Alternatives to ANOVA and Regression Amidst Non-normality: Relative Hypothesis Test Performance

Anthony J. Bishara[1], Bo Kai[2], & James B. Hittner[1]
[1]Dept. of Psychology, [2]Dept. of Mathematics

COLLEGE of CHARLESTON

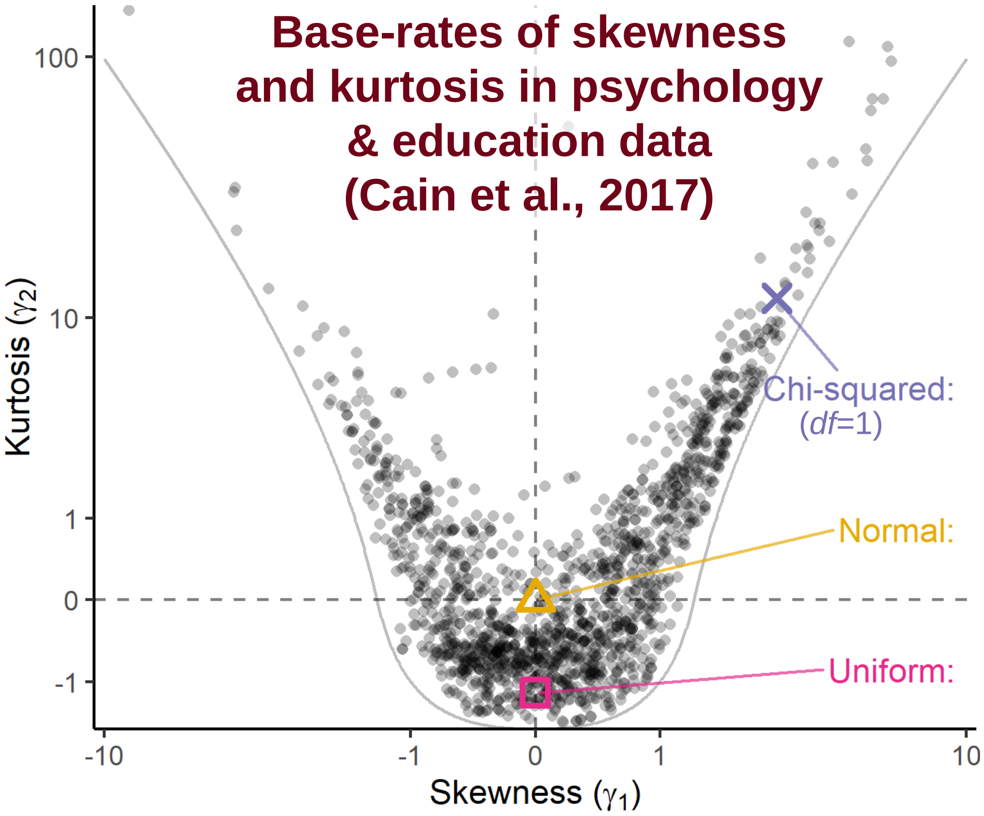
Background

- OLS Regression and ANOVA are perceived as "robust" to non-normal residuals
  - However, non-normality can lead to an opportunity cost whereby alternative models are more powerful
- Literature unclear on relative power of alternatives

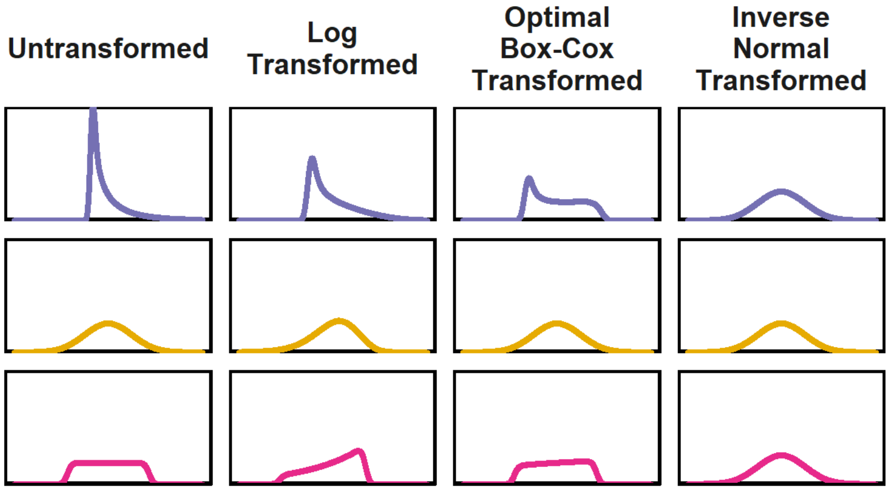
- We considered nonparametric, robust, and transformation alternatives
- Emphasis on Inverse Normal Transformations (INTs)
  - popular in genome-wide association studies
  - $INT(y_i) = \phi^{-1}\left(\frac{rank(y_i) - .5}{n}\right)$

