Estimating Transition Models with Misclassification

Nicola Torelli
Department of Economics and Statistics, University of Trieste

Adriano Paggiaro
Department of Statistics, University of Padua
paggiaro@stat.unipd.it

1. Introduction

Longitudinal survey data are widely used to study micro-level dynamics of social and economic phenomena. Interest often focuses on the use of longitudinal data to study the transition of units among a finite set of states. For instance, when studying labor market dynamics, the estimation of some descriptive measures, like gross flows among labor force states, is crucial. The effects of misclassification of labor force states in estimating gross flows have been analyzed by many authors, and it is well known that estimates may be affected by severe bias, leading to a totally erroneous picture of labor force dynamics (for a review, see Skinner and Torelli, 1993). Statistical models have been proposed to estimate gross flows by taking into account classification errors; these models can either use validation data from re-interview studies or auxiliary variables (see, among others, Chua and Fuller, 1987; Pfefferman, Skinner and Humphrey, 1998).

Similar types of data may be used to estimate transition models aimed at explaining how the time spent in a state affects the probability of leaving it. Analysis of unemployment duration is a classic example of application of these models in the context of studying labor force dynamics. A review of the uses of data coming from different sampling schemes in estimating transition models may be found in Lancaster (1990).

Data on time spent in each state are often obtained by observing the states occupied by the same unit in successive waves of a panel survey. Considering again the analysis of labor market dynamics, if labor force states are misclassified, either observed transitions actually refer to situations in which units are still in the same state, or those observed in the same state have changed it. Parameter estimates of transition models when states are misclassified may be expected to be biased.

The problem is obviously more general, and similar consequences may be expected whenever longitudinal data from follow-up studies or panel surveys are available for estimating models for survival data.

In this paper, the effect of misclassification in estimating transition models is assessed and a strategy is proposed to obtain estimates of model parameters together with estimates of misclassification probabilities. As a motivating example, transition models in the context of analysis of unemployment duration, with data typically obtained from labor force surveys, are introduced in section 2. Section 3 presents a simple version of a transition model polluted by

misclassification in the destination state, and the effect of ignoring misclassification is evaluated by means of a simulation study. Sections 4 and 5 examine the problem of estimating the model, under different assumptions, by using standard strategies. Section 6 proposes estimation of a transition model adjusting for misclassification using a bayesian approach. Section 7 considers an application of the proposed model on a real data set. Some final comments and directions for future work are made in section 8.

2. Transition models and analysis of unemployment duration

Let us start by considering the simplest situation, in which data are obtained from two waves of a panel survey. Our interest lies in estimating the probability of transition between two states (hereafter E and U) during time k separating the two interviews; more precisely, we focus on estimating the probability of occupying state E for those in state U at the first survey, as a function of time already spent in state U. It is also assumed that data on time T spent in state U before the first interview are available.

The probability of transition from U to E depends on time T spent in state U and on a vector of covariates X. Let $S(t)$ and $f(t)$ denote respectively the survivor function and the density function of T, and let $d = 1$ if a transition to E is observed and $d = 0$ otherwise.

The dependence of the survivor function on X may be modeled in different ways (for instance, by assuming a proportional hazard specification), so that the effect of the covariates is measured by a set of parameters β. We assume that S belongs to a given parametric family S_γ ; estimation of β and γ may be obtained by maximizing the following likelihood function (assuming that a sample of n independent observations is available):

$$L(\beta,\gamma)=\prod_{i=1}^{n}\left[1-\frac{S(t_i+k;x_i,\beta,\gamma)}{S(t_i;x_i,\beta,\gamma)}\right]^{d_i}\left[\frac{S(t_i+k;x_i,\beta,\gamma)}{S(t_i;x_i,\beta,\gamma)}\right]^{1-d_i} . \qquad (1)$$

Under the assumption of proportional hazards, hazard function $h(t)=f(t)/S(t)$ depends on X only through parameters β, while its shape is defined by baseline hazard $h_0(t;\gamma)$. In this framework, the survivor function may be written as $S(t)=\exp\left[-H_0(t;\gamma)\exp(x'\beta)\right]$, where $H_0(t)=\int_0^t h_0(z)dz$ is the integrated hazard. This makes likelihood function (1) formally equivalent to that obtained in a generalized linear model context for binary dependent variables; the link function specification depends on assumptions made on $h_0(t;\gamma)$.

In the following, to simplify computation and to help understanding the main results, we assume a Weibull proportional hazard model, so that $H_0(t;\alpha)=t^\alpha$, and a simple specification for (1) is obtained (note, however, that the results do not depend on this specific assumption). In this case, interpretation of shape parameter α is easy, and is related to how time spent in a state may influence the hazard of transition to the destination state: the value $\alpha = 1$ discriminates between negative duration dependence ($0 < \alpha < 1$) and positive duration dependence ($\alpha > 1$).

