Session Number: 2b
Session Title: Improving Estimates from Survey Data
Session Organizer(s): Holly Sutherland and Stephen Jenkins
Session Chair:

Paper Prepared for the 29th General Conference of
The International Association for Research in Income and Wealth

Joensuu, Finland, August 20 – 26, 2006

Measurement Error in the Dating of Income Receipt: Reducing Bias in Duration Models through Dependent Interviewing

Annette Jäckle

For additional information please contact:

Author Name(s) : Annette Jäckle

Author Address(es) : University of Essex, Colchester, Essex CO4 3SQ, UK
Author E-Mail(s) : aejack@essex.ac.uk
Author FAX(es) : 0044 1206 873151
Author Telephone(s) : 0044 1206 873896

Summary

Data on income sources collected in panel surveys typically display a concentration of transitions at the seam between waves of data collection. This concentration is caused by *constant wave response* (reporting receipt for 'all' or 'none' of the months in the reference period) and *wave under-reporting* (reporting receipt in some but not all relevant waves). The resulting 'seam effect' is likely to lead to errors in estimated durations of benefit receipt, attenuation of the estimated effects of explanatory factors on conditional exit probabilities and biases in estimated duration dependence. Little is however known about the nature of errors in histories from panel data, or about their effect on estimates. This paper uses benefit histories from survey reports and matched administrative records covering a four-year period to assess the extent of bias in key estimates, such as the distribution of spell lengths, their determinants and duration dependence. The paper also evaluates the effectiveness of dependent interviewing techniques, where information collected in a previous interview is used to remind the respondent of sources reported previously, or to verify that sources no longer reported have truly ended, at reducing bias.

Keywords: seam effect, constant wave response, under-reporting, validation, record check, benefits, tax credits.

Acknowledgements

This paper uses data from a project on 'Improving Survey Measurement of Income and Employment' (ISMIE), which was funded by the ESRC Research Methods Programme (H333250031). I am grateful to Peter Lynn and Steve Pudney for comments, to Mark Bryan for providing the travel-to-work unemployment data, to ISMIE colleagues Stephen Jenkins and Emanuela Sala and to other ISER colleagues for their assistance in producing the ISMIE dataset, in particular Nick Buck, Jon Burton, John Fildes, Heather Laurie, Mike Merrett, and Fran Williams. The ISMIE project is also indebted to the Information and Analysis Directorate, DWP Information Centre, especially Catherine Bundy, Katie Dodd and Judith Ridley, for implementing the data linkages. The opinions expressed and errors committed in this paper are my own.

1 Introduction

Data on income sources collected in panel surveys typically display a concentration of transitions at the seam between waves of data collection. This concentration is caused by *constant wave response* (reporting receipt for ‘all’ or ‘none’ of the months in the reference period) and *wave under-reporting* (reporting receipt in some but not all relevant waves). Under-reporting implies that spells frequently appear to end at an interview date, possibly only to be reported again in the following wave. Constant wave response implies that spells that started early in the reference period tend to be lengthened back to the start of the reference period. The resulting ‘seam effect’ is likely to lead to errors in estimated durations of benefit receipt (Boudreau 2003), attenuation of the estimated effects of explanatory factors on conditional exit probabilities (see, Hill 1994 for an examination of the effect of spurious seam transitions on job duration models) and biases in estimated duration dependence (the notion that the recipient’s behaviour may change as a result of benefit receipt, also referred to as ‘welfare dependency’).

Dependent interviewing techniques, where information collected in a previous interview is used to remind the respondent of sources reported previously, or to verify that sources no longer reported have truly ended, have been shown to reduce under-reporting (Dibbs, Hale, Loverock and Michaud 1995; Lynn, Jäckle, Jenkins and Sala 2004) and are therefore likely to improve survey estimates of spell durations and conditional exit probabilities. Constant wave reporting however remains a problem and so it is not clear to what extent dependent interviewing can improve estimates (Jäckle 2006).

Little is known about the nature of errors in histories from repeated panel data or their effects on estimates, let alone about ways of mitigating these. To my knowledge, the record check study reported by Marquis, Moore and Huggins (1990) is the only study of benefit income that was not purely cross-sectional in nature (for an extensive review of validation studies, see Bound, Brown and Mathiowetz 2001). Marquis and colleagues focused on bias in estimates of prevalence and change and provided valuable information and recommendations for survey design. The study was limited, however, in that it did not examine the effects of errors on estimates of durations or the determinants of durations, possibly because the matched records only covered a period of 8 months. Possibly as a consequence of the lack of information, analysts using duration models seem rather oblivious of seam errors and tend to either ignore them, refer to the possibility of their existence in a footnote, use only information from periods closest to the interview date or include a dummy variable to account for the seam month.

This paper makes several contributions using a unique data set, which contains benefit histories from survey reports and matched administrative records covering a 4 year period. The survey data also include benefit histories collected with dependent interviewing techniques (and matched administrative records) for a period of one and a half years. I assess the extent of bias in key estimates, such as the distribution of spell lengths, their determinants and duration dependence and the effectiveness of dependent interviewing at reducing bias.

The key findings are that the nature of errors appears to depend on the length of spells relative to the length of the interval between interviews. Spells for which the length is a multiple of the reporting period, are susceptible to wave under-reporting and reported spell lengths tend to be underestimated compared to administrative records. Shorter spells with a length similar to the reporting period are more susceptible to constant wave responses and reports tend to be lengthened. The errors in the reporting of spells appear to attenuate estimated effects of determinants of exits and of duration dependence. Dependent interviewing did not impact on constant wave reporting but did reduce wave under-reporting to some extent, improving estimated effects of determinants of exit and duration dependence.

The survey and matched administrative data are described in Section 2 and the analytical framework is set out in Section 3. The findings are presented in Section 4 and Section 5 discusses the implications for data collection and analysis.

2 The Survey and Administrative Data

The data used stem from a project on 'Improving Survey Measurement of Income and Employment' (ISMIE) funded by the UK Economic and Social Research Council Research Methods Programme. This project followed up respondents to the former low-income sub-sample of the UK European Community Household Panel Survey, who had been interviewed annually since 1994 and since 1997 jointly with the British Household Panel Survey (BHPS) activities. Respondents to the final interview in 2001 were eligible for the ISMIE survey in spring 2003 and asked for permission to obtain their benefit records from the Department for Work and Pensions (DWP), the government department in charge of administering benefits and tax credits. The consent rate for the record linkage was 77.4% (N=799) (lower for 25-39 year olds (v-shape), single person households, or if the interviewer had recorded problems with the previous interview; higher if lived in London or the South East, received means-tested benefits, moved into Income Support receipt or had had a longer previous interview (see, Jenkins, Cappellari, Lynn, Jäckle and Sala 2004). Of respondents who gave consent for matching, 74.1% (N=592) were successfully linked to DWP records. Non-matched respondents are likely to mainly be respondents without DWP records, who had not received benefits during the time frame of interest, although some non-matches due to problems with the identifying information used for the linkage cannot be excluded (Jenkins, Lynn, Jäckle and Sala 2004).

Both data sources include information about retirement related benefits (National Insurance Retirement Pension RP, Widows Benefit WB) disability related benefits and tax credits (Disability Living Allowance DLA, Incapacity Benefit IB, Attendance Allowance AA, Invalid Care Allowance ICA, Industrial Injuries Disablement Benefit IID, Severe Disablement Allowance SDA, Disabled Person's Tax Credit DDPT), earnings related benefits and tax credits (Housing Benefit HB, Income Support IS, Job Seeker's Allowance JSA, Working Families' Tax Credit WFTC) and Child Benefit CB.

Dependent interviewing. The ISMIE survey also included an experiment comparing different methods of dependent and independent interviewing. Respondents were randomly allocated to one of three treatment groups: 1) independent interviewing, the standard BHPS method of questioning without references to income sources reported previously, 2) proactive dependent interviewing, where respondents were reminded of income sources reported in the previous interview, and 3) reactive interviewing, where respondents were first asked the independent questions and then prompted if they did not report a source they had reported in the previous interview. For the comparisons with administrative records, respondents allocated to reactive dependent interviewing are included in both the independent data (but excluding any benefit information they gave in response to the reactive follow-up question) and in the dependent interviewing data (including responses to the follow-up). Respondents allocated to proactive and reactive interviewing are combined following a comparison of the effects both methods had. The administrative records are selected accordingly.

Window of observation. The survey data contain benefit histories for the period from 1st September 1996 until the final interview in spring 2003. Since dependent interviewing was only used in the final experimental survey, benefit histories based on dependent interviewing are only available for the period starting on 1st September 2000. The administrative records include benefit histories for the period from January 1999 until October 2003, although the start dates are recorded for ongoing spells that started before January 1999. For comparability I treat spells in the administrative data that were ongoing at the time of the 2003 interview as right censored.

Dates of receipt in the administrative records were exact claim dates, except for Housing Benefit, for which the exact end date was not known. In this case the end date was the ‘scan’ (data extract) date at which the claim was last observed live. In the survey, however, benefit receipt was recorded in months (the question was “For which months since <*start of reference period*> have you received <*benefit_x*>?”). For comparability I have converted the administrative records to monthly data, including both the start and end month, regardless of day of the month (assuming that respondents would report both in answers to the ‘in which months’ question in the survey).

Edits to the survey data had to be made to deal with respondents who did not take part in the survey in all waves or reports with missing dates of receipt. If the dates were missing in an intermediate wave and the spell was reported for the previous interview date and the first month of the following reference period, then the source was treated as having been received in all months and thus treated as continuous across 3 waves. Spells for which the start date was unknown due to either wave non-response or missing dates were treated as left censored, while spells for which the end date was unknown were treated as right censored. Income sources only reported in one consecutive wave without date information were dropped.

The survey reports from subsequent interviews were combined to create continuous benefit histories in the following way. The bulk of interviews took place in September/October each year, although fieldwork continued until February. The reference period for reporting of income sources went back to 1st September of the previous calendar year rather than to the previous interview date. As a result, there is usually an overlap in reporting periods from the earliest months in the current reference period and the most recent months in the previous reference period. There are at least two ways of dealing with this overlap. One is to assume that the report closest to the actual reporting period is more likely to be correct, and therefore to discard information from the later interview about the period already covered by the earlier interview. The second approach, which I have followed here, is to include all information as provided by the respondent, regardless of apparent inconsistencies (for example, if in the first interview in December the respondent did not report receipt of benefit x, and in the second wave the respondent reports receiving benefit x in all months of the reference period, so including the months September to December covered in the previous interview). The reason for including all information here, is to examine respondent reports as they are, with as little editing as possible. The ‘seam’ for spell starts is then the start month of each reference period (September of the previous year), rather than the month after the previous interview. The relevant ‘seam’ for spell ends is however the interview month.

