Measurement Errors in Retrospective Reports of Event Histories

Marjo Pyy-Martikainen[1] and Ulrich Rendtel[2]

[1] Åbo Akademi University and Statistics Finland, Finland
e-mail: marjo.pyy-martikainen@stat.fi

[2] Freie Unversität Berlin, Germany
e-mail: rendtel@wiwiss.fu-berlin.de

Abstract

It is well known that survey reports on event histories are affected by measurement errors. Yet surprisingly little is known about their effects in an event history analysis. Making use of longitudinal register data on unemployed job seekers, we conduct a complete record-check validation study of survey reports of unemployment spells in the Finnish subset of the European Community Household Panel. The measurement errors in survey spells lead to erroneous conclusions from an event history analysis. Dummies for the heaping months of spell starts and ends, commonly used in an attempt to control for the heaping effects, have a negligible effect on estimated covariate effects and baseline hazard function. A semi-parametric Cox model did not prove to be more robust with respect to measurement errors than a fully parameterized Weibull model. Our results call for caution with respect to conclusions drawn from an event history analysis based on survey data.

1 Introduction

Longitudinal survey data is increasingly used to analyze individuals' event history processes such as fertility, poverty and labor market transitions. Event history data typically consist of information about type and timing of events occurred, as well as a set of covariates, over some time period.

It is well-known that survey reports on events are affected by measurement errors (Bound, Brown and Mathiowetz 2001). Measurement errors in events are manifested by omission or underreporting (failure to report an event), overreporting (reporting an event that did not occur) and misdating (misreporting the timing of an event) (Mathiowetz 1986; Holt, McDonald and Skinner 1991). In longitudinal surveys, misdating is typically manifested by the heaping of events at the seam between two reference periods, a phenomenon called a seam effect.

Despite the recognition of the existence of measurement errors in survey-based data on events and, thereby, on durations between them, surprisingly little is known about their effects in an event history analysis. Some evidence exists regarding duration of occupational spells (Hill 1994), time to benefit receipt or to nonemployment (Pierret 2001) and duration of benefit receipt (Jäckle 2008). The findings from these studies are mixed: in some studies, the covariate effects were attenuated by measurement errors while in others, a strengthening of covariate effects was detected. In the study by Jäckle (2008), the estimated baseline hazards based on survey data showed a somewhat weaker duration dependence than those based on register data. The studies by Hill (1994) and Pierret (2001) are not able to provide precise information about the effects of measurement errors as they are based on the comparison of two survey data sets having different data collection methods or recall periods. The study by Jäckle (2008) is a validation study but suffers from low case numbers.

Our study provides novel evidence on the existence and effects of measurement errors in event history analysis by using high-quality longitudinal register data matched at micro-level to longitudinal survey data. The combined survey-register longitudinal data enables us to 1) provide precise information on the type and magnitude of measurement errors in survey reports of durations and 2) show how measurement errors affect the results of an event history analysis.

We hypothesize that covariate estimates of duration models with a flexible baseline hazard such as the Cox proportional hazards model are less affected by measurement errors than estimates from fully parametric models. For example, it may well be that the effect of heaping of spell starts and ends is absorbed by a flexible baseline hazard. Van den Berg et al. (2006) found that covariate estimates of a Cox proportional hazards model were less biased by non-response than estimates of an exponential or a Weibull model. Also theoretical research suggests that estimates of duration models with flexible baseline hazards are more robust to mispecification than more restrictive models (van den Berg et al. 2006, van den Berg 2001). We test our hypothesis by comparing distances of survey estimates from the corresponding register estimates of a Cox proportional hazards model and a more restrictive Weibull proportional hazards model.

Sometimes dummies for the heaping months are included as covariates in an attempt to correct for the heaping effect (e.g. Hujer and Schneider 1989, Hunt 1995, Kraus and

Steiner 1998). We estimate models both with and without dummies for January and December in order to see whether the dummies give protection against measurement errors in covariate effects or in the baseline hazard.

The next section discusses the details of our research design. Section 3 presents the main results. To save space, only results describing the effect of measurement errors in an event-history analysis are presented. Section 4 concludes by discussing the research findings.

