Statistical and Econometric Methods for Transportation Data Analysis

Chapter 10 – Duration Models

Example 10.2
Fully Parametric Duration Models

You are given 204 observations from a travel survey conducted in the spring of 1988, in the Seattle area. While the purpose of the survey was to study the number of times (per week) commuters' changed their departure time on their work-to-home trip to avoid traffic congestion, we also have information on the length of time that they delay their trips to avoid congestion. The length of time commuters' delay is ideally suited to duration models.

Your task is to estimate, Weibull, Weibull model with gamma heterogeneity and log-logistic hazard models. Please note that LIMDEP actually estimates the parameter vector -β instead of just β so that the effect of the covariates on the hazard is:

$$EXP(-\beta X)$$

This means that a negative parameter in LIMDEP increases the hazard and thus decreases the duration. So the negative sign gives the effect on duration instead of on the hazard.

In your analysis include:

1. The results of your best model specification.

2. Show and discuss the shape of the hazard function of your best specifications.

2. A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include *t*-statistics and justify the sign of your variables.

Variables available for your specification are (file Ex10-2.txt)

Variable Number	Explanation
x1	Household number
x2	Do you ever delay work-to-home departure to avoid traffic congestion? 1-yes, 0-no
x3	If sometimes delay, on average how many minutes do you delay?
x4	If sometimes delay, do you 1-perform additional work, 2-engage in non-work activities, or 3-do both?
x5	If sometimes delay, how many times have you delayed in the past week?
x6	Mode of transportation used work-to-home: 1-car SOV, 2-carpool, 3-vanpool, 4-bus, 5 other.
x7	Primary route (work-to-home): 1-I90, 2-I5, 3-SR520, 4-I405, 5-other
x8	Do you generally encounter traffic congestion on you work-to-home trip? 1-yes, 2-no
x9	Age in years: 1-(<25), 2-(26-30), 3-(31-35), 4-(36-40), 5-(41-45), 6-(46-50), 7-(>50)
x10	Gender: 1-male, 0-female
x11	Number of cars in household
x12	Number of children in household
x13	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - 50000 to 59999, 6 - >60000
x14	Do you have flexible work hours? 1-yes, 0-no
x15	Distance from work to home (in miles)
x16	Face LOS D or worse? 1-yes, 0-no
x17	Ratio of actual travel time to free-flow travel time
x18	Population of work zone
x19	Retail employment in work zone
x20	Service employment in work zone
x21	Size of work zone (in acres)

```
--> RESET
--> sample;1-204$
--> read;nvar=21;nobs=204;file=D:Ex10-2.txt$
--> reject;x3=0$
--> dstat;rhs=x3$
                             Descriptive Statistics
              All results based on nonmissing observations.
Variable        Mean         Std.Dev.        Minimum         Maximum
Cases
--------------------------------------------------------------------------------
----
X3        51.2916667      37.4671552      4.00000000      240.000000
96
```

```
--> create;if(x6=1)car=1$
--> create;ltime=log(x3)$
--> create;if(x9>6)old=1$
--> dstat;rhs=car$
                             Descriptive Statistics
              All results based on nonmissing observations.
Variable        Mean         Std.Dev.        Minimum         Maximum
Cases
--------------------------------------------------------------------------------
----
CAR       .718750000      .451969375      .000000000      1.00000000
96
```

```
--> survival;lhs=ltime;rhs=one,x15,x17,x12;model=weibull$

+----------------------------------------------------------------------+
| Log-linear survival regression model: WEIBULL                        |
| Least squares is used to obtain starting values for MLE.             |
| Ordinary    least squares regression    Weighting variable = none   |
| Dep. var. = LTIME    Mean=   3.706843804    , S.D.=   .6997312277    |
| Model size: Observations =      96, Parameters =   4, Deg.Fr.=     92 |
| Residuals:  Sum of squares= 39.91702844    , Std.Dev.=         .65870 |
| Fit:        R-squared=  .141832, Adjusted R-squared =          .11385 |
| Model test: F[  3,     92] =    5.07,    Prob value =          .00271 |
| Diagnostic: Log-L =    -94.0959, Restricted(b=0) Log-L =    -101.4378 |
|             LogAmemiyaPrCrt.=    -.794, Akaike Info. Crt.=      2.044 |
+----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  1.619000121      .54359827        2.978   .0029
 X15       .4225283769E-01  .15618491E-01    2.705   .0068  7.7083333
 X17       .9020960070      .24142168        3.737   .0002  1.9593750
 X12      -.6645707142E-02  .60944914E-01    -.109   .9132  .81250000

Normal exit from iterations. Exit status=0.
```

```
+---------------------------------------------+
| Loglinear survival model: WEIBULL            |
| Maximum Likelihood Estimates                 |
| Dependent variable                 LTIME     |
| Weighting variable                   ONE     |
| Number of observations                96     |
| Iterations completed                  11     |
| Log likelihood function        -96.28262     |
+---------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
         RHS of hazard model
 Constant  1.732270225      .65862735         2.630   .0085
 X15       .3273725360E-01  .19402531E-01     1.687   .0916  7.7083333
 X17       1.055416934      .27856852         3.789   .0002  1.9593750
 X12      -.3865858378E-01  .57807767E-01     -.669   .5037  .81250000
         Ancillary parameters for survival
 Sigma     .5872525538      .55008811E-01    10.676   .0000
```

```
+----------------------------------------------------------------+
| Parameters of underlying density at data means:                |
| Parameter   Estimate   Std. Error     Confidence Interval      |
| ------------------------------------------------------------   |
| Lambda        .01793       .00121       .0156 to       .0203   |
| P            1.70284       .15951      1.3902 to      2.0155   |
| Median      44.96713      3.03851     39.0116 to     50.9226   |
| Percentiles  of  survival  distribution:                       |
| Survival       .25       .50       .75       .95               |
| Time         67.56     44.97     26.83      9.75               |
+----------------------------------------------------------------+
```

