```
 1 . capture drop sample20plus
        name:  <unnamed>
         log:  E:\16GBBACKUPUSB\BACKUP_USB_SEPTEMBER2014\May Baydoun_folder\NHANES_NFL_MORTALITY_PAPER\OUTPUT\FIGURE1.smcl
    log type:  smcl
   opened on:   2 Nov 2022, 07:30:25

 2 .
 3 .
 4 .
 5 . ****************************STEP 11: SAMPLE SELECTION FLOWCHART AND 2-STAGE HECKMAN SELECTION: MONDAY************
 6 .
 7 .
 8 . **////////////GENERATE FINAL SAMPLE SELECTION VARIABLE////////////////////
 9 . use  NHANES_NFL_MORTALITY_PAPER,clear

10 .
11 . *************SAMPLE 20+*******************
12 .
13 . capture drop sample20plus

14 . gen sample20plus=.
   (10,175 missing values generated)

15 . replace sample20plus=1 if AGE>=20 & AGE~=.
   (5,769 real changes made)

16 . replace sample20plus=0 if sample20plus~=1 & AGE~=.
   (4,406 real changes made)

17 .
18 .
19 . *************FINAL SAMPLE*****************
20 . capture drop SAMPLE_FINAL

21 . gen SAMPLE_FINAL=.
   (10,175 missing values generated)

22 . replace SAMPLE_FINAL=1 if LNNFL~=.
   (2,071 real changes made)

23 . replace SAMPLE_FINAL=0 if SAMPLE_FINAL~=1
   (8,104 real changes made)

24 .
25 .
26 . tab SAMPLE_FINAL
```

SAMPLE_FINAL	Freq.	Percent	Cum.
0	**8,104**	**79.65**	**79.65**
1	**2,071**	**20.35**	**100.00**
Total	**10,175**	**100.00**	

```
27 .
28 . save, replace
   file NHANES_NFL_MORTALITY_PAPER.dta saved

29 .
30 .
31 . *************FINAL SAMPLE2, WITH COMPLETE COVARIATES*****
32 .
33 .
34 . **GET FROM MED4WAY ANALYSIS**
35 .
36 . save, replace
   file NHANES_NFL_MORTALITY_PAPER.dta saved

37 .
38 .
39 .
40 . **/////////////////CREATE INVERSE MILLS RATIO//////////////////////////////
41 .
42 .
43 . xi:probit SAMPLE_FINAL AGE SEX i.RACE_ETHN if  RIDAGEYR>=20
   i.RACE_ETHN       _IRACE_ETHN_0-3     (naturally coded; _IRACE_ETHN_0 omitted)

   Iteration 0:   log likelihood = -3766.1966
   Iteration 1:   log likelihood = -3732.2593
   Iteration 2:   log likelihood = -3732.2474
   Iteration 3:   log likelihood = -3732.2474

   Probit regression                                       Number of obs =  5,769
                                                           LR chi2(5)    =  67.90
                                                           Prob > chi2   = 0.0000
   Log likelihood = -3732.2474                             Pseudo R2     = 0.0090

   ------------------------------------------------------------------------------
    SAMPLE_FINAL | Coefficient  Std. err.      z    P>|z|     [95% conf. interval]
   -------------+----------------------------------------------------------------
             AGE |  -.0072819   .0009889    -7.36   0.000    -.0092201   -.0053436
             SEX |   .0005022    .033964     0.01   0.988     -.066066    .0670704
   _IRACE_ETHN_1 |  -.1561952   .0461589    -3.38   0.001     -.246665   -.0657254
   _IRACE_ETHN_2 |   .0113419    .044252     0.26   0.798    -.0753905    .0980743
   _IRACE_ETHN_3 |  -.0712786   .0514953    -1.38   0.166    -.1722076    .0296504
           _cons |   .0314909   .0760259     0.41   0.679    -.1175171     .180499
   ------------------------------------------------------------------------------

44 . capture drop p1

45 . predict p1, xb

46 .
47 . capture drop phi

48 . capture drop caphi

49 . capture drop invmills
```

```
50 .
51 . gen phi=(1/sqrt(2*_pi))*exp(-(p1^2/2))

52 .
53 . egen caphi=std(p1)

54 .
55 . capture drop invmills

56 . gen invmills=phi/caphi

57 .
58 . save, replace
   file NHANES_NFL_MORTALITY_PAPER.dta saved

59 .
60 .
61 . /*SURVEY COMMANDS*/
62 .
63 . sum WTSSNH2Y
```

Variable	Obs	Mean	Std. dev.	Min	Max
WTSSNH2Y	**2,085**	**101916.6**	**72591.71**	**0**	**384148**

```
64 .
65 . svyset [pweight=WTSSNH2Y], strata(SDMVSTRA) psu(SDMVPSU)

   Sampling weights: WTSSNH2Y
                VCE: linearized
       Single unit: missing
          Strata 1: SDMVSTRA
    Sampling unit 1: SDMVPSU
              FPC 1: <zero>

66 .
67 .
68 . save, replace
   file NHANES_NFL_MORTALITY_PAPER.dta saved

69 .
70 . ******************AGE AT ALL-CAUSE DEATH****
71 .
72 . capture drop AGE_DEATH

73 . gen AGE_DEATH=AGE+PERMTH_EXM/12
   (4,262 missing values generated)

74 .
75 . save, replace
   file NHANES_NFL_MORTALITY_PAPER.dta saved

76 .
77 .
78 . capture log close
```