Lavoro: partecipazione, dinamica e valutazione di politiche

Misura, metodi, modelli

Progetto anno 2003

The latent class approach to estimating gross flows affected by correlated classification errors, with application to data from the French Labour Force Survey

F. Bassi, U. Trivellato

Dipartimento di Scienze Statistiche, Università di Padova

Working Paper n. 70, novembre 2005

Metodi e studi di valutazione degli effetti di politiche del lavoro, di aiuto alle imprese e di welfare

Cofinanziamento MIUR, anno 2003

Unità locali del progetto:

Dip. di Economia "S. Cognetti De Martiis", Univ. di Torino	(coord. B. Contini)
Dip. di Scienze Economiche, Univ. "Ca' Foscari" di Venezia	(coord. G. Tattara)
Dip. di Metodi Quantitativi, Univ. di Siena	(coord. A. Lemmi)
Dip. di Scienze Statistiche, Univ. di Padova	(coord. U. Trivellato)
Dip. di Scienze Economiche, Univ.di Salerno	(coord. S. Destefanis)
Dip. di Politiche Pubbliche e Scelte Collettive, Univ. del Piemonte Orientale	(coord. A. Martini)

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[♦]

Panel data allow the estimation of labour force gross flows – *i.e.*, transitions in time between different states, which are an important indicator in labour market analyses. Whereas net flows measure variations in time in the stock of employed and/or unemployed people, gross flows provide information on the dynamics of the labour market (see, *e.g.*, Atkinson & Micklewright, 1991).

Panel data may be obtained by means of various survey strategies, among which are genuine panel surveys – *i.e.*, surveys repeated on a number of occasions on the same units – and cross-section surveys with retrospective questions, or some combination of the two (see Duncan & Kalton, 1987, and Trivellato, 1999, among many others).

Measurement (≡classification) errors in the observed state can induce substantial bias in the estimation of gross flows, thus leading to erroneous conclusions about labour market dynamics (see Bound, Brown, & Mathiowetz, 2001, pp. 3792-3802, for a comprehensive survey).

A large body of literature on classification errors and their impact on gross flows estimation is based on the assumption that errors are uncorrelated over time: the so-called Independent Classification Errors (ICE) assumption. ICE implies that: (i) classification errors referring to two different occasions are independent of each other conditionally on the true states; and (ii) errors only depend on the present true state. According to that assumption, classification errors produce spurious transitions and consequently induce overestimation of changes (Abowd & Zellner, 1985; Kuha & Skinner, 1997). Many researchers proposed methods which adjust estimates of gross flows for measurement error based on the ICE assumption. Frequently, this assumption is coupled with the use of external information on misclassification rates, typically provided by re-interview data (see, *e.g.*, Abowd & Zellner, 1985; Chua & Fuller, 1987; Poterba & Summers, 1986).

However, for many labour force surveys no such re-interview data are available. Also, the ICE assumption is not realistic in many contexts, in which survey design and data collection strategies indicate that measurement errors are correlated over time (see, *e.g.*, Skinner & Torelli, 1993; Singh & Rao, 1995). This is especially relevant when panel data are collected by retrospective interrogation, because of the effects of memory inaccuracy (Sudman & Brandburn, 1973; Bernard *et al.*, 1984; O'Muircheartaigh, 1996; Tourangeau, Rips, & Rasinski, 2000). The main implication of correlated classification errors is that observed gross flows show lower mobility than true ones (van de Pol & Langeheine, 1997).

In this paper, we use a model-based approach to adjusting observed gross flows for correlated classification errors. It combines a sub-model for unobserved true transition rates and a measurement sub-model relating true states to observed states. A convenient framework for formulating our model is provided by latent class analysis.

We illustrate our approach by applying it to data on young people's observed gross flows among the usual three labour force states – Employed (*E*), Unemployed (*U*) and Out of the labour force (*O*) – taken from the French Labour Force Survey (FLFS),

[♦] ADDRESS FOR CORRESPONDENCE: francesca.bassi@unipd.it

March 1990-March 1992. The model is shown to correct flows in the expected direction: estimated true transition rates show higher mobility than observed ones. In addition, the measurement part of the model has significant coefficient estimates, and the estimated response probabilities show a clear, interpretable pattern.

Our approach provides a means of accounting for correlated classification errors across panel data which is less dependent on multiple indicators than previous formulations of latent class Markov models (see, *e.g.*, Bassi, 1997; van de Pol & Langeheine 1997; Bassi *et al.* 2000;). The application to the FLFS demonstrates that. As it will be seen, a peculiar feature of the FLFS data set is that it has multiple indicators – actually, two indicators – of labour force state at only one time point.

We compare our results with those obtained on the same data by Magnac and Visser (1999) (M&V from now on). M&V adopt a quite different approach, which aims at obtaining robust estimates of true transition rates while paying marginal attention to the (mis)measurement process. The main finding by M&V is that neglecting measurement error leads to an observed labour market which is more dynamic than the true one and, consequently, to underestimation of the average duration spent in labour force states. We compare the two approaches by discussing the reasonableness of the model specifications adopted and by contrasting the results.

The paper proceeds as follows. Section 2 presents our approach to the problem of modelling correlated classification errors using latent class analysis. In Section 3 the FLFS and sample data used are described, and evidence on the patterns of gross flows and classification errors is explored. Section 4 illustrates our approach by applying it to quarterly gross flows in the French labour market. In Section 5, our methodology is compared with that of M&V. Section 6 offers some concluding remarks.

2. Correlated classification errors and latent class modelling

Latent class analysis has already been applied in a number of studies on panel data to separate true changes from observed ones affected by unreliable measurements (recent contributions include van de Pol & Langeheine, 1997; Bassi, Torelli, & Trivellato, 1998; Bassi *et al*., 2000; Biemer & Bushery, 2000).

The true labour force state is treated as a latent variable, and the observed one as its indicator. The model consists of two parts:

(a) structural, which describes true dynamics among latent variables (*e.g.*, by means of Markov structures);

(b) measurement, which links each latent variable to its indicator(s). Some restrictions incorporating *a priori* information and/or assumptions on the error-generating mechanism are imposed on the parameters of the measurement part.

As a starting point, let us consider the simplest formulation of latent class Markov (LCM) models (Wiggins, 1973), which assumes that true unobservable transitions follow a first-order Markov chain. As in all standard latent class model specifications, local independence among the indicators is assumed, *i.e.*, indicators are independent conditionally on latent variables[1].

Let $X(t)$ denote the true labour force state at time t for a generic sample individual, $Y(t)$ the corresponding observed state, $v_{l_1}(1) = P(X(1) = l_1)$ the initial state of the latent Markov chain, and $\mu_{k_t\ l_{t+1}}(t) = P(X(t+1) = l_{t+1} \mid X(t) = k_t)$ the true transition probability between state k_t and state l_{t+1} from time t to t+1, with t=1,...,T-1, where T represents the total number of consecutive, equally spaced time-points over which an individual is observed. Besides, let $a_{l_t j_t}(t) = P(Y(t) = j_t \mid X(t) = l_t)$ be the probability of observing state j_t at time t, given that the true state is l_t.

