ISER Working Paper Series

Measurement Error and Data Collection Methods: Effects on Estimates from Event History Data

Annette Jäckle
Institute for Social and Economic Research
University of Essex

No. 2008-13
April 2008

INSTITUTE FOR SOCIAL & ECONOMIC RESEARCH

www.iser.essex.ac.uk

Non-technical summary

Panel surveys mainly collect information about the situation of respondents around the time of each interview. For some topics, panel surveys in addition collect information about the occurrence and dates of events between interviews. Event history data can be used to answer questions such as how long people receive certain State benefits, which characteristics distinguish people who are at risk of receiving benefits over long periods of time, and whether how long someone receives benefits has any causal effect on the probability that they will move off benefits.

Event history data are often reported with errors, either because respondents forget to report events, or because they misreport dates. This study used matched survey and administrative data to examine the effects of errors on estimates from event history data. The survey data are from the 'Improving Survey Measurement of Income and Employment' study. The study included an experiment testing different methods of dependent interviewing, which is a data collection method used to reduce measurement error. With this method respondents are reminded of sources they have reported previously, or asked check questions to verify that sources no longer reported have truly ended. For the survey respondents, administrative records on benefit receipt were in addition obtained from the Department for Work and Pensions.

The analysis examined estimates of the durations of benefit receipt, of predictors of spell lengths and of how exit probabilities varied with the length of receipt. If the estimates from the survey data differed from those from the administrative records, this was judged as bias caused by measurement errors. Estimates from the administrative records were also compared to estimates from dependent interviewing, to test whether this data collection method reduced biases caused by measurement error.

The results showed that measurement errors did bias estimates from event history data. The durations of long spells tended to be under-estimated, probably because respondents correctly reported them in some but not all interviews. Estimates for predictors of spell durations were weaker in the survey data than the administrative data. The patterns of how exit probabilities changed with the length of benefit receipt were also weaker in the survey data. There was also some evidence suggesting that respondents with lower educational qualifications were more prone to misreport their income sources. Dependent interviewing improved estimates of spell durations for long spells, as well as the predictions of spell durations and patterns of how exit probabilities varied with the length of receipt. For short spells, dependent interviewing however increased biases in estimated spell durations.

Measurement Error and Data Collection Methods: Effects on Estimates from Event History Data

Annette Jäckle

April 2008

ABSTRACT

Event history data from panel surveys typically display a concentration of transitions at the seam between waves of data collection. This 'seam effect' is likely to bias estimated durations of benefit receipt, attenuate the estimated effects of explanatory factors on conditional exit probabilities and bias estimated duration dependence. This paper uses benefit histories from survey reports and matched administrative records to assess the extent of bias in key estimates. The paper also evaluates the effectiveness at reducing bias 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.

Keywords: dependent interviewing, survival analysis, duration analysis, bias, seam effect, validation, record check, State benefit, welfare.

JEL codes: C41, C42

Acknowledgements:
Funding for this work from the ESRC (RES-000-22-2323) is gratefully acknowledged. The data are from a project on 'Improving Survey Measurement of Income and Employment' (ISMIE), funded by the ESRC Research Methods Programme (H333250031). I am grateful to Peter Lynn, Steve Pudney, Bernhard Schimpl-Neimanns, Marjo Pyy-Martikainen and seminar attendants at the University of Essex, DIW Berlin, ZUMA, the Survey Research Centre at the University of Michigan and the Bureau of Labor Statistics 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.

Contact: Annette Jäckle, Institute for Social and Economic Research, University of Essex, Colchester CO4 3SQ, UK. Email: aejack@essex.ac.uk

1 Introduction

Event history data are used to answer important policy questions, such as how long, on average, people receive unemployment benefit, which characteristics distinguish people at risk of receiving benefits over long periods of time or whether the length of receipt has any causal effect on the probability of leaving benefits. Accurate data on the occurrence and dates of events are difficult to collect in surveys. In panel surveys, errors in event history data are visible as biased transition rates and flows between states and are likely to bias estimates from event history models. Histories of income receipt are usually collected by asking respondents, at each wave, to report their receipt status for every week or month of the period since the previous interview. When the reports from different waves are joined, a disproportionate number of transitions are typically observed at the seams between reference periods. The concentration of transitions at the seams, referred to as the "seam effect", is mainly caused by *constant wave response* (reporting receipt for 'all' or 'none' of the weeks or months in the reference period) and *wave under-reporting* (reporting receipt in some but not all relevant waves). These errors are mostly due to problems respondents have in recalling the required information, but can also be due to interviewer errors in recording responses or due to data processing errors (see Jäckle 2008). Different data collection methods, such as Event History Calendars and dependent interviewing, have been developed to reduce measurement errors that produce seam effects in panel surveys. With dependent interviewing, the technique investigated here, information collected in a previous interview is used to remind the respondent of sources reported previously, or to verify that the receipt of sources no longer reported has truly ended. This paper examines the effect of measurement errors in event history data on estimates and evaluates the effectiveness of dependent interviewing at reducing measurement errors and biases in estimates. The evaluation uses data on histories of income receipt and compares estimates based on a panel survey with estimates based on linked administrative records.

Under-reporting receipt at some but not all waves means that spells frequently appear to end at an interview date, possibly only to be reported again in the following wave. Reporting the same receipt status for all weeks or months of a reference period means that the start dates of spells tend to be pushed back to the beginning of the reference period. These errors are likely to bias estimated durations of benefit receipt (Boudreau 2003), attenuate the estimated effects of explanatory factors on conditional exit probabilities (see Hill 1994 for an example of job history duration models) and bias estimates of duration dependence (the notion that the recipient's behaviour may change with the duration of benefit receipt, also referred to as ‘welfare dependency’). Little is however known about the nature of errors in event history data from panel surveys or their effects on estimates, let alone about ways of mitigating these. The record check study reported by Marquis, Moore and Huggins (1990) appears to be the only study of measurement error in survey estimates of benefit income that was not purely cross-sectional in nature (for an extensive review of validation studies of survey data, see Bound et al. 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, maybe because the matched records only covered a period of 8 months. Possibly as a consequence of the lack of information, analysts using event history 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 (see Section 7 in Lynn et al. 2005).

Several panel surveys have attempted to improve the quality of event history data by introducing dependent interviewing (for example, the Survey of Income and Program Participation, the Survey of Labour and Income Dynamics and the British Household Panel Survey). Record check studies have shown that dependent interviewing reduces under-reporting of income receipt, hence improving cross-sectional estimates of prevalence (e.g. Dibbs et al. 1995; Lynn et al. 2004), and somewhat reduces the concentration of transitions at the seam (e.g. Moore et al. in press). The effects on other (longitudinal) estimates have however not been tested. The assumption seems to be that reduced under-reporting unambiguously improves data quality. The results of the present study, however, show that this is not necessarily the case. Although the dependent interviewing designs tested here improved reporting (Lynn et al. 2004), they did not improve the dating of receipt. Constant wave reporting remained a problem (Jäckle in press) and the net effect of dependent interviewing on estimates from event history models is therefore not predictable.

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. The individual reports were evaluated by Lynn et al. (2004), who showed that there was considerable under-reporting of income sources and that dependent interviewing improved reporting. The present paper focuses on the effects of errors in individual reports on aggregate estimates, for which this type of data are typically used. The survey and administrative data were compared to assess the extent of bias in key estimates, such as the distribution and determinants of spell lengths and patterns of duration dependence, and to assess the effectiveness of dependent interviewing at reducing bias.

The survey and matched administrative data are introduced in Section 2. Section 3 sets out the hypotheses about the expected effects of errors and dependent interviewing on estimates from event history models. Section 4 discusses the effects of errors and dependent interviewing on spell distributions and Section 5 the effects on estimates from multivariate models. Section 6 concludes with a summary and discussion of implications for data collection.

2 The survey and administrative data

The data 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 records from the Department for Work and Pensions (DWP), the government department in charge of administering benefits and tax credits. Both the survey and administrative records contained information about receipt of 14 different State benefits and tax credits, including Child Benefit, pensions and benefits related to disability and income. The consent rate for the record linkage was 77.4% (N=799). Of these, 74.1% (N=592) were successfully linked to DWP records. Non-matched respondents were probably mainly 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 (see Jenkins et al. 2008). The low-income sample, attrition during prior panel waves and potentially selective consent for linkage (see Jenkins et al. 2006) meant that the respondent sample was not representative of the general population. Comparing the duration of spells in progress in August 2003 in our administrative records with records for the entire population suggests that our sample under-represented shorter and over-represented longer spells (see Figure 1). Ideally the conclusions about measurement error and biases from this study would extend to analyses of general population data. This is possible only under certain assumptions, where the nature of those assumptions depends on the type of analysis. If sample selection is independent of measurement error conditional on true duration and the explanatory covariates, then an analysis of the properties of the survey response error is generalisable from the selected sample. This requires, for example, that low willingness to participate in the study should not be directly related to the ability to give accurate answers, although they may both be influenced by other observable factors, such as educational attainment. This is an important example, because the low-income sample over-represented respondents with low qualifications and the results suggest that respondents with low qualifications may have been more likely to misreport. In contrast with this, the biases in a full survival analysis like that presented here will generally be affected by selection unless one makes the further assumption that true duration is independent of selection, conditional on the covariates. This is equivalent to Rubin’s “selection on observables” assumption.

The survey included an experiment comparing different data collection methods, for which respondents were randomly allocated to one of three treatment groups: independent interviewing (INDI), proactive dependent interviewing (PDI) or reactive dependent interviewing (RDI). The INDI group received the standard BHPS questions. Respondents were shown a series of showcards listing a total of 34 different unearned income sources (including benefits and tax credits, but also private transfers and rents) and asked which ones they had received during the reference period: “Please look at this card and tell me if, since September 1st 2001, you have received any of the types of income or payments shown, either just yourself or jointly?” For each reported source they were then asked “And for which months since September 1st 2001 have you received…?”

The RDI group were first asked the same INDI questions. The CAPI script then checked responses against the sources reported in the previous interview. For each source reported in the previous but not the current interview, respondents were asked “Can I just check, according to our records you have in the past received <*INCOME SOURCE*>. Have you received <*INCOME SOURCE*> at any time since <*DATE OF INTERVIEW*>?” If yes, they were then asked “For which months since <*MONTH OF INTERVIEW*> have you received <*INCOME SOURCE*>?” Lynn et al. (2006) found that response distributions for the INDI questions were no different for respondents in the INDI and RDI groups. The RDI group could therefore be treated as INDI, if answers to the RDI follow-up were ignored, or as dependent interviewing, if answers to the follow-up were incorporated.

The PDI group were reminded of each income source reported in the previous interview and asked whether they had continued receiving it: “According to our records, when we last interviewed you, on <*DATE OF INTERVIEW*>, you were receiving <*INCOME SOURCE*>, either yourself or jointly. For which months since <*MONTH OF INTERVIEW*> have you received <*INCOME SOURCE*>?" They were then shown the showcards and asked the INDI questions to collect information about any new income sources.

Dates of receipt in the survey were recorded in calendar months. The administrative records contained 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.) For comparability, the administrative records were converted to monthly data.

The survey data contained 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 were only available for the period starting on 1st September 2000. The administrative records covered benefit histories from January 1999 until October 2003. For comparability, spells in the administrative data were treated as right censored if they were ongoing at the time of the 2003 interview.

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 were not recorded separately in the DWP data. For comparability I combined the survey spells by joining overlapping spells into longer spells and ignoring multiple concurrent benefit components.

The survey reports from subsequent interviews were combined to create benefit histories in the following way (illustrated in Figure 2). The bulk of interviews took place in September/October each year, although fieldwork continued until February. The reference period for reporting income sources started on 1st September of the previous calendar year, rather than on the previous interview date. As a result, there was usually an overlap in reporting periods and the earliest part of the reference period had often already been reported on in the previous interview. For the overlapping periods, I assumed that the report from the interview closer in time was more likely to be correct and discarded the information for this period from the following interview. This editing rule meant that the interview months marked the seams between reference periods. (All analyses reported here were also carried out using the alternative editing rule, where the information from the interview closest in time was discarded and the start of the reference period on 1st September each year formed the seam. The results were basically unchanged.)

