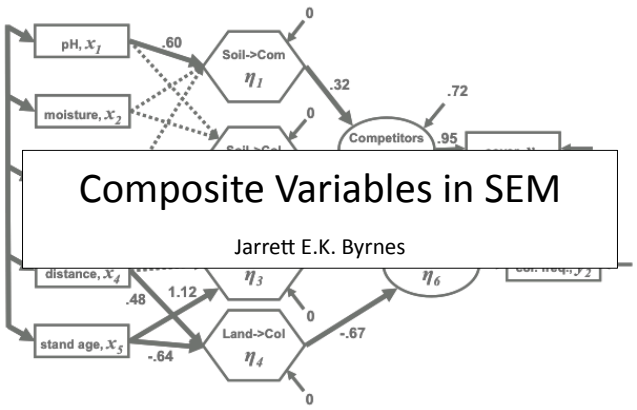


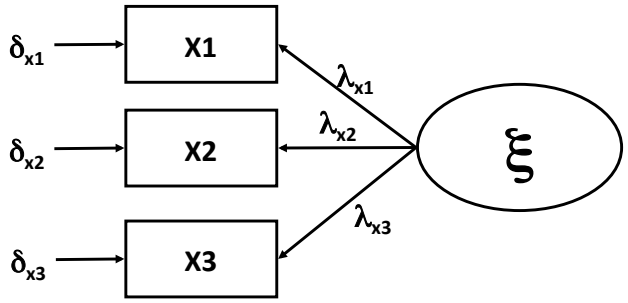
Composite Variables in SEM

Jarrett E.K. Byrnes

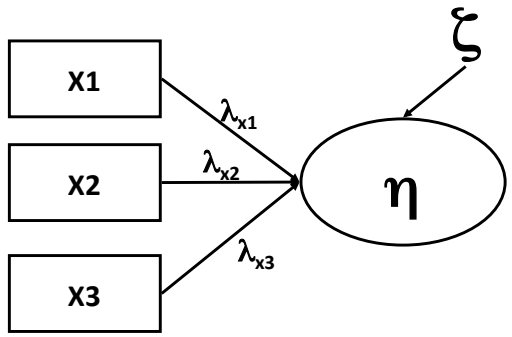
To SEM and Beyond!

1. What is a composite variable?
2. Using Composites for nonlinear variables
3. Composites v. Latents - when and why?
4. Comparison in context
5. Treatment as a Composite variable

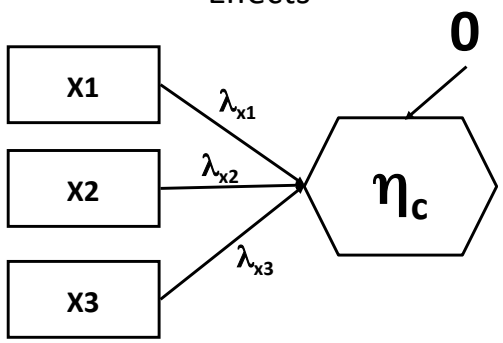
We're Comfortable with Latent Exogenous Variables…



Latent Composite

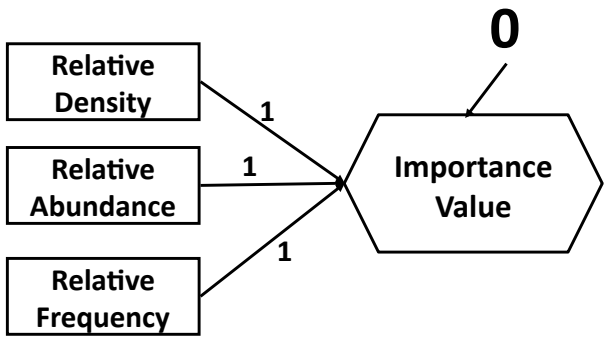


Composite Variables Reflect Joint Effects



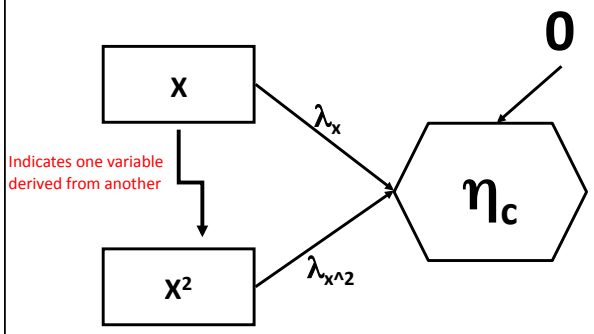
- **Coefficients can be statistically estimated.**
- **Fixing error to 0 aids in identification (otherwise it's a latent composite)**
- **Specifying a scale is often helpful.**