It follows that $P(Y(1),...,Y(T))$ is the proportion of units observed in a generic cell of the T-way contingency table. For a generic sample individual, a first-order LCM model is defined as:

$$P(Y(1) = j_1,...,Y(T) = j_T) = \sum_{l_1=1}^{s} ... \sum_{l_T=1}^{s} v_{l_1}(1) a_{l_1 j_1}(1) \prod_{t=2}^{T} a_{l_t j_t}(t) \mu_{l_{t-1} l_t}(t-1), \qquad (2.1)$$

where j_t and l_t vary over the possible labour force states – in our case, E, U and O, with s=3.

In order to proceed in formulating our model, it is essential to consider the equivalence between latent class and log-linear models. Any latent class model may be expressed as a log-linear one with some unobservable variables (Haberman, 1979). A LCM model may also be specified in the log-linear context through the so-called "modified LISREL approach" proposed by Hagenaars (1990), which extends Goodman's modified path analysis (Goodman, 1973) to include latent variables[2]. Each conditional probability in equation (2.1) may be expressed as a function of the log-linear parameters. For example, $a_{l_1 j_1}(1)$ can be written as:

$$a_{l_1 j_1}(1) = \frac{\exp(\beta_{l_1}^{X(1)} + \beta_{l_1 j_1}^{X(1)Y(1)})}{\sum_{l_1=1}^{3} \exp(\beta_{l_1}^{X(1)} + \beta_{l_1 j_1}^{X(1)Y(1)})}, \qquad (2.2)$$

where $\beta_{l_1}^{X(1)}$ and $\beta_{l_1 j_1}^{X(1)Y(1)}$ denote the first- and second-order effects in log-linear parameterisation, respectively (according to the conventional notation for log-linear models: see, *e.g.*, Hagenaars, 1990).

---

[1] In the LCM with one indicator per latent variable, the assumption of local independence coincides with the ICE condition.

[2] For an extensive presentation of this approach to categorical causal modelling and for technical details, see Hagenaars (1998).

From equation (2.2), it is apparent that any restriction on conditional probabilities may be equivalently imposed on the log-linear parameters. In general, the specification of a model as the product of conditional probabilities has the advantage of more direct interpretation, whereas log-linear parameterisation is more flexible and allows more parsimonious models to be specified. The breakdown of conditional probabilities does imply estimation of the entire set of interaction parameters, while the modified LISREL approach allows us to specify models in which some higher-order interactions among variables may be omitted or conveniently constrained.

For our model, we exploit the above equivalence: the higher flexibility of log-linear parameterisation allows us to model the correlation of classification errors over time parsimoniously.

Our approach has the advantage that it does not require external information, either on misclassification probabilities from re-interview data or in the form of auxiliary variables affecting transition and/or classification probabilities (when auxiliary variables are available at sample unit level, observed heterogeneity – in the structural and/or the measurement processes – can be introduced by conditioning on them: see Pfeffermann *et al.*, 1998). It rests on sound formulation of the overall model: chiefly, on a sensible and parsimonious specification of the measurement process, allowing for classification errors which correlate over time. For this purpose, it is important to assess carefully the evidence provided by the data and to make use of the suggestions about patterns of measurement error, particularly in retrospective surveys, provided by a fairly large body of literature (see again Bound, Brown, & Mathiowetz, 2001, pp. 3743-3748).

As the model involves many unobservables, the issue of identification deserves attention. From the literature on latent Markov chains, it is well known that further information is needed in order to assure (global) identification, in the form of restrictions on the parameters and/or availability of multiple indicators of the true state. In a seminal paper, Lazarsfeld and Henry (1968) showed that, under the ICE assumption, a first order LCM model with one indicator per latent variable is identified if the Markov chain is stationary or response probabilities are restricted to be equal across time. The introduction of systematic classification errors may cause additional identification problems. Bassi (1997) demonstrates that a first order LCM model is identified without extra restrictions if two indicators per latent variable are available, even if errors are correlated. Identification criteria have been proven for such (and other) special cases, but no general results on latent class models identifiability have been provided yet. In the case of fairly complex models, it is advisable to check at least local identification, *i.e.*, identifiability of the unknown parameters in the neighbourhood of the maximum likelihood solution. A sufficient condition for local identifiability of a latent class model is that the information matrix be full rank (Goodman, 1974). In practice, one has to work with the estimated information matrix: for a model to be identifiable, all its eigenvalues must be positive.

There are also some potential disadvantages to our approach. The number of parameters to be estimated may be high, and estimation computation intensive. This indicates that the size of the model should be kept reasonably moderate, first of all in terms of dimensions of the state and time space (*i.e.*, of the number of labour force states and the number of time units, respectively). Furthermore, since the approach involves a model jointly for unobservable true states and classification errors, the results are sensitive to possible model misspecification. In addition to the observations made above about the attention to be given to a well-founded specification, it is important to check empirical model validity. Two lines of attack are useful. First, one may look at diagnostics on the components of the model, to ascertain if they are in line with expectations: parameter estimates have the expected signs, they are statistically significant, *etc*.. Second, the validity of the overall model may be tested, comparing the estimated expected frequencies with the observed ones in the complete table by means of the log-likelihood ratio L^2 and the Pearson X^2 statistics. However, the typical pattern of labour force transitions results in a sparse and unbalanced contingency table. Thus, the usual X^2 and L^2 criteria must be used only as a general indication of fit, since their asymptotic χ^2 distribution is no longer guaranteed. One way to improve model evaluation is to specify a restricted model nested within a larger one: it can be tested with the conditional test, *i.e.*, considering the difference in the L^2 values of the two models, which is asymptotically distributed as χ^2 under weaker conditions (Hagenaars, 1990)[3].

3. The data and preliminary evidence from them

3.1 The French Labour Force Survey

The FLFS, *Enquête Emploi*, is conducted yearly by INSEE, the French national statistical agency. Its reference population is all members of French households aged 15 years or more in the calendar year in which the survey is carried out. It is of rotating panel type, one-third of the sample being replaced each year.

Information on labour force participation is collected by means of two sets of questions: retrospective interrogation on a reference period composed of the 12 months preceding the interview month, and a question on actual state during the interview month[4]. Respondents are asked to report their monthly labour condition by filling in a grid in which they classify themselves in one of the following eight categories: self-employed, fixed-term employed, permanently employed, unemployed, on a vocational training program, student, doing military service or other (retired, housewife, *etc.*).

We use the information collected in the surveys of March 1991 and March 1992 on a sub-group of the cohort sampled in 1990, made up of young people. The sample consists of 5,247 individuals, who were aged between 19 and 29 years in March 1992. At each wave, respondents were asked to report their monthly labour force history from the corresponding month of the previous year to the current month. Thus, there is only one

[3] This procedure can be extended to testing a sequence of nested models against a larger, maintained one. In order to compare alternative non-nested models, it can be useful to resort to indices based on the information criterion, such as AIC or BIC, which weight the goodness of fit of a model against its parsimony, considering the model degrees of freedom and the sample size.
[4] Our brief description of the survey is restricted to the panel data set dealt with here.

(retrospectively) observed labour force state at each month, except for March 1991, where two distinct pieces of information are available: one is the concurrent information collected with the March 1991 survey; the other is the retrospective information collected with the March 1992 survey.

Our data-set does not include information on auxiliary variables. Thus, we will specify a model for an homogeneous population. In this respect, it is worth noting that our reference population consists of a fairly restricted age group – 19-29 years. This makes the assumption tenable.

In order to keep the model at a manageable size, we carry out our analyses according to a simplified formulation of state and time space[5]. We consider a state space restricted to three labour force states – Employed (*E*), Unemployed (*U*) and Out of the labour force (*O*) – and analyse quarterly transitions from March 1990 to March 1992. Obviously, it is important to ascertain that this choice does not appreciably alter the pattern of observed transitions[6].

