

How can I perform mediation with multilevel data using Method 2 in SPSS?

Authored by
stats writer

June 30, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I perform mediation with multilevel data using Method 2 in SPSS?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=162532>

Mediation with multilevel data using Method 2 in SPSS is a statistical approach for analyzing the relationships between multiple variables at different levels. This method involves using a hierarchical linear modeling technique to examine the indirect effects of an independent variable on a dependent variable through a mediator variable. It allows for the incorporation of nested data structures and can provide a more comprehensive understanding of the underlying mechanisms of a relationship. By conducting mediation analysis with multilevel data using Method 2 in SPSS, researchers can gain insights into the complex interactions between variables and make more informed conclusions about their findings.

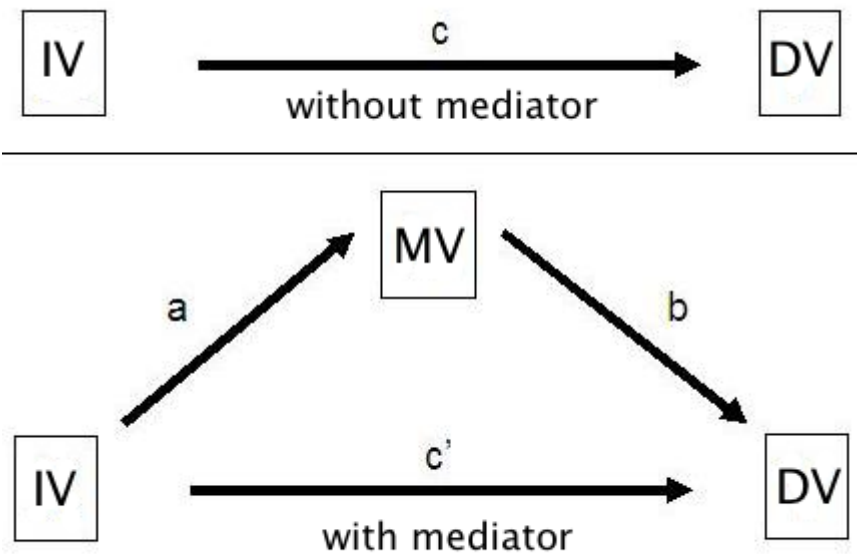
How can I perform mediation with multilevel data? (Method 2) | SPSS FAQ

Attention

See this FAQ by Bauer that discusses the need to decompose within- and between-group effects when using this approach to ensure valid results (<https://dbauer.web.unc.edu/wp-content/uploads/sites/7494/2015/08/Centering-in-111-Mediation.pdf>).

FAQ starts here

Mediator variables are variables that sit between the independent variable and dependent variable and mediate the effect of the IV on the DV. A model with one mediator is shown in the figure below.



The idea, in mediation analysis, is that some of the effect of the predictor variable, the IV, is transmitted to the DV

through the mediator variable, the MV. And some of the effect of the IV passes directly to the DV. That portion of of the effect of the IV that passes through the MV is the indirect effect.

An earlier approach to multilevel mediation suggested by Krull & MacKinnon

(2001) was method 1. This page will demonstrate an alternative approach given in the 2006 paper

by Bauer, Preacher & Gil. This approach combines the

dependent variable and the mediator into a single stacked response variable and runs one mixed model with indicator variables for the DV and mediator to obtain all of the values needed for the analysis.

We will begin by loading in a synthetic data set and reconfiguring it for our analysis.

All of the variables in this example (id the cluster ID, x the predictor variable, m the mediator variable, and y the dependent variable) are at level 1

Here is how the first 16 observations look in the original dataset. The dataset is available as a comma separated values (CSV) file here: ml_sim.csv.

Let's start by reading in the data and looking at a few descriptive statistics.

```
get data  
/type=txt  
/file="D:ml_sim.csv"  
/delimiters=", "  
/firstcase=2
```

/variables= id F2.0 x F m F y F.

execute.

dataset name ML_SIM.

compute fid = \$casenum.

execute.

descriptives variables=id fid x m y

/statistics=mean stddev range min max.

