

SC706: Longitudinal Data Analysis
Instructor: Natasha Sarkisian

Event History Analysis: Semi-parametric Models

The semi-parametric models, and specifically Cox proportional hazards model, are the most commonly used continuous-time models.

$$\text{Log}(h(t)) = a(t) + b_1X_1 + b_1X_2 + \dots + e$$

Here, $a(t)$ can be any function of time – we don't need to specify what it is. That's why it is called semi-parametric – because for our independent variables we specify parameters, but for the baseline hazard, we don't. That's similar to having dummy variables for baseline hazard in discrete-time models.

We continue to work with dropout.dta. Let's estimate Cox's model:

```
. stcox sex grd prt
      failure _d:  evt
      analysis time _t:  dur

Iteration 0:  log likelihood = -561.45327
Iteration 1:  log likelihood = -554.70159
Iteration 2:  log likelihood = -545.11669
Iteration 3:  log likelihood = -544.88744
Iteration 4:  log likelihood = -544.88694
Refining estimates:
Iteration 0:  log likelihood = -544.88694

Cox regression -- Breslow method for ties

No. of subjects =          264          Number of obs   =          264
No. of failures =          106
Time at risk    =          8086
Log likelihood  = -544.88694          LR chi2(3)       =          33.13
                                          Prob > chi2    =          0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
sex	1.351249	.2723425	1.49	0.135	.910284	2.005829
grd	1.341144	.1140381	3.45	0.001	1.135265	1.584359
prt	3.434491	.894678	4.74	0.000	2.061229	5.72267

Note that Cox model does not have an intercept – because it has the nonparametric baseline hazard, $h_0(t)$.

The coefficients are automatically displayed as hazard ratios, exponentiated coefficients interpreted the same way as odds ratios (i.e. they are multiplicative coefficients).

$$h(t)=h_0(t)*\exp(b_1X_1 + b_1X_2 + \dots)$$

The model is called proportional hazards because our coefficients (hazard ratios) do not change over time. That is, the hazard ratio for part-time versus full-time students is always 3.43, regardless of the time period, so the hazards functions of a part-time and a full-time student will be proportionate.

To get the non-exponentiated coefficients:

```
. stcox sex grd prt, nohr

      failure _d:  evt
      analysis time _t:  dur

Iteration 0:  log likelihood = -561.45327
Iteration 1:  log likelihood = -554.70159
Iteration 2:  log likelihood = -545.11669
Iteration 3:  log likelihood = -544.88744
Iteration 4:  log likelihood = -544.88694
Refining estimates:
Iteration 0:  log likelihood = -544.88694

Cox regression -- Breslow method for ties

No. of subjects =          264          Number of obs   =          264
No. of failures =           106
Time at risk   =          8086
Log likelihood = -544.88694          LR chi2(3)       =          33.13
                                          Prob > chi2    =          0.0000

-----+-----
      _t |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
      sex |   .3010295   .2015487     1.49   0.135    - .0939986   .6960576
      grd |   .2935231   .0850305     3.45   0.001     .1268665   .4601797
      prt |   1.233869   .2604979     4.74   0.000     .7233022   1.744435
-----+-----
```

Technical note: Our model specifies that it used Breslow method for ties. Cox model assumes that the hazard function is continuous, and therefore there are no tied survival times (two cases experiencing a failure simultaneously). But because we always record time imprecisely, such tied events do occur. Stata offers several alternative methods for dealing with ties; this is the most commonly used method (see Stata documentation for more options).

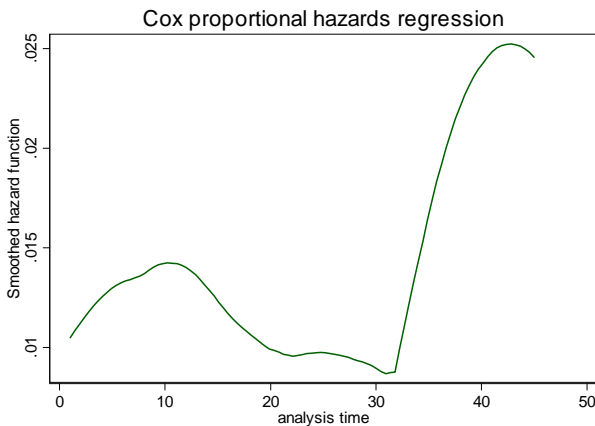
Obtaining estimates of baseline functions

Cox model doesn't estimate the actual baseline hazard, $h_0(t)$, but we can examine it after the fact, along with the survival function. To be able to obtain those graphs, we should first generate three variables using basehc, basec, and bases options:

```
. quietly stcox sex grd prt, basehc(H0) basec(C0) bases(S0)
```

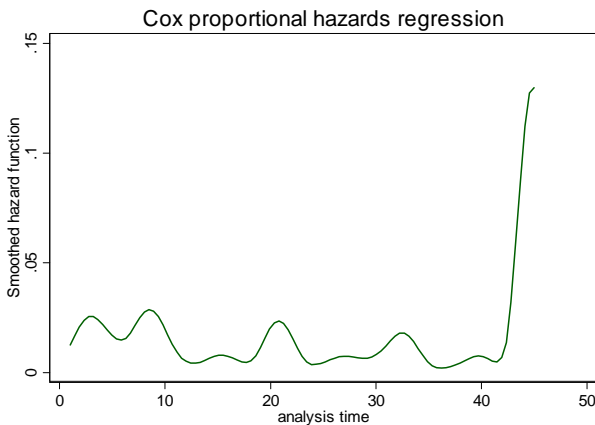
Now we can use stcurve command to generate such graphs.

