IVEware Analysis Example Replication C10

```
* IVEware Analysis Examples Replication for ASDA 2nd Edition
* Berglund April 2017
* Chapter 10 ;

libname ncsr "P:\ASDA 2\Data sets\ncsr\" ;

*set options and location to call IVEware from SAS session ;
options set=srclib "C:\iveware_15feb2017\sas" sasautos=('!srclib' sasautos) mautosource ;
title ;
options ls=119 ps=67 ;

data c10_ncsr ;
 set ncsr.ncsr_sub_5apr2017 ;

* prepare variables for analysis ;
  if mde=1 then ageonsetmde=mde_ond ; else ageonsetmde=age ;
  intwage=age ;
  ncsrwtsh100=ncsrwtsh*100 ;
 * reverse coding for correct omitted group ;
 r_ag4cat=5-ag4cat ;
 r_mar3cat=4-mar3cat ;
 r_sex=3-sex ;
 r_ald=2-ald ;
 r_mde=2-mde ;
 r_ed4cat=5-ed4cat ;
 r_racecat=5-racecat ;

 * create series of dummy variables as alternative to reversed categorical variables ;
 if sex=1 then male=1 ; else male=0 ;
 if ed4cat=1 then ed011=1 ; else ed011=0 ;
 if ed4cat=2 then ed12=1 ; else ed12=0 ;
 if ed4cat=3 then ed1315=1 ; else ed1315=0 ;
 if ed4cat=4 then ed16=1 ; else ed16=0 ;
 if racecat=1 then other=1 ; else other=0 ;
 if racecat=2 then hispanic =1 ; else hispanic=0 ;
 if racecat=3 then black=1 ; else black=0 ;
 if racecat=4 then white=1 ; else white=0 ;
 currmar=(mar3cat=1) ;
 prevmar=(mar3cat=2) ;
 nevermar=(mar3cat=3) ;
run ;

ods rtf style=normalprinter bodytitle  ;

title "Distribution of Age of Onset of MDE or Censor" ;
proc freq ;
 tables ageonsetmde ;
run ;

ods text="Section 10.3.3 KM Example: Not Available in IVEware" ;

* Cox Model, Not currently matching SAS or Stata but very close, will use but
   update as software bugs are fixed (May 2017) ;
%regress (setup=new, name="Example 10.4.5", dir=P:\ASDA 2\Analysis Example Replication\IVEware\IVEware files) ;
title Example 10.4.5 Proportional Hazards Cox Model using NCSR data. ;
 datain c10_ncsr ;
 stratum sestrat ; cluster seclustr ; weight ncsrwtsh ;
 class sex /*r_ed4cat*/ /*r_mar3cat r_racecat*/ ;
 dependent ageonsetmde ;
 censor mde(0) ;
 predictor intwage sex ed12 ed1315 ed16 hispanic black white prevmar nevermar ;
 link phreg ;
run;

title "10.5.5 Fitting a Discrete Time Model to Complex Sample Survey Data" ;

data c10_expanded ;
 set c10_ncsr ;
* prepare data for model, shape into "long" file ;
 do pyr= 1 to intwage ;
   output ;
 end ;
run ;
```

```
data c10_expanded1 ;
 set c10_expanded ;
 if pyr=mde_ond then mdetv=1 ; else mdetv=0 ;
 if pyr <=ageonsetmde ; * select person years up to and including age of onset/censor of MDE for discrete time model ;
 if mdetv=0 then mdetv_r=2 ; else mdetv_r=1 ;
run ;

title "Print out of CASEID=1" ;
proc print ;
 where caseid=1 ;
 var caseid intwage ncsrwtsh sestrat seclustr pyr mdetv ageonsetmde ;
run ;

%regress (setup=new, name="Example 10.5.5", dir=P:\ASDA 2\Analysis Example Replication\IVEware\IVEware files) ;
title Example 10.5.5 Discrete Time Logistic Model using NCSR data. ;
 datain c10_expanded1 ;
 stratum sestrat ; cluster seclustr ; weight ncsrwtsh ;
 class sex ;
 dependent mdetv_r ;
 predictor pyr intwage sex ed12 ed1315 ed16 hispanic black white prevmar nevermar ;
 link logistic ;
run;

ods text="Discrete Time Logistic Regression with Link=CLOGLOG not available in IVEware" ;

ods rtf close ;
```

