.

Miller R.A. (1984), "Job matching and occupation choice", *Journal of Political Economy*, 92, 1086-1120.

Neal D. (1999), "The complexity of job mobility among young men", *Journal of Labour Economics*, 17 (2), 237-261.

Sala E., P. Lynn (2006), "Measuring change in employment characteristics: The effects of dependent interviewing", *International Journal of Public Opinion Research*, 18, 500-509.

Shimer R. (2005), "The cyclicality of hires, separations, and job-to-job transitions", *Federal Reserve Bank of St Louis Review*, 87, 493-597.

Steven D.J., J. Haltiwanger (1999), "Measuring gross worker and job flows", in *Labor Statistics Measurement Issues*, eds. J. Haltiwanger, M. Manser and R. Topel, University of Chicago Press, Chicago.

Trivellato U. (1997), "Le misure della partecipazione al lavoro nel quadro comunitario", in *Le informazioni sul lavoro in Italia: significato e limiti delle informazioni provenienti da indagini sulle famiglie*, a cura di L. Frey, Quaderni di Economia del Lavoro n. 59, Franco Angeli, Milano.

Trivellato U., A. Paggiaro, R. Leombruni, R. Rosati (2005), "La dinamica recente della mobilità dei lavoratori, 1998-2003", in *Eppur si muove. Dinamiche e persistenze nel mercato del lavoro italiano*, a cura di B. Contini and U. Trivellato, Il Mulino, Bologna.

Inconsistencies in Reported Employment Characteristics among Employed Stayers: a Case-study with Two-wave Panels from the Italian Labour Force Survey, 1993-2003

Summary

The paper deals with measurement error, and its potentially distorting role, in information on industry and occupation collected by the Italian Quarterly Labour Force Survey. We use yearly panel data for the period from April 1993 to April 2003. The focus of our analyses is on inconsistent information on employment characteristics –industry and occupation – resulting from yearly transition matrices for workers who where continuously employed over the year and did not change job. The analysis goes trough four steps: (i) descriptive indicators of (dis)agreement; (ii) testing whether the consistency of repeated information significantly increases when the number of categories is collapsed; (iii) examination of the pattern of inconsistencies among response categories using Goodman's quasi-independence model; (iv) comparisons of alternative classifications jointly by occupation and industry. Results document sizable measurement error, which is only moderately reduced by more aggregated classifications. They suggest that even cross-section estimates of employment by industry and/or occupation are affected by non-random measurement error.

Keywords: Industry, Occupation, Measurement Errors, Survey Data.

JEL classification: C33; C42; J60

Appendix *Prospect summarizing categories aggregation process*

Table A1 *Industry*

<table>
<tr><td>12 categories</td><td>6 categories</td><td>5 categories</td><td>3 categories</td></tr>
<tr><td>Agriculture</td><td>Agriculture</td><td>Agriculture</td><td>Agriculture</td></tr>
<tr><td>Mining and raw material extraction</td><td rowspan="2">Manufacturing and mining</td><td rowspan="2">Manufacturing and mining</td><td rowspan="3">Industrial sector</td></tr>
<tr><td>Manufacturing</td></tr>
<tr><td>Construction</td><td>Construction</td><td>Construction</td></tr>
<tr><td>Wholesale and retail trade</td><td>Wholesale and retail trade</td><td>Wholesale and retail trade</td><td rowspan="10">Services</td></tr>
<tr><td>Accommodation and food services</td><td rowspan="4">Services</td><td rowspan="7">Other activities</td></tr>
<tr><td>Transportation and communication</td></tr>
<tr><td>Financial and real estate activities</td></tr>
<tr><td>Professional and support service activities</td></tr>
<tr><td>Public Administration, defence and compulsory social security</td><td rowspan="3">Public Administration</td></tr>
<tr><td>Education, health and other services</td></tr>
<tr><td>Other public, social and personal service activities</td></tr>
</table>

Table A2 *Occupation*

11 categories	6 categories	2 categories
Manager	White-collar	Employee
Executive		
Clerk		
Workman	Blue-collar	
Apprentice		
Outworker	Outworker	
Entrepreneur	Self-employed	Self-employed
Professional		
Own-account worker		
Member of a producers' cooperative	Member of a producers' cooperative	
Contributing family worker	Contributing family worker	

