Analytic Strategies: Simultaneous, Hierarchical, and Stepwise Regression

This discussion borrows heavily from Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences, by Jacob and Patricia Cohen (1975 edition).

The simultaneous model. In the simultaneous model, all K IVs are treated simultaneously and on an equal footing. Such a research strategy is clearly most appropriate when we have no logical or theoretical basis for considering any variable to be prior to any other, either in terms of a hypothetical causal structure of the data or in terms of its relevance to the research goals.

The hierarchical model. An alternative strategy to the simultaneous model is one in which the K IVs are entered cumulatively according to some specified hierarchy which is dictated in advance by the purpose and logic of the research. The hierarchical model calls for a determination of R^2 and the partial coefficients of each variable at the point at which it is added to the equation. Note that with the addition of the ith IV, the MRC analysis *at that stage* is simultaneous in i variables.

Perhaps the most straightforward use of the hierarchical model is when IVs can be ordered with regard to their temporally or logically determined priority. Thus, for example, when two IVs are sex (male or female) and an attitudinal variable, sex must be considered the causally prior variable since it antedates attitude.

The hierarchical MRC analysis may proceed by entering the IVs in the specified order and determining R^2 after each additions. Suppose, for example, when we regress Y on X1 $R^2_{Y1} = .10$. Then, when we regress Y on X1 and X2, $R^2_{Y12} = .15$. We can say that, after controlling for X1, X2 accounts for 5% of the variance in Y, i.e. $sr^2_{2\bullet1} = R^2_{Y12} - R^2_{Y1} = .15 - .10 = .05$.

Note that the R^2 contribution of any variable depends upon what else is in the equation.

Hierarchical analysis of the variables typically adds to the researcher's understanding of the phenomena being studied, since it requires thoughtful input by the researcher in determining the order of entry of IVs, and yields successive tests of the validity of the hypotheses which determine that order.

Stepwise regression. *Forward* stepwise regression programs are designed to select from a group of IVs the one variable at each stage which has the largest sr^2, and hence makes the largest contribution to R^2. (This will also be the variable that has the largest T value.) Such programs typically stop admitting IVs into the equation when no IV makes a contribution which is statistically significant at a level specified by the user. Thus, the stepwise procedure defines an a posteriori order based solely on the relative uniqueness of the variables in the sample at hand.

Backwards stepwise regression procedures work in the opposite order. The dependent variable is regressed on all K independent variables. If any variables are statistically insignificant, the one making the smallest contribution is dropped (i.e. the variable with the smallest sr^2, which will also be the variable with the smallest T value). Then the K - 1 remaining variables are regressed on Y, and again the one making the smallest contribution is dropped. The procedure continues until all remaining variables are statistically significant.

Note that forwards and backwards regression need not produce the same final model.

Problems with stepwise regression.

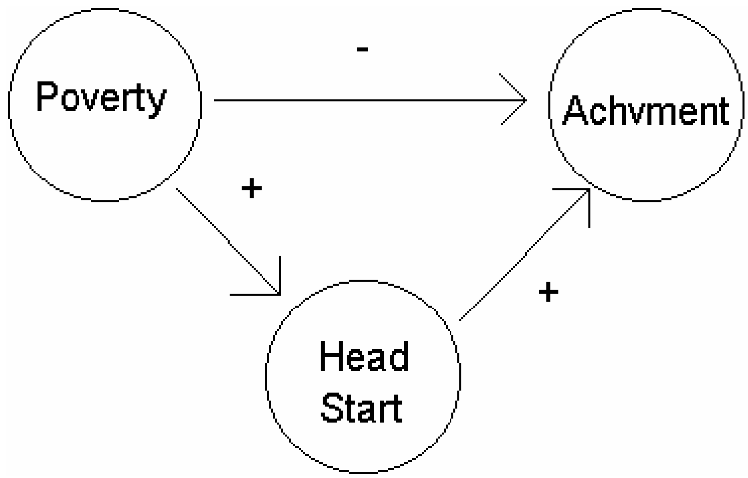
✓ When an investigator has a large pool of potential IVs and very little theory to guide selection among them, stepwise regression is a sore temptation. If the computer selects the variables, the investigator is relieved of the responsibility of making decisions about their logical or causal priority or relevance before the analysis. However, this atheoritical approach tends not to be viewed kindly. Most social scientists believe that more orderly advances are made when researchers armed with theories provide a priori hierarchical ordering which reflects causal hypotheses rather than when computers order IVs post and ad hoc for a given sample.

