Точность коэффициентов множественной регрессии

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2 \right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n \mathrm{MSD}\left(X_2 \right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Эта последовательность исследует различия и средние квадратические ошибки наклонных коэффициентов в модели с двумя независимыми переменными.

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum\left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}\left(X_2\right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Выражение для дисперсии $\hat{\beta}_2$ показано выше. Выражение для дисперсии $\hat{\beta}_3$ одинаковы, с индексами 2 и 3 меняются местами.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2 \right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\,\mathrm{MSD}\left(X_2 \right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Первый коэффициент в выражении идентичен коэффициенту дисперсии коэффициента уклона в простой регрессионной модели.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\,\mathrm{MSD}\left(X_2\right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Дисперсия зависит от дисперсии возмущающего члена $\hat{\beta}_2$, количества наблюдений и среднего квадратного отклонения X2 точно по тем же причинам, что и в простой регрессионной модели.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum\left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}\left(X_2\right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Разница в том, что при множественном регрессионном анализе выражение умножается на коэффициент, который зависит от корреляции между X2 и X3.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2 \right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n \mathrm{MSD}\left(X_2 \right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

Чем выше корреляция между объясняющими переменными, положительными или отрицательными, тем больше будет дисперсия.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2 \right)^2} \times \boxed{\frac{1}{1 - r^2_{X_2, X_3}}} = \frac{\sigma_u^2}{n \mathrm{MSD}\left(X_2 \right)} \times \boxed{\frac{1}{1 - r^2_{X_2, X_3}}}$$

Это легко понять интуитивно. Чем больше корреляция, тем труднее различать влияния независимых переменных на Y и, тем менее точными будут оценки регрессии.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum\left(X_{2i} - \bar{X}_2\right)^2} \times \boxed{\frac{1}{1 - r^2_{X_2,X_3}}} = \frac{\sigma_u^2}{n\mathrm{MSD}\left(X_2\right)} \times \boxed{\frac{1}{1 - r^2_{X_2,X_3}}}$$

Обратите внимание, что приведенное выше выражение дисперсии допустимо только для модели с двумя независимыми переменными. Когда их больше двух, выражение становится гораздо более сложным и имеет смысл переключиться на матричную алгебру.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}\left(X_2\right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

$$\text{standard deviation of } \hat{\beta}_2 = \sqrt{\frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}}}$$

Стандартное отклонение распределения $\hat{\beta}_2$, конечно, задается квадратным корнем его дисперсии.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}(X_2)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

standard deviation of $\hat{\beta}_2 = \sqrt{\frac{\sigma_u^2}{\sum (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2, X_3}}}$

За исключением дисперсии u, мы можем вычислить компоненты стандартного отклонения от выборочных данных.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}(X_2)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

$$E\left(\frac{1}{n} \sum \hat{u}_i^2\right) = \frac{n-k}{n} \sigma_u^2$$

Дисперсия u должна быть оценена. Средний квадрат остатков обеспечивает последовательную оценку , но он смещен вниз фактором (n – k) / n, где k-число параметров в конечной выборке.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum \left(X_{2i} - \bar{X}_2 \right)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n\mathrm{MSD}\left(X_2 \right)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

$$E\left(\frac{1}{n} \sum \hat{u}_i^2 \right) = \frac{n-k}{n} \sigma_u^2$$

$$\hat{\sigma}_u^2 = \frac{1}{n-k} \sum \hat{u}_i^2$$

Очевидно, что мы можем получить несмещенную оценку, разделив сумму квадратов остатков на n – k вместо n. Обозначим эту непредвзятую оценку $\hat{\sigma}_u^2$.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

True model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u$$

Fitted model

$$\hat{Y} = \hat{\beta}_1 + \hat{\beta}_2 X_2 + \hat{\beta}_3 X_3$$

$$\sigma^2_{\hat{\beta}_2} = \frac{\sigma_u^2}{\sum (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2, X_3}} = \frac{\sigma_u^2}{n \mathrm{MSD}(X_2)} \times \frac{1}{1 - r^2_{X_2, X_3}}$$

$$E\left(\frac{1}{n} \sum \hat{u}_i^2 \right) = \frac{n-k}{n} \sigma_u^2$$

$$\hat{\sigma}_u^2 = \frac{1}{n-k} \sum \hat{u}_i^2$$

$$\text{s.e.}(\hat{\beta}_2) = \sqrt{\frac{\hat{\sigma}_u^2}{\sum (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2, X_3}}} = \sqrt{\frac{\hat{\sigma}_u^2}{n \mathrm{MSD}(X_2)} \times \frac{1}{1 - r^2_{X_2, X_3}}}$$

