```
1    *** AIDS Chicago Project in Care Demo ***
2    *** Illustrates FIML missing data handling ***
3    *** for linear regression analysis ***
4    *** AFC_Demo_Part1_FIML_edits1.do :   ***
5    ***  shorten var names; recode 0/1 vars ***
6
7    version 12.1
8
9    *** Read Data File ***
10
11   use AFC_demo_compressed_V2.dta, replace
12
13   *** Describe the variables in the data set and missing data patterns ***
14
15   describe, short
16
17   summ gay_harr race sex_orn age ///
18       visit6m hiv_meds hiv_mnths ///
19       stigma hetrosxism dss1 dss2 ///
20       med_worries soc_supp
21
22   misstable summarize gay_harr race sex_orn age ///
23       visit6m hiv_meds hiv_mnths ///
24       stigma hetrosxism dss1 dss2 ///
25       med_worries soc_supp
26
27   misstable patterns gay_harr race sex_orn age ///
28       visit6m hiv_meds hiv_mnths ///
29       stigma hetrosxism dss1 dss2 ///
30       med_worries soc_supp
31
32   *** Perform the typical analyses with listwise deletion ***
33
34   * Unadjusted regressions for gay-related harrassment *
35
36   // HC3 (heteroskedastic-consistent robust standard errors - type 3)
37   // included as a just-in-case to guard against possible assumption
38   // violations
39
40   regress gay_harr age, hc3
41
42   regress gay_harr i.visit6m, hc3
43
44   // Age is significantly negatively related to harassment; doctor visits
45   // were unrelated to harassment
46
47   tab race
48   regress gay_harr i.race, hc3
49   test 2.race 3.race 4.race
50
51   tab sex_orn
52   regress gay_harr i.sex_orn, hc3
53   test 2.sex_orn 3.sex_orn 4.sex_orn
54
55   // There are significant overall differences in means for both
56   // race and sexual orientation. Pairwise comparisons showed that
57   // Blacks and Hispanics reported less harassment relative to whites.
58   // Straight-identified reported less harassment than gay-identified.
59
60   * Adjusted regression for gay-related harrassment with listwise deletion *
61
62   regress gay_harr i.race i.sex_orn age ///
63               visit6m hiv_meds hiv_mnths ///
64               stigma hetrosxism dss1 dss2 ///
65               med_worries soc_supp, hc3
66   test 2.race 3.race 4.race
67   test 2.sex_orn 3.sex_orn 4.sex_orn
68
69   // No significant overall effects. Only Black vs. white in pairwise
70   // comparisons is statitically significantly lower (p = .03)
```

```
71
72     * Adjusted regression via -sem- with listwise deletion *
73
74     // You most likely would not use -sem- with listwise deletion
75     // in a real regression analysis context; regress has more
76     // regression-specific bells and whistles. Here we include -sem-
77     // with listwise deletion as a bridge from -regress- to -sem-
78     // with FIML missing data handling.
79
80     xi: sem (gay_harr <- i.race i.sex_orn age ///
81                 visit6m hiv_meds hiv_mnths ///
82                 stigma hetrosxism dss1 dss2 ///
83                 med_worries soc_supp), vce(robust)
84     test _Irace_2 _Irace_3 _Irace_4
85     test _Isex_orn_2 _Isex_orn_3 _Isex_orn_4
86
87     // Similar results to OLS adjusted regression results from -regress-
88
89     * Adjusted regression via -sem- with FIML missing data handling *
90
91     xi: sem (gay_harr <- i.race i.sex_orn age ///
92                 visit6m hiv_meds hiv_mnths ///
93                 stigma hetrosxism dss1 dss2 ///
94                 med_worries soc_supp), ///
95                 method(mlmv) vce(robust)
96     test _Irace_2 _Irace_3 _Irace_4
97     test _Isex_orn_2 _Isex_orn_3 _Isex_orn_4
98
99     // Note new N reported on output - uses full N
100
101    // Compare results with listwise output. What is different?
102    //     1. Marginally significant overall race effect (p =.05)
103    //     2. Black (p = .007) and Hispanic (p = .031) ppts lower than whites
104    //        (only Black ppts were significantly different in the listwise analysis)
105    //     3. Significant overall sexual orientation effect (p = .02)
106    //     4. Straight (p = .005) and Bi (p .059) report less harassment
107    //        (marginal for Bi)
108
109    * What about unadjusted analyses via FIML? A little more involved *
110
111    // Essentially, this analysis generates covariances between the variables
112    // at the full N. You can get results in the correlation metric by using
113    // the -standardized- option if you want
114
115    xi: sem (<- gay_harr i.race i.sex_orn age ///
116                visit6m hiv_meds hiv_mnths ///
117                stigma hetrosxism dss1 dss2 ///
118                med_worries soc_supp), ///
119                method(mlmv) vce(robust)
120
121    // The covariances between gay harrassment and age and between
122    // gay harrassment and six month doctor visit are shown in the
123    // default -sem- output. What do we do to obtain multi-DF tests
124    // of categorical explanatory variables with more than 2 levels?
125
126    // First, replay results, but show coefficient labels so these can be
127    // copy-pasted into this do file in the -test- statements below
128
129    sem, coeflegend
130
131    // Now do the test statements for race/ethnicity and sexual identity
132    // Note how much more finicky Stata is about the syntax for -test-
133    // in this context
134
135    test (_b[cov(gay_harr,_Irace_2):_cons] = 0) ///
136         (_b[cov(gay_harr,_Irace_3):_cons] = 0) ///
137         (_b[cov(gay_harr,_Irace_4):_cons] = 0)
138
139    test (_b[cov(gay_harr,_Isex_orn_2):_cons] = 0) ///
140         (_b[cov(gay_harr,_Isex_orn_3):_cons] = 0) ///
```

```
141            (_b[cov(gay_harr,_Isex_orn_4):_cons] = 0)
142
143     // Both overall tests are significant, as was the case with the corresponding listwise
        analyses
144
145     // From the -sem- output we see that Blacks reported less harassment relative to whites, but
146     // Hispanics didn't, which is different from the listwise results above.
147
148     // Straight-identified reported less harassment than gay-identified.
149
150     // Results for age and six month doctor visits were also unchanged
151     // between the bivariate FIML analyses and the bivaritae listwise analyses
152     // Age is significantly negatively related to harassment; doctor visits
153     // were unrelated to harassment
154
155     // Replay results using standardized option to display covariances
156     // in the correlation metric. This is useful for reporting purposes
157
158     sem, standardized
159
160     // End of Part I
161
162     exit
163
```