2 Research design

We use unemployment spells as study variables of interest. We conduct a complete record-check validation study of reports of unemployment spells in Finnish subset of European Community Household Panel (FI ECHP) data by making use of longitudinal register data matched at micro-level to FI ECHP survey data. We use the first five waves of FI ECHP covering the years 1996-2000. Register data are assumed to contain true, error-free information about unemployment spells. The magnitude and type of measurement errors are evaluated by comparisons of survey reports and register data. The effects of measurement errors in an unemployment duration analysis are assessed by comparing estimates based on the two data sources. No survey weights are used in the analysis. Also, no attempts are made to correct for the non-response bias. Although estimates based on both survey and register data are affected by non-response, the differences in the estimates cannot be attributed to non-response bias as both the survey and register data contain the same individuals.

In the ECHP, questions related to individual labour market histories are asked in the form of a month-by-month main activity status calendar obtained retrospectively for the preceding year. By combining calendar information from consecutive waves it is possible to construct information on individuals' unemployment histories. The respondent is asked for the main activity status for each month, given the following instructions: if a person's weekly working hours are 15 or more, an option related to employment should be chosen. If a person has had various activity statuses during a month, employment should be preferred over other statuses.

Our analysis was based on FI ECHP sample persons aged 16 or over at the beginning of 1996. Initial non-respondents were excluded because no survey information is available for them. Also temporary drop-outs were excluded because their inclusion would have posed the problem of left-censored spells. Left-censored spells are not only a source of bias in an unemployment duration analysis but they also would have artificially increased the heaping of spell starts in January. These restrictions left us with 7574 sample persons, of which 4364 responded in each of the five interviews and 3210 attrited during years 1997 to 2000. For the total respondents, information about unemployment spells is obtained for the five-year period covering years 1995-1999. For the attriters, information is obtained until the end of the year that precedes the last interview. Unemployment spells ongoing at the end of the relevant reference period were right-censored. Spells ongoing in January 1995 were dropped. The survey data has 2719 unemployment spells belonging to 1482 persons.

Validation data was obtained from administrative registers. The data contains day-level information about starting and ending dates of unemployment spells as well as the outcome of each spell. Our aim was to construct a register spell data covering, for each individual, the same time span as the survey data. For total respondents, this means using register spells ongoing between 1 January 1995 and 31 December 1999. For attriters, register spells ongoing between 1 January 1995 and the end of the year preceding the last interview were used. Spells ongoing at the end of the relevant reference period were right-censored. Spells ongoing at 1 January 1995 were dropped[1]. The register data has 6050 spells belonging to 1854 persons.

1377 persons had at least one unemployment spell according to both survey and register. 478 persons reported having no spells at all even though at least one spell was found in the register. 105 persons reported at least one spell although none was found in the register.

[1] Also spells lasting at most two days were dropped as they were not considered as "true" unemployment spells but just registrations into the records of the employment office for some legislative reason.

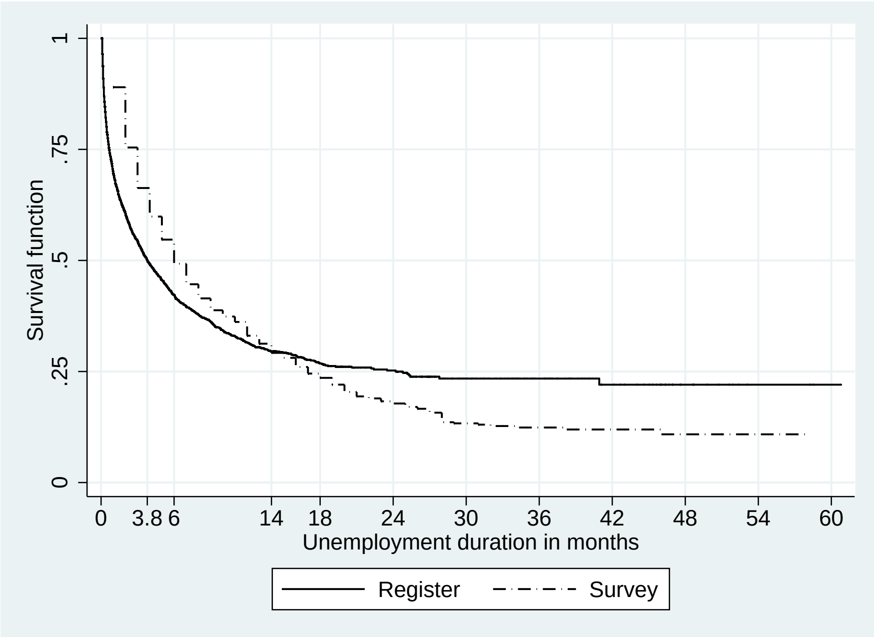

Figure 1: Kaplan-Meier survival function estimates for the register and survey data.

