TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

In quest'ultima sequenza, viene illustrato come il livello di istruzione dipenda dalla capacità cognitiva e dal background familiare, con l'istruzione del padre e della madre come proxy di quest'ultimo.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SM |   .0492424   .0390901     1.26   0.208     -.027546    .1260309
          SF |   .1076825   .0309522     3.48   0.001       .04688    .1684851
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------
```

Comunque, quando facciamo girare un modello di regressione usando il Data Set 21, osserviamo che il coefficiente dell’istruzione della madre non è significativo.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SM |   .0492424   .0390901     1.26   0.208     -.027546    .1260309
          SF |   .1076825   .0309522     3.48   0.001       .04688    .1684851
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------
```

```
. cor SM SF
(obs=540)
        |       SM        SF
--------+------------------
      SM|   1.0000
      SF|   0.6241    1.0000
```

Come abbiamo già visto, ciò potrebbe essere un effetto della multicollinearità, perché l'istruzione della madre è correlata con quella del padre.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

Come già trattato in precedenza, vi sono diverse misure per alleviare il problema della multicollinearità, tra cui imporre un’appropriata restrizione teorica.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

In particolare, nel caso del presente modello, si suggerisce che l'impatto dell'istruzione dei genitori potrebbe essere lo stesso per entrambi i genitori, ovvero β_3 e β_4 potrebbero essere uguali.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

$$\begin{aligned} S &= \beta_1 + \beta_2 ASVABC + \beta_3 (SM + SF) + u \\ &= \beta_1 + \beta_2 ASVABC + \beta_3 SP + u \end{aligned}$$

$$SP = SM + SF$$

Sotto tale ipotesi, il modello può essere riscritto come sopra. Adesso abbiamo una nuova variabile che esprime l'istruzione dei genitori, *SP*, e la multicollinearità causata dalla correlazione tra *SM* e *SF* è stata eliminata.

TESTARE UNA RESTRIZIONE LINEARE

```
. g SP=SM+SF

. reg S ASVABC SP

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  2,   537) =  156.04
       Model |  1177.98338     2  588.991689           Prob > F      =  0.0000
    Residual |  2026.99996   537  3.77467403           R-squared     =  0.3675
-------------+------------------------------           Adj R-squared =  0.3652
       Total |  3204.98333   539  5.94616574           Root MSE      =  1.9429

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1253106   .0098434    12.73   0.000     .1059743    .1446469
          SP |   .0828368   .0164247     5.04   0.000     .0505722    .1151014
       _cons |    5.29617   .4817972    10.99   0.000     4.349731    6.242608
------------------------------------------------------------------------------
```

Sopra vengono riportati i risultati della regressione rimpiazzando *SM* e *SF* con *SP*.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SM |   .0492424   .0390901     1.26   0.208    -.027546     .1260309
          SF |   .1076825   .0309522     3.48   0.001      .04688     .1684851
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------

. reg S ASVABC SP

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1253106   .0098434    12.73   0.000     .1059743    .1446469
          SP |   .0828368   .0164247     5.04   0.000     .0505722    .1151014
       _cons |    5.29617   .4817972    10.99   0.000     4.349731    6.242608
------------------------------------------------------------------------------
```

Il confronto tra le due regressioni mostra che lo standard error del coefficiente di *SP* è più piccolo dello standard error di *SM* e *SF*, e conseguentemente la sua statistica *t* è più alta. Il suo coefficiente è un compromesso tra quello di *SM* e quello di *SF*.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF
```

S	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ASVABC	.1257087	.0098533	12.76	0.000	.1063528	.1450646
SM	.0492424	.0390901	1.26	0.208	-.027546	.1260309
SF	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. reg S ASVABC SP
```

S	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ASVABC	.1253106	.0098434	12.73	0.000	.1059743	.1446469
SP	.0828368	.0164247	5.04	0.000	.0505722	.1151014
_cons	5.29617	.4817972	10.99	0.000	4.349731	6.242608

Comunque, l'uso di una restrizione porta ad avere un guadagno nell'efficienza solo se l'efficienza è valida. Se non è valida, il suo utilizzo porta ad avere degli stimatori distorti quindi standard error and test non validi.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF
```

S	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ASVABC	.1257087	.0098533	12.76	0.000	.1063528	.1450646
SM	.0492424	.0390901	1.26	0.208	-.027546	.1260309
SF	.1076825	.0309522	3.48	0.001	.04688	.1684851
_cons	5.370631	.4882155	11.00	0.000	4.41158	6.329681

```
. reg S ASVABC SP
```

