Constructing data on unemployment spells from the PSID and the BHPS

Richard Upward

June 1999

Abstract

In Haynes, Upward & Wright (1999) we analyse unemployment durations for a sample of heads of households from the US, from the Panel Study of Income Dynamics, and from the UK, from the British Household Panel Survey. The construction of the data is complex, and is described in detail in this document.

1 Original data

The US data comes from waves 21 to 26 of the Panel Study of Income Dynamics (PSID), covering the period 1988 to 1993. This data is described in detail in Hill (1992), and on the PSID web site (Institute for Social Research, University of Michigan 1999). The UK data comes from waves 1 to 6 of the British Household Panel Survey (BHPS), covering the period 1991 to 1996, described in Taylor, Brice, Buck & Prentice (1998) and on the BHPS web site (ESRC Research Centre on Micro-Social Change 1998). Some familiarity with these data sets is assumed. The time-span of the data may increase as more waves become available, provided that the appropriate information is available for these years.

2 Structure of the unemployment data

The data is constructed so that there is a single observation for each *calendar month* within the sample period for each sample member. Thus durations and transitions were recorded to the nearest month. The data were constructed in this way for two reasons:

1. The PSID data on events between interviews is collected as a sequence of months, and therefore precise dates are not available.

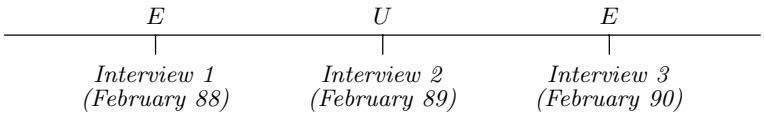
2. Although data on precise dates are available from the BHPS, it was felt that the use of these dates would suggest a spurious accuracy of recall. It is perhaps unrealistic to expect recall of events over a 12 month period to be accurate to the nearest day. In addition, in many cases the precise date of events is not recorded, while the month is recorded.

3 Overview of the data construction process

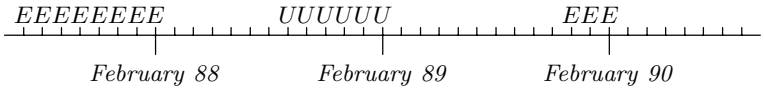
To illustrate the basic process, suppose that we wish to only collect information on employment status over a two year period. Employment status is either E (Employed) or U (Unemployed).

3.1 PSID

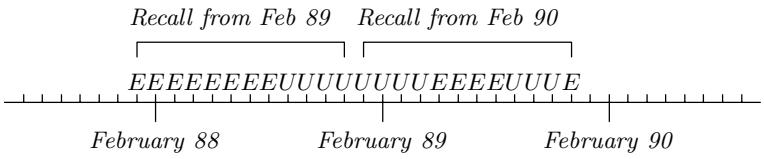
Information on the current employment status is available in the month in which the interview takes place, as illustrated below:



However, because information is also collected about the start date of the current spell, we can fill in information about each month between the interview date and the start of that spell. Note that in the first interview the spell start date is before the sample period.



At each interview the individual is also asked about employment status in each month in the *previous calendar year*, as illustrated below.

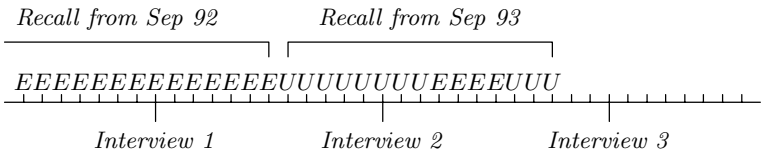


By comparing these two sets of information it is possible to check whether the recall information is consistent with the contemporaneous information, and to construct a sequence of monthly data. Note that the final month in the sample period will be the month of the final interview (in this example February 1990), and the employment status in this month will be censored. This implies that it is not possible to collect information for all months in the final year of the survey (in the case of the PSID, we will lose information relating to most of 1993).

3.2 BHPS

It is more complicated to get BHPS data into a sequence of months, because the data is collected as a sequence of spells. As with the PSID, data on the current spell (at the time of the interview) includes the date when that spell started, and this can be used to construct a sequence of months going back to the spell start date, as illustrated below.

If a spell started after the previous wave's interview date (as is the case here), the BHPS also contains information on all previous spells back to the first spell which starts *before* the previous wave. In this example:



In this example the spell before the current spell in September 1992 is the spell in progress in September 1991, which starts before the sample period. There are

three recall spells in September 1993, the earliest of which is the spell in progress in September 1992.

As with the PSID, by comparing the contemporaneous information with the recall information we can check for consistency. In this example the two are consistent: the earliest recall spell from interview 3 starts in the same month as the current spell in interview 2.

Of course, in the real data there will tend to be inconsistencies because no cross-wave checking is done at the time of interview.

4 Detailed description of data construction

4.1 PSID

4.1.1 Matching waves

The process of matching waves is described in detail in the PSID documentation (Hill 1992). A single file containing all the required family-level variables for the years 1988–1993 was created, with one observation for each family-year (TN observations in total). Each family is identified by a unique family id number. Another file containing all the required individual-level variables was then matched to this using family id and year as identifiers.

4.1.2 Selection of the sample

Total number of individuals in any year 1988–1993	37,218
Heads of household	12,806
In every year 1988–1993 (balanced panel)	5,371
Under 65 and not retired in 1988	4,526
Employment status known for every month in sample	4,460
Valid covariates for every month in the sample	3,904

4.1.3 Creating monthly employment status data

1. Create a 'long' data set with one row for each interview, i.e. a simple yearly panel.

2. Calculate the start date and status of the current spell.

3. Create a series of variables recording what each individual is doing in each month in the previous year. For example, the variables `jan1` to `dec1` record whether the individual was working for their main employer in each month in the previous year.

