Cameron and Trivedi 18-3, 18-4,

Robert Ackerman
University of North Carolina

July 18, 2014

18-3

Consider the exponential-gamma mixture. This model is a special case of a MPH model. The survivor function, conditional on a multiplicative heterogeneity factor ν, for the exponential model is $S(t|\nu) = exp(-\mu t\nu), \lambda > 0$. The unconditional survivor function is given by the average survivor function. Averaging is across the heterogeneous population using $g(\nu)$, the density of ν, as the weighting function so $S(t) = \int_0^\infty S(t|\nu)g(\nu)\partial\nu$. Assume that ν is (two-parameter) gamma distributed with $g(\nu) = \delta^k\nu^{k-1}exp(-\delta\nu)/\Gamma(k)$.

(a)

Show that, given gamma heterogeneity, $S(t) = (1 + \mu t/\delta)^{-k}$.

$$\begin{aligned}
S(t) &= \mathbb{E}_\nu\left[S(t|\nu)\right] \\
&= \int_0^\infty S(t|\nu)g(\nu)\partial\nu \\
&= \int_0^\infty exp(-\mu t\nu)\frac{\delta^k\nu^{k-1}exp(-\delta\nu)}{\Gamma(k)}\partial\nu \text{ (take out non-}\nu\text{ parts)} \\
&= \frac{\delta^k}{\Gamma(k)}\int_0^\infty exp(-\mu t\nu)\nu^{k-1}exp(-\delta\nu)\partial\nu \\
&= \frac{\delta^k}{\Gamma(k)}\int_0^\infty \nu^{k-1}exp(-\nu(\mu t+\delta))\partial\nu \text{ (let } U = \nu(\mu t+\delta) \text{ \& integrate)} \\
&= \frac{\delta^k}{\Gamma(k)}\int_0^\infty U^{k-1}exp(-U)\partial U \\
&= \frac{\delta^k}{\Gamma(k)} \times \frac{\Gamma(k)}{\mu t+\delta)^k} \\
&= \frac{\delta^k}{\mu t+\delta)^k} = (1+\mu t/\delta)^{-k} \;\square
\end{aligned}$$

(b)

Derive expressions for the unconditional duration density function $f(t)$ and the unconditional hazard function $\lambda(t)$. These general expressions can be specialized by setting the mean of nu at 1; that is set $k = \delta$, which leads to the exponential-gamma mixture. Compare the mean and variance properties of this mixture distribution with those of the original exponential distribution.

We need to differentiate the result from (a) w.r.t t, and then take advantage of the relationship between $f(t)$, $S(t)$, and $\lambda(t)$

$$f(t) = \mu k/\delta left[1 + \mu t/\delta right]^(- k - 1) \text{ and,}$$
$$\lambda(t) = f(t)/S(t) \implies$$
$$\lambda(t) = \mu k/(\delta + \mu t)$$

(c)

Suppose that the random variable ν has a two-point distribution such that with probability π it takes the value ν_1 and with probability 910π) it takes value ν_2. What are the implications of this assumption for the specification of the unconditional survivor function? Explain your answer.

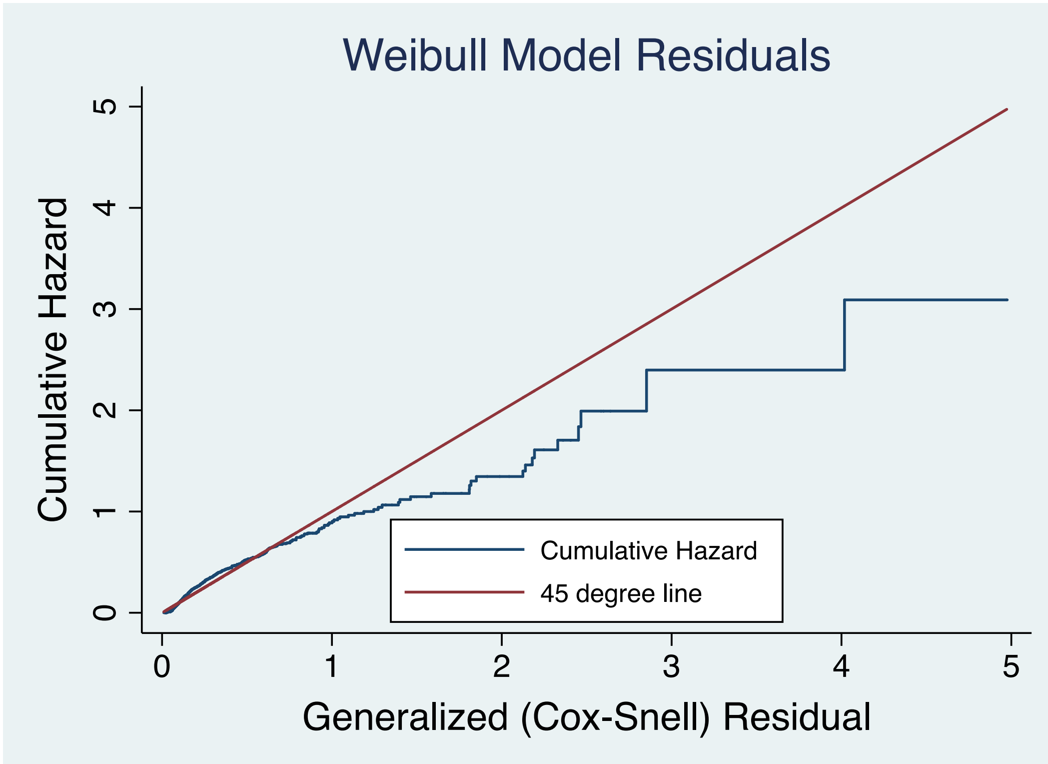
When we make this assumption, we lose the analytical solution from above that we derived from making a specific assumption about the distribution. Now we will have to estimate the integration numerically instead.

