. // NULL MODELS

. // binary outcome

. logit dv\_indicator, vce(cl mineid) offset(lnhours) or

Iteration 0: log pseudolikelihood = -1978.0504

Iteration 1: log pseudolikelihood = -1893.3905

Iteration 2: log pseudolikelihood = -1893.3681

Iteration 3: log pseudolikelihood = -1893.3681

Logistic regression Number of obs = 6,253

Wald chi2(0) = .

Log pseudolikelihood = -1893.3681 Prob > chi2 = .

(Std. Err. adjusted for 1,238 clusters in mineid)

------------------------------------------------------------------------------

| Robust

dv\_indicator | Odds Ratio Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

\_cons | .0000987 5.20e-06 -174.94 0.000 .000089 .0001094

lnhours | 1 (offset)

------------------------------------------------------------------------------

.

. pause "next"

.

. lfit

Logistic model for dv\_indicator, goodness-of-fit test

number of observations = 6253

number of covariate patterns = 1

Pearson chi2(0) = 0.00

Prob > chi2 = .

.

. pause "next"

.

. linktest

Iteration 0: log likelihood = -2828.3302

Iteration 1: log likelihood = -1965.1216

Iteration 2: log likelihood = -1826.6339

Iteration 3: log likelihood = -1780.1485

Iteration 4: log likelihood = -1778.5091

Iteration 5: log likelihood = -1778.5084

Iteration 6: log likelihood = -1778.5084

Logistic regression Number of obs = 6,253

LR chi2(2) = 2099.64

Prob > chi2 = 0.0000

Log likelihood = -1778.5084 Pseudo R2 = 0.3712

------------------------------------------------------------------------------

dv\_indicator | Coef. Std. Err. z P>|z| [95% Conf. Interval]

-------------+----------------------------------------------------------------

\_hat | 1.102418 .0515933 21.37 0.000 1.001297 1.203539

\_hatsq | .2453823 .0240955 10.18 0.000 .1981559 .2926087

\_cons | -.6749075 .0655041 -10.30 0.000 -.8032932 -.5465217

------------------------------------------------------------------------------

.

. pause "next"

.

. estat classification

Logistic model for dv\_indicator

-------- True --------

Classified | D ~D | Total

-----------+--------------------------+-----------

+ | 5115 747 | 5862

- | 89 302 | 391

-----------+--------------------------+-----------

Total | 5204 1049 | 6253

Classified + if predicted Pr(D) >= .5

True D defined as dv\_indicator != 0

--------------------------------------------------

Sensitivity Pr( +| D) 98.29%

Specificity Pr( -|~D) 28.79%

Positive predictive value Pr( D| +) 87.26%

Negative predictive value Pr(~D| -) 77.24%

--------------------------------------------------

False + rate for true ~D Pr( +|~D) 71.21%

False - rate for true D Pr( -| D) 1.71%

False + rate for classified + Pr(~D| +) 12.74%

False - rate for classified - Pr( D| -) 22.76%

--------------------------------------------------

Correctly classified 86.63%

--------------------------------------------------

.

. pause "next"

.

. predict b\_null\_yhat

(option pr assumed; Pr(dv\_indicator))

. summ dv\_indicator b\_null\_yhat

Variable | Obs Mean Std. Dev. Min Max

-------------+---------------------------------------------------------

dv\_indicator | 6,253 .8322405 .3736824 0 1

b\_null\_yhat | 6,253 .8322405 .176821 .0062748 .9964138

.

. pause "complete: null model - binary outcome"