The unit of analysis in this study was a spell of income receipt. The sample of spells was restricted to reports from respondents who had given permission for the linkage to administrative records, but included respondents for whom data were not available for all waves. Of the respondents included in the analysis, 91.5% were interviewed in all four waves. For the other sample members there were no data for one or both of the first two waves, either because they refused, were not contacted, joined a sample household or turned 16 and became eligible for the interviews after January 1999. For respondents for whom less than four waves of survey data were available, the DWP records were edited to mirror the time periods covered in the survey data. The sample included some under-reporters (respondents for whom there was a benefit record in the administrative data but not the survey data) and over-reporters (which could include correct reports for respondents for whom the linkage to the administrative records had failed). In total, data from 633 respondents were included in the analysis. Of these, 8.4% were only observed in the survey and 2.7% had not reported any benefits in the survey, but were observed in the administrative records. As a robustness check, I examined the spell distributions using only those respondents for whom there were income spells in both the survey and administrative data. The results were similar to those reported here.

The sample of spells only included spells starting after January 1999 (or September 2000 for the comparison with dependent interviewing). This inflow sample included repeated spells and right censored spells, which could be censored at the date of the final interview or during the panel due to unit non-response and item non-response to date questions. All left censored spells were dropped, including spells that were ongoing in January 1999 (or September 2000) and spells for which the start date was unknown due to unit non-response. The numbers of left censored spells are documented in Tables 1 to 4. 19 spells in the INDI sample and 6 spells in the dependent interviewing (DI) sample were left censored due to unit non-response. For comparability with the survey data, DWP spells which started during a reference period for which the respondent was not interviewed were treated as left-censored and dropped.

The comparison of INDI survey data with administrative records was based on the sample of respondents allocated to INDI and RDI where, in the survey data, any incomes reported in response to the reactive follow-up question were excluded from the analysis. The comparison of dependent interviewing with administrative records was based on the sample of respondents allocated to PDI and RDI. In this case, the responses to the reactive follow-up question were included in the survey data. The proactive and reactive groups were pooled after examining each group separately. Tables 1 to 4 show the sample sizes for the different income sources in the survey and administrative data.

3 Hypotheses: effects of errors and dependent interviewing on biases

The main errors respondents make when reporting their income histories are under-reporting and misdating of receipt. Three scenarios can be distinguished, depending on the nature of the spell to be reported on. The first scenario applies to income sources received during all months of the reference period. In this case, the most likely type of reporting error is for respondents to under-report receipt for all months (referred to as 'wave under-reporting' below). The second scenario applies to income sources received only in the earlier part of the reference period. In this case, the most likely type of error is for respondents to report their current status of non-receipt correctly, but to falsely report the same status for all months in the reference period (referred to as 'constant wave under-reporting' below). This can be thought of as a dating error, where respondents backward telescope the end of receipt to before the start of the reference period. The third scenario refers to income sources received at the time of the interview, but which were not received during the entire reference period. In this case, the most likely type of errors is for respondents to correctly report their current receipt status, but to falsely report receipt for all months in the reference period (referred to as 'constant wave over-reporting' below). This can again be thought of as a dating error, where the start date is backward telescoped to before the start of the reference period. The analysis here does not test for these errors, which could be done by validating each survey response with the corresponding administrative record. Instead, the question is what effect these types of errors have on aggregate estimates, for which the income histories are typically used.

Wave and constant wave under-reporting both mean that spells are 'chopped off' at the seams between reference periods. This will lead to an under-estimation of benefit spell durations and an over-estimation of the proportion of transitions off benefits at the seams. Wave under-reporting will in addition lead to an over-estimation of the proportion of spells completed during the window of observation. For income sources spanning several waves, respondents may correctly report receipt again in future waves. This would lead to an over-estimation of the number of repeated spells and of the proportion of transitions onto benefits at the seams. Compared to the administrative data, these errors in the survey data were expected to produce the following biases:

H1: For long spells types, spanning several reference periods, the survey data under-estimate spell durations and over-estimate the number of completed spells and the proportion of transitions off benefits at a seam. The survey data may also over-estimate the number of repeated spells and the proportion of transitions onto benefits at a seam.

H2: For short spell types the survey data may either over- or under-estimate spell durations, depending whether constant wave over- or under-reporting occurred more frequently. The survey data will also over-estimate the proportion of transitions onto and off benefits at the seams.

Dependent interviewing was expected to reduce under-reporting and the associated biases, by reminding respondents of sources reported previously. Lynn et al. (2004) showed that DI was particularly successful at reducing under-reporting of spells which had ended before the interview date, that is, at reducing constant wave under-reporting. DI was however not expected to reduce constant wave over-reporting, since the DI questions did not provide any additional cues to help respondents recall the start date and prevent backward telescoping to the start of the reference period. The effect of DI on constant wave response was therefore asymmetrical: the DI questions queried apparent transitions off benefit receipt at the earlier interview date (caused by constant wave under-reporting), but did not query apparent transitions onto benefit receipt at the interview date (caused by constant wave over-reporting). This leads to the following hypotheses about the effects of DI:

H3: For long spell types the DI survey data produces estimates of spell durations and proportions of seam transitions that are closer to the estimates from the administrative data than the INDI survey data are.

H4: For short spell types the proportion of transitions off benefits at a seam is closer to the administrative records in the DI than the INDI survey data, but the proportion of transitions onto benefits at a seam is not improved. DI also reduces the under-estimation of spell durations due to constant wave under-reporting, but does not reduce the over-estimation of durations due to constant wave over-reporting. As a result, the DI survey data over-estimate spell durations compared to the administrative records, and more so than the INDI survey data.

The errors in entry and exit dates caused by under-reporting and constant wave responses are likely to attenuate the estimated relationships of explanatory factors with conditional exit probabilities. This will especially be the case for time-varying covariates, for which the association between changes in their values during a spell and the conditional exit probability will be weakened. The errors in reported spell durations are also likely to bias estimates of duration dependence, that is, of how the conditional exit probability varies with the duration of the spell. Constant wave reporting is likely produce under-estimates of conditional exit probabilities for spell lengths that are shorter than one reference period. Both wave under-reporting and constant wave reporting are likely to produce over-estimates of conditional exit probabilities at the seams, that is for spell lengths that equal the duration of the reference period (roughly twelve months in this case):

H5: Estimated coefficients from multivariate duration models based on the survey data are smaller and less significant than estimates based on the administrative records.

H6: Compared to the administrative data, in the survey data the conditional exit probabilities are under-estimated for durations shorter than the reference period and over-estimated for durations that are multiples of the length of the reference period.

To the extent that DI improved the dating of events by reducing under-reporting, the attenuation bias in multivariate estimates is reduced and the estimates of duration dependence improved. Biases associated with constant wave over-reporting are not likely to be reduced:

H7: DI produces estimates of coefficients and standard errors in multivariate duration models that are closer to those from the administrative records, than those from the INDI survey data are.

H8: Conditional exit probabilities based on the DI survey data are closer to estimates from the administrative data than estimates based on the INDI data: both the under-estimation for durations shorter than the length of the reference period and the over-estimation for durations that are multiples of the reference period are reduced.

To test these hypotheses, the histories of income receipt were used to estimate characteristics and distributions of spells, determinants of spell durations and patterns of duration dependence. Conclusions about biases caused by measurement error were drawn by comparing estimates based on the INDI survey data and administrative records, using the inflow of spells after January 1999. Conclusions about the effectiveness of dependent interviewing were drawn by replicating this analysis with the DI survey data and administrative records, using the inflow of spells after September 2000. Since the survey and administrative data in each comparison were from the same sample of respondents, standard hypotheses tests assuming independence of samples could not be used to test for differences of estimates across data sources. The DWP records were generated as a by-product of the actual payments and were therefore treated as the gold standard. Any differences in the survey estimates were interpreted as bias.

4 Distribution of spells

I first compared the characteristics of spells in the survey and administrative data: the number of spells, their mean durations, proportion of completed spells, number of repeated spells and the proportion of transitions onto and off income receipt in seam months (Tables 1 to 4). This first descriptive analysis was carried out separately for all income sources and provided an aggregate indicator of errors in reporting. The comparison of different income sources was also used to guide decisions on how to pool sources for the multivariate analyses.

The distributions of spell lengths were then compared using lifetable estimates. This estimator accounts for the fact that spell dates were recorded in months, although in reality the start and end dates could have been on any day of the month. The estimates are based on the assumption that transitions are spread evenly over the month and treat half of the exits as having 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, $n_j = N_j - d_j/2$.

The estimated survivor functions for the different comparisons are presented graphically in Figure 3. For the nine income sources with large enough sample sizes, the survivor functions are shown separately. All income related and all disability related sources were then pooled and separate survivor functions are shown for these two groups.

4.1 Long versus short spell types

The expected effects of errors and DI on estimates depended on the length of different spell types relative to the length of the reference period (about 12 months in the present survey). The expectation was that long spell types, which on average spanned multiple reference periods, were more likely to be affected by wave under-reporting, while spell types which were on average shorter than the length of a reference period were more likely to be affected by constant wave reporting. Long and short spell types were distinguished based on the mean and median durations of spells in the administrative records. Both the mean durations (column 3 in Table 2) and median durations (where the graphs cross the horizontal 0.5 line in Figure 3) show that the disability related benefits, Retirement Pensions and Child Benefit tended to span more than one reference period and are therefore referred to as ‘long benefit types’; the income related benefits tended to be about the length of one reference period or shorter and are referred to as ‘short benefit types’.

4.2 Spell distributions: biases in independent survey data

Hypothesis H1 seems to be supported. For the longer spell types, spell durations were under-estimated in the survey compared to the administrative data: the mean duration including right censored spells (column 3 in Tables 1 and 2) was, for example, under-estimated by about 5 months for Disability Living Allowance (DLA), by about 8 months for Child Benefit (CB) and by about 10 months for Attendance Allowance (AA). Retirement Pension (RP) seemed to be reported with less error, as the survivor function from the survey matched the administrative records relatively well. This may be because RP is a more salient income source, since for pensioners it is typically the main source of income. Incapacity Benefit (IB) looked more like the short-term benefits and is discussed below.

For the DLA and AA spells, the survivor functions from the administrative data reflect the fact that none of the spells ended during the window of observation: 100% of spells were longer than 49 months (Figure 3). The survey data however painted an entirely different picture. According to the survey, 31.4% of DLA spells and 44.4% of AA spells ended during the window of observation. The survivor functions show a steep kink at around 12 months. This probably reflects wave under-reporting: long-term spells that were correctly reported in one interview and under-reported in the following interview. As result over 30% of DLA spells and nearly 50% of AA spells were shorter than 14 months according to the survey estimates. The survivor function for CB spells from the survey data showed a similar kink at about 12 months, although the bias compared to the administrative records was less stark.

The other expectations summarized in H1 also appear to be supported. The proportion of transitions off benefit receipt at a seam in the survey data far exceeded seam transitions in the administrative data. In the absence of any errors producing seam effects, and assuming that transitions were uniformly distributed, one would expect 6.1% of transitions at a seam, since 3 of the 49 months during the window of observation were seam months. In the administrative data there were no exits from DLA or AA. In the survey data, 24.4% of CB spells, 28.6% of DLA spells and 37.0% of AA spells apparently ended at a seam. This is again consistent with the hypothesis that spells were ‘cut off’ at seams by wave under-reporting. There also appeared to be some evidence that respondents who under-reported a source in one wave sometimes reported it again in a later wave: while there were no repeated spells in the administrative data, the average number of spells in the survey data was 1.03 per CB recipient, 1.06 per DLA recipient and 1.23 per AA recipient. In addition, more than two-thirds of spells in the survey data were reported as having started at a seam, compared to 0% of CB spells, 14.3% of DLA spells and 8.3% of AA spells in the administrative data.