S	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
ASVABC	.1253106	.0098434	12.73	0.000	.1059743	.1446469
SP	.0828368	.0164247	5.04	0.000	.0505722	.1151014
_cons	5.29617	.4817972	10.99	0.000	4.349731	6.242608

I coefficienti di *SM* e *SF* nel modello non vincolato sembrano soddisfare tale restrizione? Non sembra! Il coefficiente di *SM* è molto più piccolo di quello di *SF*, ma si nota anche che gli standard error sono abbastanza elevati.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

. reg S ASVABC SP

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  2,   537) =  156.04
       Model |  1177.98338     2  588.991689           Prob > F      =  0.0000
    Residual |  2026.99996   537  3.77467403           R-squared     =  0.3675
-------------+------------------------------           Adj R-squared =  0.3652
       Total |  3204.98333   539  5.94616574           Root MSE      =  1.9429
```

Adesso possiamo effettuare un test appropriato. L'imposizione di una restrizione induce ad avere un modello con un fit più basso. Infatti c'è un incremento di *RSS* (e decremento di R^2) quando c'è l'imposizione.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

. reg S ASVABC SP

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  2,   537) =  156.04
       Model |  1177.98338     2  588.991689           Prob > F      =  0.0000
    Residual |  2026.99996   537  3.77467403           R-squared     =  0.3675
-------------+------------------------------           Adj R-squared =  0.3652
       Total |  3204.98333   539  5.94616574           Root MSE      =  1.9429
```

Se la restrizione fosse valida il deterioramento del modello dovrebbe essere molto piccolo, un ammontare casuale. Comunque se la restrizione non è valida la distorsione causata attraverso la restrizione può portare ad un deterioramento significativo nel fit.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

. reg S ASVABC SP

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  2,   537) =  156.04
       Model |  1177.98338     2  588.991689           Prob > F      =  0.0000
    Residual |  2026.99996   537  3.77467403           R-squared     =  0.3675
-------------+------------------------------           Adj R-squared =  0.3652
       Total |  3204.98333   539  5.94616574           Root MSE      =  1.9429
```

In questo caso, possiamo notare che l'incremento di *RSS* è molto piccolo, e quindi sarà molto improbabile rifiutare la restrizione.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 (SM + SF) + u$$
$$= \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$SP = SM + SF$$

$$H_0 : \beta_4 = \beta_3, \quad H_1 : \beta_4 \neq \beta_3$$

L'ipotesi nulla è che la restrizione sia valida, mentre l'alternativa è che non sia vera.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 (SM + SF) + u$$
$$= \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$SP = SM + SF$$

$$H_0 : \beta_4 = \beta_3, \quad H_1 : \beta_4 \neq \beta_3$$

$$F(1, n-k) = \frac{(RSS_R - RSS_U)/1}{RSS_U/(n-k)} = \frac{2027.00 - 2023.61}{2023.61/536} = 0.90$$

La statistica test si distribuisce come una *F* con 1 e *n-k* gradi di libertà dove il numeratore rappresenta il miglioramento del fit rilassando la restrizione, diviso per il costo di tale rilassamento (un grado di libertà, perché un ulteriore parametro deve essere stimato).

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 (SM + SF) + u$$
$$= \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$SP = SM + SF$$

$$H_0 : \beta_4 = \beta_3, \quad H_1 : \beta_4 \neq \beta_3$$

$$F(1, n-k) = \frac{(RSS_R - RSS_U)/1}{RSS_U/(n-k)} = \frac{2027.00 - 2023.61}{2023.61/536} = 0.90$$

Il denominatore della statistica test è *RSS* dopo aver "migliorato" il modello (cioè *RSS* per il modello non vincolato), diviso per $n - k$, il numero di gradi di libertà rimanenti. k è il numero di parametri nel modello non vincolato.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$\beta_4 = \beta_3$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 (SM + SF) + u$$
$$= \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$SP = SM + SF$$

$$H_0 : \beta_4 = \beta_3, \quad H_1 : \beta_4 \neq \beta_3$$

$$F(1, n-k) = \frac{(RSS_R - RSS_U)/1}{RSS_U/(n-k)} = \frac{2027.00 - 2023.61}{2023.61/536} = 0.90$$

La statistica *F* è 0.90. Un valore di *F* sotto 1 non è mai significativo (guarda le tavole della *F*), quindi non si può rifiutare H_0. La restrizione sembra essere valida dall'evidenza empirica.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SM SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SM |   .0492424   .0390901     1.26   0.208     -.027546    .1260309
          SF |   .1076825   .0309522     3.48   0.001       .04688    .1684851
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------
. test SM =SF

 ( 1)  SM - SF = 0

       F(  1,   536) =    0.90
            Prob > F =    0.3440
```