3.2 Descriptive Evidence on Gross Flows and the Pattern of Classification Errors

First, we consider descriptive evidence on observed gross flows and the pattern of measurement error.

Table 1 presents the observed quarterly transition rates for our sample of young people. Let us disregard, for the moment, the distinction of transitions by type, and just look at the transitions of the same type – those denoted by WW – over the two years. Observed transition rates exhibit a moderate, but neat seasonal pattern, presumably related to the school calendar. From June to September, for example, we observe a proportion of people who enter the labour market (*OE* and *OU* rates) greater than the average; instead, a peak of exit rates from employment, chiefly towards inactivity (*EO*), is documented from September to December[7].

As regards the patterns of survey responses and classification error, interesting evidence comes from Table 2 and, again, Table 1. Table 2 exploits the double information on the labour force state in March 1991, and presents a cross-classification of people resulting from the concurrent and retrospective survey responses. It provides some crude evidence on response error: 7.8% of respondents declare a different state in the two surveys. While this is by no means conclusive evidence that the actual reported information is error-free and that retrospective information is contaminated by classification errors, it clearly points to the role of memory effects in respondents' accuracy. Inspection of the percent row distributions also indicates that response errors

[5] The dimensions of the observed contingency table grow dramatically when many polytomous variables are considered simultaneously, and may become very demanding for estimation algorithms.

[6] Results supporting this proposition are available form the authors on request.

[7] The seasonal pattern is more apparent from observed monthly flows, with higher polarisation of the transition rates from June to July and from August to September, respectively. Observed monthly flows are available from the authors on request.

vary according to the true state (under the mild assumption that it is measured with less imprecision by the concurrent survey response).

Table 1: *Observed Quarterly Transition Rates (%) by Type, March 1990 - March 1992**

Quarterly transitions **		***EE***	***EU***	***EO***	***UE***	***UU***	***UO***	***OE***	***OU***	***OO***
March90 – June90	WW	93.03	2.03	1.34	19.94	77.46	2.60	1.37	0.32	98.40
June90 – Sept.90	WW	94.08	4.32	1.60	18.99	79.43	1.58	3.87	3.79	93.34
Sept.90 – Dec.90	WW	93.93	4.39	1.98	24.00	72.47	3.53	1.91	0.80	97.30
Dec.90 – March91	WW	94.77	4.25	0.98	24.53	72.40	3.07	0.98	0.66	98.36
	BW	91.50	4.86	3.64	31.60	56.84	11.56	4.40	2.10	93.50
March91- June91	WW	96.03	3.02	0.95	23.21	74.32	2.47	1.28	0.68	98.04
	BW	91.48	4.63	3.89	35.01	54.20	10.79	4.84	2.14	93.02
June91 – Sept.91	WW	94.29	3.94	1.77	20.93	78.29	0.78	4.71	2.95	92.34
Sept.91 – Dec.91	WW	93.73	4.48	1.79	23.63	74.89	1.48	3.22	1.65	95.13
Dec.91- March92	WW	93.90	4.80	1.30	21.67	76.74	1.59	1.70	0.59	97.71

* States: *E* = employment; *U* = unemployment; *O* = out of the labour force.
** Transitions: WW = within-wave; BW = between-waves.

Table 2: *Compatibility Between Declarations of March 991 and March 1992 (absolute figures; % overall in italics; % within row in brackets)**

	March 1992		
	E	***U***	***O***
E	2,026 *37.3 (91.3)*	89 *1.7 (4.0)*	103 *1.9 (4.7)*
U	65 *1.2 (16.0)*	287 *5.3 (70.9)*	53 *1.0 (13.1)*
O	69 *1.3 (2.4)*	41 *0.7 (1.5)*	2,694 *49.6 (96.1)*

* States: *E* = employment; *U* = unemployment; *O* = out of the labour force.

Returning to Table 1, let us focus on the December 1990-March 1991 and March-June 1991 transitions. Two types of quarterly flows are observed: (i) within-wave (WW), when information about labour force states is collected in the same survey, and (ii) between-waves (BW), when information is collected in two different surveys. The differences between WW and BW transition rates appear to be substantial, and are consistently stable across the two quarters. WW transition rates describe a remarkably more stable labour market than BW ones.

It is worth noting that interpretation of this evidence, with respect to the pattern of classification errors, is not straightforward – as it would have been if BW transitions had resulted from a combination of retrospective and concurrent information respectively, collected in two subsequent survey waves, and if WW transitions had resulted from retrospective information collected within the same survey wave and extending backwards roughly for the same period time (see, *e.g.*, Martini, 1988). Given the design of the FLFS, the picture is less clear. Here:

(a) WW transition rates result: (a1) for December 1990-March 1991 from the survey wave of March 1991, *i.e.***,** from a combination of concurrent (for March) and three-month retrospective information (for December); (a2) for March-June 1991 from retrospective information quite a way back in time (twelve and nine months respectively), collected from the survey wave of March 1992.

(b) Instead, BW transition rates are estimated on the basis of information collected at the two survey waves of March 1991 (for the initial month) and March 1992 (for the final month), that is to say: (b1) for December-March, from a combination of three-month retrospective and twelve-month retrospective information; (b2) for March-June, from a combination of concurrent and nine-month retrospective information.

Nevertheless, the overall evidence provides a clear indication that the bias towards stability documented by WW transition rates is caused by survey responses affected by highly correlated classification errors.

In panel surveys with retrospective questions, memory decay is probably the main cause of response error, hence classification error. Among the most common effects of memory decay is the tendency for respondents to forget past events and/or to place them wrongly along the time axis (Sudman & Bradburn, 1973). Abundant evidence from the cognitive psychology and survey methodology literature indicates that the quality of recall declines as the length of the recall period increases, although this relationship is far from being general and stable[8]. The literature on measurement errors in reporting labour market histories documents that short spells are often forgotten and that events (*i.e.*, changes of state) are anticipated or postponed on the time axis towards the boundaries of the reference period (see, for example, Martini, 1988; Ryscavage & Martini, 1990)[9]. More specifically, there is a respondents' tendency to shift reported changes in labour force towards the interview time – the so-called "forward telescoping effect" – and/or to mechanically report the same condition throughout the whole reference period – the so-called "conditioning effect" – (Eisenhower & Mathiowetz, 1991). The overall result of these effects is to induce correlated classification errors, the magnitude of which increases as the recall period extends.

---

[8] The recall period interacts with other factors, such as salience of the event, type of response task, *etc.* (for broad reviews, see Jabine *et al.*, 1984; Biemer & Trewin 1997; Tourangeau, Rips, & Rasinski 2000; Bound, Brown, & Mathiowetz, 2001, pp. 3743-3748; for meta-analysis questioning the importance of the recall period on data quality, see Mathiowetz, 2000).

[9] Proper transitions – *i.e.*, movements from one state to a different one – may easily be wrongly placed in time, because in some cases events really may be difficult to place along the time axis. For example, employees who lost their jobs or retired (transitions *EU* and *EO*) generally take the paid holidays they are entitled to, and may not recall exactly when they left employment. The moment individuals entered the labour force (transitions *OU* and *OE*) may also be hard to recall, especially when they left school (van de Pol & Langeheine, 1997).