Table A3: *Joint classification by occupation and industry*

<table>
<tr><td>13 categories</td><td>7 categories</td><td>6 categories</td><td>4 categories</td><td>4 categories alternative class.*</td></tr>
<tr><td>Self-employed</td><td>Self-employed</td><td>Self-employed</td><td>Self-employed</td><td>Self-employed</td></tr>
<tr><td>Employee in:</td><td>Employee in:</td><td>Employee in:</td><td>Employee in:</td><td>Employee in:</td></tr>
<tr><td>Agriculture</td><td>Agriculture</td><td>Agriculture</td><td>Agriculture</td><td>Agriculture</td></tr>
<tr><td>Mining and raw material extraction</td><td rowspan="2">Manufacturing and mining</td><td rowspan="2">Manufacturing and mining</td><td rowspan="3">Industrial sector</td><td rowspan="6">Industrial sector and private services</td></tr>
<tr><td>Manufacturing</td></tr>
<tr><td>Construction</td><td>Construction</td><td>Construction</td></tr>
<tr><td>Wholesale and retail trade</td><td>Wholesale and retail trade</td><td>Wholesale and retail trade</td><td rowspan="6">Services</td></tr>
<tr><td>Accommodation and food services</td><td rowspan="4">Services</td><td rowspan="6">Other activities</td></tr>
<tr><td>Transportation and communication</td></tr>
<tr><td>Financial and real estate activities</td></tr>
<tr><td>Professional and support service activities</td></tr>
<tr><td>Public Administration, defence and compulsory social security</td><td rowspan="3">Public Administration</td><td rowspan="3">Public Administration and social services</td></tr>
<tr><td>Education, health and other services</td></tr>
<tr><td>Other public, social and personal service activities</td></tr>
</table>

*Classification used by Trivellato *et al.* (2005).

Working Papers

77. F. Bassi, A. Padoan, U. Trivellato, Inconsistencies in Reported Employment Characteristics among Employed Stayers: a Case-study with Two-wave Panels from the Italian Labour Force Survey, 1993-2003, Luglio 2008

76. D. Bondonio, Evaluating the Effects of Business Incentives Policies: A Discussion Note on Identification Strategies, Luglio 2008

75. B. Anastasia, D. Maurizio, Salari e stipendi in Veneto 2001-2002, Febbraio 2007

74. G. Tattara, M. Valentini, Labour market segmentation, flexibility and precariousness in the Italian North East, Aprile 2006

73. B. Anastasia, D. Maurizio, I lavoratori dipendenti in Veneto 1998-2003: profili e percorsi. Statistiche sistematiche da "Giove 2005", Dicembre 2005

72. A. Fossaluzza, Evaluation of the Effects of Two Classes of Active Labour Market Policies for Welfare Recipients: a Danish Study Case, Dicembre 2005

71. P. Cipollone, A. Guelfi, Financial support to permanent jobs. The Italian case, Novembre 2005

70. F. Bassi, U. Trivellato, The latent class approach to estimating gross flows affected by correlated classification errors, with application to data from the French Labour Force Survey, Novembre 2005

69. M. Fort, Education and the timing of births: evidence from a natural experiment in Italy, Novembre 2005

68. D. Contini, N. Negri, Would declining exit rates from welfare provide evidence of welfare dependence in homogeneous environments?, Ottobre 2005

67. E. Battistin, E. Rettore, U. Trivellato, Choosing among alternative classification criteria to measure the labour force state, Ottobre 2005

66. G. Tattara, M. Valentini, Evaluating the Italian training on the job contract (CFL), Novembre 2005

65. A. Paggiaro, E. Rettore, U. Trivellato, The impact of the Italian "Mobility Lists" on employment chances: new evidence from linked administrative archives, Ottobre 2005

64. D. Bondonio, R.T. Greenbaum, Do Tax Incentives Affect Local Economic Growth? What Mean Impacts Miss in the Analysis of Enterprise Zone Policies, Ottobre 2005

63. D. Bondonio, The employment impact of business incentive programs in declining areas. Mean impacts versus impacts by degrees of economic distress, Ottobre 2005

62. U. Trivellato, F. Bassi, A. Discenza, A. Giraldo, Transizioni e mobilità nel mercato del lavoro italiano, 1979-2003, Giugno 2005

61. A. De Angelini, A. Giraldo, La mobilità dei lavoratori nel Veneto.Confronto fra misure su dati RTFL e su dati NETLABOR, Settembre 2003

60. F. Bassi, E. Salvan, Dinamiche di ricollocamento per lavoratori che perdono un'occupazione stabile, Settembre 2003

59. M. Romano, Abbinamento tra gli archivi INPS e Netlabor: il caso degli iscritti alle liste di mobilità, Settembre 2003