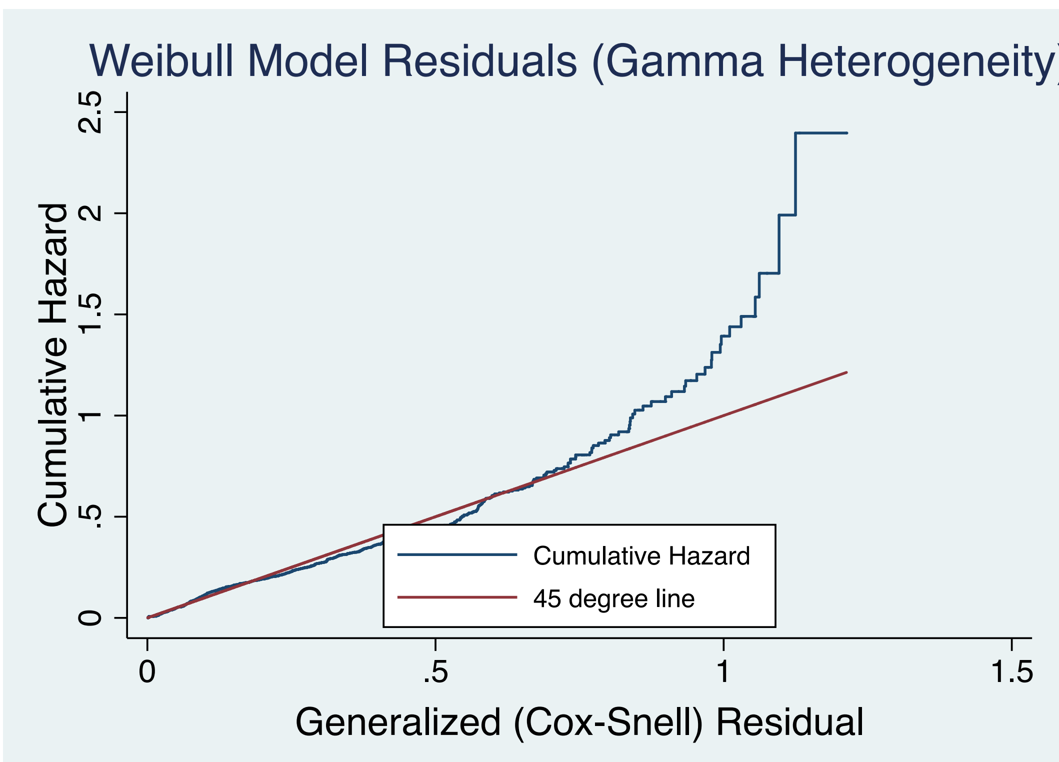
18-4

Using the sample of the McCall data set from the empirical exercise in the previous chapter, reestimate the Weibull model for this truanting to full time employment (CENSOR1 = 1) under the assumption that unobserved heterogeneity (also called frailty in some computer packages, which may also have subcommand for specifying it) has gamma distribution.

(a)

Using generalized residuals as in Section 18.7.2 test the hypothesis of model misspecification.





The two graphs above show the model residuals with no heterogeneity, and with the gamma heterogeneity assumption. Both graphs have a 45-degree line superimposed to allow for a graphical analysis of the model specification in regards to assumptions. If there were no heterogeneity *we'd expect the residuals to closely follow this 45 degree line, because there would be no relationship between $X's$ and the size of the residual. Clearly it looks like for the homoskedasticy assumption, that there is a relationship between the residuals and regressors. However, the gamma heterogeneity assumption actually makes things far worse as can be seen in the second graph.*

(b)

