How To Do
Piecewise Exponential Survival Analysis in Stata 7
(Allison 1995:Output 4.20)

revised 4-25-02

```
.
. *
. * This document can function as a "how to" for setting up data for
. * piecewise exponentional regression.  For that reason, I have
. * separated it from the other analyses for Chapter 4 of Allison
. * (1995).
. *
.
.
. *
. *-----------------------------------------------------------------
. * Output 4.20 Results for Piecewise Exponential Model Applied to
. * Recidivism Data.
. *
. * IMPORTANT! In this section I show how to use Stata's -stsplit-
. * command to generate a new data set suitable for estimation of the
. * piecewise exponential model.
. *
. * I will start with a clean slate.
. *
.
. clear

. set more off

. set mem 10m

Current memory allocation

                    current                                 memory usage
    settable          value     description                 (1M = 1024k)
    --------------------------------------------------------------------
    set maxvar         5000     max. variables allowed           1.733M
    set memory           10M    max. data space                 10.000M
    set matsize         400     max. RHS vars in models          1.254M
                                                            -----------
                                                                12.987M


.
. use c:\courses\soc213b\data\recid.dta
(RECID: Arrest Times for Released Prisoners (Allison, 1995))

.
. *
. * To estimate the piecewise exponential model we need to create a
. * particular "stacked" data set.  This is a commonly used, and
. * powerful, trick in several kinds of analysis.  Stata provides us
. * with a way to create a stacked data set that is suited to
. * estimation of a piecewise exponential model.
```

```
. *
. * To begin with, we need to create an ID variable.  The reason for
. * this is that we are going to create duplicate cases.  In order
. * to keep track of the duplicates, we need an ID variable.  Here
. * is a simple way to create an ID.  First sort on week and arrest
. * We don't need to do this, but I thought that since I was
. * going to the trouble of creating an ID variable, I might as well
. * create one with known meaning.
. *
.
. sort week arrest

.
. gen ID = _n

.
. *
. * Now -stset- the data, this time incorporating the ID variable,
. * even though doing so has absolutely no implications for
. * the use of -streg- at this point.  The "ID" piece here is
. * crucial for subsequent reformatting of the data.
. *
.
. stset week, id(ID) failure(arrest)

                id:  ID
     failure event:  arrest ~= 0 & arrest ~= .
obs. time interval:  (week[_n-1], week]
 exit on or before:  failure

------------------------------------------------------------------------------
      432  total obs.
        0  exclusions
------------------------------------------------------------------------------
      432  obs. remaining, representing
      432  subjects
      114  failures in single failure-per-subject data
    19809  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        52

.
. *
. * To display what stset has done, I next list certain variables
. * for the first 20 and last 20 cases.  Notice the new variables,
. * _st, _d, _t, and _t0.  The -stset- command "quietly" (that is,
. * without telling the user) created these variables.  They are
. * defined deep into the -stset- entry in the Stata 7 manual, on
. * p. 435, in the midst of example 7(!):
. *
. * _st  1 if the record is to be used, 0 if ignored
. * _t0  analysis time when record begins
. * _t   analysis time when record ends
. * _d   1 if failure, 0 if censored
. *
.