58. P. Baretta, U. Trivellato, La mobilità dei lavoratori da fonti amministrative e da surveys sulle famiglie: un'analisi comparata, Settembre 2003

57. D. Bondonio, R.T. Greebaum, A comparative evaluation of spatially targeted economic revitalization programs in the European Union and the United States, Settembre 2003

56. F. Berton, L. Pacelli, G. Segre, Tra lavoro dipendente e lavoro parasubordinato: chi sono, da dove vengono e dove vanno i lavoratori parasubordinati, Settembre 2003

55. G. Betti, V. Verma, M. Natilli, F. Ballini, Statistical imputation in conjunction with micro-simulation of income data, Settembre 2003

54 V. Verma, G. Betti, F. Ballini, M. Natilli, S. Galgani, Personal income in the gross and net forms: applications of the Siena micro-simulation model (SM2), Settembre 2003

53. R. Canu, G. Tattara, Quando le farfalle mettono le ali. Osservazioni sull'ingresso delle donne nel lavoro dipendente, Settembre 2003

52. D. Favaro, S. Magrini, Gender wage differentials among young workers: methodological aspects and empirical results, Settembre 2003

51. F. Devicienti, Downward nominal wage rigidity in Italy: evidence and consequences, Novembre 2002

50. V. Verma, G. Betti, Longitudinal measures of income poverty and life-style deprivation, Novembre 2002.

49. A. Borgarello, F. Devicienti, Trends in the Italian earnings distribution, 1985-1996, Novembre 2002

48. S. Bragato, F. Occari, M. Valentini, I problemi di contabilizzazione dei lavoratori extracomunitari. Una verifica nelle province di Treviso e Vicenza, Novembre 2002

47. B. Anastasia, D. Maurizio, Misure dell'occupazione temporanea: consistenza, dinamica e caratteristiche di uno stock eterogeneo, Novembre 2002

46. E. Rettore, U. Trivellato, A. Martini, La valutazione delle politiche del lavoro in presenza di selezione: migliorare la teoria, i metodi o i dati?, Novembre 2002

45. U. Trivellato, A. Giraldo, Assessing the 'choosiness' of the job seekers. An exploratory approach and evidence for Italy, Novembre 2002

44. E. Battistin, E. Rettore, Another look at the regression discontinuity design, Novembre 2002

43. B. Contini, F. Cornaglia, C. Malpede, E. Rettore, Measuring the impact of the Italian CFL programme on the job opportunities for the youths, Novembre 2002

42. F. Devicienti, Estimating poverty persistence in Britain, Novembre 2001

41. A. Giraldo, E. Rettore, U. Trivellato, Attrition bias in the bank of Italy's survey on household income and wealth, Novembre 2001

40. A. Giraldo, E. Rettore, U. Trivellato, The persistence of poverty: true state dependence or unobserved heterogeneity? Some evidence form the Italian survey on household income and wealth, Novembre 2001

39. G. Barbieri, P. Gennari, P. Sestito, Do public employment services help people in finding a job? An evaluation of the italian case, Novembre 2001

38. N. Torelli, A. Paggiaro, Estimating transition models with misclassification, Novembre 2001

37. B. Anastasia, M. Gambuzza, M. Rasera, Le sorti dei flussi: dimensioni della domanda di lavoro, modalità di ingresso e rischio disoccupazione dei lavoratori extracomunitari in Veneto, Novembre 2001

36. L. Pacelli, Fixed term contracts, social security rebates and labour demand in Italy, Novembre

2001

35. A. Borgarello, F. Deficienti, C. Villosio, Mobilità retributiva in Italia 1985-1996, Novembre 2001

34. B. Contini, C. Villosio, Job changes and wage dynamics, Novembre 2001

33. B. Contini, Earnings mobility and labor market segmentation in Europe and USA: preliminary explorations, Novembre 2001

32 A. Borgarello, F. Devicienti, Trend nella distribuzione dei salari italiani 1985-1996, Novembre 2001

31. F. Bassi, U. Trivellato, Gross flows from the French labour force survey: a reanalysis, Novembre 2001

30. F. Bassi, M. Gambuzza, M. Rasera, Imprese e contratti di assunzione: prime analisi da Netlabor, Novembre 2001

29. N. Rosati, Further results on inequality in Italy in the 1980s and the 1990s, Aprile 2001

28. A. Ichino, M. Polo, E. Rettore, Are judges biased by labor market conditions?, Novembre 2000

27. E. Battistin, E. Rettore, Testing for the presence of a programme effect in a regression discontinuity design with non compliance, Novembre 2000

