Measuring temporary employment. Do survey or register data tell the truth?

2 authors:

Dimitris Pavlopoulos
Vrije Universiteit Amsterdam

Jeroen K. Vermunt
Tilburg University

Measuring temporary employment. Do survey or register data tell the truth?

Dimitris Pavlopoulos* Jeroen K. Vermunt†

Abstract

One of the main variables in the Dutch Labour Force Survey is the variable measuring whether a respondent has a permanent or a temporary job. The aim of our study is to determine the measurement error in this variable by matching the information obtained by the longitudinal part of this survey with unique register data from the Dutch Institute for Employee Insurance. Contrary to previous approaches confronting such datasets, we take into account that also register data are not error-free and that measurement error in these data is likely to be correlated over time. More specifically, we propose the estimation of the measurement error in these two sources using an extended hidden Markov model with two observed indicators for the type of contract. Our results indicate that none of the two sources should be considered as error-free. For both indicators, we find that workers in temporary contracts are often misclassified as having a permanent contract. Particularly for the register data, we find that measurement errors are strongly autocorrelated, as, if made, they tend to repeat themselves. In contrast, when the register is correct, the probability of an error at the next time period is almost zero. Finally, we find that temporary contracts are more widespread than the Labour Force Survey suggests, while transition rates between temporary to permanent contracts are much less common than both datasets suggest.

Keywords: temporary contracts, measurement error, hidden Markov model, register data.

JEL-code: C23, J31.

*Dimitris Pavlopoulos, VU University Amsterdam, Department of Sociology, correspondence address: De Boelelaan 1081, 1081 HV Amsterdam, the Netherlands tel. +31 (0)20 59 89254, email: d.pavlopoulos@vu.nl.

†Jeroen K. Vermunt, Tilburg University, Department of Methodology and Statistics, PO Box 90153, 5000 LE Tilburg, the Netherlands, email: j.k.vermunt@tilburguniversity.edu.

1 Introduction

The issue of temporary employment is receiving increased attention in the economic and political debate. Temporary contracts allow employers to circumvent strict hiring and firing regulations (Bentolila & Bertola, 1990; Booth, 1997; Cahuc & Postel-Vinay, 2002) and some times even regulations concerning wage rigidity (OECD, 2002). Especially during economic recessions, temporary contracts are used by employers to adjust their labour force for product demand fluctuations.

The Netherlands has been a pioneer in flexible employment since the beginning of the 1990's. Contractual flexibility is an important feature of the Dutch labour market. Temporary employment rose sharply from 5.9% in 1991 to 17.1% in 2010 (OECD, 2012), while growth in temporary employment contributed 9.9 percentage points to the total employment growth from 1990 to 2000 (OECD, 2002). Employers have typically a 'minimum capacity' personnel strategy (Sels & Van Hootegem, 2001), meaning that companies employ their 'core' workers with permanent contracts and offer temporary contracts to the rest to be able to adjust in times of an economic slump.

Whereas, in the Netherlands, statistics on temporary contracts were until recently based exclusively on data from household and labour force surveys, high-quality register data has become available that may be used in conjunction with - or even replace - the survey data. The first confrontation of the two data sources revealed some severely diverging figures in the size of temporary employment. In 2009, the share of all types of temporary contracts was 15.4% according to the Labour Force Survey (LFS), while 23.6% according to the 'Polisadministratie' (PA) data, which are register data provided by the Institute for Employee Insurance (UWV) (Hilbers et al., 2011). As the size of temporary employment is very important for the design of labour market policies, Statistics Netherlands undertook the task of resolving the discrepancies between the two data sources. The results of the further investigation of the data were not very promising. Preliminary results indicate that 15.6% of those having a permanent contract according to the LFS appear to have a temporary contract according to the PA, while 18.3% of those having a temporary contract with duration shorter than one year according to the LFS appear to have a permanent contract according to the PA (Mars, 2011). Although part of the inconsistencies can be explained by the somewhat different definitions of temporary employment in the two data sources, large discrepancies remain even when both using a matched sample and selecting the cases where no definitional differences exist.

As previous research suggests, measurement error can account for the encountered inconsistencies between the survey and register data. As far as survey data are concerned, measurement error has been recognized as an important source of bias (Rodgers et al., 1993; Pischke, 1995; Bollinger, 1996; Rendtel et al., 1998; Bound et al., 2001; Biemer, 2011). Although no research exists on the error in the measurement of the contract type, research on other labour market characteristics, such as employment participation, wages, working hours, industry and occupation, indicates that survey data may contain large amounts of measurement error, which may severely bias the results of statistical analyses. For example, Biemer (2004) suggests that in the surveys of 1992-1994 of the Current Population Survey, 20.9% of the unemployed respondents were incorrectly classified to other states. Gottschalk (2005) indicates that two-thirds of the observed nominal-wage reductions without a job change were due to measurement error. Specifically, 17% of the workers report a nominal wage reduction from year to year while remaining with the same employer. However, when controlling for measurement error, yearly nominal wage reductions are faced by no more than 4-5% of the workers that remain with the same employer. Using the Panel Study of Income Dynamics (PSID) validation study, Mathiowetz (1992) suggests that company registers and survey responses in occupational classification agreed by 87.3%. Brown and Medoff (1996) find a 0.82 correlation of company registers and survey responses on the establishment size and a 0.86 on company size.

Research on measurement error in register data is clearly scarcer than on survey data. Register data are typically treated as error free and are used as a 'golden standard' when confronted with survey data. For example, most research using the PSID validation study relies on this assumption (Duncan & Hill, 1985; Rodgers et al., 1993; Bound et al., 1994; Pischke, 1995). However, there is also research showing that the 'golden standard' assumption may not be always plausible. Kapteyn and Ypma (2007) study measurement error in earnings and, although they retain the assumption that register data are error-free, they allow for errors in the matching of survey with register data. Specifically, they assume that a record in the register is identical to a record in the survey with a certain probability. They conclude that introducing this extra source of error changes the pattern of the measurement error in the survey. Abowd and Stinson (2005) compare earnings' reports from the Survey of Income and Program Participation (SIPP) and the Detailed Earnings Records (DER). Measurement error is found to be larger in the administrative DER data (20%-27%) than in the SIPP data (13%-15%). Comparing the same data sets, Gottschalk and Huynh (2010) suggest that measurement error can severely bias measures of income inequality.