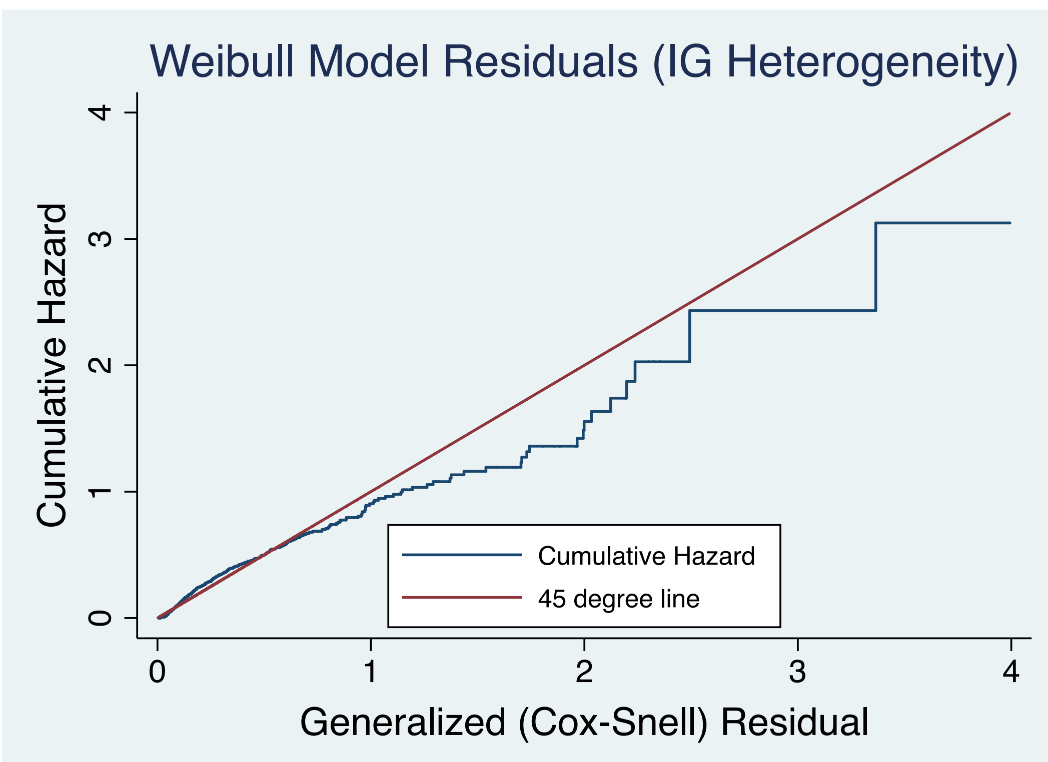
Does the new model display a duration dependence property? Does it provide a better fit to the data? Explain the results by reference to the interaction between unobserved heterogeneity and duration dependence.

It does display duration dependence, although not in the direction we would expect. With the gamma assumption it appears that the longer the duration there is an increase in the hazard rate which is the opposeite effect we would expect.

(c)

Repeat the exercise of part (a) under the assumption of log-normal heterogeneity. Are the results about duration dependence significantly different from those for the gamma heterogeneity?

Stata does not allow for log-normal specification of frailty, so instead the inverse-gaussian specification is chosen as a close substitute. As we can see from the following graph, this specification corrects for the heterogeneity by moving the latter part of the cumulative hazard closer to the 45-degree line. This is a far better specification than the gamma assumption was.



```
**** ************************************************************************************
**** Spring 2014 *********************************************************************************
**** Robert Ackerman *****************************************************************************
**** HW6 *****************************************************************************************
**** Cameron and Trivedi Microeconometrics Exercise 18-4 *****************************************
**** Note The Following Benefits Heavily from mimicking CT's mma18p1heterogeneity.do *************

**** Initial Settings ****
clear
clear matrix
capture cd "/Users/robertackerman/Desktop/Dropbox
log using "HW6_Ackerman.log", replace

set more off
pause on

**** Load Data ****
use "/Users/robertackerman/Desktop/Dropbox/
   /mma10252005/ema1996.dta", clear

**** Data Description from CT mma18p1heterogeneity.do ********************************************

* The data is from
* B.P. McCall (1996), "Unemployment Insurance Rules, Joblessness,
*                      and Part-time Work," Econometrica, 64, 647-682.

* There are 3343 observations from the CPS Displaced Worker Surveys
* of 1986, 1988, 1990 and 1992 on 33 variables including
*    spell = length of spell in number of two-week intervals
*  CENSOR1 = 1 if re-employed at full-time job

* See program mma17p4duration.do for further description of the data set
**************************************************************************************************