Distribution of Age of Onset of MDE or Censor

The FREQ Procedure

ageonsetmde	Frequency	Percent	Cumulative Frequency	Cumulative Percent
4	20	0.22	20	0.22
5	18	0.19	38	0.41
6	19	0.20	57	0.61
7	19	0.20	76	0.82
8	23	0.25	99	1.07
9	16	0.17	115	1.24
10	34	0.37	149	1.61
11	28	0.30	177	1.91
12	76	0.82	253	2.73
13	70	0.75	323	3.48
14	54	0.58	377	4.06
15	66	0.71	443	4.77
16	88	0.95	531	5.72
17	64	0.69	595	6.41
18	214	2.31	809	8.72
19	212	2.28	1021	11.00
20	222	2.39	1243	13.39
21	200	2.15	1443	15.55
22	195	2.10	1638	17.65
23	184	1.98	1822	19.63
24	176	1.90	1998	21.53
25	203	2.19	2201	23.71
26	159	1.71	2360	25.43
27	194	2.09	2554	27.52
28	162	1.75	2716	29.26
29	152	1.64	2868	30.90
30	233	2.51	3101	33.41
31	148	1.59	3249	35.00
32	186	2.00	3435	37.01
33	159	1.71	3594	38.72
34	180	1.94	3774	40.66
35	194	2.09	3968	42.75
36	171	1.84	4139	44.59
37	182	1.96	4321	46.55
38	215	2.32	4536	48.87
39	153	1.65	4689	50.52
40	209	2.25	4898	52.77
41	158	1.70	5056	54.47
42	194	2.09	5250	56.56
43	193	2.08	5443	58.64
44	162	1.75	5605	60.39
45	151	1.63	5756	62.01
46	126	1.36	5882	63.37
47	160	1.72	6042	65.09
48	142	1.53	6184	66.62
49	160	1.72	6344	68.35
50	150	1.62	6494	69.96
51	124	1.34	6618	71.30
52	141	1.52	6759	72.82

ageonsetmde	Frequency	Percent	Cumulative Frequency	Cumulative Percent
53	122	1.31	6881	74.13
54	115	1.24	6996	75.37
55	94	1.01	7090	76.38
56	121	1.30	7211	77.69
57	100	1.08	7311	78.77
58	112	1.21	7423	79.97
59	96	1.03	7519	81.01
60	103	1.11	7622	82.12
61	76	0.82	7698	82.93
62	82	0.88	7780	83.82
63	74	0.80	7854	84.62
64	90	0.97	7944	85.59
65	79	0.85	8023	86.44
66	70	0.75	8093	87.19
67	71	0.76	8164	87.96
68	90	0.97	8254	88.92
69	70	0.75	8324	89.68
70	64	0.69	8388	90.37
71	75	0.81	8463	91.18
72	63	0.68	8526	91.86
73	65	0.70	8591	92.56
74	75	0.81	8666	93.36
75	59	0.64	8725	94.00
76	75	0.81	8800	94.81
77	57	0.61	8857	95.42
78	64	0.69	8921	96.11
79	45	0.48	8966	96.60
80	54	0.58	9020	97.18
81	54	0.58	9074	97.76
82	32	0.34	9106	98.10
83	30	0.32	9136	98.43
84	32	0.34	9168	98.77
85	17	0.18	9185	98.95
86	19	0.20	9204	99.16
87	19	0.20	9223	99.36
88	13	0.14	9236	99.50
89	10	0.11	9246	99.61
90	15	0.16	9261	99.77
91	6	0.06	9267	99.84
92	4	0.04	9271	99.88
93	5	0.05	9276	99.94
94	2	0.02	9278	99.96
95	1	0.01	9279	99.97
98	2	0.02	9281	99.99
99	1	0.01	9282	100.00

Section 10.3.3 KM Example: Not Available in IVEware

Distribution of Age of Onset of MDE or Censor

```
IVEware Setup Checker, Wed May 10 11:53:26 2017                                  1

Setup listing:

title Example 10.4.5 Proportional Hazards Cox Model using NCSR data. ;
 datain c10_ncsr ;
 stratum sestrat ; cluster seclustr ; weight ncsrwtsh ;
 class sex   ;
 dependent ageonsetmde ;
 censor mde(0) ;
 predictor intwage sex ed12 ed1315 ed16 hispanic black white prevmar nevermar ;
 link phreg ;
run;
```

IVEware Jackknife Regression Procedure, Wed May 10 11:53:27 2017 1

Example 10.4.5 Proportional Hazards Cox Model using NCSR data.