The aim of the current paper is to estimate the amount of error in the measurement of contract type in the Dutch LFS. For this purpose, the survey data are matched with register data from the PA. The register data are not treated as error-free, as we model simultaneously the measurement error in both sources. We use an extended hidden Markov model with two indicators for the type of contract (temporary or permanent), each coming from one of our data sources.

The rest of the paper is organized as follows: in section 2, we elaborate further on the problem of the measurement of temporary employment in the Netherlands by presenting the relevant details on the two data sources and showing some descriptive statistics. In section 3, we present the hidden Markov model that was used in this study. Section 4 discusses the results of our analysis. The conclusions of our study are presented in section 5.

2 Description of the two data sources

The two data sources providing information on temporary contracts are the Labour Force Survey (in Dutch: *Enquête Beroepsbevolking*) administered by Statistics Netherlands (in Dutch: *Centraal Bureau voor de Statistiek* - CBS) and the 'Polisadministratie'-dataset of the Institute for Employee Insurance (UWV). The LFS is a rotating trimonthly survey on individual labour-market characteristics that is representative for the Dutch population older than 15 years of age. The survey was launched in 1987, while its longitudinal component was introduced in 1999. Since 1999, respondents are interviewed at 5 consecutive panel waves, which makes it possible to study short-term individual developments in the labour market. The information that is collected refers to the moment of the interview. The interviews are spread rather evenly within the trimester.

Errors in the measurement of the contract type in the LFS are, as is typical in surveys, the result of misreporting by respondents or mistakes in the recording of responses by interviewers. An additional error source is the use of proxy interviews. Typically, in the LFS, a single household member provides responses for all household members included in the sample, which increases the measurement error. In our LFS-sample, 40.1% of all observations refer to proxy interviews. A further possible cause of measurement error is that workers may confuse the legal employment contract with the implicit or psychological contract with their employer. Especially in younger cohorts where flexible contracts are widespread and in sectors with large job mobility and changing employment conditions, such as the health sector, workers may report that they have a permanent contract based on promises of the employer, while in reality they are employed on a temporary contract.

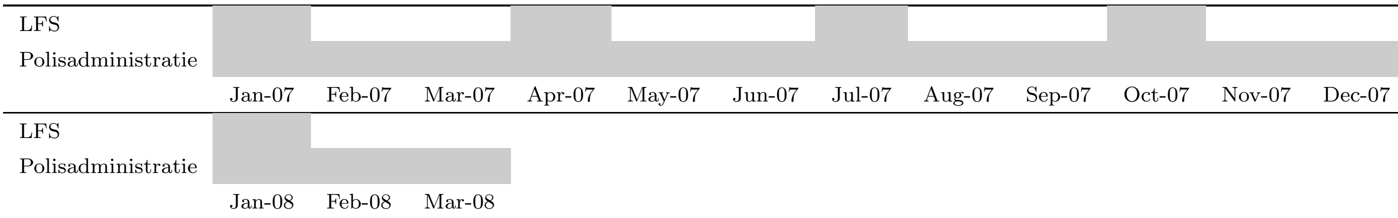
The PA is a unique register dataset containing labour market and income information for all insured workers in the Netherlands. This dataset is constructed by collecting and matching information from various sources, such as the Tax Office (in Dutch: *Belastingdienst*) - including data from individual tax-reporting statements (in Dutch: *jaaropgave*), declarations from temporary work agencies (in Dutch: *weekaanleveringen*) and the Population Register (in Dutch: *Gemeentelijke BasisAdministratie persoonsgegevens* - GBA). The PA is administered by the Dutch Institute for Employee Insurance (UWV).

The UWV has a strong interest in maintaining the high quality and accuracy of the PA as this data source is used by several governmental institutions. For example, the social security contributions, the housing allowance (in Dutch: *huurtoeslag*), and the health care allowance (in Dutch: *zorgtoeslag*) are determined using information from this dataset. To improve the data quality, the PA has undergone several revisions since 2006. There is no missing data as the submission of tax-reporting statements is compulsory for employers. However, whereas the dataset contains monthly information, employers typically submit the relevant information only once per year.[1] This may create possible mistakes for the period between two consecutive submissions, especially in the measurement of the type of contract, which is clearly not the most important variable for the users of the PA. Therefore, we may expect that if a mistake is made in the contract type, it persists till the moment that the employer submits the following report to the UWV. This means that the

[1] The moment of submission is not possible to be retrieved.

measurement error in the PA can be expected to be serially correlated.

Table 1: An illustration of our sample



NOTE: This illustrates how the rotation panel of the LFS corresponds to monthly observations from the Polisadministratie. This table refers to individuals that were interviewed every first month of the trimester. A cell that is shaded gray indicates a valid observation.

For our study, we select the LFS-respondents that were interviewed for the first time in the first trimester of 2007. Since we focus on employed individuals, we retained in the sample individuals aged from 25 to 55. After implementing the age restriction, we ended up with a sample size of 11,632 individuals. For all these individuals, the information from the LFS was matched with the monthly information from the PA by Statistics Netherlands using the social security number of individuals. The achieved matching level was 98% and all relevant inconsistencies were resolved.[2] Our final dataset has the form of a person-month file for 11,632 individuals with 15 observations corresponding to the period January 2007 - March 2008 and containing full information from the PA and partially observed information (5 observations - one response per 3 months) from the LFS. The matched dataset is illustrated in table 1. This panel dataset is unbalanced for the LFS as our survey data suffer from some attrition. More specifically, from the 11,632 individuals that

[2]The matching and the quality control was done by Statistics Netherlands.

responded to the first interview, 9,970 were left in the LFS-sample in the second interview, 9,113 for the third, 8,953 for the fourth and 8,629 for the last interview. In the PA-data for this sample there is no attrition, so the sample is fully balanced.

The variable of main interest for our study is the contract type, which takes on three possible values: permanent contract, temporary contract, and 'other'.

Table 2: Distribution of contract types according to the survey and the register

	Survey	Register
Permanent	**0.659**	**0.602**
Temporary	**0.080**	**0.123**
Other	**0.261**	**0.275**
Total	**1.0**	**1.0**
Cases	**3,887**	**11,632**

NOTE: These frequency distributions refer to the first month of the reference period, January 2007. The LFS-sample is smaller than the PA-sample as only 3,887 LFS-respondents were interviewed for the first time in January 2007. The remaining respondents were interviewed in February and March 2007.

