Responsible Partialling

John D. Morris **Mary Lieberman**
Florida Atlantic University

Partialling correlations (partial, semipartial or bipartial) is built upon regression equation(s). It is imperative that the assumptions of those models be considered. Those assumptions rest on the OLS residuals of those models (e.g., linearity, homoscedasticity, normality, independence, absence from data with extreme influence). Partial correlations of all types depend on residuals, but consideration of fit of data in respect to those residuals is seldom done. Although this can be accomplished with repeated subsequent analyses, partialling in commercial software does not directly render these assumption diagnostics. As these are mandatory for use of least squares, examples are produced and an Excel program that automatically performs an exhaustive set of diagnostic analyses and plots for partialling is offered.

P**arsing effects** is a fundamental part of science, from physics (**partial** pressure) to statistics (**partial** correlation). Although the same model fit questions arise from all mathematical models (e.g., maximum likelihood), only least squares will be considered here. Although partialling is always a function of residuals, typical commercial software [SAS, SPSS, Stata] relies on formulas, the residuals are never calculated. [Minitab is an outlier in that it has no partialling program.] The difficulty with this is that the usual assumptions of least squares model(s), which rely on residuals, are not tested. In addition, partialling always reduces variance, so it important to see the degree of that reduction and how it manifests in a scatterplot of the relevant residuals.

Perspectives - Theoretical Framework

All least squares partialling is done by a simple procedure. First, predict each variable from the variable we which to "partial out" with least squares. Assume two variables, **t** and **u**, and the "controlled for" variable is **v**. In standard score form this means creating the equations below and transformations for each subject.

$$\hat{t}_I = r_{tv}v_i,\ t_{i\ res} = t_i\text{-}\hat{t}_i, \text{ with "res" for residual.}$$

This subtraction gives the part of t that is not predictable from v, thus $r(v,t_{res})=0$. And,

$$\hat{u}_I = r_{uv}v_i,\ u_{i\ res} = u_i\text{-}\hat{u}_i,$$

where, again, this gives the part of u that is not predictable from v, thus $r(v,u_{res}) = 0$.
Thus, we now have two residuals, one from t and one from u, both uncorrelated with v. We now correlate those two residuals for a partial correlation.

$$r_{partial} = r_{tu.v} = r(t_{res},u_{res}).$$

That correlation is spoken of as the correlation between t and u, "controlling" for v. "Controlling" has a very limited meaning -- only that the two residuals are now uncorrelated with v. Although not the point of the present paper, much has been written about appropriate caution in interpretation of such a partial correlation (e.g., Lynam, Hoyle, & Newman, 2006). Remembering that these correlations only regard the linear relationship of v with t and u, we should understand that if there are other forms of relationship (e.g., quadratic) they remain. Moreover, one can't claim that this is the same as the relationship between t and u for subjects who have the same v score.

As the fundamental mathematics behind partialling is OLS regression, one can do the same with multiple zs. As well, one can elect to only partial from one variable (say t), leaving u as raw. This is a **semipartial** or **part** correlation. Tests of semipartial correlations are automatically rendered with any multiple regression program; partialling the remainder of the variables from x, but not from y, for each x in turn. And, given the flexibility of the underlying mathematics, one can choose to partial a different set of variables from x than from y. Although this is done less frequently, it is often called a **bipartial** correlation. Although the correlation of residuals from multiple regression equation(s) is the foundation of partialling, we can derive formulas that render partial correlations directly from the relevant bivariate correlations (and recursive "chaining" their application for higher order partial correlations) or use matrix algebra. Therein, avoiding calculating residuals for each subject from regression equation(s). These formulas are found in most intermediate statistics texts; for a good presentation see Cohen & Cohen (1983).

This mathematical shortcut gives us the following for first order partial and semipartial correlations:

$$r_{ut.v}=(r_{ut}-r_{uv}r_{tv})/\sqrt{((1-r_{uv}^2)(1-r_{tv}^2))}, \text{ and}$$

$$r_{u(t.v)}=(r_{ut}-r_{uv}r_{tv})/\sqrt{(1-r_{tv}^2)}).$$

However, the problem with this approach is that we apparently forget where the residuals came from. As with all multiple regression models, it is mandatory to check the assumptions of the multiple regression model(s) that produce the residuals. Thus, consideration of linearity, homoskedasticity, independence, and normality of the residuals, as well as data points with excessive leverage or influence, is, as in any regression model, our responsibility. In addition, collinearity in the regression model(s) needs to be considered. These typical regression examinations are essentially never done in partialling. [It is worthwhile to note that diagnostics for a regression model including **all** constituent variables is not the same – wrong model.] To accomplish these diagnostics, one would need to run regression analyses and diagnostics separately for each regression model used to calculate residuals– two in the case of partial or bipartial correlation, and one in the case of semipartial correlation.

Method

As with all model assumptions, rendering such regression diagnostic information should be part of commercial partialling software, but it is not. The purpose of this paper is to point out this necessity, but also to offer a flexible Excel program that automatically yields all of these diagnostics for any partial, semipartial, or bipartial correlation requested. As well, the reduction in variance due to partialling must be considered, and a scatterplot of the original and partialled residuals should be compared. One can then see what partialling did at the subject level. SAS (an option in proc "corr") is the only commercial software platform known that offers this scatterplot comparison of original and partialled data.

The Excel program used herein has a very simple interface for specifying any partial, semipartial, or bipartial correlation one wishes among variables entered. [Note that the Excel statistical "Analysis Toolpak" that has received some flack in the literature (e.g., McCullough & Berry, but see amelioration with more recent versions, Melard, 2014) is not used – all programming is original. Moreover, with a very large set of examples, an exact match with results obtained with the multiple runs necessary to obtain the same in SPSS, SAS, or Systat is always obtained.] Correlations from partialling are calculated from actual residuals with relevant residual plots ($\hat{t}$ with $t_{res,}$ and $\hat{u}$ with u_{res}, each predictor with respective residuals, Index, and P-P and Q-Q normal plots), leverage and influence statistics for total t and u models as well as DFBETAS for individual predictors are all automatically generated. In addition, VIFs for each regression model (if the number of predictors is >1) are automatically rendered.

