Illustration 6

Aim: Explaining unemployment duration, measured in number of two weeks intervals (*spell*), as a function of potential determinants such as:

- ui (UI) = 1 if filed UI claim
- reprate (RR) = eligible replacement rate
- disrate (DR) = eligible disregard rate
- tenure (TENURE) = years tenure in lost job
- logwage (LOGWAGE) = log weekly earnings in lost job (1985$)

The duration is complete when the individual is re-employeed at a full-time job, that is, when CENSOR1 = 1

Details: Cameron and Trivedi (2005), ch. 17.11

Illustration 6 – Question 1

```
. sum spell censor1 censor2 censor3 censor4 ui reprate disrate tenure logwage
```

Variable	Obs	Mean	Std. Dev.	Min	Max
spell	3343	6.247981	5.611271	1	28
censor1	3343	.3209692	.4669188	0	1
censor2	3343	.1014059	.3019106	0	1
censor3	3343	.1717021	.3771777	0	1
censor4	3343	.3754113	.4843014	0	1
ui	3343	.5527969	.4972791	0	1
reprate	3343	.4544717	.1137918	.066	2.059
disrate	3343	.1094376	.0735274	.002	1.02
tenure	3343	4.114867	5.862322	0	40
logwage	3343	5.692994	.5356591	2.70805	7.600402

- Non-negative outcome
- 32.01% of the observations present complete durations, which means that they are not censored on the right (censored observations present *censor1=0*)

Illustration 6 – Question 2

```
.stset spell, fail(censor1=1)

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

------------------------------------------------------------------------------
      3343  total observations
         0  exclusions
------------------------------------------------------------------------------
      3343  observations remaining, representing
      1073  failures in single-record/single-failure data
     20887  total analysis time at risk and under observation
                                                 at risk from t =         0
                                      earliest observed entry t =         0
                                           last observed exit t =        28
```

Illustration 6 – Question 2

```
. sts list

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

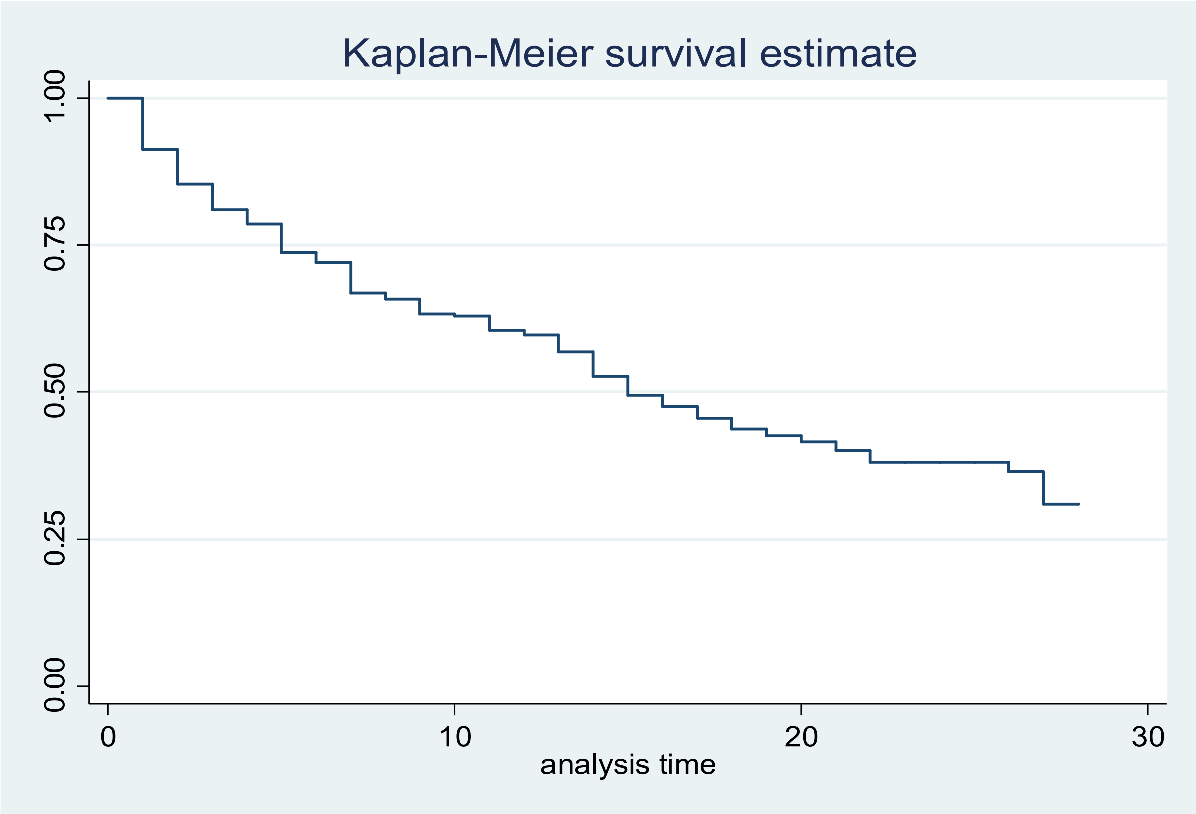
Time	Beg. Total	Fail	Net Lost	Survivor Function	Std. Error	[95% Conf.	Int.]
1	3343	294	246	0.9121	0.0049	0.9019	0.9212
2	2803	178	304	0.8541	0.0062	0.8415	0.8659
3	2321	119	305	0.8103	0.0071	0.7960	0.8238
4	1897	56	165	0.7864	0.0076	0.7712	0.8008
5	1676	104	233	0.7376	0.0085	0.7206	0.7538
6	1339	32	111	0.7200	0.0088	0.7023	0.7369
7	1196	85	178	0.6688	0.0098	0.6492	0.6876
8	933	15	70	0.6581	0.0100	0.6380	0.6773
9	848	33	98	0.6325	0.0106	0.6113	0.6528
10	717	3	55	0.6298	0.0106	0.6086	0.6503
11	659	26	77	0.6050	0.0113	0.5825	0.6267

(…)

- (1-0.9121)*100%=8.8% of the individuals exit unemployement whithin the two first weeks and (1-0.8541)*100%= 14.6% exit unemployement within one month

Illustration 6 – Question 2

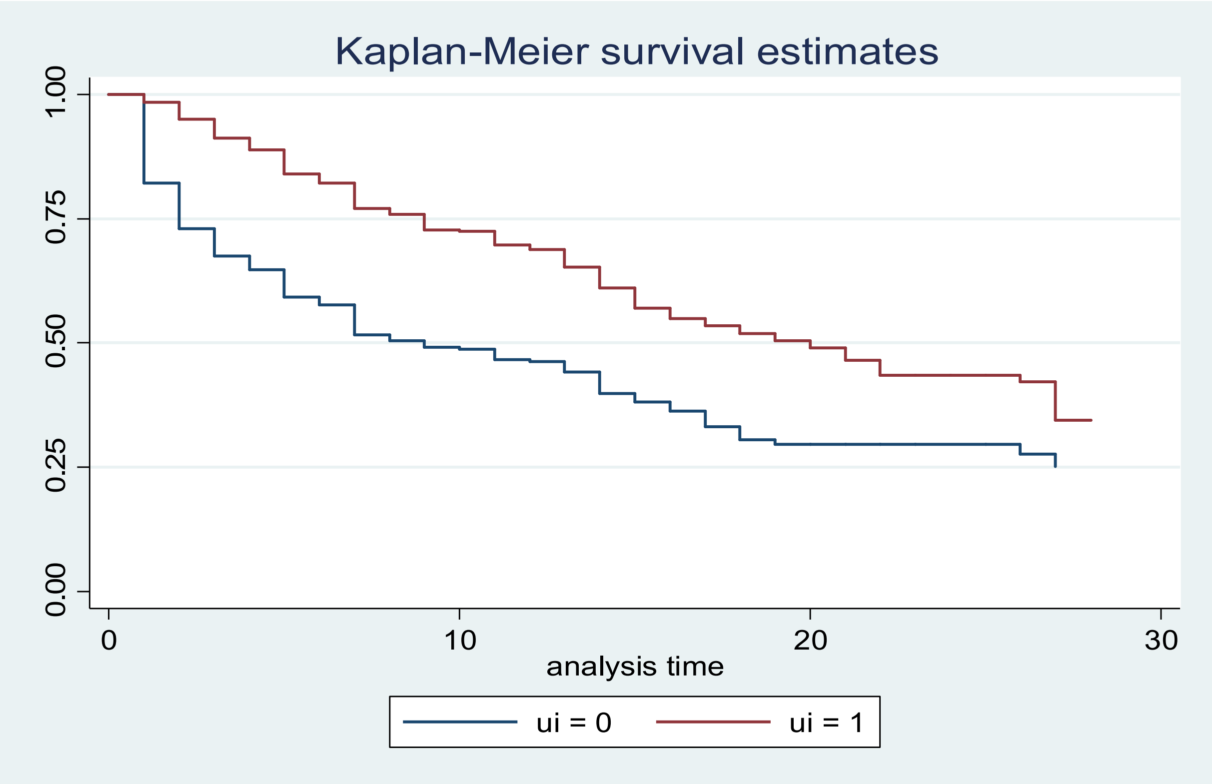
```
. sts graph
```



- Survival decreases more rapidly at first and then slowly

Illustration 6 – Question 2