In STATA il comando per fare i test è *test* e ad esso devo far seguire l’ipotesi di base.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

Le restrizioni lineari possono essere testate usando anche il test *t*. Riscriviamo il modello nella versione vincolata e aggiungiamo il termine che dovrebbe convertirlo alla versione non vincolata. Il test valuta se è necessario il termine aggiuntivo.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$0 = \beta_3 SM + \beta_4 SF - \beta_3 SP$$

Per trovare il termine di conversione, riscriviamo il modello vincolato e quello non vincolato e sottraiamo membro a membro.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$\begin{aligned} 0 &= \beta_3 SM + \beta_4 SF - \beta_3 SP \\ &= \beta_3 SM + \beta_4 SF - \beta_3 (SM + SF) \\ &= (\beta_4 - \beta_3) SF \end{aligned}$$

Notiamo che il termine che converte il modello vincolato a quello non vincolato è (β_4 – β_3)*SF*.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$\begin{aligned} 0 &= \beta_3 SM + \beta_4 SF - \beta_3 SP \\ &= \beta_3 SM + \beta_4 SF - \beta_3 (SM + SF) \\ &= (\beta_4 - \beta_3) SF \end{aligned}$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + (\beta_4 - \beta_3) SF + u$$

Se aggiungiamo questo termine otteniamo una riparametrizzazione del modello non vincolato.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$\begin{aligned} 0 &= \beta_3 SM + \beta_4 SF - \beta_3 SP \\ &= \beta_3 SM + \beta_4 SF - \beta_3 (SM + SF) \\ &= (\beta_4 - \beta_3) SF \end{aligned}$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + (\beta_4 - \beta_3) SF + u$$

$$H_0 : \beta_4 - \beta_3 = 0, \quad H_1 : \beta_4 - \beta_3 \neq 0$$

L’ipotesi nulla è che il coefficiente del termine di conversione sia 0, mentre l’alternativa è che sia differente da 0.

TESTARE UNA RESTRIZIONE LINEARE

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SM + \beta_4 SF + u$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + u$$

$$\begin{aligned} 0 &= \beta_3 SM + \beta_4 SF - \beta_3 SP \\ &= \beta_3 SM + \beta_4 SF - \beta_3 (SM + SF) \\ &= (\beta_4 - \beta_3) SF \end{aligned}$$

$$S = \beta_1 + \beta_2 ASVABC + \beta_3 SP + (\beta_4 - \beta_3) SF + u$$

$$H_0 : \beta_4 - \beta_3 = 0, \quad H_1 : \beta_4 - \beta_3 \neq 0$$

Ovviamente l'ipotesi nulla è che la restrizione sia valida. Se è valida, non c'è bisogno del termine di conversione e quindi la versione vincolata è un'adeguata rappresentazione dei dati.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SP SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SP |   .0492424   .0390901     1.26   0.208     -.027546    .1260309
          SF |   .0584401   .0617051     0.95   0.344    -.0627734    .1796536
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------
```

Sopra riportiamo i risultati della regressione corrispondente. Osserviamo che il coefficiente di *SF* non è significativamente diverso da zero, indicando che la versione vincolata è un'adeguata rappresentazione dei dati.

TESTARE UNA RESTRIZIONE LINEARE

```
. reg S ASVABC SP SF

      Source |       SS       df       MS              Number of obs =     540
-------------+------------------------------           F(  3,   536) =  104.30
       Model |  1181.36981     3  393.789935           Prob > F      =  0.0000
    Residual |  2023.61353   536  3.77539837           R-squared     =  0.3686
-------------+------------------------------           Adj R-squared =  0.3651
       Total |  3204.98333   539  5.94616574           Root MSE      =   1.943

------------------------------------------------------------------------------
           S |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
      ASVABC |   .1257087   .0098533    12.76   0.000     .1063528    .1450646
          SP |   .0492424   .0390901     1.26   0.208     -.027546    .1260309
          SF |   .0584401   .0617051     0.95   0.344    -.0627734    .1796536
       _cons |   5.370631   .4882155    11.00   0.000      4.41158    6.329681
------------------------------------------------------------------------------
```

Può essere mostrato matematicamente che il test *F* e il test *t* sono equivalenti. La statistica *F* è il quadrato della statistica *t* e la soglia critica di *F* è il quadrato del valore di *t*.