Data Source

A venerable data set from Kerlinger and Pedhazur (1973, p. 292) including **GPA**, **GRE** (**Q**uantitative and **V**erbal), the **M**iller **A**nalogies **T**est, and the **A**verage **R**ating of faculty is used to demonstrate model assumption testing that should be done, but also the program that can aid in that process.

Results

To allow variable choice and role, the program simply lists the names of all variables that were included in the Input page twice. From those two lists, all one must do is to place a "t," a "u," and "p"s (variables to partial out) under the desired variables in the respective lists. If no variables are given "p"s for either t or u, this then necessarily becomes a semipartial correlation. If the same variable(s) are given ps for the "t" and "u" variable, the result is a partial correlation and, finally, if different variables are given ps for t and u, the result is a bipartial correlation. The program parses this as a partial, semipartial, or bipartial correlation, and provides labels, and statistics appropriately.

The following screenshot shows the input to obtain the partial correlation between GRE-Q and GPA, with GRE-V partialled out of both. Given appropriate assignation of variables, results are immediately rendered. Here we see that the partial r is .471 [.126,.714], with p=.01. The CI is calculated in respect to the α entered on the Input sheet; in this case .05, thus a 95% CI.

Place a t below one variable of interest and a P below each variable to be partialled from it:

GPA	GRE-Q	GRE-V	MAT	AR
	t	p		

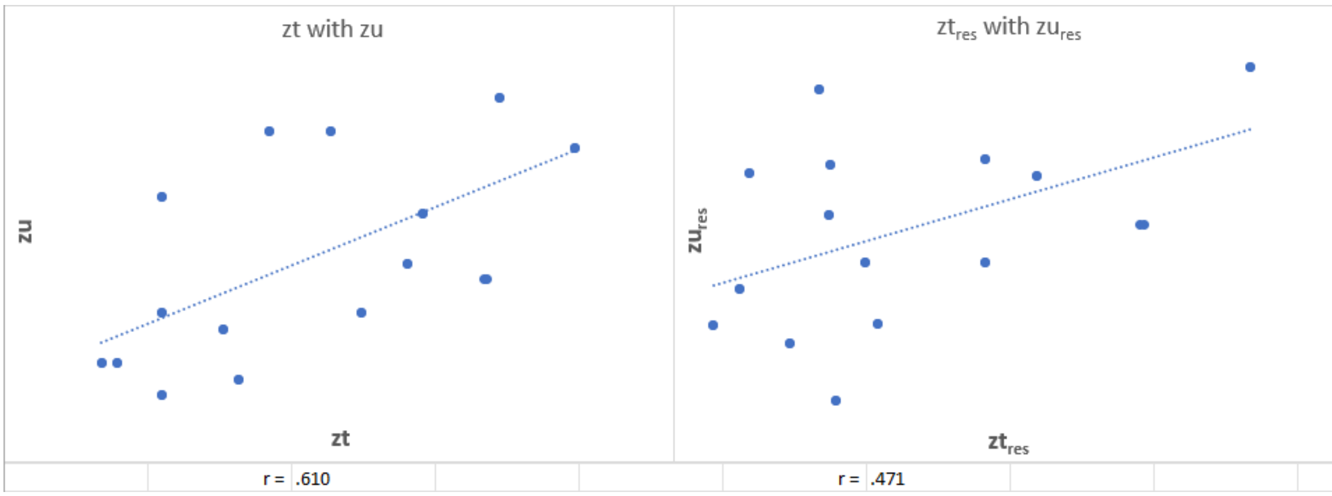
Place a u below the other variable of interest and a P below each variable to be partialled from it:

GPA	GRE-Q	GRE-V	MAT	AR
u		p		
Partial r =	**.471, [.126, .714], t(27) = 2.774, p = .010.**			

	Variance		
	Original	Partialled	% of Original
t	2367.826	1850.653	78.16%
u	.360	.238	66.19%