4. A model for correlated classification errors in retrospective surveys

4.1 A Latent Class Model for Quarterly Labour Gross Flows of Young People from the FLFS

Our purpose is to specify and estimate a model which (i) allows for a seasonal component in true transitions and, based on the empirical findings above and suggestions from the literature on response errors in surveys, (ii) parsimoniously describes the pattern of correlated errors.

Figure 1 presents the path diagram of a simplified specification of our model[10]. Here and in the sequel, indicators *Y(5)*, *Y(6)*, *Y(7)*, *Y(8)* and *Y(9)* represent observed states in the five quarters covered by the March 1992 survey (March, June, September and December 1991, and March 1992, respectively); *W(1)*, *W(2)*, *W(3)*, *W(4)* and *W(5)* refer to the sequence of states observed in the preceding survey (March, June, September and December 1990, and March 1991, respectively). As usual, $X(t)$ denotes the true labour force state at time *t*. It is apparent that, for eight quarters out of nine, there is just one indicator for each latent state: only for *X(5)* are there two indicators, *Y(5)* and *W(5)*.

Figure 1: Basic LCM Model of Gross Flows for 9 time-points

W(1)   W(2)   W(3)   W(4)   W(5)
↑ ↖ ↑ ↖ ↑ ↖ ↑ ↖ ↑
X(1)→X(2)→X(3)→X(4)→X(5)→X(6)→X(7)→X(8)→X(9)
↓ ↙ ↓ ↙ ↓ ↙ ↓ ↙ ↓
Y(5)   Y(6)   Y(7)   Y(8)   Y(9)

Figure 1 clearly distinguishes the first-order Markov process among true states *X*(*t*) and the measurement sub-model linking indicators *Y*(*t*) and *W*(*t*) to latent states and transitions. Oblique arrows are meant to indicate that responses for time *t* depend on the true transition between *t* and *t*+1; this specification overcomes the traditional ICE assumption (van de Pol & Langeheine, 1997).

The relationships described by Figure 1 may be formulated as in equation (4.1), which breaks down the observed proportion in the generic cell of the 10-way contingency table into the following product of conditional probabilities:

$$P(W(1)=i_1, W(2)=i_2, W(3)=i_3, W(4)=i_4, W(5)=i_5, Y(5)=j_5, Y(6)=j_6, Y(7)=j_7, Y(8)=j_8 Y(9)=j_9)=$$
$$=\sum_{l_1,\ldots l_9=1}^{3} v_{l_1}(1)\prod_{t=1}^{8}\mu_{l_t l_{t+1}}(t)a_{l_9 j_9}(9)a_{l_8 l_9 j_8}(8)a_{l_7 l_6 j_7}(7)a_{l_6 l_7 j_6}(6)a_{l_5 l_6 j_5}(5)z_{l_5 i_5}(5)z_{l_4 l_5 i_4}(4)z_{l_3 l_4 i_3}(3)z_{l_2 l_3 i_2}(2)z_{l_1 l_2 i_1}(1) \quad , \tag{4.1}$$

[10] Figure 1 does not correspond to our final model, which is given immediately in the sequel and in Appendix A. It serves as a stylised version of it, useful to highlight some of its basic features.

where $z_{l_5 i_5}(5) = P(W(5) = i_5 \mid X(5) = l_5)$ is the probability of observing state i_5 for March 1992 (*W(5)*), given that the true state was l_5; $z_{l_1 l_2 i_1}(1) = P(W(1) = i_1 \mid X(1) = l_1, X(2) = l_2)$ is the probability of observing state i_1 for March 1990 (*W(1)*), given that there was a transition from state l_1 to state l_2 from March 1990 (*X(1)*) to June 1990 (*X(2)*), and similarly for the other conditional response probabilities – those denoted by $z_{l_t l_{t+1} i_t}(t) = P(W(t) = i_t \mid X(t) = l_t, X(t+1) = l_{t+1})$, *t*=1,…4, referring to states observed during the March 1991 interview and those denoted by $a_{l_t l_{t+1} i_t}(t) = P(Y(t) = j_t \mid X(t) = l_t, X(t+1) = l_{t+1})$, *t*=5,…8, referring to states observed during the March 1992 interview.

Equation (4.1) may be formulated also in terms of a system of multinomial logit equations. Note, however, that in such parameterisation more parsimonious models may be specified (*e.g.*, by imposing that the hierarchical log-linear model does not contain third-order interaction parameters, which are implied by equation (4.1)).

It is important to stress again that we have no supplementary information on the measurement process, except for the double observation on labour force state in March 1991. Thus, identification is not assured unless plausible restrictions are imposed on the parameters. In particular, some assumptions are needed to identify the measurement error mechanism. The final specification of our model results from combining various pieces of information: knowledge of the design and measurement characteristics of the FLFS; theoretical considerations and empirical evidence about the pattern of reporting errors in retrospective surveys, as well as on the true dynamic process, reviewed in section 3.2; results from specification searches aimed at obtaining a sensible and parsimonious formulation.

The set of assumptions of the final model, together with some further comments on the reasons motivating them, may be summarised as follows (for mathematical formulation, see Appendix A).

(a) Quarterly flows among true labour force states follow a first-order non-homogeneous Markov chain, with transition probabilities among the same calendar quarters in different years set to be equal. This specification amounts to assuming a dominant constant seasonal component for labour market dynamics – a pattern convincingly suggested by previous analyses.

(b) As regards the measurement part of the model, it is worth emphasizing again that only for March 1991 are two distinct observations of the labour force state available – concurrent information *W(5)* and retrospective information *Y(5)*. Thus, we cannot explicitly model dependencies between indicators[11]. One way of accounting for correlated classification errors in our data is to let the indicators depend on latent transitions. That is, we assume that a response given for time *t* depends on the true transition which occurred between times *t* and *t*+1 (this is the meaning of the

[11] With essentially one indicator per latent state, a model postulating direct effects between indicators would be trivially under-identified.

oblique arrows in Figure 1 and of probabilities $z_{l_t l_{t+1} i_t}(t)$ and $a_{l_t l_{t+1} i_t}(t)$; for precise specification, see Appendix A). In other words, a sort of forward telescoping effect is postulated, which is sensible in retrospective surveys.

(c) The literature on memory effects on respondent behaviour suggests that classification errors correlation occurs also because response errors depend on the length of recall period (see section 3.2 above). Thus, we impose some additional assumptions on response probabilities, as follows[12].

For second-order interaction parameters $\beta_{l_t i_t}^{X(t)W(t)}$, t=1,…,5, and $\beta_{l_t j_t}^{X(t)Y(t)}$, t=5,…,9, describing the association between each latent state and its indicator, we start by assuming a flexible functional form, so that the probability of erroneously reporting the labour force state increases with the distance between reference and survey months. The specification we move from is:

$$\begin{aligned} \beta_{l_t i_t}^{X(t)W(t)} &= \omega_{l_t i_t}^{X(t)W(t)} + \delta(\omega_{l_t}^{X(t)} f(\Delta t)) \\ \beta_{l_t j_t}^{X(t)Y(t)} &= \omega_{l_t j_t}^{X(t)Y(t)} + \delta(\omega_{l_t}^{X(t)} f(\Delta t)) \end{aligned} , \tag{4.2}$$

where $\omega_{l_t i_t}^{X(t)W(t)}$ and $\omega_{l_t j_t}^{X(t)Y(t)}$ are parameters measuring the association between each latent state $X(t)$ and its indicator, which depend on the combination between observed and true states; δ is an indicator function having a value of 1 if the true state is correctly reported, and 0 otherwise; $\omega_{l_t}^{X(t)}$ are proportionality factors depending on the true state − they account for the fact that the three states *E, U* and *O* may be perceived differently by respondents; $f(.)$ is an increasing monotone function of time distance Δt between reference and survey months.