Coefficients Can be Fixed

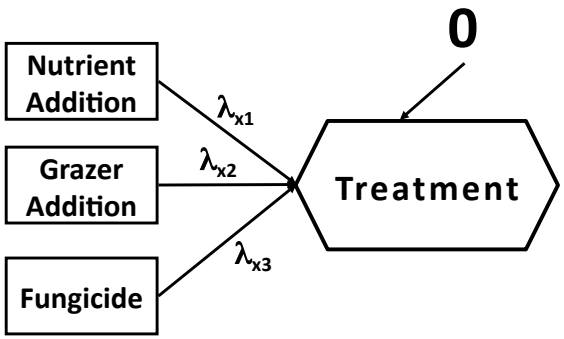


Easy way to incorporate concepts into models, particularly if exogenous variables have effects beyond the composite variable.

Composites and Nonlinearities



Composites and Categorical Variables



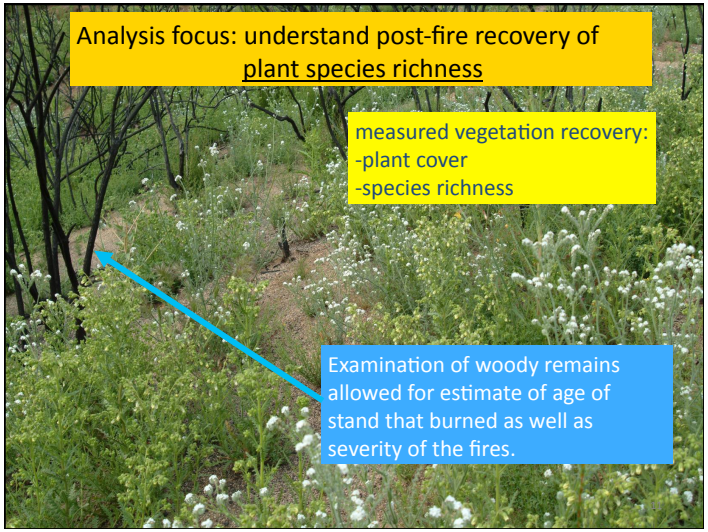
Categories coded as 0 or 1

To SEM and Beyond!

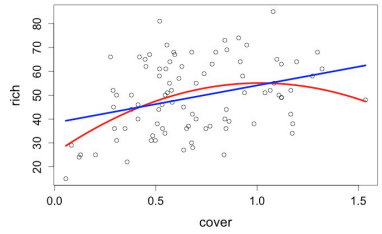
1. What is a composite variable?
2. Using Composites for nonlinear variables
3. Composites v. Latents - when and why?
4. Comparison in context
5. Treatment as a Composite variable

Mediation in Analysis of Post-Fire Recovery of Plant Communities in California Shrublands*

*Five year study of wildfires in Southern California in 1993. 90 plots (20 x 50m), (data from Jon Keeley et al.)

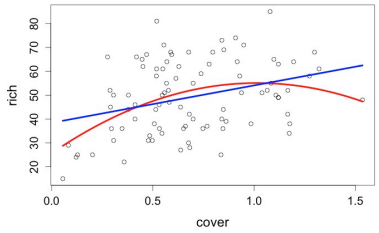


Linear or Nonlinear?



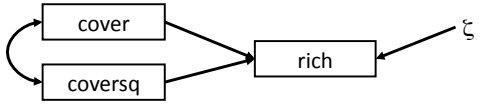

```
linear<-lm(rich ~ cover, data=keeley)

nonlinear<-lm(rich ~ cover+I(cover^2), data=keeley)

aictab(list(linear, nonlinear), c("linear",
  "squared"))
```

Linear or Nonlinear?



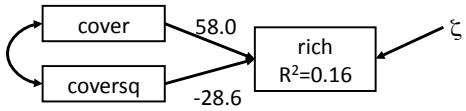

```
Model selection based on AICc :

          K   AICc Delta_AICc AICcWt Cum.Wt      LL
squared   4 735.92       0.00   0.83   0.83 -363.72
linear    3 739.08       3.15   0.17   1.00 -366.40
```