```
. stcurve, hazard
```



stcurve automatically smooths the function to better see the picture by estimating the approximate density function for the hazard variable (so called kernel density estimators). This is done by assigning weights to observations based to a specified function, so the results will look different depending on the function -- see "help kdensity" for the list of possible functions (kernels). We saw what the default produces, and here's another version:

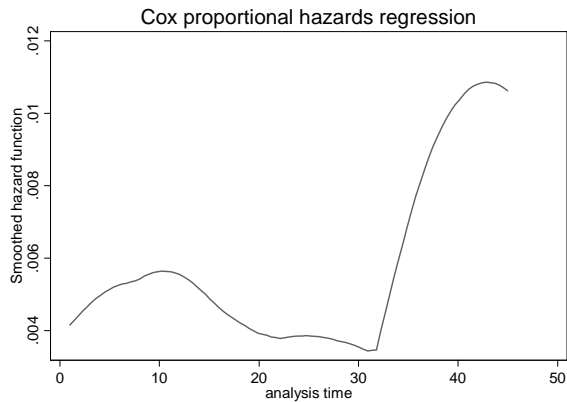
```
. stcurve, hazard kernel(cosine)
```



Here, it appears that the baseline hazard looks somewhat constant but there are a couple of large baseline hazard values at the large values of t . That's normal – for large t values, the baseline hazard will often be more variable due to the reduced effective sample size caused by prior failures and censoring.

By default, stcurve produces the functions evaluated at means of the independent variables, so if you want to get actual baseline functions, you need to specify the values of all independent variables as zero.

```
. stcurve, hazard at(prt=0 sex=0 grd=0)
```

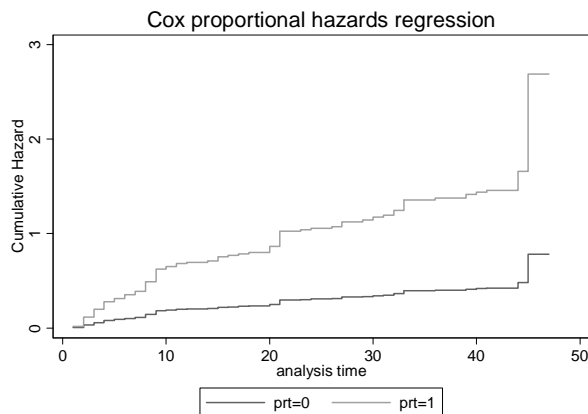


In our case, that's an unrealistic situation because our grades variable varies from 1 to 5 and never takes on a value of zero. If we wish to examine the baseline functions, we might want to recode the independent variables so that all zeroes are realistic. What zeros are will especially affect how your survival function will look like.

stcurve also allows us to examine the cumulative hazard, $H_0(t)$ and the survival function, $S_0(t)$. The cumulative hazard function is what we would think a cumulative function would be – it sums up all the hazard values up to that point in time. The survival function reflects the probability of surviving beyond time T : $S(t) = \Pr(T > t)$.

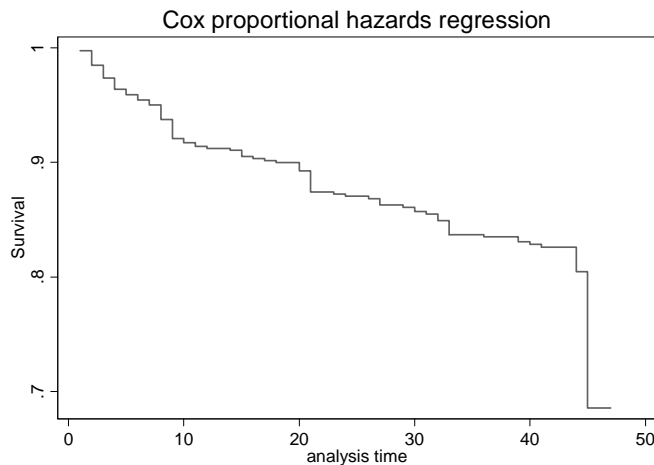
So, e.g., if we want to examine the cumulative hazard for part-time and full-time students separately:

```
. stcurve, cumhaz at(prt=0) at1(prt=1)
```



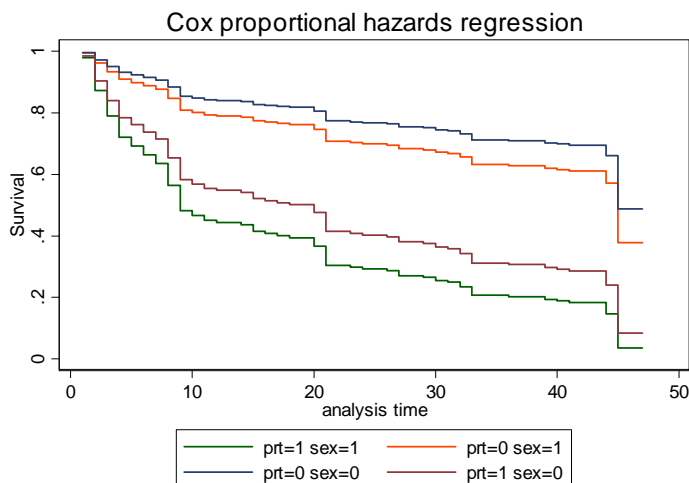
Examining the baseline hazard and cumulative hazard can be a useful diagnostic tool if one considers using parametric models as a next step. But survival functions can be very useful in interpreting the results of your Cox model itself. By selecting meaningful combinations of independent variable values, we can graph multiple survival functions for a model. Such examination of the survival functions can be very useful when interpreting the results of your analyses. To get baseline survival function:

```
. stcurve, surv at(prt=0 sex=0 grade=0)
```



And again, we can obtain survival functions for any combinations of independent variable values: $S(T|x) = S_0(t)^{\exp(Xb)}$

```
. stcurve, surv at1(prt=1 sex=1) at2(prt=0 sex=1) at3(prt=0 sex=0) at4(prt=1 sex=0)
```



Model Fit

To check the overall model fit, we can use Cox-Snell residuals. In order to obtain these, we first need to obtain another type of residuals – martingale residuals. Martingale residuals are the difference over time of the observed number of failures minus the number of failures predicted by the model.

