

Modelling Error Dependence in Categorical Longitudinal Data

*Dimitris Pavlopoulos, Vrije Universiteit Amsterdam, Paulina Pankowska,
Vrije Universiteit Amsterdam, Bart Bakker, Statistics Netherlands;
Vrije Universiteit Amsterdam, and Daniel Oberski, Utrecht University*

Introduction

Measurement error poses a considerable threat to the analysis of discrete longitudinal phenomena. Even for datasets of high quality, a small amount of unaccounted-for measurement error can severely bias estimates of over-time change, such as employment mobility (Hagenaars, 1990, 1994). While there is a growing consensus regarding the strong potential biasing effects of measurement error and the necessity for measurement-error correction, most of the current research in the social sciences ignores this potential problem.

The studies that do account for measurement error in longitudinal data typically assume that this error is random. However, such an assumption is often highly unrealistic (Skinner et al., 2002) as there are several commonly occurring situations in which the error in longitudinal data is systematic rather than random. That is, error dependency emerges when (i) some population subgroups are more prone to error than others; (ii) certain data-collection processes (e.g. proxy interviewing in surveys and administrative delays in the collection of register data) create systematic error; or (iii) error autocorrelation occurs, that is, the presence of an error at one point in time increases the probability of repeating that error in the future.

This chapter shows how hidden Markov models (HMMs) can be used to correct for error dependency in the study of longitudinal categorical phenomena. Moreover, it discusses several common error dependency scenarios that often occur in data used for social science research. We look at situations whereby the error depends on: (i) covariates related to individual level characteristics (such as age), (ii) the data-collection processes (i.e. the use of proxy interviews and dependent interviewing), and (iii) whether an error occurred previously (i.e. the presence of autocorrelated errors). We then show how such dependencies can be modelled using extended, multiple indicator HMMs. In doing so, we focus on the illustrative example of employment mobility in the Netherlands and more

specifically on transitions between different types of employment contracts on the Dutch labour market. For this purpose, we use linked data from the Dutch Labour Force Survey and the Employment Register. In this application, we also show how to interpret the results of HMMs that correct for systematic measurement error.

Hidden Markov models (HMMs), which are a special case of latent class modelling applied to longitudinal processes, are a useful tool to estimate and correct for measurement error in longitudinal categorical data. These models were developed by Wiggins (1973), Poulsen (1990), and van der Pol and de Leeuw (1986). Their increasing popularity can be attributed to the fact that HMMs do not require the use of a gold-standard error-free data source as a benchmark. In its basic structure, the HMM relies on two assumptions: the *Markov* assumption and the *local independence* assumption. The Markov assumption implies that the latent state in time point t , X_t , only depends on the latent state in $t - 1$, X_{t-1} . Conditional on X_{t-1} , the latent value at t , X_t , is independent of all other realizations of the latent state X . This rather restrictive assumption (Shorrocks, 1976) of an autoregressive process of order 1, AR(1), is relaxed with the inclusion of covariates controlling for (un)observed heterogeneity in the way proposed by Vermunt et al. (1999).

The local independence assumption, which is also referred to as independent classification error (ICE), is the second crucial assumption of the HMM. Its importance stems from the fact that it ensures model identification for the HMMs with one indicator per time point for the phenomenon of interest. This assumption implies that conditional on the true state X_t , the observed state Y_t is independent of all other realizations of the observed and true states (Biemer and Trewin 1997; Oberski et al., 2015). In practice, this means that the (standard) HMM treats all measurement error as random and does not model any dependent error.

The ICE assumption is problematic as, while it plays an important role in the identification of one-indicator HMMs (Hagenaars, 1988; Oberski, 2016), it is often unrealistic and sometimes even harmful. It is unrealistic as it ignores commonly occurring situations that lead to correlated errors. In surveys, this might be caused by respondents having a personal 'style' in answering questions (Billiet and Davidov, 2008). In register data, errors are often carried over time (e.g. as a result of software systems introducing these errors repeatedly), and are only adjusted if the company maintaining the register has a trigger or incentive to do so. The ICE assumption may be harmful as ignoring its violation leads to bias in the HMM parameter estimates (Vacek, 1985; Torrance-Rynard and Walter, 1997; Georgiadis et al., 2003; Qu and Hagdu, 1997). For example, Pankowska et al. (2019b) illustrate that ignoring the violation of this assumption in HMMs leads to a relative bias of up to 310%.

Relaxing this assumption, without having to impose additional restrictions and without risking poor model identifiability, requires the use of extended HMMs with multiple indicators per time-point for the same phenomenon. All indicators may contain measurement error and this error can be correlated for each of the

indicators. However, while error dependency is possible within each of the indicators, the error should be independent between the indicators. Such a situation is further explained in Section 8.2.

While research using extended, multiple-indicator HMMs has been limited due to the scarce availability of such data, this approach has provided some interesting results with respect to employment mobility (Pavlopoulos and Vermunt, 2015; Pankowska et al., 2018, 2019b; Bassi et al., 2000; Manzoni et al., 2010), health status (Bartolucci et al., 2009), and financial products (Paas et al., 2007). For example, Pavlopoulos and Vermunt (2015) used such an extended HMM specification with two indicators for the type of employment contract, coming from linked survey and register data, and found that almost half of the 3-monthly transitions from temporary to permanent employment are due to measurement error.

The rest of the chapter is organized as follows: Section 8.1 introduces hidden Markov models and illustrates how they can be used for measurement error correction in categorical longitudinal data; Section 8.2 presents an extended hidden Markov model that corrects for measurement error that is correlated over time; Section 8.3 presents an application of this model in the study of employment mobility in the Netherlands; and Section 8.4 presents some conclusions and recommendations for further research.

8.1 Measurement Error Correction with Hidden Markov Models

Consider a discrete random variable Y with k categories observed for I individuals each at T time points. For the i -th person at the t -th time point, we then obtain an observation Y_{it} ; combining all time points per individual leads to an observed 'path' $Y_i = \{Y_{it} : t \in 1 \dots T\}$. The HMM assumes that the variable Y is the observed realization of the true (latent) variable X (also with k categories¹), which is measured with an error. As mentioned in the previous section, the model relies on two main assumptions: the Markov assumption and the ICE assumption. The former assumption implies that, conditional on the latent state at $t - 1$ ($X_{i(t-1)}$), the latent state of an individual i at the time point t (X_{it}) is independent of all other true and observed values $X_{it'}$ and $Y_{it'}$ where $t' \neq t - 1$. In other words, this assumption indicates that the latent states follow an autoregressive process of order 1 (AR(1)). The latter assumption implies that, at each time point and for each individual the observed state Y_{it} only depends on the latent state X_{it} ; conditional on that it is independent of all other latent and observed states.