Then, we perform a specification search within a set of possible functions (linear, squared, exponential). We end up by choosing $f(\Delta t)=exp(\Delta t)$ as the best-fit specification, on the basis of L^2 [13].

(d) Lastly, we set up the following further restrictions on parameters of the measurement sub-model, motivated both by model parsimony and by evidences on the error generating mechanism.

(d1) Parameters $\omega_{l_t i_t}^{X(t)W(t)}$, $\omega_{l_t j_t}^{X(t)Y(t)}$, and $\omega_{l_t}^{X(t)}$ are set constant over time, as well as the association between *W(t)* and *X(t*+1*)* for *t*=1,...,4, and between *Y(t)* and *X(t*+1*)* for *t*=5,…,8. These restrictions are consistent with the notion that the measurement properties of properly designed survey instruments are fairly stable over time − a result well established in the literature.

[12] In a previous, rather crude version of the model, correlation among classification errors was accounted for by only letting response probabilities depend on latent transitions (Bassi, Torelli, & Trivellato 1998, pp. 116-119).

[13] This choice is in accordance with findings from experimental psychology literature and social research, which indicate that, within short or medium recall periods, the process of memory decay is approximated reasonably well by an exponential function (see, *e.g.*, Sudman & Bradburn, 1973).

(d2) The probability of making mistakes is assumed to be constant for the same month across the various years. The rationale for this assumption again has to do with the concept of the time stability of the measurement properties of the survey instruments, but in a slightly different, more specific sense. It is based on the fact that the survey waves were carried out in the same month of consecutive years – March 1991 and March 1992 – and on empirical evidence that response errors mainly depend on the period of time elapsing between survey time and the event to be recalled, rather than on the calendar month in which the event took place.
(d3) In the hierarchical log-linear model formulation, all third order interaction parameters are excluded.

4.2 Results

The model has been estimated by maximum likelihood. It is locally identified: estimated information matrix eigenvalues are all positive. Estimated quarterly transition rates from our model are listed in Table 3 and estimated conditional response probabilities in Table 4[14]. L^2 is 3,612.12 (with associated p-value = 1)[15].

Let us look, first, at the structural part of the model. When estimated transition rates are compared with the corresponding observed rates in Table 1, it emerges that observed transitions are corrected according to expectations. As implied by serially correlated measurement errors in retrospective surveys, true mobility in the French youth labour market is higher than observed mobility.

As regards the measurement part of our model, further results are given in Appendix B, which reports its estimated log-linear parameters. It is worth noting that all of them are statistically significant. Specifically, parameters $\omega_{l_t}^{X(t)}$ for t=1,…,9, $\omega_{l_t j_t}^{X(t)Y(t)}$ for t=1,…,5, and $\omega_{l_t j_t}^{X(t)Y(t)}$ for t=5,…,9, which specify classification errors as a function of time, are significantly different from zero. This evidence neatly corroborates the hypothesis of correlated classification errors.

---

[14] Software *lEM* (Vermunt, 1996) was used. We carried out the estimation procedure using a few different sets of starting values. No estimation problems, such as local maxima or parameter estimates outside the admissible space, were encountered.

[15] As noted in section 2, one should be very cautious in interpreting a conventional L^2 computed from a contingency table of gross flows, because many cells show quite low observed frequencies. On the other hand, as our model is fairly complex, it is problematic to assess its validity by means of a conditional test. Partial, but convincing indications of the plausibility of our model come from some conditional testing of nested models, in the context of comparisons with the model by M&V: simplified versions of our model – relative to its structural and measurement part, respectively – play the role of the maintained hypothesis, H_1; restricted specifications are tested against it and are clearly rejected. (Results are available from the authors on request.)

Table 3: *LCM Model: Estimated Quarterly Transition Rates (%), March 1990 - March 1992**

Quarterly transitions	***EE***	***EU***	***EO***	***UE***	***UU***	***UO***	***OE***	***OU***	***OO***
March – June	92.76	3.97	3.27	32.60	58.15	9.25	3.39	1.39	95.22
June – Sept.	92.83	5.37	1.80	28.83	70.07	1.09	4.63	2.95	92.42
Sept. – Dec.	93.56	4.53	1.91	24.32	73.10	2.57	2.58	1.21	96.23
Dec. – March	94.31	4.55	1.15	22.98	74.76	2.27	1.32	0.63	98.05

* States: *E* = employment; U = unemployment; *O* = out of the labour force.

Table 4: *LCM Model: Estimated Conditional Response Probabilities (%), March 1990 - March 1992**

Reference months	**Conditional response probabilities**								
	E/e	***U/e***	***O/e***	***E/u***	***U/u***	***O/u***	***E/o***	***U/o***	***O/o***
March 1990 and 1991 **	95.60	2.06	2.34	17.09	72.42	10.49	2.13	0.87	97.00
June 1990 and 1991 **	97.89	1.34	0.77	7.56	90.48	1.96	1.51	0.77	97.72
Sept.1990 and 1991 **	99.90	0.06	0.14	0.34	99.48	0.18	0.05	0.02	99.93
Dec. 1990 and 1991 **	100.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	100.00
March 1991 and 1992 ***	100.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	100.00

* Observed states: *E* = employment; *U* = unemployment; *O* = out of the labour force.
True states: *e* = employment; *u* = unemployment; *o* = out of the labour force.
** Retrospective information collected in March 1991 and March 1992, respectively.
*** Concurrent information collected in March 1991 and March 1992, respectively.

Table 4 reveals various interesting results. First, no appreciable errors are made in reporting the concurrent condition (Table 4, last row). Second, classification errors are unnoticeable also for the quarter immediately preceding the survey (Table 4, penultimate row)[16]. A comment is in order, to properly read that result. The opportunity of detecting classification errors depends on the information available about the measurement process. As repeatedly noted, in our case-study such information is quite limited. Thus, the evidence of no appreciable response errors does not entail that concurrent and one-quarter retrospective responses are intrinsically error-free. Rather, it means that, conditional on available information, potential classification errors in concurrent and one-quarter retrospective responses do not emerge (though the parameters of the measurement sub-model are all statistically significant).

As for the other reference months – for which retrospective reporting extends backwards for more than one quarter, as implied by (4.2) the estimated conditional response probabilities show that the longer the recall period, the greater the probability of answering incorrectly. For labour force states involving recall periods of up to one year (Table 4, first row), the size of classification errors is substantial – especially for the case of response probabilities conditional to true unemployment.

[16] Conditional response probabilities are obtained by inserting estimated parameters in the relevant formulas for $z_{l_t l_{t+1} i_t}(t)$ and $a_{l_t l_{t+1} i_t}(t)$, given in Appendix A.

We obtain further insights on some features of the latent Markov model by jointly considering the parameter estimates of the measurement and structural parts of the model. Since reported states for December and March are not affected by discernible measurement errors, it comes as no surprise that the December-March estimated transitions basically coincide with observed ones, apart from the averaging induced by the assumption that they do not vary over the two years. In addition, the March-June estimated transitions largely reflect the heavy weight assigned to the information in March 1991 in the BW observed transitions – concurrent, which is thus taken as practically error-free. For the other transitions, the general pattern outlined above is clear: there is more true dynamics than appear from the observed rates; the more the reference month extends back in time with respect to the survey month, the higher the correction towards mobility; transitions exhibit a definite seasonal pattern.