```
. quietly stcox sex grd prt, mgale(mg)
. predict cs, csnell
(1 missing value generated)
```

If the Cox model fits the data, these Cox-Snell residuals should have an exponential distribution with a hazard rate of 1. We can verify by using Cox-Snell residuals as the time variable along with the original indicator of failure and graphing the Nelson-Aalen cumulative hazard function. If the model fits the data, then the plot of the cumulative hazard versus Cox-Snell residuals should be a straight line with slope 1 (i.e. 45 degree line).

```

. stset cs, failure(evt)

      failure event:  evt != 0 & evt < .
obs. time interval:  (0, cs]
exit on or before:  failure
-----
      265 total obs.
        1 event time missing (cs>=.)                PROBABLE ERROR
-----
      264 obs. remaining, representing
      106 failures in single record/single failure data
      106 total analysis time at risk, at risk from t =      0
                                     earliest observed entry t =      0
                                     last observed exit t = 2.637641

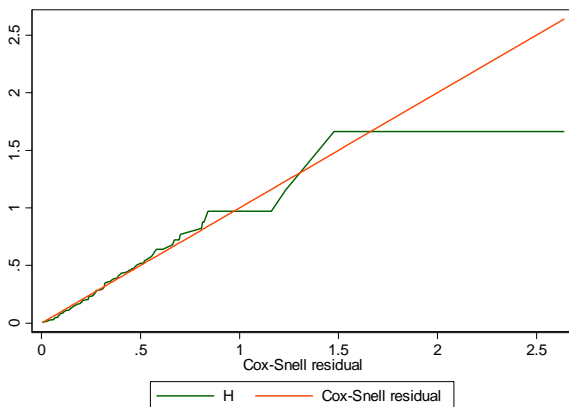
. sts gen H = na
(1 missing value generated)

```

```

. line H cs cs, sort

```



Note: We type cs twice in order to generate a 45° line. Some deviations from the line are acceptable, especially for small sample sizes and in the right-hand tail of the distribution, where the baseline hazard is more variable due to the reduced effective sample size caused by prior failures and censoring.

Here we see that the hazard function follows the 45 degree line very closely except for very large values of time. It is very common for models with censored data to stray away at large values of time and it should not cause much concern. Overall we conclude that the model fits the data well.

Another measure of model fit is the Harrell’s C concordance statistic. It indicates the ability of our model to predict who will experience the event first out of any two randomly selected individuals in the sample.

```

. estat concordance
Harrell's C concordance statistic
      failure _d:  evt
      analysis time _t:  dur

Number of subjects (N)          =      264
Number of comparison pairs (P)  =     21709
Number of orderings as expected (E) =    12774
Number of tied predictions (T)  =      2214

      Harrell's C = (E + T/2) / P =    .6394

```

Somers' D = .2788

Harrell's C ranges from 0 to 1 and indicates the proportion of all pairs where the order of event happening was identified correctly; here that is 64% of cases. We also get a Somer's D rank correlation which presents the same information as a correlation coefficients (ranging from -1 to 1).

Fitstat command does not work with stcox, so we cannot obtain BIC-based tests if we want to omit variables. We can, however, test hypotheses about joint significance of multiple variables:

```
. test sex grd
( 1) sex = 0
( 2) grd = 0
      chi2( 2) =    13.32
      Prob > chi2 =    0.0013
```

This kind of test can help you make a decision about jointly omitting a group of dummies or a set of variables that are not significant in the current model in order to make your model more parsimonious.

Diagnostics for Cox models

As is the case for OLS or logistic regression, we can examine the residuals from Cox models in order to do diagnostics. Different types of residuals that we can obtain for a Cox model are used for different diagnostic purposes.

1. Linearity.

First, we will assess the linearity assumption for our independent variables (similar to testing the linearity assumption for OLS, we will check whether we specified the appropriate functional form for our independent variables or whether we need some transformations). Here, we will use martingale residuals, but we need to obtain them for a model that omits the independent variable in question. Here, the variable we'll look at is grade – that's the only one that's acting as a continuous variable.

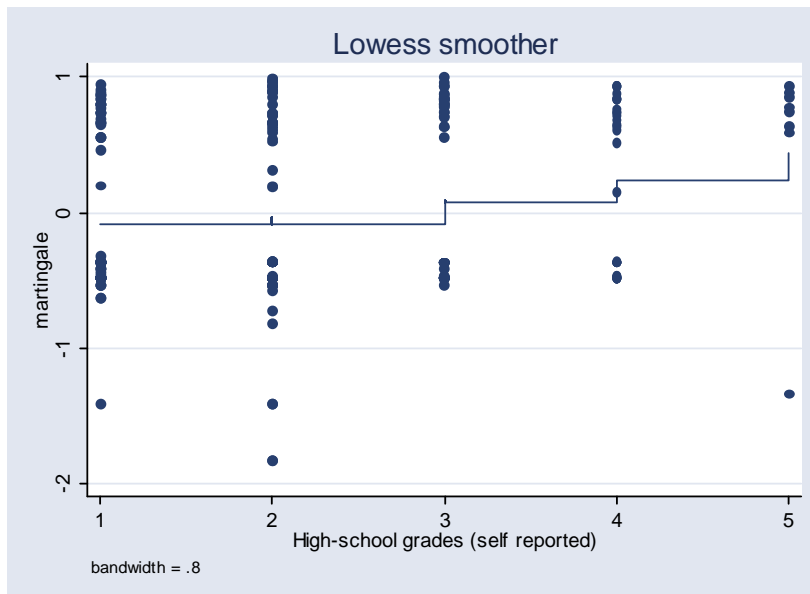
```
. stset dur, failure(evt)

