https://community.jmp.com/t5/Discovery-Summit-Europe-2015/Linear-Mixed-Models-With-JMP-Pro-One-Face-Many-Flavours/ta-p/22333

Linear Mixed Models with JMP® Pro: One Face, Many Flavours*

Jian Cao, PhD
SAS Institute, Inc.
March 24, 2015
Brussels, Belgium

(Updated by Patrick Giuliano on 11/11/2022)

(Changes included updated screenshots and some additional instruction to clarify model specification)

Abstract

One of the new features introduced in JMP Pro 11 is mixed models. This new modelling personality in the Fit Model platform enables one to fit a variety of regression models with fixed and random effects along with an appropriate covariance structure. What's a mixed model? When and why should one fit a mixed model? And how does JMP fit such a model? In this paper I will try to dispel myths about the mixed models by 1) briefly reviewing the statistical background, 2) discussing why mixed models provide better estimates and consequences of fitting traditional regression models to data where measurements of a response variable are correlated or a key explanatory variable is missing, and 3) illustrating JMP® Pro's mixed models by fitting different flavours of mixed models that are widely employed in real life applications.

Acknowledgement: I would like to thank Christopher Gotwalt and Laura Lancaster for their help.

**The material presented in this Paper applies to JMP Pro 12 as well. The key enhancement in the new JMP Pro 12 is improved optimization algorithm that enables JMP to run faster than in JMP Pro 11 for large data.*

JMP Pro 11 has added a new modeling personality, **Mixed Model**, to its **Fit Model** platform. What's a mixed model? How does JMP fit such a model? What are the key applications where mixed models can be applied? In this paper, I will try to dispel myths about the mixed models and demonstrate JMP's capability with real-life examples.

What's a Linear Mixed Model
Linear mixed models are a generalization of linear regression models, $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$. Extending the model to allow for random effects, **Z**, the new model becomes $\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z\gamma} + \boldsymbol{\varepsilon}$**.** This is the linear mixed model as there are both fixed effects, **X,** and random effects, $\boldsymbol{Z}$.

The following assumptions are made for the random effect parameters, $\boldsymbol{\gamma}$ and random error **ε:** (1) $\boldsymbol{\gamma}$ and **ε** are normally distributed, and (2) there are no correlations between $\boldsymbol{\gamma}$ and **ε**. JMP provides several commonly used structures for **ε.** The fixed effect coefficients,$\boldsymbol{\beta}$**,** and covariance matrices for $\boldsymbol{\gamma}$ and **ε** are jointly estimated by the restricted maximum likelihood method. Fitting mixed models requires additional data on each cross-section unit or, in case of modeling spatial data, dimensions of measurements. There are mixed models for non-normal distributed responses or non-linear mixed models; however, I limit the scope of my discussion to the linear mixed models that are supported in JMP Pro.

Why Mixed Models
When there exists correlation among responses or an important explanatory variable is missing, failure to account for that leads to biased estimates of the effects of treatment and other factors.

Here are some common use cases for mixed models:

- Allowing coefficients (e.g., intercept and slopes) to vary randomly across subjects (i.e., random coefficient models). A variant is individual growth model, which can be applied to predict individual growth trajectory and stability analysis;
- Analysis of randomized block designs, and split-plot designs where hard-to-change and easy-to-change factors result in multiple error terms;
- Controlling for unobserved individual heterogeneity in the form of random effects (i.e., panel data models);
- Analysis of repeated measures where within-subject errors are correlated;
- Multiple responses that are correlated because measures are taken from the same subjects;
- Subjects are hierarchical (i.e., students within schools). This is known as Hierarch Linear Model or Multi-level Models;
- Spatial variability (i.e., geospatial regression);

With JMP Pro you can easily specify and fit all of these models using the point-and-click interface and review the results in a user-friendly way.

Steps to Specify a Mixed Model in JMP Pro

1. Select **Analyze =>Fit Model**, and choose **Mixed Model Personality**;

2. Select a continuous response variable from you data table as **Y** and construct fixed effects as you normally would do with a standard least squares fit;

3A. Use **Random Effects** tab to specify random coefficients or random effects;

3B. Use **Repeated Structure** tab to select a covariance structure for model errors;

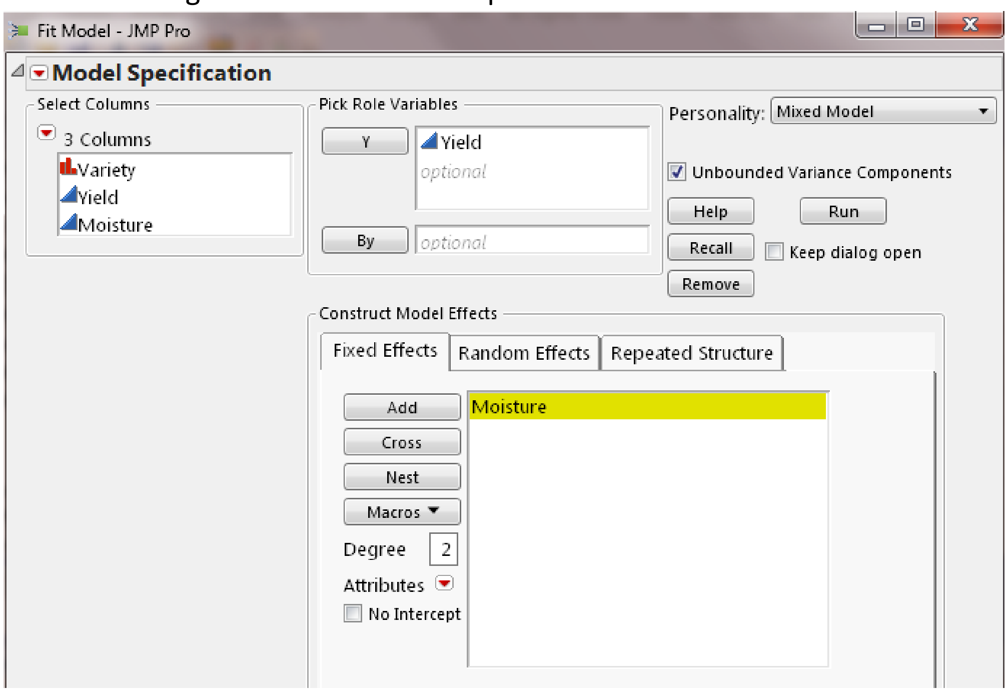
4. Click **Run**.