The contract type is derived from the main job, which means that information on other jobs that individuals may hold is ignored. Individuals who are not in paid employment are classified as belonging to the 'other' state. It should be noted that the latter state is rather heterogeneous as it includes among others the categories self-employed, unemployed, and in full-time education. However, the inclusion of this state in our analysis is necessary as, in Markov models, latent states should be mutually exclusive and exhaustive.

Table 2 presents the observed contract type distribution for the first month of the reference period according to the survey and the register data. The largest discrepancies occurs in the percentages of individuals holding permanent and temporary contracts, and less in the 'other' category. According to the survey data, in January 2007, 8% of the labour force was employed with a temporary contract, whereas in the register data this percentage is quite larger (11.8%).

Table 3: Cross-tabulation of contract type according to the survey and the register

Register	**Survey data**			
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**

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 sample.

Table 3 cross-tabulates the contract type from the two sources for the pooled sample. This table confirms the large discrepancies between the two data sources reported by Statistics Netherlands. These discrepancies concern primarily individuals that are recorded as working on temporary contracts. More specifically, 50.2% of the individuals who are recorded as having a temporary contract in the register data appear to have a permanent contract in the survey. Smaller, but still existent, inconsistencies emerge for individuals that are recorded as having a permanent contract or as being in another state.

The inconsistencies in the classification of individuals that were presented in table 3 have severe implications on the transitions between the different states. Table 4 presents the 3-month transition rates for the cases with a valid observation from the LFS. This table indicates that the register data contain more transitions than the survey data. Specifically, from individuals that have a temporary contract in month $t-3$, 5.7% have a permanent contract in month t according to the survey data and 8.5% according to the register data.

3 The hidden Markov model used to estimate the measurement error in the contract type

The model we use to estimate the error in the measurement of the contract type is a hidden or latent Markov model. This model has been used for the estimation of measurement error in variables from employment surveys (see, among others, van der Pol & Langeheine, 1990; Rendtel et al., 1998; Bassi et al., 2000; Biemer & Bushery, 2000; Biemer, 2011; Pavlopoulos et al., 2012). Our application differs somewhat from these applications in that we have two measurements instead of a single one for the outcome variable; that is, the contract type from the PA and from the LFS. Other examples of applications of latent Markov models

Table 4: Observed 3-month transitions in LFS and PA

Observed transitions from the survey data

		Contract in t		
		Permanent	Temporary	Other
	Permanent	0.981	0.009	0.010
Contract	Temporary	0.057	0.889	0.054
in t-3	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	Permanent	0.967	0.018	0.015
in t-3	Temporary	0.085	0.860	0.055
	Other	0.018	0.036	0.946
	Total	0.624	0.128	0.247

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.

using multiple response variables are Langeheine (1994), Paas et al. (2007), Bartolucci et al. (2009) and Manzoni et al. (2010).

Let C_{it} and E_{it} denote the observed state of person i at time point t according to the register and the survey, respectively, where $i = 1, ..., N$ and $t = 0, ..., T$. To deal with the fact that E_{it} is observed only every third month, we use the indicator variable δ_{it} which equals 1 if the survey information is available for the month concerned and 0 otherwise. In addition to the measurements from the register and survey, the hidden Markov model contains an unobserved variable representing an individuals' true contract type at time point t. We denote this latent state by X_{it}. Note that C_{it}, E_{it}, and X_{it} can take on three values representing the categories permanent, temporary, and other. We refer to a particular category of these variables by c_t, e_t, and x_t, respectively.

$$
\begin{tikzcd}
...E(t-3) & . & . & E(t)... \\
...X(t-3) \arrow[u] \arrow[r] \arrow[d] & X(t-2) \arrow[r] \arrow[d] & X(t-1) \arrow[r] \arrow[d] & X(t)... \arrow[u] \arrow[d] \\
...C(t-3) & C(t-2) & C(t-1) & C(t)...
\end{tikzcd}
$$

Figure 1: Path diagram for the hidden Markov model with two (partially) observed indicators

The path diagram for the hidden Markov model of interest is depicted in Figure 1. For simplicity reasons, this path diagram refers only to individuals that entered the LFS-sample in a specific month. For this reason, from the four observations that are illustrated in the diagram, only those in months $t-3$ and t are non-missing for the LFS. As can be seen, the latent contract type X_{it} follows a first-order Markov process; that is, the true contract at time point t, X_{it}, is independent of the contract at time point t', $X_{it'}$, for $t' < t-1$, conditionally on the state at $t-1$, $X_{i(t-1)}$. Another assumption is that the observed states are independent of one another within and between time points, which is referred to as the local independence assumption or the assumption of independent classification errors (ICE). It can also be seen that E_{it} is observed only each third time point.

As indicated in the previous section, we use data for 15 months, which means that t runs from 0 to $T = 14$. The probability of following a certain observed path over the $T+1$

months period can be expressed as follows:

$$P(\mathbf{C}_i = \mathbf{c}_i, \mathbf{E}_i = \mathbf{e}_i) = \sum_{x_0=1}^{3}\sum_{x_1=1}^{3}...\sum_{x_T=1}^{3} P(X_{i0} = x_0)\prod_{t=1}^{T} P(X_{it} = x_t|X_{i(t-1)} = x_{t-1})$$
$$\prod_{t=0}^{T} P(C_{it} = c_t|X_{it} = x_t)\prod_{t=0}^{T} P(E_{it} = e_t|X_{it} = x_t)^{\delta_{it}} \quad (1)$$

The relevant probabilities appearing in this equation are the initial state probabilities $P(X_{i0} = x_0)$, the time-specific transition probabilities $P(X_{it} = x_t|X_{i(t-1)} = x_{t-1})$, the measurement error probabilities for the register $P(C_{it} = c_t|X_{it} = x_t)$, and the measurement error probabilities for the survey $P(E_{it} = e_t|X_{it} = x_t)$.