A Simple Nonlinear Model



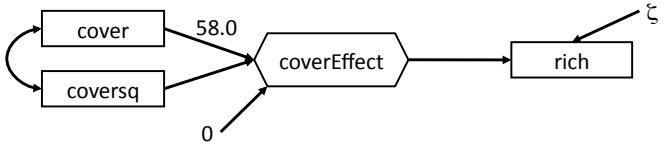

```
#Create a new nonlinear variable in the data
keeley<-within(keeley, coverSQ<-cover^2)

#Now, for a model
noCompModel <- 'rich ~ cover + coverSQ'

noCompFit <- sem(noCompModel, data=keeley)
```

A Simple Nonlinear Model



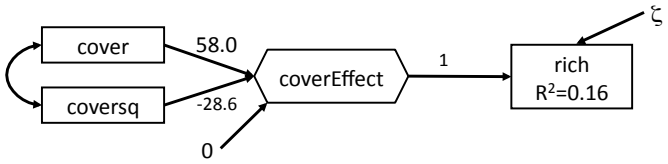

```
> summary(noCompFit)
…
                  Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  rich ~
    cover           57.999   18.613    3.116    0.002
    coverSQ        -28.577   12.176   -2.347    0.019
```

A Simple Composite Model



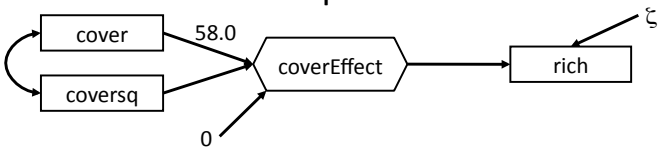

```
compModel<-'

          coverEffect <~ 58*cover + coverSQ

         rich ~ coverEffect'

compFit <- sem(compModel, data=keeley)
```

A Simple Composite Model



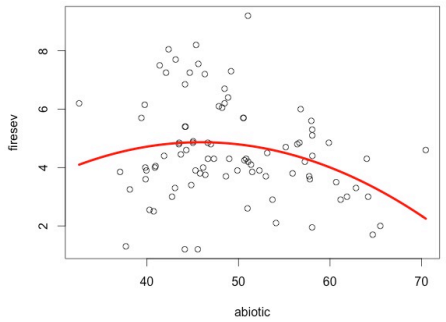

```
                Estimate  Std.err  Z-value  P(>|z|)
Composites:
  coverEffect <~
    cover         58.000
    coverSQ      -28.577    3.996   -7.152    0.000

Regressions:
  rich ~
    coverEffect     1.000    0.321    3.116    0.002

Variances:
    rich          189.597   28.263
```

A Note About the Latent Nature of Composite

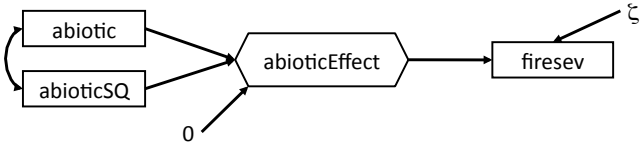


- Response variables act like latent variable indicators
- Therefore, responses must *share* some variance.
- Rules that applied to identifiably of latent variables also apply to composites.
- One composite per response if composite error = 0. If composite has multiple responses, error variance should be free.

Exercise: A Nonlinear Relationship Between abiotic and firesev

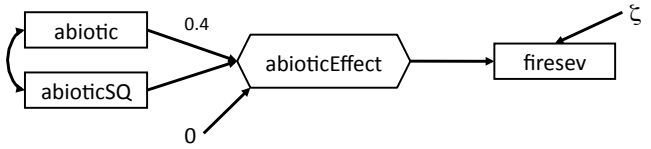


Exercise: An Abiotic Composite Model



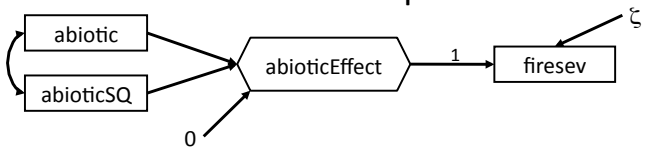
1. Fit this model – start with a regression
2. Compare the effect of fixing the abiotic loading on abiotic effect to the coefficient from the regression to fixing the abioticEffect on firesev to 1.

For some reason, this model fails



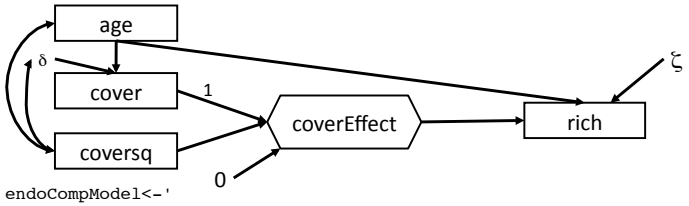

```
keeley$abioticSQ <- keeley$abiotic^2

abioticCompModelBad<-'
                    abioticEffect <~ 0.4 * abiotic +
                                     abioticSQ

                   firesev ~ abioticEffect'

abioticCompFitBad <- sem(abioticCompModelBad, data=keeley)
```

This model does not: try multiple methods with composites!



