Examples of Multiple regression: R=1.0 and R<1.0
Confidence & Concern

Variables included for analysis were first-year GPA and the Analytic, Quantitative, and Verbal subscales of the GRE. Table 1 shows the univariate statistics for each variable and their correlation, as well as the resulting regression model (Formula 1)

SPSS Code to Simulate regression data for 100 cases:

```
compute x1 = trunc(normal(10)+50).
compute x2 = trunc(abs(normal(2)+4)).
compute x3 = trunc(normal(100)+500).
```

← computed three "predictors"

```
X1          48.930    9.585
X2           3.770    1.969
X3         491.910   93.448
```

← variables will not have exact mean & std
- are a sample from the specified population
- have had decimals truncated

#1 Confidence Builder -- Properly specified model recaptures errorless criterion

```
compute y1 = 2.0*x1 + 18*x2 + .25*x3.
```

← errorless model

```
    Variable       Mean       Std
       Y1         288.698   43.157


dependent variable..   y1


     Multiple R          1.00000    Analysis of Variance
     R Square            1.00000
     Adjusted R Square   1.00000       F is undefined (no error variation)
     Standard Error      0.00000


     ------------------ Variables in the Equation ------------------
     Variable              B        SE B       Beta         T  Sig T
     X3               .25000 5.16275E-10     .54133         .    .
     X1              2.00000 5.03660E-09     .44418         .    .
     X2             18.00000 2.46530E-08     .82112         .    .
     (Constant) 4.010681E-14 3.80284E-07                    .    .
```

Notice: 1) The regression weights exactly re-capture the construction of Y1.
2) The constant = 0 means that there is no adjustment necessary for the mean of y' to equal mean y.

#2 Confidence Builder - Rescaling y produces predictable changes in weights

```
     compute y2 = y1/5.                          ← linear transform of errorless model

    Variable       Mean       Std
       Y2          57.740    8.631    ← mean & std  1/5 or Y1

dependent variable..   y2


     Multiple R          1.00000       Analysis of Variance
     R Square            1.00000
     Adjusted R Square   1.00000          F is undefined (no error variation)
     Standard error      0.00000



     ------------------ Variables in the Equation ------------------ -
     Variable               B        SE B       Beta         T  Sig T
     X3                .05000      .00000     .54133      .     .
     X1                .40000      .00000     .44418      .
     X2               3.60000      .00000     .82112      .     .
     (Constant) -1.35551E-14      .00000                 .     .
```

Notice: 1) linear transformation does not change R^2
2) linear transformation does not change Betas
3) linear transformation predictable changes raw regression weights
- bs are 1/5 or 1st model – fitting a Y2 with 1/5 the mean & standard deviation

#3 How accurate are the regression weights of an underspecified model ???

Our models are (almost?) always underspecified. So, the variables that are in the model are "trying to make up for" or "trying to do without" the variables that have been left out of the model.

dependent variable.. y1

Multiple R	**.84280**	←	**Model not "perfect" because is "underspecified"**
R Square	**.71031**		
Adjusted R Square	**.70434**		
Standard Error	**23.46633**	←	**Imperfect prediction leads to residuals**

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

Variable	B	SE B	Beta	T	Sig T	
X2	**16.75495**	**1.20550**	**.76433**	**13.899**	**.0000**	← weights not "correct"
X1	**3.02890**	**.24762**	**.45059**	**8.194**	**.0000**	← weights not "correct"
(Constant)	**126.25709**	**13.61083**		**9.276**	**.0000**	← adjusting y' mean

Notice: 1) Model less accurate because of underspcification
2) Raw and standardized weights are "off" -- but not horribly

Ran 1000 simulations to see the range of b and Beta values …

X2	B (18) 12.232 – 22.321	Beta (.82)	.53 – .96
X1	B (2) .145 – 4.51	Beta (.44)	.35 - .67

#4 How accurate are the regression weights of an properly specified model – with error ???

compute y3 = trunc(y1 + normal(10)). ← error added to model

Usually the criterion is measured "with error" and so, isn't perfectly predictable

dependent variable.. y3

Multiple R .96889 ← Error added to Y3 reduces fit f regression model
R Square .93874
Adjusted R Square .93683
Standard Error 11.15072