Survey reports of child benefit and disability living allowance had to be edited for compatibility with the administrative records. The survey collected separate information about lone parent benefit and the components of disability living allowance (care, mobility, unknown), while these subcategories are not recorded separately in the DWP data. For comparability I combined the survey spells by joining overlapping spells into longer spells and dropping multiple concurrent benefit components.

The sample includes only respondents who gave permission for linkage to administrative records, although it does include respondents who did not respond in all waves (including children who turned 16 and thus became eligible for the interview after 1997). The sample also currently includes some survey under-reporters (benefit record in administrative but not survey data) and survey over-reporters/not successfully linked to administrative records.

The sample of spells includes only spells starting after January 1999 (or September 2000 for the comparison with dependent interviewing). This ‘inflow’ sample includes repeated spells (although few) and right censored spells, which can be censored at the date of the final interview or during the panel due to wave non-response and item non-response to date questions. All left censored spells are dropped, including spells that were ongoing in January 1999 (or September 2000) and spells for which the start date is unknown due to wave non-response. For comparability with the survey data, DWP spells which start during a reference period for which the respondent was not interviewed due to wave non-response are treated as left-censored and dropped.

Tables 1 to 4 show the sample sizes for the different income sources in the survey and administrative data.

3 Analytical Framework

I first examine the characteristics and distributions of spell lengths for the different income sources. The main comparisons are 1) the survey reports derived from independent interviewing (INDI) compared to administrative (DWP) records, using the inflow of spells after January 1999 and 2) the survey reports derived with dependent interviewing (DI) compared to DWP records, using the inflow of spells after September 2000. I also compare survey reports with proactive (PDI) and reactive (RDI) dependent interviewing and, since the time periods for 1) and 2) are not comparable I show some statistics for the INDI and DWP groups using a restricted sample of inflow spells after September 2000.

I then examine the determinants of spell durations and patterns of duration dependence using multivariate models, which focuses on comparisons 1) and 2).

Since the survey and administrative data in each comparison are from the same sample of respondents, standard hypotheses tests assuming independence of samples cannot be used to test the differences of estimates across data sources. I therefore assume that the administrative records are the gold standard and that any differences in the survey data can be interpreted as bias.

The following describes the different approaches; the findings are discussed in Section 4.

3.1 Summary Statistics for Spells

In the first step I compare the prevalence of spells, their mean durations and the prevalence of completed and repeated spells and of transitions onto and off benefits in seam months in the two data sources. This first descriptive analysis is carried out separately for all income sources and provides an initial indicator of differences in reporting. The comparisons between income sources are also used to guide decisions on how to combine sources for which there are very few observations.

3.2 Distribution of Spell Lengths

I then compare the distributions of spell lengths using lifetable estimates. This approach accounts for the fact that spells are recorded in months (interval censored), although in reality they can start and end on any day of the month. The estimates are based on the assumption that transitions are spread evenly over the month, that is, half the exits observed for a month have occurred by the middle of the month.

The distribution of spell lengths can be represented by the survivor function, which is the probability that a spell lasts until the end of month M_j. This is the product of the probabilities of the spell lasting until the end of each previous month up to and including the current month:

$$\hat{S}_{(j)} = \prod_{k=1}^{j} \left(\frac{n_k - d_k}{n_k} \right)$$

- where M_j are intervals of time (months), j = 1, …, J and
  - d_j is the number of exits observed during month M_j,
  - N_j is the number of spells at risk of ending at the start of the month,
  - n_j is the adjusted number of spells at risk of ending at the midpoint of the month, $n_j = N_j - d_j/2$,

I present the estimated survivor functions for the different comparison groups graphically.

3.3 Determinants of Spell Durations and Duration Dependence

The third step is then to examine determinants of spell durations and estimates of duration dependence for the two data sources.

3.3.1 Factors Associated with Moves onto and off Benefits

The explanatory variables were selected based on eligibility criteria for the different income sources and informed by other studies (for example, Ashworth, Walker and Trinder 1997; Blank 1989; Hoynes and MaCurdy 1994; Long 1990; O'Neill, Bassi and Wolf 1987; Ruggles 1989). The variables derived from the survey data were merged to the spell information from both the survey and the administrative data. Covariates are time-varying unless stated otherwise and include information about the following:

Personal characteristics and social background: gender (fixed), age, martial status (date of change in legal marital status known, but changes in whether partner living in benefit unit only observed at interview), number of own children under 16 in the household and age of the youngest (birthdates known, but whether children left household only known at interview), highest educational qualification (at date of each interview) and region of residence (date moved known as long as only moved once since previous interview). Ethnic group is not controlled for, since 98% of respondents included in this analysis were of white origin.

Factors related to eligibility for health related benefits: whether long term sick or reported chronic health problems (at interview).

Factors related to eligibility for income related benefits: current labour market activity and number of months unemployed to date during the panel period, housing tenure (at interview) as a measure of wealth relevant for the means-tested benefits. No measure of wages (either previous or predicted) is included, since the factors typically used to predict income are already included in their own right. The partner's employment status (at the time of interview) is included, since eligibility for some sources is determined at the level of the benefit unit. This information is available for all partners; the interviews would provide more detailed information but are constrained by non-respondent partners. Outside opportunities are captured by local unemployment rates (at the date of interview). The local unemployment rate is based on the travel-to-work areas (using 1998 boundaries), of which there are about 300 in the UK. For 1997-2000 the rate is the proportion of unemployed in the labour force. Because the Office for National Statistics (ONS) discontinued the labour force measure, the 2001 and 2003 rates are based on the proportion of claimants in the resident working age population from the 2001 census, which tends to produce lower unemployment rates.

3.3.2 Specification of Multivariate Duration Models

The duration models are estimated using a discrete time proportional hazards specification (cloglog), firstly including only substantive explanatory variables and then including different specifications for the baseline hazard. I first fit a fully flexible model, including a dummy for every month in which an exit is observed, excluding the first month as the reference category. This non-parametric specification should give an idea of the pattern of duration dependence. If the number of exits observed per month is small, it will however lead to a loss in precision. I therefore also test two parametric specifications, the discrete time equivalent of the continuous time Weibull model and a polynomial specification.

The analysis so far does not account for unobserved heterogeneity, which may lead to downwards biased estimates of duration dependence and to bias in estimated effects of covariates (Kiefer 1988). The principal objective of this study, the comparison of estimates from survey and administrative data, should not be affected, but this remains to be tested. All models do, however allow for clustering to adjust for multiple observations per sample member, since they are likely to contribute more than one month to the analysis and possibly more than one spell, especially when different income sources are grouped for the analysis.

The coefficients of the models are presented, as well as predicted hazard rates to compare the estimated patterns of duration dependence estimated from the survey and administrative data.

4 Results

4.1 Distribution of Spells

4.1.1 Independent Interviewing Compared to Administrative Records

Contrary to my expectations, the survey reports do not contain more repeated spells than the administrative records (I had expected that wave under-reporting would lead to spells being ‘chopped up’ into disjoint shorter spells). Overall, the average number of spells is slightly higher in the administrative records at 1.33 compared to 1.11 in the survey data (Tables 1 and 2). For the individual sources, the average number of reports of the health related benefits IID, AA and DLA are higher in the survey data (1.25, 1.14 and 1.06 compared to 1.00 in the administrative records), while the average number of means-tested and unemployment related spells of WFTC, HB and JSA is higher in the administrative records (1.33, 1.76 and 1.59) compared to the survey data (1.13, 1.13 and 1.26). Note that the averages do not include respondents with zero spells of a given type.

Transitions off benefits at the seam are also less prevalent in the survey data than I had expected: overall 27.0% of survey spells end at a seam (again I had expected that more spells would be ‘cut off’ at a seam due to wave under-reporting). Transitions onto spells at the seam are, however, much more prevalent than I had expected: overall 52.5% of survey spells start at a seam and 17.4% of spells start and end at a seam. In the administrative data, only 5.1% of spells start at a seam and 3.7% end at a seam. These proportions are lower than expected if transitions were uniformly distributed. If this were the case, 6.12% of transitions would be observed at a seam. (Calculation: 3/49*100=6.12; 49 month-pairs of which 3 are seams.) Compared to the expected transition rates, the proportion of seam starts in the survey data is inflated by 8.6; the proportion of seam ends is inflated by 4.4.

Judging by the characteristics presented in Tables 1 and 2, the income sources seem to roughly fall into two groups. AA, DLA and CB have larger average numbers of spells, higher proportions of seam starts (inflated by more than factor 10 compared to the expected rate) and ends, larger proportions of completed spells and shorter average durations in the survey than the administrative records. This suggests that spells of these types are indeed ‘chopped off’ by wave under-reporting. One aspect these benefits have in common is that the average spell length in the administrative data (including right censored spells) is between 24.5 and 30.4 months, so roughly the length of two reference periods or longer in the case of CB. The only other benefit with such a long average duration is retirement pension. Reporting of pension income however appears to be much closer to the administrative records, although the proportion of seam starts is also inflated by factor 5.

For the remainder sources the mean spell durations according to the administrative records are around the length of one reference period for IB, IS and HB or shorter for JSA. Mean spell durations for these sources (except HB) are consistently lengthened by .5 to 1.5 months in the survey data, if right censored spells are included. There are also fewer repeated spells and seam starts (although still inflated by more than factor 5), while the inflation of seam ends is similar to that for the longer spells.

WFTC appears to be in between the two groups. The mean duration is slightly longer than one reference period (18.8 months) and like the longer spells under-reported in the survey (17.4). The proportion of seam transitions is however more comparable to that of the shorter spells. HB behaves somewhat strangely: the high proportion of seam transitions is comparable to that of the longer spell types, although the survey data contain a lower mean number of HB spells and the mean duration is lengthened by 7.2 months. This may in part be due to the fact that the administrative HB data are not very reliable and the exact end dates are not known. End dates are substituted with the date of the last scan at which a claim was observed, which is likely to lead to an underestimation of HB durations in the administrative records.

The graphs of the empirical survivor functions (Figure 1) also suggest this distinction between long and short term benefit types. DLA, AA, CB and RP are the only benefits for which median spell duration is not observed, because the window of observation is not long enough for 50% of spells to end. (In Table 2 the mean spell duration, including censored spells, for these sources is around 25 months in the administrative data, but because there are no exits, the estimated proportion surviving is 1 for DLA, AA and RP). In comparison, the median spell duration is between about 16 and 20 months for IB, IS and WFTC, just below 12 months for HB and around 4 months for JSA.

For all the sort-term benefit types (except HB), the survivor functions based on the survey data trace those based on the administrative records quite closely. Although the mean spell durations tended to be lengthened for these sources when the right censored spells were included, median durations (excluding right censored spells) are only longer in the survey data for IB and HB. RP also maps the administrative data closely.