✓ Another serious problem arises when a relatively large number of IVs is used. Since the significance test of an IV's contribution to R^2 proceeds in ignorance of the large number of other such tests being performed at the same time for the other competing IVs, there can be very serious capitalization on chance. The result is that neither the statistical significance tests for each variable nor the overall tests on the multiple R^2 at each step are valid. (For example, if you have 20 IVs, by chance alone one of them is likely to be significant at the .05 level.)

✓ A related problem is that the ad hoc order produced from a set of IVs in one sample is likely not to be found in other samples from the same population. When there are trivial differences between variables, the computer will dutifully choose the largest for addition at each step. In other samples and, more important, in the population, such differences may well be reversed. When the competing IVs are substantially correlated with each other, the problem is likely to be compounded, since the losers in the competition may not make a sufficiently large unique contribution to be entered at any subsequent step before the problem is terminated by "nonsignificance."

✓ Sometimes with a large number of IVs, variables that were entered into the equation early no longer have nontrivial (i.e. significant) relationships after other variables have been added. However, many programs (e.g. SPSS) allow for the removal of such variables.

✓ Although it is not a common phenomenon, it is possible that neither of two variables alone would reach the criterion for acceptance into the equation, yet if both were entered they would make a useful contribution to R^2. This can occur when *suppression* is present, e.g.



Conditions under which concerns are minimized

✓ The research goal is entirely or primarily predictive (technological), and not at all, or only secondarily, explanatory (scientific). (Still, even research which is only concerned with prediction will likely do better with a theoretically guided model.)

✓ N is very large, and the original K (that is, before forward stepwise selection) is not too great; a K/N ratio of at least 1 to 40 is prudent.

✓ Particularly if the results are to be substantively interpreted, a cross-validation of the stepwise analysis in a new sample should be undertaken, and only those conclusions that hold for both samples should be drawn. Alternatively, the original sample may be randomly divided in half, and the two half-samples treated in this manner.

✓ Stepwise and hierarchical regression can be combined. An investigator may be clear that some groups of variables are logically, causally, or structurally prior to others, and yet have no basis of ordering variables within such groups. Hence, the researcher can decide what order the groups of variables should be entered in, and then let the computer decide within each group what the sequencing should be. This type of analysis is likely to be primarily hierarchical (between classes of IVs) and only incidentally stepwise (within classes).

Stepwise regression example

In this section, I will show how stepwise regression could be used with the Education, Occupation and Earnings example from Sewell and Hauser (1975).

As you look through the handout, make sure you can confirm the different claims that are made. In particular, for each model, check whether the predictions for the next variable to be entered or removed are correct; also check to see whether the R^2 is as predicted.

NOTE!!! SPSS's old style of formatting output is better for purposes of my presentation, ergo I am continuing to use it. Later on I give SPSS's current printout.

```
10-Dec-91   SPSS RELEASE 4.1 FOR IBM OS/MVS

00:49:38    UNIVERSITY OF NOTRE DAME       IBM 370/3081 U   OS/MVS

   1  0  Set width = 125 .
   2  0  Title 'Stepwise Regression Example' .
   3  0  * Adapted from Education, Occupation, and Earnings:
   4  0  * Achievement in the Early Career, by William H. Sewell
   5  0  * and Robert M. Hauser.  1975, Academic Press.
   6  0  * See especially pages 72 and 79.
   7  0  MATRIX DATA VARIABLES=ROWTYPE_ V M X I Q U W Y.

There are 2,951,800 bytes of memory available.
The largest contiguous area has 2,944,880 bytes.
MATRIX DATA has already allocated 440 bytes.
More memory will be allocated to store the data to be read.
  20  0  END DATA.

Preceding task required .04 seconds CPU time;  .59 seconds elapsed.
  21  0  VARIABLE LABELS
  22  0     V "Father's educational attainment"
  23  0     M "Mother's educational attainment"
  24  0     X "Status of F's Occ when son grad from HS"
  25  0     I "Parent's average income, 1957-1960"
  26  0     Q "Son's score on Henmon-Nelson IQ test"
  27  0     U "Son's educational attainment"
  28  0     W "Son's 1964 Occupation (Duncan SEI)"
  29  0     Y "Son's average earnings, 1965-1967" .
  30  0
  31  0  Subtitle  "Stepwise forward selection" .
  32  0  REGRESSION  /MATRIX IN(*)
  33  0              /DESCRIPTIVES DEF
  34  0              /VARIABLES V M X Q U W
  35  0              /CRITERIA=POUT(.051)
  36  0              /STATISTICS DEF CHANGE OUTS ZPP TOL
  37  0              /DEPENDENT W
  38  0              /METHOD FORWARD V M X Q U.
  39  0

There are 2,954,104 bytes of memory available.
The largest contiguous area has 2,947,640 bytes.
```

```
10-Dec-91   Stepwise Regression Example
00:49:43    Stepwise forward selection

                                 * * * *   M U L T I P L E   R E G R E S S I O N   * * * *


Listwise Deletion of Missing Data

             Mean  Std Dev  Label

V          10.200    3.008  Father's educational attainment
M          10.410    2.848  Mother's educational attainment
X          33.110   22.410  Status of F's Occ when son grad from HS
Q          99.860   14.570  Son's score on Henmon-Nelson IQ test
U          13.220    1.676  Son's educational attainment
W          42.110   23.340  Son's 1964 Occupation (Duncan SEI)

N of Cases =  2069

Correlation:

                 V          M          X          Q          U          W

V            1.000       .524       .425       .254       .316       .263
M             .524      1.000       .288       .211       .260       .211
X             .425       .288      1.000       .184       .291       .262
Q             .254       .211       .184      1.000       .448       .377
U             .316       .260       .291       .448      1.000       .621
W             .263       .211       .262       .377       .621      1.000
```

