

```

001
002 . * Fix mistakes data, clarify estimates
003 . use heck, clear
004
005 . corr eps1 eps2
006 (obs=1000)
007
008          |          eps1          eps2
009 -----+-----
010     eps1 |          1.0000
011     eps2 |          0.9362          1.0000
012
013
014 . * Mistake 1. data was generated with var(eps1-eps2)=1
015 . gen eps2x = eps1-eps2
016
017 . corr eps1 eps2x, cov
018 (obs=1000)
019
020          |          eps1          eps2x
021 -----+-----
022     eps1 |          5.53102
023     eps2x |          1.83129          .955024
024
025
026 . replace eps2x = eps2x / sqrt(.955024)
027 (1000 real changes made)
028
029 . corr eps1 eps2x, cov
030 (obs=1000)
031
032          |          eps1          eps2x
033 -----+-----
034     eps1 |          5.53102
035     eps2x |          1.87392          1
036
037
038 . corr eps1 eps2x
039 (obs=1000)
040
041          |          eps1          eps2x
042 -----+-----
043     eps1 |          1.0000
044     eps2x |          0.7968          1.0000
045
046
047 . gen mst = xb2 + eps2x
048
049 . * Mistake 2. wking was not calculated according to canonical model
050 . gen byte m = mst>0
051
052 . gen yx = cond(m,yst1,.)
053 (345 missing values generated)
054
055 . * Clarification1. Non-selected sample regressions are the TRUE params
056 . regress yst1 x2 x3 x4
057
058 Source |          SS          df          MS          Number of obs =          1000
059 -----+-----+-----+-----+-----+-----
060 Model  | 5893.06667          3 1964.35556          F( 3, 996) =          354.35
061 Residual | 5521.36743         996   5.5435416          Prob > F          =          0.0000
062 -----+-----+-----+-----+-----+-----
063 Total  | 11414.4341         999   11.42586          R-squared          =          0.5163
064                                     Adj R-squared      =          0.5148
065                                     Root MSE          =          2.3545
066
067 -----+-----+-----+-----+-----+-----
068          |          Coef.   Std. Err.      t    P>|t|          [95% Conf. Interval]
069 -----+-----+-----+-----+-----+-----
070     x2   | 1.951828        .074414        26.229  0.000          1.805802          2.097854
071     x3   | -1.465202       .0751222       -19.504  0.000          -1.612618         -1.317786
072     x4   | .0266098        .0757442         0.351  0.725          -.1220267          .1752463
073     _cons | 2.898139        .0745214        38.890  0.000          2.751902          3.044376
    
```

```
072 -----
073
074 . regress mst x2 x3 x4
075
076 Source |          SS      df      MS                Number of obs =      1000
077 -----+-----                F( 3, 996) = 1894.36
078 Model   |  5678.36402      3  1892.78801          Prob > F      = 0.0000
079 Residual |  995.172203    996  .999168878          R-squared     = 0.8509
080 -----+-----                Adj R-squared  = 0.8504
081 Total   |  6673.53622    999  6.68021644          Root MSE     = .99958
082
083 -----
084      mst |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
085 -----+-----
086      x2 |  -0.0252742   .0315922    -0.800   0.424   -0.0872692   .0367207
087      x3 |  -1.945355    .0318929   -60.996   0.000   -2.00794   -1.88277
088      x4 |   1.482703    .032157    46.108   0.000    1.4196    1.545807
089     _cons |   .9533604    .0316378    30.134   0.000    .891276    1.015445
090 -----
091
092 . heckman yx x2 x3 x4, select(x2 x3 x4)
093
094 Iteration 0:  log likelihood = -1698.2726
095 Iteration 1:  log likelihood = -1696.9803
096 Iteration 2:  log likelihood = -1696.9745
097 Iteration 3:  log likelihood = -1696.9745
098
099 Heckman selection model                Number of obs      =      1000
100 (regression model with sample selection)  Censored obs       =       345