For the other long-term spells, especially DLA and AA, the differences are marked: while there are no exits in the administrative data, close to 50% of AA spells end by around 12 months and around 40% of DLA spells end by around 30 months. These differences are mainly caused by sharp increases in the proportion of exits roughly at multiples of the reference period, probably caused by correct reporting of ongoing spells in one wave, followed by under-reporting in the next wave. These kinks were apparent as the high proportion of seam transitions in Table 1 and are also visible for the shorter spell types, although less pronounced.

4.1.2 Dependent Interviewing Compared to Administrative Records

Before comparing the reports obtained with dependent interviewing with the administrative records, I first examined the extent to which the proactive and reactive data were comparable and whether they could indeed be grouped. The percentage of seam *starts* was similar with both types of DI at around 27%, although the rate expected with a uniform distribution of transitions was just 3.23% (calculated as 1/31*100=3.23; 31 month pairs of which one is a seam). The inflation of seam starts compared to the expected rate corresponded to the inflation in the INDI data by around 8.4 and suggests that neither DI method reduced the lengthening of spells back to the start of the reference period. This makes sense, since DI was asymmetric in that it did not remind respondents that they had *not* reported a source in the previous interview. The percentage of seam *ends* with PDI was 2.2%, which was lower than the expected rate. With RDI the rate was higher at 9.3%, but the inflation by 2.9 compared to the expected rate was still lower than with INDI (4.4). This suggests that PDI reduced wave under-reporting more than RDI did, for which Table 5 in Lynn et al. (2004) also provides some limited evidence.

Regardless of these differences, the estimated survivor functions did not differ significantly according to likelihood ratio tests or log-rank tests of homogeneity (at the 5% level) across the two samples. The sample sizes were however very small for some sources, providing little power to detect differences. Estimating separate survivor functions (not shown) did not yield a clear pattern of which method tracked the administrative records better. PDI produced estimates closer to the administrative data for DLA and JSA, while RDI appeared better for IB and AA.

In the combined DI samples, survey respondents did not report any repeated spells of health related sources. The mean number of earnings related sources were also lower than in the administrative data (Tables 2 and 3). Correspondingly, the mean spell durations (including right censored spells) were longer with DI (16.5 months) than in the records (9.6 months) for all income sources except IID and JSA. This suggests that DI did reduce under-reporting of sources, and thereby the occurrence of exits at the seam, but that constant wave reporting led to consecutive spells being reported as one (long) spell. For example, a respondent may have (correctly) reported *current* receipt of IS at two consecutive interviews, but erroneously reported receipt for *all* months, even though the spell in progress at the previous interview ended early in the reference period and the current spell started thereafter.

The survivor functions (Figure 1) show that DI improved estimated distributions of spell lengths for those sources for which INDI lead to under-estimated spell durations: the long-term health related benefits DLA and AA as well as CB. The kinks around month 12 are however still visible, suggesting that some under-reporting remained with DI (this is also reported by Lynn et al. 2004). For the sources where INDI lead to an over-estimation of mean spell durations, that is, the shorter term earnings related sources HB, IS and WFTC, but also IB, the over-estimation was exacerbated with DI. For these sources, the survivor functions based on the INDI data tracked the administrative records well and DI led to worse estimates.

4.2 Determinants of Spell Durations and Duration Dependence

Due to the small numbers of spells for each income source, it is necessary to group sources for the multivariate duration analysis. Ideally, the grouped sources should have similar characteristics in terms of the distribution of durations, the factors related to exit probabilities and the nature of reporting errors. The previous section showed that sources roughly fall into two categories depending on their durations. The factors associated with moves onto and off income sources are likely to differ between those sources related to health and those related to earnings. In the following analysis I have therefore grouped the health and disability related benefits, which are also longer term and subject to wave under-reporting on the one hand, and the earnings related benefits on the other hand, which tend to be shorter term and subject to constant wave responding leading to spells being lengthened. Retirement Pension and Child Benefit are not included, since they are universal benefits and exit should only be related to death or age of the youngest child.

Housing Benefit had to be dropped from the analysis, because the errors in spell durations in the administrative records caused by the lack of exact end dates distorted the multivariate estimates. For the comparison with DI none of the determinants were significant in the models based on the administrative data. The findings for the survey data were similar with and without the inclusion of Housing Benefit.

The last two panels of Figure 1 plot the estimated survivor functions for the combined sources. For the health related sources, the combined INDI data under-estimated spell durations, while DI over-estimated spell durations. Likelihood-ratio tests of homogeneity between the survey and administrative data, however, suggest that the survey and administrative estimates are no different (INDI compared to records: $\chi^2(1)=2.06$, P=0.152; DI compared to records: $\chi^2(1)=1.84$, P=0.175). For the earnings related benefits, both INDI and DI lead to over-estimates of spell durarations (INDI compared to records: $\chi^2(1)=3.53$, P=0.060; DI compared to records: $\chi^2(1)=6.94$, P=0.008).

Table 5 shows the summary statistics for the explanatory variables included in the models, measured in the first month of each spell from the survey data. For the local unemployment rate the mean across all spell months is reported. Both the INDI and the DI samples were predominantly female, with low qualifications (just under one third of recipients of earnings related benefits did not have any qualifications, compared to up to two thirds of recipients of health related benefits), around half were married or cohabiting and between 10 and 20% lived in London or the South East. Among recipients of earnings related benefits, 28% had a spouse who was in work, the mean number of own children in the household aged younger than 16 was one and their average age around 3, 18 to 25% were owner occupiers, 60% were active in the labour force, either in work or looking for work, respondents experienced on average around 7 months of unemployment during the panel period and the local unemployment rate averaged just over 3%. Among recipients of health related benefits, around 90% reported chronic health problems and around 35% reported their labour market activity status as being long term sick.

Tables 6 to 9 report the results from the duration models, comparing different specifications, including only the explanatory variables (model 1), or including a fully flexible specification for the baseline hazard (model 2), a polynomial specification (model 3) or a Weibull specification (model 4). The first two tables report the results for income and health related spells from the independent survey data and administrative records. The latter tables report the results for both sets of income sources from the dependent interviewing and administrative data. For all models, the inclusion of time in any form did not alter the estimated effects of the remainder determinants.

Figures 2 and 3 then present the predicted hazard rates for models 2 to 4, at the means of the continuous variables in the corresponding sample of records and setting the binary indicators equal to their most prevalent value. The resulting predictions are for females, without qualifications, married or cohabiting, who do not have a spouse in work, with 1 child, not living in London or the South East, not owner occupier and neither in work nor looking for work. The values of age, age of youngest child, local unemployment rate, months unemployed during the length of the panel and the income source dummies are set to the sample mean based on the administrative data.

4.2.1 Independent Interviewing Compared to Administrative Records

For earnings related benefit spells (Table 6), the estimates from the administrative records suggest a larger hazard rate, and hence shorter spells, for those in work, married or cohabiting and with a spouse in work, while exit hazards decrease with the number of children. The polynomial model suggests that the hazard changes non monotonically with time. Judging by the Akaike Information Criterion, AIC, the fully flexible model fits the records best, followed by the polynomial specification. For the survey data, there is no clear distinction in fit between models. The hazard also increased for those in work, although the effect was smaller and less significant. The remainder determinants were however not significant and neither were the polynomial effects of time. Instead, exit hazards were significantly reduced for respondents without qualifications.

The first three graphs in Figure 2 show that the predicted hazard rates based on the records is non-monotonic, with a sharp rise and fall during the first 12 months. This pattern of duration dependence is reflected in the polynomial specification, but not the Weibull which does not allow for non-monotonic changes in hazard rates. In the INDI survey data, the hazard is relatively constant at lower levels than the record estimates, except for a large spike at month 13, which is roughly the date of the first seam. As a result, the polynomial specification does not reflect the non-monotonic duration dependence. Instead the prediction suggests a slightly monotonically decreasing hazard, similar to the Weibull prediction.

These findings suggest that errors in the reporting of spells (which led to the lengthening of earnings related spells observed in the descriptive analysis in section 4.1) attenuate both the effects of substantive explanatory variables and also the pattern of duration dependence. The survey estimates also suggest that the reporting errors are related to levels of qualifications. (Note: there will also be errors in the reporting of the explanatory factors, but since the variables derived from the survey were also used as covariates for the administrative spells, these should not lead to differences in the estimated effects in this analysis.)

For the health related benefits similar conclusions hold. According to the administrative records exit hazards increase with age at a decreasing rate and decrease for the long-term sick. The results again suggest significant non-linear effects of time. For both the records and the survey, the fully flexible specification fits the data best, while all other specifications have similar AIC values.

In the survey data age has no effect, the effect being long term sick is both smaller and weaker and there is no significant duration dependence. The predicted hazard rates in the first three graphs of Figure 3, show a non-monotonic pattern in the record data, with hazards increasing after month 12, falling and then increasing again after month 34. This pattern is reflected in the polynomial predictions. The survey data again display a large spike in month 13, where the hazard rate rises to about 0.22. In comparison, the seam spike in the INDI data for earnings related spells is around 0.08. Nonetheless, with the polynomial specification the survey data match the non-monotonic hazard from the records better than for the income related sources.

4.2.2 Dependent Interviewing Compared to Administrative Records

When the window of observation for the record data is restricted for the comparison with dependent interviewing, a slightly different set of predictors are significant for exits from earnings related spells. Exit hazards again increase for those in work, but the partner and number of children variables are not significant. Instead hazard rates increase with the age of the youngest child and with months in unemployment. The patterns of duration dependence nonetheless appear similar and the fully flexible model again fits the data best.

With dependent interviewing, exit hazards increase with age at a decreasing rate and with being active in the labour market. Both the coefficients and significance levels of the indicator for being in work and months unemployed are similar in magnitude to those estimated from the records. Unlike in the model based on INDI survey data, having low qualifications is not significant. Although the estimates of the effects of predictors seems to be improved with DI compared to INDI, the effects of time are still not significant. In fact model (1) which does not allow for duration dependence fits the data best.

Comparing the predicted hazard rates (the second row of graphs in Figure 2) suggests that the DI data map the increasing and then falling hazard rates found in the record data more closely than the INDI data. The DI data do not display the spike in hazard rates in month 13, reflecting the reduction in transitions out of spells at the seam with DI. At the same time the within wave hazard rates are higher with DI than with INDI, although still consistently lower than with the record data. This reflects the over-estimation of spell lengths and under-representation of repeated spells, possibly due to constant wave response coupled with lower levels of under-reporting discussed in the previous section.

For the health related spells, none of the determinants have significant effect in the record data, most likely due to the small number of cases. In the survey data, exit hazards decrease for those long term sick, as was the case with the INDI data, although the coefficient is now roughly twice as large. There are significant non-monotonic effects of time, as in the record data over the longer period of observation for the comparison with INDI, and the polynomial specification fits the data best. The predicted hazard rates (the second row of graphs in Figure 3) show that the survey data no longer display the spike in month 13 and the polynomial predictions follow the record data, although as for the earnings related spells, the hazard rates with the DI data are consistently lower than with the records.

