lccsm.ado -- Source Code (version 2.2)

Roseanne McNamee, Matthew Carr 12oct2013

---------- START OF PROGRAM ----------

```
**// lccsm fits a two-stage survival model to longitudinal data with a
time-to-event outcome.
**// Data should be provided in a longitudinal format that is ready to be
specified as survival-type data using Stata's stset command.


program lccsm
version 12.0
clear ado
set more off


**// Defining the survival outcome
local survivalvar="`1'"

**!! The routine assumes that 'start' and 'end' variables are provided in
the data specification !!

**// Defining the longitudinal variable
local longitvar="`2'"

**// Defining any additional covariates
local marker=0
forvalues k=3(1)12 {
   local j=`k'-2
   capture local covariate`j'="``k'''"
   if `marker'==0 & "`covariate`j'''"==""" {
      local marker=1
      local numbase=`j'-1
   }
}


**// Preserve the original data
preserve


**// Rename key variables
rename `survivalvar' survivalvar
rename `longitvar' longitvar
forvalues j=1(1)10 {
   capture rename `covariate`j''' covariate`j'
}


**// Addressing potential data sequence issues
sort id start
qui replace survivalvar=1 if survivalvar[_n+1]==1 & start==start[_n+1] &
id==id[_n+1]
qui drop if start==start[_n-1] & id==id[_n-1]
qui replace survivalvar=1 if survivalvar[_n+1]==1 & end[_n+1]<=start[_n+1]
qui drop if end<=start
```

```
**// Censor at the time of the first event
by id: gen sum=sum(survivalvar)
by id: gen sum2=sum(sum)
qui drop if sum2>1
qui drop sum*


**// Generate time variables
gen time=start
gen time2=start^2
gen time3=start^3


**// Remove individuals with <3 measurements
sort id observation
gen count=1
qui replace count=count[_n-1]+1 if id==id[_n-1]
by id: egen maxcount=max(count)
qui drop if maxcount<3
qui drop maxcount
**!! Retain the count variable for CC matching !!


**// Fit the longitudinal model
if `numbase'==0 {
   xtmixed longitvar time* || id: time time2, cov(unstructured)
   qui predict blup, fitted
}
if `numbase'>0 {
   xtmixed longitvar time* covariate1-covariate`numbase' || id: time time2,
cov(unstructured)
   qui predict blup, fitted
}


**// Calculate the standard residual (time-dependent, trend adjusted)
sort id count
gen devsq=(longitvar-blup)^2
by id: gen sum=sum(devsq)
gen var=sum/count
gen sr=var^0.5
qui drop devsq var


**// Specify as survival-type data
stset end, id(id) failure(survivalvar)


**// Case-control matching
**!! Default number of controls (n=5) !!
**!! Matching on time and the number of observations (count) !!
sttocc, match(count) n(5)


**// Prepare the matched data
qui drop start end
gen start=0
gen end=1
gen new_id=_n
qui stset end, id(new_id) failure(_case==1)
```

```
**!! Using last value (of longitvar) carried forward !!
**!! Improve using continuous time estimation of trend at case event time
points !!


**// Analyse the matched data
stcox longitvar sr, nohr


**// Restore the original data
restore


end
```

---------- END OF PROGRAM ----------