Note that U has the biggest correlation with W (.621). Ergo, U will be the first variable entered into the equation. R^2 will equal $.621^2 = .3856$.

Forward selection (Continued)

```
10-Dec-91   Stepwise Regression Example
00:49:44    Stepwise forward selection

                                  * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE

Block Number  1.  Method:  Forward      Criterion   PIN  .0500       V        M        X        Q        U


Variable(s) Entered on Step Number  1..    U         Son's educational attainment


Multiple R           .62100
R Square             .38564          R Square Change    .38564
Adjusted R Square    .38534          F Change       1297.48233
Standard Error     18.29856          Signif F Change    .0000

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           1        434445.63509     434445.63509
Residual          2067        692108.94571        334.83742
F =    1297.48233       Signif F =  .0000


------------------------------------------ Variables in the Equation ------------------------------------------

Variable              B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U              8.648055     .240086    .621000  .621000  .621000  .621000   1.000000      1.000     36.021  .0000
(Constant)   -72.217286    3.199336                                                                 -22.573  .0000


------------------------ Variables not in the Equation ------------------------

Variable     Beta In  Partial  Tolerance       VIF  Min Toler         T  Sig T
V            .074170  .089779    .900144     1.111    .900144     4.097  .0000
M            .053132  .065455    .932400     1.073    .932400     2.982  .0029
X            .088809  .108401    .915319     1.093    .915319     4.956  .0000
Q            .123599  .140980    .799296     1.251    .799296     6.473  .0000
```

(1) As "Variables not in the equation" shows, Q would have the biggest T value if it were added to the equation, which means that adding Q would produce the greatest increase in R^2. Ergo, Q gets entered next.

(2) sr_Q is not reported but is easily computed. Using the information from the "Beta in" and "Tolerance" columns in the "Variables not in the equation" section, we get

$$sr_Q = b'_Q * \text{sqrt}(Tol_Q) = .123599 * \text{sqrt}(.799296) = .1105,$$

$$sr_Q^2 = .1105^2 = .01221$$

so adding Q will cause R^2 to increase from its current value of .38564 to .39785.

(3) Note that, in a bivariate regression, b' (BETA), the bivariate correlation (CORR), the semipartial correlation (PART CORR), and the partial correlation (PARTIAL) are all identical.

Forward selection (continued)

```
10-Dec-91   Stepwise Regression Example
00:49:44    Stepwise forward selection

                                 * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE


Variable(s) Entered on Step Number  2..    Q         Son's score on Henmon-Nelson IQ test


Multiple R           .63075
R Square             .39785          R Square Change    .01221
Adjusted R Square    .39727          F Change         41.89505
Standard Error     18.12019          Signif F Change    .0000

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           2        448201.50796     224100.75398
Residual          2066        678353.07284        328.34127
F =     682.52386       Signif F =  .0000


------------------------------------------ Variables in the Equation ------------------------------------------

Variable              B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U              7.876940     .265925    .565628  .621000  .505690  .545976    .799296      1.251     29.621  .0000
Q               .197996     .030590    .123599  .377000  .110501  .140980    .799296      1.251      6.473  .0000
(Constant)   -81.794985    3.496676                                                                -23.392  .0000


------------------------ Variables not in the Equation ------------------------

Variable     Beta In  Partial  Tolerance        VIF  Min Toler         T  Sig T
V           .059783  .072449    .884329      1.131    .755588     3.301  .0010
M           .041095  .050830    .921223      1.086    .770639     2.313  .0208
X           .081889  .100764    .911720      1.097    .754271     4.602  .0000
```

(1) As "Variables not in the equation" shows, X would have the largest T value if added to the equation, i.e. it would produce the greatest increase in R^2. Ergo, X gets entered next.