. list ID _st _d _t _t0 week arrest if ID < 21 | ID > 412
```

```
            ID        _st         _d           _t          _t0       week    arrest
  1.         1          1          1            1            0          1  arrested
  2.         2          1          1            2            0          2  arrested
  3.         3          1          1            3            0          3  arrested
  4.         4          1          1            4            0          4  arrested
  5.         5          1          1            5            0          5  arrested
  6.         6          1          1            6            0          6  arrested
  7.         7          1          1            7            0          7  arrested
  8.         8          1          1            8            0          8  arrested
  9.         9          1          1            8            0          8  arrested
 10.        10          1          1            8            0          8  arrested
 11.        11          1          1            8            0          8  arrested
 12.        12          1          1            8            0          8  arrested
 13.        13          1          1            9            0          9  arrested
 14.        14          1          1            9            0          9  arrested
 15.        15          1          1           10            0         10  arrested
 16.        16          1          1           11            0         11  arrested
 17.        17          1          1           11            0         11  arrested
 18.        18          1          1           12            0         12  arrested
 19.        19          1          1           12            0         12  arrested
 20.        20          1          1           13            0         13  arrested
413.       413          1          0           52            0         52     other
414.       414          1          0           52            0         52     other
415.       415          1          0           52            0         52     other
416.       416          1          0           52            0         52     other
417.       417          1          0           52            0         52     other
418.       418          1          0           52            0         52     other
419.       419          1          0           52            0         52     other
420.       420          1          0           52            0         52     other
421.       421          1          0           52            0         52     other
422.       422          1          0           52            0         52     other
423.       423          1          0           52            0         52     other
424.       424          1          0           52            0         52     other
425.       425          1          0           52            0         52     other
426.       426          1          0           52            0         52     other
427.       427          1          0           52            0         52     other
428.       428          1          0           52            0         52     other
429.       429          1          1           52            0         52  arrested
430.       430          1          1           52            0         52  arrested
431.       431          1          1           52            0         52  arrested
432.       432          1          1           52            0         52  arrested

.
. *
.
. tab _st, missing

        _st |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        432      100.00      100.00
------------+-----------------------------------
      Total |        432      100.00

.
. *
.
```

```
. tab _d arrest, missing
```

_d	Arrested after release vs. not: other	arrested	Total
0	318	0	318
1	0	114	114
Total	318	114	432

```
.
. *
.
. tab _t, missing
```

_t	Freq.	Percent	Cum.
1	1	0.23	0.23
2	1	0.23	0.46
3	1	0.23	0.69
4	1	0.23	0.93
5	1	0.23	1.16
6	1	0.23	1.39
7	1	0.23	1.62
8	5	1.16	2.78
9	2	0.46	3.24
11	2	0.46	3.94
12	2	0.46	4.40
13	1	0.23	4.63
14	3	0.69	5.32
15	2	0.46	5.79
16	2	0.46	6.25
17	3	0.69	6.94
18	3	0.69	7.64
19	2	0.46	8.10
20	5	1.16	9.26
21	2	0.46	9.72
22	1	0.23	9.95
23	1	0.23	10.19
24	4	0.93	11.11
25	3	0.69	11.81
26	3	0.69	12.50
27	2	0.46	12.96
28	2	0.46	13.43
30	2	0.46	13.89
31	1	0.23	14.12
32	2	0.46	14.58
33	2	0.46	15.05
34	2	0.46	15.51
35	4	0.93	16.44
36	3	0.69	17.13
37	4	0.93	18.06
38	1	0.23	18.29
39	2	0.46	18.75

```
         40 |          4        0.93       19.68
         42 |          2        0.46       20.14
         43 |          4        0.93       21.06
         44 |          2        0.46       21.53
         45 |          2        0.46       21.99
         46 |          4        0.93       22.92
         47 |          1        0.23       23.15
         48 |          2        0.46       23.61
         49 |          5        1.16       24.77
         50 |          3        0.69       25.46
         52 |        322       74.54      100.00
------------+-----------------------------------
      Total |        432      100.00

.
. *
.
. correlate _t week
(obs=432)

             |       _t     week
-------------+------------------
          _t |   1.0000
        week |   1.0000   1.0000


.
. *
.
. tab _t0, missing

        _t0 |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        432      100.00      100.00
------------+-----------------------------------
      Total |        432      100.00

.
. *
. * Notice that _st is a constant = 1.  That is, all records are to
. * be used.
. *
. * _d is identical to arrest.
. *
. * _t is identical to week.
. *
. * t0 is the time origin, and in these data it is always 0.
. *
.
. *
. * We are now ready to use the -stsplit- command.  When we do so,
. * we have to indicate how we are going to create a stacking
. * of the data.  In his example, Allison (1995) partitions 52
. * weeks into four equal length segments of 13 weeks each.  We will

. * called "J".  We will use "stsplit, syntax one."
. *
```

```
.
. stsplit J, at(13,26,39)
(1141 observations (episodes) created)

.