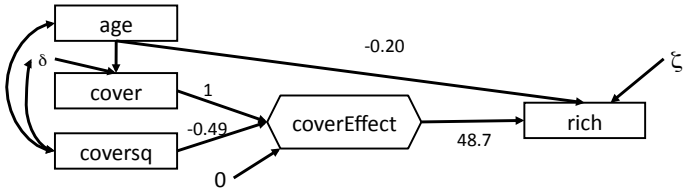

```
keeley$abioticSQ <- keeley$abiotic^2

abioticCompModel<-'
                    abioticEffect <~ abiotic + abioticSQ

                   firesev ~ 1*abioticEffect'

abioticCompFit <- sem(abioticCompModel, data=keeley)
```

Endogenous Composite Variables




```
endoCompModel<-'
                coverEffect <~ 1*cover + coverSQ

                cover ~~ coverSQ
                age ~~ coverSQ

                cover ~ age
                rich ~ age + coverEffect'


endoCompFit <- sem(endoCompModel, data=keeley, fixed.x=F)
```

Endogenous Composite Variables

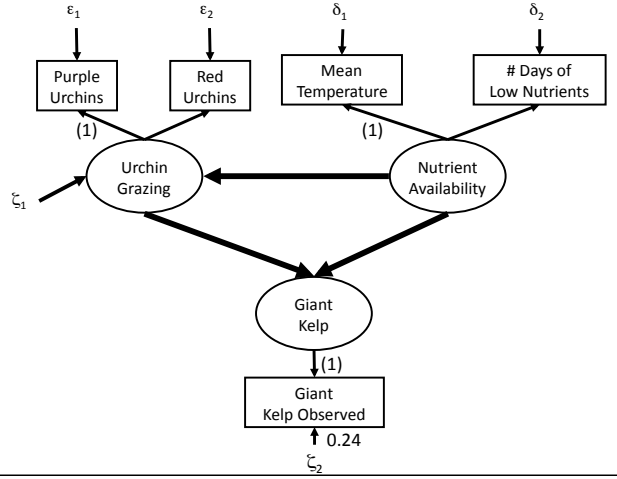



```
                     Estimate  Std.err  Z-value  P(>|z|)
Composites:
  coverEffect <~
    cover               1.000
    coverSQ            -0.497    0.078   -6.378    0.000

Regressions:
  cover ~
    age                -0.002    0.001   -3.129    0.002
  rich ~
    age                -0.201    0.121   -1.667    0.095
    coverEffect        48.705   19.246    2.531    0.011
```

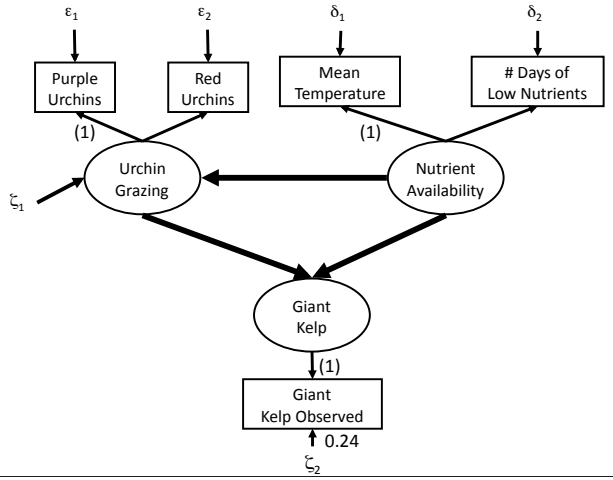
To SEM and Beyond!