--> survival;lhs=ltime;rhs=one,x15,x17,x12;model=weibull;heterogeneity$

```
+-----------------------------------------------------------------------+
| Log-linear survival regression model: WEIBULL                         |
| Least squares is used to obtain starting values for MLE.              |
| Weibull Model with Gamma Heterogeneity                                |
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = LTIME    Mean=   3.706843804    , S.D.=   .6997312277     |
| Model size: Observations =      96, Parameters =   4, Deg.Fr.=     92 |
| Residuals:  Sum of squares= 39.91702844    , Std.Dev.=         .65870 |
| Fit:        R-squared=  .141832, Adjusted R-squared =          .11385 |
| Model test: F[  3,     92] =    5.07,    Prob value =          .00271 |
| Diagnostic: Log-L =    -94.0959, Restricted(b=0) Log-L =    -101.4378 |
|             LogAmemiyaPrCrt.=    -.794, Akaike Info. Crt.=      2.044 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  1.619000121      .54359827         2.978   .0029
 X15       .4225283769E-01  .15618491E-01     2.705   .0068  7.7083333
 X17       .9020960070      .24142168         3.737   .0002  1.9593750
 X12      -.6645707142E-02  .60944914E-01     -.109   .9132  .81250000
```

Normal exit from iterations. Exit status=0.

```
+-----------------------------------------------+
| Loglinear survival model: WEIBULL             |
| Maximum Likelihood Estimates                  |
| Dependent variable                 LTIME      |
| Weighting variable                   ONE      |
| Number of observations                96      |
| Iterations completed                  16      |
| Log likelihood function        -93.88402      |
| Weibull Model with Gamma Heterogeneity        |
+-----------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
         RHS of hazard model
 Constant  1.870386758      .58870206        3.177   .0015
 X15       .3375074414E-01  .17842561E-01    1.892   .0585  7.7083333
 X17       .8579132493      .25730277        3.334   .0009  1.9593750
 X12      -.1044830246E-01  .57608312E-01    -.181   .8561  .81250000
         Ancillary parameters for survival
 Theta     .6141476031      .39135931        1.569   .1166
 Sigma     .4212482203      .71720253E-01    5.873   .0000
```

```
+------------------------------------------------------------------+
| Parameters of underlying density at data means:                  |
| Parameter   Estimate   Std. Error     Confidence Interval        |
| -------------------------------------------------------------    |
| Lambda        .02230       .00226       .0179 to       .0267     |
| P            2.37390       .40417      1.5817 to      3.1661     |
| Median      42.16025      4.27718     33.7770 to     50.5435     |
| Percentiles  of  survival  distribution:                         |
| Survival       .25       .50       .75       .95                 |
| Time         62.34     42.16     27.55     12.92                 |
+------------------------------------------------------------------+
```