¹ While the HMMs used to account for measurement error often assume that the number of categories of Y and X is the same, the number of categories of the true and observed state in fact may differ, i.e. there may be measurement error.

The joint probability of observing a certain path can be specified as:

$$P(Y_i = y_i) = \sum_{x_0=1}^k \sum_{x_1=1}^k \dots \sum_{x_T=1}^k P(X_{i0} = x_0) \prod_{t=1}^T [P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})] \prod_{t=0}^T P(Y_{it} = y_{it} | X_{it} = x_t). \quad (1)$$

where $P(X_{i0} = x_0)$ represents the initial latent state probability, and $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})$ are the latent transition probabilities. These two terms represent the structural part of the model (which is often the quantity of scientific interest). $P(Y_{it} = y_{it} | X_{it} = x_t)$ are the error probabilities and represent the measurement part of the model.

The basic structure of this model (van der Pol and Langeheine, 1990) is illustrated in Figure 8.1. In this Figure, the absence of an arrow between two states indicates that we assume that these two states are conditionally independent. Thus, the fact that there are no arrows between the realizations of the observed state Y illustrates how the ICE assumption works: the observed states Y_{t-1}, Y_t, Y_{t+1} are independent to each other conditional on the latent state X .

The identification of this HMM—with one observed indicator Y_{it} for the variable of interest X_{it} —requires that either the measurement error probabilities or the latent transition probabilities are time homogeneous. Time-homogeneity of the measurement error parameters means that $P(Y_{i(t-1)} = s | X_{i(t-1)} = r) = P(Y_{it} = s | X_{it} = r)$ for every t , while the homogeneity of latent transition probabilities would imply the same for $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})$. With either of these restrictions, the model is identified with at least three time points (Vermunt et al., 1999).

It is important to note that the Markov assumption rarely holds in practice (Shorrock, 1976). Therefore, a realistic model should also include covariates to control for observed heterogeneity. Vermunt et al. (1999) have proposed such an HMM that includes time-constant and time-varying covariates. In this model, covariates Z are allowed to affect the latent transition probabilities. Additionally, these probabilities may also be affected by unobserved characteristics. To correct for such unobserved heterogeneity both parametric and non-parametric

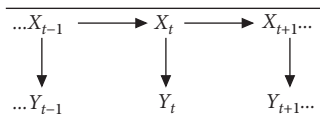


Figure 8.1 Path diagram for the HMM with one indicator.

approaches have been proposed. A parametric approach involves integrating out the time-invariant individual effects F_i . The joint probability then takes the following form:

$$\begin{aligned}
 P(Y_i = y_i|Z_i) = & \int \sum_{x_0=1}^3 \sum_{x_1=1}^3 \dots \sum_{x_T=1}^3 P(X_{i0} = x_0) \\
 & \prod_{t=1}^T [P(X_{it} = x_t|X_{i(t-1)} = x_{t-1}, Z_{it}, F_i)] \\
 & \prod_{t=0}^T P(Y_{it} = y_{it}|X_{it} = x_t) f(F_i) dF_i. \tag{2}
 \end{aligned}$$

When using a non-parametric approach to control for unobservable characteristics it is usually assumed that individuals belong to different Markov chains. The most representative example of this approach is the mover-stayer model (Blumen et al., 1966). In this model, individuals are assumed to belong either to a Markov chain where no change of states is allowed or to a chain where the changing of states is unrestricted. More sophisticated non-parametric approaches are discussed by van der Pol and Langeheine (1990). In all these models, the L different Markov chains are represented by classes of the latent variable C . The general (unrestricted) form of the joint probability of such a model is:

$$\begin{aligned}
 P(Y_i = y_i|Z_i) = & \sum_{l=1}^L \sum_{x_0=1}^K \sum_{x_1=1}^K \dots \sum_{x_T=1}^K \pi_l P(X_{i0} = x_0) \\
 & \prod_{t=1}^T [P(X_{it} = x_t|X_{i(t-1)} = x_{t-1}, Z_{it}, l)] \\
 & \prod_{t=0}^T P(Y_{it} = y_{it}|X_{it} = x_t). \tag{3}
 \end{aligned}$$

where $P(X_{i0} = x_0)$ represents the initial latent state probability, and $P(X_{it} = x_t|X_{i(t-1)} = x_{t-1}, Z_{it}, l)$ are the latent transition probabilities, which depend on covariates and latent class membership. Factor π_l represents the probability of belonging to latent class (Markov chain) $C = l$. $P(Y_{it} = y_{it}|X_{it} = x_t)$ are the measurement error probabilities. Equation 3 defines a mixed hidden Markov model with L Markov chains. All the models presented above can be extended to include control variables for observed and unobserved heterogeneity in the initial state probabilities.

8.2 Correcting for Error Dependency

In the (mixed) hidden Markov models that we have presented so far, we assumed that the measurement error is uncorrelated over time. In other words, we have retained the ICE assumption, which, as mentioned before, is often highly unrealistic. Ignoring this problem and fitting models that assume error independence can lead to bias in the structural parameter estimates (Skinner and Torelli, 1993). That said, relaxing the assumption might cause poor model identifiability. It is possible to avoid this issue by using two or more indicators per time-point. Such an extended model is identifiable if the measurement error within each of the indicators is dependent over time as long as the indicators are independent of each other conditional on the true state. To illustrate, when using survey and administrative data that were collected separately, while it is likely that both sources suffer from systematic error (in the survey due to e.g. individuals having specific responding styles and in the register as certain companies are using more error-prone software), it is fair to assume that there is no correlation between the survey and administrative errors. The path diagram of Figure 8.2 describes a hidden Markov model with two indicators— Y_1, Y_2 ; X stands for the latent state and t for time.

There are various processes that may lead to the violation of the ICE assumption. First, certain population subgroups might be more likely to have an error. For example, Pavlopoulos and Vermunt (2015) investigate whether young workers have a higher measurement error when reporting the type of their employment contract in the Dutch Labour Force Survey, as they may confuse the informal agreement they have with their employer with their formal contract. Similarly, interview characteristics, such as the use of proxy interviews, might be associated with higher error probabilities (Blair et al., 2004). The path diagram of Figure 8.3 describes a hidden Markov model with two indicators (Y_1, Y_2), one of which (Y_1) is subject to dependent/systematic error that is introduced by an observed time-constant variable V . This time-constant variable affects the observed indicators without influencing the latent states or the latent transition probabilities.