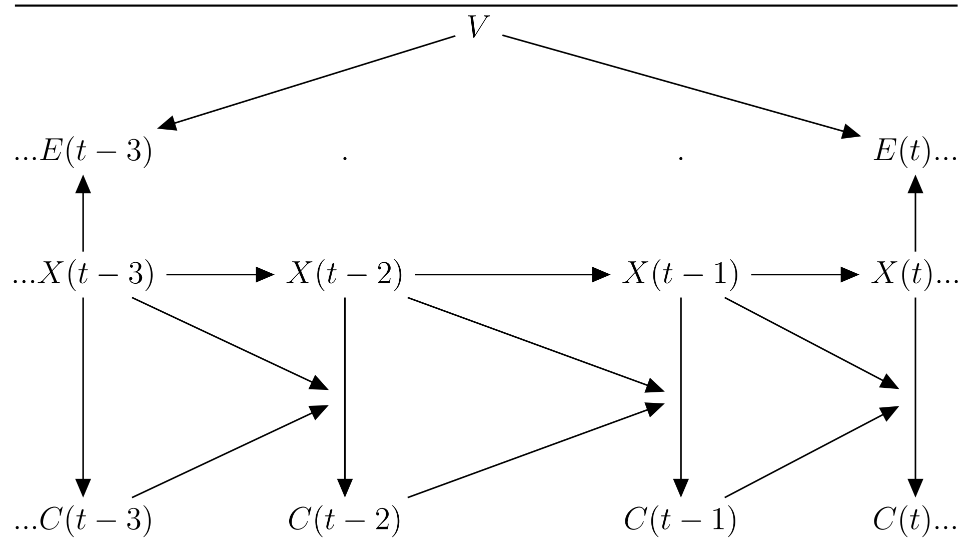


Figure 2: Path diagram for the hidden Markov model with two indicators and correlated errors

So far, we assumed that the measurement error is uncorrelated across time points - that the ICE assumption holds - which may be unrealistic in our application. First of all, as indicated in the previous section, the measurement error in the register data is likely to be serially correlated; that is, when there is a mismatch between X_{it} and C_{it}, this increases the likelihood of having the same error at time point $t + 1$. This is the result of the fact that employers make mistakes in their registers which are not adapted until a regular control takes place. In the survey data and especially since we have prospective and not retrospective data, we have no reason to justify a similar 'direct' autocorrelated error structure. However, the errors in the survey data may be correlated over time as a result of the fact that the probability of making an error may differ across groups of individuals, which is sometimes referred to as differential measurement error. Specifically, measurement error in the survey data is likely to be higher in sectors where mobility is common and ambiguity exists regarding the agreements between employers and workers, such as the health sector. Moreover, errors may be larger for young workers that care less about long-term employer relationships and therefore may have a less clear view than older respondents with respect to the formal arrangements they have on their contract. Figure 2 depicts the path diagram of the model correcting for possible heterogeneity and autocorrelation in the measurement error, where V represents the observed variables that introduce across-time correlation in the measurement error in the survey data.

Because it is also important to control for the heterogeneity in the structural part of a Markov model (Shorrocks, 1976), the model is further expanded with – possibly time-varying – observed variables affecting the initial state and latent transition probabilities, following the approach of Vermunt et al. (1999). We denote these control variables by $\mathbf{Z}_{it}$. However, these observed control variables cannot fully capture heterogeneity in the latent transition probabilities as these may be also affected by unobserved personal traits, such as motivation and ability. Following the most standard approach in the framework of hidden Markov models, we correct for unobserved heterogeneity by assuming that the population consists of a small number of latent classes with different initial state and transition probabilities (Poulsen, 1990). In this way, we avoid the unattractive distributional assumptions on the latent variable that are adopted by continuous random-effects models (Heckman & Singer, 1984; Vermunt, 1997). The number of latent classes K can be determined using model fit indices.

In our mixed hidden Markov model, the joint probability of having a particular observed state path conditionally on predictor values can be expressed as:

$$
\begin{aligned}
P(\mathbf{C}_i = \mathbf{c}_i, \mathbf{E_i} = \mathbf{e}_i | \mathbf{V}_i, \mathbf{Z}_i) \quad = \quad & \sum_{k=1}^{K} \sum_{x_0=1}^{3} \sum_{x_1=1}^{3} ... \sum_{x_T=1}^{3} \pi_k P(X_{i0} = x_0 | \mathbf{Z}_{i0}, k) \\
& \prod_{t=1}^{T} P(X_{it} = x_t | X_{i(t-1)} = x_{t-1}, \mathbf{Z}_{it}, k) \\
& P(C_{i0} = c_0 | X_{i0} = x_0) \\
& \prod_{t=1}^{T} P(C_{it} = c_t | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, C_{i(t-1)} = c_{t-1}) \\
& \prod_{t=0}^{T} P(E_{it} = e_t | X_{it} = x_t, \mathbf{V}_{it})^{\delta_{it}} \ , \qquad (2)
\end{aligned}
$$

Equation 2 specifies a finite mixture models with K latent classes to account for unobserved heterogeneity in the initial latent state and in the latent transition probabilities. π_k is the probability of belonging to the latent class k, $\mathbf{V}_{it}$ is the vector of covariates affecting the measurement error in the survey data (age and proxy interview) and $\mathbf{Z}_{it}$ is the vector of the covariates affecting the latent transition probabilities (gender, age, education and country of origin). $\mathbf{Z}_{i0}$ is the vector of the values of these covariates in the initial time point.

Compared to equation 1, in equation 2, the error probabilities in the survey data are allowed to depend on covariates ($\mathbf{V}_{it}$). The covariate effects on these error probabilities are modelled using a logit model. Moreover, the error probabilities in the register data are allowed to depend on the lagged observed and lagged true contract type. Note that $X_{i(t-1)}$ and $C_{i(t-1)}$ can take on 3 values, which implies that there are 9 (3 times 3) different sets of error probabilities in the register data, one for each possible combination of lagged observed and latent contract. Because it is not meaningful to estimate all these error probabilities freely, we used a more restricted model. More specifically, we define a logit model for $P(C_{it} = c_t | X_{it} = x_t, X_{i(t-1)} = x_{t-1}, C_{i(t-1)} = c_{t-1})$ of the form $\alpha_{c_t,x_t} + \beta_{c_t,c_{t-1},x_t,x_{t-1}}$, with $\beta_{c_t,c_{t-1},x_t,x_{t-1}}$ being a free parameter when $c_t = c_{t-1} \neq x_t = x_{t-1}$ (when the same error is made between adjacent time points) and otherwise being equal to 0. This model, which contains 6 additional parameters compared to a model without lagged effects on the misclassification probabilities, expresses that the likelihood of making a specific error depends on whether *the same error* was made at the previous time point. Similar restricted correlated error structures were used by Manzoni et al. (2010) in a latent Markov model for retrospectively collected responses.