Evidence for hypothesis H2 was mixed. Of the short spell types, spell durations were only over-estimated in the survey for Housing Benefit (HB). The median duration was about 11 months according to the administrative data and 34 months according to the survey data. Incapacity Benefit (IB) spells showed similar biases, although this was classified as a long benefit type. For both HB and IB the proportion of completed and the number of repeated spells were smaller in the survey data than the administrative records, and the proportion of start dates at a seam was larger. This suggests that for these spells, constant wave over-reporting may have led to reports for shorter spells being joined together and appearing as fewer but longer spell in the survey data.

For the other short spell types the estimated survivor functions from the survey data matched the estimates from the administrative records more closely. For Working Families’ Tax Credit (WFTC) the survey data only slightly over-estimated spell durations; for Income Support (IS) and Job Seeker’s Allowance (JSA) the survey data somewhat under-estimated spell durations. This could either imply that reporting for these income sources was less affected by errors, or that the effects of constant wave over- and under-reporting on estimated spell durations cancelled each other out. The proportion of start transitions at a seam however again far exceeded the administrative estimates, although less so than for HB spells, suggesting that these income sources were not reported without error.

4.3 Spell distributions: effects of dependent interviewing

Before comparing the reports obtained with dependent interviewing with the administrative records, I first examined whether the proactive and reactive data could be pooled. The percentage of seam starts was similar with both types of DI at around 27%. The percentage of seam ends with PDI was 2.2% and 9.3% with RDI. The estimated survivor functions did not differ significantly according to likelihood ratio tests or log-rank tests of homogeneity across the two DI samples. The sample sizes were however very small for some sources, providing little power to detect differences. Neither group produced estimates of the survivor function that were consistently closer to those from the administrative data. PDI estimates were closer to the administrative data for DLA and JSA, while RDI appeared better for IB and AA.

The combined DI samples provided support for H3, that DI would reduce the under-estimation of spell durations for long benefit types. The survivor functions for DLA, AA and CB spells from the survey matched those from the administrative data more closely than the INDI estimates. The DI estimates still displayed kinks in the survivor functions at around 12 months, suggesting that some under-reporting remained (this was also the conclusion by Lynn et al. 2004), but the effects were small compared to the INDI data: 5.0% of DLA spells, 6.7% of CB spells and 13.3% of AA spells ended in the DI data, compared to 31.4%, 29.3% and 44.4% in the corresponding INDI data (and none in the DI sample of administrative records). The reduction in under-reporting was also visible in the reduction of the proportion of seam transitions: only 5.0% of DLA spells and 6.7% of CB and AA spells in the DI data ended at a seam. This was still higher than the proportion in the corresponding administrative data, where none of the spells ended, but much lower than in the INDI survey data. In the DI survey data, the observed proportion of seam transitions was roughly twice the expected rate of 3.2% (1 of 31 months was a seam in the period covered by the DI data); in the INDI survey data the observed proportion of seam transitions was between four and six times the expected rate of 6.1%. Although DI reduced the excess of transitions *off* income sources at the seams, it did not appear to have any impact on the excess of transitions *onto* income sources at the seams. Both with INDI and DI, the proportion of start dates at a seam were more than 10 times the expected rates (except for CB, where the proportion with DI was only four times the expected rate). If DI was used in more than one wave it would, however, reduce the excess of start dates at seams for long spells. Since DI reduced the under-reporting of sections of long spells in the current wave, it would reduce the number of repeated spells appearing to start at a seam in the following wave.

Hypothesis H4, that DI would increase the net over-estimation of spell durations for short benefit types was supported. For IS and WFTC, the mean durations including right censored spells were about 4 months longer with DI than in the corresponding administrative records. This increase in spell lengths was reflected in the survivor functions, where the median duration of WFTC spells, for example was twice as long with DI as with the corresponding administrative data, although the INDI estimate was very close to the administrative data. A similar worsening of estimated spell durations with DI occurred for IB, where DI worsened the over-estimation of spell durations in the survey data. These findings are consistent with the hypothesis that DI reduced the under-estimation of spell durations caused by constant wave under-reporting, but did not reduce the over-estimation of durations caused by constant wave over-reporting.

The JSA spells showed a completely different effect, whereby DI increased the under-estimation of spell durations in the survey data. The JSA spells were the only spell type for which DI worsened the under-estimation and it is not clear why the effect was different for this income source.

5 Determinants of spell durations and duration dependence

Hypotheses H5 to H8 were tested using multivariate duration models to examine the determinants of spell durations and patterns of duration dependence. The explanatory variables were based on eligibility criteria for the different income sources and similar to those used in other studies (for example, Ashworth et al. 1997; Blank 1989; Hoynes and MaCurdy 1994; Long 1990; O'Neill et al. 1987; Ruggles 1989). The explanatory variables were derived from the survey data and merged with the spells from both the survey and the administrative data. Explanatory variables were time-varying, unless stated otherwise. For some time-varying variables the exact dates of changes were known. For others changes were only observed at the time of each interview. The fact that the timing of change for some of the time-varying covariates was measured imperfectly is likely to bias the associations of these variables with conditional exit hazards. Any such attenuation bias should, however, be the same in the estimates from the survey and the administrative data. Similarly, any errors in the measurement of the explanatory variables should affect the survey and administrative data in the same way. The comparison of estimates from the two data sources should therefore not be affected by potential problems with the measurement of the explanatory variables, although the potential measurement errors may reduce power to detect differences in coefficients between estimates from the two data sources. The variables included as predictors of spell durations were:

Personal characteristics and social background: age, gender (fixed), highest educational qualification (at the interview date), whether married or cohabiting (the exact dates of changes in legal marital status were known, but changes in whether the partner was living in the household were only observed at each interview) and region of residence (the exact moving date was known, as long as respondents moved only once between interviews). Ethnic group was not controlled for, since 98% of respondents included in the analysis were of white origin.

Factors related to eligibility for disability related benefits: whether the respondent was long-term sick or reported chronic health problems (observed at each interview).

Factors related to eligibility for income related benefits: Some of the covariates were predictors of labour market attachment, including the number of own children under 16 in the household, the age of the youngest (the exact birthdates were known, but whether a child had left the household was only known at the interview), current labour market activity and number of months unemployed to date during the panel period. No measure of wages (either previous or predicted) was included, since the factors typically used to predict wages were already included in their own right. Outside opportunities were captured by local unemployment rates (at the date of interview).[1] The covariates also included predictors of eligibility for means-tested benefits, some of which are assessed at the level of the benefit unit (defined as an adult living with their spouse and any dependent children). These were housing tenure (at each interview), as a measure of wealth, and the partner's employment status (at the time of interview) from the household grid. The partners' full interviews would have provided more detailed information, but their use was limited since not all partners had given interviews.

The duration models were estimated using a discrete time proportional hazards specification (cloglog), which is appropriate for event history data where the underlying dates are continuous, but only measured in discrete intervals (see Allison 1982). The estimates included different specifications for the baseline hazard. The first model included only substantive explanatory variables. The second specification was a fully flexible model, including a dummy for every month in which an exit was observed, excluding the first month as the reference category. This non-parametric specification should give an idea of the pattern of duration dependence. Since the number of exits observed per month was small, it would however lead to a loss in precision. I therefore also tested two parametric specifications, the discrete time equivalent of the continuous time Weibull model and a polynomial specification.

The analysis does not account for unobserved heterogeneity, which may lead to downward 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. All models do, however allow for

[1] The local unemployment rate was based on the travel-to-work areas (using 1998 boundaries), of which there are about 300 in the UK. For 1997-2000 the rate was the proportion of unemployed in the labour force. Because the Office for National Statistics discontinued the labour force measure, the 2001 and 2003 rates were based on the proportion of claimants in the resident working age population from the 2001 Census, which tended to produce lower rates.

clustering to adjust for multiple spells per sample member, especially since different income sources were pooled for the analysis.

The small numbers of spells for each income source meant that it was necessary to pool sources for the multivariate analyses. Ideally, the pooled sources should have similar characteristics in terms of the distribution of durations, the predictors of exit probabilities and the nature of reporting errors. The previous section showed that sources roughly fell into two categories depending on their durations. The factors associated with moves onto and off income sources are likely to differ between those sources for which eligibility is related to disability and those for which eligibility is related to income. I therefore pooled the disability related benefits, which were also longer term and subject to wave under-reporting on the one hand, and the income related benefits on the other hand, which tended to be shorter term and subject to constant wave reporting. Retirement Pension and Child Benefit were not included in the multivariate models, since they are universal benefits and exits should be determined by death or the age (or educational status) of the youngest child. Housing Benefit had to be dropped from the analysis, because the lack of exact end dates in the administrative records distorted the multivariate estimates. The last two panels of Figure 3 plot the estimated survivor functions for the pooled sources. For the disability related sources, the INDI data under-estimated spell durations, while DI over-estimated spell durations. For the income related benefits, both INDI and DI lead to over-estimates of spell durations.

Table 5 shows the summary statistics for the explanatory variables included in the models. The statistics are based on the administrative data and reflect the characteristics of recipients in the first month of every spell. For the local unemployment rate the mean across all spell months is reported. The statistics are shown separately for the income and disability related benefits, and within these, separately for the respondents allocated to INDI and DI. For each benefit type, the characteristics of recipients may differ slightly between the administrative data corresponding to the DI and the INDI samples. Since both were random samples of the ECHP sample, any differences are most likely due to sampling variation. Some differences may however have occurred because the samples cover different time periods.

Both the INDI and the DI samples were predominantly female, with low qualifications, around half were married or cohabiting and between 10 and 20% lived in London or the South East. Among recipients of income related benefits, about 30% 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, about 30% were owner occupiers, nearly 70% were 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 around 3.5%. Among recipients of disability related benefits, around 84% reported chronic health problems and around 30% 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), 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 disability related spells from the INDI survey and administrative records. The latter tables report the results for both sets of income sources from the DI survey and administrative data.

Figures 4 and 5 then present the predicted hazard rates for models 2 to 4, at the means of the continuous variables in the corresponding sample of administrative records and setting the binary indicators equal to their modal value. The resulting exit probabilities, predicted for different spell durations, were for females, without qualifications, married or cohabiting, who did 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 were set to the sample mean based on the administrative data (see Table 5).

5.1 Multivariate duration models: biases in independent survey data

Hypothesis H5, that coefficients and standard errors based on the survey data would tend to be smaller and less significant than those based on the administrative data, found some support. For income related spells (Table 6), the estimates from the administrative records suggested a larger hazard rate, and hence shorter spells, for those in work or married/cohabiting and, in model 1, the exit hazard decreased with the number of children. In the survey data, the hazard was also larger for those in work, although the coefficients in all models were smaller and less significant. Whether a respondent was married or cohabiting had a similar effect in the survey as in the administrative records, but the number of children had no effect. The main difference compared to the administrative records, was that educational qualifications had a significant effect in all models, reducing the exit hazard. That is, according to the survey data, respondents with lower educational qualifications tended to have longer income related spells. In the administrative data, spell length was not related to qualifications. This suggests that compared to respondents with more education, those with lower education were more likely to constant wave over-report their short-term income sources.

For disability related benefits the estimates based on the administrative data suggested that exit hazards increased with age, but at a decreasing rate. Respondents who were long-term sick had significantly lower exit hazards, and hence longer spells. In the survey data none of the covariates were significant, except for being long-term sick in model 2. In this case, estimated effects of determinants of spell durations were clearly attenuated. (For both data sources, the tiny number of exits from disability related benefits was however a potential problem; see Tables 1 and 2.)