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

-----
      265  total obs.
      1  obs. end on or before enter()
-----
      264  obs. remaining, representing
      106  failures in single record/single failure data
      8086 total analysis time at risk, at risk from t =          0
              earliest observed entry t =          0
              last observed exit t =          47

. quietly stcox  sex prt, mgale(mgl)

. lowess mgl grd
```



If the relationship is linear, then the mean smoother will also look approximately linear; if it doesn't, we need to consider transformations. With each transformation attempted, we try a similar operation – we transform the variable and use lowess to graph it against the martingale residuals.

2. Additivity.

To look for interactions in a Cox model, we can use fitint:

```
. fitint stcox sex grd prt, twoway(sex grd prt)
      failure _d:  evt
      analysis time _t:  dur
Iteration 0:  log likelihood = -561.45327
Iteration 1:  log likelihood = -546.85785
Iteration 2:  log likelihood = -544.62671
Iteration 3:  log likelihood = -544.60647
Iteration 4:  log likelihood = -544.60646
Refining estimates:
Iteration 0:  log likelihood = -544.60646
Cox regression -- Breslow method for ties
No. of subjects =          264                Number of obs   =          264
No. of failures =          106
Time at risk    =          8086
Log likelihood  = -544.60646                LR chi2(6)         =          33.69
                                                Prob > chi2        =          0.0000
```

_____	_____	_____	_____	_____	_____	_____
_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	
sex	1.221958	.5718396	0.43	0.668	.488337	3.057687
grd	1.330113	.1824939	2.08	0.038	1.016488	1.740504
prt	2.465719	1.873259	1.19	0.235	.5562471	10.92998
__4_4	1.013361	.1741993	0.08	0.938	.7235031	1.419344
__5_4	1.480426	.8065807	0.72	0.471	.508893	4.306723
__6_4	1.041654	.2133798	0.20	0.842	.6971997	1.556287

Fitting and testing any interactions and any main effects not included in interaction terms using the change in deviance from the full model when each term is removed in turn to obtain the likelihood ratio chi square statistic

```

-----
Model summary
Number of observations used in estimation:    264
Regression command:      cox
Dependent variable:      _t
Full model deviance:    1089.21
degrees of freedom:      6
-----

```

Term	Model deviance	Chi-square	df	P>Chi
sex*grd	1089.22	0.01	1	0.9385
sex*prt	1089.74	0.53	1	0.4678
grd*prt	1089.25	0.04	1	0.8424

No significant interactions emerged.

3. Outliers and Influential Observations.

To detect outliers, we can use martingale residuals and plot them against the predicted values:

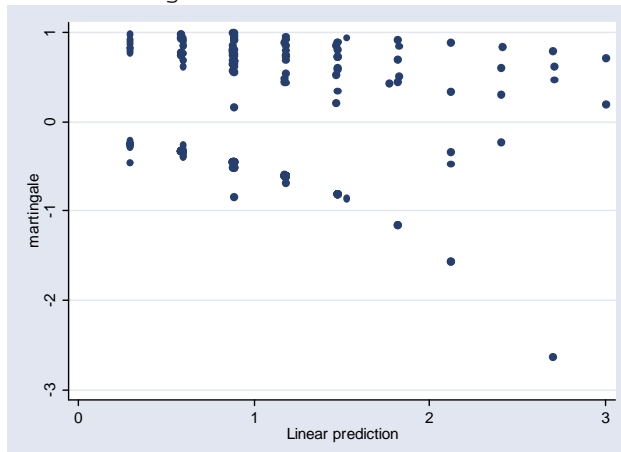
```

. quietly stcox sex grd prt, mgale(mg2)

. predict xb, xb
(1 missing value generated)

. scatter mg2 xb

```



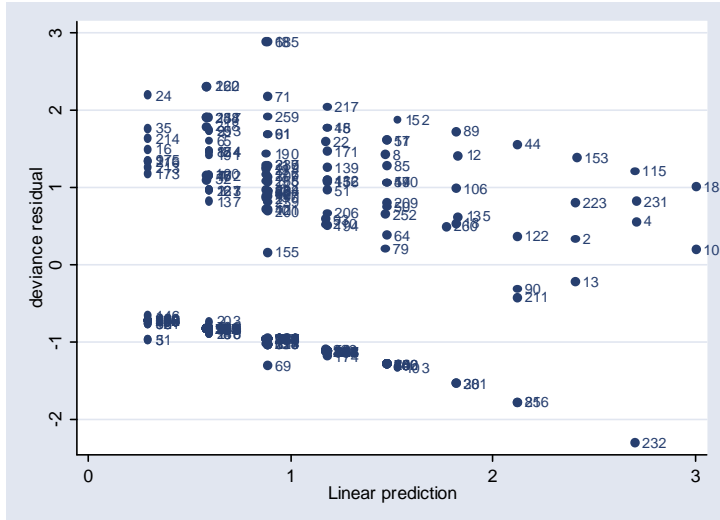
But because these residuals are skewed (they vary from minus infinity to 1), their plots can be difficult to interpret. So we use deviance residuals instead. Deviance residuals are a rescaling of martingale residuals so that they are symmetric around zero. Plot of deviance residuals against the linear predictor, survival time, rank order of survival time, or observation number can be useful in identifying outliers and assessing model fit. Note that to calculate the deviance residuals, martingale residuals have to be obtained and stored in memory.