4. Lag these values by one year so that the answers for 1989 match the 1988 interview. This means there will be no information for individual months in 1993.

5. Expand the data so there is one row for each month.

6. Create a new status variable which records the monthly answer.

7. Calculate for each month what the original `status` variable implies.

8. Compare the two estimates of `status` and use the most plausible, using a series of rules defined below.

4.1.4 Problems

Most of the problems related to the fact that the current employment status may not be consistent with the recall employment status. Current employment status (`status`) can take the following values:

1. Working
2. Temporarily laid-off
3. Unemployed
4. Retired
5. Disabled
6. Keeping House
7. Student
8. Other

Recall employment status, however, has only four values:

1. Working for current employer
2. Working for different employer
3. Unemployed
4. Out of the labour force

The following table shows how consistent the two measures are. In total the data consists of 271,560 months of information for 4,526 individuals. There are five years of information (January 1988 to December 1992; recall that 1993 cannot be used because we have no recall information from 1994), 12 months in each year, which gives a sample of 4,526 individuals ($4,526 \times 12 \times 5 = 271,560$).

			Recall status		
	NA	Working	Working	Unemployed	OLF
Current status		(Same emp)	(Different emp)		
Working	6477	186922	7722	1422	672
Temp. laid off	427	1447	68	327	193
Unemployed	1325	1033	472	3632	2849
Retired	933	1270	146	174	6933
Disabled	267	174	4	209	7701
Keeping House	318	261	37	634	5719
Student	190	447	94	207	1242
Other	80	28	25	66	584
Not available	7893	964	10966	3925	5081

Of the months where information about both contemporaneous and recall status is available, only about 6.5% are definitely inconsistent. A further 2% are uncertain, because it is conceivable that an individual reports being temporarily laid off, and subsequently reports being in work, unemployed or out of the labour force the next year. Given that these two responses are made one year apart, this level of consistency provides some justification for assuming that we can construct retrospective data using this method.

Note, however, that it is not possible to construct a complete series of employment states for all months. For example, it is only possible to code someone as 'Temporarily Laid Off' if that was their current state at the time of the interview. Information about temporary lay-offs will therefore be incomplete.

The following rules are used to code a 'best guess' of employment status in each month. In the following `status1` is the status recorded at the time of the interview and `status2` is recorded retrospectively.

1. If `status1` and `status2` agree, then there is no problem.
2. If `status1` is TLO, and `status2` is U, then assume TLO.
3. If `status1` is U, and `status2` is OLF, assume U.
4. If `status2` is not known, use `status1`.

5. If `status1` is not known, use `status2`.

6. If current month is the month of the interview, assume `status1` is correct.

7. If current month is not month of the interview, assume `status2` is correct.

4.1.5 Matching in other data

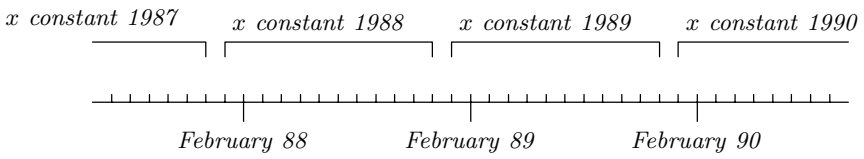
Once the data containing monthly employment status has been created, other information about the individuals is merged in. There are three types of covariates: time invariant (e.g. gender), time-varying (e.g. age) and spell-varying (e.g. occupation). Covariates which vary by spell are assumed to be constant within a spell, whereas covariates which vary by time may vary within spells. Table 1 shows the sources of all covariates used in the analysis.

	1988	1989	1990	1991	1992	1993
Time invariant						
Gender	15131	16632	18050	19350	20652	22407
Ethnic origin	16086	17483	18814	20114	21420	23276
Whether graduated from high school	16089	17486	18817	20117	21423	23279
Whether graduated from college	16104	17501	18832	20132	21438	23294
Time-varying						
Age	15130	16631	18049	19349	20651	22406
State	14803	16303	17703	19003	20303	21603
County unemployment rate	16185	17563	18915	20215	21521	23335
Own or rent accommodation	15140	16641	18072	19372	20672	22427
Mortgage	15141	16642	18073	19373	20673	22428
Number of children	15133	16634	18052	19352	20654	22409
Age of youngest child	15134	16635	18053	19353	20655	22410
Marital status	16187	17565	18916	20216	21522	23336
Years worked since 18	16126	17523	18854	20154	21460	23316
Spell-varying						
Why left last job	15328	16843	18267	19567	20867	22655
Receiving unemployment benefit	15010–	16510–	17926–	19226–	20526–	22096–
(Recorded monthly)	15021	16521	17937	19237	20537	22107
Income from unemployment benefit	14956	16456	17872	19172	20472	22108
Occupation	15162	16663	18101	19401	20701	22456
Industry	15163	16664	18102	19402	20702	22457
Whether self-employed	15157	16658	18096	19396	20696	22451
Wage	15165	16666	18104	19404	20704	22464
Tenure with present employer	15181	16682	18120	19420	20720	22489
Union coverage	15160	16661	18099	19399	20699	22454
Union membership	15161	16662	18100	19400	20700	22455

Table 1: Variables used from PSID

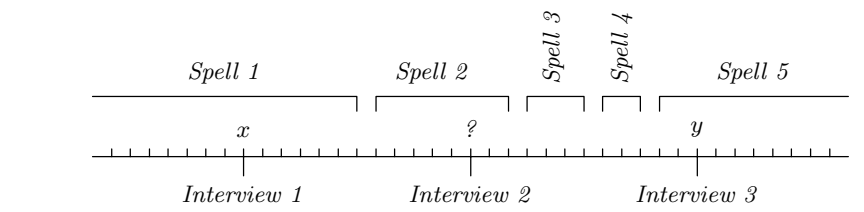
Time invariant covariates are straightforward to match in using each individual's

identifying code. Time-varying covariates are only available for the month in which the interview takes place, so it is necessary to assume that their value is constant for each calendar year. Obviously for age this is an accurate assumption, but may be less accurate for region, for example. If an individual moves in March 1989 in the example below, the coded value for region will be inaccurate until February 1990.

