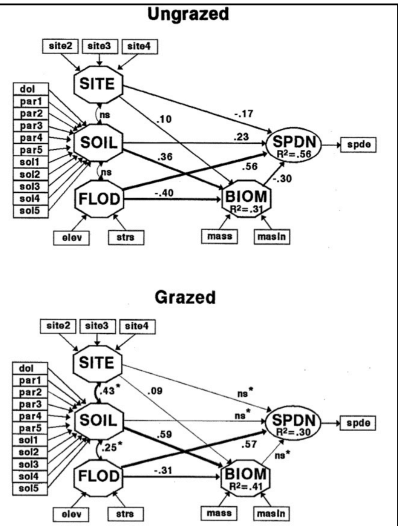
Multigroup Models in SEM

Jarrett E. K. Byrnes



Multigroup Outline

1. Pooling data from multiple sources for more power.
2. Multigroup analysis as model-wide interaction effect
3. Example from Finland

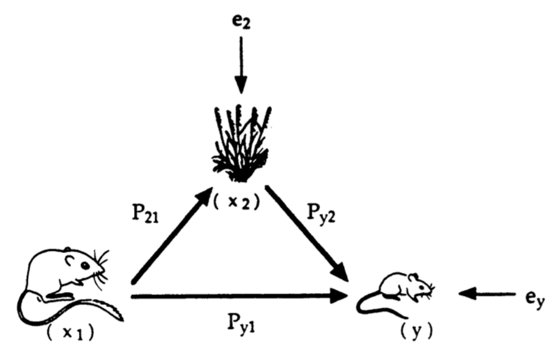
Multigroup Analysis: Constraining Variables Across Groups

Let's say you have data from two sites.

2 possibilities:

or

Example: The Effects of Kangaroo Rats on Other Rodents



Smith et al. 1997 Am. Nat.

Smith et al. – Path Analytic Results Wildly Vary due to Technique

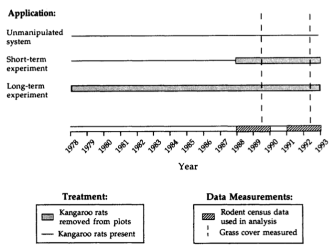


TABLE 3

CALCULATED PATH EQUATIONS OF THE EFFECT OF KANGAROO RATS ON OTHER RODENTS

Treatment	Path Equation	e_y
Harvest mice:		
A	$-.303 = -.230 - .073$	.85
B	$-.710 = -.723 + .014$	.66
C	$-.760 = -.450 - .310$	.57
Pocket mice:		
A	$-.212 = -.150 - .062$	.91
B	$-.554 = -.561 + .006$	.82
C	$-.362 = .045 - .407$	.67

- Total standardized Coefficients varied wildly
- Indirect effects not detected in 2 of 3 path models

Grace and Pugesek: Multigroup Analysis Disagrees

Harvest mice	*Pocket mice*
A. Multigroup goodness-of-fit statistic	
N for multigroup analysis = 38	N for multigroup analysis = 38
χ^2 with 9 df = 7.8400 (P = .5503)	χ^2 with 8 df = 9.6308 (P = .2919)
Group A χ^2 =3.2771	Group A χ^2 =4.6998
Group B χ^2 =1.7419	Group B χ^2 =2.2934
Group C χ^2 =2.8210	Group C χ^2 =2.6376

• Differences in std. coefs were due to differences in range of data.

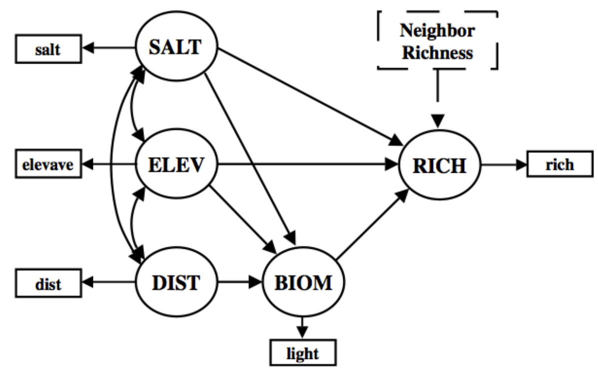
• Indirect effects detected, largely due to data from long-term observations.

• Smith et al. replied that this ***still*** meant that Path Analysis could provide bogus results without the proper data.   Oy.