Hypothesis H6, that duration dependence would be attenuated in the survey data, was also supported. For income related benefits, the polynomial model suggested that the hazard changed non-monotonically with the duration of a spell. This is illustrated in the first three graphs in Figure 4, which show that the predicted hazard rates based on the records were non-monotonic, with a sharp rise and fall during the first 12 months. This pattern of duration dependence was reflected in the polynomial specification, but not the Weibull which does not allow for non-monotonic changes in hazard rates. This is confirmed by the Akaike Information Criterion, AIC, according to which the fully flexible model fit the records best, followed by the polynomial specification. For the survey data, there was no clear distinction in fit between models and the polynomial time variables were not significant. Figure 4 shows that the estimated hazard rates were relatively constant at lower levels than the estimates based on the administrative records, except for a spike at month 13, which roughly corresponded to the length of the reference period. As a result, the polynomial specification did not reflect the non-monotonic duration dependence.

For the disability related spells, the estimates based on the administrative data also suggested significant non-linear effects of spell duration and, according to the AIC, the fully flexible specification again fit the data best. The predicted hazard rates in the first three graphs of Figure 5, reflect the non-monotonic pattern in the record data, with hazards increasing after month 12 and then falling. This pattern is reflected in the polynomial predictions. In the survey data the polynomial time variables were not significant. The predicted hazard rates again display a large spike in month 13.

5.2 Multivariate duration models: effects of dependent interviewing

Support for hypothesis H7 was weak and it is not clear to what extent DI reduced the attenuation bias in estimated coefficients. For the income related spells, the coefficients for being in work were slightly larger with the survey than the administrative data and months unemployed during the panel was significant in all models, unlike in the administrative data. Being married or cohabiting and the age of the youngest child, both of which were significant in all models from the administrative data, were not significant in any of the models based on the DI survey data. Age, however, was significant, unlike in the administrative data. From these variables it is therefore not clear that DI reduced bias in estimates. The main difference compared to the INDI survey data, was however that the educational qualifications were no longer significant in the DI data. This suggests that DI particularly helped respondents with low education and reduced the extent to which they constant wave over-reported short-term income sources. As a result the estimates were closer to those from the administrative records, in that respondents without educational qualifications no longer appeared to have longer spells than those with qualifications. In this sense, DI did appear to reduce an important bias in the predictors of exit hazards. For the disability related spells the only significant predictor in the administrative data was whether respondents reported chronic health problems. In the DI survey data, the only significant predictor was whether they reported their labour market status as being long-term sick. For both data sources, the tiny numbers of exits during the window of observation were probably a problem and so these estimates should be interpreted with caution.

The data also show some support for H8, that DI would reduce bias in estimated duration dependence. For income related spells the polynomial effects of spell duration were again not significant. Nonetheless, comparing the predicted hazard rates in Figure 4 suggests that the DI data mapped the increasing and then falling hazard rates in the record data more closely than the INDI data. The DI data did not display the spike in hazard rates in month 13, which reflects the smaller proportion of transitions out of spells at the seam. At the same time the hazard rates for durations shorter than one reference period were higher with DI than with INDI, although still consistently lower than with the administrative data. For the disability related benefits there were significant non-monotonic effects of time, as in the administrative data over the longer period of observation for the comparison with INDI, and the polynomial specification fit the data best. The predicted hazard rates in Figure 5 show that the survey data no longer displayed the spike in month 13 and the polynomial predictions followed the record data, although as for the income related spells, the hazard rates with the DI data were consistently lower than with the records.

6 Summary and conclusions

This paper has provided new evidence on the effects of errors in event history data from panel surveys, using a unique data set of matched survey and administrative data, which included an experiment comparing different data collection methods. The first contribution this paper makes is to spell out the hypotheses about how different error scenarios are likely to affect estimates from event history models. The second contribution is to illustrate the effects errors in the reporting of events and dates can have on estimated spell distributions, determinants of exit probabilities and patterns of duration dependence. The third contribution is to assess how effective dependent interviewing was at reducing biases in estimates from event history data. Previous studies of the effects of dependent interviewing focused on estimates of prevalence and monthly transition rates and did not evaluate the implications for the types of (multivariate) analyses for which event history data are typically used. The findings illustrated the importance of evaluating data collection methods by examining their effects on the types of analyses for which the data are likely to be used. For the present study, previous tests had shown that dependent interviewing reduced under-reporting and the excess of seam transitions. The estimates from event history models showed that these improvements in data quality did reduce bias for most estimates. For some estimates dependent interviewing however increased biases, since it reduced the under-reporting of events, but probably did not improve the dating of events.

The findings suggested support for the hypothesis that under-reporting receipt for some but not all reference periods, in the case of sources received continuously over several waves, led to an under-estimation of spell durations and an over-estimation of exits in the seam months between reference periods and the number of repeated and completed spells. The effects of errors in reports of spells which were likely to have started or ended during a reference period were less clear. It seemed likely that the effects on estimated spell durations, of falsely reporting the receipt status at the time of the interview for all months in the reference period, cancelled out for some income sources; for others the errors led to either over- or under-estimation of spell durations. In all cases, the errors produced an excess of seam transitions, for both start and end dates.

The hypothesis that errors in the reporting and dating of receipt would attenuate the estimated effects in multivariate duration models was also supported. The results were somewhat mixed, but coefficients and their significance levels tended to be smaller in the survey than the administrative data. The main difference was that according to the survey data, respondents with lower educational qualifications had significantly longer durations of receipt for income related spells, although qualifications were not related to durations in the administrative data. This finding warrants further investigation. It suggests that respondents with lower cognitive ability may have been more likely to falsely report receipt for all months in the reference period than other respondents, leading to an over-estimation of spell lengths for this group.

The findings also supported the hypothesis that errors would attenuate estimated patterns of duration dependence. The survey data did not reflect the non-linear estimate from the administrative records. Instead there were spikes in the conditional exit probabilities for spell durations of about 12 months, corresponding to the length of the reference period. Exit probabilities for durations shorter than 12 months were under-estimated.

The hypotheses about the effects of dependent interviewing on the descriptive and multivariate estimates were also supported. Dependent interviewing seemed to reduce under-reporting and as a result reduced the under-estimation of spell durations for spells that spanned several reference periods. As a result, the over-estimation of the number of repeated and completed spells and of exit transitions at the seam were also reduced. Dependent interviewing did however not have any impact on over-reporting of income sources early in the reference period. As a result, the excess of transitions onto benefits at the seams was not reduced and estimates of spell durations were worsened in some cases: for income sources for which the independent survey data had produced estimates of spell durations that were close to those from the administrative data, the estimates based on the DI data grossly over-estimated the durations.

Whether the hypotheses about the effects of DI on biases in estimates from the multivariate models were supported was less clear. The main effect was that the duration of income related income sources was no longer related to educational qualifications. This was possibly an important reduction of bias. The estimates of duration dependence matched the administrative data slightly better: the over-estimation of exit hazards at around 12 months was reduced, as was the under-estimation of exit hazards for durations shorter than 12 months. Conclusions about the effects of DI on multivariate models were however hampered by small sample sizes.

These findings suggest that the extent and nature of reporting errors depend on the mean lengths of spells relative to the length of the reference period between interviews. Spells with long durations relative to the reference period (in this case disability related benefits) seemed more affected by 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 income 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 US 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 of spells which start or end during the reference period.

In this study dependent interviewing was only used in a single interview. Used in successive interviews, dependent interviewing may lead to increasing over-estimation of spell durations, if consecutive short spells are falsely merged because of constant wave over-reporting. Survey designers may therefore 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. On a brighter note, using DI over more than one wave would begin to reduce the over-estimation of start transitions at the seams, since reducing the under-reporting of long income spells in intermittent waves would reduce the number of spurious repeat spells starting at seams.

References:

Allison, P. D. (1982) 'Discrete-Time Methods for the Analysis of Event Histories', *Sociological Methodology*, 13: 61-98.

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. (2008) 'The Causes of Seam Effects in Panel Surveys', *ISER Working Paper, 2008-14,* Colchester: University of Essex.

— (in press) 'Dependent Interviewing: A Framework and Application to Current Research', in P. Lynn (ed) *Methodology of Longitudinal Surveys.* Chichester: Wiley.

Jenkins, S. P., Cappellari, L., Lynn, P., Jäckle, A. and Sala, E. (2006) 'Patterns of Consent: Evidence from a General Household Survey', *Journal of the Royal Statistical Society, Series A*, 169(4): 701-722.

Jenkins, S. P., Lynn, P., Jäckle, A. and Sala, E. (2008) 'The Feasibility of Linking Household Survey and Administrative Record Data: New Evidence from Britain', *International Journal of Social Research Methodology,* 11(1): 29-43.

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', Washington, DC: The Urban Institute.

Lynn, P., Buck, N., Burton, J., Jäckle, A. and Laurie, H. (2005) 'A Review of Methodological Research Pertinent to Longitudinal Survey Design and Data Collection', *ISER Working Paper, No.* 2005-29, Colchester: University of Essex.

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

— (2006) 'The Effects of Dependent Interviewing on Responses to Questions on Income Sources', *Journal of Official Statistics*, 22(3): 357-384.

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.

Moore, J. C., Bates, N., Pascale, J. and Okon, A. (in press) 'Tackling seam bias through questionnaire design', in P. Lynn (ed) *Methodology of Longitudinal Surveys.* Chichester: Wiley.

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 income and within each group are sorted in order of prevalence according to the administrative records. The following abbreviations are used:

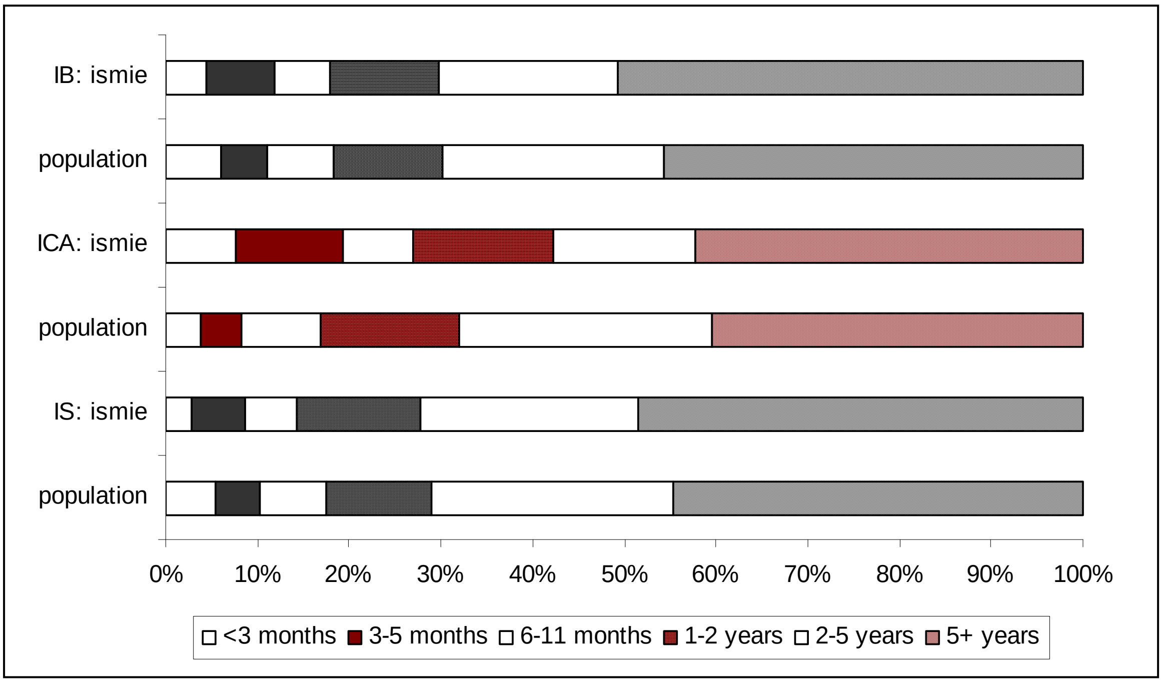
INDI Independent Interviewing
DI Dependent Interviewing

Disability related benefits:
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

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

Other:
CB Child Benefit (including One Parent Benefit)
RP Retirement Pension
WB Widows Benefit