The initial state and latent transition probabilities are also restricted using logit models, while for the latent transitions we use models with separate coefficients per origin state. The same set of covariates ($\mathbf{Z_{i0}}$ and $\mathbf{Z_{it}}$, respectively) are introduced in the models estimating the initial state and latent transition probabilities. Note that the mixed hidden Markov model described in equation 2 assumes a first-order Markov process for the true states conditionally on the individuals' covariate values and time-constant unobserved effects, but this assumption does not need to hold after marginalizing over covariate values and latent classes. A simple first-order Markov model would be inappropriate for employment transitions especially at the month level. The reason is that there is duration dependence in unemployment. For example, it is unlikely to assume that an individual that was unemployed in months 3 to 9 has the same probability of being in a particular labour market state in month 10 as an individual that was unemployed only in month 9. However, in a hidden Markov model, the bias in the classification error due to the violation of the Markov assumption is minimal. Using simulations, Biemer and Bushery (2000) show that even in cases of a severe violation of the Markov assumption, in a hidden Markov model, the bias in the estimation of classification error in unemployment does not exceed 3%.

Maximum likelihood estimates of the model parameters are obtained using a variant of the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) referred to as the forward-backward or Baum-Welch algorithm (Baum et al., 1970). We use an extension of this algorithm for mixed latent Markov models with covariates as described - among others - in Vermunt et al. (2008) and Pavlopoulos et al. (2012). In the E-step, the expected complete data log-likelihood is computed, which involves computing the relevant marginal posterior probabilities for the latent classes and latent states. In the M-step, the model parameters are updated using standard algorithms for logistic regression analysis, where the marginal posterior probabilities are used as weights. This algorithm is implemented in the program Latent GOLD (Vermunt & Magidson, 2008), which also provides standard errors for the model parameters.[3]

Missing values due to the survey construction (as respondents are interviewed once per 3 months) are Missing Completely At Random (MCAR). Missing values due to attrition in the survey are treated as Missing At Random (MAR). More specifically, following the standard manner within the ML estimation procedure, we maximize the log-likelihood for the incompletely observed data, which is obtained by integrating out the missing values. This procedure is valid under MAR.

As the LFS has a complex sampling design, the model has used the sampling weights of the survey, namely a single weight per observation. These weights are used in a pseudo ML estimation procedure, where the standard errors are adjusted for the weighting using a linearization estimator (Skinner et al., 1989). Since these are trimester weights, they are not suitable for estimating population totals at the monthly level. However, as we use information from the register for all the LFS respondents that entered the survey in a certain trimester, these weights are appropriate for the estimation of hidden Markov models.

Table 5: Fit measures for eight models estimated with the matched LFS and PA data

Model	Log-likelihood	BIC (LL)	AIC (LL)	Parameters	L^2	df	P-value
A': ICE survey	-286,814	574,118	573,716	44	240543.4	69327	1.6e-18454
A": ICE register	-454,196	908,882	908,480	44	575307.7	69327	8.5e-78021
A: ICE both	-284,413	569,384	568,926	50	235742.1	69321	4.8e-17717
B': A + non-ICE survey	-283,573	567,748	567,254	54	426966.7	69317	6.6e-50302
B": A + non-ICE register	-246,054	492,732	492,220	56	435025.8	69315	2.9e-51771
B: A + non-ICE both	-246,000	492,669	492,120	60	477741.8	69311	7.6e-59639
C': B" + predictors transitions	-245,282	491,590	490,748	92	486186.8	69279	1.8e-61222
C": B" + predictors initial & transitions	-241,990	485,140	484,189	104	479603.4	69267	4.9e-60003
C: B + predictors initial & transitions	-242,006	485,217	484,229	108	479635.2	69263	1.2e-60010

Note: Models A', A" and A specify independent classification errors (ICEs) for the survey, the register and both datasets, respectively. Model B' specifies the error in the survey to depend on age and proxy interview, Model B" specifies serially correlated errors in the register, while Model B combines these two specifications. Models C' and C" extend Model B" by introducing gender, age, education and country of origin as predictors for the transitions and for both the initial state and the transitions, respectively. Model C extends Model B by introducing the same predictors. All models are finite mixture models with 3 latent classes to correct for 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.

4 Results for the matched LFS and PA data

In total, we estimate the nine models that are presented in table 5. All these models are first order hidden Markov models with 2 indicators for the contract type as presented in the section 3. The error probabilities are time homogeneous. The (latent) transition probabilities are assumed to be time heterogeneous; that is, the transition logits are allowed

[3]Other popular programs for estimating latent Markov models are MPLUS, LEM and PANMARK.

to depend on time and time squared. These models are also finite mixture model that include three latent classes to control for unobserved heterogeneity in the initial latent state and in the latent transition probabilities. This number of latent classes was selected by comparing variants of Models B" and C with different number of latent classes.[4]

Models A', A" and A specify independent classification errors (ICEs) for the survey, the register and both datasets, respectively. Model B' specifies the error in the survey to depend on covariates V_{it} age and proxy interview, Model B" specifies serially correlated errors in the register, while Model B combines these two specifications. Models C' and C" extend Model B" by introducing predictors Z_{it} (gender, age, education and country of origin) for the transitions and for both the initial state and the transitions, respectively. Model C extends Model B by introducing the same predictors.

Table 5 presents the log-likelihood, the Bayesian Information Criterium (BIC), the Akaike Information Criterium (AIC) values and the number of parameters for nine of the models that were estimated with the matched LFS and PA data. In all models, the (latent) transition probabilities are assumed to be time heterogeneous; that is, the transition logits are allowed to depend on time and time squared.

Model A specifies that both the survey and the register data contain (independent) classification errors. As this model fits better than the restricted Models A' and A", which assume that only the survey (Model A') or only the register (Model A") contains errors, we conclude that there is evidence that both sources contain classification errors.

[4]The results of these tests are available on request.

Models B', B", and B relax the ICE assumption for the survey, the register, and both the survey and the register, respectively. More specifically, the measurement error in the survey data is allowed to depend on the respondent's age and on whether the information was obtained using a proxy interview, and the measurement error in the register data is allowed to depend on the lagged latent and observed contract type. The latter is achieved by estimating a separate set of error probabilities for repeating *the same error* across occasions. Restricted versions of Model B are estimated as well to examine whether the violation of the ICE assumption applies to the measurement error of only the survey data (Model B') or only the register data (Model B"). The fact that Model B" fits better than Models A and B' indicates that the ICE assumption should be relaxed for the indicator of the register data. Model B improves marginally the fit compared to Model B", which indicates that the ICE assumption for the survey indicator has also to be relaxed in a model without predictors for the transitions and for the initial state.

