. // NULL MODELS

. // binary outcome

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

Iteration 0: log pseudolikelihood = -15243.646

Iteration 1: log pseudolikelihood = -15243.423

Iteration 2: log pseudolikelihood = -15243.423

Logistic regression Number of obs = 30,289

Wald chi2(0) = .

Log pseudolikelihood = -15243.423 Prob > chi2 = .

(Std. Err. adjusted for 1,582 clusters in mineid)

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

| Robust

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

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

\_cons | .0000846 2.33e-06 -339.73 0.000 .0000801 .0000893

lnhours | 1 (offset)

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

.

. pause "next"

.

. lfit

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

number of observations = 30289

number of covariate patterns = 1

Pearson chi2(0) = 0.00

Prob > chi2 = .

.

. pause "next"

.

. linktest

Iteration 0: log likelihood = -20828.862

Iteration 1: log likelihood = -14825.046

Iteration 2: log likelihood = -14607.871

Iteration 3: log likelihood = -14592.489

Iteration 4: log likelihood = -14592.466

Iteration 5: log likelihood = -14592.466

Logistic regression Number of obs = 30,289

LR chi2(2) = 12472.79

Prob > chi2 = 0.0000

Log likelihood = -14592.466 Pseudo R2 = 0.2994

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

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

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

\_hat | 1.23305 .0151438 81.42 0.000 1.203369 1.262731

\_hatsq | .1740168 .0041334 42.10 0.000 .1659156 .1821181

\_cons | -.2342933 .0155529 -15.06 0.000 -.2647764 -.2038102

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

.

. pause "next"

.

. estat classification

Logistic model for dv\_indicator

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

Classified | D ~D | Total

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

+ | 13909 4288 | 18197

- | 2819 9273 | 12092

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

Total | 16728 13561 | 30289

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

True D defined as dv\_indicator != 0

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

Sensitivity Pr( +| D) 83.15%

Specificity Pr( -|~D) 68.38%

Positive predictive value Pr( D| +) 76.44%

Negative predictive value Pr(~D| -) 76.69%

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

False + rate for true ~D Pr( +|~D) 31.62%

False - rate for true D Pr( -| D) 16.85%

False + rate for classified + Pr(~D| +) 23.56%

False - rate for classified - Pr( D| -) 23.31%

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

Correctly classified 76.54%

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

.

. 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 | 30,289 .5522797 .4972675 0 1

b\_null\_yhat | 30,289 .5522797 .268938 .0000846 .983727

.

. pause "complete: null model - binary outcome"

.