I'll now turn to example to show four different flavours of mixed models: random coefficient model, analysis of repeated measures, panel data model and geospatial regression.

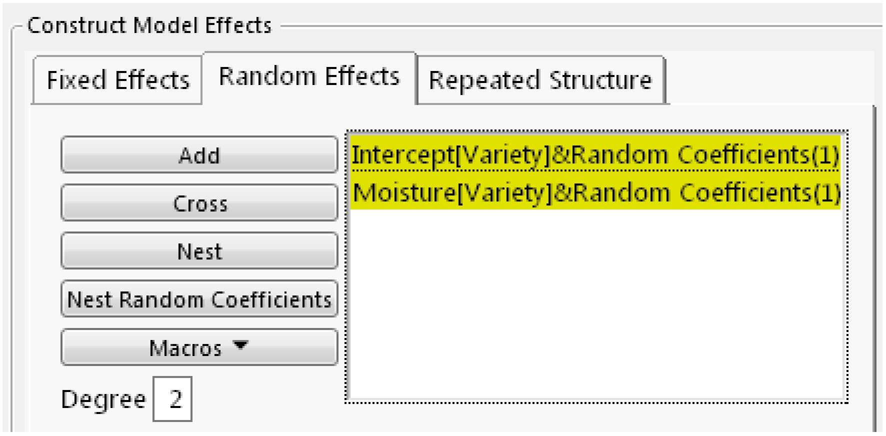
Example 1: Random Coefficient Models—allowing intercept and slopes to vary randomly across subjects

In this example we are interested in estimating the effect on wheat yield of pre-planting moisture in the soil while allowing each wheat variety to have random deviation from population effects. So, a random coefficient model is called for. The experiment randomly selects 10 varieties from wheat population and assigns each to six plots of land. In total, 60 observations with 6 measurements of yield for each variety is collected. (The data, "Wheat", is available in JMP's Sample Data directory.)

I followed the steps laid out above to specify the model. From **Fixed Effects** tab, specify Moisture along with a default intercept as fixed effects.

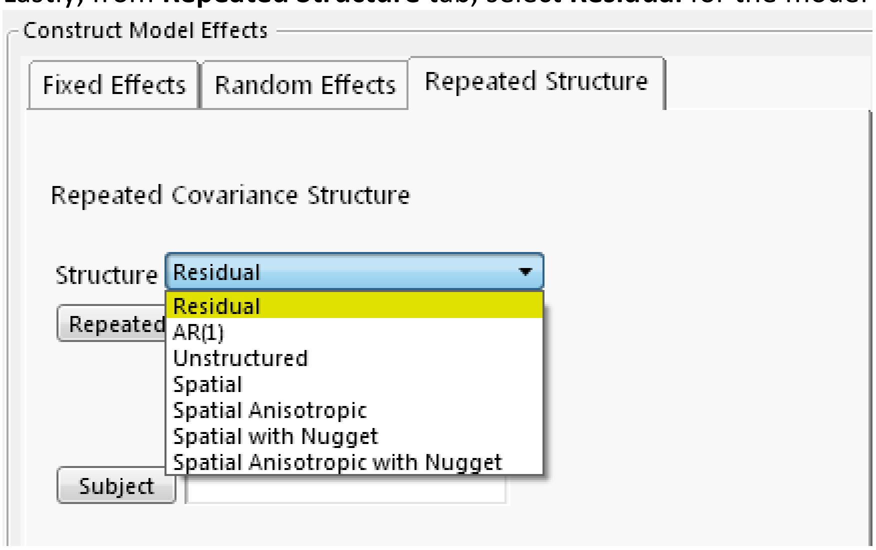


Next, from the **Random Efects** tab,using **Nest Random Coefficients** button to request random intercept and Moisture effect for each variety.



Add **Moisture** to Random Effects Tab.
Then select Moisture where it has been added under Construct Model EFfects, then Select **Variety i**n the Select Columns Pane, and then click the "Nest Random Coefficients" button. Then you will get what is shown at left.

Lastly, from **Repeated Structure** tab, select **Residual** for the model error term.



The following screenshot shows **Random Effects Covariance Parameter Estimates, Fixed Effects Parameter Estimates** and **Random Coefficients**. Let's discuss them in turn.

Random Effects Covariance Parameter Estimates

Covariance Parameter	Subject	Estimate	Std Error	95% Lower	95% Upper
Var(Intercept)	Variety	18.894659	9.1110743	1.0372813	36.752036
Cov(Moisture,Intercept)	Variety	-0.072717	0.08242	-0.234257	0.0888234
Var(Moisture)	Variety	0.0023942	0.0013492	-0.00025	0.0050385
Residual		0.3520553	0.0790171	0.2369917	0.5775592

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	33.433883	1.3996989	9.2	23.89	<.0001*	30.278007	36.589759
Moisture	0.6616554	0.0168282	8.7	39.32	<.0001*	0.623361	0.6999498

Random Coefficients

Variety

Variety	Intercept	Moisture
1	0.9577955	-0.049211
2	-2.284277	-0.066697
3	-0.40812	0.0672228
4	0.696021	-0.023306
5	1.1159079	-0.019904
6	4.6391469	0.0238888
7	-10.73005	0.0564236
8	2.401166	0.0224337
9	-0.176212	0.0233568
10	3.7886181	-0.034207

Covariance Matrix

Random Effect	Intercept	Moisture
Intercept	18.89466	-0.07272
Moisture	-0.07272	0.002394

The variance estimate for Intercept is 18.89 with a standard error estimate of 9.11, so the z-score is 2.07 (=18.89/9.11). Using the Normal Distribution function from JMP Formula Editor we can find the p-value to be 0.0192, indicating that the variation in baseline yield across varieties is statistically significant. Similarly, the *p*-value for *Cov(Moisture, Intercept)*, 0.3777, and p-

The **Random Coefficients** report gives the BLUP (Best Linear Unbiased Predictor) values for how each variety is different from the population intercept and population Moisture effect reported in **Fixed Effects Parameter Estimates.** For Variety 1, the estimated moisture effect on its yield is 0.61 (=0.66-0.05), and baseline yield is 34.39 (=33.43+0.96) and the predicted yield equation is $Yield = 34.39 + 0.61 * Moisture$.