(2) sr_X is not reported but is easily computed. Using the information from the "Beta in" and "Tolerance" columns in the "Variables not in the equation" section, we get

$sr_X = b'_X * sqrt(Tol_X) = .081889 * sqrt(.911720) = .07182,$

$sr_Q^2 = .07182^2 = .00612$

so adding X will cause R^2 to increase from its current value of .39785 to .40397.

Forward selection (continued)

```
10-Dec-91   Stepwise Regression Example
00:49:44    Stepwise forward selection

                                 * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE


Variable(s) Entered on Step Number  3..    X         Status of F's Occ when son grad from HS


Multiple R           .63558
R Square             .40397          R Square Change    .00611
Adjusted R Square    .40310          F Change         21.18189
Standard Error     18.03233          Signif F Change    .0000

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           3        455089.11434     151696.37145
Residual          2065        671465.46646        325.16487
F =     466.52140       Signif F =  .0000


------------------------------------------ Variables in the Equation ------------------------------------------

Variable              B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U              7.579367     .272420    .544260  .621000  .472683  .522162    .754271      1.326     27.822  .0000
Q               .189193     .030501    .118104  .377000  .105381  .135244    .796153      1.256      6.203  .0000
X               .085288     .018531    .081889  .262000  .078191  .100764    .911720      1.097      4.602  .0000
(Constant)   -79.805965    3.506456                                                                 -22.760  .0000


------------------------ Variables not in the Equation ------------------------

Variable     Beta In  Partial  Tolerance        VIF  Min Toler         T  Sig T
V           .034126  .038740    .768120      1.302    .735815     1.761  .0783
M           .023996  .029068    .874679      1.143    .740276     1.321  .1866


End Block Number   1   PIN =     .050 Limits reached.
```

As "Variables not in the equation" shows, V would have the greatest T value (and hence produce the greatest increase in R^2) if it were added to the equation. HOWEVER, the T value for V is not significant at the .05 level, so we do not want to include it. Ergo, STOP NOW!

Backwards Selection

```
10-Dec-91   Stepwise Regression Example
00:49:44    Stepwise forward selection

Preceding task required .14 seconds CPU time;  3.16 seconds elapsed.
  40  0  Subtitle  "Stepwise backward selection".
  41  0  REGRESSION  /MATRIX IN(*)
  42  0              /VARIABLES V M X Q U W
  43  0              /CRITERIA=POUT(.051)
  44  0              /STATISTICS DEF CHANGE OUTS ZPP TOL
  45  0              /DEPENDENT W
  46  0              /METHOD ENTER V M X Q U
  47  0              /METHOD BACKWARDS .
  48  0

There are 2,954,096 bytes of memory available.
The largest contiguous area has 2,947,640 bytes.

     1956 bytes of memory required for REGRESSION procedure.
        0 more bytes may be needed for Residuals plots.
```

```
10-Dec-91   Stepwise Regression Example
00:49:47    Stepwise backward selection

                                        * * * *   M U L T I P L E   R E G R E S S I O N   * * * *


Listwise Deletion of Missing Data

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE
Block Number  1.  Method:  Enter      V        M        X        Q        U


Variable(s) Entered on Step Number  1..    U         Son's educational attainment
                                    2..    M         Mother's educational attainment
                                    3..    X         Status of F's Occ when son grad from HS
                                    4..    Q         Son's score on Henmon-Nelson IQ test
                                    5..    V         Father's educational attainment


Multiple R           .63637
R Square             .40497          R Square Change    .40497
Adjusted R Square    .40353          F Change        280.81132
Standard Error     18.02585          Signif F Change    .0000

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           5        456221.67830      91244.33566
Residual          2063        670332.90250        324.93112
F =     280.81132       Signif F =  .0000

------------------------------------------ Variables in the Equation ------------------------------------------

Variable             B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U             7.490510     .276501    .537879  .621000  .460081  .512245    .731644      1.367     27.090  .0000
M              .102339     .165107    .012488  .211000  .010527  .013645    .710609      1.407       .620  .5354
X              .071806     .019918    .068945  .262000  .061227  .079124    .788648      1.268      3.605  .0003
Q              .181830     .030746    .113508  .377000  .100438  .129116    .782980      1.277      5.914  .0000
V              .220009     .166816    .028354  .263000  .022399  .029025    .624039      1.602      1.319  .1874
(Constant)  -80.759043    3.560080                                                               -22.685  .0000

End Block Number   1   All requested variables entered.
```

(1) We begin by estimating the model in which all independent variables are included. Note that some are not statistically significant. M has the smallest semipartial correlation, which means that dropping it would produce the smallest decrease in R^2. So, the next step is to remove M from the equation.