```

```
**** Data Description from CT mma17p4duration.do
*************************************************
** The data is from **
** B.P. McCall (1996), "Unemployment Insurance Rules, Joblessness, **
** and Part-time Work," Econometrica, 64, 647-682. **
**
** There are 3343 observations from the CPS Displaced Worker Surveys **
** of 1986, 1988, 1990 and 1992  **
** 1. spell  is length of spell in number of two-week intervals **
** 2. CENSOR1 = 1 if re-employed at full-time job **
** 3. CENSOR2 = 1 if re-employed at part-time job **
** 4. CENSOR3 = 1 if re-employed but left job: pt-ft status unknown **
** 5. CENSOR4 = 1 if still jobless **
** 6. ui (UI) = 1 if filed UI claim **
** 7. reprate (RR) = eligible replacement rate **
** 8. disrate (DR) = eligible disregard rate **
** 9. tenure (TENURE) = years tenure in lost job **
** 10. logwage (LOGWAGE) = log weekly earnings in lost job (1985$) **
** 11.-43. other variables listed in McCall (1986) table 2 p.657 **
*********************************************************************
*****************************

** Fix Seed, so  the Same Subsample is Drawn Each Time **
set seed 123456789

** 50% Randomly Generated Subsample **
sample 50

** Rename Variables to Match CT **
rename (reprate disrate ui logwage) (RR DR UI LOGWAGE)

** Generate Interaction Terms **
gen RRUI = RR*UI
gen DRUI = DR*UI

**** (a)/(b) Model specification? ****

** Define the dependent variable and the censoring variable **
stset spell, fail(censor1=1)
stdes

** Define $xlist = list of regressors used in subsequent regressions **
global xlist RR DR UI RRUI DRUI LOGWAGE

** First the Weibull w/o heterogeneity as the base **
```

```
stset spell, fail(censor1=1)
streg $xlist, nohr dist(weibull) vce(robust)
estimates store bweib

** Graph residuals and 45 degree line **
predict resid, csnell
stset resid, fail(censor1)
sts generate survivor=s
generate cumhaz = -ln(survivor)
sort resid
graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i) msize(small)
clstyle(p2)), scale (1.2) plotregion(style(none)) title("Weibull
Model Residuals") xtitle("Generalized (Cox-Snell) Residual", size(
medlarge)) xscale(titlegap(*5)) ytitle("Cumulative Hazard", size(
medlarge)) yscale(titlegap(*5)) legend(pos(6) ring(0) col(1)) legend
(size(small)) legend( label(1 "Cumulative Hazard") label(2 "45
degree line"))
graph export ECON870HW6GraphA.pdf, replace
drop resid survivor cumhaz


** Now, Weibull with gamma heterogeneity **

stset spell, fail(censor1=1)
streg $xlist, nohr dist(weibull) frailty(gamma) vce(robust) iterate(
25)
estimates store bweibG

** graph residuals and 45 degree line **
predict resid, csnell
stset resid, fail(censor1)
sts generate survivor=s
generate cumhaz = -ln(survivor)
sort resid
graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i) msize(small)
clstyle(p2)), scale (1.2) plotregion(style(none)) title("Weibull
Model Residuals (Gamma Heterogeneity)") xtitle("Generalized
(Cox-Snell) Residual", size(medlarge)) xscale(titlegap(*5)) ytitle(
"Cumulative Hazard", size(medlarge)) yscale(titlegap(*5)) legend(pos
(6) ring(0) col(1)) legend(size(small)) legend( label(1 "Cumulative
Hazard") label(2 "45 degree line"))
graph export ECON870HW6GraphB.pdf, replace
drop resid survivor cumhaz


**** (c) log normal specification ****
```

```
** Now, Weibull but with log normal heterogeneity instead **
stset spell, fail(censor1=1)
streg $xlist, nohr dist(weibull) frailty(i) vce(robust)
estimates store bweibIG

** Graph residuals and 45 degree line **
predict resid, csnell
stset resid, fail(censor1)
sts generate survivor=s
generate cumhaz = -ln(survivor)
sort resid
graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i) msize(small)
clstyle(p2)), scale (1.2) plotregion(style(none)) title("Weibull
Model Residuals (IG Heterogeneity)") xtitle("Generalized
(Cox-Snell) Residual", size(medlarge)) xscale(titlegap(*5)) ytitle(
"Cumulative Hazard", size(medlarge)) yscale(titlegap(*5)) legend(pos
(6) ring(0) col(1)) legend(size(small)) legend( label(1 "Cumulative
Hazard") label(2 "45 degree line"))
graph export ECON870HW6GraphC.pdf, replace
drop resid survivor cumhaz

** estimates for the table **
estimates table bweibIG bweib, t(%9.3f) stats(N ll) b(%9.3f) keep(RR
 DR UI RRUI DRUI LOGWAGE _cons)

log close
```

```
----------------------------------------------------------------------
-----------------------------------------------
      name:  <unnamed>
       log:  /Users/robertackerman/Desktop/Dropbox
HW6_Ackerman.log
  log type:  text
 opened on:   7 Mar 2014, 21:10:43

.
. set more off

. pause on

.
. **** Load Data ****
. use "/Users/robertackerman/Desktop/Dropbox/        /mma10252005/
ema1996.dta", clear
(Sample for 1996 EMA paper: part-time= worked part-time last week)


.
. **** Data Description from CT mma18p1heterogeneity.do
********************************************
.
. * The data is from
. * B.P. McCall (1996), "Unemployment Insurance Rules, Joblessness,
. *                      and Part-time Work," Econometrica, 64,
647-682.
.
. * There are 3343 observations from the CPS Displaced Worker Surveys
. * of 1986, 1988, 1990 and 1992 on 33 variables including