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	33.433881	1.3996984	9.2	23.89	<.0001*	30.278007	36.589754
Moisture	0.6616554	0.016828	8.7	39.32	<.0001*	0.6233617	0.6999492

Random Coefficients

Variety

Variety	Intercept	Moisture
1	0.9577924	-0.049211
2	-2.284289	-0.066697
3	-0.408109	0.0672225
4	0.6960205	-0.023306
5	1.115904	-0.019904
6	4.639151	0.0238887
7	-10.73004	0.0564233
8	2.4011709	0.0224336
9	-0.176207	0.0233566
10	3.7886051	-0.034207

Combining both the fixed effects and random coefficient estimates, we find a significant overall effect on wheat yield of moisture, and discover significant variation in the moisture effect across different varieties (if we assume the p-value that we calculated by hand in this case is accurate, which isn't necessarily the case -- consider the Wald p-values that JMP16 and up now generates, and also consider sample size in making your determination of statistical signifcance).

Other Applications of Random Coefficient Models

Individual Growth Model is a type of random coefficient model in which random *time* effect is estimated for each individual. This is done by specifying a continuous time variable such as day or month as a random effect, and using **Nest Random Coefficients** button to request separate slope (i.e., growth) and intercept for each individual.

In educational research, subjects are often nested in a hierarchical order. By adding multiple groups of random effect statements you can fit **Hierarchical Linear Models/Multi-level Models**.

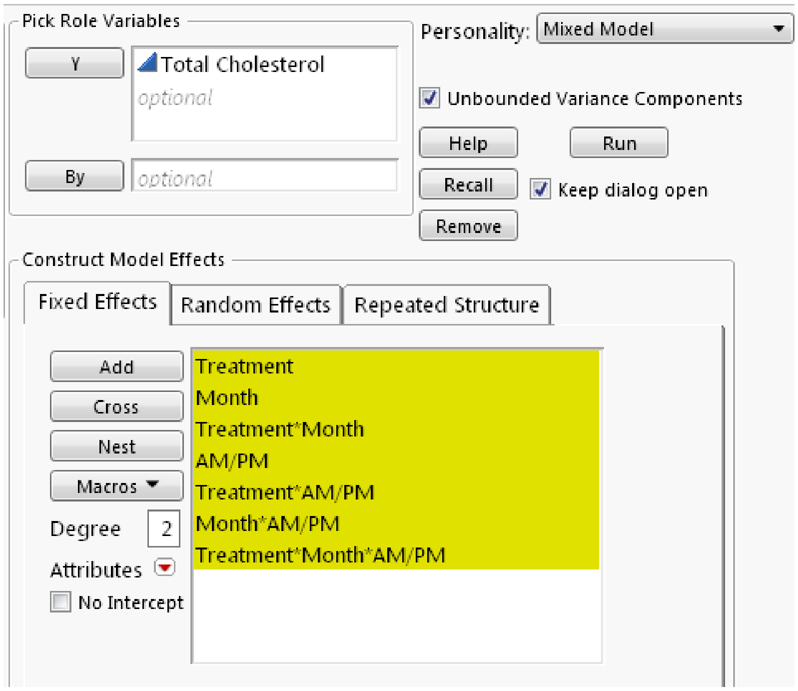
Example 2: Analysis of Repeated Measures—accounting for correlated errors

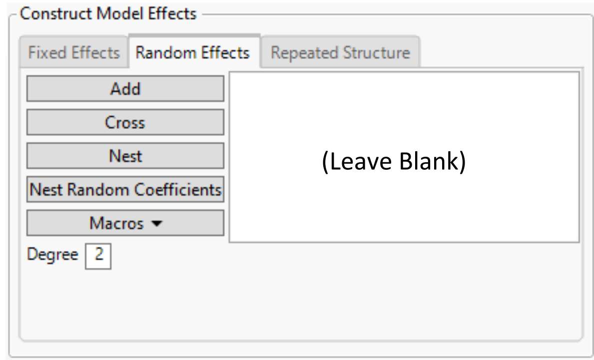
Repeated measures are the multiple measurements of a response collected from the same subjects over time. In this clinical trial, subjects (i.e., patients) were randomly assigned to different treatment groups. Each subject's total cholesterol level was measured several times during the trial. The objective of the study is to test whether new drugs are effective at lowering cholesterol. What makes the analysis of repeated measures distinct is the correlation of the measurements within a subject. Failure to account for it often leads to incorrect conclusion about the treatment effect. (The data, **Cholesterol Stacked**, is available in JMP's Sample Data directory.

JMP Pro offers three commonly used covariance structures: (and many others)

- **Unstructured** provides a flexible structure that estimates covariance for all pairs of measurement times. In this example of six repeated measures, 15 covariance parameters as well as 6 variance estimates will be estimated. This structure is most lenient but not without risk of over-fitting**.**
- **AR(1)** (first-order autoregressive) estimates correlation between two measurements that are one unit of time apart. The correlation declines as the time difference increases. AR(1) is a parsimonious structure with only two variance parameters to be estimated.
- **CS** (compound symmetry) postulates that the covariance is constant regardless of how far apart the measurements are. The # of parameters to be estimated is 2.