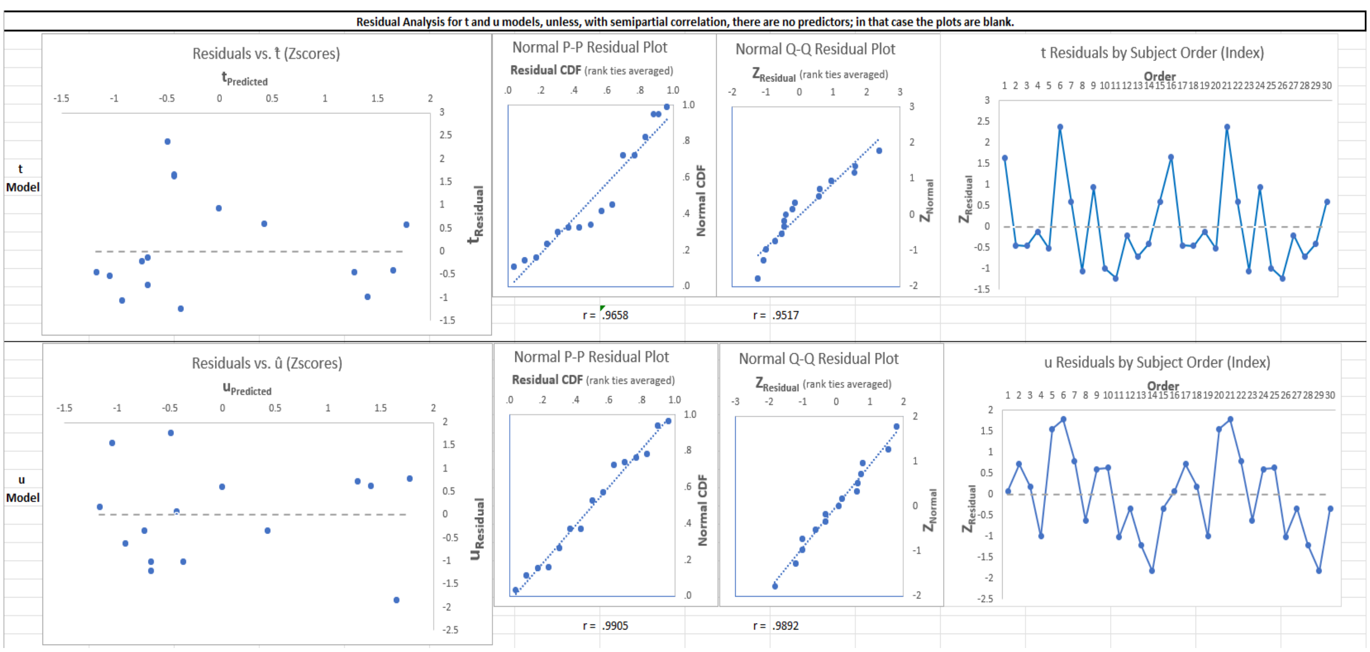
So, you can see that the reduction in variance for GRE-Q and GPA was about 78% and 66%, respectively. We believe that this is too infrequently considered in partialling.

To consider what the partialling did to the correlation between GRE-Q and GPA, we can simply ponder the respective correlations; it was .610 in raw form, and .471 after "partialling out" GRE-V. But even more information is available by considering what this did to the respective scatterplots. An "un-partialled" and "partialled" scatterplot is automatically generated to allow consideration of how partialling affected the relationship at the individual data point level. A screenshot for this example is:



One might argue that, in this case, partialling (plot to right) introduced a bit of heteroskedasticity as it appears that the spread of the GPA residuals is larger at lower levels of GRE-Q residual and less at higher levels, in contrast to the raw score scatterplot (plot to left).

The usual requisite regression residual diagnostic plots are automatically generated for each regression model. [If you have requested a semipartial correlation, the plots for the "uncontrolled" variable will be blank, as there are no residuals.] The residual plots (reduced in size from their usual portrayal for this proposal) for these data are below.



For both the GRE-Q (t) and GPA (u) residuals, there appear to be no patterns (curvilinear, etc.) as a function of predicted scores (plots to left). As well, both sets of P-P and Q-Q plots appear to satisfy the model assumptions of residual normality relatively well for both models. Finally, Index plots that are often used in consideration of independence show nothing of note.

Let's say that you would like to know the correlation between GRE-Q with GRE-V, MAT, and AR partialled out, with GPA, with no variables partialled out. This is a semi-partial correlation; the results (less produced plots to reduce space) look like this.

Place a t below one variable of interest and a P below each variable to be partialled from it:				
GPA	GRE-Q	GRE-V	MAT	AR
	t	p	p	p
Place a u below the other variable of interest and a P below each variable to be partialled from it:				
GPA	GRE-Q	GRE-V	MAT	AR
u				
Semipartial r =	**.262, [—, —], t(25) = 2.182, p = .039.**			
		Variance		
	Original	Partialled	% of Original	
t	2367.826	1550.061	65.46%	
u	.360	.360	100.00%	

Thus, the semipartial correlation (.274) is significant with α=.05. What we have is the correlation of a predictor (GRE-Q) with the variance of all other predictors (GRE-V, MAT, and AR) partialled out with a criterion (GPA). That this is equivalent to a test of the contribution of GRE-Q to the regression model predicting GPA using all four predictors; this is important illustrate to students.

Finally, let's say that you would like to know the correlation between GRE-Q with GRE-V and MAT partialled out, and GPA with GRE-V, MAT, and AR partialled out. This is a bipartial correlation. Excluding the plots again, results look like this:

Place a t below one variable of interest and a P below each variable to be partialled from it:				
GPA	GRE-Q	GRE-V	MAT	AR
	t	p	p	
Place a u below the other variable of interest and a P below each variable to be partialled from it:				
GPA	GRE-Q	GRE-V	MAT	AR
u		p	p	p
Bipartial r =	**.367, [-.015, .656], t(25) = 1.975, p = .059.**			
		Variance		
	Original	Partialled	% of Original	
t	2367.826	1837.454	77.60%	
u	.360	.154	42.81%	

Thus, given an α of .05, this bipartial correlation of .367 is not significant. From the CI, note just how close to exclusion of zero (-.015), thus significance, this correlation was; CIs are important for students to consider. Also note that nearly 80% of the variability of GRE-Q has been removed. Because there are multiple predictors in these models, an additional table is automatically triggered (as it was, but not shown, in the aforementioned semipartial model) – VIFs associated with each of these models. If no regression equation has more than one variable, this table is not produced. It is below with no VIF problems for either model.

Table 1. Leverage and Influence Statistics for the t model.