26. E. Battistin, E. Rettore, U. Trivellato, Measuring participation at work in the presence of fallible indicators of labour force state, Giugno 2000

25. A. D'Agostino, G. Ghellini, L. Neri, Percorsi di ingresso dei giovani nel mercato del lavoro, Giugno 2000

24. F. Bassi, M. Gambuzza, M. Rasera, E. Rettore, L'ingresso dei giovani nel mercato del lavoro: prime esplorazioni dall'archivio Netlabor, Giugno 2000

23. A. Paggiaro, U. Trivellato, Assessing the effects of the "Mobility List" programme in an Italian region: do (slightly) better data and more flexible models matter?, Marzo 2000

22. A. D'Agostino, G. Ghellini, L. Neri, Modelli statistici per l'analisi dei comportamenti di transizione scuola lavoro, Marzo 2000

21. G. Betti, B. Cheli, A. Lemmi, Analisi delle dinamiche di povertà e disoccupazione su uno pseudo panel italiano, Marzo 2000

20. N. Rosati, Permanent and Temporary Inequality in Italy in the 1980s and 1990s, Marzo 2000

19. F. Bassi, M. Gambizza, M. Rasera, Struttura e qualità delle informazioni del sistema NETLABOR. Una verifica sui dati delle Scica delle province di Belluno e Treviso, Marzo 2000

18. B. Anastasia, M. Gambuzza, M. Rasera, La durata dei rapporti di lavoro: evidenze da alcuni mercati locali del lavoro veneti, Marzo 2000

17. G. Betti, B. Cheli, A. Lemmi, Occupazione e condizioni di vita su uno pseudo panel italiano: primi risultati, avanzamenti e proposte metodologiche, Ottobre 1999

16. A. D'Agostino, G. Ghellini, L. Neri, A Multiple Imputation Method for School to Work Panel Data, Ottobre 1999

15. A. Paggiaro, N. Torelli, Una procedura per l'abbinamento di record nella rilevazione trimestrale delle forze di lavoro, Ottobre 1999

14. S. Campostrini, A. Giraldo, N. Parise, U. Trivellato, La misura della partecipazione al lavoro in Italia: presupposti e problemi metodologici di un approccio "time use", Ottobre 1999

13. T. Di Fonzo, P. Gennari, Le serie storiche delle forze di lavoro per il periodo 1984.1-92.3: prospettive e problemi di ricostruzione, Giugno 1999

12. A. Paggiaro, Un modello di mistura per l'analisi della disoccupazione di lunga durata, Maggio 1999

11. M. Lalla, F. Pattarin, Alcuni modelli per l'analisi delle durate complete e incomplete della disoccupazione: il caso Emilia Romagna, Maggio 1999

10 F. Bassi, M. Gambuzza, M. Rasera, Il sistema informatizzato NETLABOR. Caratteristiche di una nuova fonte sul mercato del lavoro, Maggio 1999

9. B. Contini, M. Filippi, L. Pacelli, C. Villosio, Working careers of skilled vs. unskilled workers, Gennaio 1999

8. B. Contini, L. Pacelli, C. Villosio, Short employment spell in Italy, Germany and Great Britain: testing the "Port-of-entry" hypothesis, Gennaio 1999

7. B. Contini, Wage structures in Europe and in the USA: are they rigid, are they flexible?, Gennaio 1999

6. M. Forni, S. Paba, Industrial districts, social environment and local growth. Evidence from Italy, Novembre 1998

5. Ginzburg, M. Scaltriti, G. Solinas, R. Zoboli, Un nuovo autunno caldo nel Mezzogiorno? Note in margine al dibattito sui differenziali salariali territoriali, Ottobre 1998

4. F. Bassi, Un modello per la stima di flussi nel mercato del lavoro affetti da errori di classificazione in rilevazioni retrospettive, Ottobre 1998

3. U. Trivellato, Il monitoraggio della povertà e della sua dinamica: questioni di misura e evidenze empiriche, Settembre 1998

2. N. Rosati, E. Rettore, G. Masarotto, A lower bound on asymptotic variance of repeated cross-sections estimators in fixed-effects models, Agosto 1998

1. E. Battistin, A. Gavosto, E. Rettore, Why do subsidized firms survive longer? An evaluation of a program promoting youth entrepreneurship in Italy, Agosto 1998