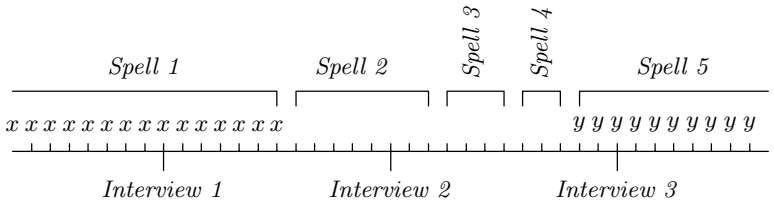


Matching in spell-varying covariates is more difficult. A few of these variables are collected from the retrospective monthly questions, and these are therefore already available for each month. For example, receipt of unemployment benefit is recorded for each month in the year before the interview, and therefore corresponds to the monthly data exactly. For spell-varying covariates which are only available in the month of the interview, however, it is necessary to assume that the value of the covariate at the time of the interview applies to all months in that spell. For example, if an individual is in occupation n in February 1990, it is assumed that they remain in that occupation for the duration of that employment spell. For some variables, such as the wage, this may be a strong assumption.

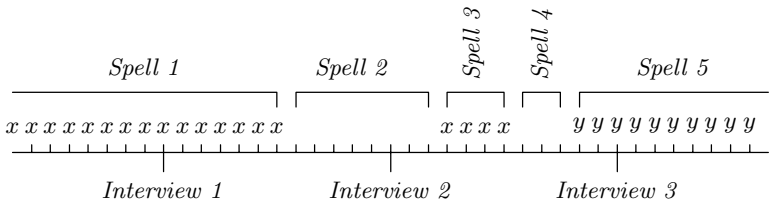
The following diagrams illustrate the method. Suppose there are five spells, the second and fourth of which are unemployment. A spell-varying covariate is recorded at interviews 1 and 3, with values x and y. The covariate is missing in interview 2 because the individual did not have a job.



First, the value of the covariate is copied to every month in that spell:



Although spell 3 is a spell of employment, we have no information about the value of the covariate. We are therefore forced to assume that it is the same as in the previous spell of employment, which is a strong assumption.



In this example we now have spell-varying information copied across to every month in the sample for which the individual is employed. Note, however, that if there is an additional spell *before* the first interview date, it will be impossible to match any spell-varying information to that spell, because we cannot use an earlier spell to copy forward information.

4.1.6 Creating a sample of unemployment spells

Once a complete data set of spells and covariates has been created, a subset of unemployment spells is created. The sample consists of spells of unemployment or 'Out of the Labour Force' (OLF) which (a) have a preceding spell of employment; (b) which do not end in retirement and (c) which have valid industry information from the preceding spell. This process is straightforward, and results in the following sample:

Total number of individuals with complete record	3,904
Individuals :	
with a spell of unemployment or OLF	1,506
whose spells finish before retirement age	1,350
whose spells start after the start of the sample period	1,182
whose spells are preceded by a spell of employment	1,055

The remaining sample of 1,055 individuals experience 1,839 spells of unemployment, distributed as follows:

No. of spells	*Frequency*	*Percentage*
1	660	56.87
2	244	23.13
3	126	11.94
4	58	5.50
5	23	2.18
6	3	0.28
8	1	0.09
Total	1055	100.00

4.2 BHPS

4.2.1 Matching waves

The process of matching waves in the BHPS is straightforward, because each individual is identified by a cross-wave identifier (Taylor *et al.* 1998).

4.2.2 Selection of the sample

Total number of individuals in any year 1991–1996	13,473
Heads of household	7,410
In every year 1988–1993 (balanced panel)	3,184
Under 65 and not retired in 1991	2,615
Employment status known for every month in sample	2,527
Valid covariates for every month in the sample	2,440

4.2.3 Creating monthly employment status data

As with the PSID, the data set should consist of every labour market spell for each individual for each month over the sample period. The time period covered should be as long as possible: at present this covers waves 1-6 of the BHPS, or approximately 1/9/90 to the end of 1995. To construct this data we need to use the job-history records (`jobhist`) from each wave, as well as the main individual panel responses (`indresp`). In addition, any variable which is not on the individual response record (such as age of youngest child) has to be retrieved from the household response file (`hhresp`) and matched to the individual record.

An individual's most recent spell starts at the date given in the latest `indresp` record type. Similarly, there will be a record of when the current spell started in `aindresp`, `bindresp` and so on. These are calculated from interview date and the derived length of current labour market spell.

The time that the current spell started determines whether there are records on the `jobhist` record types. For example, if the current spell in wave 3 starts between 01sep92 and 01sep93, there will be a record on `cjobhist`. These are crucial for determining which variables make up the eventual list of spells.