5 Summary and Conclusions

This paper has provided new evidence on the effects of errors in the reporting of income receipt on estimates of spell length distributions, determinants of exit probabilities and patterns of duration dependence. The assessment of the effectiveness of dependent interviewing at reducing reporting errors and thereby bias in such estimates is also novel. Existing studies of the effects of dependent interviewing are limited to estimates of prevalence and monthly transition rates, but do not evaluate the implications for other types of (multivariate) analyses for which these survey data are used.

The findings suggest that the extent and nature of reporting errors for income sources depend on the length of the spell relative to the length of the interval between interviews. Spells with long durations relative to the reporting period (in this case health and disability related benefits) show evidence of wave under-reporting, leading to shortened spell durations, more repeated spells and marked kinks in estimates of empirical survivor functions roughly at multiples of the interval. Spells with short durations relative to the reporting period (in this case the earnings related benefits), were consistently lengthened and fewer repeated spells were reported in the survey, suggesting that constant wave reporting may have led to (short) consecutive spells being combined to longer spells.

This implies that wave under-reporting might be more problematic in panel studies with short intervals between interviews, such as the Survey of Income and Program Participation which takes place every four months. On the other hand, a survey with short reference periods may be less sensitive to constant wave reporting.

In duration models the errors in the reporting of spells attenuated the effects of explanatory variables and of duration dependence. In the case of income related spells, exit hazards instead increased with qualifications, suggesting that respondents with low qualifications were more likely to give constant wave responses leading to over-estimation of spell durations. For both earnings and health related spells, predicted hazard rates based on the survey data did not match the non-monotonic duration patterns in the record data. Instead, hazard rates corresponding to within wave periods were consistently lower, while hazard rates at the seam showed a marked spike.

Dependent interviewing appeared to reduce the extent of wave under-reporting, leading to improved estimates of the determinants of exit probabilities. The non-monotonic patterns of duration dependence also matched the record data more closely than the independent interviewing data: the spike in hazard rates at the seam disappeared and hazard rates during the reporting period were higher, although still not as high as in the record data. Although the multivariate estimates improved, dependent interviewing also had some adverse effects on estimates of spell distributions. For the health related sources for which independent interviewing led to under-reporting of spell lengths, dependent interviewing improved estimates of survivor functions, although the kink at around month 12 was still visible, suggesting that under-reporting was not completely eliminated. For the income related sources, for which independent interviewing led to over reporting of spell lengths, dependent interviewing yielded worse estimates, exacerbating the over-reporting. This suggests that although dependent interviewing reduced under-reporting, it did not affect constant wave response, but instead contributed to it, since short spells which might have been under-reported with independent interviewing were more likely to be reported as having been received in all months of the reference period.

While the reduction of under-reporting is clearly an important task and dependent interviewing therefore a valuable tool, survey designers may need to think about different question designs which query whether receipt of a source reported in reaction to a dependent interviewing reminder or edit check really was for all months in the reference period, or whether the current receipt is part of a new spell compared to the spell in progress at the previous interview date.

References:

Ashworth, K., Walker, R. and Trinder, P. (1997) 'Benefit Dynamics in Britain: Routes On and Off Income Support', *Centre for Research in Social Policy Working Paper, No.* CRSP253S, Loughborough:

Blank, R. M. (1989) 'Analyzing the Length of Welfare Spells', *Journal of Public Economics,* 39: 245-273.

Boudreau, C.(2003) *Duration Data Analysis in Longitudinal Surveys.*University of Waterloo: Department of Statistics and Actuarial Science.

Bound, J., Brown, C. and Mathiowetz, N. (2001) 'Measurement Error in Survey Data', in J. J. Heckman and E. Leamer (eds) *Handbook of Econometrics*. Vol. 5

Dibbs, R., Hale, A., Loverock, R. and Michaud, S. (1995) 'Some Effects of Computer Assisted Interviewing on the Data Quality of the Survey of Labour and Income Dynamics', *SLID Research Paper, No.* 95-07, Ottawa: Statistics Canada.

Hill, D. H. (1994) 'The Relative Empirical Validity of Dependent and Independent Data Collection in a Panel Survey', *Journal of Official Statistics,* 10(4): 359-380.

Hoynes, H. and MaCurdy, T. (1994) 'Has the Decline in Benefits Shortened Welfare Spells?' *The American Economic Review,* 84(2): 43-48.

Jäckle, A. (2006) 'Dependent Interviewing: A Framework and Application to Current Research'.Colchester: University of Essex (mimeo).

Jenkins, S. P., Cappellari, L., Lynn, P., Jäckle, A. and Sala, E. (2004) 'Patterns of Consent: Evidence from a General Household Survey', *ISER Working Paper No. 2004-27, No.*, Colchester: University of Essex. http://www.iser.essex.ac.uk/pubs/workpaps/pdf/2004-27.pdf

Jenkins, S. P., Lynn, P., Jäckle, A. and Sala, E. (2004) 'Linking Household Survey and Administrative Record Data: What Should the Matching Variables Be?' *ISER Working Paper No. 2004-23, No.*, Colchester: University of Essex. http://www.iser.essex.ac.uk/pubs/workpaps/pdf/2004-23.pdf

Kiefer, N. M. (1988) 'Economic Duration Data and Hazard Functions', *Journal of Economic Literature,* 26(2): 646-679.

Long, S. K. (1990) 'Children and Welfare: Patterns of Multiple Program Participation', *No.*, Washington, DC: The Urban Institute.

Lynn, P., Jäckle, A., Jenkins, S. P. and Sala, E. (2004) 'The Impact of Interviewing Method on Measurement Error in Panel Survey Measures of Benefit Receipt: Evidence from a Validation Study', *ISER Working Paper, No.* 2004-28, Colchester: University of Essex. http://www.iser.essex.ac.uk/pubs/workpaps/pdf/2004-28.pdf

Marquis, K. H., Moore, J. C. and Huggins, V. J. (1990) 'Implications of SIPP Record Check Results for Measurement Principles and Practice', *Proceedings of the Survey Research Methods Section,* Alexandria, VA, American Statistical Association, 564-569.

O'Neill, J. A., Bassi, L. J. and Wolf, D. A. (1987) 'The Duration of Welfare Spells', *The Review of Economics and Statistics,* 69(2): 241-248.

Ruggles, P. (1989) 'Welfare Dependency and Its Causes: Determinants of the Duration of Welfare Spells', *SIPP Working Paper, No.* 8909, Washington, DC: US Bureau of the Census.

Appendix

In the following tables and graphs, income sources are sorted according to whether they relate to disability or earnings and within each group are sorted in order of prevalence according to the administrative records. The following abbreviations are used:

INDI Independent Interviewing
DI Dependent Interviewing

<u>Disability related benefits:</u>
DLA Disability Living Allowance (care and or mobility component)
IB Incapacity Benefit
AA Attendance Allowance
ICA Invalid Care Allowance (now known as Carer's Allowance)
IID Industrial Injuries Disablement Benefit
SDA Severe Disablement Allowance
DPT Disabled Person's Tax Credit

<u>Earnings related benefits:</u>
HB Housing Benefit
IS Income Support
JSA Job Seeker's Allowance
WFTC Working Families' Tax Credit
UB/IS Unemployment Benefit/Income Support

<u>Other:</u>
CB Child Benefit (including One Parent Benefit)
RP Retirement Pension
WB Widows Benefit

Table 1: Independent interviewing – inflow after January 1999

Source	Left censored spells	Inflow spells	Average duration (mths)	Complete spells	Average duration (mths)	Average spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	61	34	20.3	11	12.8	1.06	67.65	29.41	23.53
IB	51	34	17.0	14	9.6	1.03	41.18	26.47	14.71
AA	25	25	16.2	11	8.1	1.14	72.00	32.00	20.00
ICA	13	15	18.2	7	12.7	1.00	53.33	33.33	26.67
IID	16	5	14.8	4	14.0	1.25	100.00	80.00	80.00
SDA	19	10	20.3	4	17.3	1.00	100.00	40.00	40.00
DPT	5	0	–	0	–	–	0.00	0.00	0.00
HB	200	102	20.2	34	13.4	1.13	61.76	23.53	15.69
IS	100	85	15.3	40	9.7	1.20	51.76	28.24	18.82
JSA	41	53	6.5	41	5.1	1.26	28.30	15.09	5.66
WFTC	39	70	17.4	37	9.7	1.13	34.29	38.57	12.86
CB	134	38	22.6	13	11.5	1.00	73.68	26.32	26.32
RP	144	29	28.8	1	26.0	1.00	34.48	3.45	3.45
WB	9	6	13.5	4	10.8	1.00	50.00	50.00	50.00
Total	863	507	17.7	221	10.0	1.11	52.47	27.02	17.36

Table 2: Administrative records – inflow after January 1999

Source	Left censored spells	Inflow spells	Average duration (mths)	Complete spells	Average duration(mths)	Average spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	45	14	24.5	0	–	1.00	14.29	0.00	0.00
IB	39	46	15.5	24	11.0	1.12	6.52	6.52	0.00
AA	15	12	25.4	0	–	1.00	8.33	0.00	0.00
ICA	8	9	17.0	1	36.0	1.00	0.00	0.00	0.00
IID	7	3	24.7	0	–	1.00	0.00	0.00	0.00
SDA	2	1	37.0	0	–	1.00	0.00	0.00	0.00
DPT	0	2	10.0	2	10.0	1.00	0.00	0.00	0.00
HB	101	132	13.0	88	9.7	1.76	3.03	3.03	0.00
IS	66	90	14.6	38	9.2	1.20	5.56	3.33	2.22
JSA	11	102	6.0	87	5.3	1.59	7.84	5.88	0.00
WFTC	3	77	18.8	49	13.0	1.33	3.90	5.20	0.00
CB	100	13	30.4	2	20.0	1.00	0.00	0.00	0.00
RP	139	31	26.2	0	–	1.00	3.23	0.00	0.00
WB	4	3	10.3	0	–	1.00	0.00	0.00	0.00
Total	540	535	14.9	291	9.2	1.33	5.05	3.74	0.37

[1]Mean number of spells, excluding sample members with zero spells of a given type.

Notes: The window of observation covered on average 50 months, so 49 potential month-to-month transitions of which 3 were seams. With a uniform distribution of transitions we would therefore expect 3/49*100=6.12% of transitions to be at the seam.