Regression type: Proportional Hazard
Dependent variable: ageonsetmde
Censor variable: mde Major Depressive Episode 1=Yes 0=No
Predictors: intwage
SEX Sex 1=Male 2=Female
ed12
ed1315
ed16
hispanic
black
white
prevmar
nevermar
Cat. var. ref. codes: SEX 2
mde 0
Stratum variable: SESTRAT SAMPLING ERROR STRATUM
Cluster variable: SECLUSTR SAMPLING ERROR CLUSTER
Weight variable: NCSRWTSH NCSR sample part 1 weight

Valid cases 9282
Sum weights 9282.000152
Replicates 42

Degr freedom 42

Variable	Estimate	Std Error	T Test	Prob > |T|
intwage	-0.0516490	0.0017684	-29.20729	0.00000
SEX	-0.6076499	0.0458839	-13.24321	0.00000
ed12	-0.0118694	0.0605448	-0.19604	0.84552
ed1315	0.1869083	0.0731408	2.55546	0.01431
ed16	0.0951094	0.0891842	1.06644	0.29232
hispanic	-0.4258101	0.1446752	-2.94321	0.00527
black	-0.4769191	0.1290337	-3.69608	0.00063
white	-0.1427443	0.1172641	-1.21729	0.23029
prevmar	0.5518830	0.0524691	10.51825	0.00000
nevermar	-0.0632688	0.0775846	-0.81548	0.41940

Variable	Risk Ratio	95% Confidence Interval Lower	Upper
intwage	0.9496621	0.9462791	0.9530572
SEX	0.5446293	0.4964626	0.5974692
ed12	0.9882008	0.8745433	1.1166295
ed1315	1.2055167	1.0400870	1.3972586
ed16	1.0997792	0.9186304	1.3166495
hispanic	0.6532404	0.4878360	0.8747263
black	0.6206928	0.4783947	0.8053172
white	0.8669758	0.6842770	1.0984543
prevmar	1.7365197	1.5620456	1.9304819
nevermar	0.9386911	0.8026467	1.0977943

IVEware Jackknife Regression Procedure, Wed May 10 11:53:40 2017 2

Example 10.4.5 Proportional Hazards Cox Model using NCSR data.

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
intwage	0.69860	-0.0492614	-4.62273
SEX	0.84175	-0.5059537	-16.73599
ed12	0.59804	-0.0293528	147.29867
ed1315	0.87518	0.0314787	-83.15821
ed16	1.19722	-0.0857356	-190.14422
hispanic	1.31637	-0.2080602	-51.13779
black	1.08676	-0.5176008	8.53012
white	1.30890	0.0286337	-120.05947
prevmar	0.80515	0.5271150	-4.48790
nevermar	1.39187	0.1597778	-352.53804

