STATISTICAL SOFTWARE

We describe here statistical software for the use of the generalized gamma distribution as a framework

for parametric survival analysis. The three most commonly used statistical software packages (Stata,

SAS and S-Plus) all have programs for fitting parametric survival models. We first present the essential

commands for Stata 9, SAS 9.1 and S-Plus 7 for conventional (AFT) regression models and indicate

their limitations. To overcome the limitations of currently available software (e.g., to fit an *ammag*

model), we discuss the use of SAS PROC NLMIXED for model fitting and subsequent computation of

relative times and relative hazards, with standard errors computed by the delta method. We also provide

the key functions in S-Plus for S and R users, including a general purpose function for computation of

maximum likelihood estimates. The full data set for the application in the manuscript along with the

Stata, SAS and S-Plus programs to reproduce the analyses and figures shown in the paper can be

downloaded from the web site: *http://statepi.jhsph.edu/software*.

1. Software in Stata, SAS and S-plus for parametric models

The data types that can be handled by Stata, SAS and S-plus were described in Table 1 of the

manuscript. The major limitation of the Stata streg command is the inability to handle interval censored

observations. If the widths of the censoring intervals are relatively narrow, then interval censored

observations may be treated as events occurring at the mid-point of the interval. The major limitation of

the SAS lifereg procedure is the inability to handle left truncated observations. Only in cases where the

entry times are relatively small (i.e., well before events start to occur) can left truncation be ignored.

Although the S-plus censorReg function can handle all types of censoring and truncation, it will only fit

special cases of the GG (e.g., log normal). However, the user contributed S-Plus packages GFCURE [1]

and ACCFLF (freely downloadable from StatLib at *http://lib.stat.cmu.edu/*) can be used to fit the

generalized gamma AFT model.

The key commands to implement conventional GG (log normal for S-Plus) regression are shown

in the table below (Table 3 of manuscript). In these commands, the time until event or censoring is the

variable `exit` (>0), the entry time for left truncated observations is `entry` (≥ 0), and the censoring

variable is `event` (0 = censored, 1=event).

Basic Stata, SAS, and S-PLUS programs for parametric conventional regression analyses for time-to-event data with right censoring and left truncation.

Software	Commands
Stata[*]	`Stset `*exit*`, failure(`*event*`) enter(`*entry*`)` `Streg `*covariate list*`, distribution(gamma)`
SAS[**]	`proc lifereg;` ` model `*exit*`*`*event*`(0)=`*covariate list*`/distribution=gamma;`
S-Plus[**]	`censorReg(censor(`*exit,event*`)~`*covariate list*`, distribution="lognormal",` `truncation=censor(`*entry*`, NA, 2))`

[*] The options anc(covariate list) and anc2(covariate list) allow regression models for the log scale and shape parameters.
[**] Left truncation coded in the variable *entry* is not permitted in SAS.
[***] Generalized gamma not available.

Substituting `weibull` or `lnormal` for `gamma` in Stata and SAS; and `weibull` for `log normal` in S-

Plus will fit Weibull and log normal models, respectively. For non-conventional GG regression models,

Stata has the options *anc* and *anc2* which allow separate models for the logarithm of the scale parameter

and the shape parameter, respectively. For additional discussion of the application of parametric models

in survival analysis using Stata and SAS see [2] and [3], respectively.

The following listing contains selected observations from the data set used in the application.

The listing includes each of the data types (data with and without right censoring along with data with

and without left truncation) occurring in each of the four eras of therapy, with events preceding censored

observations. Left truncation occurs only in the third and fourth periods. The variables `p1`-`p4` are

indicator variables for the four periods and the variable `period` takes values 1 to 4. In these data, the

entry and exit times are measured in years. Note that the final two records in the listing were contributed

by the same individual, who entered period three after an initial AIDS diagnosis (late entry), was

censored at the end of period three at 1.335 years after diagnosis, then entered period four (late), and

was finally censored at the end of period four.

