In our study of MLR analyses to this point we have considered two different strategic approaches with regard to inclusion of the IVs in the regression model.

**Simultaneous strategy:**
Given $k$ IVs, we include all IVs in the model simultaneously. We obtain a regression equation expressing $Y$ as a function of the $k$ IVs, along with the squared multiple correlation $R^2_{Y.12...k}$, and a variety of additional information.

**Stepwise strategy:**
Given $k$ IVs, we seek to select a subset of those IVs for inclusion in the model, such that IVs that are included each make a significant unique contribution to explaining variance in $Y$, and IVs not included would not make such a contribution.

The simultaneous strategy is useful in studies focusing either on *prediction* or *explanation*. The stepwise strategy is useful only in *prediction* studies when the goal is to obtain an optimally efficient prediction model.
We now consider some alternative strategies, beginning with an approach called *hierarchical* regression.

**Hierarchical regression**

Given $k$ IVs, suppose that we have an *a priori* basis for defining an ordering or sequence of these IVs.

The ordering defines the sequence in which the IVs will be entered into a regression model. Criteria and approaches for establishing such an ordering, along with examples, will be discussed shortly. For now, assume that the IVs can be placed in a justifiable sequence denoted $X_1, X_2, X_3, \ldots, X_k$.

Hierarchical regression analysis then proceeds via a sequence of separate regression analyses. At each step in the sequence the next IV in the ordering is introduced to the model. And at each step we assess the increment in $R^2$ over the previous step.

**Step 1:** Enter $X_1$ into model.

- Equation: $\hat{Y} = B_0 + B_1 X_1$
- SMC: $R^2_{Y,1}$
- Interpret effect size and carry out inferential analyses.
Step 2: Enter $X_2$ into model.
Equation: $\hat{Y} = B_0 + B_1X_1 + B_2X_2$
SMC: $R^2_{Y.12}$
Increment: $R^2_{Y.12} - R^2_{Y.1}$
Note that this increment is equivalent to $sr^2_2$ removing the effect of $X_1$.

Interpret this effect size and test significance. Significance test for increment is equivalent to test of $B_2$.

Step 3: Enter $X_3$ into model.
Equation: $\hat{Y} = B_0 + B_1X_1 + B_2X_2 + B_3X_3$
SMC: $R^2_{Y.123}$
Increment: $R^2_{Y.123} - R^2_{Y.12}$
Note that this increment is equivalent to $sr^2_3$ removing the effect of $X_1$ and $X_2$.

Interpret this effect size and test significance. Significance test for increment is equivalent to test of $B_3$.

This process continues until step $k$:
Step $k$: Include all $k$ IVs in the model. Enter $X_k$ into model.

Equation: $\hat{Y} = B_0 + B_1X_1 + B_2X_2 + \cdots + B_kX_k$

SMC: $R^2_{Y.12..k}$

Increment: $R^2_{Y.12..k} - R^2_{Y.12..(k-1)}$

Note that this increment is equivalent to $sr^2_k$ removing the effect of $X_1, X_2, \ldots, X_k$.

Interpret this effect size and test significance. Significance test for increment is equivalent to test of $B_k$.

This last step is equivalent to a full simultaneous analysis.

Note that there is no consideration of stopping prior to step $k$. Even if an increment in $R^2$ is not significant at a given step, the process continues. Increments at later steps may be larger than increments at earlier steps.

The information at each step provides a type of ordered partitioning of the variance accounted for. That is, the increments at each step will sum to the total $R^2$. 
Critical point: The pattern of results is highly dependent on the ordering of the IVs. If the ordering is changed, the entire pattern of results may change substantially. Thus, the ordering of the IVs must be carefully considered and justified.

Determining the order of the IVs

General principle: The ordering of the IVs in hierarchical regression is determined on an *a priori* basis. It is not data-driven, not based on any analysis of sample data. It can and should be determined before any data are collected or analyzed.

Criteria for determining the order of the IVs:

1) *Causal priority*: Based on prior theory and logic it may be possible to establish a causal sequence of relationships among the variables. Suppose we hypothesize a sequence of the form

\[ X_1 \rightarrow X_2 \rightarrow X_3 \rightarrow X_4 \rightarrow Y \]

Then that hypothesized sequence defines the ordering of the IVs for the steps of the hierarchical regression. The IVs are entered in the sequence defined by the causal order.
In general, when the order of the IVs is defined for hierarchical regression, it should be the case that no IV is influenced by an IV that is entered at a later step.

Example: Study of attitude-behavior relationships. 
DV: Future behavior
IVs: Prior behavior, attitude about behavior, intentions to carry out behavior

Causal sequence:
Prior Beh. $\rightarrow$ Attitude $\rightarrow$ Intentions $\rightarrow$ Future Beh.

The variables should be entered in this sequence for hierarchical regression steps.

2) *Temporal precedence:*
This criterion is often related to or equivalent to causal priority. In general, if IVs occur in some natural progression in time, then the ordering for hierarchical regression should not violate that progression.