								Influence						
						Leverage			Studentized Residuals					
					PRESS or "Deleted"	"Hat" Diagonal	Mahalanobis D^2	Cook's D	Internal	External		DFFITS	COVRATIO	Welsch D
Case #	Criterion	Predicted	Residual	PRESS Predicted	Residual	>2k/n, >3k/n	Prob(χ2(p))<.001	>4/n, 1		||>2, >3	DFFIT	||>2√(k/n)	|CR-1|≥3k/n	||>3√k
	Maximum:	605.483	100.718	602.773	105.076	.171	3.995	.077	2.316	2.538	4.358	.528	1.321	2.904
	Minimum:	540.351	-49.759	542.167	-57.041	.035	.056	.000	-1.199	-1.209	-7.282	-.463	.601	-2.668
	Mean:	565.367	.000	566.355	-.989	.100	1.933	.025	-.011	.008	-.989	-.042	1.130	-.251
	SD:	22.643	42.145	22.321	45.299	.040	1.168	.023	.983	1.024	3.525	.274	.181	1.547
	Median:	554.282	-13.609	557.041	-16.175	.093	1.721	.014	-.334	-.328	-1.816	-.141	1.189	-.791
1	625.000	555.549	69.451	552.683	72.317	.040	.183	.035	1.595	1.645	2.866	.334	.866	1.836
2	575.000	595.352	-20.352	597.430	-22.430	.093	1.721	.008	-.481	-.474	-2.079	-.151	1.203	-.856
3	520.000	540.351	-20.351	542.167	-22.167	.082	1.409	.007	-.478	-.471	-1.816	-.141	1.189	-.791
4	545.000	546.141	-1.141	546.256	-1.256	.092	1.700	.000	-.027	-.026	-.116	-.008	1.233	-.048
5	520.000	547.226	-27.226	551.703	-31.703	.141	3.129	.024	-.661	-.654	-4.477	-.265	1.242	-1.542
6	655.000	554.282	100.718	549.924	105.076	.041	.236	.077	2.316	2.538	4.358	.528	.601	2.904
7	630.000	605.483	24.517	601.537	28.463	.139	3.054	.019	.595	.587	3.946	.236	1.250	1.367
8	500.000	549.759	-49.759	557.041	-57.041	.128	2.736	.070	-1.199	-1.209	-7.282	-.463	1.089	-2.668
9	605.000	564.414	40.586	562.930	42.070	.035	.056	.011	.930	.928	1.484	.177	1.053	.973
10	555.000	597.885	-42.885	602.773	-47.773	.102	2.001	.039	-1.019	-1.020	-4.889	-.344	1.109	-1.957
11	505.000	552.473	-47.473	557.334	-52.334	.093	1.727	.043	-1.122	-1.128	-4.860	-.361	1.070	-2.040
12	540.000	544.874	-4.874	545.372	-5.372	.093	1.721	.000	-.115	-.113	-.498	-.036	1.232	-.204
13	520.000	546.141	-26.141	548.788	-28.788	.092	1.700	.013	-.618	-.610	-2.647	-.194	1.182	-1.097
14	585.000	598.609	-13.609	601.175	-16.175	.159	3.634	.007	-.334	-.328	-2.566	-.143	1.315	-.837
15	600.000	581.963	18.037	578.240	21.760	.171	3.995	.014	.446	.439	3.723	.200	1.321	1.180
16	626.000	555.549	70.451	552.641	73.359	.040	.183	.036	1.618	1.671	2.907	.339	.859	1.865
17	575.000	595.352	-20.352	597.430	-22.430	.093	1.721	.008	-.481	-.474	-2.079	-.151	1.203	-.856
18	520.000	540.351	-20.351	542.167	-22.167	.082	1.409	.007	-.478	-.471	-1.816	-.141	1.189	-.791
19	545.000	546.141	-1.141	546.256	-1.256	.092	1.700	.000	-.027	-.026	-.116	-.008	1.233	-.048
20	520.000	547.226	-27.226	551.703	-31.703	.141	3.129	.024	-.661	-.654	-4.477	-.265	1.242	-1.542
21	655.000	554.282	100.718	549.924	105.076	.041	.236	.077	2.316	2.538	4.358	.528	.601	2.904
22	630.000	605.483	24.517	601.537	28.463	.139	3.054	.019	.595	.587	3.946	.236	1.250	1.367
23	500.000	549.759	-49.759	557.041	-57.041	.128	2.736	.070	-1.199	-1.209	-7.282	-.463	1.089	-2.668
24	605.000	564.414	40.586	562.930	42.070	.035	.056	.011	.930	.928	1.484	.177	1.053	.973
25	555.000	597.885	-42.885	602.773	-47.773	.102	2.001	.039	-1.019	-1.020	-4.889	-.344	1.109	-1.957
26	505.000	552.473	-47.473	557.334	-52.334	.093	1.727	.043	-1.122	-1.128	-4.860	-.361	1.070	-2.040
27	540.000	544.874	-4.874	545.372	-5.372	.093	1.721	.000	-.115	-.113	-.498	-.036	1.232	-.204
28	520.000	546.141	-26.141	548.788	-28.788	.092	1.700	.013	-.618	-.610	-2.647	-.194	1.182	-1.097
29	585.000	598.609	-13.609	601.175	-16.175	.159	3.634	.007	-.334	-.328	-2.566	-.143	1.315	-.837
30	600.000	581.963	18.037	578.240	21.760	.171	3.995	.014	.446	.439	3.723	.200	1.321	1.180

Note: "Large" values are generated in **Lavender** and Red, respectively.