Print out of CASEID=1

Obs	CASEID	intwage	NCSRWTSH	SESTRAT	SECLUSTR	pyr	mdetv	ageonsetmde
1	1	41	2.02426	1	2	1	0	34
2	1	41	2.02426	1	2	2	0	34
3	1	41	2.02426	1	2	3	0	34
4	1	41	2.02426	1	2	4	0	34
5	1	41	2.02426	1	2	5	0	34
6	1	41	2.02426	1	2	6	0	34
7	1	41	2.02426	1	2	7	0	34
8	1	41	2.02426	1	2	8	0	34
9	1	41	2.02426	1	2	9	0	34
10	1	41	2.02426	1	2	10	0	34
11	1	41	2.02426	1	2	11	0	34
12	1	41	2.02426	1	2	12	0	34
13	1	41	2.02426	1	2	13	0	34
14	1	41	2.02426	1	2	14	0	34
15	1	41	2.02426	1	2	15	0	34
16	1	41	2.02426	1	2	16	0	34
17	1	41	2.02426	1	2	17	0	34
18	1	41	2.02426	1	2	18	0	34
19	1	41	2.02426	1	2	19	0	34
20	1	41	2.02426	1	2	20	0	34
21	1	41	2.02426	1	2	21	0	34
22	1	41	2.02426	1	2	22	0	34
23	1	41	2.02426	1	2	23	0	34
24	1	41	2.02426	1	2	24	0	34
25	1	41	2.02426	1	2	25	0	34
26	1	41	2.02426	1	2	26	0	34
27	1	41	2.02426	1	2	27	0	34
28	1	41	2.02426	1	2	28	0	34
29	1	41	2.02426	1	2	29	0	34
30	1	41	2.02426	1	2	30	0	34
31	1	41	2.02426	1	2	31	0	34
32	1	41	2.02426	1	2	32	0	34
33	1	41	2.02426	1	2	33	0	34
34	1	41	2.02426	1	2	34	1	34

```
IVEware Setup Checker, Wed May 10 11:53:41 2017                                 1

Setup listing:

title Example 10.5.5 Discrete Time Logistic Model using NCSR data. ;
 datain c10_expanded1 ;
 stratum sestrat ; cluster seclustr ; weight ncsrwtsh ;
 class sex ;
 dependent mdetv_r ;
 predictor pyr intwage sex ed12 ed1315 ed16 hispanic black white prevmar nevermar
 ;
 link logistic ;
run;
```

```
IVEware Jackknife Regression Procedure, Wed May 10 11:53:57 2017                1

Example 10.5.5 Discrete Time Logistic Model using NCSR data.

Regression type:        Logistic
Dependent variable:     mdetv_r
Predictors:             pyr
                        intwage
                        SEX  Sex 1=Male 2=Female
                        ed12
                        ed1315
                        ed16
                        hispanic
                        black
                        white
                        prevmar
                        nevermar
Cat. var. ref. codes:   SEX  2
                        mdetv_r  2
Stratum variable:       SESTRAT  SAMPLING ERROR STRATUM
Cluster variable:       SECLUSTR  SAMPLING ERROR CLUSTER
Weight variable:        NCSRWTSH  NCSR sample part 1 weight

Valid cases             385696
Sum weights        386866.0469
Replicates                 42

Degr freedom               42

-2 LogLike         21657.91151

Variable              Estimate         Std Error            T Test        Prob > |T|
Intercept           -3.4355246         0.1630052         -21.07617           0.00000
pyr                  0.0327978         0.0020831          15.74497           0.00000
intwage             -0.0583337         0.0024527         -23.78330           0.00000
SEX                 -0.4448692         0.0633800          -7.01908           0.00000
ed12                -0.0201363         0.0670542          -0.30030           0.76543
ed1315               0.0929188         0.0579647           1.60302           0.11642
ed16                -0.0194510         0.0634037          -0.30678           0.76053
hispanic            -0.2484217         0.1349829          -1.84039           0.07278
black               -0.4569678         0.1505834          -3.03465           0.00412
white                0.0739955         0.1178912           0.62766           0.53363
prevmar              0.4942501         0.0617942           7.99832           0.00000
nevermar            -0.0353462         0.0880183          -0.40158           0.69003

Variable                   Odds             95% Confidence Interval
                          Ratio             Lower              Upper
Intercept
pyr                     1.0333416         1.0290067         1.0376946
intwage                 0.9433351         0.9386773         0.9480160
SEX                     0.6409081         0.5639583         0.7283574
ed12                    0.9800651         0.8560240         1.1220803
ed1315                  1.0973726         0.9762285         1.2335500
ed16                    0.9807370         0.8629448         1.1146079
hispanic                0.7800309         0.5940286         1.0242743
```

Example 10.5.5 Discrete Time Logistic Model using NCSR data.

Variable	Odds Ratio	95% Confidence Interval Lower	Upper
black	0.6332007	0.4672659	0.8580621
white	1.0768020	0.8488115	1.3660306
prevmar	1.6392685	1.4470753	1.8569879
nevermar	0.9652712	0.8081771	1.1529014

Variable	Design Effect	SRS Estimate	% Diff SRS v Est
Intercept	1.34068	-3.4209702	-0.42364
pyr	1.18835	0.0329478	0.45733
intwage	1.25033	-0.0577725	-0.96205
SEX	1.59559	-0.4961003	11.51599
ed12	0.72923	0.0066541	-133.04516
ed1315	0.54601	0.0782755	-15.75920
ed16	0.60420	-0.0161028	-17.21346
hispanic	1.13643	-0.2190248	-11.83347
black	1.46987	-0.5001870	9.45782
white	1.31095	0.0150557	-79.65318
prevmar	1.09388	0.5206784	5.34715
nevermar	1.81968	0.0656526	-285.74163

Discrete Time Logistic Regression with Link=CLOGLOG not available in IVEware