(2) Let H = the set of all variables currently in the equation. Let G_k = the set of all variables in the equation except the variable that is to be dropped. Then,

$$R_{YGm}^2 = R_{YH}^2 - sr_m^2 = .40497 - .010527^2 = .40486.$$

Ergo, after dropping M, R^2 will equal .40486.

Backwards selection (continued)

```
10-Dec-91   Stepwise Regression Example
00:49:47    Stepwise backward selection

                                 * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE
Block Number  2.  Method:  Backward     Criterion   POUT  .0510


Variable(s) Removed on Step Number  6..    M         Mother's educational attainment


Multiple R           .63629
R Square             .40486          R Square Change   -.00011
Adjusted R Square    .40371          F Change           .38420
Standard Error     18.02316          Signif F Change    .5354

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           4        456096.84110     114024.21028
Residual          2064        670457.73970        324.83418
F =     351.02283       Signif F =  .0000


------------------------------------------ Variables in the Equation ------------------------------------------

Variable              B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U              7.503414     .275675    .538806  .621000  .462185  .513934    .735815      1.359     27.218  .0000
X               .072599     .019874    .069706  .262000  .062031  .080149    .791912      1.263      3.653  .0003
Q               .182813     .030700    .114121  .377000  .101116  .129961    .785070      1.274      5.955  .0000
V               .264791     .150336    .034126  .263000  .029909  .038740    .768120      1.302      1.761  .0783
(Constant)   -80.445490    3.523431                                                                  -22.832  .0000


----------------------- Variables not in the Equation -----------------------

Variable     Beta In  Partial  Tolerance        VIF  Min Toler         T  Sig T
M            .012488  .013645    .710609      1.407    .624039      .620  .5354
```

(1) As the "variables in the equation" shows, V has the smallest semipartial correlation, which means that dropping it would produce the smallest decrease in R^2. Since the effect of V is not statistically significant, remove it next.

(2) Let H = the set of all variables currently in the equation. Let G_k = the set of all variables in the equation except the variable that is to be dropped. Then,

$$R_{YGv}^2 = R_{YH}^2 - sr_v^2 = .40486 - .029909^2 = .40397.$$

Ergo, after dropping V, R^2 will equal .40397.

Backwards selection (continued)

```
10-Dec-91   Stepwise Regression Example
00:49:47    Stepwise backward selection

                                 * * * *   M U L T I P L E   R E G R E S S I O N   * * * *

Equation Number 1    Dependent Variable..   W   Son's 1964 Occupation (Duncan SE


Variable(s) Removed on Step Number  7..    V         Father's educational attainment


Multiple R           .63558
R Square             .40397          R Square Change   -.00089
Adjusted R Square    .40310          F Change          3.10228
Standard Error     18.03233          Signif F Change    .0783

Analysis of Variance
                    DF      Sum of Squares      Mean Square
Regression           3        455089.11434     151696.37145
Residual          2065        671465.46646        325.16487
F =     466.52140       Signif F =  .0000


------------------------------------------ Variables in the Equation ------------------------------------------

Variable              B        SE B       Beta   Correl Part Cor  Partial  Tolerance        VIF          T  Sig T
U              7.579367     .272420    .544260  .621000  .472683  .522162    .754271      1.326     27.822  .0000
X               .085288     .018531    .081889  .262000  .078191  .100764    .911720      1.097      4.602  .0000
Q               .189193     .030501    .118104  .377000  .105381  .135244    .796153      1.256      6.203  .0000
(Constant)   -79.805965    3.506456                                                                -22.760  .0000


------------------------ Variables not in the Equation ------------------------

Variable     Beta In  Partial  Tolerance        VIF  Min Toler         T  Sig T
V            .034126  .038740    .768120      1.302    .735815     1.761  .0783
M            .023996  .029068    .874679      1.143    .740276     1.321  .1866


End Block Number   2   POUT =     .051 Limits reached.
```

As the "Variables in the equation" shows, X has the smallest semipartial correlation; HOWEVER, the effect of X is significant at the .05 level, so we do not want to exclude it. Ergo, STOP NOW!