3 Results

3.1 Kaplan-Meier estimates of survival function

Survey spells tend to end in employment at a slower rate than register spells at durations less than 14 months. Thereafter, the situation is reversed. The median time to getting employed is 3.8 months in the register and 6 months in the survey.

3.2 Estimates of proportional hazards models

Register covariates were used in both the register and survey based models. By using the same source of covariates in the two data sets, differences in estimates can only be attributed to differences in register and survey spells. The covariates are spell-specific and they are usually measured at the end of the year preceding the start of the unemployment spell.

The magnitude and direction of measurement errors in estimated covariate effects is similar in all estimated models. The effects of age, sex and education are larger, i.e. further from 1, in the survey-based models. Having a higher education has an effect about 20 percentage points larger in the survey-based models than in the register-based models. The difference in the effect of being a female is even slightly larger. The proportion of unemployment time and being a female do not have a statistically significant effect in the register-based models whereas in the survey-based models, their effect is statistically significant at 5 % risk level. The opposite is true for the effect of getting earnings-related unemployment benefit. The statistical grouping of municipalities and the area dummies are not important determinants of the hazard of getting employed in either the register-based or the survey-based models.

Both the Cox and the Weibull models show similar effects of January and December dummies. The register spells are less likely to end in employment in December than in another months. This seasonal variation effect in spell ends is masked in the survey estimate by the heaping of spell ends in December. Survey spells beginning in January have a lower hazard of exit into employment, implying longer spell durations while the January dummy has no effect in the register data. This is an indication of backward telescoping of survey spell starts.

The inclusion of January and December dummies do not protect against heaping effects. The effect of January and December dummies on other estimated covariate effects is negligible except for the year dummies of the survey models. A more flexible baseline hazard does not give protection against measurement errors in estimated covariate effects, either. The survey based estimates from the Cox models are in fact *further* from the register estimates than those from the Weibull models.

Figure 2 shows the estimated baseline hazard function for each model. For the baseline hazard functions of the Cox proportional hazards models, a kernel smoother with Epanechnikov kernel function was applied. The hazard function estimates were calculated setting the continuous variables at their mean values and the dummy variables to zero. For all of the estimated models, the hazard of exit into employment based on survey spells exceeds at most durations the hazard based on register spells. The baseline hazards from the survey-based Weibull models are nearly constant, while the register-based baseline hazards are decreasing. The January and December dummies have virtually no effect on the baseline hazards.