```
. sts graph, by(ui)
```



- Those claiming unemployement insurance (ui=1) are more likely to remain unemployed

Illustration 6 – Question 3

```
. gen RR = reprate
. gen DR = disrate
. gen UI = ui
. gen RRUI = RR*UI
. gen DRUI = DR*UI
. gen LOGWAGE = logwage
```

```
. global xlist RR DR UI RRUI DRUI LOGWAGE tenure slack abolpos explose stateur
houshead married female child ychild nonwhite age schlt12 schgt12 smsa bluecoll
mining constr transp trade fire services pubadmin year85 year87 year89 midatl
encen wncen southatl escen wscen mountain pacific
```

Illustration 6 – Question 3

```
. streg $xlist, nohr robust dist(exponential)
         failure _d:  censor1 == 1
   analysis time _t:  spell
 (…)
Exponential regression -- log relative-hazard form
 No. of subjects      =         3343                Number of obs    =      3343
No. of failures      =         1073
Time at risk         =        20887
                                                    Wald chi2(40)    =    565.24
Log pseudolikelihood =   -2700.6903                 Prob > chi2      =    0.0000
 ------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          RR |   .4720235   .6005534     0.79   0.432    -.7050396    1.649087
          DR |  -.5756396   .7624489    -0.75   0.450    -2.070012    .9187327
          UI |  -1.424561   .2493917    -5.71   0.000     -1.91336   -.9357622
        RRUI |   .9655904   .6118408     1.58   0.115    -.2335956    2.164776
        DRUI |  -.1990635   1.019118    -0.20   0.845    -2.196498    1.798371
     LOGWAGE |   .3508005    .115598     3.03   0.002     .1242327    .5773684
…
       _cons |  -4.079107   .8767097    -4.65   0.000    -5.797426   -2.360788
------------------------------------------------------------------------------

estimates store bexponential
```

Illustration 6 – Question 3

```
. streg $xlist, nohr robust dist(weibull)
Weibull regression -- log relative-hazard form
No. of subjects      =          3343                Number of obs    =       3343
…
                                                    Wald chi2(40)    =     501.65
Log pseudolikelihood =    -2687.5995                Prob > chi2      =     0.0000
 ------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          RR |   .4481156   .6381895     0.70   0.483    -.8027127    1.698944
          DR |  -.4269187   .8086983    -0.53   0.598    -2.011938    1.158101
          UI |  -1.496066   .2639679    -5.67   0.000    -2.013434   -.9786984
        RRUI |   1.015226   .6455611     1.57   0.116    -.2500501    2.280503
        DRUI |  -.2988417   1.065384    -0.28   0.779    -2.386956    1.789272
     LOGWAGE |   .3655253     .12212     2.99   0.003     .1261745    .6048761
(…)
       _cons |  -4.357886   .9196792    -4.74   0.000    -6.160424   -2.555347
-------------+----------------------------------------------------------------
       /ln_p |   .1215314   .0194374     6.25   0.000     .0834348    .1596281
-------------+----------------------------------------------------------------
           p |   1.129225   .0219492                      1.087014    1.173075
         1/p |   .8855632   .0172131                      .8524608    .9199511
------------------------------------------------------------------------------
 . estimates store bweibull
```

Illustration 6 – Question 3

```
. streg $xlist, nohr robust dist(gompertz)
(…)
Gompertz regression -- log relative-hazard form
No. of subjects      =         3343              Number of obs   =      3343
No. of failures      =         1073
Time at risk         =        20887
                                                 Wald chi2(40)   =    529.75
Log pseudolikelihood =    -2700.605              Prob > chi2     =    0.0000
------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          RR |    .472405   .6033813     0.78   0.434    -.7102005    1.655011
          DR |  -.5627894   .7646131    -0.74   0.462    -2.061404    .9358247
          UI |  -1.428355   .2508349    -5.69   0.000    -1.919982   -.9367272
        RRUI |   .9689413   .6144464     1.58   0.115    -.2353514    2.173234
        DRUI |  -.2112495   1.021112    -0.21   0.836    -2.212593    1.790094
     LOGWAGE |   .3524722   .1162698     3.03   0.002     .1245876    .5803567
(…)
       _cons |   -4.09733   .8802997    -4.65   0.000    -5.822686   -2.371975
-------------+----------------------------------------------------------------
      /gamma |    .002658   .0067759     0.39   0.695    -.0106225    .0159386
------------------------------------------------------------------------------

. estimates store bgompertz
```

Illustration 6 – Question 3

```
. estimates table bexponential bweibull bgompertz, b star(0.1 0.05 0.01)
--------------------------------------------------------------
    Variable | bexponential      bweibull        bgompertz
-------------+------------------------------------------------
          RR |  .47202347        .4481156        .47240504
          DR | -.57563962      -.42691874      -.56278942
          UI | -1.4245611***    -1.496066***    -1.4283547***
        RRUI |  .96559044       1.0152264        .96894134
        DRUI | -.19906351      -.29884166      -.21124952
     LOGWAGE |  .35080054***    .36552527***     .35247218***
(…)
       _cons | -4.0791071***   -4.3578855***   -4.0973303***
-------------+------------------------------------------------
ln_p   _cons |                  .12153144***
-------------+------------------------------------------------
gamma  _cons |                                   .00265803
--------------------------------------------------------------
                           legend: * p<.1; ** p<.05; *** p<.01
```

- Weibull is preferred to exponential ($\alpha = \exp(0.122) = 1.129$): the probability of the spell terminating increases for longer spells
- In all the models only UI and LOGWAGE are significant, with little variation across the three models:
  - For those claiming insurance the hazard rate changes $[exp(-1.496) - 1] * 100\% = -77{,}6\%$

Illustration 6 – Question 4