MODERN SPSS OUTPUT. SPSS has changed its formatting in recent years. The new style is prettier but makes it harder to do the sort of step by step discussion given above. Here is what modern output looks like. You should have little trouble seeing how the information in both sets of printouts is the same.

Stepwise Regression Example

```
* Adapted from Education, Occupation, and Earnings:
* Achievement in the Early Career, by William H. Sewell
* and Robert M. Hauser.  1975, Academic Press.
* See especially pages 72 and 79.
MATRIX DATA VARIABLES=ROWTYPE_ V M X I Q U W Y.
BEGIN DATA.
MEAN   10.200  10.410  33.110   6.443  99.860  13.220  42.110   7.538
STDDEV  3.008   2.848  22.410   3.141  14.570   1.676  23.340   2.589
N      2069    2069    2069     2069   2069    2069    2069     2069
CORR    1.000
CORR    0.524   1.000
CORR    0.425   0.288   1.000
CORR    0.320   0.238   0.457   1.000
CORR    0.254   0.211   0.184   0.184   1.000
CORR    0.316   0.260   0.291   0.279   0.448   1.000
CORR    0.263   0.211   0.262   0.237   0.377   0.621   1.000
CORR    0.093   0.083   0.102   0.184   0.162   0.203   0.220   1.000
END DATA.
VARIABLE LABELS
   V "Father's educational attainment"
   M "Mother's educational attainment"
   X "Status of F's Occ when son grad from HS"
   I "Parent's average income, 1957-1960"
   Q "Son's score on Henmon-Nelson IQ test"
   U "Son's educational attainment"
   W "Son's 1964 Occupation (Duncan SEI)"
   Y "Son's average earnings, 1965-1967" .

Subtitle  "Stepwise forward selection" .
```

Stepwise Regression Example

Stepwise forward selection

```
REGRESSION  /MATRIX IN(*)
            /DESCRIPTIVES DEF
            /VARIABLES V M X Q U W
            /CRITERIA=POUT(.051)
            /STATISTICS DEF CHANGE OUTS ZPP CI SES TOL
            /DEPENDENT W
            /METHOD FORWARD V M X Q U.
```

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
V Father's educational attainment	10.200000	3.0080000	2069
M Mother's educational attainment	10.410000	2.8480000	2069
X Status of F's Occ when son grad from HS	33.110000	22.4100000	2069
Q Son's score on Henmon-Nelson IQ test	99.860000	14.5700000	2069
U Son's educational attainment	13.220000	1.6760000	2069
W Son's 1964 Occupation (Duncan SEI)	42.110000	23.3400000	2069

Correlations

		V Father's educational attainment	M Mother's educational attainment	X Status of F's Occ when son grad from HS	Q Son's score on Henmon-Nelson IQ test	U Son's educational attainment	W Son's 1964 Occupation (Duncan SEI)
Pearson Correlation	V Father's educational attainment	1.000	.524	.425	.254	.316	.263
	M Mother's educational attainment	.524	1.000	.288	.211	.260	.211
	X Status of F's Occ when son grad from HS	.425	.288	1.000	.184	.291	.262
	Q Son's score on Henmon-Nelson IQ test	.254	.211	.184	1.000	.448	.377
	U Son's educational attainment	.316	.260	.291	.448	1.000	.621
	W Son's 1964 Occupation (Duncan SEI)	.263	.211	.262	.377	.621	1.000

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	U Son's educational attainment	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
2	Q Son's score on Henmon-Nelson IQ test	.	Forward (Criterion: Probability-of-F-to-enter <= .050)
3	X Status of F's Occ when son grad from HS	.	Forward (Criterion: Probability-of-F-to-enter <= .050)

a. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics				
					R Square Change	F Change	df1	df2	Sig. F Change
1	.621[a]	.386	.385	18.2985633	.386	1297.482	1	2067	.000
2	.631[b]	.398	.397	18.1201897	.012	41.895	1	2066	.000
3	.636[c]	.404	.403	18.0323286	.006	21.182	1	2065	.000

a. Predictors: (Constant), U Son's educational attainment

b. Predictors: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test

c. Predictors: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test, X Status of F's Occ when son grad from HS