. *    spell = length of spell in number of two-week intervals
. *  CENSOR1 = 1 if re-employed at full-time job
.
. * See program mma17p4duration.do for further description of the data
set
.
**********************************************************************
****************************
.
. **** Data Description from CT mma17p4duration.do
*************************************************
. ** The data is from **
. ** B.P. McCall (1996), "Unemployment Insurance Rules, Joblessness,
**
. ** and Part-time Work," Econometrica, 64, 647-682. **
. **
. ** There are 3343 observations from the CPS Displaced Worker Surveys
**
. ** of 1986, 1988, 1990 and 1992  **
. ** 1. spell  is length of spell in number of two-week intervals **
```

```
. ** 2. CENSOR1 = 1 if re-employed at full-time job **
. ** 3. CENSOR2 = 1 if re-employed at part-time job **
. ** 4. CENSOR3 = 1 if re-employed but left job: pt-ft status unknown
**
. ** 5. CENSOR4 = 1 if still jobless **
. ** 6. ui (UI) = 1 if filed UI claim **
. ** 7. reprate (RR) = eligible replacement rate **
. ** 8. disrate (DR) = eligible disregard rate **
. ** 9. tenure (TENURE) = years tenure in lost job **
. ** 10. logwage (LOGWAGE) = log weekly earnings in lost job (1985$)
**
. ** 11.-43. other variables listed in McCall (1986) table 2 p.657 **
.
**********************************************************************
****************************
.
. ** Fix Seed, so  the Same Subsample is Drawn Each Time **
. set seed 123456789

.
. ** 50% Randomly Generated Subsample **
. sample 50
(1671 observations deleted)

.
. ** Rename Variables to Match CT **
. rename (reprate disrate ui logwage) (RR DR UI LOGWAGE)

.
. ** Generate Interaction Terms **
. gen RRUI = RR*UI

. gen DRUI = DR*UI

.
. **** (a)/(b) Model specification? ****
.
. ** Define the dependent variable and the censoring variable **
. stset spell, fail(censor1=1)

     failure event:  censor1 == 1
obs. time interval:  (0, spell]
 exit on or before:  failure

----------------------------------------------------------------------
--------
     1672  total observations
        0  exclusions
----------------------------------------------------------------------
--------
```

```
      1672  observations remaining, representing
       544  failures in single-record/single-failure data
     10567  total analysis time at risk and under observation
                                                        at risk from t =
0
                                             earliest observed entry t =
0
                                                  last observed exit t =
28

. stdes

         failure _d:  censor1 == 1
   analysis time _t:  spell

                                   |-------------- per subject
--------------|
Category                   total        mean         min     median
max
----------------------------------------------------------------------
--------
no. of subjects             1672
no. of records              1672           1           1          1
1

(first) entry time                         0           0          0
0
(final) exit time                   6.319976           1          5
28

subjects with gap              0
time on gap if gap             0
time at risk               10567    6.319976           1          5
28

failures                     544    .3253589           0          0
1
----------------------------------------------------------------------
--------

.
. ** Define $xlist = list of regressors used in subsequent regressions
**
. global xlist RR DR UI RRUI DRUI LOGWAGE

.
. ** First the Weibull w/o heterogeneity as the base **
. stset spell, fail(censor1=1)

     failure event:  censor1 == 1
```

```
obs. time interval:  (0, spell]
 exit on or before:  failure

------------------------------------------------------------------------
--------
      1672  total observations
         0  exclusions
------------------------------------------------------------------------
--------
      1672  observations remaining, representing
       544  failures in single-record/single-failure data
     10567  total analysis time at risk and under observation
                                              at risk from t =
0
                                   earliest observed entry t =
0
                                        last observed exit t =
28

. streg $xlist, nohr dist(weibull) vce(robust)

         failure _d:  censor1 == 1
   analysis time _t:  spell

Fitting constant-only model:

Iteration 0:   log pseudolikelihood = -1511.4344
Iteration 1:   log pseudolikelihood = -1511.4344
Iteration 2:   log pseudolikelihood = -1511.4344

Fitting full model:

Iteration 0:   log pseudolikelihood = -1511.4344
Iteration 1:   log pseudolikelihood = -1483.0118
Iteration 2:   log pseudolikelihood = -1414.1644
Iteration 3:   log pseudolikelihood = -1412.3803
Iteration 4:   log pseudolikelihood = -1412.3699
Iteration 5:   log pseudolikelihood = -1412.3699

Weibull regression -- log relative-hazard form

No. of subjects      =         1672                Number of obs   =
1672
No. of failures      =          544
Time at risk         =        10567