Obs	publicID	entry	exit	event	p1	p2	p3	p4	period
534	5083	0.000	1.914	1	1	0	0	0	1
619	7950	0.000	3.295	0	1	0	0	0	1
1282	3500	0.000	4.410	1	0	1	0	0	2
1016	4409	0.000	1.368	0	0	1	0	0	2
1371	8265	0.000	1.388	1	0	0	1	0	3
1625	90270	0.259	2.025	1	0	0	1	0	3
1447	2113	0.000	3.239	0	0	0	1	0	3
1464	77720	0.170	2.798	0	0	0	1	0	3
1821	29470	0.500	1.926	1	0	0	0	1	4
2029	57470	1.367	3.042	1	0	0	0	1	4
1954	75560	0.500	5.369	0	0	0	0	1	4
2004	18350	0.949	1.116	0	0	0	0	1	4
2312	99730	0.261	1.335	0	0	0	1	0	3
2313	99730	1.335	6.835	0	0	0	0	1	4

2. Software in Stata for the application in section 3 of the manuscript

The following edited listing contains the Stata commands (in bold) for fitting the full, three-

parameter distribution to the data from period one.

```
. stset exit, failure(event) enter(entry)

failure event:  event != 0 & event < .
obs. time interval:  (0, exit]
enter on or after:  time entry
exit on or before:  failure


-----------------------------------------------------------------------------
    2314  total obs.
       0  exclusions
-----------------------------------------------------------------------------
    2314  obs. remaining, representing
    1053  failures in single record/single failure data
    (...)

. streg if p1 == 1, distribution(gamma)

    failure _d:  event
    analysis time _t:  exit
    enter on or after:  time entry
```

```
Iteration 0:   log likelihood = -1017.8641
 (...)
Iteration 5:   log likelihood =  -783.8584

Gamma regression -- accelerated failure-time form

No. of subjects =          633                    (...)
No. of failures =          388
Time at risk    = 684.7939983
                                                  (...)
Log likelihood  =     -783.8584


--------------------------------------
         _t |      Coef.   Std. Err.    (...)
------------+-------------------------
      _cons |   .6736712    .0716252    (...)
------------+-------------------------
    /ln_sig |  -.2572981    .0878631    (...)
     /kappa |   1.368894    .1890939    (...)
------------+-------------------------
      sigma |   .7731377    .0679303    (...)
--------------------------------------
```

The values of `Coef.` and `Std. Err.` for `_cons`, `sigma` and `/kappa` correspond to β_1, σ_1 and λ_1 in

the saturated GG model for period 1 of Table 2 of the manuscript. The coefficient `/ln_sig` is the

logarithm of sigma with a standard error obtained by the delta method. For the other three periods, the

variable p1 in the **streg** command must be replaced by p2, p3 and p4. The values of the log likelihood

for periods 2, 3 and 4 were -802.16, -358.51, and -306.24. The sum for all four periods equals -2250.77,

giving a value for -2 log likelihood of 4501.5 for the saturated model (shown in Table 2 of the

manuscript).

The following results are from a Stata command to fit the six-parameter conventional general

gamma model to the full data from all four periods.

```
. streg p1 p2 p3 p4, distribution(gamma) noconstant
* Include the noconstant option because there is no intercept, i.e., no reference category.

        failure _d:  event
  analysis time _t:  exit
  enter on or after:  time entry
Fitting full model:

Iteration 0:   log likelihood = -3146.5895  (not concave)
 (...)
```

4

```
Iteration 7:    log likelihood = -2269.3491

Gamma regression -- accelerated failure-time form

No. of subjects =          2314                    (...)
No. of failures =          1053
Time at risk    =   4478.967999
                                                   (...)
Log likelihood  =    -2269.3491


---------------------------------------
        _t |     Coef.   Std. Err.   (...)
-----------+---------------------------
        p1 |   .642839    .0552769
        p2 |  .7937214    .0528871    (...)
        p3 |  1.961447     .096863    (...)
        p4 |  2.914392     .090706    (...)
-----------+---------------------------
    /ln_sig |  -.217642    .0524492    (...)
     /kappa |  1.270146     .109724    (...)
-----------+---------------------------
     sigma |  .8044134    .0421908    (...)
---------------------------------------
```

The `Coef.` and `Std. Err.` values for `p1-p4`, `sigma` and `/kappa` correspond to those for

β_1 to β_4, σ_1 and λ_1 in the generalized gamma model in Table 2 of the manuscript. Furthermore, the -2

log likelihood gives the value 4538.7 shown in Table 2. To test the goodness-of-fit of the conventional

model (six parameters) we compare this value to the saturated (12 parameter) model using the likelihood

ratio test. The likelihood ratio statistic is $4538.70 - 4501.53 = 37.17 \sim \chi_6^2 \; (p < 0.001)$. In GG AFT

regression models the relative times are constant and may be obtained using Stata by eliminating one of

the dummy variables (e.g., `p1`, making period one the reference category) and using the `tr` option.