The simple situation considered here closely corresponds to what is typically found when analyzing unemployment duration data from labor force surveys which adopt rotating sampling schemes (Trivellato and Torelli, 1989), *i.e.*, considering a state-based sample with follow-up. These models have frequently been applied by econometricians and other social scientists (see Lancaster, 1990). In the context outlined above, as regards the impact of non-sampling errors comparatively more attention has been devoted to analysis of the effect of measurement errors in duration (*e.g.*, Holt, McDonald and Skinner, 1991 documented the effect of errors when data are obtained from a state-based design with follow-up).

3. Effect of misclassification in destination state in estimating duration models

Let us assume that: (i) at the first interview the state is observed without error, (ii) at the second interview the state may be misclassified. This simple situation has in fact some practical interest, as the follow-up interview is often less accurate (*e.g.*, admitting a higher rate of proxy responses) and/or obtained using different modes that may be more error-prone (for instance, telephone instead of face-to-face interviews). Note that the same assumption is made by Poterba and Summers (1995) in a similar context. Indicator d is then measured with error, and:

(a) $P(d=1 \mid \mathrm{E})=p_E$, *i.e.*, the probability that we observe a transition from U to E when the true state at the second interview is E, is not 1 (but we expect it to be very close to 1),

(b) $P(d=1 \mid \mathrm{U})=p_U$, *i.e.*, the probability that we observe a transition from U to E but that the true state at the second interview is U, is not 0 (but we expect it to be very close to 0).

Hereafter, p_E and $(1-p_U)$ are called "reliabilities".

This simple misclassification mechanism may induce severe bias in parameter estimates. To appreciate this, we analyzed the results of a Monte Carlo study, with state-based data with follow-up simulated assuming that transition times are generated by a proportional hazard model. Here, we only report some results with a Weibull baseline. Note, however, that we obtained the same conclusions even though different (and more flexible) assumptions were made for the baseline hazard function. In designing the simulation study, we matched a real situation, so that the size of the sample (999 individuals) and data on covariates (age, gender, education, marital status) were fixed at the values actually observed in the Italian labor force survey for those unemployed in northern Italy in October 1997. Moreover, the parameters were such that the average duration was not far from those typically observed for unemployment duration in Italy.

Table 1 contains a selection of the simulation results, in which g_E and g_U denote the logit transforms, respectively, of p_E and p_U. These probabilities were fixed respectively at around 0.97 and 0.05, with misclassification errors close to those found in many empirical works (see, among others, Poterba and Summers, 1995, which contains estimates from validation data) and specifically in studies carried out with Italian data (Bassi, Torelli and Trivellato, 1998). It is

interesting to note that the average size of parameters b is always reduced, showing a sort of attenuation effect.

Table 1. Monte Carlo simulations for a Weibull proportional hazard model with misclassification; number of replications = 100, $g_E = 4$, $g_U = -3$

Parameter	True	Average	st.dev	True	Average	st.dev.	True	Average	st.dev.
Log *a*	-0.5	-0.398	0.118	0	-0.014	0.089	0.5	0.378	0.076
Intercept	-2.5	-2.179	0.362	-4	-3.372	0.402	-6	-4.854	0.464
Age	-1	-0.656	0.319	-1	-0.743	0.330	-1	-0.785	0.300
Gender (1=F)	1	0.749	0.151	1	0.755	0.157	1	0.727	0.172
Marit. (1=married)	0	-0.016	0.227	0	-0.067	0.259	0	-0.021	0.232
Educ. (1=higher)	1	0.736	0.159	1	0.777	0.183	1	0.745	0.178

Evidence on the strong biasing effect due to misclassification has clearly been confirmed by larger simulation exercises assuming different sizes for misclassification probabilities. Table 2 lists some results on a Weibull baseline with negative duration dependence and symmetric misclassification probabilities $p_U = (1-p_E)$ ranging from 0.01 to 0.27. As expected, more bias is induced by a more substantial amount of misclassification, but even moderate misclassification leads to a non-negligible bias in the parameter estimates.

More clear evidence emerges about the relation between bias and true rate of transitions to E: given a fixed p_E, the effect is greater if there is a prevalence of true transitions, whereas it is negligible if such transitions are few; the opposite is obviously true for p_U. Table 3 presents the results for a rate of transitions close to 15% and misclassification only on one side (either $p_E = 1$ or $p_U = 0$). A simple explanation is that the effect is stronger whenever there are more "candidates" for misclassification. The following sections show how this different effect influences the possibility of jointly estimating q, b and both p_E and p_U.