1. What is a composite variable?
2. Using Composites for nonlinear variables
3. Composites v. Latents - when and why?
4. Comparison in context
5. Treatment as a Composite variable

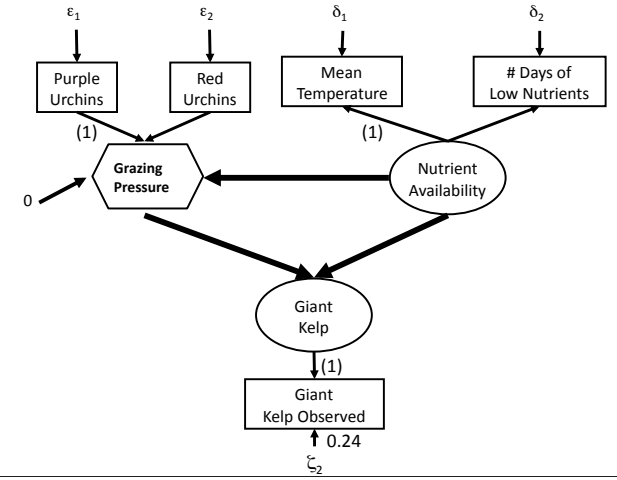
Consider This Model…



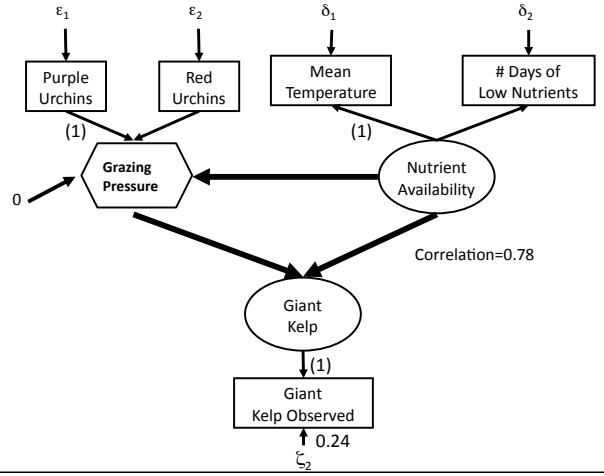
What is the direction of causality?



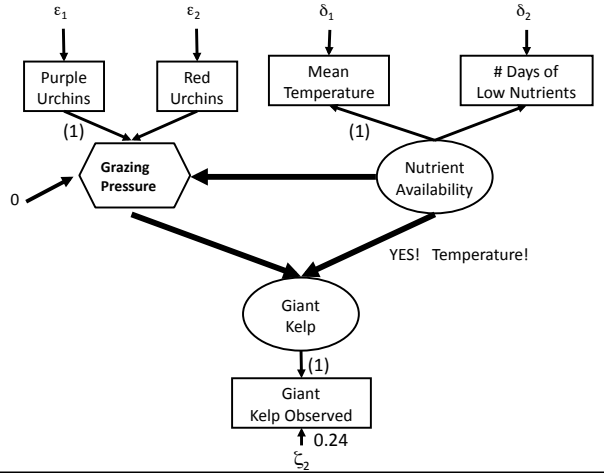
What is the direction of causality?



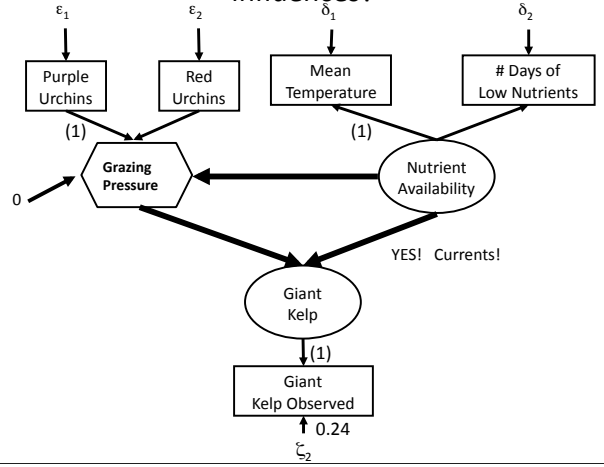
Are the indicators in a block interchangeable?



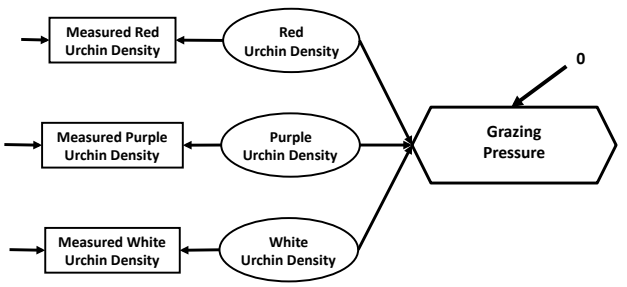
Do indicators covary because of joint causes?