`. streg p2 p3 p4, distribution(gamma) tr`

`(...)`

```
Fitting constant-only model:

Iteration 0:    log likelihood = -2740.1508   (not concave)
 (...)
Iteration 5:    log likelihood = -2579.2072

Fitting full model:

Iteration 0:    log likelihood = -2579.2072   (not concave)
 (...)
Iteration 6:    log likelihood = -2269.3491
```

```
 (...)
---------------------------------------
       _t |  Tm. Ratio   Std. Err.   (...)
-------------+-----------------------
       p2 |    1.16286    .0680693    (...)
       p3 |   3.738214    .3798404    (...)
       p4 |   9.694447    .9576734    (...)
-------------+-----------------------
(...)
```

The `Tm. Ratio` and `Std. Err.` values for `p2`, `p3` and `p4` correspond to the relative times for the

conventional *GG* model in Table 2 of the manuscript. The saturated GG model can be fit using the

single command `streg p2 p3 p4, distribution(gamma) anc(p2 p3 p4) anc2(p2 p3 p4).`

In addition to the saturated and conventional AFT general gamma models in Table 2 of the

manuscript, the conventional AFT Weibull model can also be easily obtained using Stata; a similar

option, `hr`, provides the relative hazards. With the exception of the Weibull model for period four

however, Stata can not directly provide the final general gamma models (i.e., gamma and *ammag*). This

limitation motivates the use of more general software which we provide in the next subsection.

3. Software in SAS (PROC NLMIXED) for the application in section 3 of the manuscript

In order to compute the log likelihood, relative hazards and relative times when using the general

purpose program PROC NLMIXED it is necessary to compute both the CDF and PDF, as well as the

quantile function of the GG distribution $GG(\beta, \sigma, \lambda)$. For this purpose we used supplied functions for

the standard two-parameter gamma distribution, density and quantile functions. The parameterization

used by these functions is slightly different from that we have used for $G(\beta, \sigma)$. For our

parameterization for the standard gamma, the syntax is $function('gamma', \arg, \sigma^{-2}, \sigma^2 e^{\beta})$, where the

generic term *function*, may be either *cdf*, *pdf* or *quantile*, and the generic (second) argument denoted by

arg is either time ($t > 0$) for *cdf* and *pdf*, or proportion ($0 < p < 1$) for *quantile*.

As indicated at the end of the first subsection (*Survival and density functions*) of section 2.1.1 of the manuscript, the *GG* survival function is then

$$S_{GG(\beta,\sigma,\lambda)}(t) = S_{G(\beta\lambda/\sigma,|\lambda|)}(t^{\lambda/\sigma}) = 1 - cdf('gamma', t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma}) \quad \text{if } \lambda > 0$$
$$= cdf('gamma', t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma}) \quad \text{if } \lambda < 0$$

The density function is obtained by differentiating the survival function with respect to time,

$$f_{GG(\beta,\sigma,\lambda)}(t) = \left| \frac{dt^{\lambda/\sigma}}{dt} \right| f_{G(\beta\lambda/\sigma,|\lambda|)}(t^{\lambda/\sigma}) = \frac{|\lambda|}{\sigma} t^{\lambda/\sigma-1} pdf('gamma', t^{\lambda/\sigma}, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})$$

Finally, as shown in the corresponding subsection (*Quantile function*) of section 2.1.2, the quantile function is as follows:

$$t_{GG(\beta,\sigma,\lambda)}(t) = [quantile('gamma', p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \quad \text{if } \lambda > 0$$
$$= [quantile('gamma', 1-p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \quad \text{if } \lambda < 0$$

As indicated in the equations at the beginning of section 2.1.1, these functions may be written in an equivalent way, which shows the connection with the accelerated failure time model. For example,

$$S_{GG(\beta,\sigma,\lambda)}(t) = 1 - cdf('gamma', \exp[\lambda(\log(t)-\beta)/\sigma], \lambda^{-2}, \lambda^2) \quad \text{if } \lambda > 0$$
$$= cdf('gamma', \exp[\lambda(\log(t)-\beta)/\sigma], \lambda^{-2}, \lambda^2) \quad \text{if } \lambda < 0$$

The density function is again obtained by differentiation.

$$f_{GG(\beta,\sigma,\lambda)}(t) = \frac{|\lambda|}{\sigma t} \exp[\lambda(\log(t)-\beta)/\sigma] pdf('gamma', \exp[\lambda(\log(t)-\beta)/\sigma], \lambda^{-2}, \lambda^2)$$

Finally the (log) quantile function is given by (4).