. *
. * Let's look at what we have accomplished.  Notice that -stsplit-
. * has created a "stacked" data set.
. * Before listing the cases, I'll sort on week within ID.
. *
.
. sort ID _t

.
. list ID _t _d _t0 age J if ID < 21 | ID > 428

              ID         _t        _d        _t0       age          J
    1.         1          1         1          0        20          0
    2.         2          2         1          0        44          0
    3.         3          3         1          0        30          0
    4.         4          4         1          0        18          0
    5.         5          5         1          0        19          0
    6.         6          6         1          0        19          0
    7.         7          7         1          0        20          0
    8.         8          8         1          0        28          0
    9.         9          8         1          0        21          0
   10.        10          8         1          0        40          0
   11.        11          8         1          0        20          0
   12.        12          8         1          0        23          0
   13.        13          9         1          0        26          0
   14.        14          9         1          0        30          0
   15.        15         10         1          0        21          0
   16.        16         11         1          0        19          0
   17.        17         11         1          0        19          0
   18.        18         12         1          0        22          0
   19.        19         12         1          0        27          0
   20.        20         13         1          0        23          0
 1558.       429         13         0          0        33          0
 1559.       429         26         0         13        33         13
 1560.       429         39         0         26        33         26
 1561.       429         52         1         39        33         39
 1562.       430         13         0          0        25          0
 1563.       430         26         0         13        25         13
 1564.       430         39         0         26        25         26
 1565.       430         52         1         39        25         39
 1566.       431         13         0          0        21          0
 1567.       431         26         0         13        21         13
 1568.       431         39         0         26        21         26
 1569.       431         52         1         39        21         39
 1570.       432         13         0          0        23          0
 1571.       432         26         0         13        23         13
 1572.       432         39         0         26        23         26
 1573.       432         52         1         39        23         39

.

. * Now let's provide more mnemonic names for _t0, _t, and _d,
```

```
. * and also create an elapsed time variable.
. *
.
. gen origin = _t0

.
. gen ending = _t

.
. gen event = _d

.
. gen elapsed = ending - origin

.
. *
. * There is more to be aware of here.  Consider the following.
. *
.
. tab J

          J |      Freq.     Percent        Cum.
------------+-----------------------------------
          0 |        432       27.46       27.46
         13 |        412       26.19       53.66
         26 |        378       24.03       77.69
         39 |        351       22.31      100.00
------------+-----------------------------------
      Total |       1573      100.00

.
. *
. * Notice that the number of "observations" is 1,573.  But if we
. *
.
. tab J arrest

           |    Arrested after
           |    release vs. not
         J |     other   arrested |     Total
-----------+----------------------+----------
         0 |         0         20 |        20
        13 |         0         34 |        34
        26 |         0         27 |        27
        39 |       318         33 |       351
-----------+----------------------+----------
     Total |       318        114 |       432


.
. *
. * We see that the N is 432.  To help understand what is going on
. * here, let's try
. *
.
. tab J _d
```

```
           |          _d
         J |         0          1 |     Total
-----------+----------------------+----------
         0 |       412         20 |       432
        13 |       378         34 |       412
        26 |       351         27 |       378
        39 |       318         33 |       351
-----------+----------------------+----------
     Total |      1459        114 |      1573


.
. *
. * Now we are back to an N of 1,573.  My account is that _d was
. * duplicated by -stsplit-, and that it is exactly what we want! When
. * we stack the data, we are actually replicating observations.
. * That is why I chose to label _d as event, to reinforce that this
. * is the correct dependent variable for our intended piecewise
. * exponential regression.
. *
. *
. * Finally, we need to -stset- the data again, in order to specify
. * the correct time to event and actual event correctly!
. *
.
. stset elapsed, failure(event)

     failure event:  event ~= 0 & event ~= .
obs. time interval:  (0, elapsed]
 exit on or before:  failure

------------------------------------------------------------------------------
     1573  total obs.