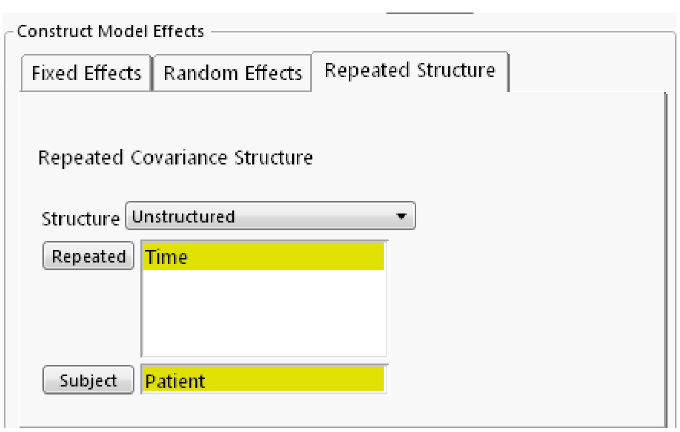
The following screenshot shows the **Fixed Effects** part of the repeated measures analysis, which includes *Treatment*, *Month*, *AM/PM*, and their interactions.





(Fixed effects part of the model)

I will consider three different covariance structures for the within-subject errors. First, let's use **Unstructured.** Apply *Time* column as **Repeated,** and *Patient* column as **Subject**--this defines the repeated measurements within a subject. It is important to note that JMP requires that **Subject** column be uniquely valued and that **Repeated** column be categorical for the **Unstructured** option.



(Unstructured Covariance Structure)

Key reports include **Repeated Effects Covariance Parameter Estimates, Fixed Effects Parameter estimates,** and **Tests Fixed Effects Tests**.

Repeated Effects Covariance Parameter Estimates

Repeated Effect: Time
Subject: Patient

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Var(June PM)	65.568482	23.181959	20.132677	111.00429
Var(June AM)	63.512277	22.454981	19.501323	107.52323
Cov(June PM,June AM)	63.878686	22.700344	19.386829	108.37054
Var(May PM)	57.058089	20.173081	17.519577	96.596602
Var(May AM)	56.603347	20.012305	17.379949	95.826744
Cov(May PM,May AM)	55.365523	19.83529	16.489068	94.241978
Var(April PM)	19.268932	6.8125963	5.9164888	32.621376
Var(April AM)	18.725	6.6202872	5.7494754	31.700525
Cov(April PM,April AM)	18.354884	6.6035621	5.4121399	31.297628
Cov(May PM,April AM)	9.4147226	8.5038584	-7.252534	26.081979
Cov(May AM,April AM)	9.2756709	8.4629182	-7.311344	25.862686
Cov(May PM,April PM)	6.6230805	8.4532303	-9.944946	23.191108
Cov(May AM,April PM)	5.5074058	8.3704001	-10.89828	21.913089
Cov(June PM,April AM)	1.5810194	8.7687993	-15.60551	18.76755
Cov(June AM,April AM)	1.1945478	8.6266098	-15.7133	18.102392
Cov(June PM,April PM)	0.7647113	8.8882627	-16.65596	18.185386
Cov(June AM,April PM)	0.3183447	8.7461245	-16.82374	17.460434
Cov(June AM,May AM)	1.106725	14.992149	-28.27735	30.490796
Cov(June AM,May PM)	0.6455101	15.050552	-28.85303	30.14405
Cov(June PM,May AM)	0.9262595	15.232066	-28.92804	30.780561
Cov(June PM,May PM)	0.6543895	15.292238	-29.31785	30.626625

Fixed Effects Parameter Estimates

Fixed Effects Tests

Source	Nparm	DFNum	DFDen	F Ratio	Prob > F
Treatment	3	3	16.0	274.96713	<.0001*
Month	2	2	15.0	340.48166	<.0001*
Treatment*Month	6	6	18.3	123.47461	<.0001*
AM/PM	1	1	16.0	360.93593	<.0001*
Treatment*AM/PM	3	3	16.0	0.6339843	0.6038
Month*AM/PM	2	2	15.0	1.1988247	0.3289
Treatment*Month*AM/PM	6	6	18.3	1.1642781	0.3671

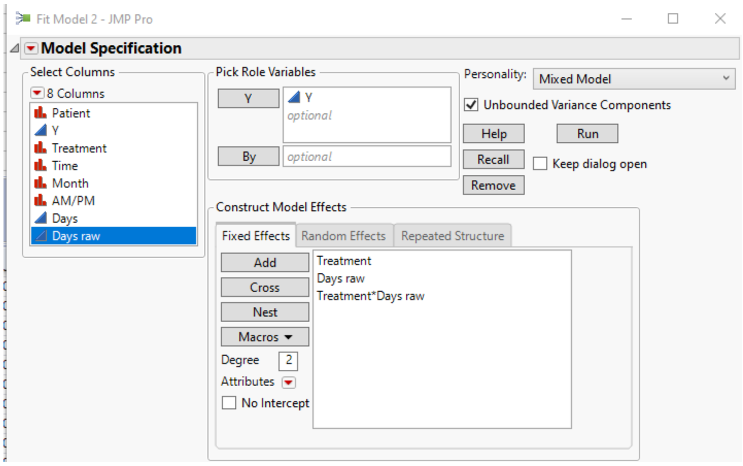
(Results using Unstructured)

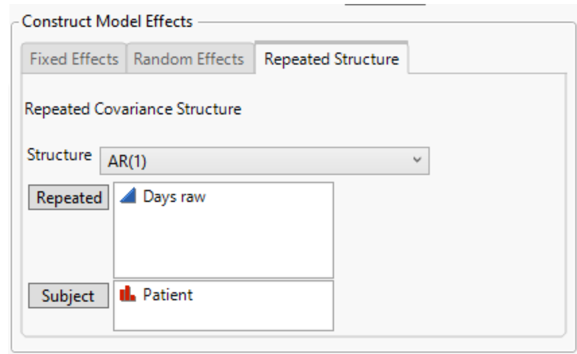
One way of testing statistical significance of the covariance estimates is to calculate the z-scores and find their p-values, as I did in the previous random coefficient model example. However, we can check the confidence limits: if the 95% confidence interval for a covariance estimate includes zero, then we can say that the estimate is not statistically significant from zero at α=5%*(1)*. As we can see, all six variance estimates are significantly different from zero but most of covariance estimates are not. This suggests that a parsimonious structures, such as AR (1), should be considered.