Initially, the data are all stacked on top of each other in the following order:

1	ajobhist
2	aindresp
3	bjobhist
4	bindresp
5	cjobhist
6	cindresp
7	djobhist
8	dindresp
9	ejobhist
10	eindresp
11	fjobhist
12	findresp

This order should result in approximately chronological spells. Within `jobhist` records, spells are sorted by start date. There may be any number of records (including zero) from each of the `jobhist` data sets, although individuals should all have a record on the `indresp` record unless they are not in the sample. In order to create a id for records which are not from `indresp` record types, it is first necessary to merge each `jobhist` record type with `indresp` record types.

The pattern of spell dates is extremely complex. The following variables were used to create and sequence of spells:

1. `pid`: unique identifier for each individual.
2. `record`: the source of the data.
3. `spno`: the spell number if the data came from the `jobhist` record type.
4. `doi`: the date of the interview.
5. `pdoi`: the date of the previous year's interview.
6. `start`: the start date of the spell, as calculated from interview date and spell tenure.
7. `status`: labour market status.
8. `hstatus`: labour market status if data came from `jobhist` record type.
9. `stpy`: reason for leaving spell, if data came from `jobhist` record type.

The following rules were used to construct the sequence of spells.

1. If spell starts before `pdoi`, and `status` is the same, and both spells are from the -indresp- records, drop the later occurrence of the spell.

2. If spell starts before `pdoi`, but `status` has changed, create a new spell. Assume that the start of that spell is in the month following `pdoi`.

3. If a spell from the `jobhist` record starts before `pdoi`, it should be the same spell as recorded in last year's `indresp` record. If `status` of this spell is the same as `status` in previous wave, copy across the value of `hstatus` and `stpy`, and then drop the information from `jobhist`. The values of -hstatus- and -stpy- should be copied across to the first occurrence of that spell.

4. If two or more records come from the same wave, then they should not be merged together.

Some examples are given below:

```
------------------------------------------------------------------------------
pid     record   spno        doi     start      status     hstatus        stpy
------------------------------------------------------------------------------
  1   aindresp      .  09sep1991     Jan61         OLF           .           .

No action needed, a single spell.

  1   aindresp      .  09sep1991     Jan61         OLF           .           .
------------------------------------------------------------------------------
  2   ajobhist      1  10sep1991     Sep68         OLF         OLF          -8
  2   aindresp      .  10sep1991     Jan91  Unemployed           .           .
  2   bindresp      .  14sep1992     Feb91  Unemployed           .           .

Spell 3 starts before -pdoi-, and is therefore assumed to be the same
spell because -status- also remains unchanged. (RULE 1).

  2   ajobhist      1  10sep1991     Sep68         OLF         OLF          -8
  2   aindresp      .  10sep1991     Jan91  Unemployed           .           .
------------------------------------------------------------------------------
  3   ajobhist      1  10sep1991     Sep70         OLF         OLF          -8
  3   aindresp      .  10sep1991     Jan91  Unemployed           .           .
  3   bindresp      .  14sep1992     Feb91  Unemployed           .           .
  3   cindresp      .  13nov1993     Dec91  Unemployed           .           .

Spells 3 and 4 both start before -pdoi- and are therefore assumed to
be the same spell. (RULE 1).

  3   ajobhist      1  10sep1991     Sep70         OLF         OLF          -8
  3   aindresp      .  10sep1991     Jan91  Unemployed           .           .
```

```
-------------------------------------------------------------------------------
  4   aindresp      .  14sep1991     Apr90    Employed           .           .
  4   bindresp      .  14oct1992     Apr90    Employed           .           .
  4   cindresp      .  30sep1993     Apr90    Employed           .           .

All spells are the same spell. (RULE 1).

  4   aindresp      .  14sep1991     Apr90    Employed           .           .
-------------------------------------------------------------------------------
  5   aindresp      .  16sep1991     Jun81    Employed           .           .
  5   bindresp      .  21sep1992     Sep90         OLF           .           .
  5   cindresp      .  11oct1993     Oct59         OLF           .           .
  5   eindresp      .  09nov1995     Oct59         OLF           .           .
  5   findresp      .  23oct1996     Jul89         OLF           .           .

Although the second spell has a start date before -pdoi-, -status- has
changed. The assumption is that current status is always accurate, and
so we assume that this individual 'forgot' they were employed in the
previous year. When did the spell of OLF start?  We have to assume it
started as close as possible to the given date which is still after
-pdoi-, which is Oct91. (RULE 2).

  5   aindresp      .  16sep1991     Jun81    Employed           .           .
  5   bindresp      .  21sep1992     Oct91         OLF           .           .
-------------------------------------------------------------------------------
  6   aindresp      .  23sep1991     Jan85    Employed           .           .
  6   bjobhist      1  21sep1992     Jan85    Employed   Same emp.    Promoted
  6   bindresp      .  21sep1992     Sep92    Employed           .           .
  6   cindresp      .  11oct1993     Dec84    Employed           .           .
  6   eindresp      .  09nov1995     Dec84    Employed           .           .
  6   findresp      .  06nov1996     Dec84    Employed           .           .

The second spell here is clearly the same as the first. Information
from -bjobhist- needs to be copied across to augment the first
spell. The remaining spells are dropped: it looks like they
subsequently ignored the promotion. (RULE 3, RULE 1)

  6   aindresp      .  23sep1991     Jan85    Employed   Same emp.    Promoted
  6   bindresp      .  21sep1992     Sep92    Employed           .           .
-------------------------------------------------------------------------------
  7   aindresp      .  23sep1991     Jun83    Employed           .           .
  7   bindresp      .  23sep1992     Mar82    Employed           .           .
  7   cindresp      .  24jan1993     Mar82    Employed           .           .
  7   djobhist      2  22sep1994     Sep92    Employed  Diff. emp.    Lost job
```

```
  7   djobhist      1  22sep1994     Apr94         OLF         OLF          -8
  7   dindresp      .  22sep1994     Jun94    Employed           .           .
```