```

. predict dev, deviance
(1 missing value generated)

```

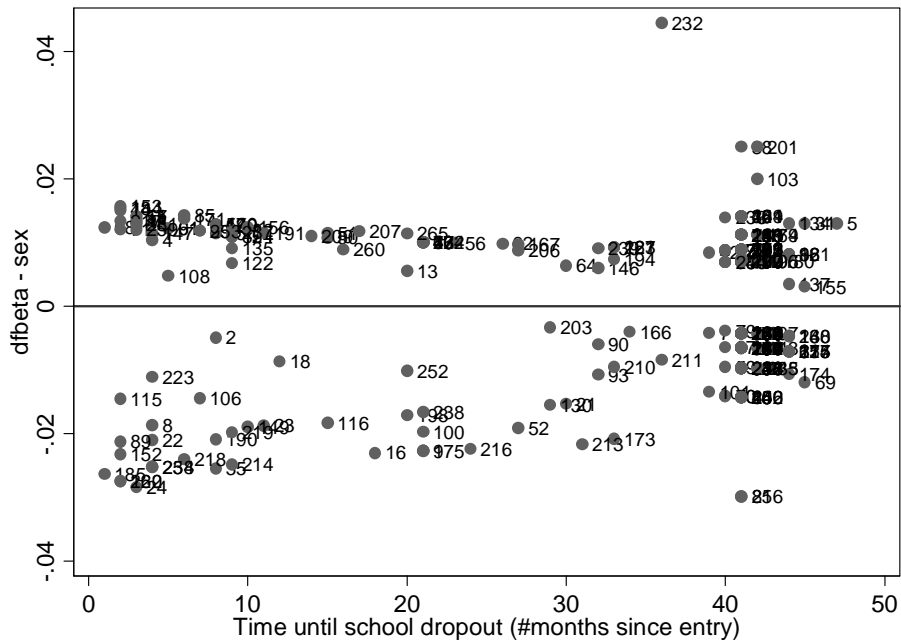
```
. scatter dev xb, mlabel(obs)
```



Deviance residuals should resemble white noise if the fit is adequate. Here, a few cases stand out as potential outliers due to their large value of residuals: #232 as well as the two on the top.

To identify influential observations (the ones that unduly influence coefficients), we can obtain `dfbetas`.

```
. predict dfbeta1 dfbeta2 dfbeta3, dfbeta
. scatter dfbeta1 dur, yline(0) mlabel(obs)
```



We would be looking for observations with unusually large `dfbetas`. For sex, it looks like case 232 might be unduly influential.

4. Error term distribution

To check whether Cox regression's assumptions about the variance of the error term are correct, we can obtain robust standard error estimates and compare them with the regular standard error estimates. If they are similar, then our results are fine. If they differ a lot, we would rather report robust standard errors as they perform better in the presence of assumption violations.

```
. stcox sex grd prt, robust
      failure _d:  evt
      analysis time _t:  dur
Iteration 0:  log pseudolikelihood = -561.45327
Iteration 1:  log pseudolikelihood = -554.70159
Iteration 2:  log pseudolikelihood = -545.11669
Iteration 3:  log pseudolikelihood = -544.88744
Iteration 4:  log pseudolikelihood = -544.88694
Refining estimates:
Iteration 0:  log pseudolikelihood = -544.88694
Cox regression -- Breslow method for ties
No. of subjects      =           264          Number of obs      =           264
No. of failures      =           106
Time at risk         =           8086

Log pseudolikelihood =   -544.88694          Wald chi2(3)      =           35.07
                                          Prob > chi2       =           0.0000
```

_t	Haz. Ratio	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
sex	1.351249	.272348	1.49	0.135	.9102767	2.005846
grd	1.341144	.1184056	3.32	0.001	1.128042	1.594504
prt	3.434491	.9379166	4.52	0.000	2.01099	5.865633

5. Assumption of proportionality.

Recall that Cox model assumes that effect of all independent variables are constant over time (e.g. the effect of being a part-time student remains the same month after month). That means that the ratio of hazards for two individuals with the same sets of independent variable values will be constant over time. This is the most important assumption that we need to test. One way to test it is to run the model with all variables specified as time-varying covariates:

```
. stcox sex grd prt, tvc(sex grd prt)

      failure _d:  evt
      analysis time _t:  dur

Iteration 0:  log likelihood = -561.45327
Iteration 1:  log likelihood = -556.94211
Iteration 2:  log likelihood = -543.91212
Iteration 3:  log likelihood = -543.85238
Iteration 4:  log likelihood = -543.85207
Iteration 5:  log likelihood = -543.85207
Refining estimates:
Iteration 0:  log likelihood = -543.85207

Cox regression -- Breslow method for ties

No. of subjects =           264          Number of obs      =           264
No. of failures =           106
Time at risk    =           8086
```

```

Log likelihood = -543.85207
LR chi2(6) = 35.20
Prob > chi2 = 0.0000

```

	_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	

rh							
	sex	1.403551	.4539459	1.05	0.295	.7446062	2.645632
	grd	1.359741	.1804695	2.32	0.021	1.048291	1.763724
	prt	5.186965	2.045356	4.17	0.000	2.394761	11.23477

t							
	sex	.9971699	.0167754	-0.17	0.866	.9648268	1.030597
	grd	.9984857	.0075134	-0.20	0.840	.9838678	1.013321
	prt	.9628441	.0281733	-1.29	0.196	.9091791	1.019677

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of _t.

Here, we examine the significance of variables in the second panel (the one labeled t). In this case, no significant interactions with time are present. We could also use some function of time instead of time itself in these interactions; ln(time) is often used:

```

. stcox sex grd prt, tvc(sex grd prt) texp(ln(_t))

      failure _d:  evt
      analysis time _t:  dur

```

```

Iteration 0:  log likelihood = -561.45327
Iteration 1:  log likelihood = -556.79532
Iteration 2:  log likelihood = -544.39802
Iteration 3:  log likelihood = -543.64391
Iteration 4:  log likelihood = -543.63962
Iteration 5:  log likelihood = -543.63962
Refining estimates:
Iteration 0:  log likelihood = -543.63962

```

Cox regression -- Breslow method for ties