Table 2. Leverage and Influence Statistics for the u model

						Leverage		Influence						
									Studentized Residuals					
					PRESS or "Deleted"	"Hat" Diagonal	Mahalanobis D^2	Cook's D	Internal	External		DFFITS	COVRATIO	Welsch D
Case #	Criterion	Predicted	Residual	PRESS Predicted	Residual	>2k/n, >3k/n	Prob(χ2(p))<.001	>4/n, 1		$\lvert\,\rvert>2, >3$	DFFIT	$\lvert\,\rvert>2\sqrt{(k/n)}$	$\lvert CR-1\rvert\geq 3k/n$	$\lvert\,\rvert>3\sqrt{k}$
	Maximum:	4.079	.640	4.276	.711	.244	6.108	.280	1.628	1.684	.071	.561	1.358	3.183
	Minimum:	2.714	-.779	2.692	-.976	.076	1.229	.001	-2.104	-2.266	-.197	-1.139	.694	-6.864
	Mean:	3.313	.000	3.325	-.011	.133	2.900	.048	-.012	-.021	-.011	-.044	1.161	-.295
	SD:	.446	.386	.461	.457	.043	1.253	.068	1.012	1.041	.074	.450	.171	2.663
	Median:	3.274	.168	3.241	.187	.125	2.658	.020	.427	.420	.018	.143	1.209	.811
1	3.200	2.984	.216	2.966	.234	.076	1.229	.006	.542	.535	.018	.153	1.209	.857
2	4.100	3.932	.168	3.913	.187	.103	2.031	.005	.427	.420	.019	.143	1.269	.811
3	3.000	2.796	.204	2.767	.233	.125	2.657	.010	.525	.517	.029	.196	1.281	1.126
4	2.600	2.846	-.246	2.872	-.272	.093	1.732	.010	-.624	-.617	-.025	-.198	1.214	-1.117
5	3.700	3.274	.426	3.204	.496	.142	3.144	.051	1.109	1.114	.070	.453	1.123	2.632
6	4.000	3.360	.640	3.289	.711	.100	1.927	.073	1.628	1.684	.071	.561	.845	3.183
7	4.300	4.049	.251	4.005	.295	.150	3.381	.019	.656	.649	.044	.272	1.287	1.591
8	2.700	3.152	-.452	3.241	-.541	.164	3.783	.070	-1.193	-1.203	-.089	-.532	1.117	-3.134
9	3.600	3.549	.051	3.541	.059	.141	3.111	.001	.132	.129	.008	.052	1.358	.304
10	4.100	3.689	.411	3.622	.478	.141	3.127	.047	1.069	1.072	.068	.435	1.138	2.527
11	2.700	3.050	-.350	3.100	-.400	.125	2.658	.029	-.902	-.898	-.050	-.339	1.177	-1.954
12	2.900	2.714	.186	2.692	.208	.102	1.995	.006	.474	.467	.021	.158	1.258	.896
13	2.500	2.846	-.346	2.882	-.382	.093	1.732	.020	-.877	-.873	-.036	-.280	1.144	-1.582
14	3.000	3.378	-.378	3.500	-.500	.244	6.108	.089	-1.048	-1.051	-.122	-.597	1.302	-3.696
15	3.300	4.079	-.779	4.276	-.976	.202	4.883	.280	-2.104	-2.266	-.197	-1.139	.694	-6.864
16	3.200	2.984	.216	2.966	.234	.076	1.229	.006	.542	.535	.018	.153	1.209	.857
17	4.100	3.932	.168	3.913	.187	.103	2.031	.005	.427	.420	.019	.143	1.269	.811
18	3.000	2.796	.204	2.767	.233	.125	2.657	.010	.525	.517	.029	.196	1.281	1.126
19	2.600	2.846	-.246	2.872	-.272	.093	1.732	.010	-.624	-.617	-.025	-.198	1.214	-1.117
20	3.700	3.274	.426	3.204	.496	.142	3.144	.051	1.109	1.114	.070	.453	1.123	2.632
21	4.000	3.360	.640	3.289	.711	.100	1.927	.073	1.628	1.684	.071	.561	.845	3.183
22	4.300	4.049	.251	4.005	.295	.150	3.381	.019	.656	.649	.044	.272	1.287	1.591
23	2.700	3.152	-.452	3.241	-.541	.164	3.783	.070	-1.193	-1.203	-.089	-.532	1.117	-3.134
24	3.600	3.549	.051	3.541	.059	.141	3.111	.001	.132	.129	.008	.052	1.358	.304
25	4.100	3.689	.411	3.622	.478	.141	3.127	.047	1.069	1.072	.068	.435	1.138	2.527
26	2.700	3.050	-.350	3.100	-.400	.125	2.658	.029	-.902	-.898	-.050	-.339	1.177	-1.954
27	2.900	2.714	.186	2.692	.208	.102	1.995	.006	.474	.467	.021	.158	1.258	.896
28	2.500	2.846	-.346	2.882	-.382	.093	1.732	.020	-.877	-.873	-.036	-.280	1.144	-1.582
29	3.000	3.378	-.378	3.500	-.500	.244	6.108	.089	-1.048	-1.051	-.122	-.597	1.302	-3.696
30	3.300	4.079	-.779	4.276	-.976	.202	4.883	.280	-2.104	-2.266	-.197	-1.139	.694	-6.864

Note: “Large” values are generated in **Lavender** and Red, respectively.

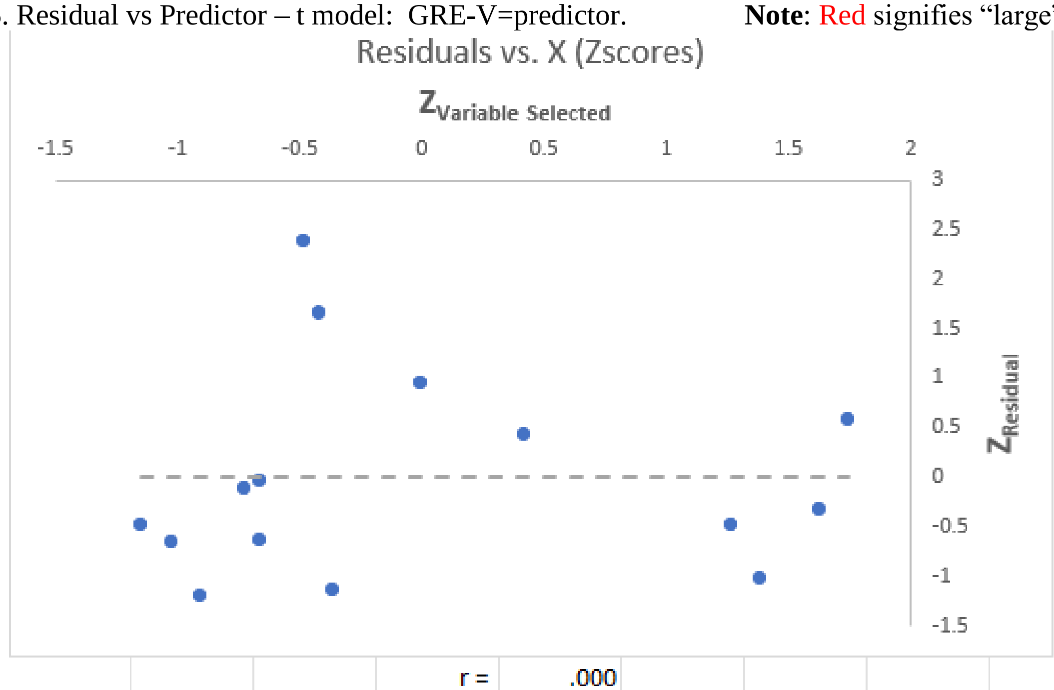
Table 3. DFBETAS for t model.