--> survival;lhs=ltime;rhs=one,x15,x17,x12;model=logistic;plot$

```
+-----------------------------------------------------------------------+
| Log-linear survival regression model: LOGISTIC                        |
| Least squares is used to obtain starting values for MLE.              |
| Ordinary    least squares regression    Weighting variable = none     |
| Dep. var. = LTIME    Mean=   3.706843804    , S.D.=   .6997312277     |
| Model size: Observations =      96, Parameters =   4, Deg.Fr.=     92 |
| Residuals:  Sum of squares= 39.91702844    , Std.Dev.=         .65870 |
| Fit:        R-squared=  .141832, Adjusted R-squared =          .11385 |
| Model test: F[  3,     92] =    5.07,    Prob value =          .00271 |
| Diagnostic: Log-L =    -94.0959, Restricted(b=0) Log-L =    -101.4378 |
|             LogAmemiyaPrCrt.=    -.794, Akaike Info. Crt.=      2.044 |
+-----------------------------------------------------------------------+
+---------+--------------+----------------+--------+---------+----------+
|Variable | Coefficient  | Standard Error |b/St.Er.|P[|Z|>z] | Mean of X|
+---------+--------------+----------------+--------+---------+----------+
 Constant  1.619000121      .54359827        2.978   .0029
 X15       .4225283769E-01  .15618491E-01    2.705   .0068  7.7083333
 X17       .9020960070      .24142168        3.737   .0002  1.9593750
 X12      -.6645707142E-02  .60944914E-01    -.109   .9132  .81250000
```

Normal exit from iterations. Exit status=0.

```
+---------------------------------------------+
| Loglinear survival model: LOGISTIC          |
| Maximum Likelihood Estimates                |
| Dependent variable                LTIME     |
| Weighting variable                  ONE     |
| Number of observations               96     |
| Iterations completed                  9     |
| Log likelihood function       -94.28102     |
+---------------------------------------------+
```

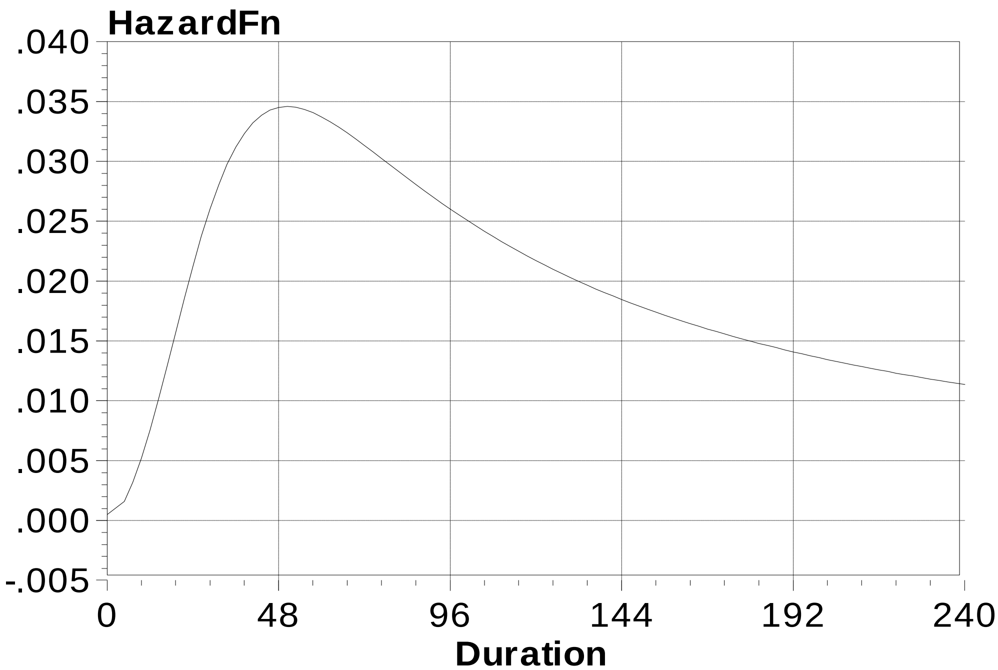
Variable	Coefficient	Standard Error	b/St.Er.	P[|Z|>z]	Mean of X
	RHS of hazard model				
Constant	1.859264488	.56577702	3.286	.0010	
X15	.3536846032E-01	.16964884E-01	2.085	.0371	7.7083333
X17	.8117401209	.24761640	3.278	.0010	1.9593750
X12	-.6399300532E-02	.55823521E-01	-.115	.9087	.81250000
	Ancillary parameters for survival				
Sigma	.3648248813	.34783222E-01	10.489	.0000	

Parameters of underlying density at data means:

Parameter	Estimate	Std. Error	Confidence Interval
Lambda	.02430	.00157	.0212 to .0274
P	2.74104	.26134	2.2288 to 3.2533
Median	41.14903	2.65177	35.9516 to 46.3465

Percentiles of survival distribution:

Survival	.25	.50	.75	.95
Time	61.44	41.15	27.56	14.06



Estimated Hazard Function