                                                   Wald chi2(6)    =
176.48
Log pseudolikelihood =   -1412.3699                Prob > chi2     =
0.0000
```

```
------------------------------------------------------------------------------
--------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------
+----------------------------------------------------------------
          RR |   .2913001   .8432765     0.35   0.730    -1.361491
1.944092
          DR |  -2.182662   1.077429    -2.03   0.043    -4.294383
-.0709405
          UI |  -2.019882   .3612006    -5.59   0.000    -2.727822
-1.311942
        RRUI |   1.856453   .8579786     2.16   0.030     .1748453
3.53806
        DRUI |   .4939231   1.438983     0.34   0.731    -2.326431
3.314277
     LOGWAGE |   .5545383   .1419545     3.91   0.000     .
2763126     .832764
       _cons |  -5.573391   1.111942    -5.01   0.000    -7.752758
-3.394024
-------------
+----------------------------------------------------------------
       /ln_p |   .0838479   .0250761     3.34   0.001     .
0346996    .1329962
-------------
+----------------------------------------------------------------
           p |   1.087463   .0272694                      1.035309
1.142246
         1/p |   .9195711   .0230593                      .
8754684    .9658955
------------------------------------------------------------------------------
--------

. estimates store bweib

.
. ** Graph residuals and 45 degree line **
. predict resid, csnell

. stset resid, fail(censor1)

     failure event:  censor1 != 0 & censor1 < .
obs. time interval:  (0, resid]
 exit on or before:  failure

------------------------------------------------------------------------------
--------
     1672  total observations
        0  exclusions
```

```
--------------------------------------------------------------------------------------
--------
     1672  observations remaining, representing
      544  failures in single-record/single-failure data
      544  total analysis time at risk and under observation
                                                    at risk from t =
0
                                         earliest observed entry t =
0
                                              last observed exit t =
4.973147

. sts generate survivor=s

. generate cumhaz = -ln(survivor)

. sort resid

. graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i)
>  msize(small) clstyle(p2)), scale (1.2) plotregion(style(none))
title("Weibull Model Residuals") xtitle("Generalize
> d (Cox-Snell) Residual", size(medlarge)) xscale(titlegap(*5))
ytitle("Cumulative Hazard", size(medlarge)) yscale(ti
> tlegap(*5)) legend(pos(6) ring(0) col(1)) legend(size(small))
legend( label(1 "Cumulative Hazard") label(2 "45 degr
> ee line"))

. graph export ECON870HW6GraphA.pdf, replace
(file /Users/robertackerman/Desktop/Dropbox/Econ 873/
ECON870HW6GraphA.pdf written in PDF format)

. drop resid survivor cumhaz

.
.
. ** Now, Weibull with gamma heterogeneity **
.
. stset spell, fail(censor1=1)

     failure event:  censor1 == 1
obs. time interval:  (0, spell]
 exit on or before:  failure

--------------------------------------------------------------------------------------
--------
     1672  total observations
        0  exclusions
--------------------------------------------------------------------------------------
--------
```

```
     1672  observations remaining, representing
      544  failures in single-record/single-failure data
    10567  total analysis time at risk and under observation
                                                   at risk from t =
0
                                       earliest observed entry t =
0
                                            last observed exit t =
28

. streg $xlist, nohr dist(weibull) frailty(gamma) vce(robust)
iterate(25)

         failure _d:  censor1 == 1
   analysis time _t:  spell

Fitting Weibull model:

Fitting constant-only model:

Iteration 0:   log pseudolikelihood = -1579.6714
Iteration 1:   log pseudolikelihood = -1513.2265
Iteration 2:   log pseudolikelihood = -1481.6809  (not concave)
Iteration 3:   log pseudolikelihood = -1479.5154
Iteration 4:   log pseudolikelihood = -1477.0543
Iteration 5:   log pseudolikelihood = -1467.4982  (backed up)
Iteration 6:   log pseudolikelihood = -1405.3517
Iteration 7:   log pseudolikelihood = -1375.1696
Iteration 8:   log pseudolikelihood = -1372.5041
Iteration 9:   log pseudolikelihood = -1371.3763
Iteration 10:  log pseudolikelihood = -1370.8551
Iteration 11:  log pseudolikelihood = -1370.7292  (backed up)
Iteration 12:  log pseudolikelihood = -1370.6669  (backed up)
Iteration 13:  log pseudolikelihood = -1370.6359  (backed up)
Iteration 14:  log pseudolikelihood = -1370.6204  (backed up)
Iteration 15:  log pseudolikelihood = -1370.6127  (backed up)
Iteration 16:  log pseudolikelihood = -1370.6088  (backed up)
Iteration 17:  log pseudolikelihood = -1370.6069  (backed up)
Iteration 18:  log pseudolikelihood = -1370.6059  (backed up)
Iteration 19:  log pseudolikelihood = -1370.6057  (backed up)
Iteration 20:  log pseudolikelihood = -1370.6056  (backed up)