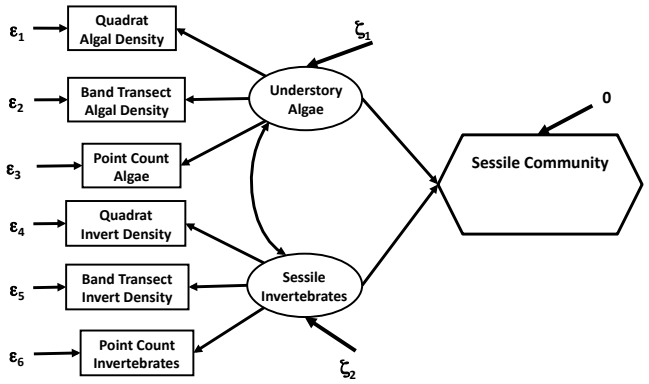
Do indicators have a consistent set of causal influences?



Latents and Composites Together: L-C Block for Measurement Error



Latents and Composites Together: L-C Block



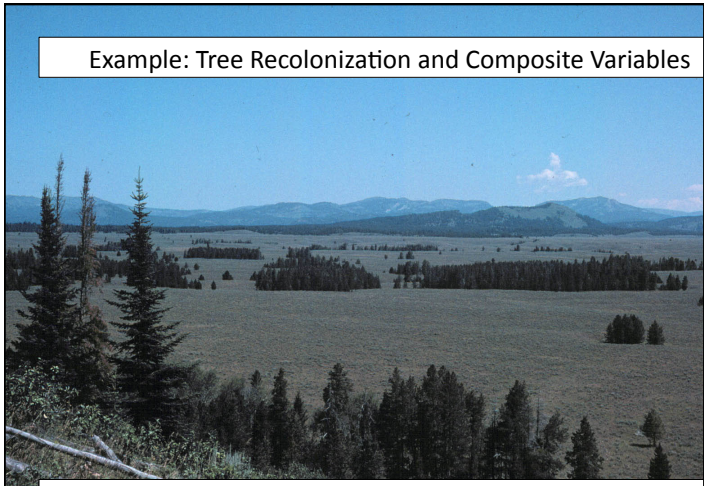
Questions to Ask of Your Latent/Composite Variables

1. What is the direction of causality?
2. Are the indicators in a block interchangeable?
3. Do indicators covary because of joint causes?
4. Do indicators have a consistent set of causal influences?

To SEM and Beyond!

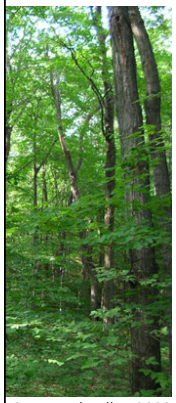
1. What is a composite variable?
2. Using Composites for nonlinear variables
3. Composites v. Latents - when and why?
4. Comparison in context
5. Treatment as a Composite variable

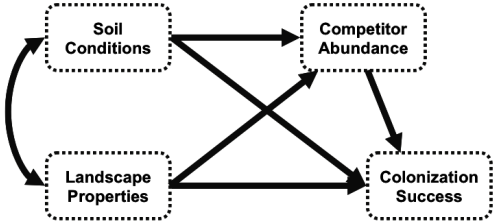
Example: Tree Recolonization and Composite Variables



Grace, J.B. & Bollen, K.A. (2008). Representing general theoretical concepts in structural equation models: the role of composite variables. *Environ. Ecol. Stat., 15, 191–213.*

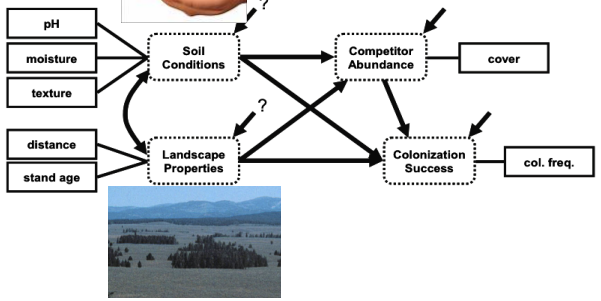
What is the Contribution of Local versus Regional Factors to Recolonization





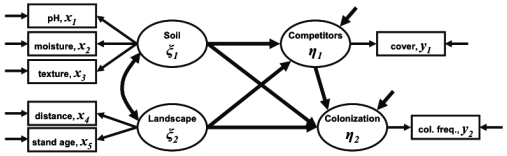
Grace and Bollen 2008

Adding Variables to the Metamodel

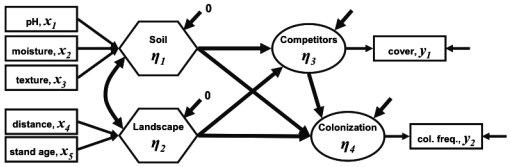


Grace and Bollen 2008

Latent or Composites?