Finally, we extended Models B" and B by including covariates (gender, age, education and country of origin) in the models for the latent transition and the initial latent state probabilities (Model C" and C, respectively). Model C' is a restricted version of Model C" in which predictors are allowed to affect only the latent transition probabilities. The fact that Model C" fits better than Model B" and Model C' indicates that covariates have a significant effect on both the transitions and the initial states. The fact that, according to 2 of the 3 measures, Model C fits worse than Model C" means that the ICE assumption in the survey data should be retained in the model including predictors for the transitions and for the initial state.[5] In what follows, we present estimates derived from Model C".[6]

We investigated various alternative non-ICE models. Specifically, we studied whether the measurement error in the survey data differs for sectors with large contract and employment mobility, such as the health sector, but this did not turn out to be the case. For the register data, we looked at alternative restricted specifications for the correlated errors, but these turned out to be worse in terms of model fit than the models from Table 5.

Table 6: The size of the measurement error in the survey data according to Model C"

	Observed contract in t		
Latent contract in t	**Permanent**	**Temporary**	**Other**
Permanent	0.998	0.001	0.002
Temporary	0.125	0.832	0.042
Other	0.004	0.005	0.991

NOTE: Standard errors are always smaller than 0.0001.

Now let us look at the amount of classification error in the two data sources. According to equation 2, for the survey and register data, this is represented by the probabilities $P(E_{it} = e_{it}|X_{it} = x_t)$ and $P(C_{it} = c_{it}|X_{it} = x_t, X_{i(t-1)} = x_{t-1}, C_{i(t-1)} = c_{t-1})$, respectively. The estimates from Model C" are presented in tables 6 and 7. Specifically, table 6 shows that permanent contracts and the other state are measured very accurately in LFS as almost all individuals are correctly classified. This is indicated by the large probabilities

[5] As the results of Model C show, the size of the measurement error in the survey data changes only marginally with age and proxy interview. This is further evidence in favor of retaining the ICE assumption for the survey indicator. Actually, the estimates for the size of the measurement error in both the survey and the register data and for the latent transition probabilities are very similar between the models C, C' and C". This shows that the results of our model are robust to small model misspecifications.

[6] The estimates from Models C and C' are available on request.

in the main diagonal of the table. Some error is found for individuals that have in reality a temporary contract. 12.5% of these individuals report that they have a temporary contract, while another 4.2% report being in another state.

Table 7 reports the estimated measurement-error probabilities for the register data, which according to equation 2 depend on the lagged observed and latent state. Due to the restrictions imposed (see section 3), separate error (logit) parameters were estimated for repeating the same error between months $t-1$ and t. These situations correspond to the shaded cells in table 7. As can be seen, the measurement errors are strongly autocorrelated; that is, if an error was made in month $t-1$ and if it was possible to repeat the same error (if one remained in the same latent state), the error almost surely persisted in month t. For instance, if an individual with a permanent contract in month $t-1$ was registered mistakenly as having a temporary contract and she had still a permanent contract in month t, then she had a 0.968 probability of being wrongly registered again as having a temporary contract in t. For the other five possible errors, the probability of a persisting measurement error is somewhat lower, but it is never below 0.84.

A different picture emerges when no error is made at time point $t-1$ or when an individual changes latent state between $t-1$ and t and therefore no error repetition is possible. In these cases, register data is almost error-free. For instance, when an individual was correctly registered as having a permanent contract in month $t-1$ and has a temporary contract at t, the contract type is registered correctly as temporary at t with a probability of 0.930. In practice, this means that the initial registration of the contract is crucial for

Table 7: Conditional probabilities of measurement error in register data according to Model C”

			Observed contract in t		
Observed contract in $t-1$	**Latent contract in t**	**Latent contract in $t-1$**	Permanent	Temporary	Other
Permanent	Permanent	Permanent	0.986	0.009	0.004
Permanent	Permanent	Temporary	0.986	0.009	0.004
Permanent	Permanent	Other	0.986	0.009	0.004
Permanent	Temporary	Permanent	0.045	0.930	0.025
Permanent	Temporary	Temporary	0.968	0.032	0.001
Permanent	Temporary	Other	0.045	0.930	0.025
Permanent	Other	Permanent	0.005	0.005	0.990
Permanent	Other	Temporary	0.005	0.005	0.990
Permanent	Other	Other	0.913	0.000	0.087
Temporary	Permanent	Permanent	0.027	0.973	0.000
Temporary	Permanent	Temporary	0.986	0.009	0.004
Temporary	Permanent	Other	0.986	0.009	0.004
Temporary	Temporary	Permanent	0.045	0.930	0.025
Temporary	Temporary	Temporary	0.045	0.930	0.025
Temporary	Temporary	Other	0.045	0.930	0.025
Temporary	Other	Permanent	0.005	0.005	0.990
Temporary	Other	Temporary	0.005	0.005	0.990
Temporary	Other	Other	0.001	0.842	0.157
Other	Permanent	Permanent	0.039	0.000	0.961
Other	Permanent	Temporary	0.986	0.009	0.004
Other	Permanent	Other	0.986	0.009	0.004
Other	Temporary	Permanent	0.045	0.930	0.025
Other	Temporary	Temporary	0.005	0.099	0.896
Other	Temporary	Other	0.045	0.930	0.025
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.

the PA. If this registration is correct, then the registered contract type of the individual can be fully trusted until some true labour market change takes place. In contrast, if the contract type of the individual is initially registered wrongly, then this error will almost surely persist until the individual changes contract.

To estimate the overall amount of error in the register data, we use the posterior probability of having a particular type of latent contract at each time point. This probability is estimated for all individuals in our sample by the hidden Markov model. These estimates are quite accurate as the classification error is only 0.016. The averages of these probabilities over individuals and time points are presented in table 8. By comparing the probabilities in the main diagonal of tables 6 and 8, we see that the error is larger in the register indicator than in the survey indicator. Specifically, individuals that are truly working on a temporary contract have a 0.237 probability of being registered as having a permanent contract (0.125 in the survey data) and a 0.079 probability of being registered as being in the other state in the PA (0.042 in the survey data). There is also some classification error for individuals that are truly working on a permanent contract, as they have a 0.081 probability of being registered as temporary workers and an 0.031 probability of being registered to another state.