Multigroup Outline

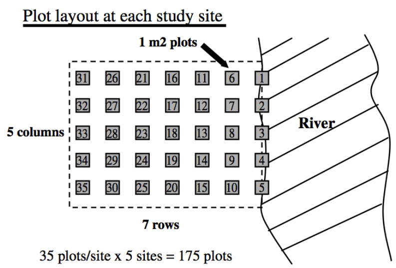
1. Pooling data from multiple sources for more power.
2. Multigroup analysis as model-wide interaction effect
3. Example from Finland

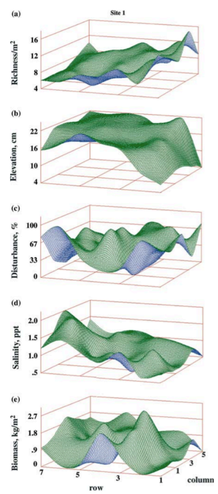
Example: Is Richness Determined by the Same Factors at Different Sites?



Mancera et al. 2005

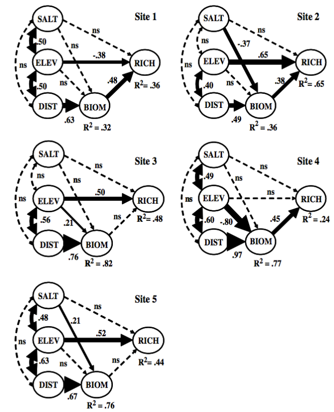
Wide Variation Within all 5 Sites





Mancera et al. 2005

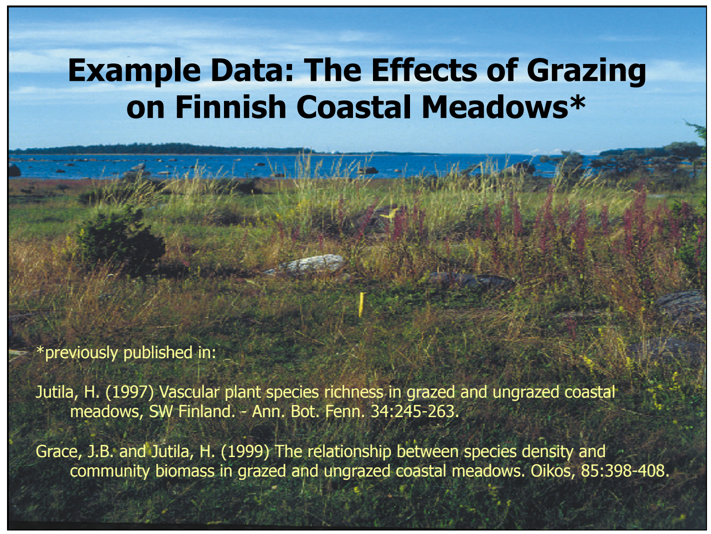
Variation in Determinants of Species Richnes – but each mechanism is important somewhere…



Mancera et al. 2005

Multigroup Outline

1. Pooling data from multiple sources for more power.
2. Multigroup analysis as model-wide interaction effect
3. Example from Finland



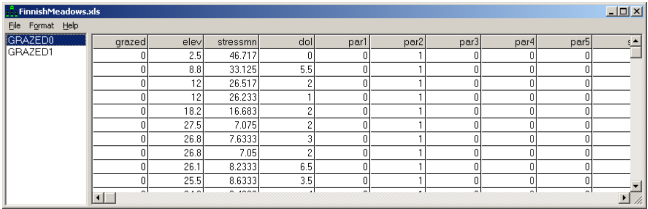
Example Data: The Effects of Grazing on Finnish Coastal Meadows*

*previously published in:

Jutila, H. (1997) Vascular plant species richness in grazed and ungrazed coastal meadows, SW Finland. - Ann. Bot. Fenn. 34:245-263.

Grace, J.B. and Jutila, H. (1999) The relationship between species density and community biomass in grazed and ungrazed coastal meadows. Oikos, 85:398-408.

View of Data in FinnishMeadows.xls



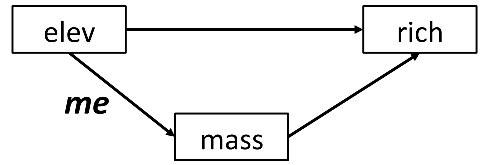
grazed	elev	stressmn	dol	par1	par2	par3	par4	par5
0	2.5	46.717	0	0	1	0	0	0
0	8.8	33.125	5.5	0	1	0	0	0
0	12	26.517	2	0	1	0	0	0
0	12	26.233	1	0	1	0	0	0
0	18.2	16.683	2	0	1	0	0	0
0	27.5	7.075	2	0	1	0	0	0
0	26.8	7.6333	3	0	1	0	0	0
0	26.8	7.05	2	0	1	0	0	0
0	26.1	8.2333	6.5	0	1	0	0	0
0	25.5	8.6333	3.5	0	1	0	0	0