Таким образом , оценка стандартного отклонения распределения вероятностей $\hat{\beta}_2$, известного как стандартная ошибка для краткости $\hat{\beta}_2$, дается выражением выше.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,     72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.       t     P>|t|      [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010     .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746    -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926    -20.60593    22.64009
----------------------------------------------------------------------------
```

Мы будем использовать это выражение, чтобы проанализировать, почему стандартная ошибка S-это больше для Союза подвыборки, чем для несоюзной подвыборки по заработной плате уравнение регрессии, используя набор данных 21.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,    72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010     .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746    -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926    -20.60593    22.64009
----------------------------------------------------------------------------
```

Чтобы выбрать подвыборку в Stata, добавьте в команду оператор 'if'. Переменная КОЛЛЬБЕРГА равна 1 Для респондентов, ставки оплаты труда которые определяются коллективным договором, а для остальных-0.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,    72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436

----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.       t    P>|t|     [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010     .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746    -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926    -20.60593    22.64009
----------------------------------------------------------------------------
```

Обратите внимание, что в тестах на равенство Stata требует, чтобы знак = дублировался.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,     72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t    P>|t|      [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010      .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746     -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926     -20.60593    22.64009
----------------------------------------------------------------------------
```

В случае союза подвыборки, стандартная ошибка s является 0.5365.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==0

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =     425
-------------+------------------------------           F(  2,   422) =   29.48
       Model |  7270.82789     2  3635.41394           Prob > F      =  0.0000
    Residual |  52043.2371   422  123.325206           R-squared     =  0.1226
-------------+------------------------------           Adj R-squared =  0.1184
       Total |   59314.065   424  139.891663           Root MSE      =  11.105
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t     P>|t|     [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |   1.866279   .2438803     7.65   0.000     1.386907     2.34565
         EXP |   1.100186   .2223238     4.95   0.000     .6631858    1.537186
       _cons |   -15.9847   4.623791    -3.46   0.001    -25.07323   -6.896172
----------------------------------------------------------------------------
```

В случае подвыборки без объединения стандартная ошибка S равна 0,2439, что в два раза меньше.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2,X_3}}}$$

$$\text{s.e.}\left(\hat{\beta}_2\right) = \sqrt{\frac{\hat{\sigma}_u^2}{\sum\left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r^2_{X_2,X_3}}} = \sqrt{\frac{\hat{\sigma}_u^2}{n\text{MSD}\left(X_2\right)} \times \frac{1}{1 - r^2_{X_2,X_3}}}$$

Мы объясним разницу, посмотрев на компоненты стандартной ошибки. Удобно начать с перестановки выражения для стандартной ошибки как произведения четырех факторов.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*					
Component	$\hat{\sigma}_u$	n	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union			0.5365		
Non-union			0.2439		
Factor			*product*		
Union					
Non-union					

Мы аранжируем компоненты стандартной ошибки как таблицу.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,     72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t     P>|t|      [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010      .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746     -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926     -20.60593    22.64009
----------------------------------------------------------------------------
```

$$\hat{\sigma}_u^2 = \frac{1}{n-k} RSS$$

Мы начнем с $\hat{\sigma}_u$. Вот RSS для соединения подвыборки .

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

------------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,    72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
------------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010     .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746    -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926    -20.60593    22.64009
------------------------------------------------------------------------------
```

$$\hat{\sigma}_u^2 = \frac{1}{n-k} RSS$$

В не состоящем в профсоюзе подобразце есть 75 наблюдений. k равен 3. Таким образом n – k равен 72.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==1

------------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =      75
-------------+------------------------------           F(  2,     72) =    4.72
       Model |  1027.91667     2  513.958336           Prob > F      =  0.0119
    Residual |  7841.35558    72  108.907716           R-squared     =  0.1159
-------------+------------------------------           Adj R-squared =  0.0913
       Total |  8869.27225    74   119.85503           Root MSE      =  10.436
------------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
           S |    1.42955    .536452     2.66   0.010     .3601522    2.498947
         EXP |   .1918676   .5901747     0.33   0.746    -.9846242    1.368359
       _cons |    1.01708   10.84695     0.09   0.926    -20.60593    22.64009
------------------------------------------------------------------------------
```

$$\hat{\sigma}_u^2 = \frac{1}{n-k} RSS$$

RSS / (n – k) равен 108.908. Для получения $\hat{\sigma}_u$ берем квадратный корень. Это 10.436.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*					
Component	$\hat{\sigma}_u$	n	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	**10.436**	**75**	**0.5365**		
Non-union			**0.2439**		
Factor			***product***		
Union					
Non-union					

Мы поместим это в таблицу вместе с количеством наблюдений.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. reg EARNINGS S EXP if COLLBARG==0