We are not only interested in the measurement error itself, but also in how much it affects the estimate of the size of temporary employment. Using again the average posterior probabilities of having a particular type of latent contract, we estimate the size of temporary employment in the Netherlands. In table 9, we compare the size of temporary

Table 8: The size of the measurement error in the register data according to Model C”

	Observed contract in t		
Latent contract in t	**Permanent**	**Temporary**	**Other**
Permanent	0.888	0.081	0.031
Temporary	0.237	0.684	0.079
Other	0.032	0.017	0.951

Note: These probabilities are the average posterior probabilities of having a particular type of latent contract as estimated by Model C” with classification error 0.016.

employment as estimated by the hidden Markov model with the observed distributions of the contract type from the LFS and the PA. The average posterior probability of being in a temporary contract is 10.9% and lies in between the values obtained from LFS and PA.

Table 9: The average size of temporary employment according to Model C”

	Observed		**Latent**
	Survey	**Register**	
Permanent	0.667	0.597	0.634
Temporary	0.087	0.130	0.109
Other	0.246	0.273	0.257
Cases	48,297	174,480	174,480

Note: The latent probabilities are the average posterior probabilities of having a particular type of latent contract as estimated by Model C” with classification error 0.016.

Table 10 presents the evolution of the size of temporary employment according to the two data sources and according to the hidden Markov model. This table confirms the finding that the size of temporary employment according to our model is in between that of the register data and that of the survey data. It can also be seen that in the period of reference, the proportion of temporary employed increased. The small drop that is observed in the register data in January 2008 (month 13) compared to December 2007 (month 12) may be explained by the fact that many temporary contracts end on December 31st, and that, moreover, some of these contract are converted into permanent contracts. The somewhat larger fluctuation in the size of temporary employment according to the survey data is due to the fact that respondents of the LFS are interviewed once per 3 months and thus the various monthly estimates come partly from different survey respondents.

Not only the aggregate change, but also the individual level change is important to be investigated; that is, the probability of making a transition from temporary to permanent employment and vice versa. These transition probabilities are presented in table 11. More specifically, table 11 presents the (average) latent transition probabilities obtained from Model C". The transition probabilities refer to a period of 3 months and are averaged over the 12 three-month periods in our data. If we compare the findings of table 11 with those of table 4, we see that the latent transitions probabilities are much smaller than those of both the register and the survey data. According to the latent transition probabilities, 3.2% of the individuals with a temporary contract were working with a permanent contract one year later, but according to the survey and register data, these percentages are 5.7% and 8.5%, respectively. This shows that measurement error inflate upwards the size of transition probabilities. Such an inflation would be clearly expected when errors are

Table 10: The evolution of the proportion of temporary employed for the period between January 2007 and March 2008

	Source		
Month	Survey	Register	Latent
1	0.080	0.123	0.102
2	0.082	0.124	0.103
3	0.085	0.123	0.102
4	0.084	0.128	0.103
5	0.084	0.129	0.103
6	0.090	0.129	0.104
7	0.089	0.130	0.105
8	0.087	0.131	0.106
9	0.091	0.135	0.110
10	0.087	0.134	0.112
11	0.088	0.135	0.114
12	0.091	0.135	0.114
13	0.090	0.131	0.116
14	0.089	0.131	0.118
15	0.096	0.132	0.121

NOTE: Survey data include trimonthly observations per individual, while register data include monthly observations per individual. The latent probabilities are the average posterior probabilities of having a particular type of latent contract as estimated by Model C" with classification error 0.016.

independent over time (Hagenaars, 1990, 1994). When errors are not independent over time, as in our case, the expectation is less clear as errors may either increase or reduce the transitions, depending on the nature and the size of the association. The same pattern of underestimation of stability can be observed for the permanent contract state: 98.1% and 96.7% stayed in this state according to the survey and the register data, respectively, while the true stability is 98.7%.

Table 11: Observed 12-month transitions in LFS and PA and latent transitions according to Model C

Latent transitions

		Permanent	Temporary	Other
Contract	Permanent	0.987	0.006	0.007
in t-12	Temporary	0.032	0.931	0.037
	Other	0.009	0.030	0.961
	Total	0.634	0.110	0.256

NOTE: The latent probabilities are the average posterior probabilities of having a particular type of latent contract as estimated by Model C" with classification error 0.016.

5 Conclusions

In this paper, we investigated the measurement error in the type of the employment contract in the Dutch LFS by matching its longitudinal component from 2007 and early 2008 with a unique register dataset, the PA. We applied several hidden Markov models, in which the true contract type is treated as a latent state and in which the survey and register information serve as observed indicators of an individual's true contract. We modeled the measurement error in the two data sources by taking into account that the error in the register is correlated across occasions.

Our results show that the register data contain more error than the survey data, and therefore cannot be used as a golden standard. However, the improvement of the initial registration in the register data can significantly improve their quality as measurement error in the indicator of the contract type that comes from this dataset is serially correlated.

The measurement error results into an underestimation of the percentage of individuals that are working on a temporary contract. In the LFS this percentage is 8.9%, whereas after correction for measurement error this percentage rises to 10.9%. Another effect of measurement error is that it yields severely overestimated transition probabilities. According to the LFS and PA, the transition probability between temporary to permanent employment in a 3-month period is 5.7% and 8.5%, respectively, whereas the corresponding latent transition probability is only 3.2%. This finding is particularly important for Dutch policy makers as it clearly indicates that there is much less mobility from temporary to permanent employment than originally thought.

The results of this study remain fairly stable across the model specifications that we tested. This shows that the results are robust to small model misspecifications. However, results remain somehow dependant on model assumptions. Further sensitivity tests and applications can further verify the validity of our results. Future research may focus particularly on sensitivity tests with the use of Monte Carlo simulations.

References

Abowd, J. M., & Stinson, M. H. (2005). Estimating measurement error in SIPP annual job earnings: A comparison of census survey and SSA administrative data. *Technical Paper, U.S. Census Bureau.*

Bartolucci, F., Lupparelli, M., & Montanari, G. E. (2009). Latent markov model for longitudinal binary data: An application to the performance evaluation of nursing homes. *Annals of Applied Statistiscs*, *3*(2), 611-636.

Bassi, F., Hagenaars, J. A., Croon, M. A., & Vermunt, J. K. (2000). Estimating true changes when categorical panel data are affected by uncorrelated and correlated classification errors. *Sociological Methods and Research*, *29*(2), 230-268.

Baum, L. E., Petrie, T., Soules, G., & Weiss, N. (1970). A maximization technique occuring in the statistical analysis of probabilistic functions of markov chains. *The Annals of Mathematical Statistics*, *41*(1), 164-171.

Bentolila, S., & Bertola, G. (1990). Firing costs and labour demand: How bad is eurosclerosis? *The Review of Economic Studies*, *57*(3), 381-402.