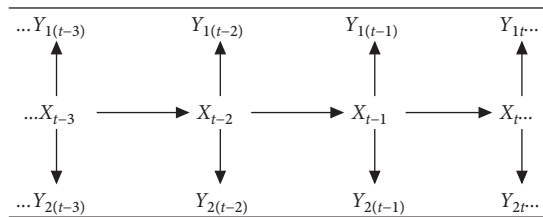


Figure 8.2 Path diagram for the HMM with two indicators.

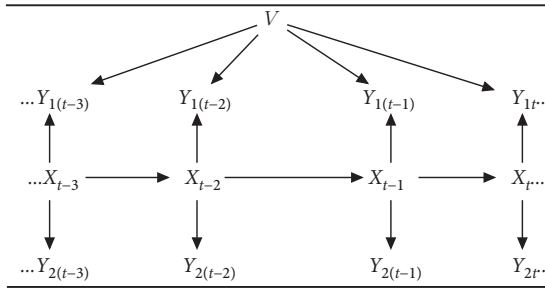


Figure 8.3 Path diagram for the HMM with two indicators and covariate-dependent errors.

The joint probability of having a particular observed path for a two-indicator HMM in which one of the indicators is subject to error dependency can be expressed as follows:

$$\begin{aligned}
 P(\mathbf{Y}_{1i} = y_{1i}, \mathbf{Y}_2 = y_{2i} | \mathbf{V}_i) &= \sum_{x_0=1}^K \sum_{x_1=1}^K \dots \sum_{x_T=1}^K P(X_{i0} = x_0) \\
 &\quad \prod_{t=1}^T P(X_{it} = x_t | X_{i(t-1)} = x_{t-1}) \\
 &\quad \prod_{t=1}^T P(Y_{1it} = y_{1t} | X_{it} = x_t, \mathbf{V}_i) \\
 &\quad \prod_{t=0}^T P(Y_{2it} = y_{2t} | X_{it} = x_t). \tag{4}
 \end{aligned}$$

As in model 3, $P(X_{i0} = x_0)$ and $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})$ represent the initial latent state probability and latent transition probabilities, respectively. $P(Y_{1it} = y_{1t} | X_{it} = x_t, \mathbf{V}_{it})$ represent the measurement error probabilities for the first indicator – Y_1 , which depend on the (time-homogenous) covariate V , and $P(Y_{2it} = y_{2t} | X_{it} = x_t)$ represent the measurement error probabilities for Y_2 , which retain the ICE assumption.

Another commonly occurring error dependency is autocorrelation. Here, we examine a simple case of such an autocorrelation—an AR(1) process where the error in time point t depends on the latent state in $t - 1$ and/or the observed stated in $t - 1$. For example, an error might be more likely to (re)occur if it was already made at the previous time-point. The path diagram of such a two-indicator HMM is illustrated in Figure 8.4.

For this two-indicator HMM which also accounts for autocorrelated error in Y_2 , the joint probability of following a specific path takes the following form:

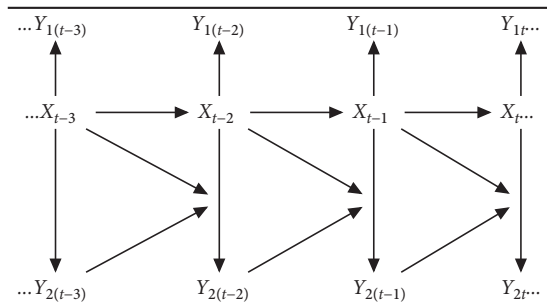


Figure 8.4 Path diagram for the HMM with two indicators and autocorrelated errors.

$$\begin{aligned}
 P(\mathbf{Y}_{1i} = \mathbf{y}_{1i}, \mathbf{Y}_{2i} = \mathbf{y}_{2i}) &= \sum_{x_0=1}^k \sum_{x_1=1}^k \dots \sum_{x_T=1}^k P(X_{i0} = x_0) \\
 &\quad \prod_{t=1}^T P(X_{it} = x_t | X_{i(t-1)} = x_{t-1}) \\
 &\quad \prod_{t=0}^T P(Y_{1it} = y_{1t} | X_{it} = x_t) \\
 &\quad \prod_{t=1}^T P(Y_{2it} = y_{2t} | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, Y_{2i(t-1)} = y_{2(t-1)}).
 \end{aligned} \tag{5}$$

The latent initial state probabilities— $P(X_{i0} = x_0)$ —and transition probabilities— $P(X_{it} = x_t | X_{i(t-1)} = x_{t-1})$ —are the same as in Equation 4. However, the measurement error probabilities for Y_1 — $P(Y_{1it} = y_{1t} | X_{it} = x_t)$ —now retain the ICE assumption while the measurement error probabilities for Y_2 — $P(Y_{2it} = y_{2t} | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, Y_{1i(t-1)} = y_{1(t-1)})$ —assume the error to be autocorrelated and depend on whether an error was made in the previous time-point. It should be noted that the models illustrated in Equations 4 and 5 can be extended to include correction for observed and unobserved heterogeneity in the ways that are explained in Section 8.1.

As a result of the introduction of autocorrelated error in Y_2 , the model has a large number of parameters. To illustrate, if the latent state X_{it} and the observed state Y_{2it} each take on three values, we will need to estimate nine different sets of error probabilities, one for each possible combination of $X_{i(t-1)}$ and $Y_{2i(t-1)}$. Therefore, it is meaningful to apply restrictions that correspond to realistic situations of error autocorrelation. To define restrictions, following Pavlopoulos and Vermunt (2015), we can define a logit model for the probability of making an error— $P(Y_{2it} = y_{2t} | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, Y_{2i(t-1)} = y_{2(t-1)})$ —that takes the

following form: $\alpha_{y_{2t}, x_t} + \beta_{y_{2t}, y_{2(t-1)}, x_t, x_{t-1}}$, whereby the first term α_{y_{2t}, x_t} represents the random component of the error and the second term $\beta_{y_{2t}, y_{2(t-1)}, x_t, x_{t-1}}$ represents the systematic component. Restrictions can be defined on the systematic component in accordance to expectations regarding the structure of the error.

