9. Bayesian Estimation With Interaction And Polynomial Effects

Moderation (Interaction)

Moderation occurs when the magnitude of an association depends on a third variable

The influence of a focal predictor on the outcome depends on a moderator variable

For whom does an effect apply?

Moderated Regression Model

Regression model with a product term

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 M_i + \beta_3 X_i M_i + \varepsilon_i = E(Y|X) + \varepsilon_i$$
$$Y_i \sim N\left(E(Y|X, M, XM), \sigma_\varepsilon^2\right)$$

β_1 is the influence of X when M equals zero (i.e., a conditional effect), and β_3 captures the change in the β_1 slope for every one-unit increase in M

Incomplete Predictor Variables

If X or M is missing, so too is the product

Older (biased) methods often attempt to fill-in the missing product variable or impose an incorrect normal distribution assumption on it

Bayesian estimation is a recent innovation, but you have to do it a particular way

Chronic Pain Data

Pain-related data for 275 chronic pain patients

The data include psychological correlates of pain severity such as depression, pain interference with daily life, perceived control over pain, stress, and psychosocial disability

pain.dat

Variable	Name	Missing %	Scaling
Patient identifier	ID	0	Integer index
Gender	MALE	0	0 = female, 1 = male
Age	AGE	0	Continuous
Education level	EDUGROUP	0	1 = Some college or less, 2 = college, 3 = Post-BA
Work hours per week	WORKHRS	11.7	Continuous
Exercise	EXERCISE	1.7	8-point ordinal scale
Pain intensity rating	PAIN	7.3	1 = none/little, 2 = moderate, 3 = severe/very severe
Anxiety	ANXIETY	6.0	Continuous
Stress	STRESS	0	7-point ordinal scale
Perceived control over	CONTROL	0	Continuous
Pain interference with life	INTERFERE	13.3	Continuous
Depression	DEPRESS	13.3	Continuous
Psychosocial disability	DISABILITY	3.0	Continuous

Substantive Example

Does the influence of depression on psychosocial disability differ for males and females?

$$DISABILITY_i = \beta_0 + \beta_1(DEPRESS_i) + \beta_2(MALE_i) \cdot$$
$$+ \beta_3(DEPRESS_i)(MALE_i) + \varepsilon_i$$

Psychosocial disability measures pain's impact on emotional behaviors such as psychological autonomy and communication, emotional stability, etc.

Incomplete Predictors Revisited

An incomplete predictor variable requires a distribution and its own regression model

Bayesian estimation with interactions is not fundamentally different from standard regression

Predictors appear in two models, and the distribution of imputations depends on the substantive analysis and a model for the covariate

Predictor Distribution

Bayes' rule says that the conditional distribution of X is the product of two distributions (models)

$$p(X|Y,X,M) = \frac{p(Y|X,M,XM) \times p(X|M) \times p(M)}{p(Y)}$$

Predictor model

$$\propto p(Y|X,M,XM) \times p(X|M)$$

Substantive model

Regression Models

Substantive model

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 M_i + \beta_3 X_i M_i + \varepsilon_i = E(Y|X) + \varepsilon_i$$

$$Y_i \sim N\left(E(Y|X,M,XM), \sigma^2_\varepsilon\right)$$

$$DISABILITY_i = \beta_0 + \beta_1(DEPRESS_i) + \beta_2(MALE_i) + \beta_3(DEPRESS_i)(MALE_i) + \varepsilon_i$$