Table 2. Monte Carlo simulations, under different levels of symmetric misclassification probabilities $p_U = (1-p_E)$; averages over 100 replications

Parameter	True	Symmetric misclassification probabilities $p_U = (1-p_E)$				
		0	**0.01**	**0.05**	**0.12**	**0.27**
Log *a*	-0.5	-0.508	-0.483	-0.382	-0.263	-0.104
Intercept	-2.5	-2.575	-2.527	-2.271	-2.035	-1.840
Age	-1	-1.030	-0.977	-0.699	-0.446	-0.154
Gender (1=F)	1	1.003	0.955	0.740	0.515	0.239
Marit. (1=married)	0	0.063	0.058	0.027	-0.000	0.039
Educ. (1=higher)	1	1.024	0.973	0.760	0.512	0.228

Table 3. Monte Carlo simulations under different levels of asymmetric misclassification probabilities (either p_U = 0 or p_E = 1); averages over 100 replications

Parameter	True	Asymmetric misclassification probabilities			
		p_U = 0		p_E = 1	
		p_E = 0.95	p_E = 0.73	p_U = 0.05	p_U = 0.27
Log *a*	-0.5	-0.518	-0.480	-0.390	-0.174
Intercept	-2.5	-2.531	-2.913	-2.201	-1.523
Age	-1	-1.033	-0.997	-0.731	-0.284
Gender (1=F)	1	0.959	0.918	0.753	0.349
Marit. (1=married)	0	0.001	0.028	-0.025	-0.020
Educ. (1=higher)	1	1.001	0.968	0.776	0.348

4. Estimation of transition models with known reliabilities

If p_E and p_U were known, it would be possible to obtain consistent estimates of the parameters of the model by maximizing the following likelihood function:

$$L(b,q) = \prod_{i=1}^{n} \left[P(d_i = 1; x)\right]^{d_i} \left[1 - P(d_i = 1; x)\right]^{1-d_i}, \tag{2}$$

where

$$P(d = 1; x) = P(d = 1|\mathrm{E})P(\mathrm{E}; x) + P(d = 1|\mathrm{U})P(\mathrm{U}; x) = \\ = p_E\left[1 - \frac{S(t_i + k; x_i, b, q)}{S(t_i; x_i, b, q)}\right] + p_U\left[\frac{S(t_i + k; x_i, b, q)}{S(t_i; x_i, b, q)}\right]$$

is the probability of observing a transition.

Table 4 lists some simulations (carried out in the same conditions stated in the previous section) by using likelihood (2) and assuming that reliabilities of indicators of states are fixed to certain values. When p_E and p_U are fixed to their true values, the parameters of interest are consistently estimated without bias, but the most striking evidence regards the sensitivity of the estimates when p_E and p_U are fixed to values which are different from the true ones but reasonably close to them.

Usually, in real applications, it is reasonable to have an idea about the range in which the probabilities of error should lie, and one could be tempted to use values obtained from validation studies, carried out in situations similar to those considered, for the reliabilities. Referring to a similar situation, *i.e.*, estimating models for a dichotomous dependent variable with misclassification, Hausman *et al*. (1998) show that, if consistent estimates of p_E and p_U are used, it is possible to obtain consistent estimates of parameters b, but their standard errors are underestimated. If p_E and p_U are not estimated consistently, neither are the estimates of b consistent.

Table 4. Monte Carlo simulations of a transition model with fixed reliabilities; averages over 100 replications

Parameter	**True**	**ML estimation with fixed g_E= -g_U (true g_E=3)**				
		4	**3.5**	**3**	**2.5**	**2**
Log *a*	-0.5	-0.426	-0.460	-0.525	-0.660	-1.141
Intercept	-2.5	-2.348	-2.409	-2.524	-2.748	-3.656
Age	-1	-0.865	-0.950	-1.098	-1.356	-1.802
Gender (1=F)	1	0.854	0.916	1.025	1.215	1.568
Marit. (1=married)	0	-0.001	0.003	0.018	0.064	0.592
Educ. (1=higher).	1	0.826	0.888	1.001	1.209	1.838

5. Estimation of transition models with unknown reliabilities

Likelihood function (2) may be used to obtain estimates of the model parameters even when misclassification probabilities are unknown. As stated in Abrevaya and Hausman (1999), a necessary condition for identification is the obvious restriction $p_E > p_U$, indicating that the probability of being classified in a state, say E, is greater if the true state is E than if it is U.