Fixed Effects Tests report shows a highly significant treatment effect. Cholesterol level is also found to vary significantly from month to month and from morning to afternoon.

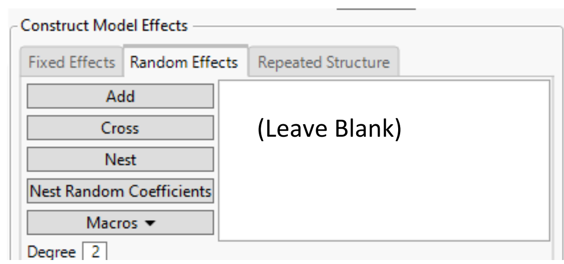
(1) This is just the "Confidence Interval version" of the 2-Sample t-test (reference: Bluman, Elementary Stats (7th Ed. 2015).

Next, we consider **AR(1)** as the covariance structure for the within-subject errors. Please note that **Repeated** column used in AR(1) by JMP must be a continuous variable. So, D*ays*—number of days from the trial start date at each measurement--is used instead of a categorical variable used for the Structure = Unstructured option.





(AR(1) Covariance Structure)



The **Repeated Effects Covariance Parameter Estimate** report shows a highly significant within-subject correlation of 0.95. Fixed effects results are similar to those in the UN option --treatment effect and time effects (in Days) are statistically significant.

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	276.17698	1.9946144	52.7	138.46	<.0001*	272.17582	280.17815
Treatment[A]	-32.98374	2.1833359	52.1	-15.11	<.0001*	-37.36478	-28.6027
Treatment[B]	-20.03827	2.1833359	52.1	-9.18	<.0001*	-24.41931	-15.65723
Treatment[Control]	26.049918	2.1833359	52.1	11.93	<.0001*	21.668876	30.430959
Days raw	-0.735803	0.0505395	53.3	-14.56	<.0001*	-0.83716	-0.634446
Treatment[A]*(Days raw-30.5833)	-0.877856	0.087537	53.3	-10.03	<.0001*	-1.053411	-0.7023
Treatment[B]*(Days raw-30.5833)	-0.643666	0.087537	53.3	-7.35	<.0001*	-0.819221	-0.46811
Treatment[Control]*(Days raw-30.5833)	0.788161	0.087537	53.3	9.00	<.0001*	0.6126056	0.9637164

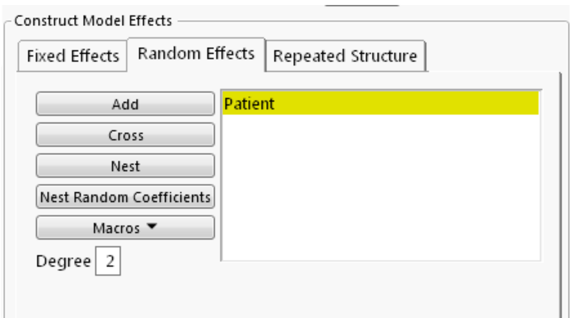
Repeated Effects Covariance Parameter Estimates

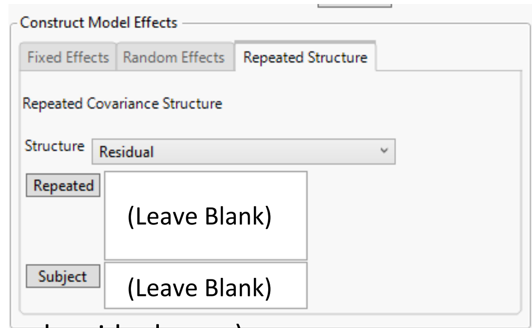
Subject: Patient (Results using AR(1))

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
AR(1) Days raw	0.7983061	0.0473103	0.7055796	0.8910327
Residual	100.5539	18.519406	72.239307	149.60244

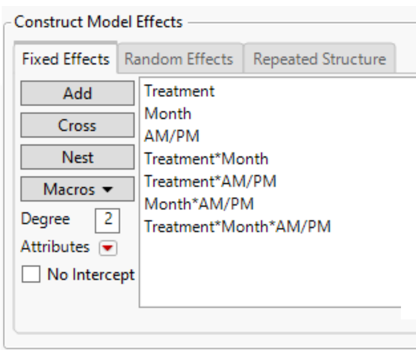
Note that in this example we have to create a new Column = "Days raw" from the sample dataset without the recoded Column formula for "Days" so that JMP will accept the input in the "Repeated" Dialog Box.

To complete our example, finally, let's fit the model with a **CS** structure. To do so, select **Residual** as the **Repeated Covariance Structure**—but no need to specify **Repeated** and **Subject** columns with this option; instead, we add the subject ID, **Patient,** as a random effect on the **Random Effects** tab. That is, within-subject covariance is modelled through the random subject effect.





(CS Covariance Structure with random subject effect and residual error)



(Results using CS)

Random Effects Covariance Parameter Estimates

Variance Component	Var Ratio	Estimate	Std Error	95% Lower	95% Upper	Wald p-Value	Pct of Total
Patient	0.33372	11.707432	6.2749012	-0.591148	24.006012	0.0621	25.022
Residual		35.081923	5.546939	26.320843	49.105827		74.978
Total		46.789355	7.7386523	34.678063	66.607248		100.000

Judged by the 95% confidence limits, the covariance between any two measures on the same subject is not statistically significant at α=0.05 (actually, p-value= 0.0621). Fixed effect test results are similar to the previous models and are thus not shown here.

So, which repeated structure should be adopted? One criterion for model comparison is AICc. From the **Fit Statistics** reported by JMP (not shown), AICcs are: **Unstructured**—703.84, **AR(1)**—652.63 and **CS**—832.55. So, **AR(1)** is the winner.