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation
id	800	99	1	100	50.50	28.884
fid	800	799.00	1.00	800.00	400.5000	231.08440
x	800	8.02	-4.15	3.87	-.1540	1.33037
m	800	11.49	-6.48	5.01	-.0248	1.48361
y	800	14.52	-8.60	5.92	-.1834	1.66918
Valid N (listwise)	800					

There are 100 level-2 units each with eight observations. fid is a row id, so when the data is not stacked, there is just 1 observation for each fid.

Let's look at the three models of a mediation analysis beginning with the model with just the IV.

mixed y with x

/fixed = x

```
/random = intercept x | subject(id) covtype(un)  
/method = reml  
/print = solution testcov.
```

ARABPSYCHOLOGY.COM

Information Criteria^a

-2 Restricted Log Likelihood	2435.086
Akaike's Information Criterion (AIC)	2443.086
Hurvich and Tsai's Criterion (AICC)	2443.136
Bozdogan's Criterion (CAIC)	2465.814
Schwarz's Bayesian Criterion (BIC)	2461.814

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: y.

Fixed Effects**Type III Tests of Fixed Effects^a**

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	99.575	.080	.778
x	1	99.871	138.315	.000

a. Dependent Variable: y.

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	-.027135	.095978	99.575	-.283	.778	-.217562	.163291
x	.689729	.058647	99.871	11.761	.000	.573374	.806084

a. Dependent Variable: y.

Covariance Parameters**Estimates of Covariance Parameters^a**

Parameter		Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Residual		.822588	.047033	17.490	.000	.735383	.920134
Intercept + x [subject = id]	UN (1,1)	.747010	.129272	5.779	.000	.532138	1.048645
	UN (2,1)	-.026055	.056865	-.458	.647	-.137509	.085400
	UN (2,2)	.217230	.047307	4.592	.000	.141759	.332880

a. Dependent Variable: y.

Next, comes the model with the mediator predicted by

the IV.

mixed m with x

/fixed = x

/random = intercept x | subject(id) covtype(un)

/method = reml

/print = solution testcov.

ARABPSYCHOLOGY.COM

Information Criteria^a

-2 Restricted Log Likelihood	2234.570
Akaike's Information Criterion (AIC)	2242.570
Hurvich and Tsai's Criterion (AICC)	2242.621
Bozdogan's Criterion (CAIC)	2265.299
Schwarz's Bayesian Criterion (BIC)	2261.299

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: m.

Fixed Effects

Type III Tests of Fixed Effects^a

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	99.606	1.080	.301
x	1	102.308	173.411	.000

a. Dependent Variable: m.

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	.095218	.091622	99.606	1.039	.301	-.086566	.277001
x	.611454	.046433	102.308	13.169	.000	.519358	.703551

a. Dependent Variable: m.

Covariance Parameters

Estimates of Covariance Parameters^a

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
Residual	.645031	.036535	17.655	.000	.577256	.720763	
Intercept + x [subject = id]	UN (1,1)	.709315	.118383	5.992	.000	.511418	.983791
	UN (2,1)	.006748	.041926	.161	.872	-.075425	.088920
	UN (2,2)	.116669	.028872	4.041	.000	.071831	.189495

a. Dependent Variable: m.

Finally, the model with both the IV and mediator

predicting the DV.

mixed y with x m

/fixed = x m

/random = intercept x m | subject(id) covtype(un)

/method = reml

/print = solution testcov.

ARABPSYCHOLOGY.COM

Information Criteria^a

-2 Restricted Log Likelihood	2044.936
Akaike's Information Criterion (AIC)	2058.936
Hurvich and Tsai's Criterion (AICC)	2059.078
Bozdogan's Criterion (CAIC)	2098.702
Schwarz's Bayesian Criterion (BIC)	2091.702

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: y.

Fixed Effects**Type III Tests of Fixed Effects^a**

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	98.826	2.281	.134
x	1	76.236	41.529	.000
m	1	92.360	177.928	.000

a. Dependent Variable: y.

Estimates of Fixed Effects^a

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	-.093740	.062074	98.826	-1.510	.134	-.216911	.029431
x	.249708	.038749	76.236	6.444	.000	.172537	.326879
m	.625066	.046860	92.360	13.339	.000	.532002	.718129

a. Dependent Variable: y.