Figure 1: Duration of caseloads – ISMIE sample compared to GB population



Source: Administrative records for ISMIE sample and population data from the Work and Pensions Longitudinal Study of 100% of claimants, available from [accessed 24.03.2008] http://83.244.183.180/100pc/tabtool.html

Figure 2: Editing rules for overlapping reference periods in income receipt histories

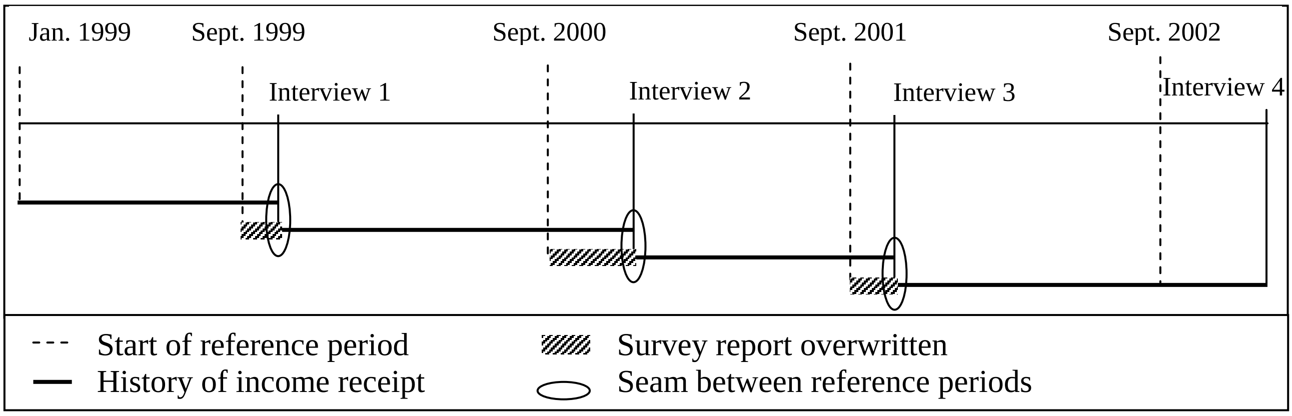


Table 1: Independent interviewing – inflow after January 1999

Source	Left censored spells	Inflow spells	Mean duration (mths)	Completed spells (%)	Mean duration (mths)	Mean spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	52	35	19.89	31.43	11.45	1.06	68.57	28.57	22.86
IB	38	34	16.79	38.24	9.46	1.03	41.18	26.47	14.71
AA	20	27	15.26	44.44	8.33	1.23	70.37	37.04	22.22
ICA	9	16	16.69	43.75	12.71	1.07	50.00	31.25	25.00
IID	12	5	14.60	80.00	13.75	1.25	100.00	80.00	80.00
SDA	15	10	20.00	40.00	16.50	1.00	100.00	40.00	40.00
DPT	4	0	–	–	–	–	–	–	–
HB	131	106	19.87	33.96	12.92	1.14	63.21	23.58	16.98
IS	78	89	14.37	48.31	9.30	1.25	48.31	26.97	17.98
JSA	23	56	6.04	75.00	4.86	1.33	26.79	21.43	5.36
WFTC	26	70	17.23	52.86	9.62	1.13	37.14	44.29	14.29
CB	116	41	22.66	29.27	12.00	1.03	68.29	24.39	21.95
RP	142	31	28.65	3.23	25.00	1.00	48.39	3.23	3.23
WB	6	6	13.17	66.67	10.25	1.00	50.00	66.67	50.00
Total	672	526	17.19	42.97	9.71	1.14	52.66	28.33	17.30

Table 2: Administrative records – inflow after January 1999

Source	Left censored spells	Inflow spells	Mean duration (mths)	Completed spells (%)	Mean duration (mths)	Mean spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	44	14	24.50	0.00	–	1.00	14.29	0.00	0.00
IB	39	44	16.45	50.00	11.91	1.07	6.82	6.82	0.00
AA	15	12	25.42	0.00	–	1.00	8.33	0.00	0.00
ICA	8	9	17.00	11.11	36.00	1.00	0.00	0.00	0.00
IID	7	3	24.67	0.00	–	1.00	0.00	0.00	0.00
SDA	2	1	37.00	0.00	–	1.00	0.00	0.00	0.00
DPT	0	2	10.00	100.00	10.00	1.00	0.00	0.00	0.00
HB	96	132	13.02	66.67	9.74	1.76	2.27	3.03	0.00
IS	65	72	15.53	41.67	10.87	1.13	5.56	2.78	1.39
JSA	10	95	6.42	82.11	6.28	1.52	5.26	6.32	0.00
WFTC	3	77	18.83	63.64	12.98	1.33	7.79	5.19	1.30
CB	94	13	30.38	15.38	20.00	1.00	0.00	0.00	0.00
RP	138	30	27.03	0.00	–	1.00	10.00	0.00	0.00
WB	4	3	10.33	0.00	–	1.00	0.00	0.00	0.00
Total	525	507	15.36	53.65	9.81	1.31	5.33	3.75	0.39

[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	Mean duration (mths)	Completed spells (%)	Mean duration (mths)	Mean spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	52	20	19.85	5.00	13.00	1.00	30.00	5.00	0.00
IB	48	22	14.09	22.73	4.60	1.00	27.27	4.55	0.00
AA	27	15	16.07	13.33	7.00	1.07	40.00	6.67	0.00
ICA	10	6	24.00	0.00	–	1.00	16.67	0.00	0.00
IID	7	2	9.50	50.00	1.00	1.00	50.00	0.00	0.00
SDA	12	6	20.50	33.33	13.00	1.00	33.33	33.33	0.00
DPT	3	1	29.00	0.00	–	1.00	0.00	0.00	0.00
HB	178	82	19.22	13.41	8.73	1.03	37.80	2.44	0.00
IS	110	55	14.71	30.91	7.88	1.08	29.09	7.27	0.00
JSA	32	31	3.23	83.87	3.46	1.63	12.90	9.68	3.23
WFTC	56	44	15.02	36.36	10.19	1.07	18.18	15.91	0.00
CB	127	15	26.67	6.67	13.00	1.00	13.33	6.67	0.00
RP	161	21	20.29	0.00	–	1.00	19.05	0.00	0.00
WB	3	4	14.75	25.00	11.00	1.00	0.00	25.00	0.00
Total	826	324	16.34	25.62	7.04	1.07	26.85	7.10	0.31

Table 4: Administrative records – inflow after September 2000

Source	Left censored spells	Inflow spells	Mean duration (mths)	Completed spells (%)	Mean duration (mths)	Mean spells/ person[1]	Seam start (%)	Seam end (%)	Seam start and end (%)
DLA	43	9	15.22	0.00	–	1.00	11.11	0.00	0.00
IB	49	24	10.67	41.67	7.40	1.09	8.33	0.00	0.00
AA	18	12	15.42	0.00	−	1.00	0.00	0.00	0.00
ICA	10	5	13.40	0.00	−	1.00	0.00	0.00	0.00
IID	6	1	12.00	0.00	−	1.00	0.00	0.00	0.00
SDA	2	0	−	0.00	−	0.00	−	−	−
DPT	2	0	−	0.00	−	0.00	−	−	−
HB	168	116	8.14	65.52	6.17	1.90	0.00	1.72	0.00
IS	89	45	10.60	35.56	7.63	1.10	0.00	0.00	0.00
JSA	32	52	6.79	78.85	6.71	1.21	9.62	1.92	0.00
WFTC	53	32	11.63	46.88	9.93	1.14	0.00	6.25	0.00
CB	105	3	23.00	0.00	−	1.00	0.00	0.00	0.00
RP	158	19	14.00	0.00	−	1.00	0.00	0.00	0.00
WB	1	4	10.50	0.00	−	1.00	0.00	0.00	0.00
Total	736	322	9.88	49.07	6.89	1.30	2.48	1.55	0.00

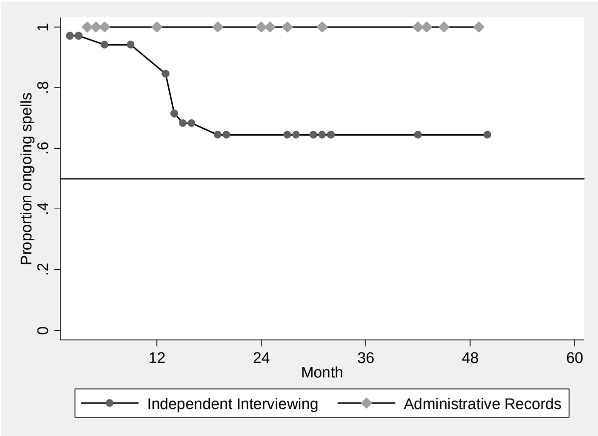
[1]Mean number of spells, excluding sample members with zero spells of a given type.

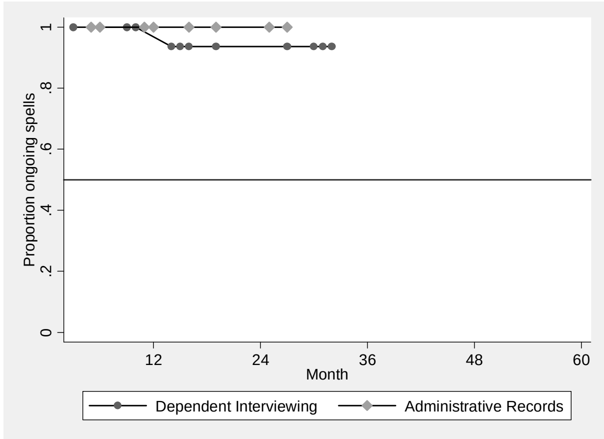
Notes: The window of observation covered 32 months, so 31 potential month-to-month transitions of which 1 was a seam. With a uniform distribution of transitions we would therefore expect 1/31*100=3.23% of transitions to be at the seam.

Figure 3: Lifetable estimates of survivor functions

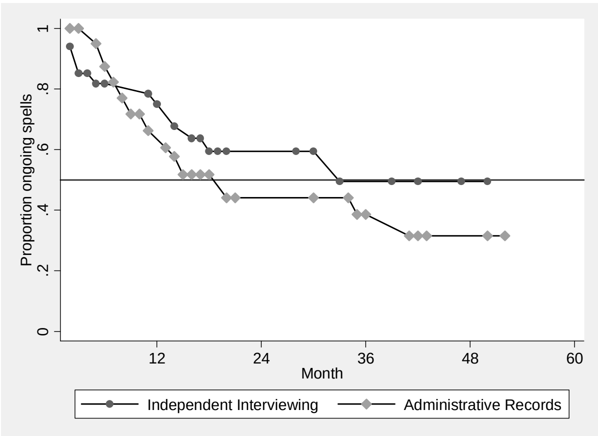
The horizontal bar at y=.5 indicates median spell duration, by which 50% of spells have ended.