        0  exclusions
------------------------------------------------------------------------------
     1573  obs. remaining, representing
      114  failures in single record/single failure data
    19809  total analysis time at risk, at risk from t =         0
                             earliest observed entry t =         0
                                  last observed exit t =        13

.
. list ID _t elapsed _d event _t0 age J if ID < 21 | ID > 428

             ID         _t    elapsed         _d      event        _t0        age          J
  1.          1          1          1          1          1          0         20          0
  2.          2          2          2          1          1          0         44          0
  3.          3          3          3          1          1          0         30          0
  4.          4          4          4          1          1          0         18          0
  5.          5          5          5          1          1          0         19          0
  6.          6          6          6          1          1          0         19          0
  7.          7          7          7          1          1          0         20          0
  8.          8          8          8          1          1          0         28          0
  9.          9          8          8          1          1          0         21          0
 10.         10          8          8          1          1          0         40          0
 11.         11          8          8          1          1          0         20          0
 12.         12          8          8          1          1          0         23          0
 13.         13          9          9          1          1          0         26          0
 14.         14          9          9          1          1          0         30          0
 15.         15         10         10          1          1          0         21          0
 16.         16         11         11          1          1          0         19          0
```

```
  17.        17          11          11         1          1          0        19          0
  18.        18          12          12         1          1          0        22          0
  19.        19          12          12         1          1          0        27          0
  20.        20          13          13         1          1          0        23          0
1558.       429          13          13         0          0          0        33          0
1559.       429          13          13         0          0          0        33         13
1560.       429          13          13         0          0          0        33         26
1561.       429          13          13         1          1          0        33         39
1562.       430          13          13         0          0          0        25          0
1563.       430          13          13         0          0          0        25         13
1564.       430          13          13         0          0          0        25         26
1565.       430          13          13         1          1          0        25         39
1566.       431          13          13         0          0          0        21          0
1567.       431          13          13         0          0          0        21         13
1568.       431          13          13         0          0          0        21         26
1569.       431          13          13         1          1          0        21         39
1570.       432          13          13         0          0          0        23          0
1571.       432          13          13         0          0          0        23         13
1572.       432          13          13         0          0          0        23         26
1573.       432          13          13         1          1          0        23         39

.
. correlate _t elapsed
(obs=1573)

             |       _t  elapsed
-------------+------------------
          _t |   1.0000
     elapsed |   1.0000   1.0000


.
. sum _t elapsed

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+-----------------------------------------------------
          _t |    1573    12.59313   1.758224          1         13
     elapsed |    1573    12.59313   1.758224          1         13

.
. sum _t0

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+-----------------------------------------------------
         _t0 |    1573           0          0          0          0

.
. correlate _d event
(obs=1573)

             |       _d    event
-------------+------------------
          _d |   1.0000
       event |   1.0000   1.0000


.
. * Notice that _t and _t0 have changed, and are now correct for
. * the stacked data approach.  Thus, the maximum value of elapsed
. * time (_t) is 13, which is what we need here for a piecewise
. * exponential approach (13 weeks per quarter), in which we will
```

```
. * include dummies to allow for a shift in the hazard from quarter
. * to quarter.
. *
.
. capture save c:\courses\soc213b\data\recid_pwexp_042302.dta, replace

.
. *
. *-------------------------------------------------------------
. * Now we are ready to try to replicate Allison's Output 4.20.
. * Before doing that, let's review the steps that need to be taken
. * in order to generate the stacked data set for piecewise
. * exponential regression.
. *
. * Step 1.  use  <pathname>recid.dta
. *
. * Step 2.  Create an ID variable of your choice if not already
. *          present.
. *
. * Step 3.  -stset- the data using the ID variable.  In the present
. *          instance we used: stset week, id(ID) failure(arrest)
. *
. * Step 4.  -stsplit- the data at the desired cut-points.  In the
. *          present instance, we used:  stsplit J, at(13,26,39)
. *
. * Step 5.  -stset- the newly stacked data.  In the present instance
. *          we used: stset elapsed, failure(event)
. *          Note that "event" is simply _d, but "elapsed" had to be
. *          created as _t-_t0  This is important!