Iteration 21:  log pseudolikelihood = -1370.6056  (backed up)
Iteration 22:  log pseudolikelihood = -1370.6056  (backed up)
Iteration 23:  log pseudolikelihood = -1370.6056  (backed up)
Iteration 24:  log pseudolikelihood = -1370.6056  (backed up)
Iteration 25:  log pseudolikelihood = -1370.6056  (backed up)
convergence not achieved

Fitting full model:
```

```
Iteration 0:   log pseudolikelihood = -1399.7502
Iteration 1:   log pseudolikelihood = -1360.2261
Iteration 2:   log pseudolikelihood = -1347.1368
Iteration 3:   log pseudolikelihood = -1345.4859
Iteration 4:   log pseudolikelihood = -1345.2673
Iteration 5:   log pseudolikelihood = -1345.2669
Iteration 6:   log pseudolikelihood = -1345.2669

Weibull regression -- log relative-hazard form
                      Gamma frailty

No. of subjects      =        1672                Number of obs   =
1672
No. of failures      =         544
Time at risk         =       10567
                                                  Wald chi2(6)    =
79.90
Log pseudolikelihood =   -1345.2669               Prob > chi2     =
0.0000

----------------------------------------------------------------------
--------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------
+----------------------------------------------------------------
          RR |   1.026475   1.821661     0.56   0.573    -2.543914
4.596865
          DR |  -4.171282   2.610563    -1.60   0.110
-9.287891    .9453278
          UI |  -5.775773   1.055135    -5.47   0.000      -7.8438
-3.707746
        RRUI |   3.290943   2.530113     1.30   0.193    -1.667988
8.249874
        DRUI |   1.189086    4.24754     0.28   0.780    -7.135939
9.514111
     LOGWAGE |   1.118262   .3788828     2.95   0.003      .375665
1.860858
       _cons |  -8.624274   2.840972    -3.04   0.002    -14.19248
-3.056071
-------------
+----------------------------------------------------------------
       /ln_p |   1.276361   .3275401     3.90   0.000     .6343944
1.918328
     /ln_the |   2.339281   .4896284     4.78   0.000     1.379627
3.298935
-------------
+----------------------------------------------------------------
           p |   3.583576   1.173765                       1.88588
```

```
6.809564
         1/p |   .2790508   .0914003                      .
1468523    .5302565
       theta |   10.37377   5.079295                      3.973419
27.08378
----------------------------------------------------------------------
--------

. estimates store bweibG

.
. ** graph residuals and 45 degree line **
. predict resid, csnell
(option unconditional assumed)

. stset resid, fail(censor1)

     failure event:  censor1 != 0 & censor1 < .
obs. time interval:  (0, resid]
 exit on or before:  failure

----------------------------------------------------------------------
--------
     1672  total observations
        0  exclusions
----------------------------------------------------------------------
--------
     1672  observations remaining, representing
      544  failures in single-record/single-failure data
      544  total analysis time at risk and under observation
                                             at risk from t =
0
                                  earliest observed entry t =
0
                                       last observed exit t =
1.213049

. sts generate survivor=s

. generate cumhaz = -ln(survivor)

. sort resid

. graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i)
>  msize(small) clstyle(p2)), scale (1.2) plotregion(style(none))
title("Weibull Model Residuals (Gamma Heterogeneity
> )") xtitle("Generalized (Cox-Snell) Residual", size(medlarge))
xscale(titlegap(*5)) ytitle("Cumulative Hazard", siz
> e(medlarge)) yscale(titlegap(*5)) legend(pos(6) ring(0) col(1))
```

```
legend(size(small)) legend( label(1 "Cumulative Haz
> ard") label(2 "45 degree line"))

. graph export ECON870HW6GraphB.pdf, replace
(file /Users/robertackerman/Desktop/Dropbox/Econ 873/
ECON870HW6GraphB.pdf written in PDF format)

. drop resid survivor cumhaz

. 
. 
. **** (c) log normal specification ****
. 
. ** Now, Weibull but with log normal heterogeneity instead **
. stset spell, fail(censor1=1)

     failure event:  censor1 == 1
obs. time interval:  (0, spell]
 exit on or before:  failure

----------------------------------------------------------------------
--------
      1672  total observations
         0  exclusions
----------------------------------------------------------------------
--------
      1672  observations remaining, representing
       544  failures in single-record/single-failure data
     10567  total analysis time at risk and under observation
                                              at risk from t =
0
                                   earliest observed entry t =
0
                                        last observed exit t =
28

. streg $xlist, nohr dist(weibull) frailty(i) vce(robust)

         failure _d:  censor1 == 1
   analysis time _t:  spell

Fitting Weibull model:

Fitting constant-only model:

Iteration 0:   log pseudolikelihood = -1576.5168  (not concave)
Iteration 1:   log pseudolikelihood = -1505.3222
Iteration 2:   log pseudolikelihood = -1498.5224