Table 3: Dependent interviewing – inflow after September 2000

Source	Left censored spells	Inflow spells	Average duration (mths)	Complete spells	Average duration(mths)	Average spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	62	19	19.7	1	13.0	1.00	31.58	5.26	0.00
IB	55	23	13.8	7	6.0	1.00	30.43	4.35	0.00
AA	31	14	17.1	4	6.5	1.00	42.86	7.14	0.00
ICA	13	6	24.0	0	–	1.00	16.67	0.00	0.00
IID	8	2	9.5	1	1.0	1.00	50.00	0.00	0.00
SDA	16	6	20.7	2	13.5	1.00	33.33	33.33	0.00
DPT	3	1	31.0	0	–	1.00	0.00	0.00	0.00
HB	262	81	19.3	11	9.0	1.03	38.27	2.47	0.00
IS	134	54	14.6	18	8.2	1.08	31.48	7.41	1.85
JSA	48	29	3.5	25	3.7	1.53	6.90	6.90	0.00
WFTC	66	43	14.8	16	10.1	1.08	18.60	11.63	0.00
CB	147	15	27.0	1	13.0	1.00	13.33	6.67	0.00
RP	160	20	20.8	0	–	1.00	15.00	0.00	0.00
WB	5	4	15.3	1	13.0	1.00	0.00	25.00	0.00
Total	1,019	317	16.5	87	7.3	1.06	27.13	6.31	0.32

Table 4: Administrative records – inflow after September 2000

Source	Left censored spells	Inflow spells	Average duration (mths)	Complete spells	Average duration(mths)	Average spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	43	9	15.2	0	–	1.00	11.11	0.00	0.00
IB	49	26	9.5	12	6.3	1.18	7.69	3.85	3.85
AA	18	13	13.9	1	1.0	1.08	0.00	0.00	0.00
ICA	10	5	13.4	0	–	1.00	0.00	0.00	0.00
IID	6	1	12.0	0	–	1.00	0.00	0.00	0.00
SDA	2	0	–	0	–	0.00	0.00	0.00	0.00
DPT	2	0	–	0	–	0.00	0.00	0.00	0.00
HB	168	116	8.1	76	6.2	1.90	0.00	1.72	0.00
IS	95	57	10.6	21	6.8	1.16	0.00	0.00	0.00
JSA	33	54	6.4	44	6.2	1.26	5.56	1.85	0.00
WFTC	53	32	11.6	15	9.9	1.14	0.00	6.25	0.00
CB	105	3	23.0	0	–	1.00	33.33	0.00	0.00
RP	158	20	13.0	1	1.0	1.00	0.00	0.00	0.00
WB	1	4	10.5	0	–	1.00	0.00	0.00	0.00
Total	743	340	9.6	170	6.5	1.32	2.06	1.77	0.29

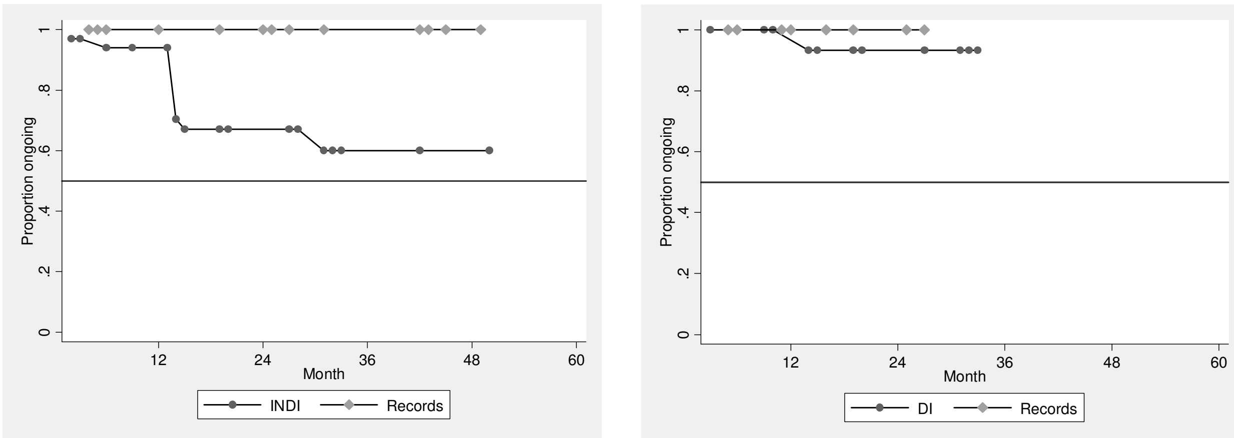
[1]Mean number of spells, excluding sample members with zero spells of a given type.

Notes: The window of observation covered 32 months, so 31 potential month-to-month transitions of which 1 was a seam. With a uniform distribution of transitions we would therefore expect 1/31*100=3.23% of transitions to be at the seam.

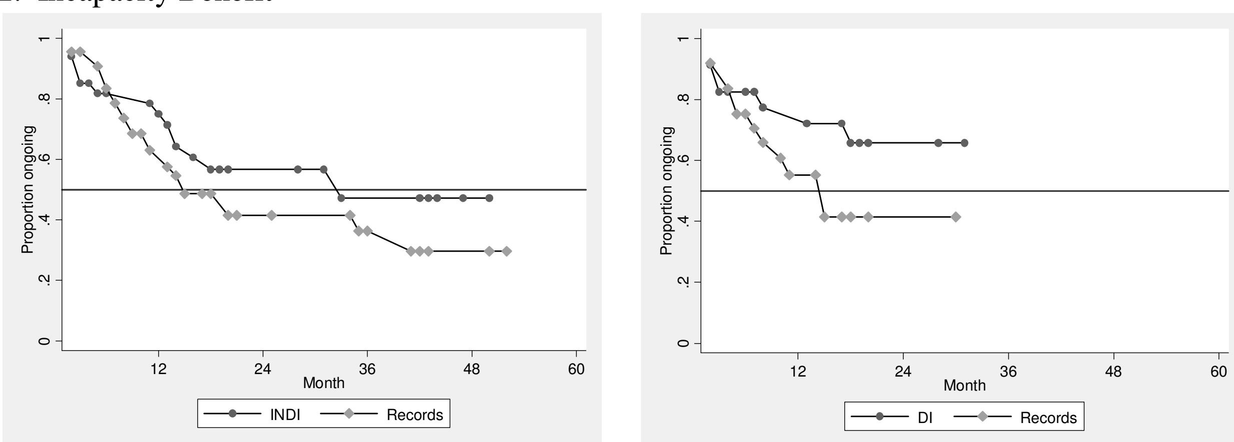
Figure 1: Lifetable estimates of Survivor Functions

The horizontal bar at y=.5 indicates median spell duration, by which 50% of spells have ended.