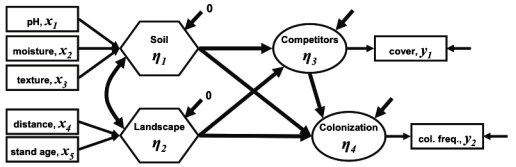


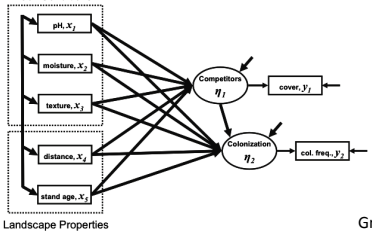
What do you think?



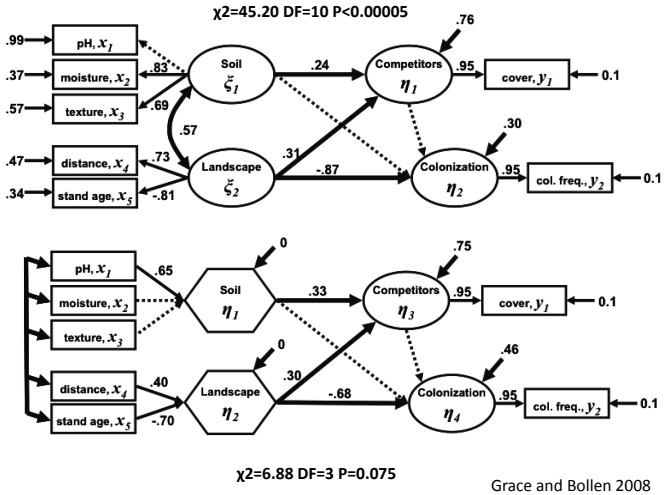
Grace and Bollen 2008

Generality v. Specificity



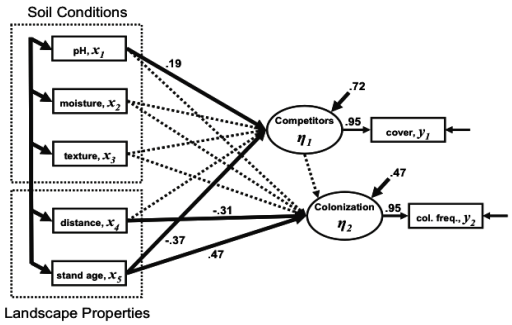


Grace and Bollen 2008



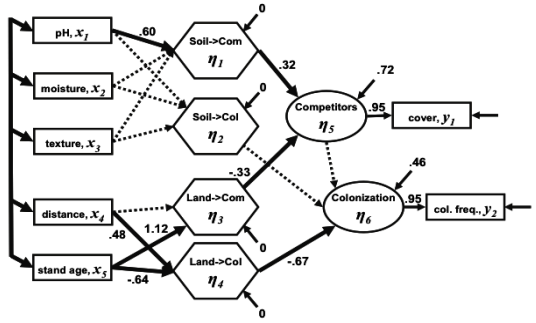
Grace and Bollen 2008

How Confident are We in Composite Loadings and their Conclusions?



Specific model without composites provides similar answers.

Testing our Confidence in Composites

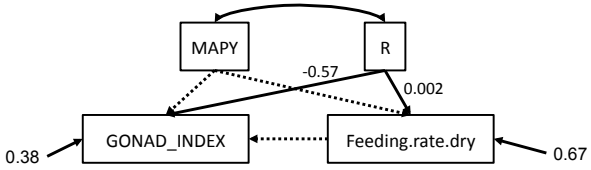


The general composite construct is not obscuring more specific relationships in the data.

To SEM and Beyond!

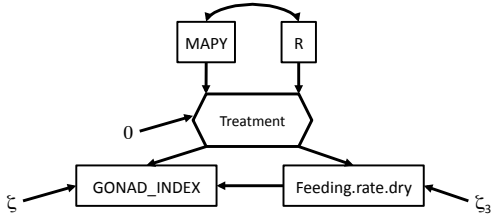
1. What is a composite variable?
2. Using Composites for nonlinear variables
3. Composites v. Latents - when and why?
4. Comparison in context
5. Treatment as a Composite variable

Previous Model Unstandardized



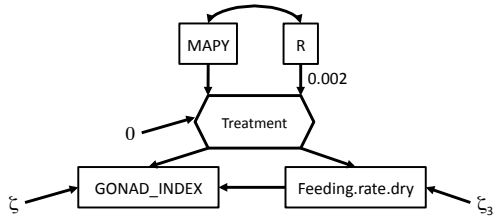
1. *Rhodymenia* is not good food.
   - Urchins eat more, but produce less gonad
2. Performance is similar with *Macrocystis* or Mixture diet