It is common for individuals to have a higher probability of making an error if an error was also made in the previous time-point. In this case, the parameter $\beta_{y_{2t}, y_{2(t-1)}, x_t, x_{t-1}}$ is freed when $y_{2(t-1)} \neq x_{t-1}$ and it is restricted to 0 otherwise. Another realistic case of error dependency is the one modelled by Pavlopoulos and Vermunt (2015), where the parameter is only estimated when an error that was made in $t-1$ can be repeated in t —that is, when the same error can be made between adjacent time-points. More specifically, the authors estimate $\beta_{y_{2t}, y_{2(t-1)}, x_t, x_{t-1}}$ as a free parameter only when $y_{2t} = y_{2(t-1)} \neq x_t = x_{t-1}$; otherwise this parameter is restricted to be 0. A similar HMM with a restricted correlated error structure was applied by Manzoni et al. (2010).

8.3 Illustration Using Dutch Labour Force Data

To illustrate how HMMs can be used to correct for systematic error, we use the example of employment mobility. Specifically, we are interested in estimating the true transitions between employment with a permanent contract, a temporary contract, and other states. The analyses presented here resemble the analyses of Pavlopoulos and Vermunt (2015), where results are discussed in greater detail. We show that measurement error inflates the transition rates from temporary to permanent employment in a way that considerably distorts our image of mobility on the labour market.

8.3.1 Data

In our analysis, we use linked data from two sources: the Dutch Labour Force Survey (LFS) and the Employment Register (ER).² The LFS is conducted by Statistics Netherlands (CBS) and is a quarterly survey on individual labour market characteristics (EUROSTAT, 2019). It is representative of the Dutch population at and above the age of 15. The survey was launched in 1987 and its longitudinal component was introduced in 1999. Since 1999, respondents have been interviewed quarterly and the survey consists of five waves, which makes it possible to study

² The dataset is not freely accessible; access can only be granted by CBS following special arrangements. More information can be found here (in Dutch): <https://www.cbs.nl/nl-nl/onzediensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/stelsel-van-sociaal-statistische-bestanden-ssb->.

short-term individual developments in the labour market. The survey inquires about the respondents' current situation only. The ER (Bakker et al., 2014) is managed by the Employee Insurance Agency (UWV). It includes information needed for tax and insurance purposes and covers all employees in the Netherlands. The ER data have been collected since 2006 and are extensively used for administrative and research purposes (Mattijssen and Pavlopoulos; 2019; Skriabikova and Smits, 2019; Mattijssen et al., 2020). Both data sources are known to contain measurement error. In the LFS, the error is a result of such commonly occurring situations as misreporting of respondents and coding mistakes of interviewers. A possible cause of systematic error in the Dutch LFS is proxy interviewing as approximately 40% of the responses are collected in this way. Another possible cause is specific to the phenomenon investigated (i.e. transitions between employment contracts) and is related to age: young workers, who are more likely than older ones to have temporary contracts, are also more likely to confuse the informal agreements they have with their employers with their legal contract.

The sources of measurement error in the ER have not been thoroughly investigated. It is, however, generally known that register data are not error free (Bakker, 2012; Oberski et al., 2017). In the particular case of the Dutch ER, the data-collection process might have introduced serially correlated errors. That is, rather than submitting new information each month, employers submit information for a few consecutive months at a specific point in time (e.g. annually). If no real change occurs during those months, the values from the first month will be copied forward for the entire period. Therefore, if the original information is wrong, the error will be carried over.

The linkage of LFS and ER data (for the years 2007–15) was performed using national identification numbers, which are unique and cover the entire population. Therefore, the linkage effectiveness of this procedure was almost perfect (Mars, 2011) and it is safe to assume that the effect of linkage error is negligible. The linked dataset included all LFS respondents with the exception of cross-border workers. Both datasets provide information on the phenomenon of interest, that is, the employment contract of the individual. We obtained quarterly information on the employment contract from the LFS and monthly from the ER (Bakker, 2012).

The diverging distributions of the employment contract type according to the LFS and ER provide a first indication of the presence of measurement error in these sources. Table 8.1 presents the cross-tabulation of the distribution of the employment contract type according to the LFS and the ER, for respondents that first took the survey in the first quarter of 2007. This table comes from Pavlopoulos and Vermunt (2015). Strikingly, this table illustrates that 50.1% of the individuals that were registered as having a temporary contract in the ER reported having a permanent contract in the LFS.

The error in the measurement of the employment contract type has biasing consequences for the estimation of the transitions between different contract types

Table 8.1 Cross-tabulation of contract type according to the survey and the register.

| Register data | Survey data | | | |
|---------------|-------------|-----------|--------|--------|
| | Permanent | Temporary | Other | Total |
| Permanent | 0.944 | 0.039 | 0.017 | 1.0 |
| Temporary | 0.502 | 0.437 | 0.061 | 1.0 |
| Other | 0.081 | 0.030 | 0.889 | 1.0 |
| Total | 0.667 | 0.087 | 0.246 | 1.0 |
| Cases | 32,225 | 4,216 | 11,856 | 48,297 |

Source: Pavlopoulos and Vermunt (2015).

Note: The frequency distributions are calculated for the pooled sample. The grand total represents the number of LFS records included in our analysis in the pooled (over all time points) sample.

Table 8.2 Observed 3-month transitions in LFS and PA.

| Observed transitions from the survey data | | | | |
|---|-----------|-----------------|-----------|-------|
| | | Contract in t | | |
| | | Permanent | Temporary | Other |
| Contract in $t-3$ | Permanent | 0.981 | 0.009 | 0.010 |
| | Temporary | 0.057 | 0.889 | 0.054 |
| | Other | 0.017 | 0.035 | 0.948 |
| | Total | 0.674 | 0.089 | 0.237 |
| Observed transitions from the register data | | | | |
| | | Contract in t | | |
| | | Permanent | Temporary | Other |
| Contract in $t-3$ | Permanent | 0.967 | 0.018 | 0.015 |
| | Temporary | 0.085 | 0.860 | 0.055 |
| | Other | 0.018 | 0.036 | 0.946 |
| | Total | 0.624 | 0.128 | 0.247 |

Source: Pavlopoulos and Vermunt (2015).

Note: For both tables, these are the transition rates over a 3-month period and for 34,820 cases of our pooled sample. These cases come from LFS respondents that appear at least twice in our sample.

over time. This is illustrated in Table 8.2, which presents the 3-month transition probabilities according to the two data sources (Pavlopoulos and Vermunt, 2015). More specifically, the table shows that according to the LFS 5.7% of the individuals who have a temporary contract in month $t-3$ have a permanent contract in month t , while according to the ER this rate amounts to 8.5%.