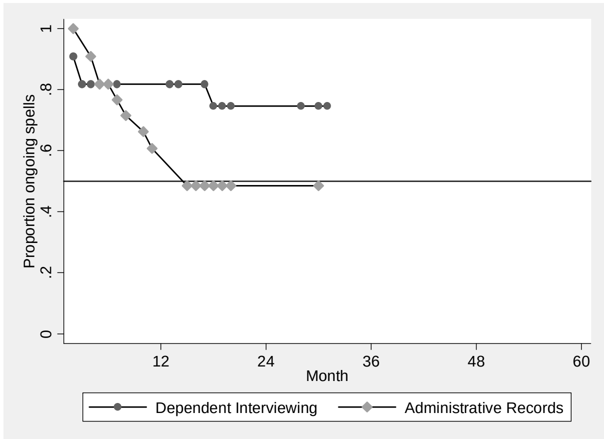
1. Disability Living Allowance



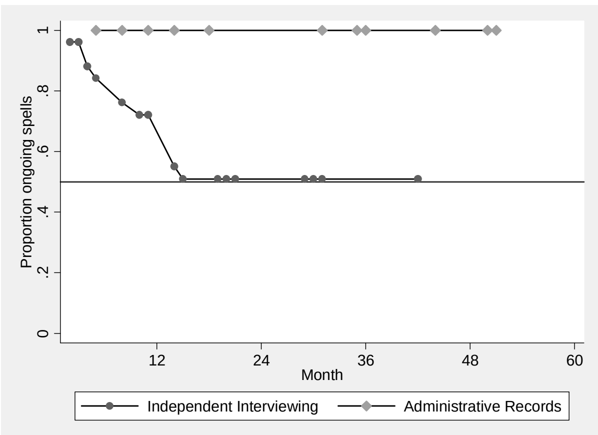


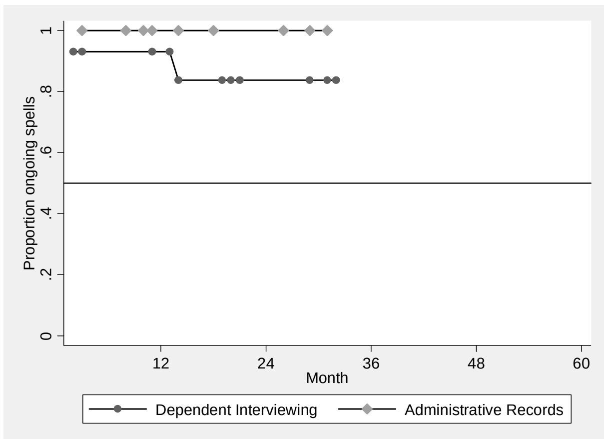
2. Incapacity Benefit



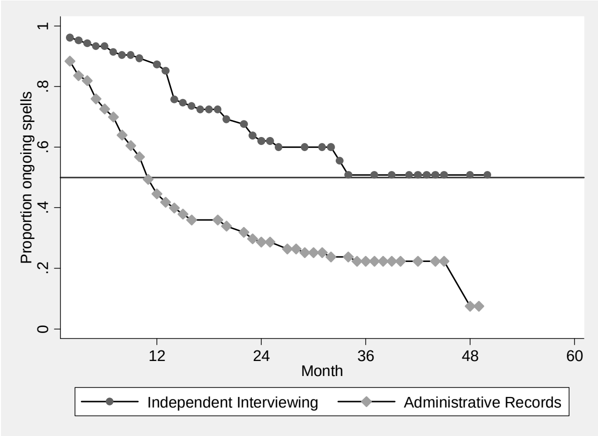


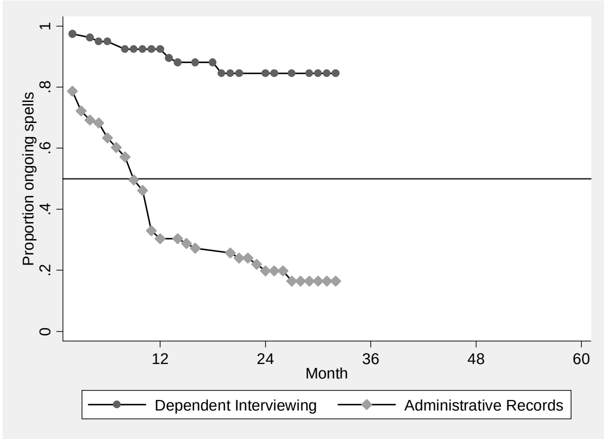
3. Attendance Allowance



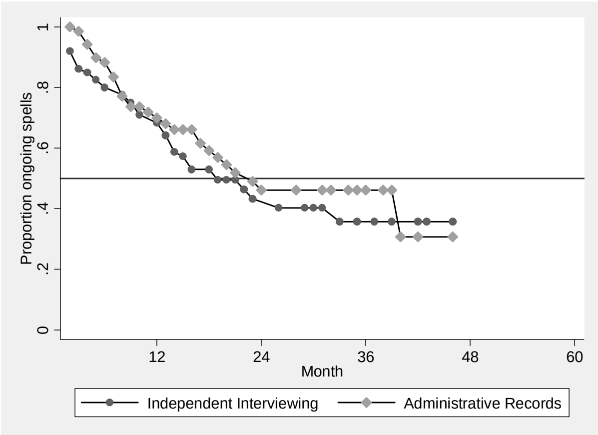


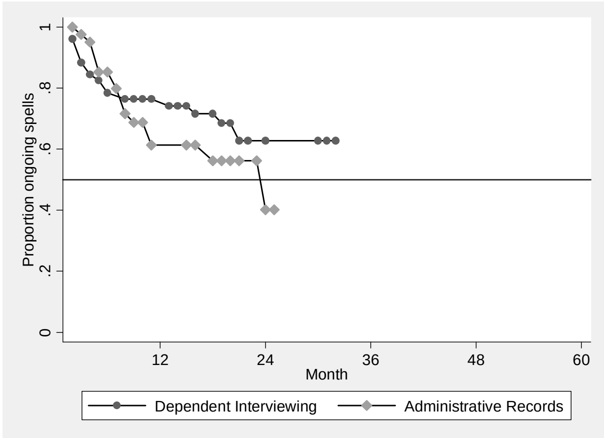
4. Housing Benefit



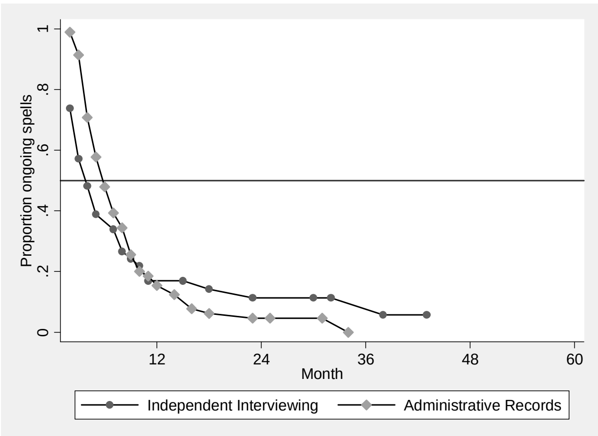


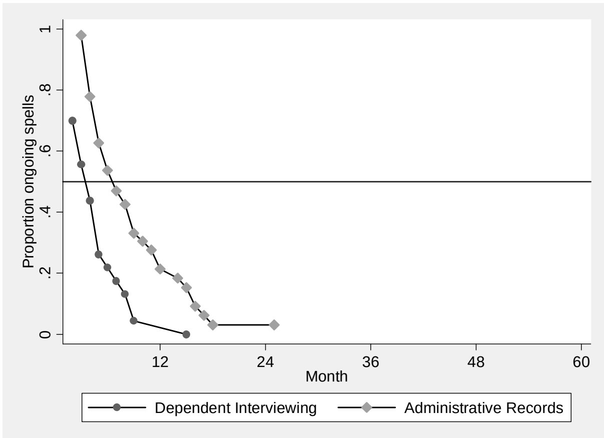
5. Income Support



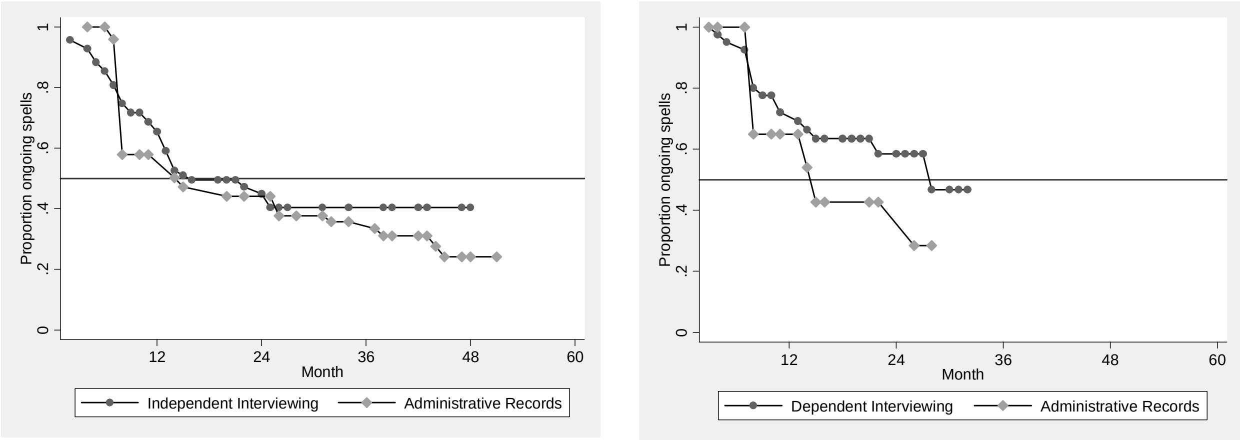


6. Job Seeker's Allowance

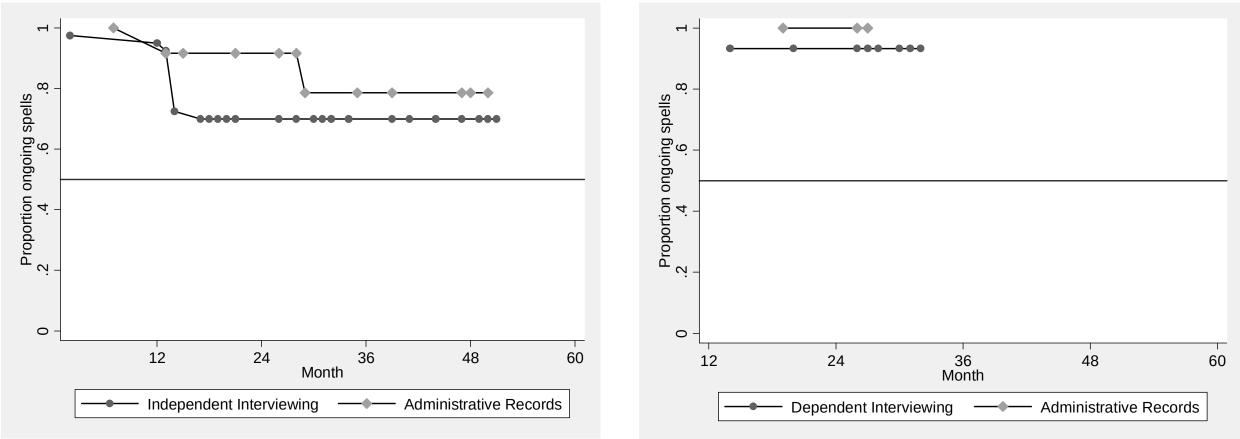




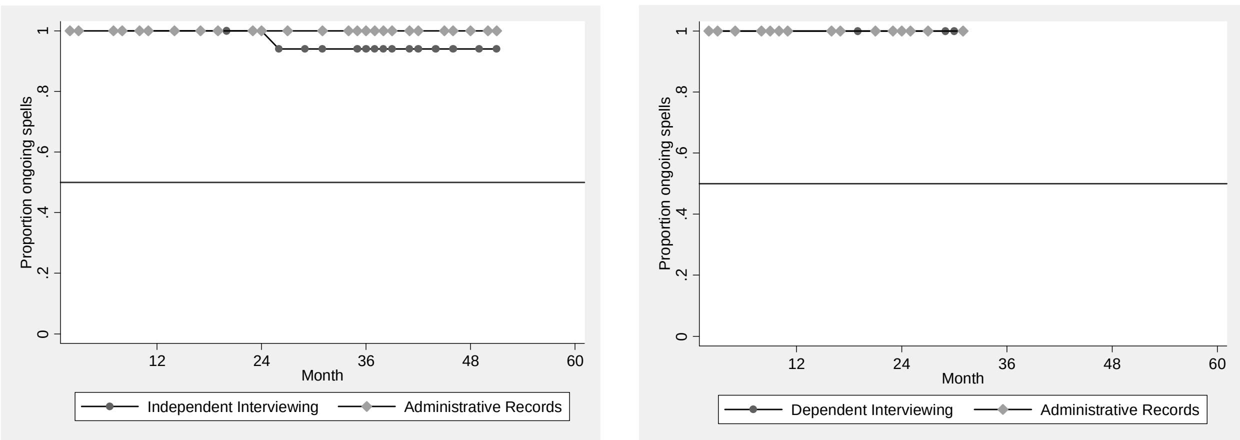
7. Working Families' Tax Credit



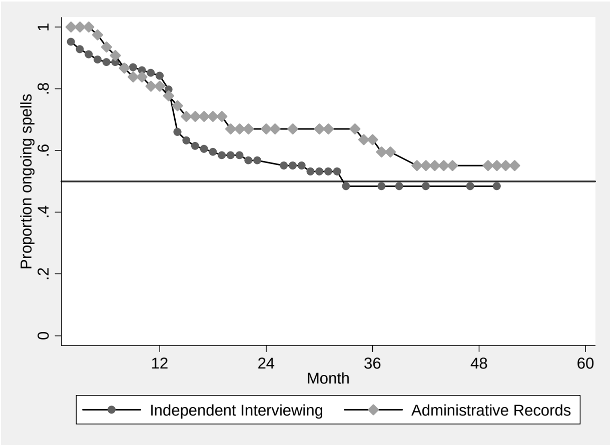
8. Child Benefit

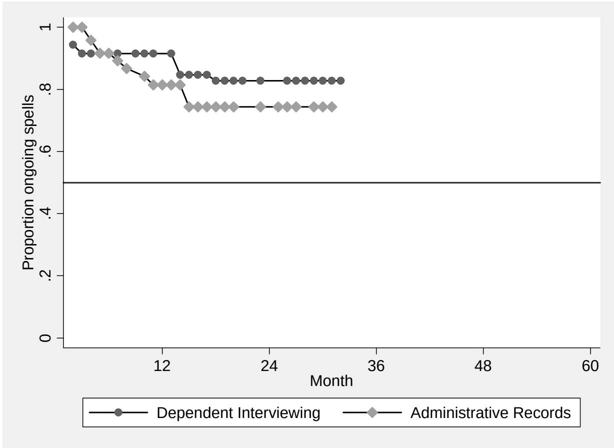


9. Retirement Pension

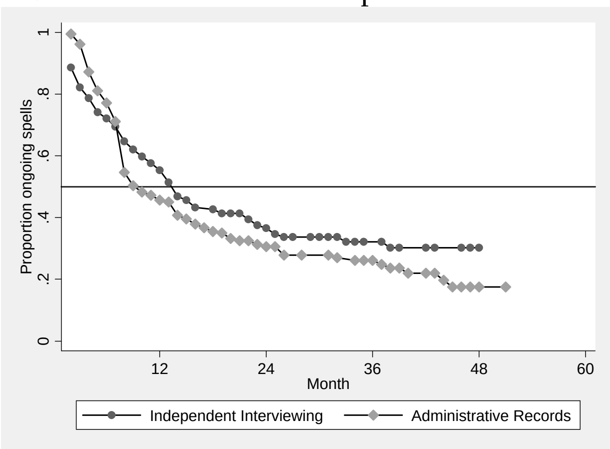


10. Disability related sources pooled





11. Income related sources pooled



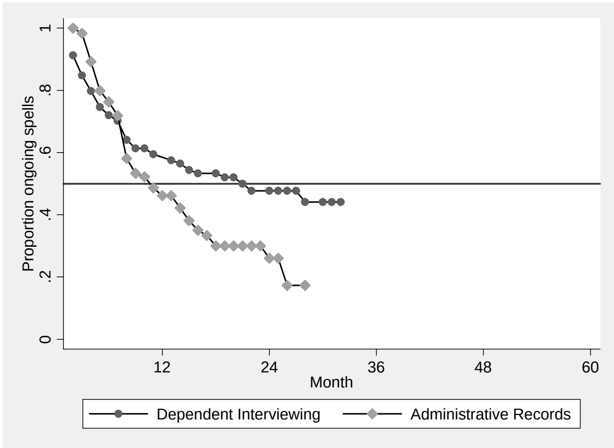


Table 5: Summary statistics of explanatory variables

	Income related benefits				Disability related benefits			
	INDI		DI		INDI		DI	
	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.	Mean	St.Dev.
age	35.419	14.592	36.953	17.300	46.776	19.047	53.451	21.368
male	0.386	0.488	0.380	0.487	0.424	0.497	0.451	0.503
no qualifications	0.258	0.439	0.271	0.446	0.424	0.497	0.490	0.505
married/cohabiting	0.427	0.496	0.442	0.499	0.576	0.497	0.490	0.505
London/South East	0.172	0.378	0.163	0.371	0.153	0.362	0.118	0.325
spouse employed	0.314	0.465	0.294	0.457	–	–	–	–
# children <16	1.124	1.285	0.953	1.292	–	–	–	–
age youngest <16	3.508	4.486	2.380	3.719	–	–	–	–
unemployment rate[1]	3.629	1.750	3.257	1.590	–	–	–	–
own house	0.324	0.469	0.287	0.454	–	–	–	–
self-/employed	0.444	0.498	0.388	0.489	–	–	–	–
unemployed	0.253	0.436	0.287	0.454	–	–	–	–
months unemployed	6.440	10.249	7.946	10.424	–	–	–	–
health problems	–	–	–	–	0.835	0.373	0.843	0.367
long-term sick	–	–	–	–	0.365	0.484	0.255	0.440
N spells	241	–	129	–	85	–	51	–

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

Table 6: Duration models for income 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	1.272**	0.402	1.331**	0.424	1.145**	0.390	1.139**	0.380
JSA	2.144***	0.420	2.321***	0.448	2.002***	0.423	1.913***	0.419
age/10	-0.215	0.349	-0.246	0.368	-0.106	0.337	-0.120	0.328
age/10 squared	-0.005	0.036	-0.002	0.038	-0.011	0.035	-0.011	0.034
male	0.182	0.231	0.137	0.248	0.056	0.214	0.150	0.209
no qualifications	-0.681*	0.280	-0.740*	0.293	-0.729**	0.275	-0.665*	0.264
married/cohabiting	0.598*	0.280	0.611*	0.297	0.517	0.274	0.551*	0.268
London/SE	0.200	0.269	0.183	0.283	0.130	0.260	0.175	0.251
spouse employed	0.090	0.265	0.076	0.266	-0.005	0.244	0.044	0.247
number of children	-0.264	0.189	-0.279	0.196	-0.249	0.179	-0.245	0.179
age of youngest	-0.019	0.046	-0.018	0.048	-0.008	0.042	-0.012	0.042
unemployment rate	0.022	0.071	0.026	0.075	-0.014	0.065	0.000	0.064
own house	0.152	0.257	0.166	0.268	0.210	0.235	0.165	0.230
(self-)employed	0.918*	0.359	1.002**	0.383	0.847*	0.355	0.810*	0.343
unemployed	-0.030	0.394	0.004	0.406	-0.019	0.380	-0.056	0.376
months unemployed	-0.006	0.010	-0.007	0.011	0.002	0.010	0.000	0.010
t	–	–	–	–	-0.007	0.078	–	–
t^2	–	–	–	–	0.000	0.006	–	–
t^3	–	–	–	–	0.000	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.237*	0.095
_cons	-3.480***	0.671	-3.823***	0.708	-3.283***	0.668	-3.101***	0.629
# spell-months	2807	–	2807	–	2807	–	2807	–
# parameters	17	–	40	–	20	–	18	–
log likelihood	-448.11	–	-438.20	–	-442.98	–	-445.14	–
AIC	930.21	–	956.39	–	925.97	–	926.28	–
Records								
IS	0.921*	0.364	0.985**	0.374	0.897*	0.369	1.010**	0.389
JSA	1.842***	0.274	1.910***	0.300	1.893***	0.291	2.042***	0.319