```
. stcox $xlist, nohr robust
(…)
Cox regression -- Breslow method for ties

No. of subjects      =         3343              Number of obs    =       3343
No. of failures      =         1073
Time at risk         =        20887
                                                 Wald chi2(40)    =     540.98
Log pseudolikelihood =   -7717.2334              Prob > chi2      =     0.0000
------------------------------------------------------------------------------
             |               Robust
          _t |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          RR |   .5222796   .5711698     0.91   0.361    -.5971926    1.641752
          DR |   -.752507     .72175    -1.04   0.297    -2.167111    .6620971
          UI |  -1.317719   .2372893    -5.55   0.000    -1.782798   -.8526409
        RRUI |   .8822462    .582115     1.52   0.130    -.2586783    2.023171
        DRUI |  -.0951357    .977774    -0.10   0.922    -2.011538    1.821266
     LOGWAGE |   .3352639   .1106483     3.03   0.002     .1183972    .5521306
(…)
------------------------------------------------------------------------------

. estimates store bcox
```

Illustration 6 – Question 4

```
. estimates table bcox, b star(0.1 0.05 0.01) keep(RR DR UI RRUI DRUI LOGWAGE)

------------------------------
    Variable |      bcox
-------------+----------------
          RR |   .52227964
          DR |  -.75250697
          UI |  -1.3177194***
        RRUI |   .88224619
        DRUI |  -.09513573
     LOGWAGE |   .33526391***
------------------------------
legend: * p<.1; ** p<.05; *** p<.01
```

- The conclusions do not change

Illustration 6 – Question 5

The Figures are produced using the code in CT

```
. streg $xlist, nohr robust dist(exponential)
. 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("Exponential 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 exp.wmf, replace

. drop resid survivor cumhaz
```

Illustration 6 – Question 5

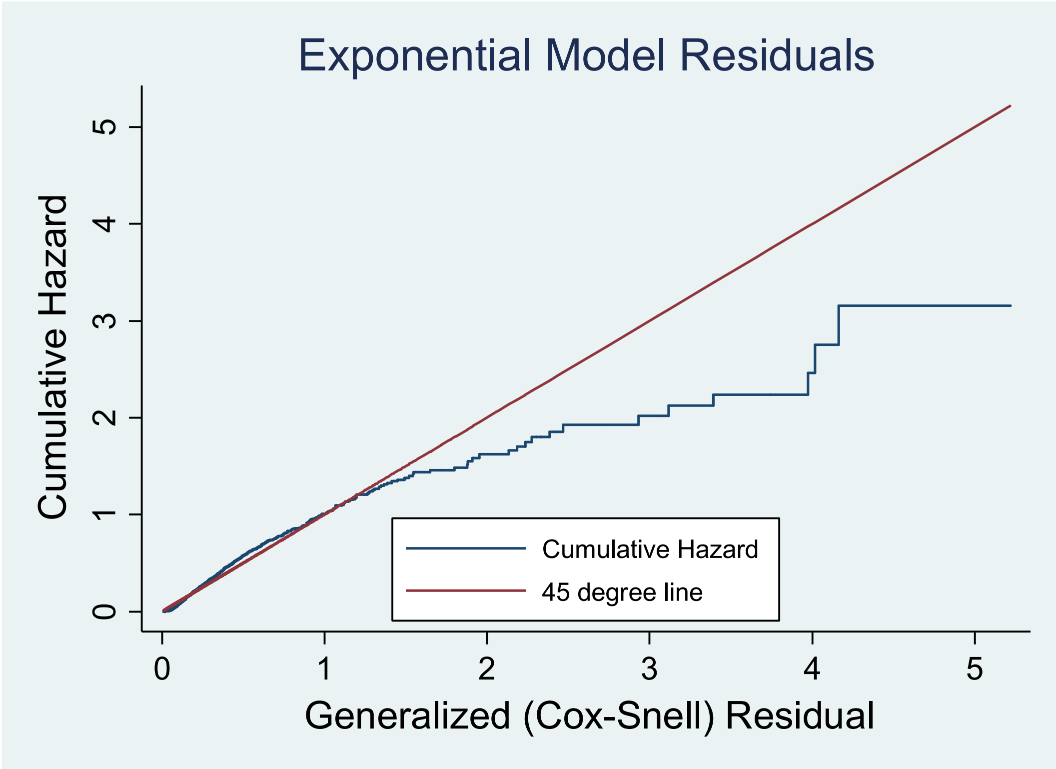
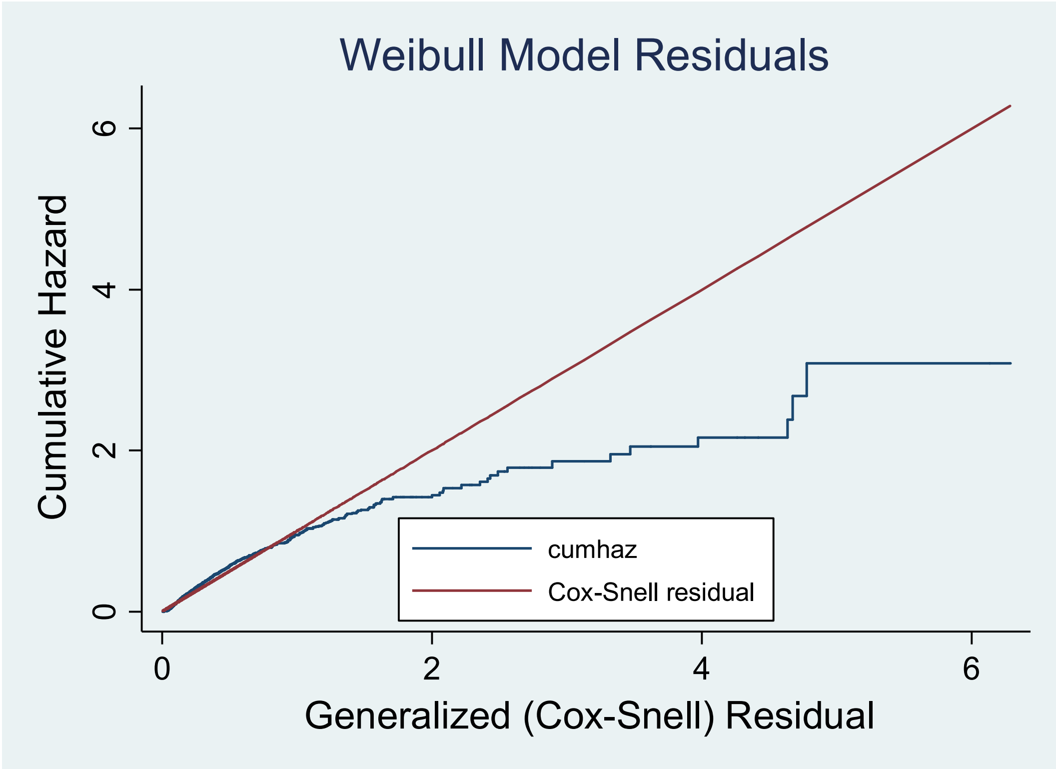


Illustration 6 – Question 5

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

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 Weibull6.wmf, replace

drop resid survivor cumhaz
```