To simplify the maximization of (2), it is convenient to use the EM algorithm. To this end, let A denote a variable indicating actual transition to employment (*i.e.*, without error). Likelihood (2) is then obtained by marginalization of the following joint likelihood:

$$L(b,q,p_E,p_U)=\prod_{i=1}^{n}\left[1-\frac{S(t_i+k;x_i,b,q)}{S(t_i;x_i,b,q)}\right]^{A_i}\left[\frac{S(t_i+k;x_i,b,q)}{S(t_i;x_i,b,q)}\right]^{1-A_i} \cdot \prod_{i=1}^{n}\left[{p_E}^{d_i}(1-p_E)^{1-d_i}\right]^{A_i}\left[{p_U}^{d_i}(1-p_U)^{1-d_i}\right]^{1-A_i} . \quad (3)$$

In the M step we maximize (3) conditionally on the expected value of A for each unit in the sample, which is easily calculated in the E step as follows:

$$E(A_i)=\frac{\left[1-\frac{S(t_i+k;b,q)}{S(t_i;b,q)}\right]\left[{p_E}^{d_i}(1-p_E)^{1-d_i}\right]}{\left[1-\frac{S(t_i+k;b,q)}{S(t_i;b,q)}\right]\left[{p_E}^{d_i}(1-p_E)^{1-d_i}\right]+\frac{S(t_i+k;b,q)}{S(t_i;b,q)}{p_U}^{d_i}(1-p_U)^{1-d_i}} . \quad (4)$$

In practice, maximizing the likelihood function for this model very often leads to totally unreasonable values for some parameters, and the algorithm converges to points at the border of the parametric space. These problems may have to do with weak identification of some parameters of the model, inducing flatness of the likelihood function. In fact, we have empirically verified that, when very large samples are available, reasonable results may be obtained more frequently. This point certainly deserves more thorough investigation.

Nonetheless, some results on the performance of maximum likelihood estimation, obtained within larger Monte Carlo studies, not presented here, may shed light on the quality of estimates of p_E and p_U obtained by using standard maximum likelihood and on how they are related to transition rates. If data include few transitions, the estimates of p_U may either go towards the border of the parametric space or show reasonable values. Conversely, as outlined in commenting Table 3, there are few transitions which are candidate for misclassification, thus unstable and often unreasonable estimates of p_E are obtained. It is important to note, however, that this lack of information seems to have no biasing effect on the other parameters of the model. When the transition rate is high, the same results are obviously obtained if we invert p_E and p_U.

6. Bayesian estimation of transition models with classification errors

In applications to real data, it is reasonable to have a more or less vague idea about which values are credible for misclassification probabilities. At least we know that they should not be very far from 0. This encourages us to formulate the transition model within a bayesian framework. In this case, the posterior distribution of the parameters is given by:

$$g(\beta, \theta, p_E, p_U, A | t, X, \delta) = L(\beta, \theta, p_E, p_U, A | t, X, \delta) g_0(\beta, \theta, p_E, p_U, A),$$

where g_0 denotes prior distribution and $L(\beta, \theta, p_E, p_U, A | t, X, \delta)$ has the same form defined in (3). At least for the most interesting parameters, *i.e.*, β, θ and the reliabilities, it is plausible to summarize the information available on them by specifying informative priors. Posterior distribution is obviously quite complex, but may be explored using MCMC methods. For the model examined here, one suitable strategy is to adopt Metropolis-Hastings within Gibbs sampler iterations (see Chib and Greenberg, 1995). The Metropolis-Hastings algorithm is used to sample from intractable full conditional distributions arising within the Gibbs sampler. At the *i*-th iteration of the algorithm, conditionally on the values drawn at the previous iteration, the main steps are the following:

1. Draw A^i from $g(A | \beta^{i-1}, \theta^{i-1}, p_E^{i-1}, p_U^{i-1}; t, X, \delta)$;
2. Draw (β^i, θ^i) from $g(\beta, \theta | A^i; t, X)$;
3. Draw p_E^i from $g(p_E | A^i; \delta)$;
4. Draw p_U^i from $g(p_U | A^i; \delta)$.

Drawing a sample from g(*A*) at step 1 is straightforward, as it is a binomial distribution whose mean is given by (4). Instead, the full conditional distributions of the other parameters are not in close form, so that a single step of the Metropolis-Hastings algorithm is used at steps 2, 3, and 4. For a generic parameter ϕ, we draw a candidate value ϕ^* from a normal distribution centered at the previous value ϕ^{i-1}, and accept the candidate with probability

$\min\left(g(f^{*})/g(f^{i-1}),1\right)$; if the candidate is rejected, we keep $f^{i} = f^{i-1}$. Variances of normal distributions are chosen in order to keep the acceptance rate around 20%.