age/10	0.147	0.394	0.029	0.409	-0.014	0.412	0.073	0.427
age/10 squared	-0.057	0.050	-0.041	0.051	-0.040	0.052	-0.053	0.054
male	-0.069	0.187	-0.048	0.209	-0.092	0.207	-0.072	0.211
no qualifications	0.077	0.242	0.207	0.249	0.150	0.247	0.095	0.267
married/cohabiting	0.509*	0.222	0.437	0.231	0.530*	0.228	0.506*	0.250
London/SE	-0.170	0.206	-0.343	0.208	-0.265	0.210	-0.133	0.224
spouse employed	0.333	0.212	0.344	0.210	0.329	0.211	0.394	0.241
number of children	-0.215*	0.106	-0.180	0.106	-0.206	0.110	-0.210	0.116
age of youngest	0.030	0.027	0.038	0.026	0.044	0.026	0.027	0.029
unemployment rate	-0.043	0.045	-0.055	0.049	-0.071	0.048	-0.033	0.049
own house	0.103	0.168	0.171	0.179	0.148	0.180	0.115	0.183
(self-)employed	0.929**	0.292	1.115***	0.310	1.001**	0.305	0.921**	0.308
unemployed	-0.184	0.278	0.006	0.300	-0.049	0.283	-0.158	0.287
months unemployed	0.007	0.007	0.002	0.009	0.003	0.009	0.002	0.008
t	–	–	–	–	0.330***	0.057	–	–
t^2	–	–	–	–	-0.021***	0.004	–	–
t^3	–	–	–	–	0.000***	0.000	–	–
ln(t)	–	–	–	–	–	–	0.206*	0.082
_cons	-4.102***	0.699	-7.608***	1.291	-4.794***	0.775	-4.436***	0.771
# spell-months	3154	–	3154	–	3154	–	3154	–
# parameters	17	–	46	–	20	–	18	–
log likelihood	-549.76	–	-477.44	–	-531.91	–	-547.17	–
AIC	1133.53	–	1046.88	–	1103.83	–	1130.35	–

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 disability related benefits – survey (INDI) versus records

Survey (INDI)	(1) Coeff.	S.E.	(2) Coeff.	S.E.	(3) Coeff.	S.E.	(4) Coeff.	S.E.
SDA	0.001	0.552	-0.036	0.576	-0.167	0.535	-0.022	0.551
IID	0.684	0.442	0.482	0.489	0.416	0.461	0.628	0.448
AA	0.408	0.550	0.346	0.591	0.265	0.546	0.365	0.546
ICA	-0.147	0.630	-0.253	0.625	-0.328	0.619	-0.140	0.615
DLA	-0.476	0.466	-0.502	0.483	-0.601	0.469	-0.484	0.462
age/10	0.339	0.685	0.350	0.765	0.424	0.686	0.330	0.660
age/10 squared	-0.042	0.069	-0.042	0.077	-0.048	0.069	-0.041	0.066
male	0.079	0.364	0.059	0.375	0.026	0.352	0.058	0.354
no qualifications	-0.481	0.404	-0.534	0.420	-0.551	0.400	-0.475	0.392
married/cohabiting	-0.215	0.305	-0.186	0.312	-0.163	0.298	-0.207	0.298
London/SE	0.185	0.578	0.025	0.627	0.096	0.587	0.172	0.562
health problems	0.202	0.680	0.075	0.704	0.213	0.691	0.251	0.699
long-term sick	-0.795	0.427	-0.859*	0.438	-0.862	0.442	-0.788	0.416
t	–	–	–	–	-0.007	0.173	–	–
t^2	–	–	–	–	0.006	0.014	–	–
t^3	–	–	–	–	0.000	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.107	0.157
_cons	-3.703**	1.416	-4.604**	1.649	-3.905*	1.548	-3.501*	1.429
# spell-months	2218	–	2218	–	2218	–	2218	–
# parameters	14	–	33	–	17	–	15	–
log likelihood	-236.53	–	-208.51	–	-232.83	–	-236.27	–
AIC	501.06	–	483.01	–	499.67	–	502.55	–
Records								
ICA	-2.346**	0.820	-3.398***	0.707	-2.884**	1.033	-2.521**	0.838
DTC	0.392	0.572	0.152	0.734	0.402	0.827	0.845	0.754
age/10	3.055***	0.923	4.438***	1.072	3.574***	1.052	3.106**	1.069
age/10 squared	-0.404***	0.105	-0.566***	0.121	-0.464***	0.118	-0.409***	0.118
male	-0.186	0.449	-0.936	0.616	-0.512	0.526	-0.140	0.528
no qualifications	-0.340	0.467	-0.675	0.517	-0.374	0.556	-0.375	0.554
married/cohabiting	-0.782	0.521	-0.988*	0.425	-0.756	0.526	-0.737	0.538
London/SE	-0.509	0.822	-1.032	0.683	-0.512	0.891	-0.258	0.862
health problems	-0.078	0.535	-0.196	0.669	-0.461	0.684	-0.452	0.638
long-term sick	-2.585***	0.736	-3.894***	0.926	-3.110***	0.820	-2.691***	0.796
t	–	–	–	–	0.501**	0.163	–	–
t^2	–	–	–	–	-0.026**	0.009	–	–
t^3	–	–	–	–	0.000***	0.000	–	–
ln(t)	–	–	–	–	–	–	0.434**	0.164
_cons	-7.249***	1.798	-11.397***	2.556	-9.820***	2.324	-8.055***	2.169
# spell-months	1656	–	1656	–	1656	–	1656	–
# parameters	11	–	23	–	14	–	12	–
log likelihood	-101.89	–	-75.45	–	-96.43	–	-100.03	–
AIC	225.77	–	196.91	–	220.87	–	224.05	–