Predictor model

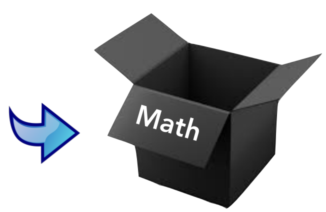
$$X_i = \gamma_0 + \gamma_1 M_i + r_i$$

$$X_i \sim N\left(E(X|M), \sigma^2_r\right)$$

$$DEPRESS_i = \gamma_0 + \gamma_1(MALE_i) + r_i$$

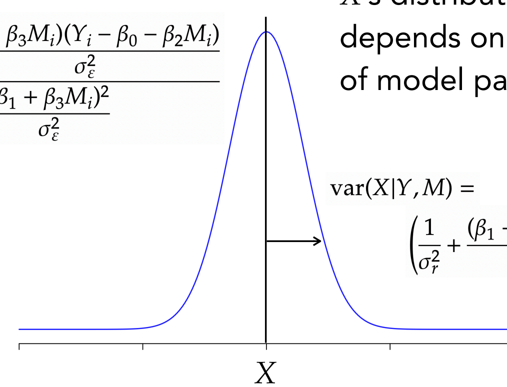
Multiplying Two Normal Distributions …

$$f(X|Y,X,M) \propto \exp\left\{-\frac{1}{2}\frac{\left(Y_i - E(Y|X,M,XM)\right)^2}{\sigma^2_\varepsilon}\right\} \times \exp\left\{-\frac{1}{2}\frac{\left(X_i - E(X|M)\right)^2}{\sigma^2_r}\right\}$$



Gives An Even Uglier Normal Distribution

$$E(X|Y,M) = \frac{\dfrac{\gamma_0 + \gamma_1 X_i}{\sigma^2_r} + \dfrac{(\beta_1 + \beta_3 M_i)(Y_i - \beta_0 - \beta_2 M_i)}{\sigma^2_\varepsilon}}{\dfrac{1}{\sigma^2_r} + \dfrac{(\beta_1 + \beta_3 M_i)^2}{\sigma^2_\varepsilon}}$$



X's distribution again depends on two sets of model parameters

$$\mathrm{var}(X|Y,M) = \left(\frac{1}{\sigma^2_r} + \frac{(\beta_1 + \beta_3 M_i)^2}{\sigma^2_\varepsilon}\right)^{-1}$$

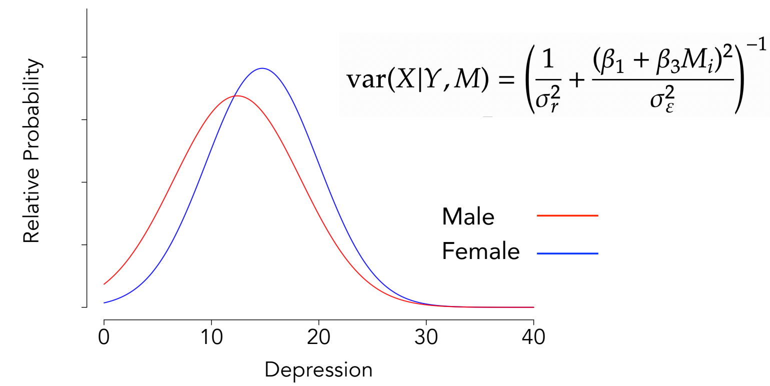
Important Point 1: Heteroscedasticity

The predictor distributions are heteroscedastic, such that the variance of the imputations depends on the moderator and interaction term

Standard imputation routines and maximum likelihood estimation generally fail to capture this nuance and lead to large biases as a result

Distributions Of Missing Values

Male depression imputations have more variation



$$\mathrm{var}(X|Y,M) = \left(\frac{1}{\sigma_r^2} + \frac{(\beta_1 + \beta_3 M_i)^2}{\sigma_\varepsilon^2}\right)^{-1}$$

Important Point 2: Product Term

The product term is never modeled directly

The distribution of X and M reflect lower-order scores that are consistent with the interaction effect in the data (if one exists)

That is, MCMC imputes X and M then computes the product from the filled-in scores

MCMC Recipe — Same As Regression!

Do for t = 1 to T iterations

1. Estimate substantive model parameters (regression coefficients, residual variance), given the filled-in data
2. Impute missing Y values, given the regression model parameters
3. Estimate predictor model regression parameters (regression coefficients, residual variances), given the filled-in data
4. Impute missing predictors, given two sets of model parameters

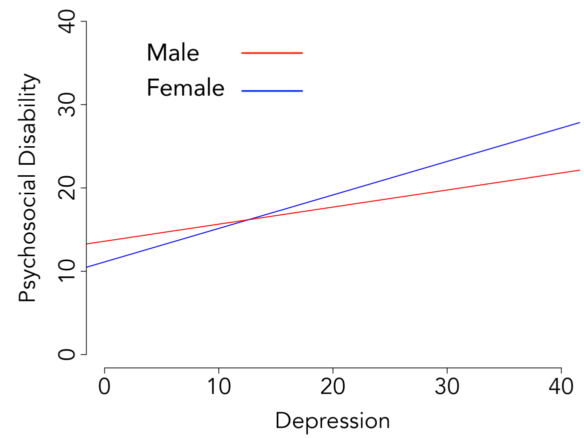
Repeat

Posterior Distribution Summary

Analysis results with depression mean centered

Parameter	Mean	Std. Dev.	Lower 2.5%	Upper 97.5%
Intercept	16.97	0.35	16.28	17.67
DEPRESS slope	0.42	0.06	0.31	0.53
MALE slope	-0.35	0.50	-1.31	0.65
DEPRESS*MALE	-0.21	0.08	-0.37	-0.06