Illustration 6 – Question 5



- In both the exponential and the Weibull case the residuals are not close to the reference line

Illustration 6 – Question 6

```
stset spell, fail(censor1)
streg $xlist, nolog nohr dist(exponential) frailty(gamma) robust
estimates store bexpgamma

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("Exponential-Gamma 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 exp_gamma.wmf, replace

drop resid survivor cumhaz
```

Illustration 6 – Question 6

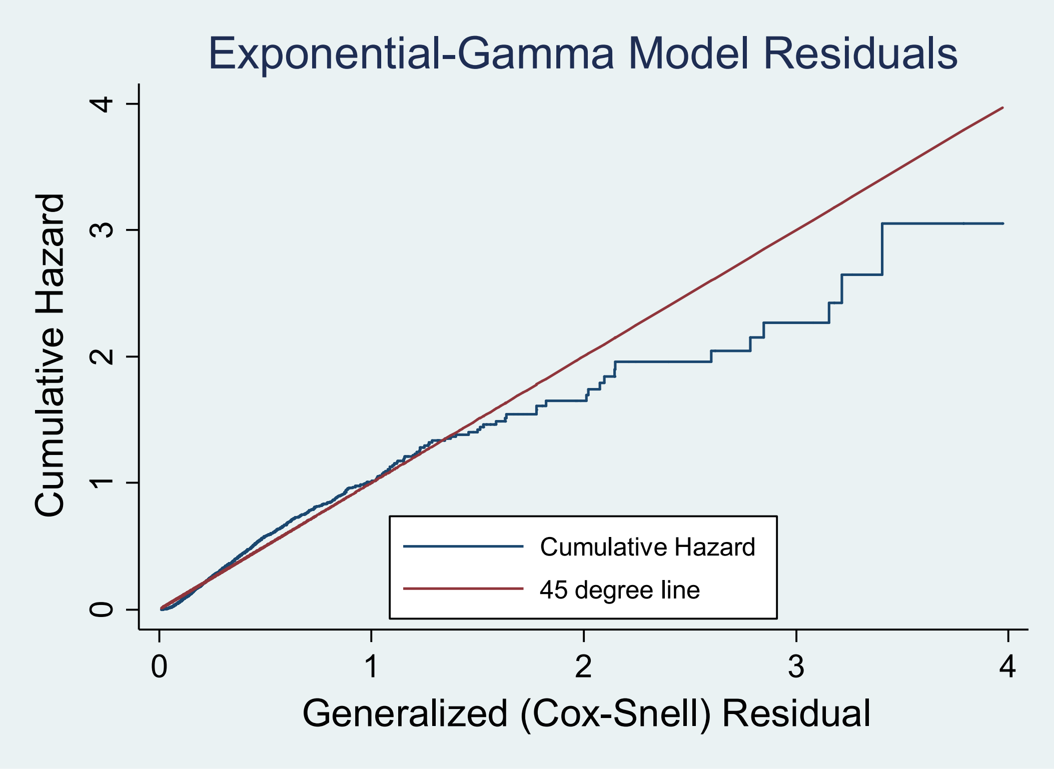
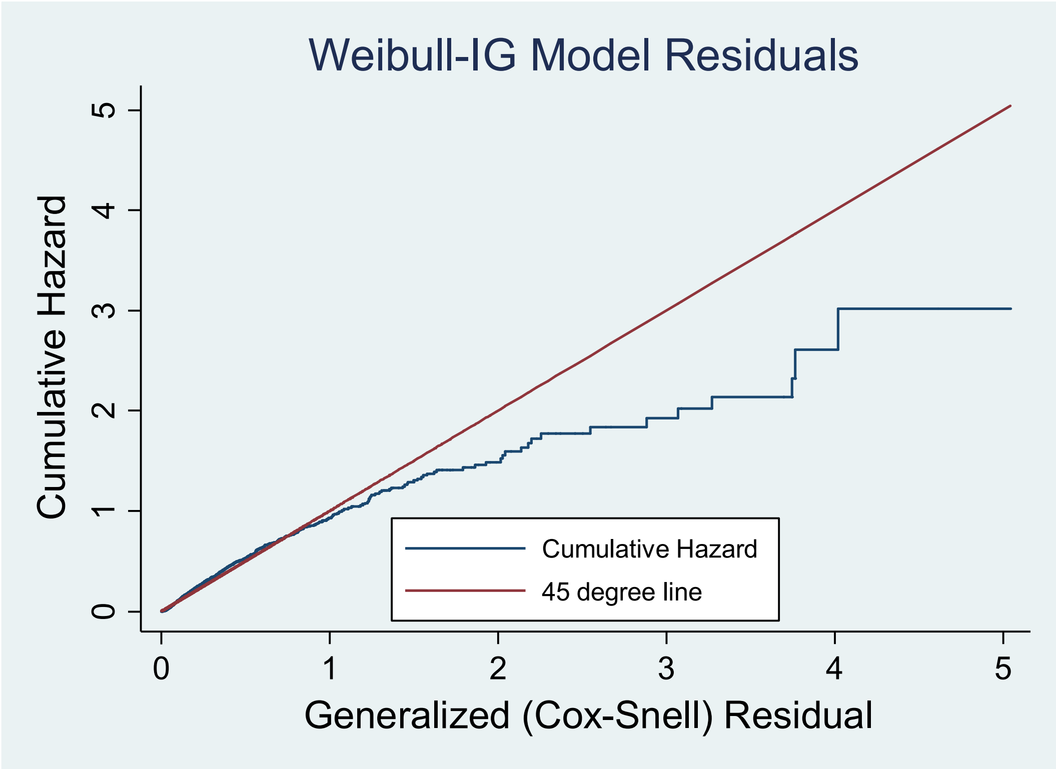


Illustration 6 – Question 6

```
stset spell, fail(censor1=1)
streg $xlist, nolog nohr dist(weibull) frailty(invgauss) robust
estimates store bweibIG

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-IG 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 Weibul16_IG.wmf, replace

drop resid survivor cumhaz
```

Illustration 6 – Question 6



- Improvements relative to baseline cases where heterogeneity is overlooked are not relevant. Mispecification is apparent. In fact, the original paper using these data uses a more flexible hazard function

Illustration 6 – Question 6

```
. estimates table bexpgamma bweibIG, b star(0.1 0.05 0.01)
----------------------------------------------
    Variable |   bexpgamma        bweibIG
-------------+--------------------------------
          RR |   .50058284       .73562769
          DR | -.88244688      -1.0725662
          UI | -1.5845375***    -2.574752***
        RRUI |  1.0911676*      1.7335706*
        DRUI |   .05740483        -.060621
     LOGWAGE |   .37928053***    .57565599***
...
-------------+--------------------------------
ln_the _cons | -1.4629948***    1.8526958***
-------------+--------------------------------
ln_p   _cons |                   .56116668***
----------------------------------------------
           legend: * p<.1; ** p<.05; *** p<.01
```

Note that p corresponds to α, which is obtained from ln(p) as exp(0.561)=1.753, σ^2 is theta and thus $\sigma^2 = \exp(1.853) = 6.379$. The duration dependence was underestimated when heterogeneity is ignored.