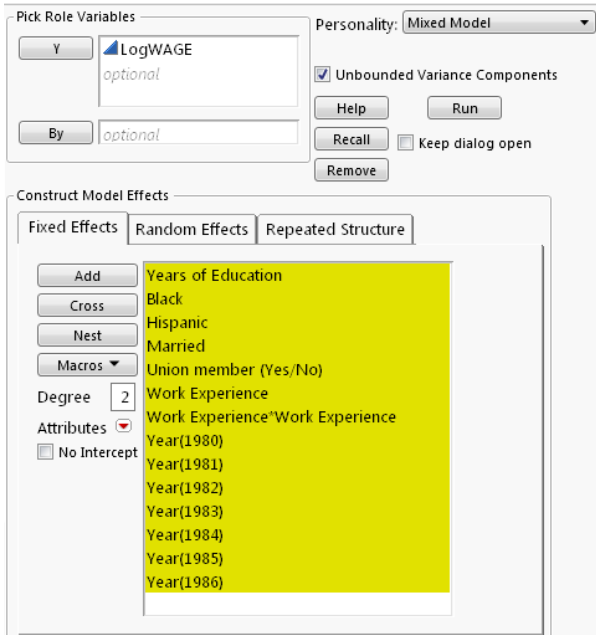
Example 3: Panel Data Models--controlling for unobserved heterogeneity

This example is taken from Vella and Verbeek (1998), which is discussed in *Introductory Econometrics* by Jeffrey Woodridge as Example 14.4. See references below for more info.

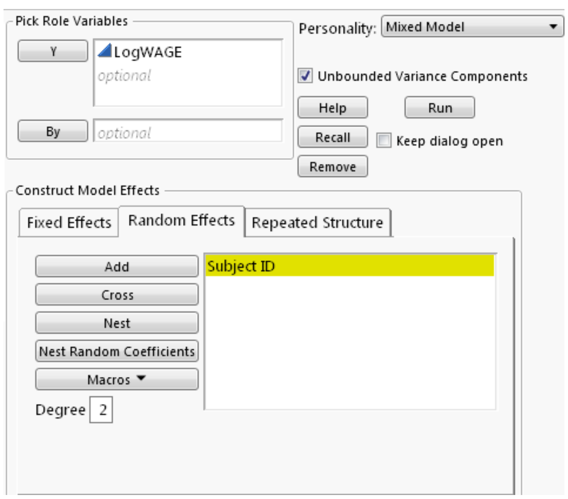
The original data came from the **National Longitudinal Survey of Youth 1979 Cohort** (NLSY79). In the data, each of the 545 male workers worked every year from 1980 through 1987. We're interested in estimating the effect on wage earnings of union membership controlling for education, work experience, ethnicity, etc.

Although NLSY79 collects detailed background information on the workers to be used as control variables, there is still individual difference that cannot be observed or measured. Panel data provides a way of accounting for individual heterogeneity: if the unobserved heterogeneity can be assumed to be uncorrelated with all the explanatory variables included in the model, we can account for it by treating it as a random effect.

Following Woodridge's discussion a Log(Wage) equation is fit in which worker's ID is entered as a random effect to capture the unobserved differences.



(Fixed effects part of the Log(Wage) Equation)



(Random effects part of the Log(Wage) Equation)

I select **Residual** for the model error term. The model is called **one-way random effect model** in econometrics**.**

The results are shown below.

Fit Mixed

Actual by Predicted Plot | Actual by Conditional Predicted Plot

Fit Statistics

-2 Residual Log Likelihood	4473.0746
-2 Log Likelihood	4373.9563
AICc	4408.0972
BIC	4516.4201

Random Effects Covariance Parameter Estimates

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Subject ID	0.1100163	0.0076783	0.0949671	0.1250654
Residual	0.1232764	0.0028279	0.1179162	0.1290123

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	0.1577073	0.2165583	559.0	0.73	0.4668	-0.26766	0.5830748
Years of Education	0.0918895	0.0108335	539.1	8.48	<.0001*	0.0706084	0.1131706
Black	-0.139383	0.0484973	543.8	-2.87	0.0042*	-0.234648	-0.044118
Hispanic	0.0217842	0.0433036	535.8	0.50	0.6151	-0.063281	0.1068498
Married	0.0634614	0.0167953	4293.5	3.78	0.0002*	0.030534	0.0963888
Union member (Yes/No)	0.1053187	0.0178703	4327.8	5.89	<.0001*	0.0702839	0.1403536
Work Experience	0.1060383	0.0154897	1602.6	6.85	<.0001*	0.0756561	0.1364204
Work Experience*Work Experience	-0.00474	0.0006886	4107.4	-6.88	<.0001*	-0.00609	-0.00339
Year(1980)	-0.134647	0.0825171	634.7	-1.63	0.1032	-0.296686	0.0273926
Year(1981)	-0.094303	0.0711689	651.1	-1.33	0.1856	-0.234051	0.0454452
Year(1982)	-0.103939	0.060326	696.5	-1.72	0.0853	-0.222382	0.0145032
Year(1983)	-0.114648	0.0499671	796.9	-2.29	0.0220*	-0.212731	-0.016565
Year(1984)	-0.091851	0.0401671	1034.7	-2.29	0.0224*	-0.170669	-0.013033
Year(1985)	-0.077197	0.0312602	1758.3	-2.47	0.0136*	-0.138508	-0.015886
Year(1986)	-0.043066	0.0242344	4019.7	-1.78	0.0756	-0.090579	0.0044465

(Panel Model Results)

From the **Random Effects Covariance Parameter Estimates** report we find that individual heterogeneity accounts for 47.8% (=0.11/(0.11+0.12)) of the total variation, indicating a large unobserved heterogeneity effect. In other words, an OLS analysis would likely yield misleading results.