101                                           Uncensored obs     =       655
102
103                                           Wald chi2(3)       =       691.96
104 Log likelihood = -1696.974              Prob > chi2        =       0.0000
105
106 -----
107      |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
108 -----+-----
109 yx
110      x2 |   1.989593   .0858265    23.182   0.000    1.821376    2.15781
111      x3 |  -1.377036   .1175197   -11.717   0.000   -1.60737   -1.146702
112      x4 |  -0.0862489  .107443    -0.803   0.422   -0.2968333  .1243355
113     _cons |   3.016034   .1246573    24.195   0.000    2.77171    3.260358
114 -----
115 select
116      x2 |   .0044422   .0594406     0.075   0.940   -0.1120592  .1209436
117      x3 |  -1.959868   .1151432   -17.021   0.000   -2.185545   -1.734192
118      x4 |   1.526502   .1006375    15.168   0.000    1.329256    1.723748
119     _cons |   1.007486   .0757667    13.297   0.000    .8589856    1.155985
120 -----
121 /athrho  |   1.092788   .1452801     7.522   0.000    .8080446    1.377532
122 /lnsigma |   .8500475   .0307585    27.636   0.000    .7897619    .910333
123 -----
124      rho  |   .7978936   .0527898     .6685101  .8803974
125     sigma |   2.339758   .0719674     2.202872  2.48515
126     lambda |   1.866878   .1578915     1.557416  2.176339
127 -----
128 LR test of indep. eqns. (rho = 0):  chi2(1) = 46.25  Prob > chi2 = 0.0000
129 -----
130
131 . * Clarification 2. Two-step estimates
132 . probit m x2 x3 x4
133
134 Iteration 0:  log likelihood = -644.29638
135 Iteration 1:  log likelihood = -358.74838
136 Iteration 2:  log likelihood = -290.65584
137 Iteration 3:  log likelihood = -272.54272
138 Iteration 4:  log likelihood = -270.6111
139 Iteration 5:  log likelihood = -270.58461
140 Iteration 6:  log likelihood = -270.58461
141
142 Probit estimates                Number of obs      =      1000
```

```
143                                     LR chi2(3)      =      747.42
144                                     Prob > chi2     =      0.0000
145 Log likelihood = -270.58461         Pseudo R2      =      0.5800
146
```

```
147 -----
148      m |      Coef.   Std. Err.      z    P>|z|      [95% Conf. Interval]
149 -----+-----
150      x2 |   -.0322762   .0629343    -0.513   0.608   - .1556251   .0910728
151      x3 |  -1.892455   .1233495   -15.342   0.000   -2.134216  -1.650695
152      x4 |   1.487229   .1072403   13.868   0.000    1.277042   1.697417
153     _cons |   .9834887   .0777608   12.648   0.000    .8310803   1.135897
154 -----
```

155
156 Note: 2 failures and 22 successes completely determined.

```
157
158 . predict mhat, xb
159
160 . mat pihat = get(_b)
161
162 . gen lamhat = -normd(mhat) / normprob(mhat)
163
164 . regress yx x2 x3 x4 lamhat
165
```

```
166 Source |      SS      df      MS                Number of obs =      655
167 -----+-----                F( 4, 650) = 185.42
168 Model | 3469.22361      4  867.305903          Prob > F      = 0.0000
169 Residual | 3040.31819    650  4.6774126          R-squared     = 0.5329
170 -----+-----                Adj R-squared = 0.5301
171 Total | 6509.5418     654  9.95342783          Root MSE     = 2.1627
172
```

```
173 -----
174      yx |      Coef.   Std. Err.      t    P>|t|      [95% Conf. Interval]
175 -----+-----
176      x2 |   1.998592   .0818465    24.419   0.000    1.837876   2.159307
177      x3 |  -1.461874   .1450389   -10.079   0.000   -1.746675  -1.177073
178      x4 |  -.0285287   .1232674    -0.231   0.817   -.2705791   .2135216
179     lamhat | -2.110151   .3546256    -5.950   0.000   -2.806501  -1.413801
180     _cons |   2.904152   .1779002   16.325   0.000    2.554824   3.253481
181 -----
```

```
182
183 .
184 .
185 end of do-file
186
187 . log close
188
```