ANOVA[d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	434445.635	1	434445.635	1297.482	.000[a]
	Residual	692108.946	2067	334.837		
	Total	1126554.581	2068			
2	Regression	448201.508	2	224100.754	682.524	.000[b]
	Residual	678353.073	2066	328.341		
	Total	1126554.581	2068			
3	Regression	455089.114	3	151696.371	466.521	.000[c]
	Residual	671465.466	2065	325.165		
	Total	1126554.581	2068			

a. Predictors: (Constant), U Son's educational attainment

b. Predictors: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test

c. Predictors: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test, X Status of F's Occ when son grad from HS

d. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients				95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	Std. Error	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-72.217	3.199			-22.573	.000	-78.492	-65.943					
	U Son's educational attainment	8.648	.240	.621	.017	36.021	.000	8.177	9.119	.621	.621	.621	1.000	1.000
2	(Constant)	-81.795	3.497			-23.392	.000	-88.652	-74.938					
	U Son's educational attainment	7.877	.266	.566	.019	29.621	.000	7.355	8.398	.621	.546	.506	.799	1.251
	Q Son's score on Henmon-Nelson IQ test	.198	.031	.124	.019	6.473	.000	.138	.258	.377	.141	.111	.799	1.251
3	(Constant)	-79.806	3.506			-22.760	.000	-86.683	-72.929					
	U Son's educational attainment	7.579	.272	.544	.020	27.822	.000	7.045	8.114	.621	.522	.473	.754	1.326
	Q Son's score on Henmon-Nelson IQ test	.189	.031	.118	.019	6.203	.000	.129	.249	.377	.135	.105	.796	1.256
	X Status of F's Occ when son grad from HS	.085	.019	.082	.018	4.602	.000	.049	.122	.262	.101	.078	.912	1.097

a. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Excluded Variables[d]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics Tolerance	VIF	Minimum Tolerance
1	V Father's educational attainment	.074[a]	4.097	.000	.090	.900	1.111	.900
	M Mother's educational attainment	.053[a]	2.982	.003	.065	.932	1.073	.932
	X Status of F's Occ when son grad from HS	.089[a]	4.956	.000	.108	.915	1.093	.915
	Q Son's score on Henmon-Nelson IQ test	.124[a]	6.473	.000	.141	.799	1.251	.799
2	V Father's educational attainment	.060[b]	3.301	.001	.072	.884	1.131	.756
	M Mother's educational attainment	.041[b]	2.313	.021	.051	.921	1.086	.771
	X Status of F's Occ when son grad from HS	.082[b]	4.602	.000	.101	.912	1.097	.754
3	V Father's educational attainment	.034[c]	1.761	.078	.039	.768	1.302	.736
	M Mother's educational attainment	.024[c]	1.321	.187	.029	.875	1.143	.740

a. Predictors in the Model: (Constant), U Son's educational attainment

b. Predictors in the Model: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test

c. Predictors in the Model: (Constant), U Son's educational attainment, Q Son's score on Henmon-Nelson IQ test, X Status of F's Occ when son grad from HS

d. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

```
Subtitle  "Stepwise backward selection".
```

Stepwise Regression Example

Stepwise backward selection

```
REGRESSION  /MATRIX IN(*)
            /DESCRIPTIVES DEF
            /VARIABLES V M X Q U W
            /CRITERIA=POUT(.051)
            /STATISTICS DEF CHANGE OUTS ZPP CI SES TOL
            /DEPENDENT W
            /METHOD ENTER V M X Q U
            /METHOD BACKWARDS .
```

Regression

Descriptive Statistics

	Mean	Std. Deviation	N
V Father's educational attainment	10.200000	3.0080000	2069
M Mother's educational attainment	10.410000	2.8480000	2069
X Status of F's Occ when son grad from HS	33.110000	22.4100000	2069
Q Son's score on Henmon-Nelson IQ test	99.860000	14.5700000	2069
U Son's educational attainment	13.220000	1.6760000	2069
W Son's 1964 Occupation (Duncan SEI)	42.110000	23.3400000	2069

Correlations

		V Father's educational attainment	M Mother's educational attainment	X Status of F's Occ when son grad from HS	Q Son's score on Henmon-Nel son IQ test	U Son's educational attainment	W Son's 1964 Occupation (Duncan SEI)
Pearson Correlation	V Father's educational attainment	1.000	.524	.425	.254	.316	.263
	M Mother's educational attainment	.524	1.000	.288	.211	.260	.211
	X Status of F's Occ when son grad from HS	.425	.288	1.000	.184	.291	.262
	Q Son's score on Henmon-Nelson IQ test	.254	.211	.184	1.000	.448	.377
	U Son's educational attainment	.316	.260	.291	.448	1.000	.621
	W Son's 1964 Occupation (Duncan SEI)	.263	.211	.262	.377	.621	1.000