Illustration 7 – Question 1

```
. drop if ltotexp==.
(109 observations deleted)


. summarize ltotexp suppins totchr age female white
    Variable |        Obs        Mean    Std. Dev.       Min        Max
-------------+---------------------------------------------------------
     ltotexp |      2,955    8.059866    1.367592   1.098612   11.74094
     suppins |      2,955    .5915398    .4916322          0          1
      totchr |      2,955    1.808799    1.294613          0          7
         age |      2,955    74.24535    6.375975         65         90
      female |      2,955    .5840948    .4929608          0          1
-------------+---------------------------------------------------------
       white |      2,955    .9736041    .1603368          0          1
```

Illustration 7 – Question 1

```
. summarize ltotexp, d
                  ln(totexp) if totexp > 0
-------------------------------------------------------------
      Percentiles      Smallest
 1%      4.394449      1.098612
 5%      5.703783      1.791759
10%      6.364751      2.197225       Obs                2,955
25%      7.267525      2.639057       Sum of Wgt.        2,955

50%      8.111928                     Mean            8.059866
                        Largest       Std. Dev.       1.367592
75%      8.922258      11.56003
90%      9.772752      11.59225       Variance        1.870308
95%      10.23785      11.72489       Skewness       -.3857887
99%      11.05072      11.74094       Kurtosis        3.842263
```

Illustration 7 – Question 1

```
. quantile ltotexp
```

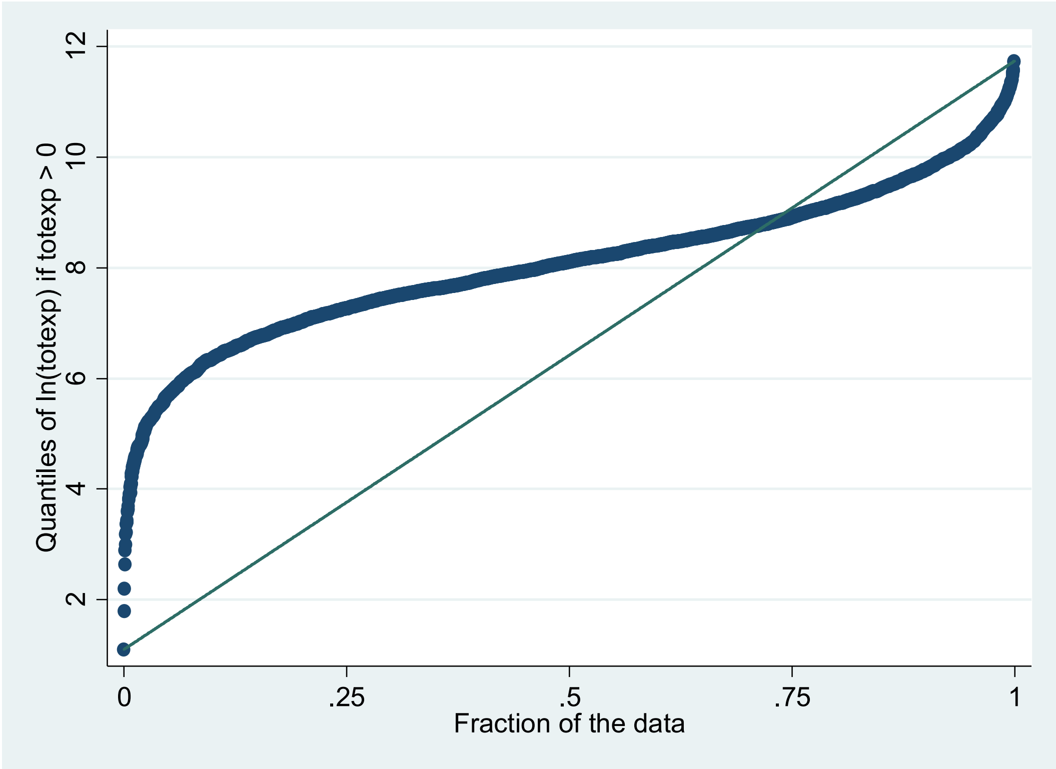


Illustration 7 – Question 2

```
. qreg ltotexp suppins totchr age female white, nolog

Median regression                                   Number of obs =      2,955
  Raw sum of deviations  1555.48 (about 8.111928)
  Min sum of deviations 1398.492                    Pseudo R2     =     0.1009

------------------------------------------------------------------------------
     ltotexp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     suppins |   .2769771   .0535936     5.17   0.000     .1718924    .3820617
      totchr |   .3942664   .0202472    19.47   0.000     .3545663    .4339664
         age |   .0148666   .0041479     3.58   0.000     .0067335    .0229996
      female |  -.0880967   .0532006    -1.66   0.098    -.1924109    .0162175
       white |   .4987457   .1630984     3.06   0.002     .1789474     .818544
       _cons |   5.648891    .341166    16.56   0.000     4.979943    6.317838
------------------------------------------------------------------------------
```

Illustration 7 – Question 2

```
. mat b = e(b)
. qui predict double xb
. qui gen double expxb = exp(xb)
. su expxb, mean
. mat b = r(mean) * b
. mat li b, ti("Marginal effects ($) on total medical expenditures")

b[1,6]:  Marginal effects ($) on total medical expenditures
        suppins      totchr         age      female       white       _cons
y1     1037.755   1477.2049   55.700813  -330.07346   1868.6593     21164.8
```

- Interpretation of some effects on the conditional median of expenditures
  - An additional chronic medical condition increases the conditional median of expenditures on 1477,20$
  - A individial with private insurante presents a conditional median of expenditures 1037,755$ higher than na individual that does not possess insurance

Illustration 7 – Question 3

```
. quietly regress ltotexp suppins totchr age female white
. estimates store OLS

. quietly qreg ltotexp suppins totchr age female white, quantile(.25)
. estimates store Q25

. quietly qreg ltotexp suppins totchr age female white, quantile(.50)
. estimates store Q50

. quietly qreg ltotexp suppins totchr age female white, quantile(.75)
. estimates store Q75

. set seed 10101
. quietly bsqreg ltotexp suppins totchr age female white, quantile(.50) reps(400)
. estimates store bsQ50
```

Illustration 7 – Question 3

```
. estimates table OLS Q25 Q50 Q75 bsQ50, b(%7.3f) star(0.1 0.05 0.01)
```

Variable	OLS	Q25	Q50	Q75	bsQ50
suppins	0.257***	0.386***	0.277***	0.149**	0.277***
totchr	0.445***	0.459***	0.394***	0.374***	0.394***
age	0.013***	0.016***	0.015***	0.018***	0.015***
female	-0.077*	-0.016	-0.088*	-0.122**	-0.088*
white	0.318**	0.338*	0.499***	0.193	0.499**
_cons	5.898***	4.748***	5.649***	6.600***	5.649***

legend: * p<.1; ** p<.05; *** p<.01

- Individual significance is similar for *suppins*, *totchr* and *age*, but differs for *female* and *white*
- Robust individual significance for LAD is equivalent to standard case
- OLS and QR coefficients differ, even when comparing OLS with LAD (both measures of central location)
- Specially for *suppins,* impact differs across quantiles: impact is higher in low quantiles

Illustration 7 – Question 4

```
. quietly regress ltotexp suppins totchr age female white

. estat hettest suppins totchr age female white, iid

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity
         Ho: Constant variance
         Variables: suppins totchr age female white

         chi2(5)      =    71.38
         Prob > chi2  =   0.0000
```

- Homoskedasticity is rejected. The previous interpretation of partial effects would need adjustments.
- Check this result with the Machado and Santos Silva command qreg2 (needs to be previously installed)…