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

Variable	B	SE B	Beta	T	Sig T
X3	.25737	.01206	.54211	21.337	.0000
X1	1.86345	.11767	.40259	15.836	.0000
X2	17.97213	.57597	.79754	31.203	.0000
(Constant)	1.17546	8.88463		.132	.8950

compute y4 = trunc(y1 + normal(50)). ← more error added to model

dependent variable.. y3

Multiple R .63446 ← more error leads to poorer fit
R Square .40254
Adjusted R Square .38387
Standard Error 48.54283 ← more residuals

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

Variable	B	SE B	Beta	T	Sig T
X3	.21521	.05251	.32519	4.098	.0001
X1	.73573	.51226	.11403	1.436	.1542
X2	18.23383	2.50740	.58046	7.272	.0000
(Constant)	70.85608	38.67778		1.832	.0701

Ran 1000 simulations to see the range of b and Beta values …

X3	B (.25) -1.1 – 3.6	Beta (.54) -.23 – .83
X1	B (2) -.214 – 5.21	Beta (.44) -.19 – .78
X2	B (18) 12.232 – 22.321	Beta (.82) .23 – .97

#5 Reality #1 - Criterion measured with Error & an Under-specified model

Equation Number 6 Dependent Variable.. Y4

Multiple R	.54589
R Square	.29800
Adjusted R Square	.28353
Standard Error	52.34670

dependent variable.. y4

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

Variable	B	SE B	Beta	T	Sig T
X2	17.16205	2.68914	.54634	6.382	.0000
X1	.76061	.55236	.11788	1.377	.1717
(Constant)	179.54230	30.36188		5.913	.0000

Notice: Again, the regression weights "adjust" in an effort to reproduce a variable that is composed of variables other than the predictors + error. Also, again there is "differential accuracy" of those adjusted weights, with the contribution of x1 being under estimated.

Ran 1000 simulations to see the range of b and Beta values …

X2	B (18) 11.446 – 25.329	Beta (.82) .483 – .98
X1	B (2) -2.156 – 7.495	Beta (.44) -.15 - .587

#6 Reality #2 - Criterion measured with Error & an Over-specified model

```
compute x4 = trunc(normal(25)+100).
```
← build a new & independent predictor

```
Dependent Variable..   Y4

   Multiple R           .54589
   R Square             .29800
   Adjusted R Square    .28353
   Standard Error     52.34670



   ------------------ Variables in the Equation ------------------
    Variable              B        SE B       Beta         T  Sig T
   X3                 .28521      .04251     .41519     4.098  .0001
   X2               21.56705     2.68914     .44634     6.382  .0000
   X1                 .02061      .55236     .01788     1.377  .1717
   X4                3.12645      .56445     .15228     5.538  .0000   ←  !!!!!!!
   (Constant)      179.54230    30.36188                5.913  .0000
```

Notice: Again, the regression weights "adjust" in an effort to reproduce a variable that is composed of variables other than the predictors + error.

Ran 1000 simulations to see the range of b and Beta values …

X3	B (.25)	-0.1 – 3.9	Beta (.54)	-.13 – .88
X2	B (18)	9.446 – 25.329	Beta (.82)	.283 – .98
X1	B (2)	-5.136 – 9.235	Beta (.44)	-.15 - .587
X4	B (0)	-6.476 – 8.395	Beta (.0)	-.35 - .487

#7 Reality #3 - Criterion measured with Error & a Mis-specified model

```
Dependent Variable..   Y4

   Multiple R           .54589
   R Square             .29800
   Adjusted R Square    .28353
   Standard Error     52.34670



   ------------------ Variables in the Equation ------------------
    Variable              B        SE B       Beta          T  Sig T

   X2              21.56705     2.68914     .44634     6.382  .0000
   X1                .02061      .55236     .01788     1.377  .1717
   X4              -2.12645      .56445     .15228    -4.538  .0000   ←  !!!!!!!
   (Constant)    179.54230    30.36188                 5.913  .0000
```

Notice: Again, the regression weights "adjust" in an effort to reproduce a variable that is composed of variables other than the predictors + error.

Ran 1000 simulations to see the range of b and Beta values …

X2	B (18)	8.34 – 25.329	Beta (.82)	.381 – .967
X1	B (2)	-3.146 – 5.213	Beta (.44)	-.20 - .467
X4	B (0)	-5.296 – 7.695	Beta (.0)	-.38 - .527