. *
. * This really amounts to three steps, because step 1 is necessary
. * for all analysis and step 2 is generally unnecessary since the
. * vast majority of data sets are created with ID variables.
. *-------------------------------------------------------------
. *
. * The "J" variable: SAS's default for a categorical variable spread
. * as dummies in a regression is to omit the last category.  Stata's
. * is to omit the first category.  Here's a trick that also
. * makes it easy to do significance testing of all contrasts.
. *
. * First, let's do Allison's regression omitting the first dummy
. * (Stata's default).
. *
.
. xi: streg fin age race wexp mar paro prio i.J, d(e) time
i.J               _IJ_0-39            (naturally coded; _IJ_0 omitted)

         failure _d:  event
   analysis time _t:  elapsed


Iteration 0:   log likelihood = -496.66912
Iteration 1:   log likelihood = -479.45999
Iteration 2:   log likelihood = -476.32072
Iteration 3:   log likelihood = -476.30169
Iteration 4:   log likelihood = -476.30168
```

```
No. of subjects =         1573                     Number of obs   =      1573
No. of failures =          114
Time at risk    =        19809
                                                   LR chi2(10)     =     40.73
Log likelihood  =   -476.30168                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fin |   .3773538   .1913391     1.97   0.049     .0023361    .7523715
         age |   .0569996   .0219639     2.60   0.009     .0139512     .100048
        race |  -.3125937   .3079846    -1.01   0.310    -.9162324     .291045
        wexp |   .1489036   .2121903     0.70   0.483    -.2669818    .5647889
         mar |   .4331109   .3817913     1.13   0.257    -.3151863    1.181408
        paro |   .0836245   .1957112     0.43   0.669    -.2999623    .4672114
        prio |  -.0908704   .0286358    -3.17   0.002    -.1469956   -.0347452
      _IJ_13 |  -.6318795   .2820101    -2.24   0.025    -1.184609   -.0791497
      _IJ_26 |  -.5067807   .2955675    -1.71   0.086    -1.086082    .0725209
      _IJ_39 |  -.8201861   .2841418    -2.89   0.004    -1.377094   -.2632784
       _cons |   4.542809   .6216911     7.31   0.000     3.324316    5.761301
------------------------------------------------------------------------------

.
. *
. * Next, let's do Allison's regression omitting the last dummy
. * (SAS's default).
. *
.
. char J[omit] 39

.
. xi: streg fin age race wexp mar paro prio i.J, d(e) time
i.J               _IJ_0-39            (naturally coded; _IJ_39 omitted)

         failure _d:  event
   analysis time _t:  elapsed


Iteration 0:   log likelihood = -496.66912
Iteration 1:   log likelihood = -479.45999
Iteration 2:   log likelihood = -476.32072
Iteration 3:   log likelihood = -476.30169
Iteration 4:   log likelihood = -476.30168

Exponential regression -- accelerated failure-time form

No. of subjects =         1573                     Number of obs   =      1573
No. of failures =          114
Time at risk    =        19809
                                                   LR chi2(10)     =     40.73
Log likelihood  =   -476.30168                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fin |   .3773538   .1913391     1.97   0.049     .0023361    .7523715
```

```
         age |   .0569996   .0219639     2.60   0.009     .0139512     .100048
        race |  -.3125937   .3079846    -1.01   0.310    -.9162324     .291045
        wexp |   .1489036   .2121903     0.70   0.483    -.2669818    .5647889
         mar |   .4331109   .3817913     1.13   0.257    -.3151863    1.181408
        paro |   .0836245   .1957112     0.43   0.669    -.2999623    .4672114
        prio |  -.0908704   .0286358    -3.17   0.002    -.1469956   -.0347452
       _IJ_0 |   .8201861   .2841418     2.89   0.004     .2632784    1.377094
      _IJ_13 |   .1883067   .2446029     0.77   0.441    -.2911061    .6677195
      _IJ_26 |   .3134054   .2595641     1.21   0.227    -.1953309    .8221418
       _cons |   3.722622   .6081572     6.12   0.000     2.530656    4.914589
------------------------------------------------------------------------------

.