Data from 1-m2 plots arrayed along an elevation gradient in each of several paired grazed and ungrazed meadows in SW Finland.

grazed = 0 is no, 1 is yes (this is our grouping variable)
elev = elevation of the plot above mean sea level
stressmn = flood stress index derived from long-term flooding records
dol = depth of litter layer in the plot
par1 - par5 = different parent materials
sol1 - sol5 = different soil types
mass, mass2, masslog = biomass in g/m2, square of biomass, and log biomass
rich, rich2, richlog = species richness per m2, square of richness and log richness

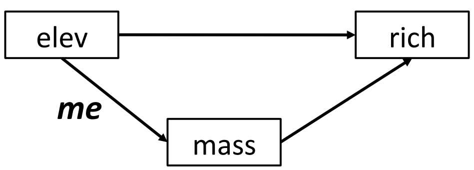
13

Giving a Path a Name



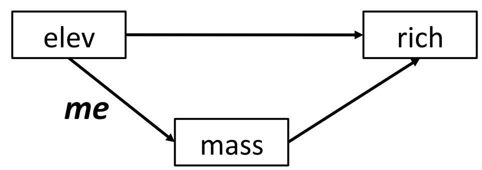

```
meadowModel<-'rich ~ elev + mass
              mass ~ me*elev'
```

Giving a Path a Name



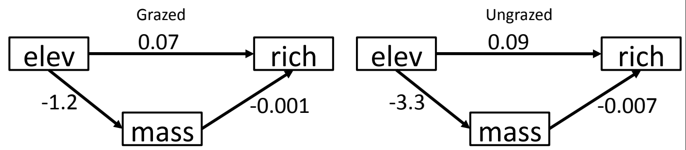

```
> meadowFit<-sem(meadowModel, data=meadows)
>
> coef(meadowFit)
 rich~elev  rich~mass         me rich~~rich mass~~mass
     0.073     -0.003     -2.733     16.275  39563.789
```

Giving a Path a Name




```
                   Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  rich ~
    elev              0.073    0.008    9.026    0.000
    mass             -0.003    0.001   -2.988    0.003
  mass ~
    elev      (me)   -2.733    0.372   -7.352    0.000
```

Multiple Groups




```
> meadowFitFree<-sem(meadowModel, data=meadows,
  group="grazed")

> coef(meadowFitFree)
   rich~elev     rich~mass            me    rich~~rich    mass~~mass        rich~1
       0.073        -0.001        -1.203        12.459     28057.591         7.169
      mass~1  rich~elev.g2  rich~mass.g2  mass~elev.g2 rich~~rich.g2 mass~~mass.g2
     260.855         0.088        -0.007        -3.274        14.384     43567.993
   rich~1.g2     mass~1.g2
      11.349       451.732
```

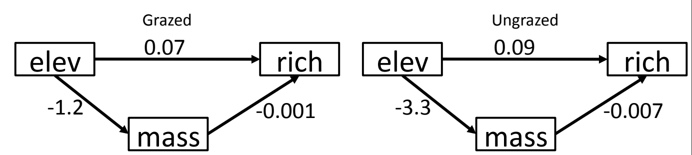
summary(meadowFitFree)

```
lavaan (0.4-11) converged normally after 65 iterations

  Number of observations per group
  1                                                165
  0                                                189

  Estimator                                         ML
  Minimum Function Chi-square                    0.000
  Degrees of freedom                                 0
  P-value                                        1.000

Chi-square for each group:

  1                                              0.000
  0                                              0.000
```

summary(meadowFitFree) continued

```
Group 1 [1]:

                   Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  rich ~
    elev              0.073    0.010    7.232    0.000
    mass             -0.001    0.002   -0.424    0.672
  mass ~
    elev      (me)   -1.203    0.470   -2.559    0.010

Intercepts:
    rich              7.169    0.708   10.126    0.000
    mass            260.855   26.764    9.746    0.000

Variances:
    rich             12.459    1.372
    mass          28057.591 3089.039
```

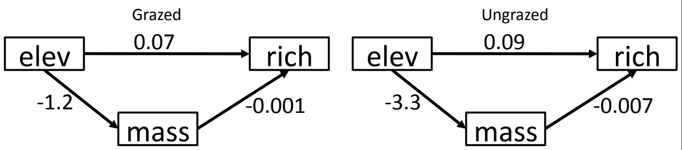
Testing Complete Constraints




```
> meadowFitEqual<-sem(meadowModel, data=meadows, group="grazed",
  group.equal=c( "intercepts", "means", "regressions",
  "residuals", "residual.covariances"))
> anova(meadowFitFree, meadowFitEqual)
Chi Square Difference Test

               Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
meadowFitFree   0 10019 10073   0.00
meadowFitEqual  6 10107 10138 100.13     100.13       6  < 2.2e-16
```