```

No. of subjects =          264
No. of failures =          106
Time at risk    =          8086
LR chi2(6)      =          35.63
Prob > chi2     =          0.0000
Log likelihood  = -543.63962

```

	_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]	

rh							
	sex	1.352932	.6919672	0.59	0.555	.4965035	3.686628
	grd	1.355588	.2783309	1.48	0.138	.9064795	2.027204
	prt	7.751015	4.60011	3.45	0.001	2.422051	24.80469

t							
	sex	.995428	.2034744	-0.02	0.982	.6668323	1.485946
	grd	.9907029	.0848779	-0.11	0.913	.8375624	1.171844
	prt	.658304	.1857756	-1.48	0.138	.3786296	1.14456

Note: Second equation contains variables that continuously vary with respect to time; variables are interacted with current values of ln(_t).

Another way to test the assumption of proportionality is by examining another set of residuals – Schoenfeld residuals (as well as scaled Schoenfeld residuals for each independent variable).

```
. quietly stcox sex grd prt, schoenfeld(sch*) scaledsch(sca*)
. estat phtest, rank detail
```

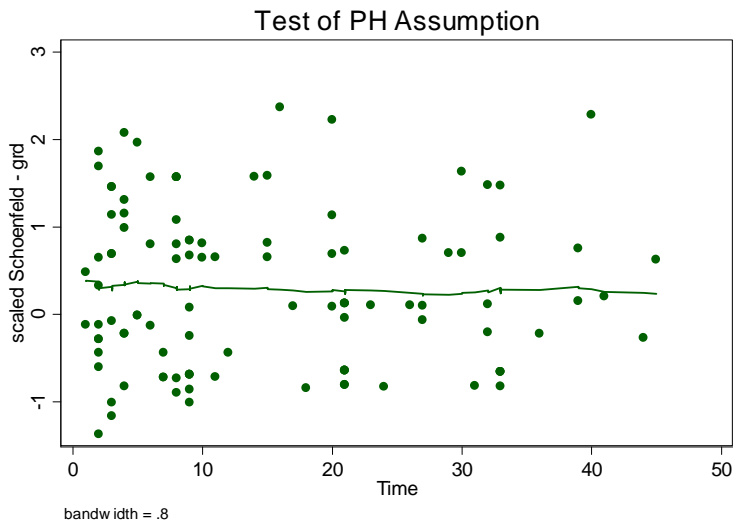
```
Test of proportional-hazards assumption
Time: Rank(t)
```

	rho	chi2	df	Prob>chi2
sex	-0.01210	0.02	1	0.9003
grd	-0.03371	0.13	1	0.7202
prt	-0.14826	2.23	1	0.1351
global test		2.56	3	0.4645

Note: the option detail provides individual variable tests; the option rank specifies that we only use the rank order of times rather than specific values; it's recommended but the results should be approximately the same with or without that option in most cases.

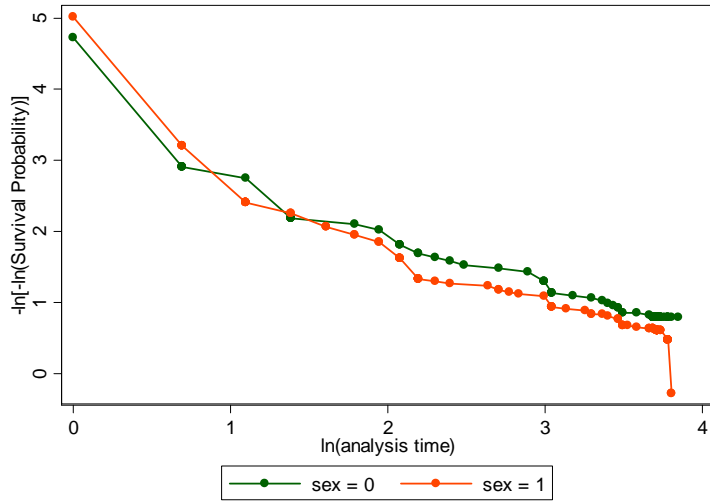
We can also obtain graphs that correspond to this test. The test checks that for each variable, the slope in the regression of residuals on time is zero.