The second and third spells are dropped. The record from -djobhist- starts before -pdoi- is must therefore be assumed to be the same spell. We therefore need to copy across the -hstatus- and -stpy- to the earliest recording of this spell, the first row. The next 2 spells start after the previous wave interview, and are therefore kept as a new spell (RULE 1, RULE 3).

```
  7   aindresp      .  23sep1991     Jun83    Employed  Diff. emp.    Lost job
  7   djobhist      1  22sep1994     Apr94         OLF         OLF          -8
  7   dindresp      .  22sep1994     Jun94    Employed           .           .
-------------------------------------------------------------------------------
  8   aindresp      .  13sep1991     Jan79    Employed           .           .
  8   bindresp      .  10sep1992     Jul87    Employed           .           .
  8   cindresp      .  30sep1993     Jul69    Employed           .           .
  8   dindresp      .  17sep1994     Jul68    Employed           .           .
  8   eindresp      .  20sep1995     Jul73    Employed           .           .
  8   fjobhist      1  25oct1996     Jan74    Employed  Diff. emp.    Personal
  8   findresp      .  25oct1996     Oct96    Employed           .           .
```

First six spells are all the same (RULE 1). Information from the -jobhist- record is copied across to the first row (RULE 3).

```
  8   aindresp      .  13sep1991     Jan79    Employed  Diff. emp.    Personal
  8   findresp      .  25oct1996     Oct96    Employed           .           .
-------------------------------------------------------------------------------
  9   aindresp      .  14sep1991     Sep87    Employed           .           .
  9   bindresp      .  12sep1992     Jan90    Employed           .           .
  9   cindresp      .  07oct1993     Nov89    Employed           .           .
  9   dindresp      .  20sep1994     Oct77    Employed           .           .
  9   eindresp      .  20sep1995     Oct77    Employed           .           .
```

Single spell (RULE 1).

```
  9   aindresp      .  14sep1991     Sep87    Employed           .           .
-------------------------------------------------------------------------------
 10   eindresp      .  20sep1995     Sep83         OLF           .           .
 10   findresp      .  23oct1996     Jul84         OLF           .           .
```

Single spell (RULE 1).

```
 10   eindresp      .  20sep1995     Sep83         OLF           .           .
```

```
-------------------------------------------------------------------------------
 11   aindresp      .  05oct1991     Mar90         OLF           .           .
 11   bindresp      .  09sep1992     Dec89         OLF           .           .
 11   cindresp      .  04oct1993     Dec89         OLF           .           .
 11   dindresp      .  22sep1994     Mar90         OLF           .           .
 11   eindresp      .  25sep1995     Mar90         OLF           .           .
 11   findresp      .  06jan1997     Mar90         OLF           .           .

Single spell (RULE 1).

 11   aindresp      .  05oct1991     Mar90         OLF           .           .
-------------------------------------------------------------------------------
 12   aindresp      .  12sep1991     Jun71    Employed           .           .
 12   bindresp      .  01oct1992     Jul70    Employed           .           .
 12   cindresp      .  08oct1993     Jul72    Employed           .           .

Single spell (RULE 1).

 12   aindresp      .  12sep1991     Jun71    Employed           .           .
-------------------------------------------------------------------------------
 13   aindresp      .  13sep1991     Jun88    Employed           .           .
 13   bindresp      .  15sep1992     May89    Employed           .           .
 13   cindresp      .  08oct1993     Jul89    Employed           .           .
 13   dindresp      .  23jan1995     Jul89    Employed           .           .

Single spell (RULE 1).

 13   aindresp      .  13sep1991     Jun88    Employed           .           .
-------------------------------------------------------------------------------
 14   aindresp      .  25sep1991     Oct88    Employed           .           .
 14   bindresp      .  14sep1992     Dec90    Employed           .           .
 14   cindresp      .  09nov1993     Oct88    Employed           .           .
 14   eindresp      .  25apr1996     Oct88    Employed           .           .
 14   findresp      .  02dec1996     Oct88    Employed           .           .

Single spell (RULE 1).

 14   aindresp      .  25sep1991     Oct88    Employed           .           .
-------------------------------------------------------------------------------
 15   aindresp      .  25sep1991     Dec88    Employed           .           .
 15   bindresp      .  16sep1992     Oct88    Employed           .           .
 15   cindresp      .  10oct1993     Oct88    Employed           .           .

Single spell (RULE 1).
```

```
 15   aindresp      .  25sep1991     Dec88    Employed           .           .
-------------------------------------------------------------------------------
 16   aindresp      .  25sep1991     Aug90    Employed           .           .
 16   bindresp      .  14oct1992     Jul90    Employed           .           .
 16   cindresp      .  13dec1993     Jul90    Employed           .           .
 16   dindresp      .  17nov1994     Apr93    Employed           .           .
 16   eindresp      .  12jan1996     Jan90    Employed           .           .

Single spell (RULE 1).

 16   aindresp      .  25sep1991     Aug90    Employed           .           .
-------------------------------------------------------------------------------
 17   aindresp      .  16sep1991     Jun89    Employed           .           .
 17   bjobhist      1  10sep1992     Oct90    Employed   Same emp.    Promoted
 17   bindresp      .  10sep1992     Oct91    Employed           .           .
 17   cjobhist      1  10sep1993     Aug89    Employed   Same emp.    Promoted
 17   cindresp      .  10sep1993     Oct92    Employed           .           .
 17   djobhist      2  19sep1994     Jan93    Employed   Same emp.    Personal
 17   djobhist      1  19sep1994     May94    Employed   Same emp.  Better job
 17   dindresp      .  19sep1994     May94    Employed           .           .
 17   eindresp      .  21sep1995     May94    Employed           .           .
 17   findresp      .  26oct1996     Jan95    Employed           .           .

The second spell, from -jobhist- starts before -pdoi- and is therefore
assumed to be the same spell. (RULE 3). The next spell starts after
-pdoi- and is therefore a new spell. The next spell from -jobhist-
looks like the same spell as spell 1. However, from RULE 3 the
information will be copied across to the previous spell. Note that the
final spell from -dindresp- is not dropped because we assume that it
is a different spell, even though both start dates are the same. (RULE 4).

 17   aindresp      .  16sep1991     Jun89    Employed   Same emp.    Promoted
 17   bindresp      .  10sep1992     Oct91    Employed   Same emp.    Promoted
 17   cindresp      .  10sep1993     Oct92    Employed   Same emp.    Personal
 17   djobhist      1  19sep1994     May94    Employed   Same emp.  Better job
 17   dindresp      .  19sep1994     May94    Employed           .           .
-------------------------------------------------------------------------------
```

