HOW REGRESSION CONTROLS

In multiple regression, the effects have a partial interpretation: they show the effect of X1, while the effects of all other X-variables has been controlled (=”kept constant”). How does multiple regression do this?

THE OBSERVED MEANS

Report

y

NX1	NX2	Mean	N	Std. Deviation
1	1	2.83	6482	1.008
	2	3.91	3767	.970
	Total	3.23	10249	1.124
2	1	4.04	3767	.967
	2	5.15	6482	1.015
	Total	4.74	10249	1.132
Total	1	3.27	10249	1.152
	2	4.70	10249	1.163
	Total	3.98	20498	1.358

Marginal effect X1: 4.74-3.23 1.51
Controlled effect X1: 4.04-2.83 and 5.15-3.91 1.22

Marginal effect X2: 4.70-3.27 1.43
Controlled effect X2: 3.91-2.83 and 5.15-4.04 1.10

CORRELATIONS

Correlations

		y	NX1	NX2
y	Pearson Correlation	1	.557	.523
	Sig. (2-tailed)		.000	.000
	N	20498	20498	20498
NX1	Pearson Correlation	.557	1	.265
	Sig. (2-tailed)	.000		.000
	N	20498	20498	20498
NX2	Pearson Correlation	.523	.265	1
	Sig. (2-tailed)	.000	.000	
	N	20498	20498	20498

MULTIPLE REGRESSION

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	1.715	.025		68.836	.000
	NX1	1.513	.016	.557	96.005	.000
2	(Constant)	.504	.027		18.571	.000
	NX1	1.222	.014	.450	84.681	.000
	NX2	1.098	.014	.404	76.101	.000

a. Dependent Variable: y

RESIDUALIZING NX1 AND NX2

regr /dep=nx1 /enter=nx2 /save=resid(cnx1).
regr /dep=nx2 /enter=nx1 /save=resid(cnx2).

cnx1 Unstandardized Residual

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-.63245	3767	18.4	18.4	18.4
	-.36755	6482	31.6	31.6	50.0
	.36755	6482	31.6	31.6	81.6
	.63245	3767	18.4	18.4	100.0
	Total	20498	100.0	100.0	

cnx2 Unstandardized Residual

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	-.63245	3767	18.4	18.4	18.4
	-.36755	6482	31.6	31.6	50.0
	.36755	6482	31.6	31.6	81.6
	.63245	3767	18.4	18.4	100.0
	Total	20498	100.0	100.0	

THE EFFECTS OF THE RESIDUALIZED X-VARIABLES

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.984	.009		466.143	.000
	cnx1 Unstandardized Residual	1.222	.018	.434	68.926	.000

a. Dependent Variable: y

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	3.984	.009		456.085	.000
	cnx2 Unstandardized Residual	1.098	.018	.390	60.606	.000

a. Dependent Variable: y

CORRELATIONS

Correlations

	y	cnx1 Unstandardized Residual	NX1	cnx2 Unstandardized Residual	NX2
y	1	.434	.557	.390	.523
		.000	.000	.000	.000
	20498	20498	20498	20498	20498
cnx1 Unstandardized Residual	.434	1	.964	-.265	.000
	.000		.000	.000	1.000
	20498	20498	20498	20498	20498
NX1	.557	.964	1	.000	.265
	.000	.000		1.000	.000
	20498	20498	20498	20498	20498
cnx2 Unstandardized Residual	.390	-.265	.000	1	.964
	.000	.000	1.000		.000
	20498	20498	20498	20498	20498
NX2	.523	.000	.265	.964	1
	.000	1.000	.000	.000	
	20498	20498	20498	20498	20498

CONCLUSION

Partial coefficients in multiple regression are the same as the effects of residualized X-variables in simple regression.

B(Y,X1) is the effect of X1 on Y, while the effect of X1 on X2 has been taken out.
B(Y,X2) is the effect of X2 on Y, while the effect of X2 on X1 has been taken out.

CNX1 is the part of X1 that is not correlated with X2. CNX2 is the part of X2 that is not correlated with X1. Note how close this logic is to experimental design.