Illustration 7 – Question 5

```
. set seed 10101
. sqreg ltotexp suppins totchr age female white, quantile(0.25,.50,0.75)
reps(400)
(…)
Simultaneous quantile regression                    Number of obs =      2,955
  bootstrap(400) SEs                                .25 Pseudo R2 =     0.1292
                                                    .50 Pseudo R2 =     0.1009
                                                    .75 Pseudo R2 =     0.0873
------------------------------------------------------------------------------
             |              Bootstrap
     ltotexp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
q25          |
     suppins |   .3856797   .0642742     6.00   0.000     .2596529    .5117065
      totchr |    .459022   .0234579    19.57   0.000     .4130265    .5050175
         age |   .0155106   .0043944     3.53   0.000     .0068941    .0241271
      female |  -.0160694   .0581328    -0.28   0.782    -.1300543    .0979155
       white |   .3375936   .1110348     3.04   0.002       .11988    .5553072
       _cons |   4.747962   .3485751    13.62   0.000     4.064487    5.431438
-------------+----------------------------------------------------------------
(cont.)
```

Illustration 7 – Question 5

(cont.)

```
-------------+----------------------------------------------------------------
q50          |
     suppins |   .2769771   .0579685     4.78   0.000     .1633144    .3906398
     totchr  |   .3942664   .0195859    20.13   0.000      .355863    .4326698
        age  |   .0148666   .0044102     3.37   0.001     .0062192    .0235139
     female  |  -.0880967   .0554863    -1.59   0.112    -.1968926    .0206992
      white  |   .4987457   .2199888     2.27   0.023     .0673984    .9300929
      _cons  |   5.648891   .3966791    14.24   0.000     4.871095    6.426687
-------------+----------------------------------------------------------------
q75          |
     suppins |   .1488548   .0649951     2.29   0.022     .0214143    .2762952
     totchr  |   .3735364   .0228424    16.35   0.000     .3287478     .418325
        age  |   .0182506   .0049533     3.68   0.000     .0085383     .027963
     female  |  -.1219365   .0562735    -2.17   0.030    -.2322759   -.0115971
      white  |   .1931923   .2045296     0.94   0.345    -.2078428    .5942275
      _cons  |   6.599972   .4247018    15.54   0.000      5.76723    7.432714
------------------------------------------------------------------------------
```

Illustration 7 – Question 5

```
. test [q25=q50=q75]:suppins

 ( 1)  [q25]suppins - [q50]suppins = 0
 ( 2)  [q25]suppins - [q75]suppins = 0

       F(  2,  2949) =    5.28
            Prob > F =    0.0051

. test [q25=q50=q75]: totchr

 ( 1)  [q25]totchr - [q50]totchr = 0
 ( 2)  [q25]totchr - [q75]totchr = 0

       F(  2,  2949) =    6.39
            Prob > F =    0.0017
```

- Equality of coefficients is rejected in both cases, at a significance level of 5%

Illustration 7 – Question 6

```
. ssc install grqreg
checking grqreg consistency and verifying not already installed...
installing into c:\ado\plus\...
installation complete.

. qreg ltotexp suppins totchr age female white, q(.50) nolog
Median regression                                   Number of obs =      2,955
  Raw sum of deviations  1555.48 (about 8.111928)
  Min sum of deviations 1398.492                    Pseudo R2     =     0.1009
------------------------------------------------------------------------------
    ltotexp |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    suppins |   .2769771   .0535936     5.17   0.000     .1718924    .3820617
     totchr |   .3942664   .0202472    19.47   0.000     .3545663    .4339664
        age |   .0148666   .0041479     3.58   0.000     .0067335    .0229996
     female |  -.0880967   .0532006    -1.66   0.098    -.1924109    .0162175
      white |   .4987457   .1630984     3.06   0.002     .1789474     .818544
      _cons |   5.648891    .341166    16.56   0.000     4.979943    6.317838
------------------------------------------------------------------------------

.  grqreg, cons ci ols olsci reps(40)
```

Illustration 7 – Question 6

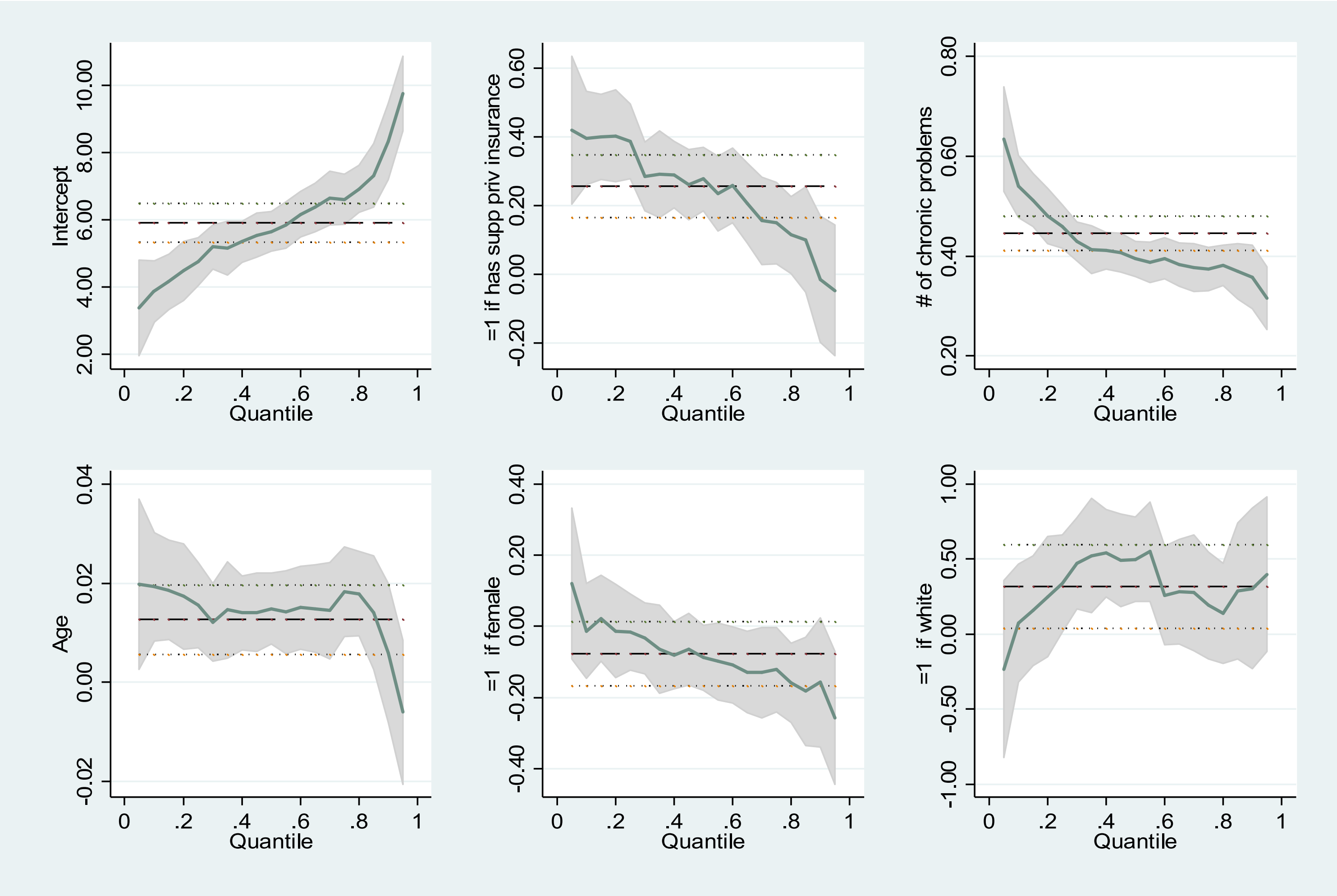


Illustration 8 – Question 1

```
. sum docvis private totchr age female white

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      docvis |      3677    6.822682    7.394937          0        144
     private |      3677    .4966005    .5000564          0          1
      totchr |      3677    1.843351    1.350026          0          8
         age |      3677    74.24476    6.376638         65         90
      female |      3677    .6010335    .4897525          0          1
-------------+--------------------------------------------------------
       white |      3677    .9709002    .1681092          0          1
```