. *
. * And in the remainder of the regressions, we will permute the
. * remaining categories as "omitted."
. *
.
. char J[omit] 26

.
.  xi: streg fin age race wexp mar paro prio i.J, d(e) time
i.J               _IJ_0-39            (naturally coded; _IJ_26 omitted)

         failure _d:  event
   analysis time _t:  elapsed


Iteration 0:   log likelihood = -496.66912
Iteration 1:   log likelihood = -479.45999
Iteration 2:   log likelihood = -476.32072
Iteration 3:   log likelihood = -476.30169
Iteration 4:   log likelihood = -476.30168

Exponential regression -- accelerated failure-time form

No. of subjects =         1573                     Number of obs   =      1573
No. of failures =          114
Time at risk    =        19809
                                                   LR chi2(10)     =     40.73
Log likelihood  =   -476.30168                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fin |   .3773538   .1913391     1.97   0.049     .0023361    .7523715
         age |   .0569996   .0219639     2.60   0.009     .0139512     .100048
        race |  -.3125937   .3079846    -1.01   0.310    -.9162324     .291045
        wexp |   .1489036   .2121903     0.70   0.483    -.2669818    .5647889
         mar |   .4331109   .3817913     1.13   0.257    -.3151863    1.181408
        paro |   .0836245   .1957112     0.43   0.669    -.2999623    .4672114
        prio |  -.0908704   .0286358    -3.17   0.002    -.1469956   -.0347452
       _IJ_0 |   .5067807   .2955675     1.71   0.086    -.0725209    1.086082
      _IJ_13 |  -.1250988   .2578852    -0.49   0.628    -.6305446     .380347
      _IJ_39 |  -.3134054   .2595641    -1.21   0.227    -.8221418    .1953309
       _cons |   4.036028   .6091373     6.63   0.000     2.842141    5.229915
------------------------------------------------------------------------------
```

```
.
. char J[omit] 13

.
. xi: streg fin age race wexp mar paro prio i.J, d(e) time
i.J               _IJ_0-39            (naturally coded; _IJ_13 omitted)

         failure _d:  event
   analysis time _t:  elapsed


Iteration 0:   log likelihood = -496.66912
Iteration 1:   log likelihood = -479.45999
Iteration 2:   log likelihood = -476.32072
Iteration 3:   log likelihood = -476.30169
Iteration 4:   log likelihood = -476.30168

Exponential regression -- accelerated failure-time form

No. of subjects =         1573                     Number of obs   =      1573
No. of failures =          114
Time at risk    =        19809
                                                   LR chi2(10)     =     40.73
Log likelihood  =   -476.30168                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fin |   .3773538   .1913391     1.97   0.049     .0023361    .7523715
         age |   .0569996   .0219639     2.60   0.009     .0139512     .100048
        race |  -.3125937   .3079846    -1.01   0.310    -.9162324     .291045
        wexp |   .1489036   .2121903     0.70   0.483    -.2669818    .5647889
         mar |   .4331109   .3817913     1.13   0.257    -.3151863    1.181408
        paro |   .0836245   .1957112     0.43   0.669    -.2999623    .4672114
        prio |  -.0908704   .0286358    -3.17   0.002    -.1469956   -.0347452
       _IJ_0 |   .6318795   .2820101     2.24   0.025     .0791497    1.184609
      _IJ_26 |   .1250988   .2578852     0.49   0.628     -.380347    .6305446
      _IJ_39 |  -.1883067   .2446029    -0.77   0.441    -.6677195    .2911061
       _cons |   3.910929   .6038977     6.48   0.000     2.727311    5.094547
------------------------------------------------------------------------------

.
. *
. * We can reset the first category of J as "omitted."
. *
.
. char J[omit] 0                   /* Back to the Stata default. */

.