Variable	1. Cox register hr (s.e)	1. Cox survey hr (s.e.)	2. Cox, time dummies register hr (s.e)	2. Cox, time dummies survey hr (s.e.)	3. Weibull register hr (s.e)	3. Weibull survey hr (s.e.)	4. Weibull, time dummies register hr (s.e)	4. Weibull, time dummies survey hr (s.e.)
Female	1.039 (0.099)	**0.798** (0.047)	1.033 (0.099)	**0.807** (0.048)	1.033 (0.102)	**0.794** (0.052)	1.028 (0.101)	**0.803** (0.053)
Age	**1.098** (0.032)	**1.129** (0.020)	**1.099** (0.032)	**1.134** (0.020)	**1.105** (0.033)	**1.131** (0.022)	**1.105** (0.033)	**1.135** (0.023)
Age squared	**0.999** (0.000)	**0.998** (0.000)	**0.999** (0.000)	**0.998** (0.000)	**0.999** (0.000)	**0.998** (0.000)	**0.999** (0.000)	**0.998** (0.000)
Upper secondary educ.	1.034 (0.118)	1.106 (0.077)	1.033 (0.119)	1.097 (0.077)	1.047 (0.124)	1.137 (0.089)	1.046 (0.124)	1.126 (0.088)
Higher education	**1.377** (0.204)	**1.558** (0.151)	**1.366** (0.203)	**1.565** (0.152)	**1.433** (0.219)	**1.690** (0.179)	**1.423** (0.217)	**1.687** (0.180)
Proportion of UE[a] time	1.002 (0.002)	**0.997** (0.001)	*1.003* (0.002)	**0.997** (0.001)	1.002 (0.002)	**0.997** (0.001)	1.002 (0.002)	**0.997** (0.001)
Semi urban municipality	1.133 (0.171)	1.114 (0.092)	1.135 (0.172)	1.107 (0.093)	1.158 (0.175)	1.114 (0.104)	1.160 (0.175)	1.098 (0.104)
Rural municipality	1.016 (0.110)	0.972 (0.074)	1.013 (0.110)	0.981 (0.075)	1.031 (0.116)	0.957 (0.080)	1.030 (0.116)	0.967 (0.081)
Southern Finland	1.345 (0.339)	1.141 (0.099)	1.354 (0.341)	1.091 (0.094)	1.380 (0.348)	*1.173* (0.112)	1.387 (0.349)	1.104 (0.104)
Eastern Finland	1.283 (0.348)	1.135 (0.118)	1.296 (0.351)	1.095 (0.113)	1.307 (0.354)	1.178 (0.134)	1.315 (0.356)	1.124 (0.126)
Central Finland	1.171 (0.291)	1.186 (0.133)	1.178 (0.292)	1.147 (0.130)	1.193 (0.297)	1.179 (0.148)	1.198 (0.298)	1.132 (0.141)
Northern Finland	*1.568* (0.414)	1.175 (0.137)	*1.574* (0.415)	1.154 (0.134)	*1.628* (0.430)	1.212 (0.154)	*1.633* (0.430)	1.186 (0.148)
Earnings-rel. UE[a] benefit	**1.319** (0.140)	1.035 (0.065)	**1.310** (0.139)	1.016 (0.063)	**1.331** (0.143)	1.046 (0.073)	**1.325** (0.143)	1.018 (0.070)
Year 1996	1.087 (0.077)	**1.151** (0.077)	1.078 (0.076)	**1.250** (0.088)	1.118 (0.083)	**1.192** (0.093)	1.110 (0.083)	**1.315** (0.108)
Year 1997	1.091 (0.091)	*1.158* (0.089)	1.081 (0.091)	**1.292** (0.105)	1.141 (0.101)	**1.219** (0.105)	1.133 (0.101)	**1.399** (0.128)
Year 1998	**1.323** (0.135)	**1.383** (0.111)	**1.304** (0.136)	**1.526** (0.129)	**1.440** (0.152)	**1.558** (0.137)	**1.423** (0.154)	**1.796** (0.164)
Year 1999	**1.321** (0.168)	1.004 (0.110)	*1.250* (0.161)	1.105 (0.147)	**1.508** (0.193)	*1.246* (0.142)	**1.434** (0.187)	**1.539** (0.199)
Begin in January			1.050 (0.079)	**0.693** (0.047)			1.036 (0.083)	**0.646** (0.049)
December			**0.536** (0.050)	1.025 (0.095)			**0.545** (0.049)	0.873 (0.077)
Weibull shape					0.616 (0.019)	1.032 (0.018)	0.615 (0.018)	1.050 (0.019)
Distance from corresponding register estimates		131.1		131.5		107.5		111.7

Estimates significant at 5% (10 %) risk level are displayed in **boldface** (*italics*).
Standard errors adjusted for clustering.

[a]UE Unemployment

Table 1: Proportional hazards models.