	DFBETAS →		
Case # ↓	Constant	GRE-V	MAT
1	0.1511	-0.1149	-0.0121
2	0.1037	-0.0911	-0.0335
3	-0.0809	0.1066	-0.0275
4	-0.0066	0.0006	0.0058
5	-0.0194	0.2023	-0.1887
6	0.2495	-0.2095	-0.0050
7	-0.1633	0.1785	0.0160
8	-0.0196	0.3390	-0.3325
9	0.0440	0.0171	-0.0416
10	0.2376	-0.2236	-0.0612
11	-0.2467	-0.0455	0.2775
12	-0.0292	0.0042	0.0241
13	-0.1533	0.0150	0.1333
14	0.0433	-0.1259	0.0666
15	-0.1187	-0.0407	0.1751
16	0.1535	-0.1167	-0.0123
17	0.1037	-0.0911	-0.0335
18	-0.0809	0.1066	-0.0275
19	-0.0066	0.0006	0.0058
20	-0.0194	0.2023	-0.1887
21	0.2495	-0.2095	-0.0050
22	-0.1633	0.1785	0.0160
23	-0.0196	0.3390	-0.3325
24	0.0440	0.0171	-0.0416
25	0.2376	-0.2236	-0.0612
26	-0.2467	-0.0455	0.2775
27	-0.0292	0.0042	0.0241
28	-0.1533	0.0150	0.1333
29	0.0433	-0.1259	0.0666
30	-0.1187	-0.0407	0.1751

Table 4. DFBETAS for u model.

	DFBETAS →			
Case # ↓	Constant	GRE-V	MAT	AR
1	0.0439	-0.0120	0.0414	-0.1056
2	-0.0897	0.0681	0.0075	0.0459
3	0.0843	-0.0894	0.0769	-0.1148
4	-0.1560	0.0199	0.1312	-0.0218
5	0.0314	-0.3281	0.3028	-0.0281
6	0.1951	-0.2406	-0.1856	0.4286
7	8.1769	0.1751	-0.0157	0.0747
8	-0.0055	0.2755	-0.4120	0.2499
9	0.0091	-0.0082	-0.0248	0.0452
10	-0.2682	0.2876	0.1567	-0.2280
11	-0.2097	0.0048	0.2769	-0.1721
12	0.1182	-0.0058	-0.0701	-0.0480
13	-0.2210	0.0281	0.1858	-0.0308
14	0.1663	-0.4963	0.0532	0.3529
15	0.5974	0.3127	-0.6439	-0.4437
16	0.0439	-0.0120	0.0414	-0.1056
17	-0.0897	0.0681	0.0075	0.0459
18	0.0843	-0.0894	0.0769	-0.1148
19	-0.1560	0.0199	0.1312	-0.0218
20	0.0314	-0.3281	0.3028	-0.0281
21	0.1951	-0.2406	-0.1856	0.4286
22	-0.1769	0.1751	-0.0157	0.0747
23	-0.0055	0.2755	-0.4120	0.2499
24	0.0091	-0.0082	-0.0248	0.0452
25	-0.2682	0.2876	0.1567	-0.2280
26	-0.2097	0.0048	0.2769	-0.1721
27	0.1182	-0.0058	-0.0701	-0.0480
28	-0.2210	0.0281	0.1858	-0.0308
29	0.1663	-0.4963	0.0532	0.3529
30	0.5974	0.3127	-0.6439	-0.4437

Table 5. Residual vs Predictor – t model: GRE-V=predictor. **Note**: Red signifies "large" values



Tables 1 & 2 show the leverage and influence statistics, and Tables 3 & 4 show the DIFFS, respectively for the t and u models. Entries are automatically highlighted in **Red,** if above a suggested criterion from the literature. Tables 5-9 contain the residual with individual predictor plots for each predictor for the t and u models, showing no signs of patterns by individual predictor.

VIFs for Models with # predictors > 1		
	t and u criterion	
Predictors	GRE-Q	GPA
GRE-V	1.221	1.294
MAT	1.221	1.492
AR		1.462

Table 6. Residual vs Predictor – t model: MAT=predictor.