Simple Slopes



Interpretations

For females, the slope of depression on disability is is .418, indicating that an increase in depression increases psychosocial disability

The negative interaction coefficient means that the male slope is lower by .214

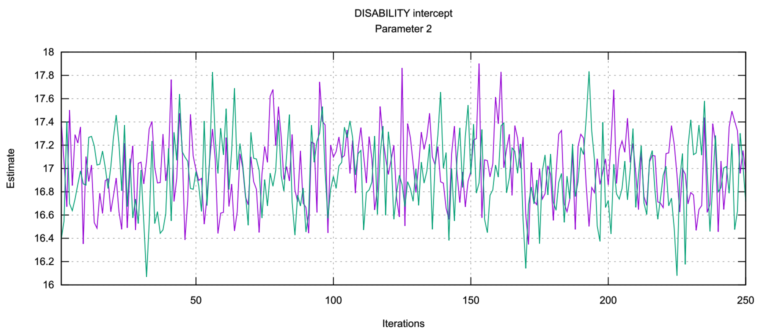
The interaction is “significant” because zero is not in the 95% credible interval

Blimp Analysis Script

```
DATA: pain.dat;
VARIABLES: id male age edugroup workhrs exercise pain anxiety
  stress control interfere depress disability;
NOMINAL: male;
MISSING: 999;
MODEL: disability ~  depress male depress*male;
CENTER: grandmean = depress;
SIMPLE: depress | male;
SEED: 90291;
BURN: 1000;
ITERATIONS: 10000;
CHAINS: 4 processors 4;
OPTIONS: psr;
```