Covariance Parameters**Estimates of Covariance Parameters^a**

Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
Residual	.507669	.030805	16.480	.000	.450745	.571782	
Intercept + x + m [subject = id]	UN (1,1)	.261471	.053506	4.887	.000	.175080	.390489
	UN (2,1)	-.031081	.024725	-1.257	.209	-.079542	.017379
	UN (2,2)	.039379	.021541	1.828	.068	.013479	.115050
	UN (3,1)	-.006042	.028718	-.210	.833	-.062328	.050245
	UN (3,2)	-.002522	.019270	-.131	.896	-.040290	.035245
	UN (3,3)	.120113	.030855	3.893	.000	.072599	.198724

We see that the IV although still significant has been reduced from .69 to .25.

Now, we need to restructure the data to stack y on m for each row and create

indicator variables for both the mediator and the dependent variables. Here's how we can do this (note that you will need to change the location of the new dataset to one on your computer).

```
vector variable = m to y.
```

```
loop sy = 1 to 2.
```

```
compute z = variable(sy).
```

```
xsave outfile 'D:ml_simlong.sav'
```

```
/drop y.
```

```
end loop.
```

```
execute.
```

```
get file'D:ml_simlong.sav'.
```

```
compute sy = sy - 1.
```

```
compute sm = ~sy.
```

```
execute.
```

The new response variable is called z and has y stacked on m.

We named the indicators for the mediator and the DV `sm` and `sy` respectively, to be consistent with Bauer et al (2006). We have also created a new `m` that contains the value for the mediator from each of the original observations.

Now we can run our mixed model for multilevel mediation using `mixed`.

Notice that because we include the `sm` and `sy` indicators

in the model that we need to use the `NOINT` option for the fixed effects

(it is not automatically included for random effects, so there is no need to suppress it).

In addition to the random effects, we use a repeated subcommand to model the heterogeneity in residual variances for `y` and `m` (which are now stacked and just in the variable `z`).

`mixed z with sm sy x m`

`/fixed = sm sm * x sy sy * m sy * x | noint`

`/random = sm sm * x sy sy * m sy * x | subject(id)
covtype(un)`

```
/repeated = sm | subject(fid id) covtype(diagDIAG)  
/method = reml  
/print = g solution testcov covb.
```

ARABPSYCHOLOGY.COM

ARABPSYCHOLOGY.COM

Model Dimension ^a						
		Number of Levels	Covariance Structure	Number of Parameters	Subject Variables	Number of Subjects
Fixed Effects	sm	1		1		
	sm * x	1		1		
	sy	1		1		
	sy * m	1		1		
	sy * x	1		1		
Random Effects	sm + sm * x + sy + sy * m + sy * x ^b	5	Unstructured	15	id	
Repeated Effects	sm	2	Diagonal	2	fid * id	800
Total		12		22		

a. Dependent Variable: z.

b. As of version 11.5, the syntax rules for the RANDOM subcommand have changed. Your command syntax may yield results that differ from those produced by prior versions. If you are using version 11 syntax, please consult the current syntax reference guide for more information.

Information Criteria ^a	
-2 Restricted Log Likelihood	4252.269
Akaike's Information Criterion (AIC)	4286.269
Hurvich and Tsai's Criterion (AICC)	4286.657
Bocagogan's Criterion (CAIC)	4394.638
Schwarz's Bayesian Criterion (BIC)	4377.638

The information criteria are displayed in smaller-is-better forms.

a. Dependent Variable: z.

Type III Tests of Fixed Effects ^a					
Source	Numerator df	Denominator df	F	Sig.	
sm	1	99.016	1.086	.300	
sm * x	1	101.266	173.174	.000	
sy	1	98.079	2.444	.121	
sy * m	1	92.064	179.782	.000	
sy * x	1	71.122	35.144	.000	

a. Dependent Variable: z.

Estimates of Fixed Effects ^a							
Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
sm	.093215	.089432	99.016	1.042	.300	-.084236	.270666
sm * x	.611857	.046495	101.266	13.160	.000	.519625	.704088
sy	-.096852	.061958	98.079	-1.563	.121	-.219805	.026101
sy * m	.610563	.045536	92.064	13.408	.000	.520125	.701001
sy * x	.220812	.037247	71.122	5.928	.000	-.146545	.295079

a. Dependent Variable: z.