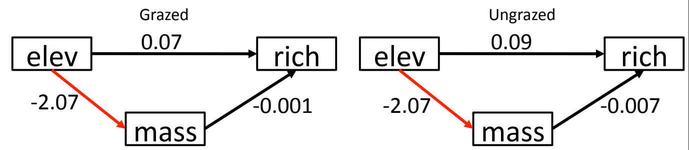
Model Differs Between Treatments

Testing One Constraint



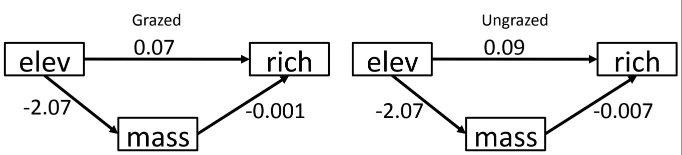

```
#constrain just the elev-mass relationship
meadowModel2<-'rich ~ elev + mass
               mass ~ c("me", "me")*elev'

meadowFitFree2<-sem(meadowModel2, data=meadows,
                    group="grazed")
```

Testing One Constraint




```
> coef(meadowFitFree2)
   rich~elev     rich~mass            me    rich~~rich    mass~~mass       rich~1 
       0.073        -0.001        -2.067        12.459     28632.459        7.169 
      mass~1 rich~elev.g2  rich~mass.g2 rich~~rich.g2 mass~~mass.g2   rich~1.g2 
     303.829         0.088        -0.007        14.384     44661.148       11.349 
   mass~1.g2 
     408.622 
```

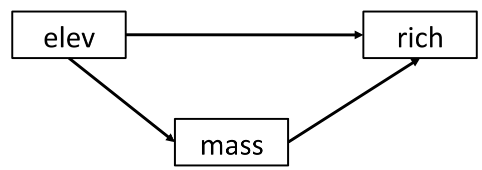
Testing One Constraint




```
> anova(meadowFitFree, meadowFitFree2)
Chi Square Difference Test

               Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
meadowFitFree   0 10019 10073 0.0000                              
meadowFitFree2  1 10025 10075 8.0301     8.0301       1   0.004601
```

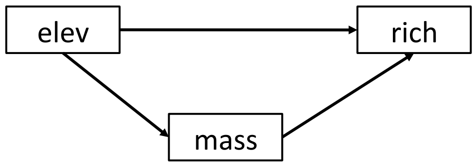
Treatment Affects elev -> mass relationship

Multigroup Exercise



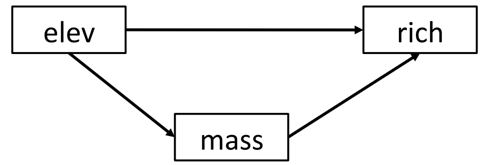
1. Fit a free and constrained version of the meadow model with par2 as your group.
2. Evaluate a single constraint of your choice.

Solution: The Fits



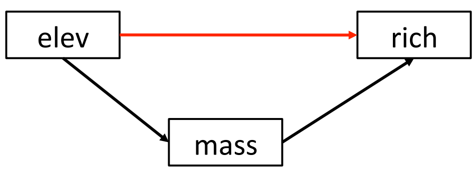

```
> meadowFitFreePar1<-sem(meadowModel, data=meadows,
    group="par2")

> meadowFitEqualPar1<-sem(meadowModel, data=meadows,
    group="par2", group.equal=c( "intercepts",
    "means", "regressions", "residuals",
    "residual.covariances"))

> anova(meadowFitEqualPar1, meadowFitFreePar1)
```

Solution: Par1 Matters!