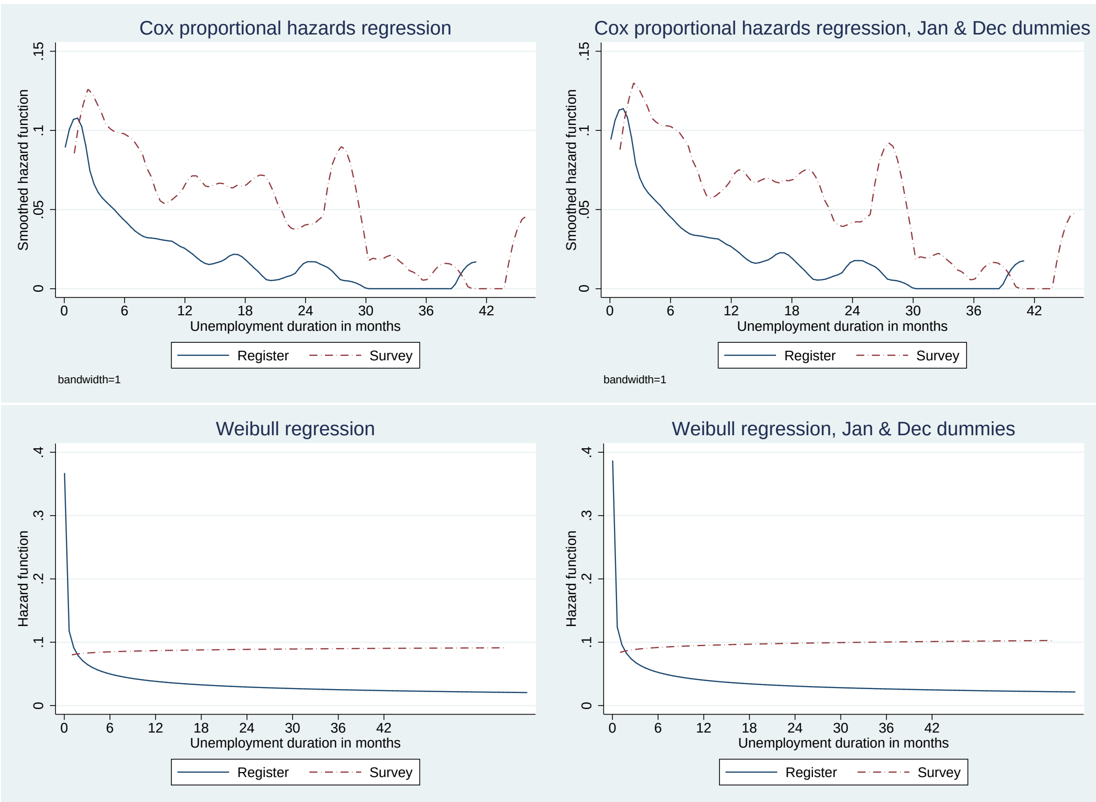

Figure 2: Baseline hazard functions

3.3 Conclusions

Measurement errors in survey-based event history data led to erroneous conclusions from an event history analysis. Our hypothesis about the robustness of a Cox proportional hazards model was not supported. The inclusion of dummies for the heaping months had a negligible effect on the estimates. These findings suggest that results from a survey-based event history analysis should be interpreted with caution. Also, no simple strategies to correct for measurement errors are likely to exist.

References

Bound, J., Brown, C. and Mathiowetz, N. (2001) Measurement Error in Survey Data. In: *Handbook of Econometrics*, (ed. J. Heckman, E. Leamer), **5**, Elsevier, Amsterdam, 3705-3843.

Hill, D. (1994) The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey. *Journal of Official Statistics*, **10**, 359-380.

Holt D., McDonald, J. and Skinner, C. (1991) The Effect of Measurement Errors on Event History Analysis. In: *Measurement Errors in Surveys*, (ed. P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, S. Sudman), Wiley, 665-685.

Hujer, R. and Schneider, H. (1989) The Analysis of Labor Market Mobility Using Panel Data. *European Economic Review*, **33**, 530-536.

Hunt, J. (1995) The Effect of Unemployment Compensation on Unemployment Duration in Germany. *Journal of Labor Economics*, **13**, 88-120.

Jäckle, A. (2008) Measurement Error and Data Collection Methods: Effects on Estimates from Event History Data. *ISER Working Paper Series* No. 2008-13.

Kraus, F. and Steiner, V. (1998) Modelling Heaping Effects in Unemployment Duration Models - With an Application to Retrospective Event Data in the German Socio-Economic Panel. In: *Jahrbücher für Nationalökonomie und Statistik*, Lucius & Lucius, Stuttgart, 550-573.

Mathiowetz, N. (1986) The Problem of Omissions and Telescoping Error: New Evidence from a Study of Unemployment. *Proceedings of the section on Survey Research Methods*, American Statistical Association.

Pierret, C (2001) Event History Data and Survey Recall. *Journal of Human Resources*, **36**, 439-466.

van den Berg, G. (2001) Duration Models: Specification, Identification and Multiple Durations. In: *Handbook of Econometrics*, (ed. J. Heckman, E. Leamer) , **5**, Elsevier, Amsterdam, 3381-3460.

van den Berg, G., Lindeboom, M., Dolton, P. (2006) Survey Non-response and the Duration of Unemployment. *Journal of Royal Statistical Society A*, **169**, 585-604.