Table 5 contains some characteristics of posterior distributions resulting from the use of MCMC techniques as defined above, applied to three different simulated samples from a proportional hazard model with Weibull baseline. The tables (first three columns) also list the results obtained by estimating parameters by maximum likelihood when misclassification is ignored. Once again, simulated data were chosen in order to resemble data from two successive waves from the Italian labor force survey, aiming at analyzing transitions from unemployment to employment. Thus, the sample size is again kept at 999 in each simulation and the values of the parameters are fixed close to the ones estimated for a real sample, the only exception being the use of different structures of duration dependence (different values for a).

The misclassification probabilities chosen for the simulation study are similar to those encountered in real situations, as far as classification of labor force states is concerned ($g_E = 4$ and $g_U = -3$, which are approximately equivalent to setting $p_E = 0.97$ and $p_U = 0.05$). The prior distributions (assumed to be independent) for parameters b are gaussian $N(0,4)$ (with the only exception being the mean of the constant term, which is fixed at 4, to match the average unemployment duration in the sample in case of exponential baseline and no effects of covariates). For g_E and g_U the priors are $N(4.5,1)$ and $N(-3.5,1)$, thus misclassification probabilities are *a-priori* assumed to vary around respectively 1% and 3% - *i.e.*, less than the true misclassification probabilities used to generate the data. 100,000 Gibbs sampler iterations were considered, after a burn-in period of 20,000 iterations, and a sample of size 1,000 from the posterior is obtained selecting values every 100 iterations.

Table 5: Some summaries of a sample obtained (by MCMC) from posterior distribution of transition models with misclassification (see the main text for details)

	Maximum likelihood			**Posterior distribution summaries and percentiles**						
	True	**MLE**	**s.d.**	**Mode**	**Mean**	**5%**	**25%**	**50%**	**75%**	**95%**
Log *a*	-0.5	-0.314	0.107	-0.400	-0.446	-0.791	-0.544	-0.419	-0.317	-0.199
Intc.	-2.5	-2.291	0.373	-2.689	-2.643	-3.540	-2.961	-2.654	-2.321	-1.820
Age	-1	-0.618	0.296	-0.758	-0.859	-1.616	-1.141	-0.840	-0.567	-0.159
Gender	1	0.815	0.162	1.163	1.110	0.682	0.909	1.097	1.267	1.607
Marit.	0	-0.108	0.225	-0.107	-0.104	-0.679	-0.338	-0.113	0.125	0.499
Educ.	1	0.601	0.163	0.814	0.819	0.448	0.629	0.803	0.979	1.263
g_E	4			4.398	4.500	3.310	4.024	4.501	4.977	5.697
g_U	-3			-2.843	-3.090	-4.208	-3.424	-3.017	-2.703	-2.298

	Maximum likelihood			Posterior distribution summaries and percentiles						
	True	**MLE**	**s.d.**	**Mode**	**Mean**	**5%**	**25%**	**50%**	**75%**	**95%**
Log *a*	0	0.057	0.087	0.062	0.062	-0.162	-0.015	0.069	0.149	0.254
Intc.	-4	-3.612	0.408	-4.810	-4.982	-6.495	-5.547	-4.930	-4.382	-3.662
Age	-1	-0.525	0.300	-0.798	-0.907	-1.820	-1.217	-0.881	-0.537	-0.125
Gender	1	0.678	0.165	1.020	1.072	0.575	0.858	1.054	1.272	1.621
Marit.	0	0.139	0.222	0.277	0.549	-0.112	0.196	0.476	0.829	1.481
Educ.	1	0.686	0.164	1.024	1.087	0.596	0.867	1.057	1.283	1.672
g_E	4			4.318	4.526	3.390	4.046	4.505	5.015	5.721
g_U	-3			-2.451	-2.708	-3.766	-2.932	-2.572	-2.351	-2.110

	Maximum likelihood			Posterior distribution summaries and percentiles						
	True	**MLE**	**s.d.**	**Mode**	**Mean**	**5%**	**25%**	**50%**	**75%**	**95%**
Log *a*	0.5	0.410	0.072	0.478	0.489	0.337	0.425	0.487	0.555	0.650
Intc.	-6	-5.106	0.436	-5.855	-6.122	-7.601	-6.640	-6.030	-5.528	-4.971
Age	-1	-0.439	0.291	-0.594	-0.610	-1.305	-0.863	-0.592	-0.347	0.015
Gender	1	0.811	0.164	1.056	1.070	0.673	0.885	1.055	1.225	1.541
Marit.	0	0.136	0.218	0.247	0.231	-0.238	0.017	0.220	0.423	0.741
Educ.	1	0.713	0.161	0.846	0.913	0.533	0.754	0.898	1.052	1.343
g_E	4			4.580	4.486	3.334	3.992	4.509	4.969	5.630
g_U	-3			-2.796	-3.092	-4.236	-3.557	-3.092	-2.718	-2.357

The results obtained are encouraging and the bayesian approach may, in this case, overcome some difficulties which emerge when using maximum likelihood. The availability of informative priors for misclassification probabilities is crucial but, as already noted, it is quite reasonable to use prior distributions which assign a substantial probability (*e.g.*, more than 0.95) to the interval (0,0.2) for misclassification probabilities.