Biemer, P. (2004). An analysis of classification error for the revised Current Population Survey employment questions. *Survey Methodology*, *30*(2), 127-140.

Biemer, P. (2011). *Latent class analysis of survey error.* New Jersey: John Wiley & Sons.

Biemer, P., & Bushery, J. (2000). On the validity of markov latent class analysis for estimating classification error in labor force data. *Survey Methodology*, *26*(2), 139-152.

Bollinger, C. R. (1996). Bounding mean regressions when a binary regressor is mismeasured. *Journal of Econometrics*, *73*(2), 387-399.

Booth, A. L. (1997). An analysis of firing costs and their implications for unemployment policy. In D. J. Snower & G. de la Dehesa (Eds.), *Unemployment policy.* Cambridge: Cambridge University Press.

Bound, J., Brown, C., Duncan, G. J., & Rodgers, W. L. (1994). Evidence on the validity of cross-sectional and longitudinal labor market data. *Journal of Labor Economics*, *12*(3), 345-368.

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

Brown, C., & Medoff, J. L. (1996). Employer characteristics and work environment. *Annales D'Economie et de Statistique*, *41*, 275-298.

Cahuc, P., & Postel-Vinay, F. (2002). Temporary jobs, employment protection and labor market performance. *Labour Economics*, *9*(1), 6391.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, *39*(1), 1-38.

Duncan, G. J., & Hill, D. H. (1985). An investigation of the extent and consequences of measurement error in labor-economic survey data. *Journal of Labor Economics*, *3*(3), 508-522.

Gottschalk, P. (2005). Downward nominal-wage flexibility: Real or measurement error.

Review of Economics and Statistics, *87*(3), 556-568.

Gottschalk, P., & Huynh, M. (2010). Are earlings inequality and mobility overstated? the impact of non-classical measurement error. *Review of Economics and Statistics*, *92*(2), 302315.

Hagenaars, J. A. (1990). *Categorical longitudinal data log-linear panel, trend and cohort analysis*. Newbury Park, CA: Sage Publications.

Hagenaars, J. A. (1994). Latent variables in log-linear models of repeated observations. In A. von Eye & C. C. Clogg (Eds.), *Latent variable analysis: Applications for developmental research* (p. 329-352). Thousand Oaks, CA: Sage Publications.

Heckman, J. J., & Singer, B. L. (1984). A method for minimising the impact of distributional assumptions in econometric models for duration data. *Econometrica*, *52*(2), 271-320.

Hilbers, P., Houwing, H., & Kösters, L. (2011). De flexibele schil - overeenkomsten en verschillen tussen CBS- en UWV-cijfers [the flexible periphery - similarities and differences between CBS and UWV-data]. In B. Hermans et al. (Eds.), *Sociaaleconomische trends, 2e kwartaal 2011 [socioeconomic trends, 2nd trimester 2011]* (p. 26-33). Den Haag/Heerlen: Statistics Netherlands.

Kapteyn, A., & Ypma, J. Y. (2007). Measurement error and misclassification: A comparison of survey and register data. *Journal of Labor Economics*, *25*(3), 513-551.

Langeheine, R. (1994). Latent variable markov models. In A. von Eye & C. Clogg (Eds.), *Latent variables analysis. applications for developmental research* (p. 373-395). Thousand Oaks, California: Sage Publications.

Manzoni, A., Vermunt, J. K., Luijkx, R., & Muffels, R. (2010). Memory bias in retrospectively collected employment careers: A model-based approach to correct for measurement error. *Sociological Methodology*, *40*(1), 39-73.

Mars, G. (2011, December). *Cijfers over flexibele arbeidsrelaties - confrontatie van bronnen en defiities [figures on flexible labour relations - confrontation of sources and definitions].* Statistics Netherlands, report nr SAH-2011-H11. The Hague/Heerlen.

Mathiowetz, N. A. (1992). Errors in reports of occupations. *Public Opinion Quarterly*, *56*(3), 352-355.

OECD. (2002). *Employment outlook 2002*. Paris: Author.

OECD. (2012). *Country statistical profiles.* OECD Database: retrieved on 2012/12/16 from `http://stats.oecd.org/`.

Paas, L. J., Vermunt, J. K., & Bijmolt, T. H. (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), 955-974.

Pavlopoulos, D., Muffels, R., & Vermunt, J. K. (2012). How real is mobility between low pay, high pay and non-employment. *Journal of Royal Statistical Society, Series A*, *175*(3), 749-773.

Pischke, J.-S. (1995). Measurement error and earnings dynamics: Some estimates from the PSID validation study. *Journal of Business and Economic Statistics*, *13*(3), 305-314.

Poulsen, C. S. (1990). Mixed markov and latent markov modelling applied to brand choice behaviour. *International Journal of Research in Marketing*, *7*(1), 5-19.

Rendtel, U., Langeheine, R., & Berntsen, R. (1998). The estimation of poverty dynamics using different measurements of household income. *Review of Income and Wealth*, *44* (1), 81-98.

Rodgers, W. L., Brown, C., & Duncan, G. J. (1993). Errors in survey reports of earnings, hours worked, and hourly wages. *Journal of the American Statistical Association*, *88* (3), 345-368.

Sels, L., & Van Hootegem, G. (2001). Seeking the balance between flexibility and security: a rising issue in the low countries. *Work, Employment and Society*, *15* (2), 327-352.

Shorrocks, A. F. (1976). Income mobility and the markov assumption. *Economic Journal*, *86*, 566-578.

Skinner, C. J., Holt, D., & Smith, T. M. F. (1989). *Analysis of complex surveys.* Wiley.

van der Pol, F., & Langeheine, R. (1990). Mixed markov latent class models. *Sociological Methodology*, *20*, 213-247.

Vermunt, J. K. (1997). *Log-linear models for event histories.* London: SAGE publications.

Vermunt, J. K., Langeheine, R., & Böckenholt, U. (1999). Discrete-time discrete-state latent markov models with time-constant and time-varying covariates. *Journal of Educational and Behavioral Statistics*, *24*, 178-205.

Vermunt, J. K., & Magidson, J. (2008). *LG - Syntax user's guide: Manual for Latent GOLD 4.5 syntax module.* Belmont Massachussetts: Statistical Innovations Inc.

Vermunt, J. K., Tran, B., & Magidson, J. (2008). Latent class models in longitudinal research. In S. Menard (Ed.), *Handbook of longitudinal research: Design, measurement, and analysis* (p. 373-385). Burlington, MA: Elsevier.