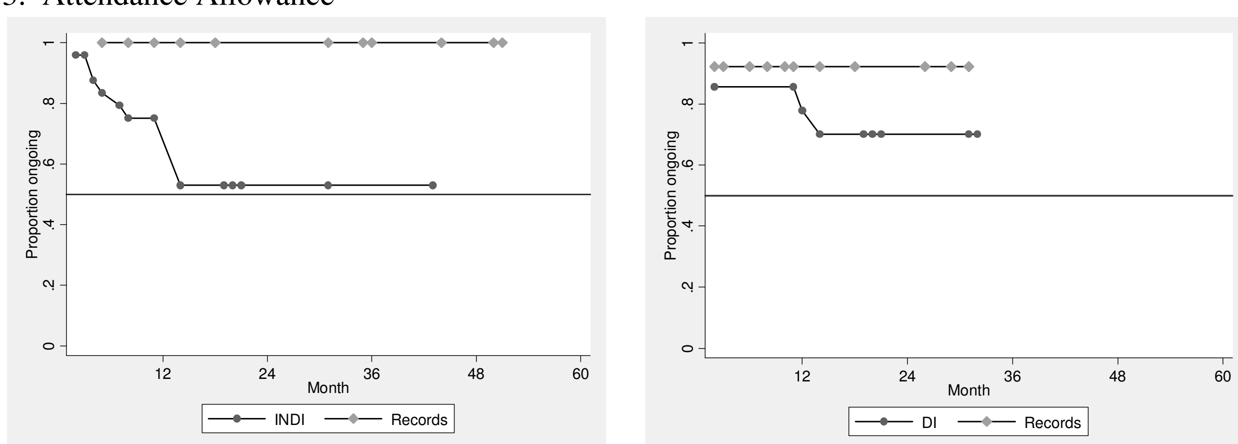
1. Disability Living Allowance



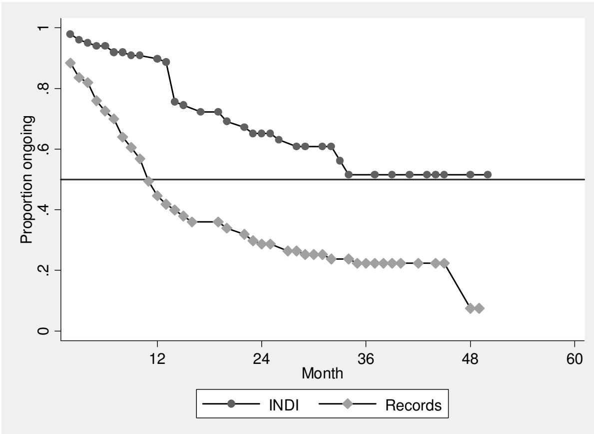
2. Incapacity Benefit

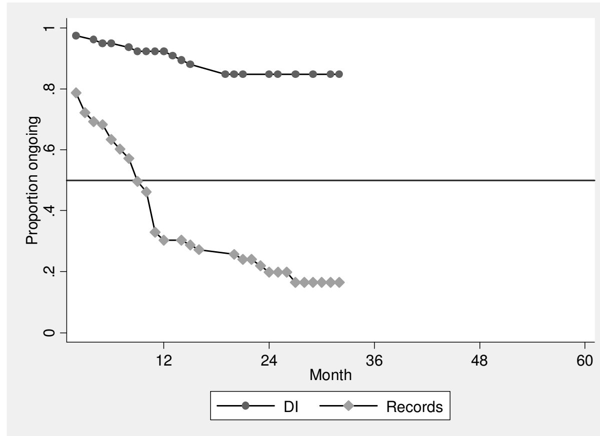


3. Attendance Allowance

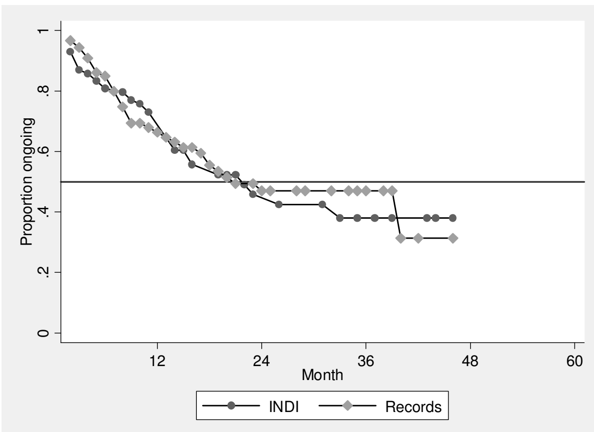


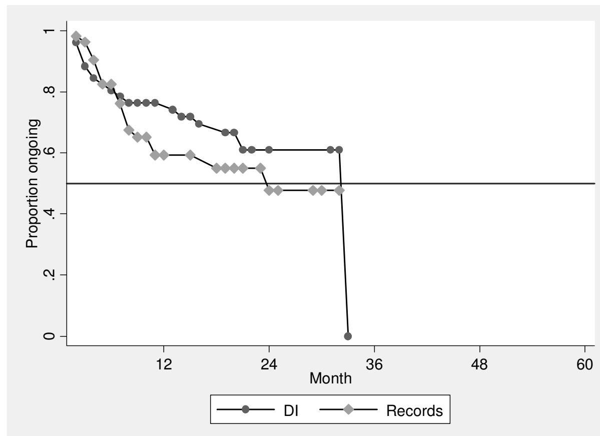
4. Housing Benefit



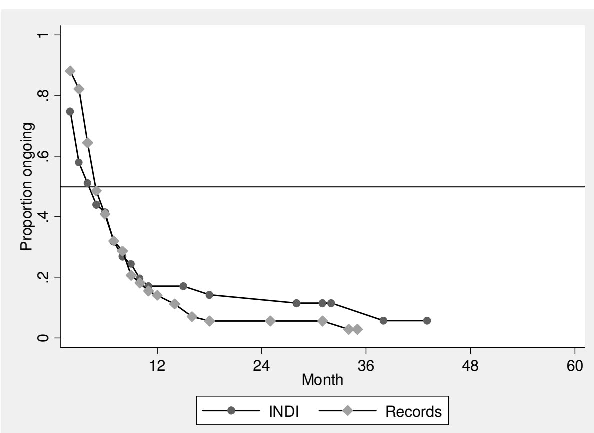


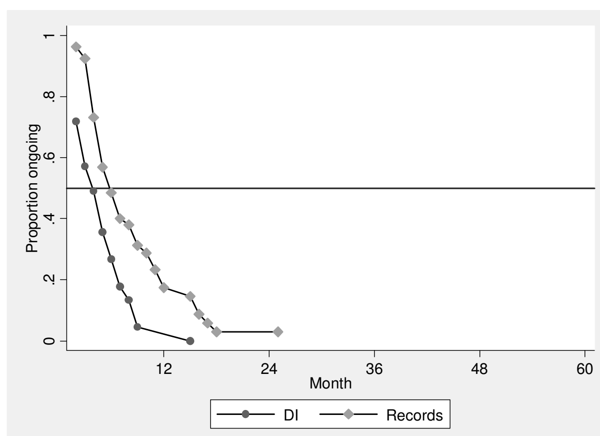
5. Income Support



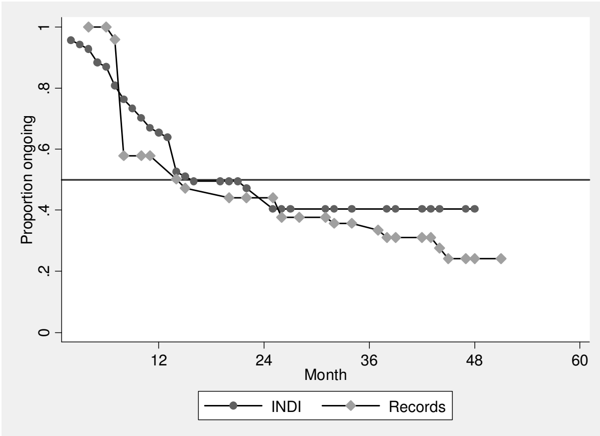


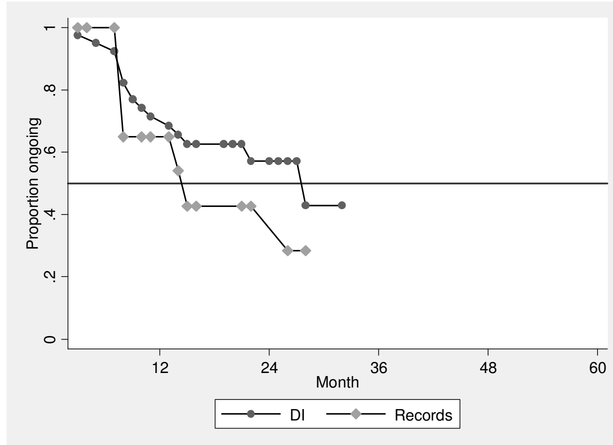
6. Job Seeker's Allowance



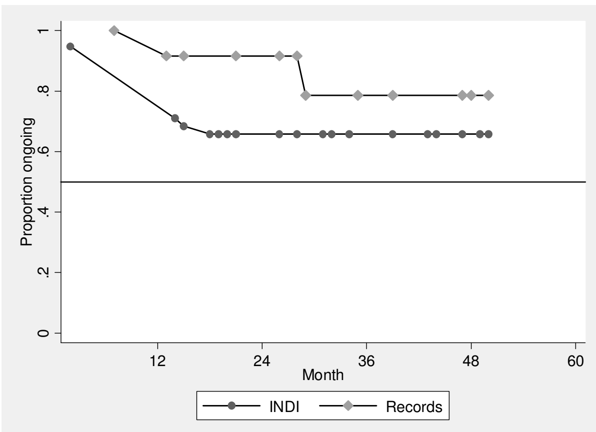


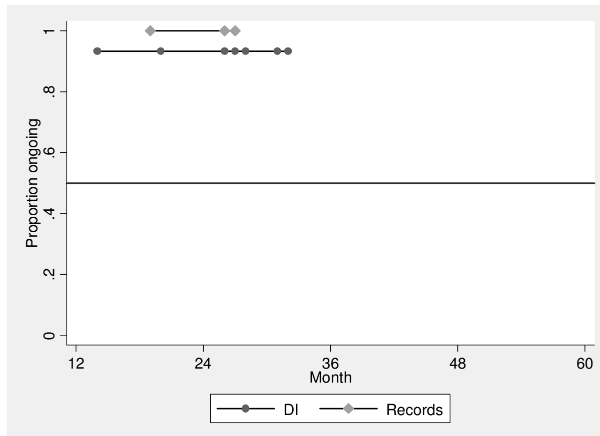
7. Working Families' Tax Credit



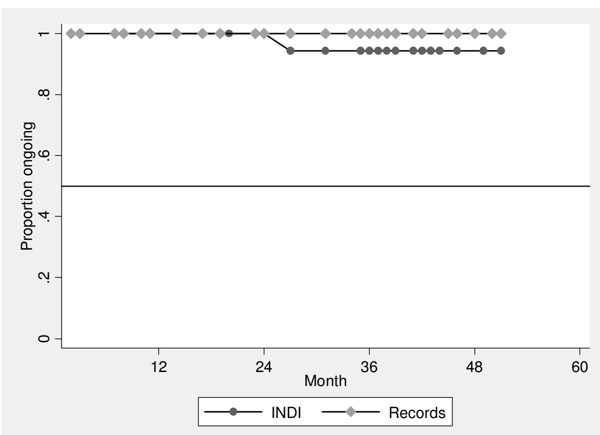


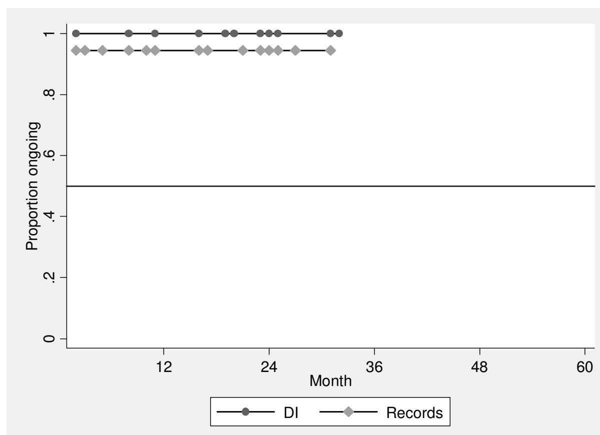
8. Child Benefit



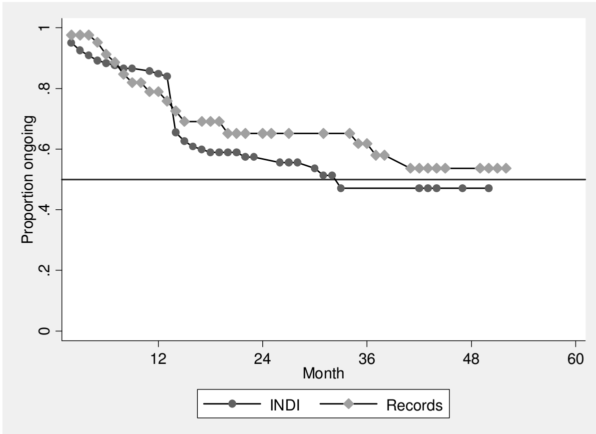


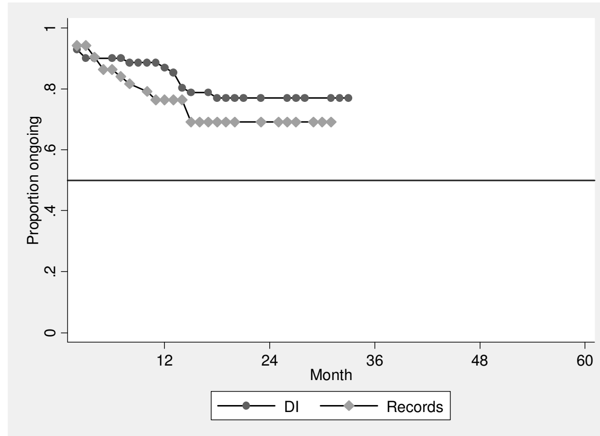
9. Retirement Pension



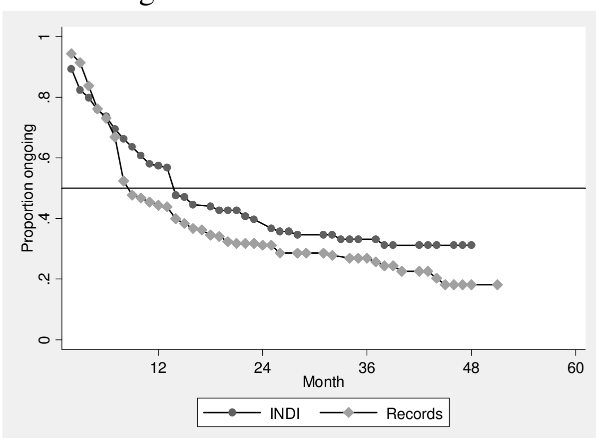


10. Disability related sources combined





11. Earnings related sources combined



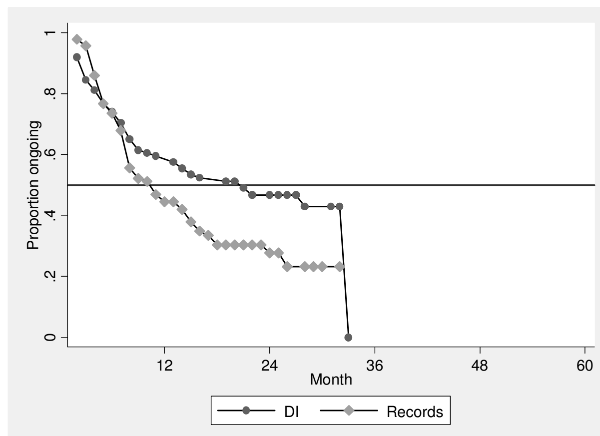


Table 5: Summary statistics of explanatory variables

	Earnings related benefits				Health related benefits			
	INDI		DI		INDI		DI	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
age	38.27	16.88	38.44	17.95	54.530	19.401	55.750	20.011
male	0.454	0.499	0.373	0.486	0.407	0.493	0.380	0.489
no qualifications	0.298	0.459	0.317	0.467	0.569	0.497	0.690	0.466
married/cohabiting	0.507	0.501	0.460	0.500	0.545	0.500	0.493	0.504
London/South East	0.173	0.379	0.206	0.406	0.106	0.309	0.169	0.377
spouse employed	0.282	0.451	0.280	0.451	−	−	−	−
# children <16	1.068	1.176	0.992	1.176	−	−	−	−
age youngest <16	3.481	4.517	2.643	3.850	−	−	−	−
unemployment rate[1]	3.482	1.631	3.069	1.413	−	−	−	−
own house	0.246	0.432	0.175	0.381	−	−	−	−
self-/employed	0.396	0.490	0.349	0.479	−	−	−	−
unemployed	0.208	0.407	0.246	0.432	−	−	−	−
mths unemployed	6.401	10.411	7.222	10.395	−	−	−	−
health problem	−	−	−	−	0.886	0.319	0.930	0.258
long term sick	−	−	−	−	0.374	0.486	0.338	0.476
N spells	207		126		123		71	

Notes: Characteristics in first month of spell according to survey reports. [1]Averaged over all spell months.

Table 6: Duration models for earnings related benefits – Survey (INDI) versus Records

Survey (INDI)	(1) Coeff.	S.E.	(2) Coeff.	S.E.	(3) Coeff.	S.E.	(4) Coeff.	S.E.
IS	1.287**	0.408	1.343**	0.437	1.175**	0.394	1.173**	0.385
JSA	2.059***	0.439	2.258***	0.475	1.957***	0.438	1.867***	0.435
age	0.009	0.036	0.007	0.039	0.018	0.035	0.016	0.034
age^2	0.000	0.000	0.000	0.000	-0.001	0.000	0.000	0.000
male	0.219	0.251	0.191	0.269	0.085	0.241	0.171	0.236
no quals	-0.729*	0.284	-0.793**	0.302	-0.777**	0.282	-0.716**	0.271
married/cohab	0.467	0.289	0.461	0.304	0.401	0.283	0.445	0.276
London/SE	0.042	0.277	0.024	0.296	-0.026	0.270	0.029	0.261
spouse emp	0.091	0.263	0.076	0.266	0.003	0.247	0.042	0.249
# children	-0.264	0.159	-0.271	0.164	-0.257	0.150	-0.253	0.150
age youngest	-0.023	0.043	-0.024	0.045	-0.013	0.040	-0.016	0.039
U rate	0.002	0.069	0.013	0.075	-0.030	0.065	-0.015	0.064
own house	0.200	0.269	0.210	0.279	0.257	0.252	0.213	0.247
emp/semp	0.846*	0.363	0.919*	0.394	0.793*	0.355	0.765*	0.345
unemployed	-0.024	0.415	0.014	0.432	-0.011	0.403	-0.041	0.399
mths unemp	-0.006	0.011	-0.009	0.012	0.002	0.011	0.000	0.010
t	–	–	–	–	0.000	0.080	–	–
t^2	–	–	–	–	0.000	0.006	–	–
t^3	–	–	–	–	0.000	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.208*	0.096
_cons	-3.811	–	-4.221	–	-3.662	–	-3.459	–
# spell-mths	2842	–	2842	–	2842	–	2842	–
# parameters	17	–	40	–	20	–	18	–
log likelihood	-431.98	–	-408.26	–	-427.76	–	-429.83	–
AIC	897.96	–	896.53	–	895.53	–	895.67	–
Records								
IS	0.915**	0.352	0.987**	0.366	0.881**	0.342	0.937**	0.356
JSA	1.765***	0.272	1.890***	0.304	1.754***	0.272	1.808***	0.286
age	-0.023	0.034	-0.038	0.036	-0.030	0.034	-0.026	0.035