As already argued, if there are few (or, conversely, too many) transitions, the data do not convey enough information to estimate one of the reliabilities, and the posterior distribution of γ_E (γ_U) essentially reflects only information summarized by prior distribution. This influences estimation of the other parameters very little, as their posterior distribution is mostly concentrated around values which are close to the true ones. More precisely, p_U is close to its true value 3, while parameters a and b are clearly moved far apart from their maximum likelihood estimates.

To appreciate the performance of the MCMC procedure, Figure 1 reports one sequence of values from the marginal posterior distribution of the parameter b associated to "education" simulated by MCMC techniques using the second data set of table 5. Figure 2, shows the marginal posterior density of the same parameter (obtained using a kernel smoothing technique). The other parameters show substantially a similar behavior.

Figure 1: Sequence of values simulated by MCMC for the parameter b associated to "education"

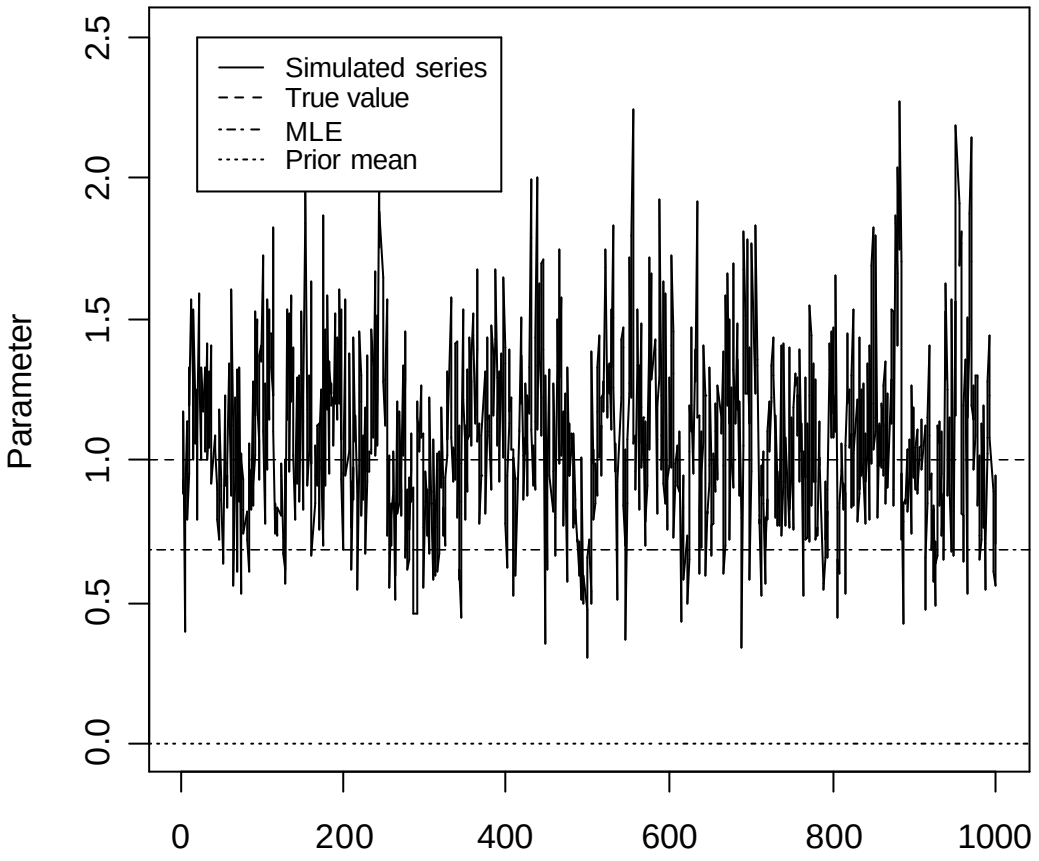
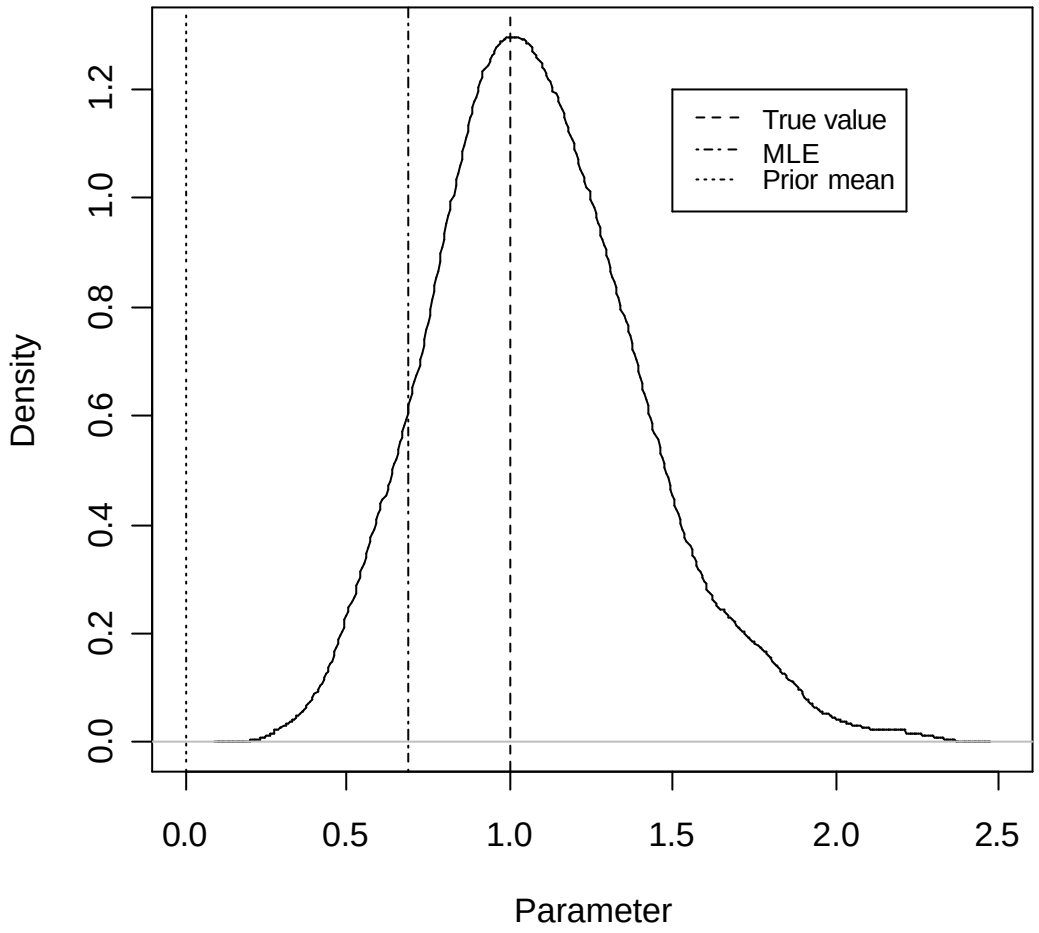