5. A comparison with Magnac and Visser's model

In this section we compare our results with those obtained by Magnac and Visser (1999), who basically analysed the same data but took a quite different approach and obtained remarkably different results. First, we summarise some differences in the data-sets used and briefly outline their modelling strategy.

M&V consider monthly transitions observed in the FLFS over the period from January 1989 to March 1992, on the same sample of 5,427 individuals aged between 19 to 29 years in March 1992. Thus, with respect to our case-study, they extend the observation period by considering information collected in three consecutive waves of the survey: January 1990, March 1991, and March 1992. Moreover, they distinguish six labour force states: (1) permanent employment, (2) fixed-term employment, (3) training, (4) unemployment, (5) education and (6) out of the labour force[17].

M&V base their model specification and estimation strategy on the following assumptions.

(a) Latent (≡true) transitions between labour force states follow a time-homogeneous first-order Markov process.
(b) The actual reported state is a perfectly reliable source of information, whereas retrospectively reported states are contaminated by classification errors, with error probabilities fixed *a priori*[18].
(c) (Mis)classifications recorded at time t and $t+d$ are conditionally independent: this is what M&V call the "d-step ICE" assumption.

[17] Correspondence with the classification we used is straightforward: E comprises states 1, 2 and 3; U consists of state 4; O comprises states 5 and 6.

[18] M&V consider state-dependent error probabilities, and get the matrix of error probabilities of one-year retrospective responses by comparing retrospectively reported states with actual reported states, taken as error-free. Besides, they assume that, within the one year recall period, the matrix of error probabilities increases linearly with time.

Let d be an integer variable assuming values between 1 and T–1; $\lambda_{ij}^{d}(t)$ the probability that at time $t+d$ the observed state is j, given that at time t the observed state was i; $b_{ik}(t)$ the probability that the true state is k at time t, given that the observed state is i (while the notation introduced in section 2 is otherwise maintained). For t=1,2,…, T–d, the d-step ICE assumption implies that:

$$\lambda_{ij}^{(d)}(t) = P(Y(t+d) = j \mid Y(t) = i) = \sum_{k,l=1}^{6} b_{ik}(t)\mu_{kl}^{(d)}(t)a_{lj}(t+d)\,. \tag{5.1}$$

In short, the rationale for the modelling strategy of M&V appears to be to obtain robust estimates of true transition rates, irrespective of possibly correlated classification errors – this is basically the purpose of the d-ICE assumption.

The general picture emerging from the models estimated by M&V is that neglecting measurement error leads to an overestimation of mobility in the labour market. In addition, the estimates of some transition probabilities turn out to be severely biased by classification errors.

Evidence from the data and arguments from the literature on measurement errors in panel data cast some doubts on M&V's assumptions, and also on their results.

First, it should be noted that the observed flows exhibit an apparent seasonal pattern (see Table 1 for quarterly rates). This calls into question the assumption of a time-homogeneous Markov process.

Second, the evidence of response error in the data (Tables 1 and 2) does not provide support to M&V's specification of the (mis)measurement process. It rests crucially on the d-step ICE assumption, which was introduced in order to attenuate the inconveniences caused by the ICE assumption and was explained in the following terms: "It only requires that (mis)classifications are conditionally independent in a d-unit period" (M&V, p. 466). The motivation for the d-ICE assumption is to obtain results robust against correlated classification errors, while paying a price in terms of efficiency – only a limited fraction of the available sample information is used. Nevertheless, we do not see a convincing rationale for it.

The fact that, in panel surveys with retrospective questions, memory decay is probably the main cause of response errors is certainly taken into account. But this is done only in part and with a rigid format. M&V do this by combining the d-ICE condition with strong assumptions about response error probabilities. Two points deserve attention and, in our view, are far from persuasive.

(a) Error probabilities are exogenously determined, on the basis of an entirely *a priori* specification.
(b) The combination of the two sets of assumptions – d-ICE and those on response error probabilities – supplies a pattern of classification errors which does not capture the component of correlation over time.

Thus, it comes as no surprise to find that the final model estimated by M&V corrects labour market dynamics towards stability, like all strategies using the classical ICE assumption.

For a closer comparative assessment with respect to our model, we also estimated one of the M&V's final models on the period March 1990-March 1992. That is, we took from them this specification: true monthly flows following a first-order time-homogeneous Markov chain, error probabilities are fixed, and d=18.

Table 5: *Estimated Quarterly Transition Rates (%) Implied by M&V's Model, March 1990 - March 1992**

EE	***EU***	***EO***	***UE***	***UU***	***UO***	***OE***	***OU***	***OO***
97.32	2.28	0.40	13.21	86.20	0.59	2.65	1.54	95.81

* Model specification: monthly flows, time-homogeneous Markov chain, d=18.
States: E = employment; U = unemployment; O = out of the labour force.

Results are listed in Table 5, which presents the quarterly transition rates implied by the model above. They describe a labour market which is much more stable than the observed one, and *a fortiori* than that emerging from our analyses. As regards the measurement sub-model, it is interesting to note that M&V's assumption of concurrent error-free information is not confuted by our estimates (see Table 4, last row). However, their d-ICE assumption and our specification of the correlated measurement error mechanism are quite divergent, and play a crucial role in producing the substantially different results just mentioned.

6. Concluding remarks

This paper presents a latent class approach to correct gross flows from correlated classification errors. The emphasis is on the capacity of this approach to account for correlated classification errors across panel data without heavily depending on multiple indicators.

A case-study serves to illustrate our modelling strategy. A model is formulated and estimated within a latent class analysis set-up, to adjust observed quarterly labour force gross flows of French young people, March 1990-March 1992. The data were taken from the French Labour Survey, which collects information on labour force conditions mainly by means of retrospective interrogation. Based on arguments from survey methodology literature and on empirical evidence from the same French data-set, the model accounts for correlated classification errors over time and allows for a seasonal component in the latent Markov process.

Our model corrects observed flows in the expected direction: estimated true transition rates show higher mobility than observed ones. Besides, estimated conditional response probabilities show a neat, sensible pattern: there is no discernible classification error for concurrent information, whereas the size of classification errors is considerable in retrospectively reported labour force states for reference months far from the interview month.

Our model was compared with the rather different one for the same data developed by Magnac and Visser (1999). The approach of M&V essentially aims at obtaining robust estimates of the transition rates, while paying marginal attention to the measurement process. Interestingly enough, our results differ substantially from those of M&V. While their model corrects observed flows towards stability, our model does just the opposite. Our results are as expected in the case of panel data collected (mainly) by means of retrospective questions. Indeed, these data are typically affected by correlated classification errors, induced by the combined effect of memory inaccuracy and interrogation strategy.

The implications of these findings for the analysis of labour market dynamics are far from negligible.

Acknowledgements

Research for this paper was supported by a grant from the Italian Ministry for Education, University and Scientific Research to the project "Metodi e studi di valutazione degli effetti di politiche del lavoro, di aiuto alle imprese e di *welfare*". We are grateful to Michael Visser and Thierry Magnac for providing us with anonymous micro-data from the French Labour Survey and with the computer programmes for estimating their models. A preliminary version of this paper was presented at SMABS2004 (Conference of the Society of Multivariate Analysis for the Behavioural Sciences), Jena, 17-22 July 2004, at which we benefited from useful comments. We thank Enrico Rettore for valuable suggestions. The usual disclaimer applies.