$$\log[t_{GG(\beta,\sigma,\lambda)}(t)] = \beta + \sigma g_\lambda(p) = \beta + \sigma \log[quantile('gamma', p, \lambda^{-2}, \lambda^2)]/\lambda \quad \text{if } \lambda > 0$$
$$= \beta + \sigma g_\lambda(1-p) = \beta + \sigma \log[quantile('gamma', 1-p, \lambda^{-2}, \lambda^2)]/\lambda \quad \text{if } \lambda < 0$$

The gamma density and distribution functions can be used to compute the log likelihood as a function of the three parameters and the exit time (using the censoring information), as well as the time of entry for truncated observations. For example, for the case of an event with late entry (data type 3 in

Table 1 of manuscript), the following code calculates the contribution to the likelihood represented by

`ll`.

```
* Log likelihood for uncensored observations, with late entry: log[f(t)t/S(e)];
 if (entry > 0 and event = 1 and lambda < 0) then
    ll = log(pdf('gamma', exit**(lambda/sigma),1/(lambda*lambda),lambda*lambda*exp(beta*lambda/sigma)))
                        + lambda*log(exit)/sigma +   log(-lambda) - log(sigma)
     - log(  cdf('gamma',entry**(lambda/sigma),1/(lambda*lambda),lambda*lambda*exp(beta*lambda/sigma)));
 if (entry > 0 and event = 1 and lambda > 0) then
    ll = log(pdf('gamma', exit**(lambda/sigma),1/(lambda*lambda),lambda*lambda*exp(beta*lambda/sigma)))
                        + lambda*log(exit)/sigma +   log( lambda)  - log(sigma)
     - log(1-cdf('gamma',entry**(lambda/sigma),1/(lambda*lambda),lambda*lambda*exp(beta*lambda/sigma)));
```

or, using a previously created variables `logexit = log(exit);` and `logentry = log(entry);`

```
* Log likelihood for uncensored observations, with late entry: log[f(t)t/S(e)];
 if (entry > 0 and event = 1 and  lambda < 0) then
    ll = log(     pdf('gamma',exp(lambda*(logexit -beta)/sigma),   1/(lambda*lambda), lambda*lambda) )
                        + lambda*(logexit -beta)/sigma + log(-lambda) - log(sigma)
       - log(     cdf('gamma',exp(lambda*(logentry-beta)/sigma),   1/(lambda*lambda), lambda*lambda) );
 if (entry > 0 and event = 1 and  lambda > 0) then
    ll = log(     pdf('gamma',exp(lambda*(logexit -beta)/sigma),   1/(lambda*lambda), lambda*lambda) )
                        + lambda*(logexit -beta)/sigma + log( lambda) - log(sigma)
       - log( 1 - cdf('gamma',exp(lambda*(logentry-beta)/sigma),   1/(lambda*lambda), lambda*lambda) );
```

Note that we exclude the term $-\log(t)$ from the log of the PDF. This is because both SAS and Stata also

omit this term in their programs for conventional *GG* models; the term does not depend on the

parameters and consequently affects only the value of the log likelihood.   To avoid large arguments for

the supplied functions that can result in numerical problems, the time variable should be appropriately

scaled (e.g., from days to years). Once all contributions are incorporated into the likelihood function, the

key commands in NLMIXED to obtain the MLE are:

```
proc nlmixed cov fd;
* The option cov requests the covariance matrix for the parameter estimates;
* The option fd requests numerical derivatives for the gradient and Hessian;

* Initial values for the parameters;
parms beta=0  sigma=1  lambda=0.5;

/* Compute ll, the log likelihood for the current observation, as a function
   of the parameters and independent variables, using programming statements
   with the gamma pdf and cdf functions. The input data set must contain the
   exit and entry (if any) times, or their logarithms, and the censoring
   variable. */
model exit ~ general(ll);
run;
```

The keyword `general` tells the program that the variable `ll` is the value of log likelihood function for the current observation. For this keyword, the variable to the left of the tilde is simply a place holder, and plays no role in the optimization. For this application we used numerical derivatives (the `fd` option in the `nlmixed` statement), which gave accurate results even for the complex functions involved. The scale parameter must of course be positive. This constraint can be enforced using an additional program statement (`bounds sigma > 0;`), or through a log linear model for this parameter (parameter `logsig` and an extra program statement, `sigma = exp(logsig);`) ; we did not need either of these options in our application.

Given a set of initial values in the `parms` statement, the program sums the values of the log likelihood over all observations, and attempts to minimize the negative log likelihood as a function of the parameters. When the convergence criteria are satisfied, the program estimates the covariance matrix as the inverse of the observed information (the Hessian) and reports the results. The initial values in this sample program are for the standard $GG(0,1,0.5)$; in our application we did not find that any refinement was needed, although this may not always be the case. One way to provide better initial values would be to use summary statistics such as the median and inter-quartile ratio; another would be to use one of the standard packages to fit an initial exponential or Weibull model. Alternatively one can perform an initial search by including ranges of values for the parameters in the `parms` statement.