Figure 2: Density of the marginal posterior distribution for the parameter b associated to "education"



7. Estimation of a transition model with misclassification using data from Italian labor force survey

In Italy, the labor force survey adopts a sample rotation scheme which provides longitudinal data. More precisely, each household is included in the sample for two consecutive surveys, then it drops out for two surveys, and re-enters the sample for two final waves. About 50% of the sample is common in two successive surveys. For each household member, information on labor force state is obtained, and those who declare themselves as unemployed also answer a question on how long they have been unemployed. These data have been used to estimate transition models of the form discussed in the previous sections. Similar data are also available from other labor force surveys in developed countries (in fact, the scheme adopted in Italy closely resembles the Current Population Survey carried out in the U.S.) and data like these have actually been used to estimate transition models to employment (see Trivellato and Torelli, 1989). The Italian data are probably also polluted by substantial classification errors (as documented by Bassi, Torelli and Trivellato, 1998). Although the model proposed here rests on some simplifying assumptions which may easily be considered

not completely realistic, we think that it is useful to gain some insights on the sensitivity of estimates to misclassification.

Table 6 contains the results of the application of MCMC methods to a real sample coming from the Italian labor force survey. A sample of 999 labor force units from Northern Italy who were unemployed in October 1997 is available, for whom we observe the duration of the ongoing unemployment spell (and other covariates) and the labor state after 3 months (January 1998). Starting from these data, we used the MCMC methods described in section 6. Table 6 shows some summaries of the sample of 1,000 values coming from posterior distribution.