Illustration 8 – Question 1

```
. sum docvis, d

                          docvis
-------------------------------------------------------------
      Percentiles      Smallest
 1%            0              0
 5%            0              0
10%            0              0       Obs                3677
25%            2              0       Sum of Wgt.        3677

50%            5                      Mean           6.822682
                        Largest       Std. Dev.      7.394937
75%            9             59
90%           15             73       Variance       54.68509
95%           20            106       Skewness       4.174335
99%           33            144       Kurtosis       49.67923

. set seed 10101
. generate docvisu= docvis+runiform()
. quantile docvis
. quantile docvisu
```

Illustration 8 – Question 2

```
. quantile docvis
```

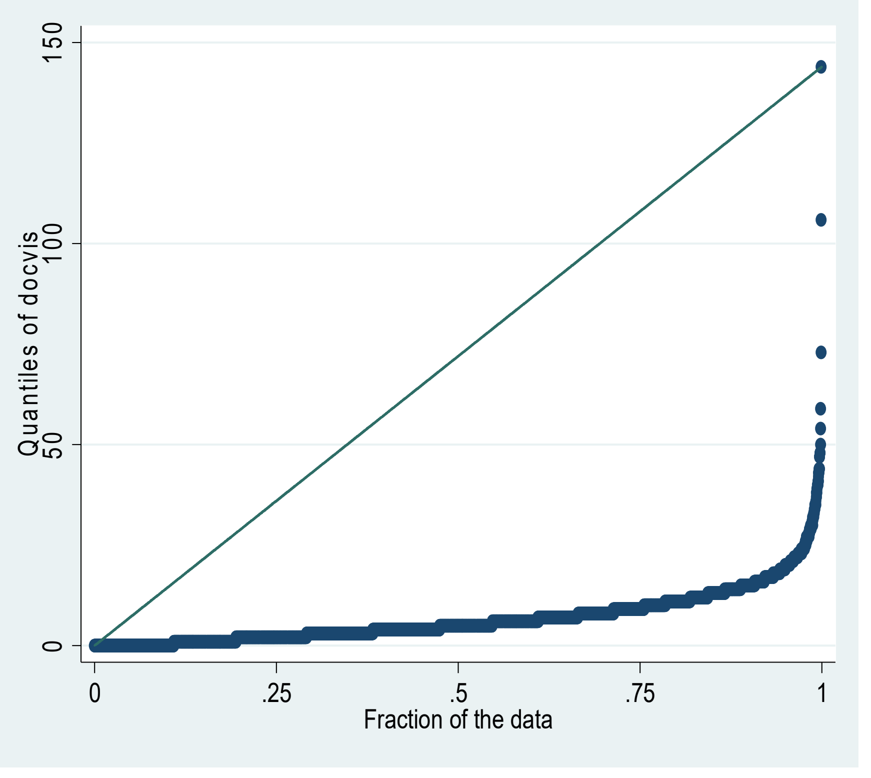
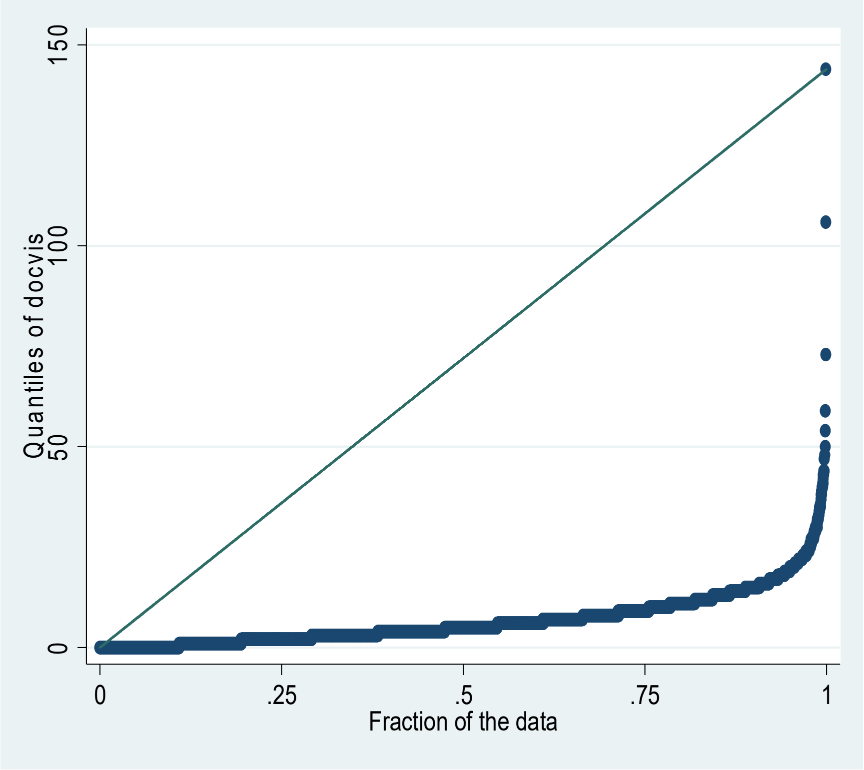


Illustration 8 – Question 2

```
. set seed 10101
. generate docvisu= docvis+runiform()
. quantile docvis
. quantile docvisu
```