Iteration 3:   log pseudolikelihood = -1485.0403
Iteration 4:   log pseudolikelihood = -1482.9004
```

```
Iteration 5:   log pseudolikelihood = -1482.8926
Iteration 6:   log pseudolikelihood = -1482.8926

Fitting full model:

Iteration 0:   log pseudolikelihood = -1398.7711
Iteration 1:   log pseudolikelihood = -1382.8715
Iteration 2:   log pseudolikelihood = -1374.6906
Iteration 3:   log pseudolikelihood = -1374.4884
Iteration 4:   log pseudolikelihood = -1374.4881
Iteration 5:   log pseudolikelihood = -1374.4881

Weibull regression -- log relative-hazard form
                      Inverse-Gaussian frailty

No. of subjects      =        1672               Number of obs   =
1672
No. of failures      =         544
Time at risk         =       10567
                                                 Wald chi2(6)    =
235.66
Log pseudolikelihood =   -1374.4881              Prob > chi2     =
0.0000

----------------------------------------------------------------------
--------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf.
Interval]
-------------
+-----------------------------------------------------------------
          RR |   .2815256   1.232454     0.23   0.819     -2.13404
2.697092
          DR |  -3.482526   1.570346    -2.22   0.027    -6.560348
-.4047044
          UI |  -3.392758   .5358074    -6.33   0.000    -4.442921
-2.342595
        RRUI |   3.046113   1.277013     2.39   0.017      .543213
5.549014
        DRUI |   .9973391   2.148549     0.46   0.643    -3.213739
5.208417
     LOGWAGE |   .8475168   .2113665     4.01   0.000      .433246
1.261788
       _cons |  -6.945322   1.643688    -4.23   0.000    -10.16689
-3.723753
-------------
+-----------------------------------------------------------------
       /ln_p |     .54545   .0283293    19.25   0.000     .
4899255    .6009744
     /ln_the |   2.044921   .0959762    21.31   0.000     1.856811
```

```
2.233031
-------------
+-----------------------------------------------------------------
          p |   1.725385    .048879                      1.632195
1.823895
        1/p |   .5795809   .0164191                      .
5482771    .6126721
      theta |   7.728548   .7417564                      6.403285
9.328095
----------------------------------------------------------------------
--------

. estimates store bweibIG

.
. ** Graph residuals and 45 degree line **
. predict resid, csnell
(option unconditional assumed)

. stset resid, fail(censor1)

     failure event:  censor1 != 0 & censor1 < .
obs. time interval:  (0, resid]
 exit on or before:  failure

----------------------------------------------------------------------
--------
     1672  total observations
        0  exclusions
----------------------------------------------------------------------
--------
     1672  observations remaining, representing
      544  failures in single-record/single-failure data
      544  total analysis time at risk and under observation
                                              at risk from t =
0
                                   earliest observed entry t =
0
                                        last observed exit t =
3.991242

. sts generate survivor=s

. generate cumhaz = -ln(survivor)

. sort resid

. graph twoway (scatter cumhaz resid, c(J) msymbol(i) msize(small)
clstyle(p1)) (scatter resid resid, c(l) msymbol(i)
>  msize(small) clstyle(p2)), scale (1.2) plotregion(style(none))
```

```
title("Weibull Model Residuals (IG Heterogeneity)")
>  xtitle("Generalized (Cox-Snell) Residual", size(medlarge))
xscale(titlegap(*5)) ytitle("Cumulative Hazard", size(m
> edlarge)) yscale(titlegap(*5)) legend(pos(6) ring(0) col(1))
legend(size(small)) legend( label(1 "Cumulative Hazard
> ") label(2 "45 degree line"))

. graph export ECON870HW6GraphC.pdf, replace
(file /Users/robertackerman/Desktop/Dropbox/Econ 873/
ECON870HW6GraphC.pdf written in PDF format)

. drop resid survivor cumhaz

.
. ** estimates for the table **
. estimates table bweibIG bweib, t(%9.3f) stats(N ll) b(%9.3f) keep(RR
DR UI RRUI DRUI LOGWAGE _cons)

--------------------------------------
    Variable |  bweibIG      bweib
-------------+------------------------
          RR |     0.282       0.291
             |     0.228       0.345
          DR |    -3.483      -2.183
             |    -2.218      -2.026
          UI |    -3.393      -2.020
             |    -6.332      -5.592
        RRUI |     3.046       1.856
             |     2.385       2.164
        DRUI |     0.997       0.494
             |     0.464       0.343
     LOGWAGE |     0.848       0.555
             |     4.010       3.906
       _cons |    -6.945      -5.573
             |    -4.225      -5.012
-------------+------------------------
           N |      1672        1672
          ll | -1374.488   -1412.370
--------------------------------------
                           legend: b/t

.
. log close
      name:  <unnamed>
       log:  /Users/robertackerman/Desktop/Dropbox/
HW6_Ackerman.log
  log type:  text
 closed on:   7 Mar 2014, 21:10:51
----------------------------------------------------------------------
-----------------------------------------------
```