```
. estat phtest, plot(grd)
```

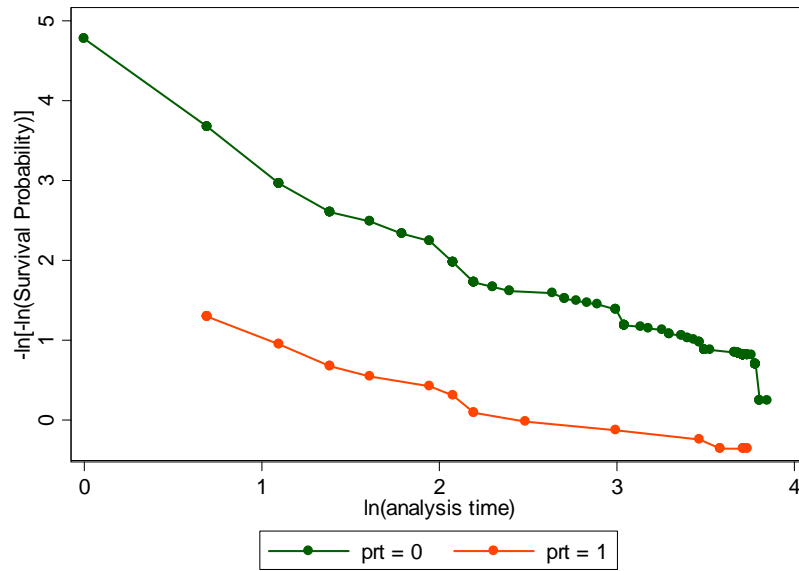


We can also test the assumption of the proportionality graphically using stphplot plots – $-\ln(-\ln(\text{survival } p))$ curves for each category of a nominal or ordinal independent variable versus $\ln(\text{analysis time})$. These are often called log-log plots. If the plotted lines are reasonable parallel, the assumption has not been violated.

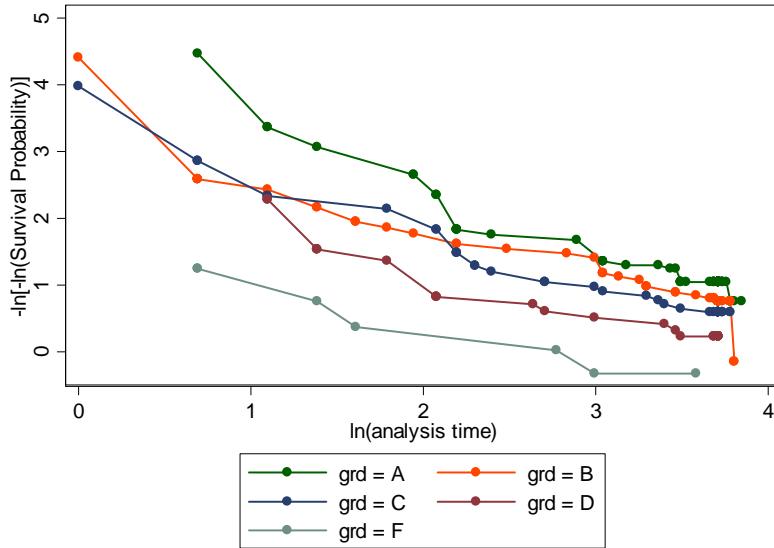
```
. stphplot, by(sex)
      failure _d: evt
      analysis time _t: dur
```



```
. stphplot, by(prt)
      failure _d: evt
      analysis time _t: dur
```



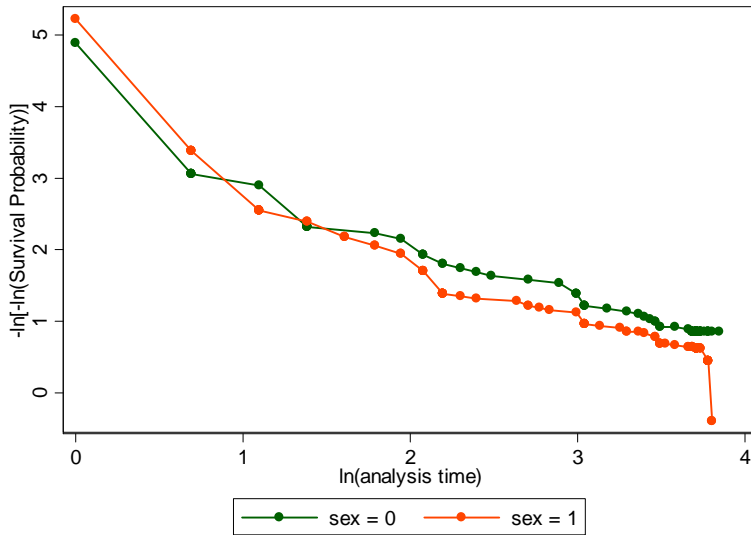
```
. stphplot, by(grd)
      failure _d: evt
      analysis time _t: dur
```



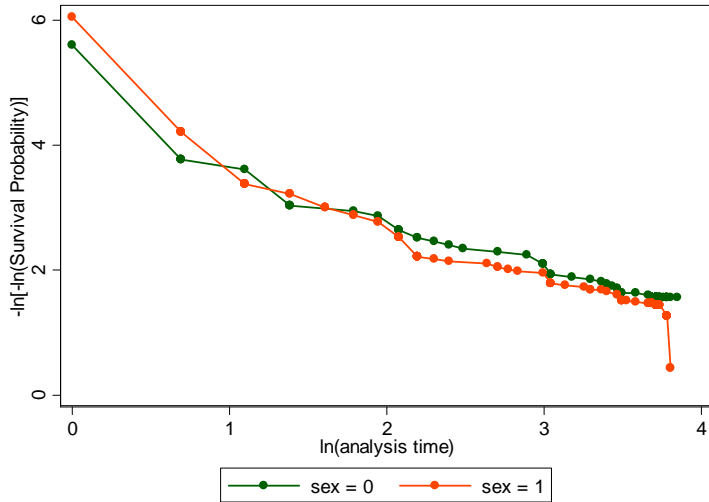
These are just estimates of survival function without other independent variables in the model. If you want to take those into account, you can either generate such plots with all other variables set to their averages, or all other variables set to zero:

```
. stphplot, by(sex) adjust( grd prt)
```

```
failure _d: evt
analysis time _t: dur
```



```
. stphplot, by(sex) adjust( grd prt) zero
failure _d: evt
analysis time _t: dur
```



Here, all the plots confirm that the assumption is reasonable. If the assumption is violated for a specific independent variable, you might want to allow the baseline hazard to vary depending on that variable. We could, of course, estimate separate models for subgroups. But that way, we allow both the baseline hazard and the individual coefficients to vary by group. If we only want to allow the baseline to vary, but think that the coefficients should be the same across groups, we use stratified Cox models. The main difference between a stratified model and an ordinary Cox model is that a stratified model allows baseline hazards to vary separately for the groups based on the strata variable. Here, the assumption is not violated, but we'll do it anyways:

```
. stcox grd prt, strata(sex)
      failure _d: evt
      analysis time _t: dur
```