age^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
male	-0.006	0.191	-0.008	0.209	-0.018	0.194	-0.002	0.197
no quals	-0.070	0.247	-0.022	0.260	-0.046	0.245	-0.069	0.252
married/cohab	0.529*	0.215	0.487*	0.222	0.533*	0.209	0.532*	0.220
London/SE	-0.279	0.230	-0.354	0.235	-0.356	0.221	-0.275	0.236
spouse emp	0.532*	0.212	0.552*	0.216	0.508*	0.203	0.555*	0.223
# children	-0.234*	0.100	-0.205*	0.099	-0.232*	0.097	-0.235*	0.103
age youngest	0.019	0.026	0.027	0.027	0.032	0.025	0.019	0.027
U rate	-0.009	0.043	0.000	0.048	-0.031	0.045	-0.005	0.045
own house	0.006	0.173	-0.042	0.187	-0.020	0.180	0.010	0.177
emp/semp	0.930**	0.288	1.068***	0.303	1.004***	0.287	0.933**	0.291
unemployed	0.250	0.268	0.451	0.280	0.348	0.267	0.256	0.269
mths unemp	0.003	0.007	-0.001	0.009	0.003	0.008	0.003	0.008
t	–	–	–	–	0.256***	0.054	–	–
t^2	–	–	–	–	-0.018***	0.004	–	–
t^3	–	–	–	–	0.000***	0.000	–	–
ln(t)	–	–	–	–	–	–	0.049	0.081
_cons	-3.493	–	-4.502	–	-3.987	–	-3.563	–
# spell-mths	3350	–	3350	–	3350	–	3350	–
# parameters	17	–	45	–	20	–	18	–
log likelihood	-600.39	–	-544.53	–	-585.33	–	-600.23	–
AIC	1234.79	–	1179.05	–	1210.65	–	1236.45	–

Notes: Model (1): time not included, (2): fully flexible (time coefficients not reported), (3): polynomial, (4): Weibull. Omitted income source: WFTC.
Standard errors adjusted for clustering at the individual level. * P<.05, ** P<.01, *** P<.001.

Table 7: Duration models for health related benefits – Survey (INDI) versus Records

Survey	(1)		(2)		(3)		(4)	
(INDI)	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
SDA	-0.043	0.536	-0.131	0.552	-0.198	0.521	-0.071	0.540
IID	0.628	0.420	0.440	0.462	0.400	0.435	0.570	0.424
AA	0.275	0.547	0.264	0.574	0.159	0.540	0.240	0.539
ICA	-0.228	0.617	-0.327	0.606	-0.364	0.606	-0.214	0.603
DLA	-0.535	0.465	-0.582	0.481	-0.645	0.464	-0.540	0.458
age	0.034	0.068	0.034	0.073	0.041	0.068	0.034	0.065
age^2	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001
male	0.078	0.371	0.100	0.395	0.039	0.360	0.058	0.360
no quals	-0.537	0.402	-0.594	0.405	-0.595	0.396	-0.533	0.390
married/cohab	-0.332	0.317	-0.338	0.327	-0.283	0.309	-0.326	0.310
London/SE	0.062	0.579	-0.101	0.623	-0.005	0.588	0.055	0.566
health prob	0.273	0.679	0.186	0.680	0.283	0.695	0.328	0.703
lt sick	-0.842*	0.414	-0.879*	0.423	-0.883*	0.419	-0.832*	0.401
t	–	–	–	–	0.005	0.133	–	–
t^2	–	–	–	–	0.003	0.009	–	–
t^3	–	–	–	–	0.000	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.112	0.160
_cons	-3.551	–	-4.489	–	-3.765	–	-3.362	–
# spell-mths	2222	–	2222	–	2222	–	2222	–
# parameters	14	–	33	–	17	–	15	–
log likelihood	-236.09	–	-196.65	–	-233.21	–	-235.81	–
AIC	500.18	–	459.30	–	500.41	–	501.62	–
Records								
ICA	-2.482**	0.857	-3.375***	0.894	-2.864**	1.060	-2.551**	0.861
DTC	0.085	0.731	-0.131	0.785	0.025	0.856	0.298	0.801
age	0.279*	0.115	0.447**	0.146	0.307*	0.123	0.293*	0.130
age^2	-0.003*	0.001	-0.005**	0.002	-0.004*	0.001	-0.004*	0.002
male	0.182	0.419	-0.230	0.603	0.032	0.464	0.248	0.452
no quals	-0.515	0.528	-0.767	0.553	-0.661	0.596	-0.499	0.572
married/cohab	-0.622	0.616	-0.644	0.550	-0.545	0.602	-0.567	0.627
London/SE	-0.774	0.870	-1.450*	0.717	-0.936	0.939	-0.686	0.899
health prob	0.282	0.481	0.157	0.592	0.098	0.601	0.111	0.546
lt sick	-2.185***	0.652	-3.158***	0.742	-2.541***	0.699	-2.218***	0.674
t	–	–	–	–	0.339*	0.170	–	–
t^2	–	–	–	–	-0.018*	0.009	–	–
t^3	–	–	–	–	0.000*	0.000	–	–
ln(t)	–	–	–	–	–	–	0.206	0.200
_cons	-7.331	–	-11.603	–	-8.939	–	-7.926	–
# spell-mths	888	–	888	–	888	–	888	–
# parameters	11	–	23	–	14	–	12	–
log likelihood	-102.71	–	-79.99	–	-99.68	–	-102.22	–
AIC	227.43	–	205.98	–	227.35	–	228.44	–

Notes: Model (1): time not included, (2): fully flexible (time coefficients not reported), (3): polynomial, (4): Weibull. Omitted income source: IB.
Standard errors adjusted for clustering at the individual level. * P<.05, ** P<.01, *** P<.001.