8.3.2 Methods

For illustration and comparison purposes we run several HMMs using the LFS-ER linked dataset. In doing so, we consider seven different model specifications. In model A1, we assume that only the LFS indicator of the employment contract type is subject to measurement error. In model A2, we assume the opposite, namely that the contract type variable in the register data is measured with error while the indicator from the LFS is error-free. Model A3 assumes that both indicators are measured with error. In all three models, when an indicator is measured with error, the ICE assumption is retained, which implies that the error is random. Furthermore, all models assume the latent transitions to be time-heterogeneous (i.e. they depend on time and time squared) and they control for unobserved heterogeneity by assuming that individuals belong to one of three time-constant latent classes. No further restrictions are imposed on these latent classes (see Equation 3).

In the B-group of models, we relax the ICE assumption, using model A3 as a starting point. In model B1, the ICE assumption is relaxed only for the survey data by assuming that the response in the survey is conditional on age and proxy interviewing. This corresponds to Equation 4 and Figure 8.1. In models B2₁ and B2₂, we relax the ICE assumption for the register data only by assuming that the error in t depends on the latent employment contract in $t - 1$ and the observed contract in $t - 1$. This corresponds to Equation 5 and Figure 8.4. Models B2₁ and B2₂ differ with regards to the restrictions imposed. More specifically, in model B2₁ the coefficient for the correlated error is estimated only when an error can be repeated between two consecutive time-points, while in model B2₂ the coefficient is estimated when an error was made in the previous time-point. The restrictions for these models are explained in Section 8.2. Model B3 combines the model specifications of B1 and B2₁ and thus it relaxes the ICE assumption for both the survey and the register data.

For all these models, we obtain maximum likelihood estimates using a version of the expectation-maximization (EM) algorithm (Dempster et al., 1977) known as the forward-backward or Baum-Welch algorithm (Baum et al., 1970). In this chapter, we use an extension of this algorithm for mixed HMMs with covariates (Vermunt et al., 2008; Pavlopoulos et al., 2012). In the E-step, the expected log-likelihood for the complete data is computed. This involves the computation of the relevant marginal posterior probabilities for the latent states. In the M-step, the model parameters are updated using standard algorithms for estimating a logistic regression, where we use the marginal posterior probabilities as weights. The estimation is carried out with the use of the Latent Gold software (Vermunt and Magidson 2015).³ Missing data are handled using an ignorable likelihood (Little

³ The syntax for the data preparation and the data analysis can be found in the online appendix on the companion website (see <http://www.oup.co.uk/companion/LongitudinalData>).

and Rubin 2002). An advantage of this approach is that differing measurement intervals between the survey and administrative data are handled automatically.

8.3.3 Results

Table 8.3 presents the model fit measures for the seven models mentioned above. To examine the presence of systematic error, we compare the B-group models with those of A-group. This comparison is based on the Akaike (AIC) and Bayesian (BIC) information criteria (Agresti, 2002). Specifically, we use the AIC and BIC that are based on the log-likelihood (Vermunt and Magidson, 2016). For both measures, a lower value indicates a better model fit. These measures are defined as:

$$\begin{aligned} AIC &= -2\log\mathcal{L} + (N)np\text{ar} \\ BIC &= -2\log\mathcal{L} + 2np\text{ar} \end{aligned}$$

where $\log\mathcal{L}$ stands for the log-likelihood and $np\text{ar}$ for the number of parameters. As can be seen from the table, in our application, relaxing the ICE assumption always leads to an improvement of the model fit as all B-group models fit the data better than those of the A-group.

The comparison of the goodness of fit of the models that correct for systematic measurement error (i.e. the B-group models) provides further insights regarding the nature of the systematic error. More specifically, models B2₁ and B2₂, which correct for error autocorrelation in the register data, fit the data better than models B1 and B3, which respectively assume that the LFS is subject to systematic error (that depends on age and proxy interviewing) and that both the survey and register data suffer from systematic error. Comparing models B2₁ and B2₂ confirms that B2₁ fits the data better. This indicates that the best fitting model is an HMM that estimates an extra parameter when an error made in $t - 1$ can be repeated in t .

To interpret the results concerning systematic measurement error, we focus on the logit coefficients as well as the conditional probabilities of making an error. Table 8.4 presents the the logit coefficient estimates for model B1, which allows the error in the LFS data to depend on age and proxy interviewing. As the coefficients are conditioned on the latent contract, we estimate six coefficients for each of these two variables.

The findings indicate that, for LFS respondents whose true employment contract is either permanent or 'other', age is negatively correlated with the likelihood of making an error. In contrast, for individuals who in reality hold a temporary contract, age is positively correlated with the probability of making an error.

Table 8.3 Fit measures for seven models estimated with the matched LFS and ER data.

| Model | Log-likelihood | BIC (LL) | AIC (LL) | Parameters | L ² | df | P-value |
|--|----------------|----------|----------|------------|----------------|--------|---------|
| A1: ICE survey | -286,549 | 573,589 | 573,186 | 44 | 240,014.4 | 69,327 | < 0.001 |
| A2: ICE register | -454,195 | 908,882 | 908,479 | 44 | 575,307 | 69,327 | < 0.001 |
| A3: ICE both | -284,413 | 569,383 | 568,926 | 50 | 235,742 | 69,321 | < 0.001 |
| B1: A3 + non-ICE survey | -283,572 | 567,747 | 567,253 | 54 | 426,966 | 69,317 | < 0.001 |
| B2 ₁ : A3 + non-ICE register (same error) | -246,054 | 492,732 | 492,220 | 56 | 435,025 | 69,315 | < 0.001 |
| B2 ₂ : A3 + non-ICE register (an error) | -257,650 | 515,924 | 515,412 | 56 | 458,218 | 69,315 | < 0.001 |
| B3: A3 + non-ICE both | -245,808 | 492,374 | 491,751 | 68 | 477,358 | 69,303 | < 0.001 |

Note: Models A1, A2, and A3 specify independent classification errors (ICEs) for the survey, the register, and both datasets, respectively. Model B1 specifies the error in the survey to depend on age and proxy interview, Models B2₁ and B2₂ specify serially correlated errors in the register, while Model B3 combines the specifications of B1 and B2₁. All models are finite mixture models with three latent classes to correct for unobserved heterogeneity in the initial latent state and in the latent transition probabilities. Moreover, all models assume time heterogeneity for the latent transition probabilities. Specifically, we condition the latent transition probabilities on a linear trend for the month of the observation as well as on its square.