References

Abowd, J.M., & Zellner, A. (1985). Estimating gross labor force flows. *Journal of Business & Economic Statistics, 3*, 254-283.

Atkinson, A.B., & Mickelwright, J. (1991). Unemployment compensation and labour market transitions: a critical review. *Journal of Economic Literature, 29*, 1679-1727.

Bassi, F. (1997). Identification of latent class Markov models with multiple indicators and correlated measurement error. *Journal of the Italian Statistical Society, 3,* 201-211.

Bassi, F., Croon, M., Hagenaars, J.A, & Vermunt, J.K. (2000). Estimating true changes when categorical panel data are affected by correlated and uncorrelated classification errors. An application to unemployment data. *Sociological Methods and Research, 29*, 230-268.

Bassi, F., Torelli, N., & Trivellato, U. (1998). Data and modelling strategies in estimating labour force gross flows affected by classification errors. *Survey Methodology, 24*, 109-122.

Bernard, H.R., Killworth, P., Krinenfeld, D., & Sailer, L. (1984). The problem of informant accuracy: the validity of retrospective data. *Annual Review of Anthropology, 13*, 495-517.

Biemer, P.P., & Bushery, J.M. (2000). On the validity of latent class analysis for estimating classification error in labor force data. *Survey Methodology, 2*, 139-152.

Biemer, P., & Trewin, D. (1997). A review of measurement error effects on the analysis of survey data. In L. Lyberg *et al.* (Eds.), *Survey measurement and process quality* (pp. 603-631). New York: Wiley.

Bound, J., Brown, C., & Mathiowetz, N.A. (2001). Measurement error in survey data. In J.J. Heckman & E. Leamer (Eds.), *Handbook of econometrics Vol.5* (pp. 3705-3843). Amsterdam: Elsevier.

Chua, T.C., & Fuller, W.A. (1987). A model for multinomial response error applied to labor flows. *Journal of the American Statistical Association, 82*, 46-51.

Duncan, G.J., & Kalton, G. (1987). Issues of design and analysis of surveys across time. *International Statistical Review, 55*, 97-117.

Eisenhower, D., & Mathiowetz, N.A. (1991). Recall errors: sources and bias reduction techniques. In P. Biemer *et al*. (Eds.), *Measurement errors in surveys* (pp. 127-144). New York: Wiley.

Goodman, L. (1973). The analysis of a multidimensional contingency table when some variables are posterior to the others. *Biometrika, 60*, 179-192.

Goodman, L. (1974). Exploratory latent structure analysis using both identifiable and unidentifiable models. *Biometrika, 61*, 215-231.

Haberman, S.J. (1979). *Analysis of qualitative data*, 2 vols. New York: Academic Press.

Hagenaars, J.A., (1990). *Categorical longitudinal data: Log-Linear panel, trend ad cohort analysis*. Newbury Park: Sage.

Hagenaars, J.A. (1998). Categorical causal modeling: latent class analysis and directed log-linear models with latent variables. *Sociological Methods and Research, 26*, 436-487.
Jabine, T., Loftus, E., Straf, M., Tanur, J., & R. Tourangeau R. (Eds.) (1984). *Cognitive aspects of survey methodology: Building a bridge between disciplines*. Washington, D.C.: National Academy Press.
Kuha, J., & Skinner, C. (1997). Categorical data and misclassification. In L. Lyberg *et al.* (Eds.), *Survey measurement and process quality* (pp. 633-670). New York: Wiley.
Lazarsfeld, P.F., & Henry, N.W. (1968). *Latent structure analysis*. New York: Houghton Muffins.
Magnac, T., & Visser, M. (1999). Transition models with measurement errors. *Review of Economics and Statistics, 81*, 466-474.
Martini, A. (1988). Retrospective versus panel data in estimating labor force transition rates: comparing SIPP and CPS. *Proceedings of the Social Science Section, American Statistical Association*, 109-114.
Mathiowetz, N.A. (2000). Retrospective recall: a review of the effects of length of recall on the quality of survey data. Paper presented at the 4th Conference on *Methodological Issues in Official Statistics*, Stockholm, October 12-13, 2000 (mimeo).
O'Muircheartaigh, C. (1996). Measurement errors in panel surveys: implications for survey design and for survey instruments. *Proceedings of the Scientific Meeting of the Italian Statistical Society, 1*, 207-218. Rimini: Maggioli.
Pfeffermann, D., Skinner, C., & Humphreys, K. (1998). The estimation of gross flows in the presence of measurement error using auxiliary variables. *Journal of the Royal Statistical Society, 161 Part 1*, 13-32.
Poterba, J., & Summers, L. (1986). Reporting errors and labor market dynamics. *Econometrica, 54*, 1319-1338.
Ryscavage, P., & Martini, A. (1990). Measuring the extent of unemployment: retrospective versus longitudinal surveys. *Proceedings of the Social Statistics Section*, *American Statistical Association*, 94-99.
Singh, A.C., & Rao, J.N.K. (1995). On the adjustment of gross flows estimates for classification errors with application to data from the Canadian Labor Force Survey. *Journal of the American Statistical Association, 90*, 1-11.
Skinner, C., & Torelli, N. (1993). Measurement error and the estimation of gross flows from longitudinal economic data. *Statistica, 3*, 391-405.
Sudman, S., & Bradburn, N.M. (1973). Effects of time and memory factors on response in surveys. *Journal of the American Statistical Association, 68*, 805-815.
Tourangeau R., Rips, L.J., & Rasinski, K.A. (2000). *The psychology of survey response*. Cambridge: Cambridge University Press.
Trivellato, U. (1999). Issues in the design and analysis of panel studies: a cursory review. *Quality & Quantity, 33*, 339-352.
Van de Pol, F., & Langeheine, R. (1997) Separating change and measurement error in panel surveys with an application to labour market data. In L. Lyberg (Eds.), *Survey measurement and process quality* (pp. 671-688). New York: Wiley.

Vermunt, J.K. (1996). *Log-linear event history analysis: A general approach with missing data, latent variables and unobserved heterogeneity*. Tilburg: Tilburg University Press.

Wiggins, L.M. (1973). *Panel analysis: Latent probability models for attitude and behavior change*. Amsterdam: Elsevier.

The Latent Class Approach to Estimating Gross Flows Affected by Correlated Classification Errors, with Application to Data from the French Labour Force Survey

Summary

Classification errors in the observed labour force state can lead to serious biases in the estimation of gross flows from panel data. We focus on correlated classification errors, which typically emerge when panel data are collected by retrospective interrogation, because of memory inaccuracy. We use latent class analysis to specify and estimate a model relating the true states and observed states at successive points in time. Our approach has the advantage that it does not require external information, either on misclassification probabilities or auxiliary variables affecting transition and classification probabilities. Besides, it does not crucially depend on the availability of multiple indicators of the true state. We apply the approach to data from the French Labour Force Survey. The model corrects flows in the expected direction: estimated true transition rates show higher mobility than observed ones. In addition, the measurement part of the model has significant coefficient estimates, and the estimated response probabilities show a clear, interpretable pattern. We compare our results with those produced by a different approach, which aims at obtaining robust estimates of true transition rates while paying marginal attention to the measurement process.