----------------------------------------------------------------------------
      Source |       SS       df       MS              Number of obs =     425
-------------+------------------------------           F(  2,   422) =   29.48
       Model |  7270.82789     2  3635.41394           Prob > F      =  0.0000
    Residual |  52043.2371   422  123.325206           R-squared     =  0.1226
-------------+------------------------------           Adj R-squared =  0.1184
       Total |   59314.065   424  139.891663           Root MSE      =  11.105
----------------------------------------------------------------------------
    EARNINGS |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+--------------------------------------------------------------
           S |   1.866279   .2438803     7.65   0.000     1.386907     2.34565
         EXP |   1.100186   .2223238     4.95   0.000     .6631858    1.537186
       _cons |   -15.9847   4.623791    -3.46   0.001    -25.07323   -6.896172
----------------------------------------------------------------------------
```

Аналогично $\hat{\sigma}_u$, в случае подвыборки без объединения, является квадратный корень из 123.325, что 11.105. Мы также отмечаем, что количество наблюдений в выборке составляет 425.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	0.5365		
Non-union	11.105	425	0.2439		

Factor	*product*
Union	
Non-union	

Мы поместим это в столбцы.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2,X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	7.6932	0.5365	
Non-union	11.105	425	7.3467		0.2439

Factor				*product*
Union				
Non-union				

Мы вычисляем среднее квадратное отклонение S для двух подмножеств из выборочных данных.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

```
. cor S EXP if COLLBARG==1
(obs=75)
             |        S      EXP
-------------+------------------
           S |   1.0000
         EXP |  -0.5866   1.0000

. cor S EXP if COLLBARG==0
(obs=425)
             |        S      EXP
-------------+------------------
           S |   1.0000
         EXP |  -0.5796   1.0000
```

Коэффициенты корреляции для S и EXP -0.5866 и -0.5796 для Союза и несоюзной проб, соответственно. (Обратите внимание, что" cor " - это команда Stata для вычисления корреляций.)

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor				***product***	
Union					
Non-union					

Эти записи завершают верхнюю половину таблицы. Теперь мы рассмотрим влияние каждого элемента на стандартную ошибку, используя математическое выражение вверху.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439

Factor					*product*
Union	10.436				
Non-union	11.105				

$\hat{\sigma}_u$ компоненты не нуждаются в модификации. Это немного больше для несоюзной подвыборки, и поэтому отрицательно сказывается на стандартную ошибку.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2,X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	n	MSD(S)	$r_{S,\ EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor					***product***
Union	10.436	0.1155			
Non-union	11.105	0.0485			

Число наблюдений значительно больше для несоюзной подвыборки, так что второй фактор намного меньше, чем для союза подвыборки.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \boxed{\frac{1}{\sqrt{\mathrm{MSD}(X_2)}}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*					
Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\ EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor					*product*
Union	10.436	0.1155	0.3605		
Non-union	11.105	0.0485	0.3689		

Возможно, удивительно различие в обучении подобно для этих двух подобразцов.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	n	MSD(S)	$r_{S,\ EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor				***product***	
Union	10.436	0.1155	0.3605	1.2348	
Non-union	11.105	0.0485	0.3689	1.2271	

Корреляция между образованием и опытом работы также похожие на две подвыборки. Обратите внимание, что знак корреляции не имеет значения, так как он квадрат.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2,X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	n	MSD(S)	$r_{S,\,EXP}$	s.e.
Union	**10.436**	**75**	**7.6932**	**–0.5866**	**0.5365**
Non-union	**11.105**	**425**	**7.3467**	**–0.5796**	**0.2439**
Factor				***product***	
Union	**10.436**	**0.1155**	**0.3605**	**1.2348**	**0.5366**
Non-union	**11.105**	**0.0485**	**0.3689**	**1.2271**	**0.2439**

Умножая четыре фактора вместе, получаем стандартные ошибки. (Расхождение в последней цифре стандартной ошибки объединения было вызвано ошибкой округления.)

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1 - r^2_{X_2, X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	*n*	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor					***product***
Union	10.436	0.1155	0.3605	1.2348	0.5366
Non-union	11.105	0.0485	0.3689	1.2271	0.2439

Мы видим, что стандартная ошибка меньше для несоюзной подвыборки заключается в том, что существует гораздо больше замечаний, чем в несоюзной подвыборке. В противном случае стандартные ошибки были бы примерно одинаковы.

ТОЧНОСТЬ МНОГОКРАТНЫХ КОЭФФИЦИЕНТОВ РЕГРЕССА

$$\text{s.e.}\left(\hat{\beta}_2\right) = \hat{\sigma}_u \times \frac{1}{\sqrt{n}} \times \frac{1}{\sqrt{\text{MSD}(X_2)}} \times \frac{1}{\sqrt{1-r^2_{X_2,X_3}}}$$

Decomposition of the standard error of *S*

Component	$\hat{\sigma}_u$	n	MSD(*S*)	$r_{S,\,EXP}$	s.e.
Union	10.436	75	7.6932	–0.5866	0.5365
Non-union	11.105	425	7.3467	–0.5796	0.2439
Factor					***product***
Union	10.436	0.1155	0.3605	1.2348	0.5366
Non-union	11.105	0.0485	0.3689	1.2271	0.2439

Согласие $\hat{\sigma}_u$, как измерено, является немного низшим для не состоящего в профсоюзе образца, и это имеет крайний эффект возмещения. Другие два фактора очень похожи для этих двух подобразцов.