For regression models additional parameters are required. The simplest way to use the basic program is to preserve the names of the three original parameters. For example, to fit a model with different location (beta) and scale (sigma) parameters and a common shape (lambda) parameter to data from four different groups (denoted by `period` = 1-4), the following code could be used.

```
parms b1=0 s1=1 b2=0 s2=1 b3=0 s3=1 b4=0 s4=1 lambda=0.5;

if (period = 1) then beta  = b1;
if (period = 1) then sigma = s1;
```

9

```
if (period = 2) then beta  = b2;
if (period = 2) then sigma = s2;
           (....)
```

The NLMIXED procedure also has the ability to compute general functions of the estimated

parameters, with standard errors computed by the delta method. For example, the statement `estimate`

`'log sigma' log(sigma);` estimates the log of the scale parameter, which may be more normally

distributed. In addition the procedure has a separate feature that allows functions that also depend on

time or covariates, with standard errors again computed by the delta method. The GG quantile function

is again as follows:

$$t_{GG(\beta,\sigma,\lambda)}(t) = [quantile('gamma', p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \qquad \text{if } \lambda > 0$$
$$= [quantile('gamma', 1 - p, \lambda^{-2}, \lambda^2 e^{\beta\lambda/\sigma})]^{\sigma/\lambda} \quad \text{if } \lambda < 0$$

Using the supplied gamma quantile function we computed log percentiles for a range of values of p

($0<p<1$), and values of the log hazard for a range of $t>0$ values. To do the former we created an

additional file of values of p, with the desired range and number of points (e.g., 100) sufficient for

plotting purposes and dummy values for the other variables. This file was then appended to the original

data file. In addition, a `weight` variable was defined to be one for each of the original observations and

zero for the additional data, and used with the `replicate` statement in the procedure to exclude the

additional observations from the fitting process. Additional statements were added to compute the log of

the pth percentile of the fitted *GG* distribution for each value of p in the appended data set. These

percentiles were included in a `predict` statement, which put them (`Pred`) in an output file (`results`)

together with the standard errors (`StdErrPred`). The following is the required code.

```
/* Compute ll and lqp, the log quantile for variable prob in the input data set */

model exit ~ general(ll);
replicate  weight;
predict lqp out=results;
```

To compute the log of the hazard for each value of time, we first computed both the density and the survival function, which are also saved in the output data set using the `id` statement. Note that the program can produce confidence bands for the survival function.

```
/* Compute ll and the pdf, ft, survival function (i.e, st), hazard and log hazard (i.e., loghz) for
   each value of time (exit) in the input data set */

 model exit ~ general(ll);
 replicate  weight;
 id ft st hazard;
 predict loghz out=results;
```

In our analysis models for different eras were fit independently. The output data sets were then merged and logs of both relative times and relative hazards computed by subtraction. In this case the delta method required only squaring the individual standard errors and adding them together. This step must be included in the NLMIXED program for parameters that are not independent, for example if different groups have a common shape parameter.

An NLMIXED program based on these components gave the following edited results for the data from the first period (selected by including the additional command `where p1 = 1;`).

```
                       The NLMIXED Procedure
                          Specifications
Data Set                                    (SAS data set)
Dependent Variable                          exit
Distribution for Dependent Variable         General
Optimization Technique                      Dual Quasi-Newton
Integration Method                          None

                            Dimensions
              Observations Used                      633
              Observations Not Used                    0
              Total Observations                     633
              Parameters                               3

                            Parameters
              beta        sigma       lambda      NegLogLike

                 0            1          0.5        830.99793

                         Iteration History
      Iter     Calls    NegLogLike        Diff      MaxGrad        Slope

         1         8     807.103492    23.89444     54.07252      -1224.24
           (...)
```

```
        9        54    783.858399    4.681E-7    0.000101    -9.59E-7

                   NOTE: GCONV convergence criterion satisfied.

                             Fit Statistics
                   -2 Log Likelihood              1567.7
                                   (...)

                           Parameter Estimates
                     Standard
Parameter   Estimate    Error    (...)

beta         0.6737    0.07156   (...)
sigma        0.7731    0.06785   (...)
lambda       1.3689    0.1889    (...)

                 Covariance Matrix of Parameter Estimates
            Row    Parameter        beta      sigma      lambda

             1      beta        0.005120   -0.00337    0.01093
             2      sigma      -0.00337    0.004603   -0.01123
             3      lambda      0.01093    -0.01123    0.03567
```