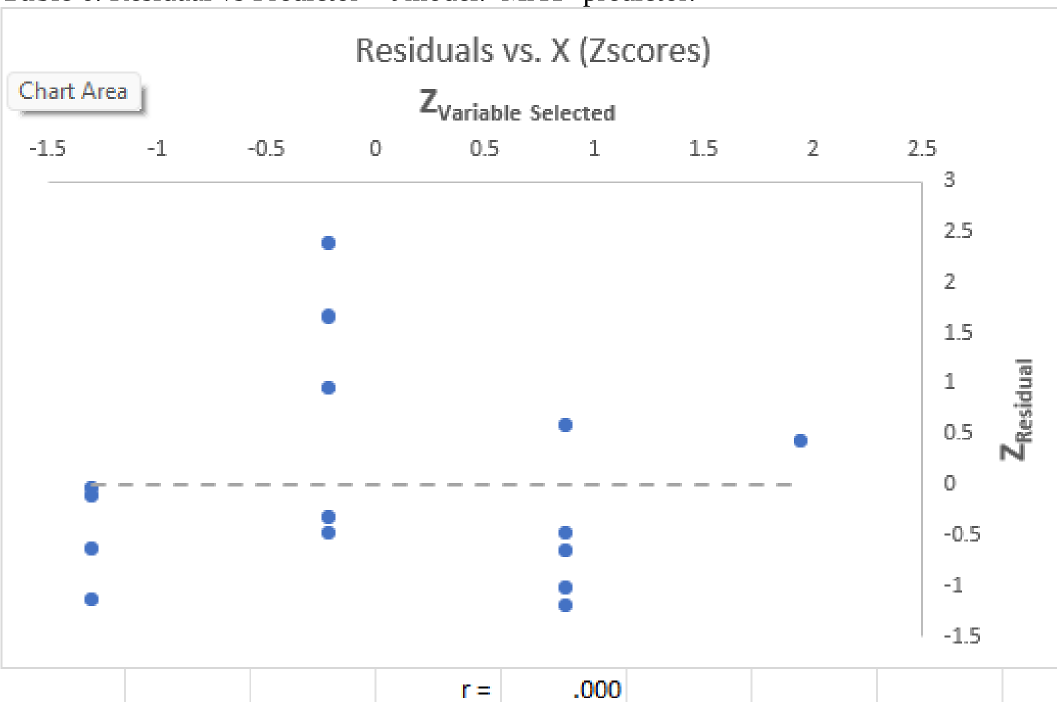
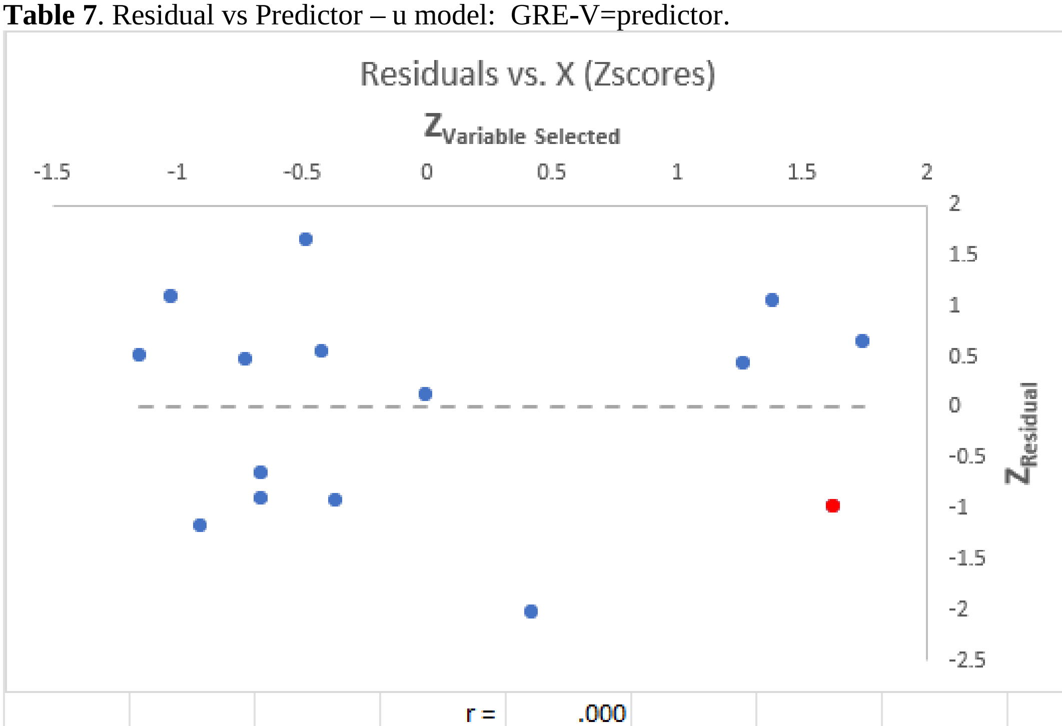
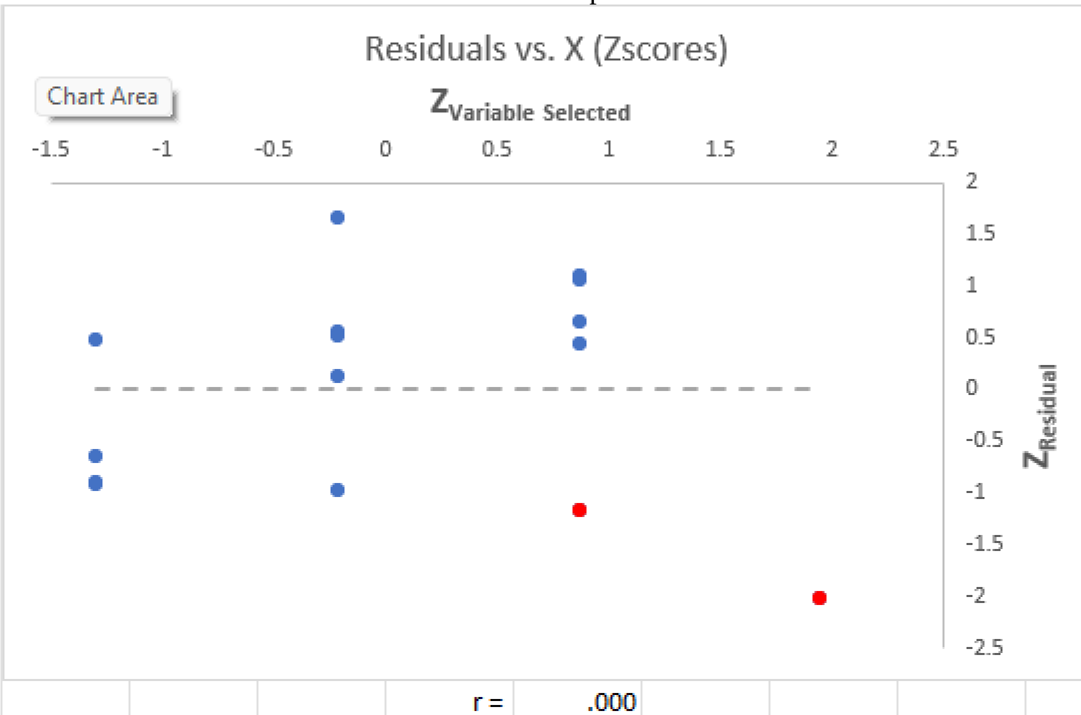