The **Fixed Effects Parameter Estimates** report shows an estimated rate of return to education at 9.2% and a union premium of 10.5%, both of which are highly statistically significant. As a comparison, a pooled OLS would estimate the union premium at 18.2%. See Woodridge (2013, Page 495).

References

Francis Vella and Marno Verbeek (1998), "*Whose Wages Do Unions Raise? A Dynamic Model of Unionism and Wage Rate Determination for Young Men*", Journal of Applied Econometrics, Vol. 13, No. 2, pp. 163-183. Data can be downloaded from Journal's website
http://qed.econ.queensu.ca/jae/1998-v13.2/vella-verbeek/

Jeffrey M. Woodridge (2013), *Introductory Econometrics: A Modern Approach (5th ed)*, CENGAGE Learning.

Example 4: Modeling geospatial data--taking spatial correlation into account

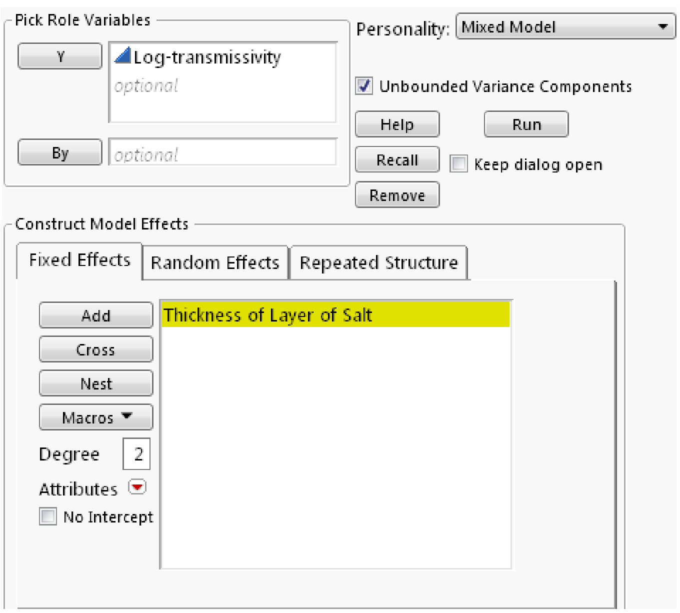
Like repeated measures are correlated over time, spatial data are likely correlated in space. That is, measurements that are relatively close together are more alike than those farther apart. Thus, we need to take spatial dependency into account in the analysis.

Spatial data are recorded along with coordinates such as latitude and longitude, positions of row and column, north-south and east-west directions. The distance between two measurements are calculated using a Euclidean distance function, which is used to form a covariance structure. If a distance function doesn't depend on the directions of measurements, then the covariance is said to be *isotropic;* otherwise it is *anisotropic*. In addition, a *nugget* effect can be added to account for abrupt changes over small distances in a local area.

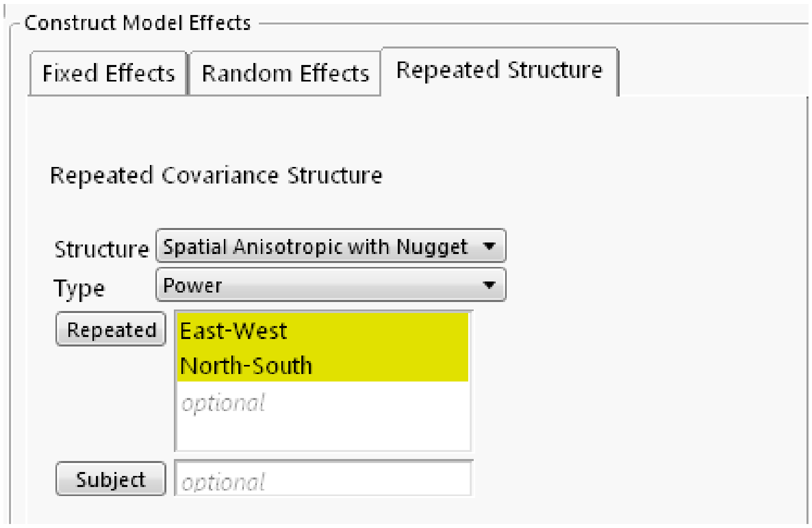
JMP Pro provides four Euclidean distance functions for isotropic structures: power, exponential, Gaussian and spherical. Various forms of anisotropic structures are available. A nugget effect can also be added to covariance structures.

The following example is taken from *SAS for Mixed Models, 2nd Edition*, 2006, pp. 457-460. (http://www.sas.com/store/prodBK_59882_en.htm). In order to investigate the water drainage at a hazardous waste disposal site, 30 samples were taken at various locations at the site and recorded by their north-south and east-west directions. A linear relationship between water drainage (measured by *log-transmissivity)* and the thickness of a layer of salt was proposed. (The data, *Hazardous Waste*, is Data Set 11.6 in the zipped file http://support.sas.com/publishing/bbu/59882/59882.zip.)

A spatial regression model is fit using a spatial anisotropic power structure with a nugget effect. This structure allows (1) distance to be a power function of spatial correlation, (2) spatial correlations to differ in different directions, and (3) variation over small distances.



(Fixed effects part of the model)



Fit Mixed (Spatial anisotropic power with nugget)

Actual by Predicted Plot

Fit Statistics

-2 Residual Log Likelihood	78.631234
-2 Log Likelihood	70.747191
AICc	86.399365
BIC	91.154375

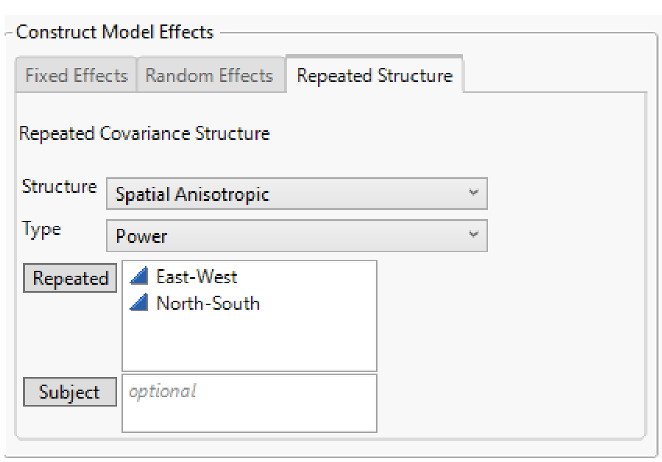
Repeated Effects Covariance Parameter Estimates