The maximum likelihood estimates of the parameters and their asymptotic standard errors are given in the listing, corresponding to β_1, σ_1 and λ_1 in the saturated general gamma model of Table 2. These results are practically identical with those produced by Stata. The following results were obtained using the `estimate` statement above, and are also very similar to those from Stata.

```
                      Additional Estimates
                Standard
 Label  Estimate    Error

logsig  -0.2573   0.08776   (...)
```

As previously indicated, the primary motivation for the use of NLMIXED is to be able to fit models not available in the standard statistical packages. Specifically, we can fit the two-parameter gamma and the *ammag* among many others, as well as allow regression in both the scale and shape parameters. Since for computational reasons the case $\lambda = 0$ is best handled separately, it seems appropriate to use the scale (σ) parameter for restricted models, and then compute a shape (λ) value that is either positive or negative, depending on the desired two-parameter distribution. Thus to fit the standard gamma and *ammag* models, the `lambda` parameter must be removed from the parameter list

and the program must contain additional statements of the form `lambda = sigma;` or `lambda = 1/sigma;` respectively. Likewise for a Weibull fit, the additional statement is `lambda=1`. For the special case of the lognormal distribution ($\lambda = 0$) the same functions are available for the normal distribution (keyword '`normal`'), and a separate program can be written. Alternatively one can simply fix `lambda` to be a small value (e.g., `lambda = 0.001;`). The key commands for the *ammag* fit to the data of period one are as follows.

```
proc nlmixed cov fd;
 parms beta=0  sigma=1;
 lambda = 1/sigma;
 /*** Compute ll, the log likelihood for the current observation ***/
 model exit ~ general(ll);
 where p1 = 1;
run;
```

For the *ammag* fit the following results were obtained.

```
                          Parameters
                  beta        sigma     NegLogLike

                     0            1     865.296989

                      Iteration History
      Iter     Calls    NegLogLike       Diff      MaxGrad       Slope

         1         6     808.559901    56.73709     81.67648    -2014.32
           (...)
         7        31     784.13085     3.821E-7     0.000119    -7.86E-7

              NOTE: GCONV convergence criterion satisfied.

                        Fit Statistics
               -2 Log Likelihood              1568.3
                           (...)

                    Parameter Estimates
                  Standard
Parameter    Estimate     Error    (...)

beta          0.6365     0.05092   (...)
sigma         0.7978     0.06052   (...)
```

The estimates and standard errors for `beta` and `sigma` correspond to β_1 and σ_1 for the final model of Table 2. The value of the likelihood ratio statistic comparing the two models is

$1568.26 - 1567.72 = 0.54 \sim \chi_1^2 \ (p = 0.46)$. To obtain the gamma model for period two, one simply needs to substitute the command `lambda = sigma;` and select the data.

The following is a partial listing of a `results` file from the estimation of log percentiles for the *ammag* fit for period one, after the actual data have been deleted. The file includes the values of the log percentiles (`Pred`), their asymptotic standard errors and 95% Wald-type confidence limits. For completeness the actual percentiles (`qp`) were also included using the `id` statement.

Obs	period	weight	prob	qp	Pred	StdErr Pred	Alpha	Lower	Upper
1	1	0	0.010	0.02264	-3.78825	0.078179	0.05	-3.94177	-3.63473
2	1	0	0.015	0.03396	-3.38257	0.077995	0.05	-3.53573	-3.22941
3	1	0	0.020	0.04529	-3.09464	0.077790	0.05	-3.24739	-2.94188
4	1	0	0.025	0.05663	-2.87120	0.077567	0.05	-3.02352	-2.71888
5	1	0	0.030	0.06798	-2.68854	0.077331	0.05	-2.84040	-2.53669
6	1	0	0.035	0.07934	-2.53403	0.077084	0.05	-2.68540	-2.38266
7	**1**	**0**	**0.040**	**0.09071**	**-2.40010**	**0.076826**	**0.05**	**-2.55096**	**-2.24923**
8	1	0	0.045	0.10209	-2.28189	0.076560	0.05	-2.43223	-2.13154
9	1	0	0.050	0.11349	-2.17607	0.076286	0.05	-2.32587	-2.02627
10	1	0	0.055	0.12490	-2.08028	0.076006	0.05	-2.22953	-1.93102

Similarly, the following lines are from the estimation of the log hazard (`Pred`) for this fit. The values of the density (`ft`), survival (`st`) and hazard (`%`) functions have also been included. Note the consistency between line 7 in the previous listing and line 3 in the following.