Table 6: Sample obtained (by MCMC) from posterior distribution of transition models with misclassification; data from Italian labor force survey, October 1997

	Maximum likelihood		**Posterior distribution summaries and percentiles**						
	MLE	**s.d.**	**Mode**	**Mean**	**5%**	**25%**	**50%**	**75%**	**95%**
Log a	-0.440	0.129	-0.535	-0.609	-1.114	-0.719	-0.561	-0.432	-0.289
Intc.	-2.078	0.366	-2.328	-2.314	-3.211	-2.632	-2.312	-1.971	-1.482
Age	-0.776	0.350	-0.735	-0.800	-1.565	-1.074	-0.774	-0.498	-0.141
Gender	0.839	0.171	1.020	1.091	0.697	0.895	1.059	1.258	1.603
Marit.	-0.130	0.270	0.011	-0.025	-0.550	-0.263	-0.024	0.180	0.574
Educ.	0.460	0.171	0.595	0.598	0.253	0.437	0.588	0.737	0.984
g_E			4.462	4.503	3.352	4.022	4.504	4.996	5.624
g_U			-3.212	-3.440	-4.479	-3.840	-3.393	-2.994	-2.563

The results confirm that, if we take into account the probability of misclassification, the absolute values of the parameters of interest tend to be higher than those obtained by standard maximum likelihood estimation assuming no misclassification. This is particularly true for the duration dependence parameter a, and for parameters b associated with gender and education. Moreover, the classification error for the unemployed is mostly concentrated between 2% and 5%; interpretation of the estimated probability of misclassification for the employed is more difficult, due to the small fraction of transition observed in the sample.

8. Concluding remarks

This paper highlights how important it is to take into account misclassification when estimating transition models. As the use of some standard methods is precluded, specifying reasonable models for misclassification may provide a solution. Misclassification probabilities entering the model are usually unknown, but we can make reasonable guesses about their probable magnitude. When this is the case, classic inferential methods are still possible but less natural and less practical than a bayesian approach. This approach gives estimates both of the parameters of the model and of the misclassification probabilities. We considered a simplified model, and the main directions for future research will extend the framework towards more realistic specifications. More specifically, the model may be extended in many directions: (i) by examining data coming from multi-wave panel surveys instead of limiting attention simply to two waves; (ii) by allowing misclassification probabilities to depend on covariates; (iii) by

considering the case of substantial misclassification also at the first interview; (iv) by extending the model to analysis of movements among a set of more than two states.

Extending the model may be of great interest when analyzing the general case of modeling event history data obtained from panel surveys. There is a clear connection of the model examined here with the more general problem of analyzing categorical data with misclassification (for a review, see Kuha and Skinner, 1997) which may be further explored to extend the results presented here. As already noted, more effort is needed to fully understand identification problems when using maximum likelihood.

REFERENCES

Abrevaya, J. and Hausman, J.A. (1999). "Semiparametric Estimation with Mismeasured Dependent Variables: an Application to Duration Models for Unemployment Spells". *Annales d'Economie et de Statistique*, 55-56, 243-275.

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

Chib, S. and Greenberg, E. (1995). "Understanding the Metropolis-Hastings Algorithm". *The American Statistician*, 49, 4, 327-335.

Chua, T. and Fuller, W.A. (1987). "A Model for Multinomial Response Errors Applied to Labor Flows". *Journal of the American Statistical Association*, 82, 46-51.

Hausman, J.A., Abrevaya, J. and Scott-Morton, F.M. (1998). "Misclassification of the Dependent Variable in a Discrete-Response Setting". *Journal of Econometrics*, 87, 239-269.

Holt, D., McDonald, J. and Skinner, C. (1991). "The effect of measurement error on event history analysis". In Biemer, P.P., Groves, R.M., Lyberg, L.E., Mathiowetz, N.A. and Sudman, S. (eds.), *Measurement Error in Surveys*, New York: Wiley, 665-686.

Kuha, J. and Skinner, C. (1997). "Categorical Data Analysis and Misclassification". In Lyberg, L.E., Biemer, P.P., De Leeuw, E., Dippo, C.S., Schwarz, N. and Trewin, D (eds.), *Survey Measurement and Process Quality*, New York: Wiley, 633-670.

Lancaster, T. (1990). *The Econometric Analysis of Transition Data.* Cambridge: Cambridge University Press.

Pfefferman, D., Skinner, C.J. and Humphrey, S.K. (1998). "The Estimation of Gross Flows in the Presence of Measurement Error Using Auxiliary Variables". *JRSS*, Series A, 161, 13-32.

Poterba, J.M. and Summers, L.H. (1995). “Unemployment Benefits and Labor Market Transitions: a Multinomial Logit Model with Errors in Classification”. *Review of Economics and Statistics*, 77, 207-216.

Skinner, C.J. and Torelli, N. (1993). ”Measurement Errors and the Estimation of Gross Flows from Longitudinal Economic Data”. *Statistica*, 3, 391-405.

Trivellato, U. and Torelli, N. (1989). “Analysis of Labor Force Dynamics from Rotating Panel Survey Data”. *Bulletin of The International Statistical Institute*, Vol. LIII, Book 2, 425-444.