Keywords

Correlated classification errors, gross flows, labour force panel data; latent class analysis.

Appendix A: Final model specification, resulting from equation (4.1) and assumptions (a)-(d) in section 4.1

$v_{l_1}(1) = P(X(1) = l_1)$ **;**

$\mu_{l_t l_{t+1}}(t) = P(X(t+1) = l_{t+1} \mid X(t) = l_t)$ *t*= 1,…,8**;**

$z_{l_t l_{t+1} i_t}(t) = P(W(t) = i_t \mid X(t) = l_t, X(t+1) = l_{t+1})$ *t*= 1,…,4**;**

$z_{l_5 i_5}(5) = P(W(5) = i_5 \mid X(5) = l_5)$ **;**

$a_{l_t l_{t+1} j_t}(t) = P(Y(t) = j_t \mid X(t) = l_t, X(t+1) = l_{t+1})$ *t*= 5,…,8**;**

$a_{l_9 j_9}(9) = P(Y(9) = j_9 \mid X(9) = l_9)$ **;**

l_t , t = 1,…,9; i_t , t = 1,…,5; j_t , t = 5,…,9 vary over *E*, *U* and *O*.

(a):

$\mu_{l_1 l_2}(1) = \mu_{l_5 l_6}(5);$

$\mu_{l_2 l_3}(2) = \mu_{l_6 l_7}(6);$

$\mu_{l_3 l_4}(3) = \mu_{l_7 l_8}(7);$

$\mu_{l_4 l_5}(4) = \mu_{l_8 l_9}(8).$

(b)-(d3):

$$z_{l_t l_{t+1} i_t}(t) = \frac{\exp(\beta_{i_t}^{W(t)} + \beta_{i_t l_t}^{W(t)X(t)} + \beta_{i_t l_{t+1}}^{W(t)X(t+1)})}{\sum_{i_t=1}^{3} \exp(\beta_{i_t}^{W(t)} + \beta_{i_t l_t}^{W(t)X(t)} + \beta_{i_t l_{t+1}}^{W(t)X(t+1)})}$$ *t*= 1,…,4**;**

$$a_{l_t l_{t+1} j_t}(t) = \frac{\exp(\beta_{j_t}^{Y(t)} + \beta_{j_t l_t}^{Y(t)X(t)} + \beta_{j_t l_{t+1}}^{Y(t)X(t+1)})}{\sum_{j_t=1}^{3} \exp(\beta_{j_t}^{Y(t)} + \beta_{j_t l_t}^{Y(t)X(t)} + \beta_{j_t l_{t+1}}^{Y(t)X(t+1)})}$$ *t*= 5,…,8

(c):

$\beta_{l_t i_t}^{X(t)W(t)} = \omega_{l_t i_t}^{X(t)W(t)} + \delta(\omega_{l_t}^{X(t)} \exp(\Delta t))$ *t*=1,…,5;

$$\beta_{l_t j_t}^{X(t)Y(t)} = \omega_{l_t j_t}^{X(t)Y(t)} + \delta(\omega_{l_t}^{X(t)} \exp(\Delta t)) \qquad t=5,\ldots,9.$$

(d1):

$$\omega_{l_1 i_1}^{X(1)W(1)} = \omega_{l_2 i_2}^{X(2)W(2)} = \omega_{l_3 i_3}^{X(3)W(3)} = \omega_{l_4 i_4}^{X(4)W(4)} = \omega_{l_5 i_5}^{X(5)W(5)};$$

$$\omega_{l_5 j_5}^{X(5)Y(5)} = \omega_{l_6 j_6}^{X(6)Y(6)} = \omega_{l_7 j_7}^{X(7)Y(7)} = \omega_{l_8 j_8}^{X(8)Y(8)} = \omega_{l_9 j_9}^{X(9)Y(9)};$$

$$\omega_{l_1}^{X(t)} = \omega_{l_2}^{X(t)} = \omega_{l_3}^{X(t)} = \omega_{l_4}^{X(t)} = \omega_{l_5}^{X(t)};$$

$$\omega_{l_5}^{X(t)} = \omega_{l_6}^{X(t)} = \omega_{l_7}^{X(t)} = \omega_{l_8}^{X(t)} = \omega_{l_9}^{X(t)};$$

$$\beta_{l_2 i_1}^{X(2)W(1)} = \beta_{l_3 i_2}^{X(3)W(2)} = \beta_{l_4 i_3}^{X(4)W(3)} = \beta_{l_5 i_4}^{X(5)W(4)};$$

$$\beta_{l_6 j_5}^{X(6)Y(5)} = \beta_{l_7 j_6}^{X(7)Y(6)} = \beta_{l_8 j_7}^{X(8)Y(7)} = \beta_{l_9 j_8}^{X(9)Y(8)};$$

$$\beta_{i_1}^{W(1)} = \beta_{i_2}^{W(2)} = \beta_{i_3}^{W(3)} = \beta_{i_4}^{W(4)} = \beta_{i_5}^{W(5)};$$

$$\beta_{j_5}^{Y(5)} = \beta_{j_6}^{Y(6)} = \beta_{j_7}^{Y(7)} = \beta_{j_8}^{Y(8)} = \beta_{j_9}^{Y(9)}.$$

(d2):

$$a_{l_9 j_9}(9) = z_{l_5 i_5}(5);$$

$$a_{l_8 l_9 j_8}(8) = z_{l_4 l_5 i_4}(4);$$

$$a_{l_7 l_8 j_7}(7) = z_{l_3 l_4 i_3}(3);$$

$$a_{l_6 l_7 j_6}(6) = z_{l_2 l_3 i_2}(2);$$

$$a_{l_5 l_6 j_5}(5) = z_{l_2 l_1 i_2}(1).$$

Appendix B: Log-linear parameter estimates of measurement part of final model

Effect $\beta_{l_{t_t}}^{W(t)}$ for t=1,…,5, and $\beta_{l_{t_t}}^{Y(t)}$ for t=5,…,9

Independent parameter*	Estimated value	Standard error	z-value
E	-0.3749	0.2146	-1.747
U	-0.3154	0.1594	-1.978

* Effect coding was applied.

Effect $\beta_{l_{t+1}i_t}^{X(t+1)W(t)}$ for t=1,…,4, and $\beta_{l_{t+1}j_t}^{X(t+1)Y(t)}$ t=5,…,8

Independent parameter*	Estimated value	Standard error	z-value
eE	2.8550	0.3863	7.391
eU	-1.5328	0.2297	-6.674
uE	-1.7428	0.1576	-11.056
uU	2.5792	0.1444	17.861

* Effect coding was applied.

Parameters $\omega_{l_t}^{X(t)}$ for t=1,…,9

Independent parameter	Estimated value	Standard error	z-value
e	0.6973	0.1834	3.801
u	0.7109	0.1732	4.104
o	0.6755	0.1427	4.732

Effect $\omega_{l_t i_t}^{X(t)W(t)}$ for t=1,…,5, and $\omega_{l_t j_t}^{X(t)Y(t)}$ t=5,…,9

Independent parameter*	Estimated value	Standard error	z-value
eE	2.2222	0.4108	5.410
eU	-1.2085	0.2206	-5.478
uE	-1.4748	0.2178	-6.770
uU	1.7407	0.1245	13.984

*Effect coding was applied.

Working Papers

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), Ottobre 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.