Covariance Matrix for Estimates of Fixed Effects ^a						
Parameter	sm	sm * x	sy	sy * m	sy * x	
sm	.007998	.000322	.000576	.000093	-.000056	
sm * x	.000322	.002162	.000127	.000985	-.000197	
sy	.000576	.000127	.003839	-.000114	-.000060	
sy * m	.000093	.000985	-.000114	.002074	-.000484	
sy * x	-.000056	-.000197	-.000060	-.000484	.001387	

a. Dependent Variable: z.

Estimates of Covariance Parameters ^a							
Parameter	Estimate	Std. Error	Wald Z	Sig.	95% Confidence Interval		
					Lower Bound	Upper Bound	
Repeated Measures	Var. [sm=0.0]	.508965	.030674	16.593	.000	.452261	.572779
	Var. [sm=1.00]	.646731	.036682	17.631	.000	.578689	.722774
sm + sm * x + sy + sy * m + sy * x [subject=id]	UN (1,1)	.679434	.113179	6.003	.000	.490179	.941757
	UN (2,1)	.018161	.041087	.442	.658	-.062368	.098690
	UN (2,2)	.120483	.029315	4.110	.000	.074785	.194105
	UN (3,1)	.056812	.064523	.880	.379	-.069651	.183276
	UN (3,2)	.011881	.029211	.407	.684	-.045372	.089133
	UN (3,3)	.270284	.054087	4.997	.000	.182593	.400089
	UN (4,1)	.009322	.039004	.239	.811	-.067125	.085768
	UN (4,2)	.098955	.022822	4.336	.000	.054224	.143687
	UN (4,3)	-.004284	.027816	-.154	.878	-.058802	.050235
	UN (4,4)	.111872	.029141	3.839	.000	.067143	.186400
	UN (5,1)	-.006674	.034505	-.193	.847	-.074303	.060955
	UN (5,2)	-.021496	.018848	-1.141	.254	-.058437	.015444
	UN (5,3)	-.018276	.023330	-.783	.433	-.064002	.027450
	UN (5,4)	.005422	.017884	.303	.762	-.029630	.040474
	UN (5,5)	.032437	.020045	1.618	.106	.009661	.108907

a. Dependent Variable: z.

Covariance Matrix for Estimates of Covariance Parameters ^a																	
Parameter	Repeated Measures		sm + sm * x + sy + sy * m + sy * x [subject=id]														
	Var. [sm=0.0]	Var. [sm=1.00]	UN (1,1)	UN (2,1)	UN (2,2)	UN (3,1)	UN (3,2)	UN (3,3)	UN (4,1)	UN (4,2)	UN (4,3)	UN (4,4)	UN (5,1)	UN (5,2)	UN (5,3)	UN (5,4)	UN (5,5)
	Repeated Measures	.000941	.000005	-.000031	.000019	.000001	-.000018	.000004	-.000011	-.000012	-.000005	-.000019	-.000121	.000035	-.000006	.000025	.000072
sm + sm * x + sy + sy * m + sy * x [subject=id]	.000005	-.001346	-.000191	.000007	-.000118	-.000001	-.000012	-.000004	-.000009	.000016	-.000005	.000005	.000008	-.000027	.000016	.000005	-.000008
UN (1,1)	-.000031	.000019	.012810	.000475	-.000143	.000945	.000043	.000079	.000118	-.000108	.000041	.000044	-.000080	.000154	-.000032	.000042	-.000032
UN (2,1)	.000019	-.000007	.000475	.001688	.000121	.000116	.000133	-.000018	.000735	.000081	.000054	.000005	-.000080	-.000029	.000011	.000004	-.000019
UN (2,2)	.000001	-.000118	-.000143	.000121	.000859	.000000	.000098	.000000	.000097	.000428	.000031	.000196	-.000005	-.000087	-.000007	-.000037	.000008
UN (3,1)	-.000018	-.000001	.000945	.000116	.000000	.004163	.000094	.000625	.000075	-.000042	.000098	.000156	-.000122	.000220	-.000088	.000134	.000061
UN (3,2)	.000004	-.000012	.000043	.000133	.000098	.000094	.000853	.000125	.000065	.000017	.000393	.000038	.000014	-.000033	-.000090	.000029	-.000017
UN (3,3)	-.000011	-.000004	.000079	-.000018	.000000	.000625	.000125	.002925	.000097	.000014	-.000033	.000091	-.000038	-.000018	-.000067	.000065	-.000036
UN (4,1)	-.000012	-.000009	.000118	.000735	.000097	.000075	.000065	.000097	.001521	.000073	.000065	.000014	-.000331	-.000008	.000013	-.000012	.000006
UN (4,2)	-.000005	.000016	-.000108	.000081	.000428	-.000042	.000017	.000014	.000073	.000521	.000015	.000398	.000023	-.000107	.000063	-.000070	.000001
UN (4,3)	-.000019	-.000005	.000041	.000054	.000031	.000098	-.000393	-.000033	.000065	.000015	.000774	-.000039	.000011	.000333	.000037	.000013	
UN (4,4)	-.000121	-.000005	.000044	.000005	.000196	-.000156	-.000038	-.000091	.000014	.000398	-.000039	.000849	-.000028	-.000065	.000041	-.000187	.000023
UN (5,1)	.000035	.000008	-.000080	-.000080	-.000122	.000014	-.000038	-.000331	.000023	.000011	-.000028	.000191	.000032	.000015	.000085	-.000061	
UN (5,2)	-.000006	-.000027	.000154	-.000029	-.000087	.000220	-.000033	-.000018	-.000008	-.000107	.000061	-.000065	.000032	.000355	-.000023	.000167	-.000049
UN (5,3)	.000025	.000016	-.000106	.000011	-.000007	-.000088	-.000090	-.000063	.000013	.000041	.000125	-.000023	.000044	.000023	.000054	.000020	-.000098
UN (5,4)	.000072	.000005	-.000032	.000004	-.000037	.000134	.000029	.000065	-.000012	-.000070	.000037	-.000187	.000055	.000167	.000020	.000320	-.000139
UN (5,5)	-.000155	-.000008	.000042	-.000019	.000006	.000061	-.000017	-.000036	.000006	.000001	.000013	.000023	-.000061	-.000049	-.000088	-.000139	.000402