Various consistency checks were then carried out. The start date of each spell should be after the previous spell's start date, for example, and no spells should have a negative duration. The end date of each spell was taken to be the start date of the next spell. If there was no following spell, the censoring date was taken to be the the date of the final interview. Once the data were compressed into a sequence of

spells, the data was expanded so that each observation corresponded to a month.

4.2.4 Matching in other data

Matching in other data uses slightly different methods than for the PSID. Table 2 lists the locations of the covariates used from the BHPS.

	Variable	Record
Time invariant		
Gender	sex	indresp
Ethnic origin	race	indresp
Whether graduated from high school	scend, qfedhi	indresp
Whether graduated from college	feend, qfedhi	indresp
Time-varying		
Age	age	indresp
Region	region	indresp
Regional unemployment rate	Matched in from external data	
Own or rent accommodation	tenure	indresp
Mortgage	tenure	indresp
Number of children	nchild	indresp
Age of youngest child	agechy	hhresp
Marital status	mastat	indresp
Years worked since 18	Age - school leaving age	
Spell-varying		
Why left last job	jhstpy	jobhist
Income from unemployment benefit	aficode	income
Occupation	jbsoc	indresp,jobhist
Industry	jbsic	indresp,jobhist
Whether self-employed	jbsemp	indresp,jobhist
Wage	paygu	indresp,jobhist
Tenure with present employer	cjsten	indresp,jobhist
Union coverage	tujbpl	indresp
Union membership	tuin1, tuin2	indresp

Table 2: Variables used from BHPS

Covariates which occur on both the cross-section (`indresp`) and retrospective (`jobhist`) records are already included in all months in the data. Variables which only occur on the cross-section records which do not vary by time are simply copied across to all months for that individual. Those which do vary by time are treated in the same way as for the PSID; they are assumed constant for each calendar year.

Data on unemployment benefit receipt is available in monthly form, and this too is matched to the main data set. Respondents are asked to list all forms of income for each month in the previous 12 months. Selecting those forms of income which refer to unemployment benefit, these are merged directly to the main data set. This provides another test of the consistency of the spells constructed. It should be the

case that months where unemployment benefit is received should ohly match to months where status is unemployment. The results from the merge are encouraging: less than 1% of months where status is not unemployment have a record for receipt of unemployment benefit.

However, only 27% of months where status is unemployment have a record for receipt of unemployment benefit. In terms of spells, 43% of unemployment spells have some positive benefit receipt at some point.

4.2.5 Creating a sample of unemployment spells

The subset of unemployment spells is chosen using the same criteria as for the PSID.

Total number of individuals with complete record	2,466
Individuals :	
with a spell of unemployment or OLF	1,160
whose spells finish before retirement age	802
whose spells start after the start of the sample period	633
whose spells are preceded by a spell of employment	563

The remaining sample of 563 individuals experience 744 spells of unemployment, distributed as follows:

No. of spells	*Frequency*	*Percentage*
1	437	77.62
2	89	15.81
3	25	4.44
4	10	1.78
6	1	0.18
8	1	0.18
Total	744	100.00

4.3 Matching industry codes between PSID and BHPS

In order to construct a comparable measure of whether individuals remain in the same industry or move between industries, it was necessary to construct an equivalent measure of industry for the two data sets. Table 3 contains my 'best guess' of the match between the 1981–1993 US SIC definitions and the 1980 SIC industry codes.