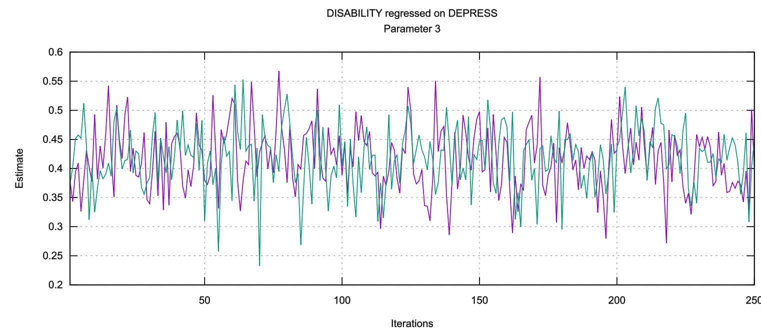
Blimp Trace Plots

Intercept estimates from 250 iterations



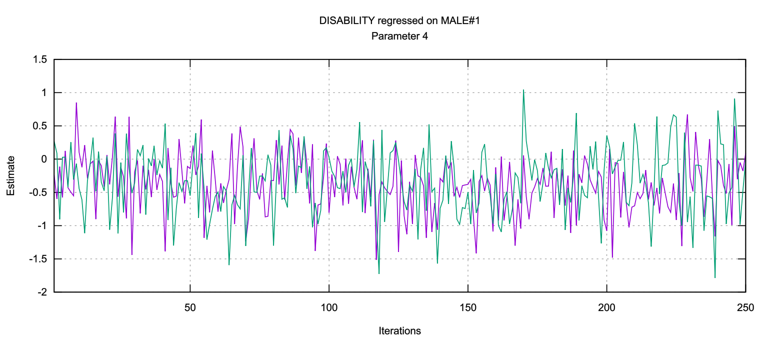
Blimp Trace Plots

Depression slope estimates from 250 iterations



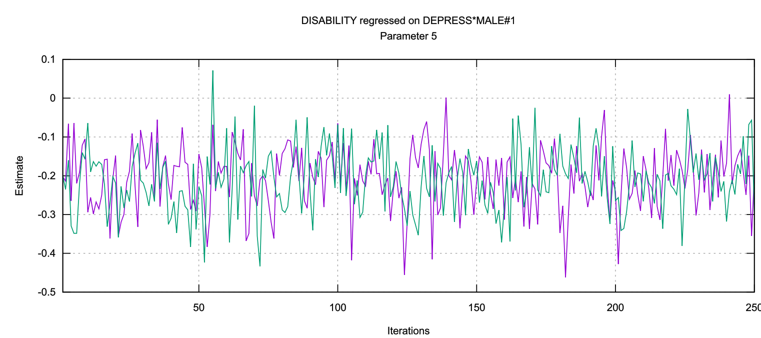
Blimp Trace Plots

Gender slope estimates from 250 iterations



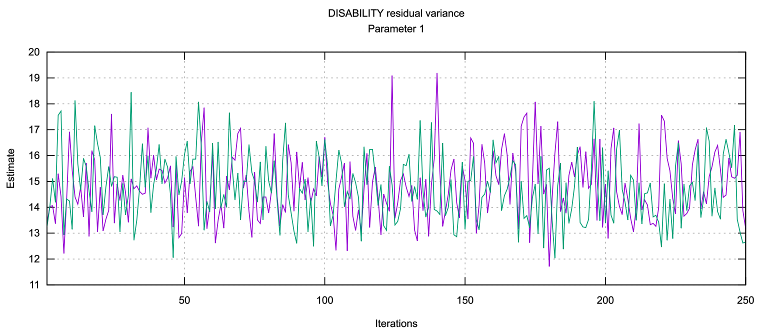
Blimp Trace Plots

Interaction term estimates from 250 iterations



Blimp Trace Plots

Residual variance estimates from 250 iterations



25

Blimp Output

POTENTIAL SCALE REDUCTION (PSR) OUTPUT:

Comparing iterations across 4 chains	Highest PSR	Parameter #
51 to 100	1.025	7
101 to 200	1.098	6
151 to 300	1.020	6
201 to 400	1.028	6
251 to 500	1.011	6
301 to 600	1.016	6
351 to 700	1.011	6
401 to 800	1.009	6
451 to 900	1.012	6
501 to 1000	1.007	6

26

Blimp Output

ANALYSIS MODEL ESTIMATES:

Missing outcome: disability
Grand Mean Centered: depress

Parameters	Mean	Median	StdDev	Lower 2.5	Upper 97.5
Variances:					
Residual Var.	14.774	14.685	1.333	12.409	17.616
Coefficients:					
Intercept	16.970	16.970	0.354	16.278	17.674
depress	0.418	0.417	0.055	0.310	0.528
male#1	-0.350	-0.349	0.497	-1.310	0.647
depress*male#1	-0.214	-0.214	0.081	-0.373	-0.055
Standardized Coefficients:					
depress	0.589	0.591	0.068	0.449	0.715
male#1	-0.039	-0.039	0.055	-0.145	0.072
depress*male#1	-0.203	-0.203	0.076	-0.351	-0.053
Proportion Variance Explained					
by Fixed Effects	0.234	0.233	0.046	0.145	0.323
by Residual Variation	0.766	0.767	0.046	0.677	0.855

Summaries based on 10000 iterations using 4 chains

27

Blimp Output

CONDITIONAL EFFECTS ANALYSIS:

Missing outcome: disability

Grand Mean Centered: depress

Conditional Effects	Mean	Median	StdDev	Lower 2.5	Upper 97.5
depress | male#1 @ 0					
Intercept	16.970	16.970	0.354	16.278	17.674
Slope	0.418	0.417	0.055	0.310	0.528
depress | male#1 @ 1					
Intercept	16.620	16.623	0.382	15.874	17.376
Slope	0.204	0.204	0.060	0.086	0.321

Summaries based on 10000 iterations using 4 chains

NOTE: Intercepts are computed by setting all predictors not involved in the conditional effect to zero.