a. Dependent Variable: z.

Random Effect Covariance Structure (G) ^a					
	sm id	sm * x id	sy id	sy * m id	sy * x id
sm id	.679434	.018161	.056812	.009322	-.006674
sm * x id	.018161	.120483	.011881	.009855	-.021496
sy id	.056812	.011881	.270284	-.004284	-.018276
sy * m id	.009322	.009855	-.004284	.111872	.005422
sy * x id	-.006674	-.021496	-.018276	.005422	.032437

a. Dependent Variable: z.

We now have access to all of the information needed to compute the average indirect effect and average total effect and their standard errors using the equations given in Bauer, et. al. (2006).

\$\$

$$\text{ind} = ab + \sigma_{\{a_{\{j\}}b_{\{j\}}\}} \quad \text{quad (EQ:A11)}$$

\$\$

\$\$

$$\begin{aligned} \text{Var}(\text{ind}) &= b^{\{2\}}\sigma^{\{2\}}_{\{\text{hat}\{a\}\}} + \\ &a^{\{2\}}\sigma^{\{2\}}_{\{\text{hat}\{b\}\}} + \\ &\sigma^{\{2\}}_{\{\text{hat}\{a\}\}}\sigma^{\{2\}}_{\{\text{hat}\{b\}\}} + \\ &2ab\sigma_{\{\text{hat}\{a\},\text{hat}\{b\}\}} + \\ &(\sigma_{\{\text{hat}\{a\},\text{hat}\{b\}\}})^2 + \\ &\sigma^{\{2\}}_{\{\text{hat}\{\sigma\}_{\{a_{\{j\}},b_{\{j\}}\}}\}} \quad \text{quad (EQ:A14)} \end{aligned}$$

\$\$

average total effect

\$\$

$$\text{tot} = ab + \sigma_{\{a_{\{j\}}b_{\{j\}}\}} + c' \quad \text{quad (EQ:A15)}$$