Table 3: US SIC and UK 1980 SIC industry definitions

1981-1993 Description	US SIC 1981-1993	UK SIC 2-digit 1980
Agricultural production (01)	017	1
Agricultural services, except horticultural (07 except 0713 and 073)	018	1
Horticultural services (073)	019	1
Forestry (08)	027	2
Fisheries (09)	028	3
Metal mining (10)	047	21
Coal mining (11, 12)	048	11
Crude petroleum and natural gas extractions (13)	049	13
Nonmetallic mining and quarrying, except fuel (14)	057	23
General building contractors (15)	067	50
General contractors, except building (16)	068	50
Special trade contractors (17)	069	50
Not specified construction	077	50
Logging (241)	107	46
Sawmills, planing mills, and mill work (242, 243)	108	46
Miscellaneous wood products (244, 249)	109	46
Furniture and fixtures (25)	118	46
Glass and glass products (321-323)	119	24
Cement, concrete, gypsum, and plaster products (324, 327)	127	24
Structural clay products (325)	128	24
Pottery and related products (326)	137	24
Miscellaneous nonmetallic mineral and stone products (328, 329)	138	24
Blast furnaces, steel works, rolling and finishing mills (3312, 3313)	139	22
Other primary iron and steel industries (3315-3317, 332, 3391, part 3399)	147	22
Primary aluminum industries (3334, part 334, 3352, 3361, part 3392, part	148	22
Other primary nonferrous industries (3331-3333, 3339, part 334, 3351,	149	22
Cutlery, hand tools, and other hardware (342)	157	31
Fabricated structural metal products (344)	158	32
Screw machine products (345)	159	32
Metal stamping (346)	167	31
Miscellaneous fabricated metal products (341, 343, 347, 348, 349)	168	31
Not specified metal industries	169	31
Engines and turbines (351)	177	32
Farm machinery and equipment (352)	178	32
Construction and material handling machines (353)	179	32
Metalworking machinery (354)	187	32
Office and accounting machines (357 except 3573)	188	33
Electronic computing equipment (3573)	189	33
Machinery, except electrical, not elsewhere classified (355, 356, 358,	197	32
Not specified machinery	198	32
Household appliances (363)	199	31
Radio, T.V., and communication equipment (365, 366)	207	34
Electrical machinery, equipment, and supplies, not elsewhere classified	208	34
Not specified electrical machinery, equipment, and supplies	209	34
Motor vehicles and motor vehicle equipment (371)	219	35
Aircraft and parts (372)	227	36
Ship and boat building and repairing (373)	228	36
Railroad locomotives and equipment (374)	229	36

Table 3: US SIC and UK 1980 SIC industry definitions

1981-1993 Description	US SIC 1981-1993	UK SIC 2-digit 1980
Mobile dwellings and campers (3791)	237	35
Cycles and miscellaneous transportation equipment (375, 3799)	238	36
Scientific and controlling instruments (381, 382)	239	37
Optical and health services supplies (383, 384, 385)	247	37
Photographic equipment and supplies (386)	248	37
Watches, clocks, and clockwork-operated devices (387)	249	37
Not specified professional equipment	257	37
Ordnance (19)	258	32
Miscellaneous manufacturing industries (39)	259	31
Meat products (201)	268	41
Dairy products (202)	269	41
Canning and preserving fruits, vegetables, and seafoods (203)	278	41
Grain-mill products (204, 0713)	279	41
Bakery products (205)	287	41
Confectionery and related products (207)	288	42
Beverage industries (208)	289	42
Miscellaneous food preparation and kindred products (206, 209)	297	42
Not specified food industries	298	42
Tobacco manufactures (21)	299	42
Knitting mills (225)	307	43
Dyeing and finishing textiles, except wool and knit goods (226)	308	43
Floor coverings, except hard surface (227)	309	43
Yarn, thread, and fabric mills (221-224, 228)	317	43
Miscellaneous textile mill products (229)	318	43
Apparel and accessories (231-238)	319	43
Miscellaneous fabricated textile products (239)	327	43
Pulp, paper, and paperboard mills (261-263, 266)	328	47
Miscellaneous paper and pulp products (264)	329	47
Paperboard containers and boxes (265)	337	47
Newspaper publishing and printing (271)	338	47
Printing, publishing, and allied industries, except newspapers (272-279)	339	47
Industrial chemicals (281)	347	25
Plastics, synthetics and resins, except fibers (282, except 2823 and	348	25
Synthetic fibers (2823, 2824)	349	26
Drugs and medicines (283)	357	25
Soaps and cosmetics (284)	358	25
Paints, varnishes, and related products (285)	359	25
Agricultural chemicals (287)	367	25
Miscellaneous chemicals (286, 289)	368	25
Not specified chemicals and allied products	369	25
Petroleum refining (291)	377	14
Miscellaneous petroleum and coal products (295, 299)	378	14
Rubber products (301-303, 306)	379	25
Miscellaneous plastic products (307)	387	25
Tanned, curried, and finished leather (311)	388	44
Footwear, except rubber (313, 314)	389	45
Leather products, except footwear (312, 315-317, 319)	397	44
Not specified manufacturing industries	398	49

Table 3: US SIC and UK 1980 SIC industry definitions

1981-1993 Description	US SIC 1981-1993	UK SIC 2-digit 1980
Railroads and railway express service (40)	407	71
Street railways and bus lines (411, 413-415, 417)	408	72
Taxicab service (412)	409	72
Trucking service (421, 423)	417	72
Warehousing and storage (422)	418	77
Water transportation (44)	419	74
Air transportation (45)	427	75
Pipe lines, except natural gas (46)	428	17
Services incidental to transportation (47)	429	76
Radio broadcasting and television (483)	447	79
Telephone (wire and radio) (481)	448	79
Telegraph and miscellaneous communication services (482, 489)	449	79
Electric light and power (491)	467	16
Electric-gas utilities (493)	468	15/16
Gas and steam supply systems (492, 496)	469	16
Water supply (494)	477	17
Sanitary services (495)	478	92
Other and not specified utilities (497)	479	92
Motor vehicles and equipment (501)	507	61
Drugs, chemicals, and allied products (502)	508	61
Dry goods and apparel (503)	509	61
Food and related products (504)	527	61
Farm products-raw materials (505)	528	61
Electrical goods (506)	529	61
Hardware, plumbing, and heating supplies (507)	537	61
Not specified electrical and hardware products	538	61
Machinery equipment and supplies (508)	539	61
Metals and minerals, not elsewhere classified (5091)	557	61
Petroleum products (5092)	558	61
Scrap and waste materials (5093)	559	62
Alcoholic beverages (5095)	567	61
Paper and its products (5096)	568	61
Lumber and construction materials (5098)	569	61
Wholesalers, not elsewhere classified (5094, 5097, 5099)	587	61
Not specified wholesale trade	588	61
Lumber and building material retailing (521-524)	607	64
Hardware and farm equipment stores (525)	608	64
Department and mail order establishments (531, 532)	609	64
Limited price variety stores (533)	617	64
Vending machine operators (534)	618	64
Direct selling establishments (535)	619	64
Miscellaneous general merchandise stores (539)	627	64
Grocery stores (541)	628	64
Dairy products stores (545)	629	64
Retail bakeries (546)	637	64
Food stores, not elsewhere classified (542-544, 549)	638	64
Motor vehicle dealers (551, 552)	639	65
Tire, battery, and accessory dealers (553)	647	65