  where Φ^{-1} = inverse normal CDF

Base-rates of skewness and kurtosis in psychology & education data (Cain et al., 2017)



Examples of Transformations



COLLEGE of CHARLESTON

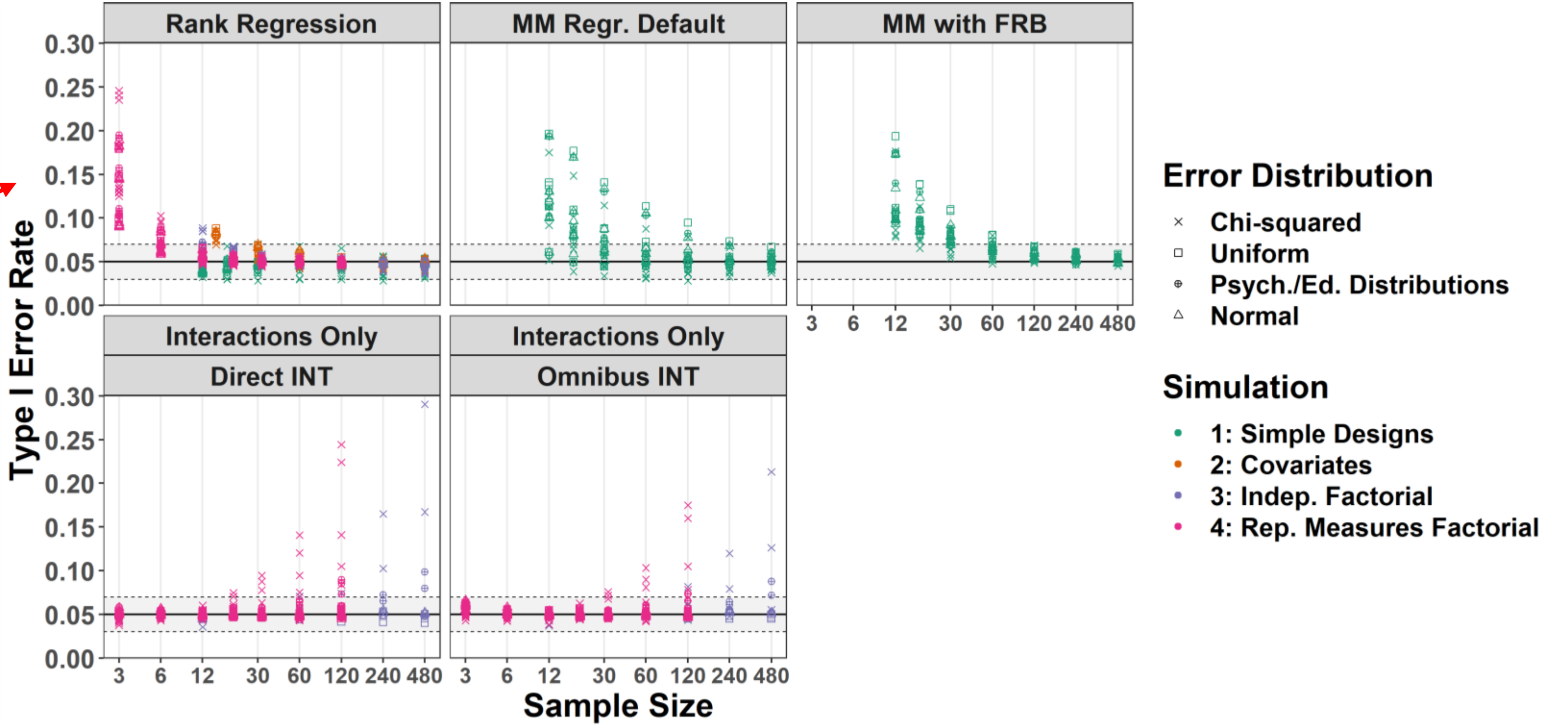
Methods

- Monte Carlo studies
  - We manipulated residual distributions, *n*, effect sizes, covariate sizes, predictor correlations, etc.
  - Examined 2,052 scenarios