Table 8.4 Logit coefficients for systematic error from model B1.

| | Latent contract | Observed contract (LFS) | | |
|-------------------------------------|-----------------|-------------------------|---------------------|---------------------|
| | | Permanent | Temporary | Other |
| Effect of age | Permanent | ref | -0.0619 (0.0016) | -0.0104 (0.0024) |
| | Temporary | 0.0103 (0.0013) | ref | 0.0152 (0.0027) |
| | Other | -0.0201 (0.0017) | -0.0697 (0.0026) | ref |
| Effect of proxy interviewing | Permanent | ref | -0.0189 (0.0253) | 0.0325 (0.0394) |
| | Temporary | 0.1418 (0.0220) | ref | 0.3240 (0.0468) |
| | Other | 0.0175 (0.0292) | 0.254 (0.0429) | ref |

Note: age is a continuous variable, so this table presents the effect of an additional year.

The effect of using proxy interviews is statistically insignificant when the true/latent contract is permanent. However, when the latent contract is temporary, this type of interviewing increases the probability of making an error. For individuals who have ‘other’ type of contract, proxy interviews increase the probability of misreporting the contract as temporary. It is also worthwhile mentioning that Pankowska et al. (2019a) investigated the measurement-error effect of another interviewing technique, namely proactive dependent interviewing (PDI),⁴ and showed that in this instance the data-collection process does not have a significant effect on the systematic error. The authors also showed that there are high levels of systematic error regardless of the interviewing regime used.

Table 8.5 provides the conditional measurement-error probabilities for the register data, which were estimated using model B2₁. In addition to the latent contract in t , these probabilities also depend on the lagged observed and latent states. The shaded cells correspond to the situations where the same error that was made in $t - 1$ can be repeated t . In these cases, an additional error parameter was estimated (see Section 8.2). Table 8.5 provides strong evidence for the presence of error-autocorrelation in the ER data. For instance, if in $t - 1$ an individual was recorded as having a permanent contract, while in reality they had a temporary contract and remained employed with temporary contract in t , they have a 0.965 probability of being observed as employed with a permanent contract in t .

⁴ PDI refers to a situation where the interviewer begins a question by reminding the respondent of the answer (s)he provided in the previous interview and then inquires about her/his current employment state.

Table 8.5 Conditional probabilities of measurement error in the ER according to Model B₂₁.

| Observed contract in $t-1$ | Latent contract in t | Latent contract in $t-1$ | Observed contract in t | | |
|----------------------------|------------------------|--------------------------|--------------------------|-----------|-------|
| | | | Permanent | Temporary | Other |
| Permanent | Permanent | Permanent | 0.986 | 0.010 | 0.004 |
| Permanent | Permanent | Temporary | 0.986 | 0.010 | 0.004 |
| Permanent | Permanent | Other | 0.986 | 0.010 | 0.004 |
| Permanent | Temporary | Permanent | 0.046 | 0.930 | 0.024 |
| Permanent | Temporary | Temporary | 0.965 | 0.034 | 0.001 |
| Permanent | Temporary | Other | 0.046 | 0.930 | 0.024 |
| Permanent | Other | Permanent | 0.005 | 0.005 | 0.990 |
| Permanent | Other | Temporary | 0.005 | 0.005 | 0.990 |
| Permanent | Other | Other | 0.918 | 0.000 | 0.082 |
| Temporary | Permanent | Permanent | 0.027 | 0.973 | 0.000 |
| Temporary | Permanent | Temporary | 0.986 | 0.010 | 0.004 |
| Temporary | Permanent | Other | 0.986 | 0.010 | 0.004 |
| Temporary | Temporary | Permanent | 0.046 | 0.930 | 0.024 |
| Temporary | Temporary | Temporary | 0.046 | 0.930 | 0.024 |
| Temporary | Temporary | Other | 0.046 | 0.930 | 0.024 |
| Temporary | Other | Permanent | 0.005 | 0.005 | 0.990 |
| Temporary | Other | Temporary | 0.005 | 0.005 | 0.990 |
| Temporary | Other | Other | 0.001 | 0.856 | 0.143 |
| Other | Permanent | Permanent | 0.035 | 0.000 | 0.965 |
| Other | Permanent | Temporary | 0.986 | 0.010 | 0.004 |
| Other | Permanent | Other | 0.986 | 0.010 | 0.004 |
| Other | Temporary | Permanent | 0.046 | 0.930 | 0.024 |
| Other | Temporary | Temporary | 0.004 | 0.091 | 0.905 |
| Other | Temporary | Other | 0.046 | 0.930 | 0.024 |
| Other | Other | Permanent | 0.005 | 0.005 | 0.990 |
| Other | Other | Temporary | 0.005 | 0.005 | 0.990 |
| Other | Other | Other | 0.005 | 0.005 | 0.990 |

Note: Standard errors are always smaller than 0.0001.

8.4 Discussion

Measurement error poses a considerable threat to the analysis of categorical longitudinal phenomena. Through various processes and mechanisms this error is often systematic in nature. In this chapter, using the illustrative example of employment mobility on the Dutch labour market, we have shown that survey data can suffer from systematic errors, whereby the probability of making an error depends on such individual-level characteristics as age and/or on the data-collection processes used such as proxy interviewing. We have also shown that register data can be subject to autocorrelated errors, whereby the probability of

obtaining an error largely depends on whether the error made in the previous time-point can be repeated.

While hidden Markov models (HMMs) are an increasingly popular method used to correct for measurement error (mainly because they do not require the availability of ‘golden standard’ error-free benchmarking data), the applicability of standard, one-indicator HMMs to situations where systematic error is prevalent is rather limited, as standard models strongly depend on the ICE assumption for identifiability. This assumption entails that the observed state at time t only depends on the latent state in time t and therefore it implies that any measurement error is random. An attractive solution to this problem is the use of extended, multiple-indicator HMMs, with indicators coming from different sources. The use of independent data sources allows for the relaxation of the ICE assumption within each of these sources while maintaining the assumption of independence between them.

It is worthwhile mentioning that this method requires performing record linkage, which might lead to linkage error—a new source of bias. However, Pankowska et al. (2019b) show that HMM (structural) estimates are largely robust to this type of error. Following this line of work, future research should also look at the sensitivity and robustness of HMM estimates to other assumption violations. In particular, it would be interesting to investigate how minor changes in the data-collection processes affect the magnitude and/or structure of the error. Such changes could include, for instance, a switch to other administrative software for the recording of register data or a transition from telephone to online surveys.