28

Blimp Output (Predictor Distributions)

```
Covariate Models
------------------
Missing covariate:  depress

                              -------------------------------------------------------------
Parameters                    |   Mean    |  Median   |  StdDev   |Lower 2.5  |Upper 97.5|
                              -------------------------------------------------------------
Grand Mean                    |    14.702|    14.697|     0.402|    13.913|    15.486|
Level 1:                      |          |          |          |          |          |
  male#1                      |     0.668|     0.670|     0.482|    -0.279|     1.606|
  Residual Var.               |    38.028|    37.815|     3.581|    31.621|    45.683|
                              |          |          |          |          |          |
                              -------------------------------------------------------------
Complete covariate: male
                              -------------------------------------------------------------
Parameters                    |   Mean    |  Median   |  StdDev   |Lower 2.5  |Upper 97.5|
                              -------------------------------------------------------------
Grand Mean#1                  |    -0.263|    -0.264|     0.077|    -0.415|    -0.111|
Level 1:                      |          |          |          |          |          |
  #1 ~ depress                |     0.018|     0.018|     0.013|    -0.008|     0.043|
  Var(#1)                     |     1.000|     1.000|     0.000|     1.000|     1.000|
                              |          |          |          |          |          |
                              -------------------------------------------------------------
```


Curvilinear Regression Model

A regression model with a squared term “bends” the regression line

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{2i}^2 + \varepsilon_i = E\left(Y \middle| X_1, X_2, X_2^2\right) + \varepsilon_i$$

$$Y_i \sim N\left(E\left(Y \middle| X_1, X_2, X_2^2\right), \sigma_\varepsilon^2\right)$$

Y is normally distributed around predicted values

Math Achievement Data

Math achievement data for 250 students

The data set includes pre-test and post-test math achievement scores and academic-related variables such as math self-efficacy, math anxiety, standardized reading scores, socio-demographic variables

math.dat

Variable	Name	Missing %	Scaling
Identifier variable	ID	0	Integer index
Gender	MALE	0	0 = female, 1 = male
Free or reduced lunch	LUNCHASST	4.3	0 = none, 1 = assistance
Achievement group	ACHIEVEGRP	2.0	1 = typically achieving, 2 = low achieving, 3 = learning disability
Standardized reading	STANREAD	10.0	Continuous
Math self-efficacy	EFFICACY	9.7	6-point ordinal scale
Math anxiety	ANXIETY	9.3	Continuous
Pre-test math achievement	MATHPRE	0	Continuous
Post-test math achievement	MATHPOST	18.0	Continuous

Substantive Example

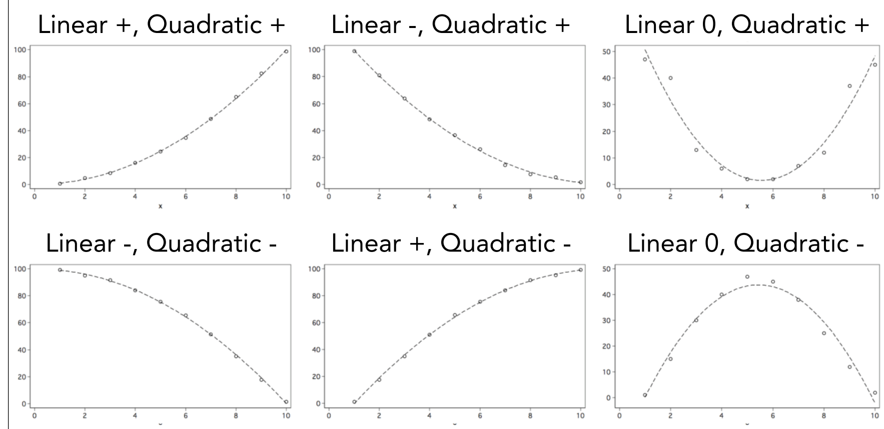
Anxiety's influence on math achievement depends on a student's anxiety level

$$MATHPOST_i = \beta_0 + \beta_1(MALE_i) + \beta_2(ANXIETY_i) + \beta_3\left(ANXIETY_i^2\right) + \varepsilon_i$$

e.g., The influence of anxiety on achievement could become more negative as anxiety increases