Table 3: US SIC and UK 1980 SIC industry definitions

1981-1993 Description	US SIC 1981-1993	UK SIC 2-digit 1980
Gasoline service stations (554)	648	65
Miscellaneous vehicle dealers (559)	649	65
Apparel and accessories stores, except shoe stores (56 except 566)	657	64
Shoe stores (566)	658	64
Furniture and home furnishings stores (571)	667	64
Household appliances, TV, and radio stores (572, 573)	668	64
Eating and drinking places (58)	669	66
Drug stores (591)	677	64
Liquor stores (592)	678	66
Farm and garden supply stores (596)	679	64
Jewelry stores (597)	687	64
Fuel and ice dealers (598)	688	65
Retail florists (5992)	689	64
Miscellaneous retail stores (593-595, 599 except 5992)	697	64
Not specified retail trade	698	64
Banking (60)	707	81
Credit agencies (61)	708	81
Security, commodity brokerage, and investment companies (62, 67)	709	81
Insurance (63, 64)	717	82
Real estate, including real estate- insurance-law offices (65, 66)	718	83
Advertising (731)	727	83
Services to dwellings and other buildings (734)	728	83
Commercial research, development, and testing labs (7391, 7397)	729	94
Employment and temporary help agencies (736, 7398)	737	83
Business management and consulting services (part 7392)	738	83
Computer programming services (part 7392)	739	83
Detective and protective services (7393)	747	83
Business services, not elsewhere classified (732, 733, 735, 7394, 7395,	748	83/84/85
Automobile services, except repair (751, 752, 754)	749	83
Automobile repair and related services (753)	757	67
Electrical repair shops (762, 7694)	758	67
Miscellaneous repair services (763, 764, 769, except 7694)	759	67
Private households (88)	769	99
Hotels and motels (701)	777	66
Lodging places, except hotels and motels (702, 703, 704)	778	66
Laundering, cleaning, and other garment services (721, 727)	779	98
Beauty shops (723)	787	98
Barber shops (724)	788	98
Shoe repair shops (725)	789	67
Dressmaking shops (part 729)	797	64
Miscellaneous personal services (722, 726, part 729)	798	98
Theaters and motion pictures (78, 792)	807	97
Bowling alleys, billiard and pool parlors (793)	808	97
Miscellaneous entertainment and recreation services (791, 794)	809	97
Offices of physicians (801, 803)	828	95
Offices of dentists (802)	829	95
Offices of chiropractors (804)	837	95
Hospitals (806)	838	95

Table 3: US SIC and UK 1980 SIC industry definitions

1981-1993 Description	US SIC 1981-1993	UK SIC 2-digit 1980
Convalescent institutions (8092)	839	95
Offices of health practitioners, not elsewhere classified (part 8099)	847	95
Health services, not elsewhere classified (807, part 8099)	848	95
Legal services (81)	849	83
Elementary and secondary schools (821)	857	93
Colleges and universities (822)	858	93
Libraries (823)	859	97
Educational services, not elsewhere classified (824, 829)	867	93
Not specified educational services	868	93
Museums, art galleries, and zoos (84)	869	97
Religious organizations (866)	877	96
Welfare services (part 867)	878	96
Residential welfare facilities (part 867)	879	95
Nonprofit membership organizations (861-865, 869)	887	96
Engineering and architectural services (891)	888	83
Accounting, auditing, and bookkeeping services (893)	889	83
Miscellaneous professional and related services (892, 899)	897	83
Postal service (part 9190)	907	79
Federal public administration (part 9190, 9490)	917	91
State public administration (9290)	927	91
Local public administration (9390)	937	91

References

ESRC Research Centre on Micro-Social Change (1998), "British Household Panel Survey", Homepage at `http://www.irc.essex.ac.uk/bhps/index.html`.

Haynes, M., Upward, R. & Wright, P. (1999), "The relative costs of intra- and inter-sectoral mobility: a comparative study", Paper prepared for Centre for Research on Globalisation and Labour Markets conference.

Hill, M. (1992), *The Panel Study of Income Dynamics: a User's Guide*, Newbury Park, CA: Sage Publications.

Institute for Social Research, University of Michigan (1999), "Panel Study of Income Dynamics", Homepage at `http://www.isr.umich.edu/src/psid/index.html`.

Taylor, M., Brice, J., Buck, N. & Prentice, E. (1998), *British Household Panel Survey User Manual Volume A: Introduction, Technical Report and Appendices.*, Colchester: University of Essex.