```
Iteration 0: log likelihood = -490.49731
Iteration 1: log likelihood = -484.85009
Iteration 2: log likelihood = -474.94071
Iteration 3: log likelihood = -474.72113
Iteration 4: log likelihood = -474.72068
Refining estimates:
Iteration 0: log likelihood = -474.72068
```

Stratified Cox regr. -- Breslow method for ties

```
No. of subjects =          264          Number of obs   =          264
No. of failures =          106
Time at risk    =          8086
Log likelihood  = -474.72068          LR chi2(2)       =          31.55
                                          Prob > chi2    =          0.0000
```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
grd	1.33618	.1137238	3.41	0.001	1.130885 1.578744
prt	3.436641	.8954549	4.74	0.000	2.062263 5.72696

Stratified by sex

We can examine the baseline curves for this model, e.g. the survival function:

```
. stcox grd prt, strata(sex) bases(Ss)
```

```

failure_d: evt
analysis time _t: dur

```

```

Iteration 0: log likelihood = -490.49731
Iteration 1: log likelihood = -484.85009
Iteration 2: log likelihood = -474.94071
Iteration 3: log likelihood = -474.72113
Iteration 4: log likelihood = -474.72068
Refining estimates:
Iteration 0: log likelihood = -474.72068

```

Stratified Cox regr. -- Breslow method for ties

```

No. of subjects =          264          Number of obs   =          264
No. of failures =          106
Time at risk    =          8086
Log likelihood  = -474.72068
LR chi2(2)      =          31.55
Prob > chi2     =          0.0000

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
grd	1.33618	.1137238	3.41	0.001	1.130885 1.578744
prt	3.436641	.8954549	4.74	0.000	2.062263 5.72696

Stratified by sex

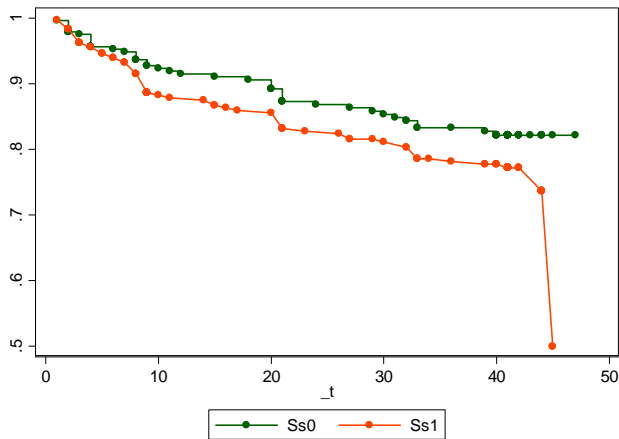
```

. gen Ss0=Ss if sex==0
(152 missing values generated)

. gen Ss1=Ss if sex==1
(114 missing values generated)

. twoway connect Ss0 Ss1 _t, c(J) sort

```



Note that such stratified model assumes that the effects of all independent variables are the same regardless of sex. We can check that assumption by using interaction terms:

```

. xi: stcox i.sex*grd i.sex*prt, strata(sex)
i.sex          _Isex_0-1          (naturally coded; _Isex_0 omitted)
i.sex*grd      _IsexXgrd_#        (coded as above)
i.sex*prt      _IsexXprt_#        (coded as above)

```

```

failure_d: evt
analysis time _t: dur

```

note: _Isex_1 dropped due to collinearity

```

Iteration 0: log likelihood = -490.49731
Iteration 1: log likelihood = -485.92293
Iteration 2: log likelihood = -474.64702
Iteration 3: log likelihood = -474.36547
Iteration 4: log likelihood = -474.36458
Iteration 5: log likelihood = -474.36458
Refining estimates:
Iteration 0: log likelihood = -474.36458

```

Stratified Cox regr. -- Breslow method for ties

```

No. of subjects =          264          Number of obs   =          264
No. of failures =          106
Time at risk   =          8086
Log likelihood = -474.36458
LR chi2(5)     =          32.27
Prob > chi2    =          0.0000

```

_t	Haz. Ratio	Std. Err.	z	P> z	[95% Conf. Interval]
grd	1.325624	.1666341	2.24	0.025	1.036149 1.695971
_IsexXgrd_1	1.033782	.1783969	0.19	0.847	.7371226 1.449835
_Isex_1	1
prt	2.739824	1.126454	2.45	0.014	1.223936 6.133193
_IsexXprt_1	1.516008	.8059959	0.78	0.434	.5347544 4.297823

Stratified by sex

Example Article:

Hill, Twyla J. 2000. Legally Extending the Family: An Event History Analysis of Grandparent Visitation Rights Laws. *Journal of Family Issues*, 21, 246-261.

Questions to answer about the article:

1. What is included in the analysis? When do cases enter and exit the sample? What constitutes a failure? Are there censored cases?
2. Does the analysis deal with repeatable or non-repeatable events? Single event type of multiple types?
3. On p. 252, the authors write, “Grandparent visitation rights laws are a subset of visitation rights legislation; most states have separate laws for parental visitation. These laws also can be considered as a sample of family laws in general. Therefore, the use of inferential statistics is appropriate.” Why do they need to justify the use of inferential statistics?
4. What are the dependent and the independent variables in this analysis? Are the independent variables time-invariant or time-variant? What do Tables 1 and 2 tell you about the variables?
5. What kind of model is used? How is time represented in this model? What is selected as the “beginning of time”? Do the effects of predictors vary by time?
6. What is presented in Figure 1? What is the relationship between this graph, survival curve, and hazard function?
7. What is reported in Table 3? How can we interpret these results? How do the authors discuss these results in the text?
8. In addition to what the authors chose to present, how else could they have presented their results?
9. What measures of model fit and model diagnostics are presented? What diagnostics and potential problems did the authors not address?