33

Linear And Quadratic Combinations



34

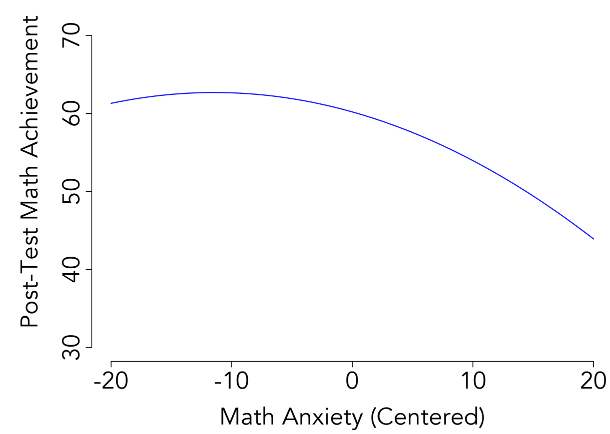
Posterior Distribution Summary

Analysis results with 10,000 MCMC iterations

Parameter	Mean	Std. Dev.	Lower 2.5%	Upper 97.5%
Intercept	57.781	0.700	56.394	59.139
MALE slope	-4.906	1.141	-7.167	-2.687
ANXIETY slope	-0.435	0.088	-0.607	-0.262
ANXIETY² slope	-0.019	0.007	-0.033	-0.006

35

Regression Line



36

Interpretations

At the anxiety mean, the slope (instantaneous rate of change) is −.435, which means that increases in anxiety reduces math achievement

Negative curvature means that the anxiety slope becomes more negative as anxiety increases

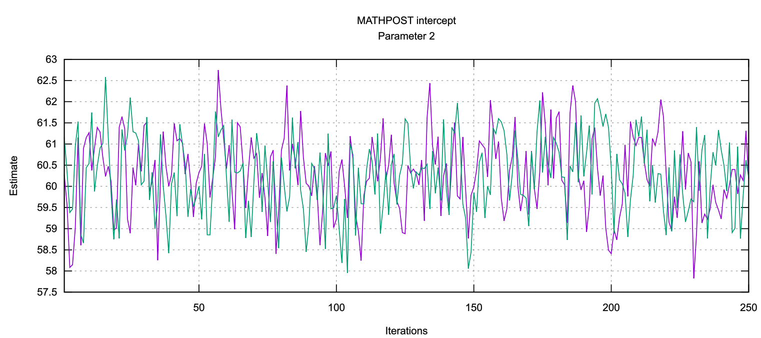
The squared term (curvature) is "significant" because zero is not in the 95% credible interval

Blimp Analysis Script

```
DATA: math.dat;
VARIABLES: id male lunchasst achievegrp stanread efficacy
  anxiety mathpre mathpost;
FIXED: male;
MISSING: 999;
MODEL: mathpost ~  male anxiety anxiety*anxiety ;
CENTER: grandmean = male anxiety;
SEED: 90291;
BURN: 1000;
ITERATIONS: 10000;
CHAINS: 4 processors 4;
OPTIONS: psr;
```