Table 7. Residual vs Predictor – u model: GRE-V=predictor.



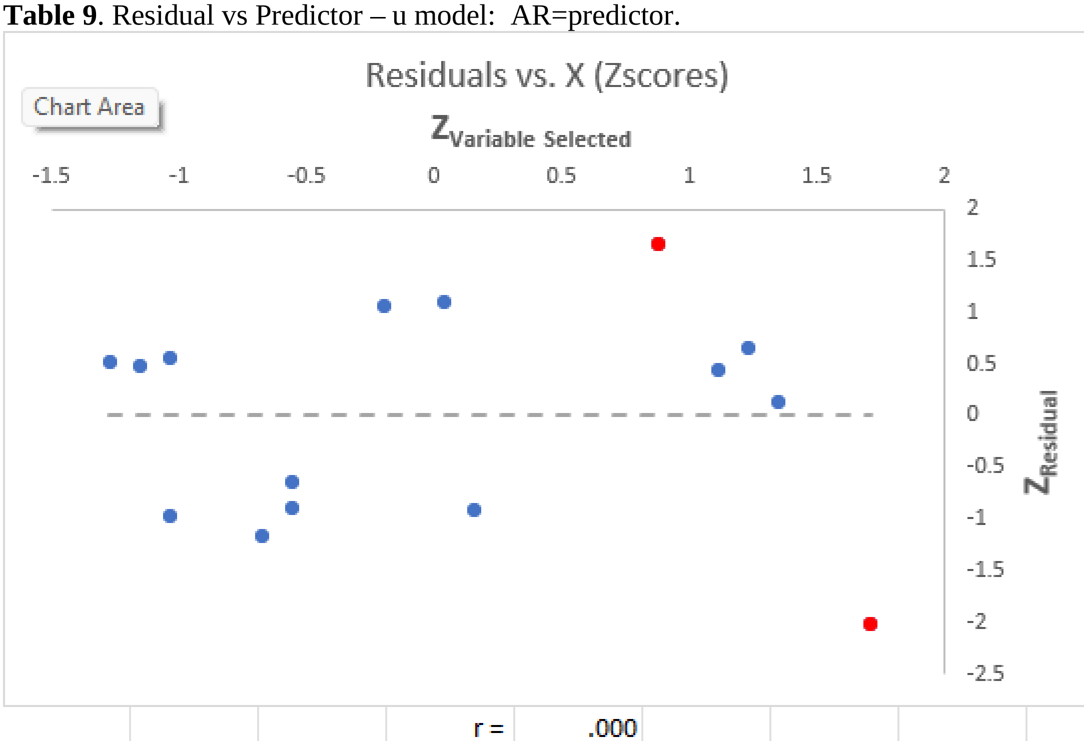
Note: Red signifies data point deemed influential by DFBETAS (Table 4) for GRE-V.

Table 8. Residual vs Predictor – u model: MAT=predictor.



Note: Red signifies data point deemed influential by DFBETAS (Table 4) for MAT.

Table 9. Residual vs Predictor – u model: AR=predictor.



Note: Red signifies data point deemed influential by DFBETAS (Table 4) for AR.

Significance

Attendance to least squares model assumptions is our responsibility, regardless of how those models are used. One might even argue that, as partialling depends **only** on residuals, which is what those assumptions regard, that our responsibility is even more so. Thus, such examination is considered mandatory. As commercial programs do not automatically provide such, a program to allow easy generation of that information is offered to all who would like it.

The **Partialling** program provides a flexible integrated treatment of partialling, and a comprehensive set of diagnostic statistics and plots, at least some of which one should consider. If you care to take a look, a guaranteed secure, anonymized, “voiceless” video demonstration of the analysis of these data with this software is here (make your browser full screen for the best view).
https://johnnysolarseed.com/Responsible_Partialling_Demo/Responsible_Partialling_Demo.mp4

References

Cohen, J., & Cohen, P. (1983). *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. Hillsdale, NJ: Erlbaum.

Kerlinger, F., N., & Pedhazur, E., J. (1973). *Multiple Regression in Behavioral Research*. New York, NY: Holt, Rinehart & Winston.

Lynam, D. R., Hoyle, R. H., & Newman, J. P. (2006). The perils of partialling. *Assessment*, *13*, 328-341.

McCullough, B. D., Wilson, B. (2002). On the accuracy of statistical procedures in Microsoft Excel 2000 and Excel XP. *Computational Statistics & Data Analysis*, *28*, 723-721.

Melard, G. (2014). On the accuracy of statistical procedures in Microsoft Excel 2010. *Computational Statistics*, *29*, 1095-1128.

Send correspondence to: Mary Lieberman
Florida Atlantic University
Email: mlieberm@fau.edu