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Spatial Power East-West	0.8052924	0.1535787	0.5042838	1.1063011
Spatial Power North-South	0.9149311	0.0656425	0.7862741	1.0435881
Nugget	0.042169	0.0475184	-0.050965	0.1353033
Residual	1.4913473	0.6745013	0.7232787	4.6686809

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	-4.946099	0.5237702	2.7	-9.44	0.0039*	-6.727095	-3.165102
Thickness of Layer of Salt	-0.02503	0.0075612	21.2	-3.31	0.0033*	-0.040746	-0.009313

Covariance Parameter Estimate report suggests highly significant spatial correlation and that two correlation coefficients are, respectively, 0.81 (East-West) and 0.91 (North-South). However, there appears to have no nugget based on the confidence limits. **Fixed Effects Parameter Estimates** show a significant negative effect (-0.025) on water drainage of thickness of salt. We refit the model by removing the nugget effect**:**



Fit Statistics Repeated Structure = Spatial Anisotropic without Nugget

-2 Residual Log Likelihood	80.346038
-2 Log Likelihood	72.197131
AICc	84.697131
BIC	89.203118

Repeated Effects Covariance Parameter Estimates

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Spatial Power East-West	0.864312	0.1113235	0.646122	1.0825021
Spatial Power North-South	0.8165802	0.1486885	0.5251561	1.1080043
Residual	1.6477606	0.7543611	0.7933988	5.251859

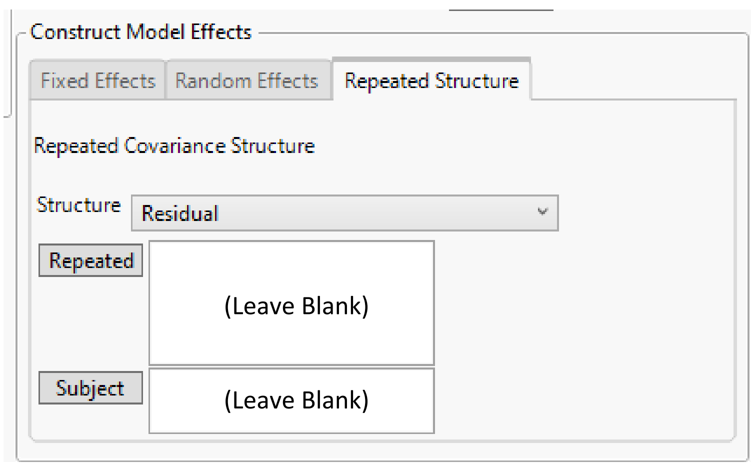
Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	-4.920759	0.5027641	2.4	-9.79	0.0053*	-6.759791	-3.081727
Thickness of Layer of Salt	-0.021282	0.007246	17.1	-2.94	0.0092*	-0.03656	-0.006003

(Results using spatial anisotropic power without nugget)

We notice some minor changes: the estimated spatial correlations are 0.86 and 0.82, respectively, and effect of thickness of salt is -0.021. Comparing the goodness of fit using AICc, the second model is slightly better as its AICs is smaller (86.4 vs. 84.7).

To formally test the existence of spatial correlation, we fit an independent errors model by selecting **Residual** as the structure (i.e. assuming no spatial correlation).



The difference in -2 Residual Log Likelihood value between the two models forms a $\chi 2$ likelihood ratio test. The -2 Residual Log Likelihood from the independent errors model is 94.07, so the difference is a 13.72 (=94.07-80.35). This yields a p-value of 0.001 for DF=2. Therefore, significant spatial correlation is found at this site.

Mixed Model for Log-transmissability Independent Errors Model (Structure = Residual)

Fit Statistics

-2 Residual Log Likelihood	94.068661
-2 Log Likelihood	84.336683
AICc	91.25976
BIC	94.540276

Repeated Effects Covariance Parameter Estimates

Covariance Parameter	Estimate	Std Error	95% Lower	95% Upper
Residual	1.0408237	0.2781718	0.6554778	1.9037972

Fixed Effects Parameter Estimates

Term	Estimate	Std Error	DFDen	t Ratio	Prob>|t|	95% Lower	95% Upper
Intercept	-5.024542	0.2016519	28.0	-24.92	<.0001*	-5.437607	-4.611477
Thickness of Layer of Salt	-0.033859	0.0065829	28.0	-5.14	<.0001*	-0.047344	-0.020375

Fixed Effects Tests

Source	Nparm	DFNum	DFDen	F Ratio	Prob > F
Thickness of Layer of Salt	1	1	28.0	26.455749	<.0001*

Summary

Hopefully, these examples have illustrated the versatility of linear mixed models and ease of fitting a mixed model with JMP. Before I close I'd like to share some general tips.

- In order to run a mixed model, data needs to be organized in a "tall & skinny" format where multiple measures of a response are stacked into a single column. If your data is in a "short & wide" format, use the JMP Tables function **Stack** to transpose.
- Follow the **JMP Repeated Covariance Structure Requirements** when entering Repeated and Subject columns. http://www.jmp.com/support/help/Launch_the_Mixed_Model_Personality.shtml#1013652
- Try different covariance structures and evaluate different models by comparing AICc or BIC. The smaller the AICc (or BIC), the better fit of a model. *Ceteris paribus*, a parsimonious model is better.
- An independent errors model (i.e. a model with only fixed effects and a Residual repeated structure) can serve as a baseline model to perform a $\chi 2$ likelihood ratio test on the existence of a covariance structure.
- Keep in mind that when both Random effects and Repeated effects are included in a model there is often insufficient data to estimate both effects.