Obs	period	weight	exit	ft	st	hazard	Pred	StdErr Pred
1	1	0	0.01	0.44182	0.99558	44.3781	-0.81242	0.078565
2	1	0	0.05	0.44096	0.97792	45.0916	-0.79647	0.078220
3	**1**	**0**	**0.09**	**0.43955**	**0.96031**	**45.7714**	**-0.78151**	**0.077308**
4	1	0	0.13	0.43772	0.94277	46.4296	-0.76723	0.076102
5	1	0	0.17	0.43555	0.92530	47.0713	-0.75351	0.074716
6	1	0	0.21	0.43308	0.90793	47.6996	-0.74025	0.073214
7	1	0	0.25	0.43033	0.89066	48.3165	-0.72740	0.071640
8	1	0	0.29	0.42735	0.87350	48.9236	-0.71491	0.070025
9	1	0	0.33	0.42414	0.85647	49.5219	-0.70276	0.068395
10	1	0	0.37	0.42073	0.83957	50.1123	-0.69090	0.066769

4. Software in S-Plus (nlminb) for the application in section 3 of the manuscript

Here we present the key functions in S-Plus to obtain the maximum likelihood estimates for any member of the general gamma distribution. Similar to SAS, S-Plus supplies the functions `pgamma`, `dgamma` and `qgamma` for the CDF, PDF and quantile functions of the standard two-parameter gamma distribution. The parameterization used by these functions is slightly different from the one we have used for $G(\beta,\sigma)$ since S-plus uses `shape` and `rate` as the parameters, where `shape` is σ^{-2} and `rate` is $\sigma^{-2}e^{-\beta}$. Based on the fact that $T \sim GG(\beta,\sigma,\lambda)$ implies $T^{\lambda/\sigma} \sim G(\beta\lambda/\sigma,|\lambda|)$, a customized function for the survival function of the general gamma, which we denote by `sggamma`, is easily defined by the following code.

```
sggamma <- function(beta, sigma, lambda, timet)
{

s <- lambda^(-2)
r <- s * exp(( - beta * lambda)/sigma)

# consider timet below 0.00001 at 0.00001 to avoid division by zero
timet[timet<0.00001] <- 0.00001

b <- lambda/sigma
surv <- (lambda>0)*(1-pgamma(timet^b, shape=s, rate=r))
surv <- surv+(lambda<0)*pgamma(timet^b, shape=s, rate=r)
surv
}
```

The commands for the density function of the general gamma, which we denote by `dggamma`, are the same as those of `sggamma` except that the last three commands should be replaced by

```
dens <- (lambda>0)*((b*(timet^(b-1))*dgamma(timet^b, shape=s, rate=r)))
dens <- dens+(lambda<0)*((-b*(timet^(b-1))*dgamma(timet^b, shape=s, rate=r)))
dens
```

Although the hazard function can simply be calculated as the ratio of `dggamma` and `sggamma`, a function to directly compute the hazard function, which we denote by `hggamma`, will have the same commands as those for `sggamma` except that the last three commands should be replaced by

```
haz <- (lambda>0)*((b*(timet^(b-1))*dgamma(timet^b,shape=s,rate=r))/(1-pgamma(timet^b,shape=s,rate=r)))
haz <- haz+(lambda<0)*((-b*(timet^(b-1))*dgamma(timet^b,shape=s,rate=r))/pgamma(timet^b,shape=s,rate=r))
haz
```

Finally, the code for the quantile function, which we denote by `qggamma` is as follows.

```
qggamma <- function(beta, sigma, lambda, p)
{

s <- lambda^(-2)
r <- s * exp(( - beta * lambda)/sigma)

b <- lambda/sigma
quan <- (lambda>0)*((qgamma(p,shape=s, rate=r))^(1/b))
quan <- quan+(lambda<0)*((qgamma(1-p,shape=s, rate=r))^(1/b))
quan
}
```

The functions `dggamma` for the PDF and `sggamma` for the survival function can be used to

compute the -2 log likelihood, whose minimum can be obtained using the `nlminb` function of S-Plus [4].