Variables Entered/Removed[b]

Model	Variables Entered	Variables Removed	Method
1	U Son's educational attainment, M Mother's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment[a]	.	Enter
2	.	M Mother's educational attainment	Backward (criterion: Probability of F-to-remove >= .051).
3	.	V Father's educational attainment	Backward (criterion: Probability of F-to-remove >= .051).

a. All requested variables entered.

b. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics: R Square Change	F Change	df1	df2	Sig. F Change
1	.636[a]	.405	.404	18.0258459	.405	280.811	5	2063	.000
2	.636[b]	.405	.404	18.0231567	.000	.384	1	2063	.535
3	.636[c]	.404	.403	18.0323286	-.001	3.102	1	2064	.078

a. Predictors: (Constant), U Son's educational attainment, M Mother's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment

b. Predictors: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment

c. Predictors: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test

ANOVA[d]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	456221.678	5	91244.336	280.811	.000[a]
	Residual	670332.903	2063	324.931		
	Total	1126554.581	2068			
2	Regression	456096.841	4	114024.210	351.023	.000[b]
	Residual	670457.740	2064	324.834		
	Total	1126554.581	2068			
3	Regression	455089.114	3	151696.371	466.521	.000[c]
	Residual	671465.466	2065	325.165		
	Total	1126554.581	2068			

a. Predictors: (Constant), U Son's educational attainment, M Mother's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment

b. Predictors: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment

c. Predictors: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test

d. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients		t	Sig.	95% Confidence Interval for B		Correlations			Collinearity Statistics	
		B	Std. Error	Beta	Std. Error			Lower Bound	Upper Bound	Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	-80.759	3.560			-22.685	.000	-87.741	-73.777					
	V Father's educational attainment	.220	.167	.028	.021	1.319	.187	-.107	.547	.263	.029	.022	.624	1.602
	M Mother's educational attainment	.102	.165	.012	.020	.620	.535	-.221	.426	.211	.014	.011	.711	1.407
	X Status of F's Occ when son grad from HS	.072	.020	.069	.019	3.605	.000	.033	.111	.262	.079	.061	.789	1.268
	Q Son's score on Henmon-Nelson IQ test	.182	.031	.114	.019	5.914	.000	.122	.242	.377	.129	.100	.783	1.277
	U Son's educational attainment	7.491	.277	.538	.020	27.090	.000	6.948	8.033	.621	.512	.460	.732	1.367
2	(Constant)	-80.445	3.523			-22.832	.000	-87.355	-73.536					
	V Father's educational attainment	.265	.150	.034	.019	1.761	.078	-.030	.560	.263	.039	.030	.768	1.302
	X Status of F's Occ when son grad from HS	.073	.020	.070	.019	3.653	.000	.034	.112	.262	.080	.062	.792	1.263
	Q Son's score on Henmon-Nelson IQ test	.183	.031	.114	.019	5.955	.000	.123	.243	.377	.130	.101	.785	1.274
	U Son's educational attainment	7.503	.276	.539	.020	27.218	.000	6.963	8.044	.621	.514	.462	.736	1.359
3	(Constant)	-79.806	3.506			-22.760	.000	-86.683	-72.929					
	X Status of F's Occ when son grad from HS	.085	.019	.082	.018	4.602	.000	.049	.122	.262	.101	.078	.912	1.097
	Q Son's score on Henmon-Nelson IQ test	.189	.031	.118	.019	6.203	.000	.129	.249	.377	.135	.105	.796	1.256
	U Son's educational attainment	7.579	.272	.544	.020	27.822	.000	7.045	8.114	.621	.522	.473	.754	1.326

a. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)

Excluded Variables[c]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics		
						Tolerance	VIF	Minimum Tolerance
2	M Mother's educational attainment	.012[a]	.620	.535	.014	.711	1.407	.624
3	M Mother's educational attainment	.024[b]	1.321	.187	.029	.875	1.143	.740
	V Father's educational attainment	.034[b]	1.761	.078	.039	.768	1.302	.736

a. Predictors in the Model: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test, V Father's educational attainment

b. Predictors in the Model: (Constant), U Son's educational attainment, X Status of F's Occ when son grad from HS, Q Son's score on Henmon-Nelson IQ test

c. Dependent Variable: W Son's 1964 Occupation (Duncan SEI)