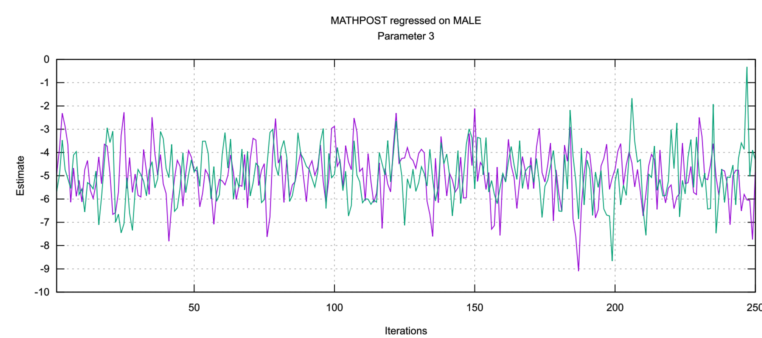

Blimp Trace Plots

Intercept estimates from 250 iterations



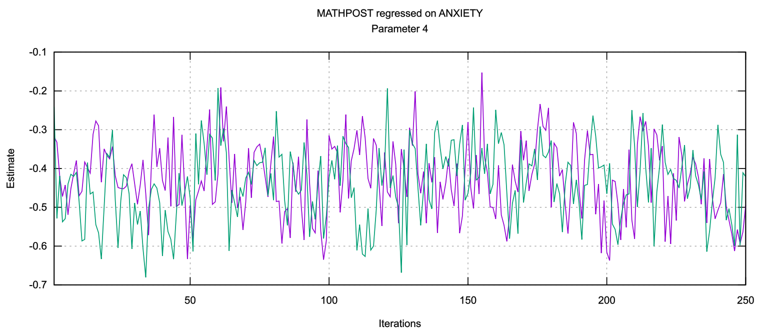
Blimp Trace Plots

Gender slope estimates from 250 iterations



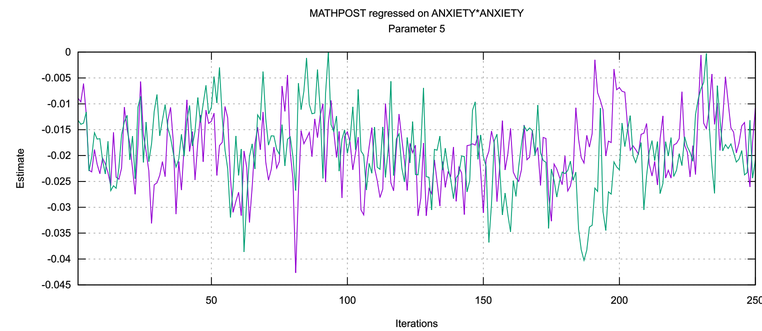
Blimp Trace Plots

Anxiety slope estimates from 250 iterations



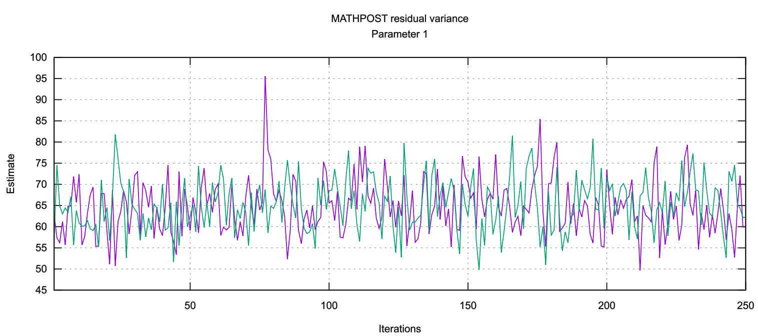
Blimp Trace Plots

Squared term estimates from 250 iterations



Blimp Trace Plots

Residual variance estimates from 250 iterations



Blimp Output

```
POTENTIAL SCALE REDUCTION (PSR) OUTPUT:

  Comparing iterations across 10 chains   | Highest PSR | Parameter # |
                            51 to 100     |       1.071 |           6 |
                           101 to 200     |       1.042 |           5 |
                           151 to 300     |       1.031 |           6 |
                           201 to 400     |       1.008 |           6 |
                           251 to 500     |       1.006 |           4 |
                           301 to 600     |       1.007 |           6 |
                           351 to 700     |       1.006 |           6 |
                           401 to 800     |       1.009 |           6 |
                           451 to 900     |       1.010 |           6 |
                           501 to 1000    |       1.009 |           6 |
```

Blimp Output

ANALYSIS MODEL ESTIMATES:

Missing outcome: mathpost
Grand Mean Centered: anxiety male

Parameters	Mean	Median	StdDev	Lower 2.5	Upper 97.5
Variances:					
Residual Var.	64.887	64.441	6.565	53.030	79.014
Coefficients:					
Intercept	57.781	57.785	0.700	56.394	59.139
male	-4.906	-4.908	1.141	-7.167	-2.687
anxiety	-0.435	-0.436	0.088	-0.607	-0.262
anxiety*anxiety	-0.019	-0.019	0.007	-0.033	-0.006
Standardized Coefficients:					
male	-0.255	-0.256	0.056	-0.360	-0.142
anxiety	-0.348	-0.349	0.068	-0.479	-0.210
anxiety*anxiety	-0.254	-0.258	0.085	-0.411	-0.081
Proportion Variance Explained					
by Fixed Effects	0.302	0.301	0.058	0.189	0.413
by Residual Variation	0.698	0.699	0.058	0.587	0.811

Summaries based on 10000 iterations using 10 chains

Blimp Output (Predictor Distributions)

Covariate Models

Missing covariate: anxiety

Parameters	Mean	Median	StdDev	Lower 2.5	Upper 97.5
Grand Mean	18.281	18.288	0.498	17.312	19.283
Level 1:					
male	-2.444	-2.444	0.999	-4.410	-0.502
Residual Var.	58.246	57.961	5.397	48.536	69.854