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 income related benefits – survey (DI) versus records

Survey	(1)		(2)		(3)		(4)	
(DI)	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.	Coeff.	S.E.
IS	1.120*	0.548	1.113	0.587	1.142*	0.556	1.117*	0.544
JSA	3.222***	0.604	3.386***	0.691	3.185***	0.605	3.197***	0.609
age/10	1.730*	0.763	1.763*	0.819	1.669*	0.750	1.725*	0.757
age/10 squared	-0.223*	0.095	-0.225*	0.102	-0.216*	0.094	-0.222*	0.094
male	-0.181	0.332	-0.250	0.404	-0.178	0.350	-0.174	0.339
no qualifications	-0.286	0.405	-0.268	0.419	-0.261	0.408	-0.286	0.402
married/cohabiting	0.528	0.325	0.610	0.402	0.580	0.356	0.521	0.336
London/SE	-0.020	0.407	0.081	0.443	0.007	0.405	-0.023	0.408
spouse employed	-0.053	0.380	-0.031	0.433	-0.050	0.389	-0.054	0.377
Number of children	-0.166	0.146	-0.170	0.165	-0.173	0.149	-0.165	0.145
age of youngest	-0.029	0.062	-0.034	0.069	-0.025	0.062	-0.028	0.062
unemployment rate	0.018	0.115	0.064	0.128	0.017	0.117	0.015	0.115
own house	0.096	0.305	0.078	0.350	0.095	0.309	0.099	0.303
(self-)employed	1.152*	0.463	1.310**	0.507	1.208*	0.506	1.144*	0.469
unemployed	0.044	0.790	0.325	0.919	0.109	0.833	0.029	0.791
months unemployed	0.040**	0.013	0.042**	0.014	0.038**	0.013	0.039**	0.013
t	–	–	–	–	0.200	0.149	–	–
t^2	–	–	–	–	-0.023	0.014	–	–
t^3	–	–	–	–	0.001	0.000	–	–
ln(t)	–	–	–	–	–	–	-0.027	0.141
_cons	-7.894***	1.313	-9.007***	1.398	-8.119***	1.325	-7.814***	1.325
# spell-months	1551	–	1550	–	1551	–	1551	–
# parameters	17	–	33	–	20	–	18	–
log likelihood	-197.16	–	-186.10	–	-195.15	–	-197.15	–
AIC	428.33	–	438.20	–	430.29	–	430.30	–
Records								
IS	1.063*	0.450	1.440**	0.529	1.511**	0.541	1.444**	0.555
JSA	1.692***	0.323	1.951***	0.364	2.031***	0.367	2.086***	0.402
age/10	0.196	0.402	0.092	0.492	0.041	0.462	0.020	0.471
age/10 squared	-0.044	0.050	-0.044	0.060	-0.040	0.056	-0.038	0.056
male	-0.177	0.242	-0.177	0.307	-0.198	0.310	-0.264	0.301
no qualifications	-0.158	0.215	0.138	0.291	0.153	0.276	0.065	0.291
married/cohabiting	0.772***	0.230	0.954**	0.313	1.008***	0.303	1.129***	0.317
London/SE	-0.271	0.330	-0.296	0.427	-0.283	0.413	-0.289	0.413
spouse employed	-0.157	0.275	-0.145	0.347	-0.182	0.334	-0.233	0.344
number of children	-0.074	0.097	-0.057	0.126	-0.069	0.121	-0.058	0.123
age of youngest	0.082**	0.030	0.091*	0.037	0.092*	0.038	0.099**	0.038
unemployment rate	-0.036	0.057	-0.009	0.071	-0.023	0.072	-0.005	0.071
own house	0.307	0.255	0.490	0.323	0.468	0.316	0.452	0.327
(self-)employed	0.898*	0.350	1.044**	0.383	1.099**	0.381	0.938*	0.372
unemployed	-0.110	0.428	0.062	0.496	0.038	0.476	0.003	0.488
months unemployed	0.022*	0.011	0.015	0.014	0.015	0.014	0.014	0.014
t	–	–	–	–	0.882***	0.154	–	–
t^2	–	–	–	–	-0.075***	0.016	–	–
t^3	–	–	–	–	0.002***	0.000	–	–
ln(t)	–	–	–	–	–	–	0.756***	0.137
_cons	-4.695***	0.816	-20.394***	1.023	-7.417***	1.078	-6.085***	0.978
# spell-months	1175	–	1175	–	1175	–	1175	–
# parameters	17	–	34	–	20	–	18	–
log likelihood	-239.19	–	-203.83	–	-224.49	–	-228.72	–
AIC	512.37	–	475.66	–	488.97	–	493.44	–

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 disability related benefits – survey (DI) versus records

Survey (INDI)	(1) Coeff.	S.E.	(2) Coeff.	S.E.	(3) Coeff.	S.E.	(4) Coeff.	S.E.
SDA	1.370	1.148	1.374	1.167	1.722	1.137	1.321	0.989
IID	3.655	2.689	3.597	3.226	2.667	1.678	2.892	1.904
AA	0.395	1.128	0.180	1.301	0.035	0.872	0.104	0.914
DLA	-1.660	1.358	-1.653	1.382	-1.423	1.131	-1.465	1.155
age/10	1.945	2.053	1.990	2.416	1.450	1.597	1.655	1.759
age/10 squared	-0.197	0.169	-0.199	0.196	-0.145	0.132	-0.166	0.145
male	-1.681	1.790	-1.743	1.995	-1.125	1.250	-1.217	1.356
no qualifications	-0.944	1.629	-0.960	1.867	-0.541	1.079	-0.687	1.163
married/cohabiting	-0.820	0.930	-1.014	1.074	-0.555	0.752	-0.711	0.751
London/SE	1.245	0.930	1.114	0.862	0.626	0.644	0.750	0.699
health problems	0.322	1.590	0.326	1.854	0.007	1.520	-0.031	1.699
long-term sick	-2.494*	1.169	-2.553*	1.179	-2.402*	1.136	-2.300*	0.978
t	–	–	–	–	-2.617**	0.970	–	–
t^2	–	–	–	–	0.318*	0.134	–	–
t^3	–	–	–	–	-0.011*	0.005	–	–
ln(t)	–	–	–	–	–	–	-0.719*	0.330
_cons	-7.237	3.865	-8.095	4.539	-2.236	3.433	-5.270	3.320
# spell-months	1263	–	1263	–	1263	–	1263	–
# parameters	13	–	16	–	16	–	14	–
log likelihood	-53.10	–	-44.80	–	-44.67	–	-50.34	–
AIC	132.20	–	121.60	–	121.33	–	128.67	–
Records								
age/10	2.502	2.363	2.254	2.479	2.393	2.721	2.462	2.689
age/10 squared	-0.278	0.240	-0.253	0.246	-0.266	0.269	-0.273	0.269
male	-0.151	0.657	-0.158	0.711	-0.198	0.726	-0.168	0.738
no qualifications	-0.579	0.826	-0.208	0.956	-0.232	0.888	-0.514	0.919
married/cohabiting	-0.667	0.726	-0.390	0.809	-0.335	0.759	-0.567	0.665
London/SE	0.782	0.911	1.193	0.923	1.056	0.962	1.157	1.005
health problems	-2.447**	0.936	-3.325*	1.346	-3.007*	1.505	-3.239*	1.453
long-term sick	0.195	0.916	-0.082	0.947	-0.140	0.931	0.365	0.987
t	–	–	–	–	0.501	0.650	–	–
t^2	–	–	–	–	-0.016	0.078	–	–
t^3	–	–	–	–	0.000	0.003	–	–
ln(t)	–	–	–	–	–	–	0.679	0.558
_cons	-6.712	5.432	-7.275	6.327	-8.344	7.182	-7.516	6.756
# spell-months	657	–	657	–	657	–	657	–
# parameters	9	–	15	–	12	–	10	–
log likelihood	-42.65	–	-35.39	–	-40.30	–	-41.62	–
AIC	103.30	–	100.78	–	104.60	–	103.24	–

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 4: Predicted hazard rates for income related benefits

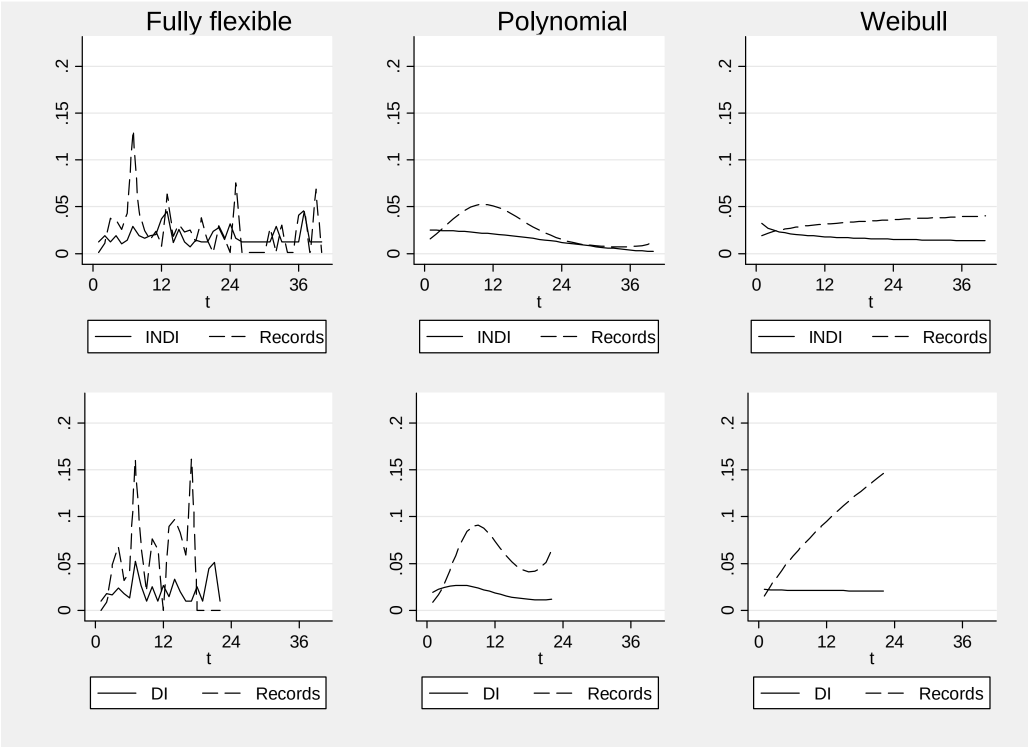


Figure 5: Predicted hazard rates for disability related benefits