Table 8: Duration models for earnings related benefits – Survey (DI) versus Records

Survey (DI)	(1) Coeff.	S.E.	(2) Coeff.	S.E.	(3) Coeff.	S.E.	(4) Coeff.	S.E.
IS	1.160*	0.545	1.196*	0.588	1.197*	0.550	1.160*	0.542
JSA	3.105***	0.611	3.312***	0.701	3.095***	0.613	3.099***	0.616
age	0.184*	0.075	0.194*	0.081	0.179*	0.074	0.184*	0.075
age^2	-0.002*	0.001	-0.002*	0.001	-0.002*	0.001	-0.002*	0.001
male	-0.323	0.398	-0.331	0.449	-0.310	0.412	-0.321	0.400
no quals	-0.183	0.392	-0.169	0.404	-0.158	0.393	-0.183	0.391
married/cohab	0.375	0.361	0.411	0.422	0.416	0.379	0.374	0.365
London/SE	-0.387	0.418	-0.388	0.458	-0.394	0.417	-0.387	0.417
spouse emp	-0.210	0.376	-0.133	0.415	-0.196	0.385	-0.210	0.375
# children	-0.116	0.145	-0.108	0.160	-0.119	0.143	-0.116	0.144
age youngest	-0.054	0.063	-0.065	0.069	-0.053	0.063	-0.054	0.063
U rate	-0.040	0.105	0.012	0.116	-0.039	0.108	-0.041	0.104
own house	0.194	0.347	0.112	0.376	0.179	0.354	0.195	0.346
emp/semp	1.102*	0.455	1.241*	0.500	1.155*	0.491	1.101*	0.456
unemployed	-0.011	0.811	0.230	0.920	0.054	0.853	-0.015	0.811
mths unemp	0.037**	0.013	0.039**	0.014	0.036**	0.013	0.037**	0.013
t	–	–	–	–	0.188	0.151	–	–
t^2	–	–	–	–	-0.020	0.014	–	–
t^3	–	–	–	–	0.001	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.008	0.136
_cons	-7.625	–	-8.775	–	-7.911	–	-7.603	–
# spell-mths	1507	–	1506	–	1507	–	1507	–
# parameters	17	–	34	–	20	–	18	–
log likelihood	-197.48	–	-187.34	–	-196.04	–	-197.48	–
AIC	428.97	–	442.68	–	432.07	–	430.97	–
Records								
IS	1.025*	0.438	1.218**	0.472	1.245*	0.490	1.152*	0.492
JSA	1.642***	0.345	1.782***	0.358	1.822***	0.358	1.801***	0.394
age	0.015	0.036	0.002	0.041	0.003	0.039	0.003	0.040
age^2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
male	-0.112	0.264	-0.087	0.310	-0.105	0.300	-0.125	0.305
no quals	-0.129	0.259	0.035	0.303	-0.001	0.292	0.013	0.313
married/cohab	0.411	0.244	0.435	0.274	0.497	0.265	0.479	0.286
London/SE	-0.594	0.349	-0.590	0.396	-0.576	0.371	-0.624	0.384
spouse emp	0.059	0.260	0.126	0.296	0.080	0.282	0.107	0.300
# children	-0.083	0.097	-0.060	0.109	-0.081	0.107	-0.063	0.109
age youngest	0.071*	0.032	0.075*	0.037	0.076*	0.036	0.077*	0.036
U rate	-0.033	0.058	-0.016	0.070	-0.032	0.067	-0.010	0.067
own house	0.078	0.272	0.149	0.309	0.109	0.305	0.139	0.316
emp/semp	0.901*	0.352	1.020**	0.363	1.044**	0.371	0.982**	0.379
unemployed	-0.100	0.455	0.013	0.489	-0.036	0.483	0.006	0.499
mths unemp	0.027*	0.011	0.023	0.013	0.024	0.012	0.027*	0.013
t	–	–	–	–	0.519***	0.147	–	–
t^2	–	–	–	–	-0.041**	0.013	–	–
t^3	–	–	–	–	0.001**	0.000	–	–
ln(t)	–	–	–	–	–	–	0.393**	0.137
_cons	-4.356	–	-6.126	–	-5.830	–	-5.080	–
# spell-mths	1294	–	1294	–	1294	–	1294	–
# parameters	17	–	34	–	20	–	18	–
log likelihood	-265.73	–	-239.56	–	-257.76	–	-261.78	–
AIC	565.46	–	547.12	–	555.53	–	559.55	–

Notes: Model (1): time not included, (2): fully flexible (time coefficients not reported), (3): polynomial, (4): Weibull. Standard errors adjusted for clustering at the individual level. * P<.05, ** P<.01, *** P<.001.

Table 9: Duration models for health related benefits – Survey (DI) versus Records

Survey	(1)		(2)		(3)		(4)	
(INDI)	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
SDA	0.361	0.883	0.303	0.854	0.413	0.858	0.361	0.801
IID	1.559	2.028	1.354	2.327	1.311	1.628	1.419	1.665
AA	-0.656	1.268	-0.938	1.369	-0.917	1.058	-0.874	1.092
DLA	-2.280	1.182	-2.386	1.218	-2.165	1.115	-2.202*	1.118
age	0.091	0.128	0.090	0.144	0.059	0.106	0.069	0.105
age^2	-0.001	0.001	-0.001	0.001	0.000	0.001	-0.001	0.001
male	-0.601	0.897	-0.646	0.940	-0.706	0.770	-0.619	0.778
no quals	-1.425	0.728	-1.359	0.783	-1.046	0.615	-1.143	0.626
married/cohab	-0.978	0.818	-1.130	0.888	-0.911	0.670	-1.022	0.709
London/SE	-0.493	1.044	-0.638	1.155	-0.354	0.736	-0.337	0.764
health prob	0.019	1.441	0.126	1.876	0.036	1.306	0.138	1.297
lt sick	-1.996*	0.997	-2.156*	1.021	-1.536	0.853	-1.709*	0.820
t	–	–	–	–	-2.039**	0.732	–	–
t^2	–	–	–	–	0.258*	0.102	–	–
t^3	–	–	–	–	-0.009*	0.004	–	–
ln(t)	–	–	–	–	–	–	-0.622*	0.262
_cons	-3.422	–	-4.074	–	-0.181	–	-2.430	–
# spell-mths	1075	–	1075	–	1075	–	1075	–
# parameters	13	–	20	–	16	–	14	–
log likelihood	-68.96	–	-63.88	–	-60.24	–	-66.10	–
AIC	163.92	–	167.76	–	152.48	–	160.20	–
Records								
AA	-2.827	1.828	-3.115	2.177	-3.150	1.899	-2.783	1.752
age	0.013	0.125	-0.001	0.125	0.016	0.111	0.017	0.116
age^2	0.000	0.001	0.000	0.001	0.000	0.001	0.000	0.001
male	-0.527	0.570	-0.704	0.670	-0.716	0.618	-0.522	0.558
no quals	-1.340	0.997	-1.270	1.017	-1.253	0.843	-1.311	0.909
married/cohab	0.040	0.854	0.107	0.790	-0.167	0.891	-0.007	0.877
London/SE	0.622	1.133	0.996	1.184	0.500	1.223	0.551	1.215
health prob	-0.469	0.969	-0.888	1.176	-0.474	1.098	-0.411	1.077
lt sick	-1.138	1.208	-1.226	1.234	-1.416	1.107	-1.132	1.169
t	–	–	–	–	-0.529	0.568	–	–
t^2	–	–	–	–	0.086	0.073	–	–
t^3	–	–	–	–	-0.004	0.003	–	–
ln(t)	–	–	–	–	–	–	-0.094	0.424
_cons	-2.324	–	-2.749	–	-1.485	–	-2.284	–
# spell-mths	428	–	428	–	428	–	428	–
# parameters	10	–	17	–	13	–	11	–
log likelihood	-49.33	–	-44.31	–	-48.14	–	-49.30	–
AIC	118.67	–	122.62	–	122.28	–	120.60	–

Notes: Model (1): time not included, (2): fully flexible (time coefficients not reported), (3): polynomial, (4): Weibull. Standard errors adjusted for clustering at the individual level. * P<.05, ** P<.01, *** P<.001.

Figure 2: Predicted hazard rates for earnings related benefits

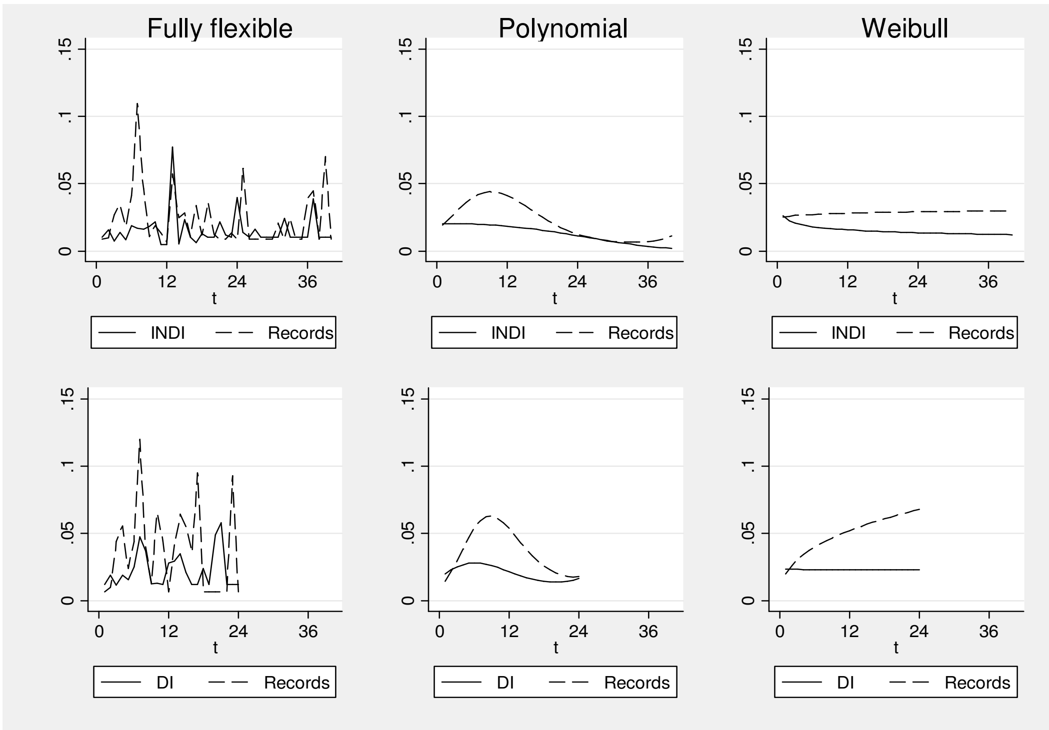


Figure 3: Predicted hazard rates for health related benefits