\$\$

\$\$

$$\text{Var}(\text{ind}) = b^{\{2\}}\sigma^{\{2\}}_{\{\text{hat}\{a\}\}} +$$

$$\begin{aligned}
 & a^2 \sigma^2_{\hat{b}} + \\
 & 2a \sigma_{\hat{a}, \hat{b}} + 2b \sigma_{\hat{a}, \hat{c}'} + \\
 & 2a \sigma_{\hat{b}, \hat{c}'} + \\
 & \sigma^2_{\hat{\sigma}_{a_j, b_j}} + \\
 & \sigma^2_{\hat{c}'} + \\
 & \sigma^2_{\hat{a}} \sigma^2_{\hat{b}} + \\
 & (\sigma_{\hat{a}, \hat{b}})^2 \quad \text{quad (EQ:A18)} \\
 & \$\$
 \end{aligned}$$

These formulae involve the fixed effects estimates, their variances and covariances, and variances and covariances from the random effects. Here are all the values

\$\$

$$a = 0.6119 \setminus$$

$$b = 0.6106 \setminus$$

$$c' = 0.2208 \setminus$$

$$\sigma_{a_j b_j} = 0.09896 \setminus$$

$$\sigma^2_{\hat{a}} = 0.002162 \setminus$$

$$\sigma^2_{\hat{b}} = 0.002074 \setminus$$

$$\sigma_{\hat{a}, \hat{b}} = 0.000985 \setminus$$

$$\sigma_{\hat{a}, \hat{c}'} = -0.00020 \setminus$$

$$\sigma_{\hat{b}, \hat{c}'} = -0.00048 \setminus$$

$$\sigma^2_{\hat{c}} = 0.001387 \setminus$$

$$\sigma^2_{\hat{\sigma}_{a_j, b_j}} = 0.02282^2$$

\$\$

To calculate this, you just need a calculator. A simple way in SPSS is using SPSS' matrix language, which essentially allows us to just declare the constants and write out the formulae.

matrix.

```
compute a = 0.611857.
compute b = 0.610563.
compute rcov_ab = 0.098955.
compute cprime = 0.220812.
compute Va = 0.002162.
compute Vb = 0.002074.
compute Vcprime = 0.001387.
compute cov_ab = 0.000985.
compute cov_ac = -0.000197.
compute cov_bc = -0.000484.
compute Vcov_ab = 0.022822**2.

compute ind_eff = a*b + rcov_ab.
```

```

compute V_ind = a**2*Vb + b**2*Va + Va*Vb +
2*a*b*cov_ab + cov_ab**2 + Vcov_ab.
compute test_ind = ind_eff/V_ind**.5.
compute tot_eff = ind_eff + cprime.
compute V_tot = b**2*Va + a**2*Vb + Va*Vb +
2*a*b*cov_ab + cov_ab**2 + Vcprime + 2*b*cov_ac +
2*a*cov_bc + Vcov_ab.
compute test_tot = tot_eff/V_tot**.5.

print ind_eff / FORMAT = F6.6 / title 'indirect effect'.
print V_ind / FORMAT = F6.6 / title 'variance of indirect
effect'.
print test_ind / FORMAT = F6.6 / title 'significance test
of indirect effect, test against standard normal'.
print tot_eff / FORMAT = F6.6 / title 'total effect'.
print V_tot / FORMAT = F6.6 / title 'variance of total
effect'.
print test_tot / FORMAT = F6.6 / title 'significance test of
total effect, test against standard normal'.
end matrix.

```

Run MATRIX procedure:

indirect effect

.47253

variance of indirect effect

.00284

significance test of indirect effect, test against standard normal

8.8597

total effect

.69334

variance of total effect

.00340

significance test of total effect, test against standard normal

11.893

----- END MATRIX -----

Finally, note that it is possible to achieve the same effect as using the repeated subcommand using a second random subcommand. We add a random slope by fid to model the additional heterogeneity in outcomes. The residual variance is the

residual variance of y, the residual variance plus the variance of the random slope is the residual variance for m. Note that this model takes some time to run.

mixed z with sm sy x m

/fixed = sm sm * x sy sy * m sy * x | noint

/random = sm sm * x sy sy * m sy * x | subject(id)

covtype(un)

/random = sm | subject(fid)

/method = reml

/print = g solution testcov covb.

ARABPSYCHOLOGY.COM

Note that $.508965 + .137766 = 0.646731$, the residual variance of m from our previous model, showing that these two approaches yield similar results, although the random slope approach is somewhat less direct.

References

ARABPSYCHOLOGY.COM