Example:
DV: Salary
IVs: gender, age, job performance
The criterion of temporal precedence would dictate the ordering gender, age, job performance.
3) Research relevance

The primary questions and goals of the research project may involve a focus on certain IVs with less interest in effects of other IVs. If the IVs can be ordered according to research relevance, then that ordering can define the steps of a hierarchical regression, with the most relevant IVs entered first.

Another sort of research question that could dictate an ordering of IVs: Given existing knowledge of the use of certain IVs to explain a DV, we may ask whether a new IV provides additional explanatory information. In such a case, the hierarchical regression would introduce the new IV at the last step.

4) Analyzing change in DV

Consider a research design where a variable of interest, call it $W$, is measured at two points in time. Call the two measures $W_{pre}$ and $W_{post}$. Suppose there is some variable $V$ that we believe may be useful to predict change in $W$.

In planning regression analyses in such a design it would be tempting to define the DV as the amount of change in $W$: $W_{change} = W_{post} - W_{pre}$

Then to use $V$ as a predictor of $W_{change}$.
However, such analyses are very problematic because the change score ($W_{\text{change}}$) tends to have a substantial negative correlation with the pre score ($W_{\text{pre}}$).

As a result, any analysis of $W_{\text{change}}$ is potentially misleading because change is not independent of initial status.

The tools of hierarchical regression can be used to resolve this problem. Consider a hierarchical regression analysis where:

- DV $Y$ is $W_{\text{post}}$
- IVs are entered in order $X_1=W_{\text{pre}}$, $X_2=V$.

The procedure described earlier for the steps of hierarchical regression would then be followed:

**Step 1:** Enter $X_1$ into model.

Equation:  \[ \hat{Y} = B_0 + B_1 X_1 \]

SMC:  \[ R_{Y.1}^2 \]

Interpret effect size and carry out inferential analyses.

This step predicts the post score from the pre score. Thus, at subsequent steps, the effect of the pre score has been partialed out.
Step 2: Enter $X_2$ into model.
Equation: $\hat{Y} = B_0 + B_1 X_1 + B_2 X_2$
SMC: $R^2_{Y.12}$
Increment: $R^2_{Y.12} - R^2_{Y.1}$
Note that this increment is equivalent to $sr^2_2$ removing the effect of $X_1$.

Interpret this effect size and test significance. Significance test for increment is equivalent to test of $B_2$.

This step provides us information as to how well $X_2$ accounts for $Y$, above and beyond the effect of the pre score.

Example:
$W$ is a measure of depression at two occasions 3 months apart.
$V$ is a dichotomous (0,1) variable indicating whether or not the individual received psychotherapy in the intervening period.
The research question is whether presence or absence of therapy affected change in depression.

An inappropriate approach:
DV: $W_{change} = W_{post} - W_{pre}$ (change in depression)
IV: $V$ (psychotherapy)
Regression analysis predicting $W_{\text{change}}$ from $V$ will be misleading because change in depression is negatively correlated with initial status on depression.

To overcome this “contamination” we employ a hierarchical sequence of regression models where:

- **DV**: $Y = W_{\text{post}}$ (depression at second occasion)
- **IVs**: $X_1 = W_{\text{pre}}$ (depression at first occasion)
  $X_2 = V$ (psychotherapy)

Step 1 will indicate the effect of pre score on post score, and partial out that effect.

Step 2 will indicate the effect of therapy on post score after removing the effect of pre score.

It should be kept in mind that there may be more IVs involved in such a design. In such situations, the pre score should be entered at Step 1, and other IVs entered subsequently in an appropriate order.

For more extensive designs and research questions about change over time, there is an extensive literature on methods for analysis of change, with a focus on representing and explaining individual differences in patterns of change.
Issues in determining the sequence of IVs

1) Justification for ordering of IVs:
Whenever a hierarchical regression analysis is being considered, it must be kept in mind, as mentioned earlier, that the pattern of results depends on the selected ordering of the IVs. In turn, interpretations and conclusions may be greatly affected by that ordering.

Great care must be taken in determining the ordering, and clear justification must be provided.

2) Ambiguous ordering:
In some (many?) cases, the ordering may be ambiguous. There may be more than one justifiable ordering, maybe several. In such cases, it is appropriate to try alternative orderings and compare results. One can evaluate the degree to which conclusions and interpretations are affected.

3) No basis for ordering:
Ordering of the IVs is not always possible or desired. It should only be considered when there is a natural or justifiable ordering and when a hierarchical regression approach would provide useful information relevant to the research questions.
4) Variables occurring in natural subsets:
It is common for IVs to be organizable into natural subsets. For example, in studying the attitude-behavior relation, we may have multiple measures of prior behavior, multiple measures of attitude, etc.

In such cases, it may be possible to order the sets of IVs, but not the separate IVs within the sets. We will study the use of regression methods for analysis of sets of IVs.

**Differences between stepwise and hierarchical regression**

Both methods involve a sequence of regression analyses, entering one additional IV at each step. Increments in $R^2$ over the previous step are of interest.

Major differences:
Stepwise regression:
Data-driven; order of IVs is determined by data. Useful in some prediction studies. No substantive interpretation.

Hierarchical regression:
Theory driven; order of IVs is determined *a priori*. Useful in prediction and explanation studies. Substantive interpretation important.