A program to obtain the fit of the general gamma and *ammag* distributions to the data of period one is as

follows.

```
full.data<-matrix(scan("C:{INCLUDE PATH WHERE fulldata.dat IS}\\fulldata.dat"), ncol=8, byrow=T)
#fulldata.dat is file with full data with 2314 records of the form
#  entry exit event p1 p2 p3 p4 period1234

data<-full.data[full.data[,4]==1,] #select period 1
data.0<-data[data[,3]==O, ] #censored
data.1<-data[data[,3]==1, ] #uncensored

#general gamma
m2loglikgralgamma<-function(para)
{
beta <- para[1]
sigma <- para[2]
lambda <- para[3]
loglik<-0
#uncensored
loglik<-loglik+sum(log(dggamma(beta,sigma,lambda,data.1[,2])/sggamma(beta,sigma,lambda,data.1[,1])))
#censored event
loglik<-loglik+sum(log(sggamma(beta,sigma,lambda,data.0[,2])/sggamma(beta,sigma,lambda,data.0[,1])))
(-2)*loglik
}

#minimize (-2) log likelihood starting at start=
gg.mle_nlminb(start=c(0,1,0.5), obj=m2loglikgralgamma)
print(c("GENERAL GAMMA FOR PERIOD 1"))
print(c("beta=",round(gg.mle$para[1],3)))
print(c("sigma=",round(gg.mle$para[2],3)))
print(c("lambda=",round(gg.mle$para[3],3)))
print(c("-2 Log Likelihood=",round(gg.mle$objective,2)))

#ammag
m2loglikammag<-function(para)
{
beta <- para[1]
sigma <- para[2]
```

```
lambda <- 1/sigma
loglik<-0
#uncensored
loglik<-loglik+sum(log(dggamma(beta,sigma,lambda,data.1[,2])/sggamma(beta,sigma,lambda,data.1[,1])))
#censored event
loglik<-loglik+sum(log(sggamma(beta,sigma,lambda,data.0[,2])/sggamma(beta,sigma,lambda,data.0[,1])))
(-2)*loglik
}

#minimize (-2) log likelihood starting at start=
am.mle_nlminb(start=c(0,1), obj=m2loglikammag)
print(c("AMMAG FOR PERIOD 1"))
print(c("beta=",round(am.mle$para[1],3)))
print(c("sigma=",round(am.mle$para[2],3)))
print(c("lambda=1/sigma=",round(1/am.mle$para[1],3)))
print(c("-2 Log Likelihood=",round(am.mle$objective,2)))
```

The following results were obtained for period one.

```
"GENERAL GAMMA FOR PERIOD 1"
"beta=" "0.674"
"sigma=" "0.773"
"lambda=" "1.369"
"-2 Log Likelihood=" "1206.7"

"AMMAG FOR PERIOD 1"
"beta=" "0.637"
"sigma=" "0.798"
"lambda=1/sigma=" "1.571"
"-2 Log Likelihood=" "1207.25"
```

For the general gamma fit, the estimates agree with those provided by Stata and NLMIXED. The `"-2 Log Likelihood="` of 1206.7 provided by S-Plus differs by 361 from the 1567.7 provided by SAS and Stata. This is because both SAS and Stata exclude from the log likelihood the negative of the sum of the logarithms of the uncensored times to event, which for period one is 180.5.

For the *ammag*, the estimates from S-Plus fully agree with those provided by NLMIXED. Again, `"-2 Log Likelihood="` of 1207.25 provided by S-Plus differs by 361 from the 1568.26 provided by SAS. Since the excluded term does not depend on the parameters, comparisons of values of -2 log likelihood are the same in both packages. Indeed, the likelihood ratio statistic comparing the *ammag* to the general gamma using S-Plus was 1207.25-1206.7= 0.55 which agrees with 0.54 from SAS.

The functions `hggamma` and `qggamma` are useful for calculating relative hazards and relative times for any two members of the general gamma family. For standard errors, the user needs to provide the expression for the Hessian matrix in the optimization routine `nlminb`; an alternative is to use the bootstrap.

Although we do not consider them, programs to perform maximum likelihood estimation for general models specified by the user are also available in other software packages. For example, the general optimization routines `optim` in R [4] and `ml` in Stata [5] and could also be used to fit all of the models considered here.

References

1. Peng WW, Dear KBG, Denham JW. A generalized F mixture model for cure rate estimation. *Statistics in Medicine* 1998; **17**, 813-830.

2. Cleves MA, Gould WW, Gutierrez RG. *An Introduction to Survival Analysis Using Stata, Revised Edition*. College Station: Stata Press, 2004.

3. Allison PA. *Survival Analysis Using SAS: A Practical Guide.* Cary, NC:SAS Institute, 1995.

4. Venables WN, Ripley BD. *Modern Applied Statistics with S*, 4th ed. New York: Springer, 2002.

5. Gould, W., Pitblado, J., Sribney, W. *Maximum Likelihood Estimation with Stata*, 3rd. ed. College Station: Stata Press, 2006