References

- Agresti, A. (2002). *Categorical Data Analysis*. New York: Wiley-Interscience.
- Bakker, B. F. (2012). ‘Estimating the Validity of Administrative Variables.’ *Statistica Neerlandica*, 66(1): pp. 8–17. Available at: <http://doi.wiley.com/10.1111/j.1467-9574.2011.00504.x>.
- Bakker, B. F., Van Rooijen, J., and Van Toor, L. (2014). ‘The System of Social Statistical Datasets of Statistics Netherlands: An Integral Approach to the Production of Register-Based Social Statistics.’ *Statistical Journal of the IAOS*, 30: pp. 411–24. Available at: <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sj100803>.
- Bartolucci, F., Lupparelli, M., and Montanari, G. E. (2009). ‘Latent Markov Model for Longitudinal Binary Data: An Application to the Performance Evaluation of Nursing Homes.’ *Annals of Applied Statistics*, 3(2): pp. 611–36. Available at: <http://arxiv.org/abs/0908.2300>.
- Bassi, F., Hagenaaars, J. A., Croon, M. A., and Vermunt, J. K. (2000). ‘Estimating True Changes when Categorical Panel Data are Affected by Uncorrelated

- and Correlated Classification Errors: An Application to Unemployment Data.' *Sociological Methods and Research*, 29(2): pp. 230–68. Available at: <https://journals.sagepub.com/doi/10.1177/0049124100029002003>.
- Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (1970). 'A Maximization Technique Occuring in the Statistical Analysis of Probabilistic Functions of Markov Chains.' *The Annals of Mathematical Statistics*, 41(1): pp. 164–71. Available at: https://projecteuclid.org/download/pdf_1/euclid.aoms/1177697196.
- Biemer, P. and Trewin, D. (1997). 'A Review of Measurement Error Effects on the Analysis of Survey Data.' In *Survey Measurement and Process Quality*. Hoboken, NJ: John Wiley & Sons, chapter 27, pp. 601–32. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1002/9781118490013.ch27>.
- Billiet, J. B. and Davidov, E. (2008). 'Testing the Stability of an Acquiescence Style Factor Behind Two Interrelated Substantive Variables in a Panel Design.' *Sociological Methods & Research*, 36(4): pp. 542–62. Available at: <https://journals.sagepub.com/doi/10.1177/0049124107313901>.
- Blair, J., Bickart, G., and Barbara, M. (2004). 'Measurement Effects in Self vs. Proxy Response to Survey Questions: An Information-Processing Perspective.' In R. M. Biemer, Paul P. Groves, L. E. Lyberg, N. A. Mathiowetz, and S. Sudman (eds.), *Measurement Errors in Surveys*. NJ: John Wiley & Sons, chapter 9. Available at: <https://onlinelibrary.wiley.com/doi/10.1002/9781118150382.ch9>.
- Blumen, I., Kogan, M., and MacCarthy, P. J. (1966). 'Probability Models for Mobility.' In P. F. Lasarsfeld and N. W. Henry (eds.), *Readings in Mathematical Social Science*. Cambridge, MA: MIT Press, pp. 318–34.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). 'Maximum Likelihood from Incomplete Data via the EM Algorithm.' *Journal of the Royal Statistical Society, Series B*, 39(1): pp. 1–38. Available at: <https://web.mit.edu/6.435/www/Dempster77.pdf>.
- EUROSTAT (2019). EU Labour Force Survey—Main Features and Legal Basis. Available at: https://ec.europa.eu/eurostat/statistics-explained/index.php/EU_labour_force_survey_%E2%80%93_main_features_and_legal_basis.
- Georgiadis, M. P., Johnson, W. O., Gardner, I. A., and Singh, R. (2003). 'Correlation-Adjusted Estimation of Sensitivity and Specificity of two Diagnostic Tests.' *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 52(1): pp. 63–76. Available at: <https://www.jstor.org/stable/3592632?seq=1>.
- Hagenaars, J. A. (1990). *Categorical Longitudinal Data Log-Linear Panel, Trend and Cohort Analysis*. Newbury Park, CA: SAGE Publications.
- Hagenaars, J. A. (1988). 'Latent Structure Models with Direct Effects between Indicators: Local Dependence Models.' *Sociological Methods & Research*, 16(3): pp. 379–405. Available at: <http://smr.sagepub.com/cgi/doi/10.1177/0049124188016003002>.
- Hagenaars, J. A. (1994). 'Latent Variables in Log-Linear Models of Repeated Observations.' In A. von Eye and C. C. Clogg (eds.), *Latent Variable Analysis: Applications for Developmental Research*. Thousand Oaks, CA: SAGE Publications, pp. 329–52.
- Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. Hoboken, NJ: John Wiley & Sons. Available at: <http://doi.wiley.com/10.1002/9781119013563>.

- Manzoni, A., Vermunt, J. K., Luijkx, R., and Muffels, R. (2010). 'Memory Bias in Retrospectively Collected Employment Careers: A Model-Based Approach to Correct for Measurement Error.' *Sociological Methodology*, 40(1): pp. 39–73. Available at: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9531.2010.01230.x>.
- Mars, G. (2011). 'Cijfers over Flexibele Arbeidsrelaties—Confrontatie van Bronnen en Definities [Figures on Flexible Labour Relations—Confrontation of Sources and Definitions]'. Statistics Netherlands, report nr SAH-2011-H11.
- Mattijssen, L. and Pavlopoulos, D. (2019). 'A Multichannel Typology of Non-standard Employment Careers.' *Social Science Research*, 77: pp. 101–14. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0049089X17310293>.
- Mattijssen, L., Pavlopoulos, D., and Smits, W. (2020). 'Occupations and the Non-Standard Employment Career: How the Occupational Skill Level and Task Types Influence the Career Outcomes of Non-Standard Employment.' *Work, Employment and Society*, 34(3): pp. 495–513. Available at: <http://journals.sagepub.com/doi/10.1177/0950017020902984>.
- Oberski, D. L. (2016). 'Beyond the Number of Classes: Separating Substantive from Non-substantive Dependence in Latent Class Analysis.' *Advances in Data Analysis and Classification*, 10(2): pp. 171–82. Available at: <https://link.springer.com/article/10.1007/s11634-015-0211-0>.
- Oberski, D. L., Hagenaars, J. A., and Saris, W. E. (2015). 'The Latent Class Multitrait-Multimethod Model.' *Psychological Methods*, 20(4): pp. 422–43. Available at: <https://doi.apa.org/doiLanding?doi=10.1037%2Fa0039783>.
- Oberski, D. L., Kirchner, A., Eckman, S., and Kreuter, F. (2017). 'Evaluating the Quality of Survey and Administrative Data with Generalized Multitrait-Multimethod Models.' *Journal of the American Statistical Association*, (520): pp. 1477–89. Available at: <https://www.tandfonline.com/doi/full/10.1080/01621459.2017.1302338>.
- Paas, L. J., Vermunt, J. K., and Bijmolt, T. H. A. (2007). 'Discrete-Time Discrete-State Latent Markov Modelling for Assessing and Predicting Household Acquisitions of Financial Products.' *Journal of the Royal Statistical Society, Series A*, 170(4): pp. 955–74. Available at: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-985X.2007.00478.x>.
- Pankowska, P., Bakker, B. F., Oberski, D. L., and Pavlopoulos, D. (2019a). 'Dependent Interviewing—A Remedy or a Curse for Measurement Error in Surveys?' Paper presented at the ESRA 2019 conference.
- Pankowska, P., Bakker, B. F., Oberski, D. L., and Pavlopoulos, D. (2019b). 'How Linkage Error Affects Hidden Markov Model Estimates: A Sensitivity Analysis.' *Journal of Survey Statistics and Methodology*, 8(3): pp. 483–512. Available at: <https://academic.oup.com/jssam/article/8/3/483/5506582>.
- Pankowska, P., Bakker, B. F., Oberski, D. L., and Pavlopoulos, D. (2018). 'Reconciliation of Inconsistent Data Sources by Correction for Measurement Error: The Feasibility of Parameter Re-Use.' *Statistical Journal of the IAOS*, 34(3): pp. 317–29. Available at: <https://content.iospress.com/articles/statistical-journal-of-the-iaos/sji170368>.