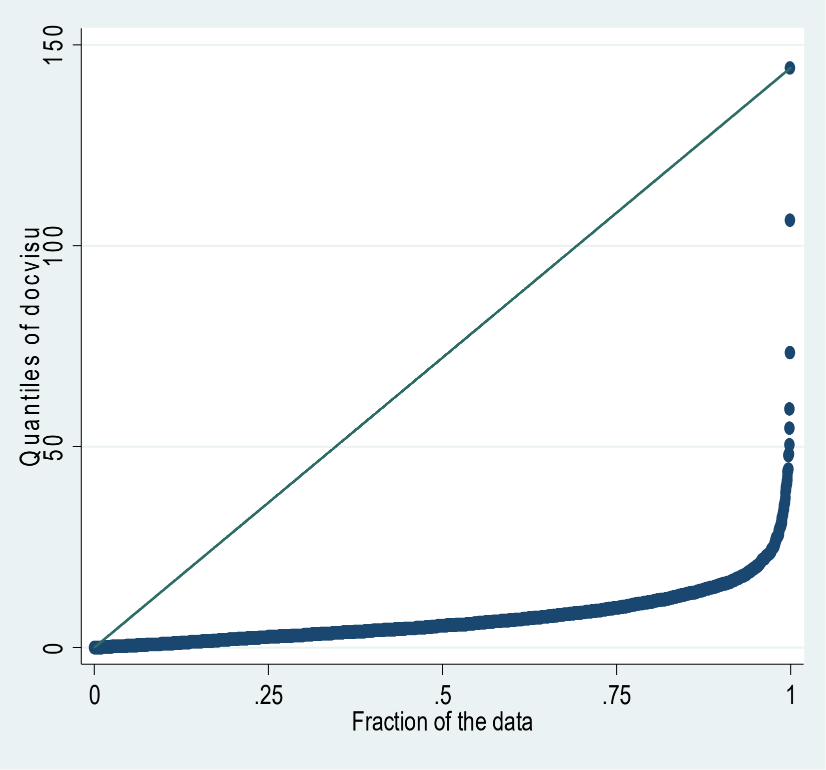


Illustration 8 – Question 3

```
. nbreg docvis private totchr age female white, vce(robust)
Negative binomial regression                       Number of obs   =       3677
Dispersion            = mean                       Wald chi2(5)    =     642.51
Log pseudolikelihood = -10634.246                  Prob > chi2     =     0.0000
------------------------------------------------------------------------------
             |               Robust
      docvis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     private |   .1721237   .0344177     5.00   0.000     .1046663    .2395811
      totchr |   .3002597   .0123475    24.32   0.000     .2760591    .3244604
         age |    .005416   .0028296     1.91   0.056    -.0001299     .010962
      female |   -.022273   .0345702    -0.64   0.519    -.0900292    .0454833
       white |   .0811637   .0981205     0.83   0.408     -.111149    .2734764
       _cons |   .7305526   .2290788     3.19   0.001     .2815665    1.179539
-------------+----------------------------------------------------------------
    /lnalpha |  -.4123734   .0387841                     -.4883889   -.3363579
-------------+----------------------------------------------------------------
       alpha |    .662077   .0256781                      .6136142    .7143674
------------------------------------------------------------------------------
```

- Only two significant variables, with the expected signs

Illustration 8 – Question 3

```
. mfx

Marginal effects after nbreg
      y  = Predicted number of events (predict)
         =  6.2779353
------------------------------------------------------------------------------
variable |      dy/dx    Std. Err.     z    P>|z|  [    95% C.I.   ]      X
---------+--------------------------------------------------------------------
 private*|   1.082549      .21481    5.04   0.000   .661523  1.50358     .4966
  totchr |   1.885011       .0771   24.45   0.000    1.7339  2.03613   1.84335
     age |   .0340016      .01767    1.92   0.054  -.000622  .068626   74.2448
  female*|  -.1401461      .21798   -0.64   0.520  -.567381  .287089   .601033
   white*|   .4905679      .57117    0.86   0.390   -.62891  1.61005     .9709
------------------------------------------------------------------------------
(*) dy/dx is for discrete change of dummy variable from 0 to 1
```

Illustration 8 – Question 4

```
. ssc install qcount
checking qcount consistency and verifying not already installed...
installing into c:\ado\stbplus\...
installation complete.

. qcount docvis private totchr age female white, q(.50) rep(500)
Count Data Quantile Regression
( Quantile 0.50 )
                                            Number of obs            =      3677
                                            No. jittered samples     =       500
------------------------------------------------------------------------------
      docvis |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
     private |   .2026718    .040978     4.95   0.000     .1223564    .2829872
      totchr |   .3465017   .0181826    19.06   0.000     .3108645     .382139
         age |   .0084275   .0033868     2.49   0.013     .0017896    .0150655
      female |   .0025244   .0413117     0.06   0.951     -.078445    .0834937
       white |   .1200437   .0980561     1.22   0.221    -.0721427    .3122302
       _cons |   .0338473   .2525817     0.13   0.893    -.4612037    .5288983
------------------------------------------------------------------------------
```

- An additional significant variable: age

Illustration 8 – Question 4

```
. qcount_mfx
 Marginal effects after qcount
        y = Qz(0.50|X)
          =  5.05847 (0.0975)
--------------------------------------------------------------------------------
                |      ME       Std. Err.      z     P>|z|   [   95% C.I   ]        X
----------------+---------------------------------------------------------------
private         |   .92609303   .18593848    4.98  0.0000   0.5617   1.2905     0.50
totchr          |   1.5795176   .07860898    20.1  0.0000   1.4254   1.7336     1.84
age             |   .03841669   .01533359    2.51  0.0122   0.0084   0.0685    74.24
female          |   .01150429   .18823153   .0611  0.9513  -0.3574   0.3804     0.60
white           |   .51745093    .4008875    1.29  0.1968  -0.2683   1.3032     0.97
--------------------------------------------------------------------------------
------------------------------------------
                | ME   [95% C. Set]     X
----------------+-------------------------
private         |  0       0   1      0.50
totchr          |  1       1   1      1.84
age             |  0       0   0     74.24
female          |  0      -1   0      0.60
white           |  0      -1   1      0.97
------------------------------------------
```

- Over Qz, similar to those of BN. In the original scale, only *totchr* imposes variation on the count

Illustration 8 – Question 5

```
. quietly qcount docvis private totchr age female white, q(.25) rep(500)
. qcount_mfx
 Marginal effects after qcount
       y = Qz(0.25|X)
         =  2.47577 (0.0666)
------------------------------------------------------------------------------
                |       ME      Std. Err.       z     P>|z|  [   95% C.I   ]        X
----------------+-------------------------------------------------------------
private         |  .76479684   .1149431    6.65  0.0000  0.5395  0.9901     0.50
totchr          |  1.0358285  .03800236    27.3  0.0000  0.9613  1.1103     1.84
age             |  .02386525  .00758687    3.15  0.0017  0.0090  0.0387    74.24
female          |  .15941462  .12641299    1.26  0.2073 -0.0884  0.4072     0.60
white           | -.14911412  .21579416   -.691  0.4896 -0.5721  0.2738     0.97
------------------------------------------------------------------------------

------------------------------------------
                | ME   [95% C. Set]     X
----------------+-------------------------
private         |  1        1  1        0.50
totchr          |  1        1  1        1.84
age             |  0        0  0       74.24
female          |  0        0  0        0.60
white           |  0       -1   0       0.97
------------------------------------------
```

Illustration 8 – Question 5

```
. quietly qcount docvis private totchr age female white, q(.75) rep(500)
. qcount_mfx
 Marginal effects after qcount
       y = Qz(0.75|X)
         =  9.06561 (0.1600)
```

	ME	Std. Err.	z	P>|z|	[95% C.I]		X
private	1.2253771	.33163708	3.69	0.0002	0.5754	1.8754	0.50
totchr	2.3236381	.13395758	17.3	0.0000	2.0611	2.5862	1.84
age	.0264796	.02548337	1.04	0.2988	-0.0235	0.0764	74.24
female	-.00437375	.32837804	-.0133	0.9894	-0.6480	0.6392	0.60
white	1.1880344	.81541512	1.46	0.1451	-0.4102	2.7862	0.97

	ME	[95% C. Set]		X
private	1	0	1	0.50
totchr	2	2	2	1.84
age	0	0	0	74.24
female	0	-1	0	0.60
white	1	-1	2	0.97

Illustration 8 – Question 5

- Note that standard errors decrease as we go from q=0.25 to q=0.75, reflecting the fact that precision is higher where more observations concentrate. Typically, standard errors are expected to be smaller at the central location (LAD), but in this case we have lots of zeros and small counts

- Comparing the AME over the quantiles for the most significant covariates: *private* and *totchar*:

	q=0.25	q=0.50	q=0.75
private	1	0	1
totchar	1	1	2

  - Different impact at different locations