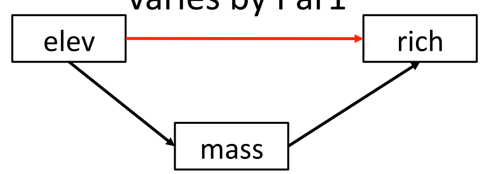

```
Chi Square Difference Test

                   Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
meadowFitFreePar1   0 10072 10127  0.000
meadowFitEqualPar1  6 10091 10122 30.774     30.774       6  2.799e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Solution: Testing elev->rich constraint



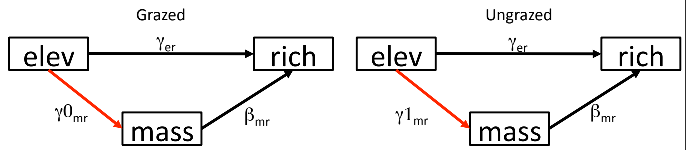

```
meadowModel3<-'rich ~ c("re", "re")*elev + mass
              mass ~ elev'

meadowFitFreePar1_3<-sem(meadowModel3, data=meadows,
    group="par2")

#does it matter?
anova(meadowFitFreePar1, meadowFitFreePar1_3)
```

Solution: Elev-> Richness relationship varies by Par1



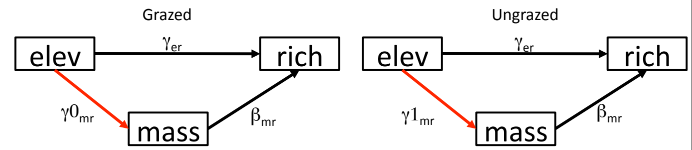

```
Chi Square Difference Test

                    Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
meadowFitFreePar1    0 10072 10127  0.000
meadowFitFreePar1_3  1 10080 10131 10.137     10.137       1   0.001453 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Testing One Release



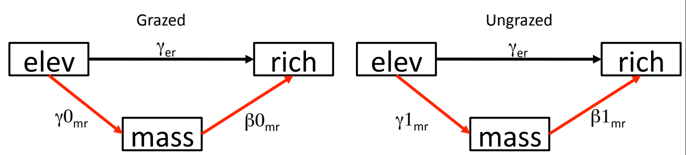

```
meadowModelNoLabel<-'rich ~ elev + mass
              mass ~ elev'

meadowFitOneFree<-sem(meadowModelNoLabel, data=meadows, group="grazed",
                      group.equal="regressions",
                      group.partial="mass ~ elev")
```

Testing One Release



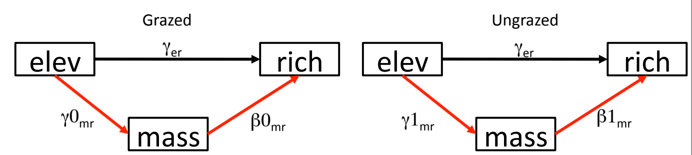

```
> meadowFitOneFree
lavaan (0.4-12) converged normally after 46 iterations

  Number of observations per group
  1                                                165
  0                                                189

  Estimator                                         ML
  Minimum Function Chi-square                   13.085
  Degrees of freedom                                 2
  P-value                                        0.001
```

Testing Two Releases



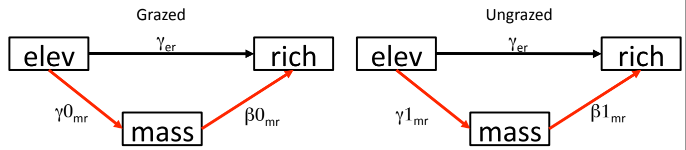

```
meadowFitTwoFree<-sem(meadowModelNoLabel, data=meadows,
        group="grazed", group.equal="regressions",
         group.partial=c("mass ~ elev", "rich ~ mass"))
```

Testing Two Releases




```
lavaan (0.4-12) converged normally after 59 iterations

  Number of observations per group
  1                                                165
  0                                                189

  Estimator                                         ML
  Minimum Function Chi-square                    0.948
  Degrees of freedom                                 1
  P-value                                        0.330
```

Fit Not Different Than Completely Unconstrained Model




```
> anova(meadowFitTwoFree, meadowFitFree)
Chi Square Difference Test

                 Df   AIC   BIC  Chisq Chisq diff Df diff Pr(>Chisq)
meadowFitFree     0 10019 10073 0.0000
meadowFitTwoFree  1 10018 10068 0.9477    0.94767       1     0.3303
```

Care and Feeding of Multigroup Analysis

- Disparate results can be produced by different groups encompassing different ranges of variability.
- Variation in one group can reinforce weak patterns in another group.
- Need to have adequate sample size for each group!

What should I constrain? What should I test?

1. What are you interested in constraining? Typically just regression parameters.
2. Test a free v. constrained model.
3. Evaluate releasing key parameters or constraining key parameters, based on questions.

Formal Multigroup Analysis Typically Considers Batches of Parameters

Typically test constraints sequentially:

1. Latent variable loadings
2. Correlated errors
3. Structural path coefficients
4. Variance parameter estimates
5. Endogenous latent variances and covariances
6. Exogenous latent variances and covariances