- Pavlopoulos, D., Muffels, R., and Vermunt, J. K. (2012). 'How Real is Mobility Between Low Pay, High Pay and Non-employment?' *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 175(3). Available at: <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-985X.2011.01017.x>.
- Pavlopoulos, D. and Vermunt, J. K. (2015). 'Measuring Temporary Employment. Do Survey or Register Data Tell the Truth?' *Survey Methodology*, 41(1): 197–214. Available at: <https://www150.statcan.gc.ca/n1/pub/12-001-x/2015001/article/14151-eng.htm>.
- van der Pol, F. and de Leeuw, J. (1986). 'A Latent Markov Model to Correct for Measurement Error.' *Sociological Methods & Research*, 15(1–2): pp. 118–41. Available at: <http://journals.sagepub.com/doi/10.1177/0049124186015001009>.
- van der Pol, F. and Langeheine, R. (1990). 'Mixed Markov Latent Class Models.' *Sociological Methodology*, 20: pp. 213–47. Available at: <https://www.jstor.org/stable/271087?seq=1>.
- Poulsen, C. S. (1990). 'Mixed Markov and Latent Markov Modelling Applied to Brand Choice Behaviour.' *International Journal of Research in Marketing*, 7(1): pp. 5–19. Available at: <https://www.sciencedirect.com/science/article/abs/pii/016781169090028L>.
- Qu, Y. and Hagdu, A. (1997). 'Modeling Correlations Between Diagnostic Tests in Efficacy Studies with an Imperfect Reference Test.' In T. Grogioire (ed.), *Modelling Longitudinal and Spatially Correlated Data*. New York: Springer, pp. 363–71. Available at: http://link.springer.com/10.1007/978-1-4612-0699-6_32.
- Shorrocks, A. F. (1976). 'Income Mobility and the Markov Assumption.' *Economic Journal*, 86: pp. 566–78. Available at: <http://www.jstor.org/stable/2230800>.
- Skinner, C., Stuttard, N., Beissel-Durrant, G., and Jenkins, J. (2002). 'The Measurement of Low Pay in the UK Labour Force Survey.' *Oxford Bulletin of Economics and Statistics*, 64: pp. 653–76. Available at: <https://onlinelibrary.wiley.com/doi/10.1111/1468-0084.64.s.5>.
- Skinner, C. and Torelli, N. (1993). 'Measurement Error and the Estimation of Gross Flows from Longitudinal Economic Data.' *Statistica*, 53(3): pp. 391–405. Available at: <https://rivista-statistica.unibo.it/article/view/944>.
- Skriabikova, O. and Smits, W. (2019). 'Wages of Flexible Employees: the Impact of Education, Occupation and Firm Characteristics.' *Tijdschrift voor Arbeidsvraagstukken*, 35(1): pp. 26–37. Available at: https://www.tijdschriftvoorarbeidsvraagstukken.nl/inhoud/tijdschrift_artikel/TA-35-1-3/Beloning-van-flexwerkers.
- Torrance-Rynard, V. L. and Walter, S. D. (1997). 'Effects of Dependent Errors in the Assessment of Diagnostic Test Performance.' *Statistics in Medicine*, 16(19): pp. 2157–75. Available at: <https://www.onlinelibrary.wiley.com/doi/abs/10.1002/sim.4444>.
- Vacek, P. M. (1985). 'The Effect of Conditional Dependence on the Evaluation of Diagnostic Tests.' *Biometrics*, 41(4): p. 959. Available at: <https://www.jstor.org/stable/2530967>.

- Vermunt, J. K., Langeheine, R., and Böckenholt, U. (1999). 'Discrete-time Discrete-state Latent Markov Models with Time-constant and Time-varying Covariates.' *Journal of Educational and Behavioral Statistics*, 24: pp. 178–205. Available at: <https://www.statisticalinnovations.com/wp-content/uploads/LGSyntaxusersguide.pdf>.
- Vermunt, J. K. and Magidson, J. (2015). *LG-Syntax User's Guide: Manual for Latent GOLD 5.0 Syntax Module*. Belmont, MA: Statistical Innovations Inc. Available at: <https://www.statisticalinnovations.com/wp-content/uploads/LGSyntaxusersguide.pdf>.
- Vermunt, J. K. and Magidson, J. (2016). *Technical Guide for Latent GOLD 5.1: Basic, Advanced, and Syntax*. Belmont, MA: Statistical Innovations Inc. Available at: <https://www.statisticalinnovations.com/wp-content/uploads/LGtechnical.pdf>.
- Vermunt, J. K., Tran, B., and Magidson, J. (2008). 'Latent Class Models in Longitudinal Research.' In S. Menard (ed.), *Handbook of Longitudinal Research: Design, Measurement, and Analysis*. Burlington, MA: Elsevier, pp. 373–85.
- Wiggins, L. M. (1973). *Panel Analysis: Latent Probability Models for Attitude and Behavior Processes*, Amsterdam/New York: Elsevier.