Treatment as a Composite Affecting Multiple Responses




```
#read in and binary-ize the treatment
urchinData<-read.csv("./urchin_ex_sem.csv")
source("./makeBinaryTreatments.R")

binTrt<-makeBinaryTreatments(urchinData, "treatment")

urchinData<-cbind(urchinData, binTrt)
```

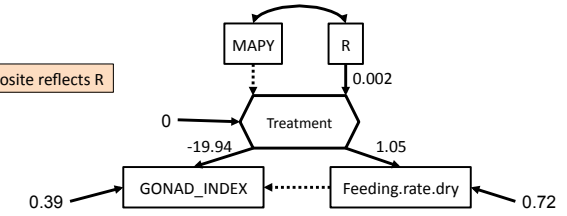
A Composite Treatment Model




```
urchinCompositeModel<-'
  Treatment <~ MAPY + .002*R

  Feeding.rate.dry ~ Treatment
  GONAD_INDEX ~Treatment + Feeding.rate.dry
'
```

MAPY Has No Effect

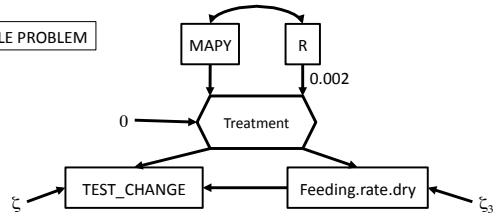
Composite reflects R




```
lavaan (0.5-12) converged normally after  71 iterations

                                                  Used       Total
  Number of observations                            20          21

  Estimator                                         ML
  Minimum Function Test Statistic                1.993
  Degrees of freedom                                 1
  P-value (Chi-square)                           0.158
```

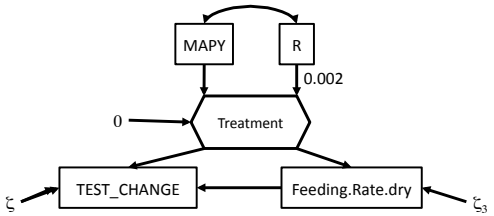
Exercise: Fit this Model

SCALE PROBLEM



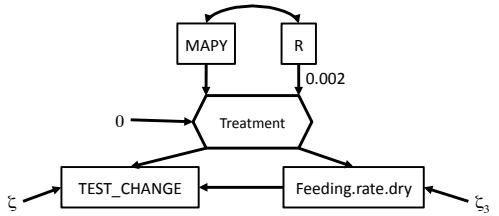

```
> urchinCompositeFit2<-sem(urchinCompositeModel2, data=urchinData)
Error in solve.default(E) :
  system is computationally singular: reciprocal condition number =
   2.09555e-16

[lavaan message:] could not compute standard errors!

 You can still request a summary of the fit to inspect
 the current estimates of the parameters.
```

Exercise: Fit this Model



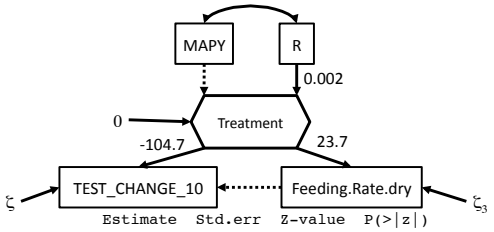

```
> head(urchinData[,c(5,17)])
  Feeding.rate.dry TEST_CHANGE
1      0.006454893        7.68
2      0.011449023        3.74
3      0.012258490        5.78
4      0.007628933        6.34
5      0.011282345        5.00
6      0.007344447        4.94
```

Transform Scales to Fit




```
urchinData$TEST_CHANGE_10<-
        urchinData$TEST_CHANGE/10
```

Fit Model




```
                   Estimate  Std.err  Z-value  P(>|z|)
Regressions:
  Feeding.rate.dry ~
    Treatment         0.888    0.319    2.785    0.005
  TEST_CHANGE_10 ~
    Treatment      -104.771   32.011   -3.273    0.001
    Feedng.rt.dry    23.670   16.784    1.410    0.158
```

A Composite Conclusion

- Composite variables are useful as variables to gather information about multiple aspects of a single effect.
- Excellent for representing nonlinearities.
- Often what ecologists think of in terms of aggregate variables.
- Provide method of incorporating complex treatment effects.