. xi: streg fin age race wexp mar paro prio i.J, d(e) time
i.J               _IJ_0-39            (naturally coded; _IJ_0 omitted)

         failure _d:  event
   analysis time _t:  elapsed
```

```
Iteration 0:   log likelihood = -496.66912
Iteration 1:   log likelihood = -479.45999
Iteration 2:   log likelihood = -476.32072
Iteration 3:   log likelihood = -476.30169
Iteration 4:   log likelihood = -476.30168

Exponential regression -- accelerated failure-time form

No. of subjects =         1573                     Number of obs   =      1573
No. of failures =          114
Time at risk    =        19809
                                                   LR chi2(10)     =     40.73
Log likelihood  =   -476.30168                     Prob > chi2     =    0.0000

------------------------------------------------------------------------------
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         fin |   .3773538   .1913391     1.97   0.049     .0023361    .7523715
         age |   .0569996   .0219639     2.60   0.009     .0139512     .100048
        race |  -.3125937   .3079846    -1.01   0.310    -.9162324     .291045
        wexp |   .1489036   .2121903     0.70   0.483    -.2669818    .5647889
         mar |   .4331109   .3817913     1.13   0.257    -.3151863    1.181408
        paro |   .0836245   .1957112     0.43   0.669    -.2999623    .4672114
        prio |  -.0908704   .0286358    -3.17   0.002    -.1469956   -.0347452
      _IJ_13 |  -.6318795   .2820101    -2.24   0.025    -1.184609   -.0791497
      _IJ_26 |  -.5067807   .2955675    -1.71   0.086    -1.086082    .0725209
      _IJ_39 |  -.8201861   .2841418    -2.89   0.004    -1.377094   -.2632784
       _cons |   4.542809   .6216911     7.31   0.000     3.324316    5.761301
------------------------------------------------------------------------------

.
. *
. * Now notice that Stata makes it easy to test the global null
. * hypothesis that the division of time into four segments makes no
. * difference.  This will be a Wald test on the coefficients of the
. * included J-dummies.  It does not matter which three of the four
. * possible dummies are included.  The Stata command is
. *
.
. test _IJ_13 _IJ_26 _IJ_39

 ( 1)  [_t]_IJ_13 = 0.0
 ( 2)  [_t]_IJ_26 = 0.0
 ( 3)  [_t]_IJ_39 = 0.0

           chi2(  3) =    8.70
         Prob > chi2 =    0.0336


.
. *
. * ---------------------------------------------------------------
. * IN CONCLUSION
. *
. * Now that we've been through this process, let's review. The
. * basic question is:  What is -stsplit- doing?
. *
```

```
. * From
. *
. * stset week, id(ID) failure(arrest)
. *
. * Stata knows that the time-to-event variable is "week", it knows
. * that individuals can be uniquely identified by "ID", and that
. * failures are indicated by "arrest".  -stset- quietly creates
. * _t, _t0, _d, and _st.
. *
. * stsplit J, at(13,26,39)
. *
. * tells Stata to make risk sets for weeks 1-13, 14-26, 27-39,
. * 40-52.  Each risk set is of successively smaller size.
. *
. * Stacking the data with the -stsplit- command requires that
. * we modify the variable that measures, in the stacked data, the
. * time to event, so that time to event is defined relative to
. * each component of the stack (quarter in this example).  Once we
. * have done this, _t0 will be a constant 0, and _t will range
. * between 1 and 13--in other words, between the first and
. * thirteenth weeks of a thirteen week time segment. Thus, _t,
. * or more generally, _t - _t0, will express duration from the
. * start of the most recent one-fourth of months (in some
. * problems, origin time would vary with individual). That is,
. * duration will be reset for every 13 week time segment.  This is
. * what we want.
. *
. * We communicate all of this to Stata with ANOTHER -stset- command.
. *
. * stset elapsed, failure(event)
. *
. * where elapsed = _t-_t0 and event = _d
. *
. * Thus, the sequence of
. *     stset
. *     stsplit
. *     stset
. * with proper arguments keeps track of everything we need,
. * but we do need to be aware that we are specifying a NEW failure
. * variable in the second -stset- command.
. *
.
. log close
       log:  c:\courses\soc213b\log\recid_pwexp_042302.log
  log type:  text
 closed on:  25 Apr 2002, 23:24:16
```