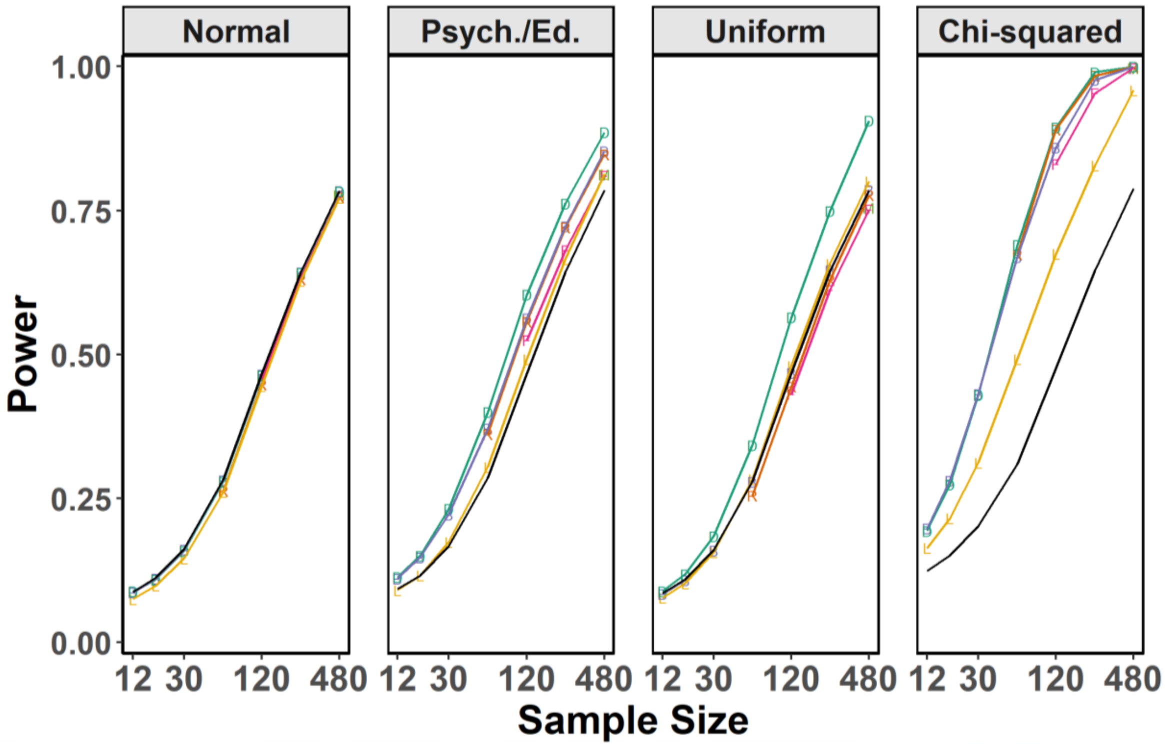
Compared Type I Error & Power:

- Parametric OLS Regression/ANOVA
- Rank-based Regression
- Robust MM Regression
  - with or without the Fast Robust Bootstrap (FRB)
- Transformation with OLS Regression
  - ln(*y*)
  - Box-Cox(*y*)
  - Direct INT(*y*)
  - Indirect INT(residuals)
  - Omnibus INT (see McCaw et al., 2020)
    - Cauchy aggregation of Direct & Indirect p-values
  - Conditional INT(*y*)
    - Transform only if significant normality test
  - Aligned Rank: rank(residuals)

- Type I error was maintained at α=.05 for most methods
- These were the exceptions



Power by residual distribution (Simulation 1)



Method	Average power in non-normal scenarios, n=480
Direct INT	.93
Rank Regression or Kruskal-Wallis	.88
Box-Cox Transform	.88
MM with FRB	.85
MM Regr. Default	.86
Log Transform	.86
OLS Regr. or ANOVA	.79

COLLEGE of CHARLESTON

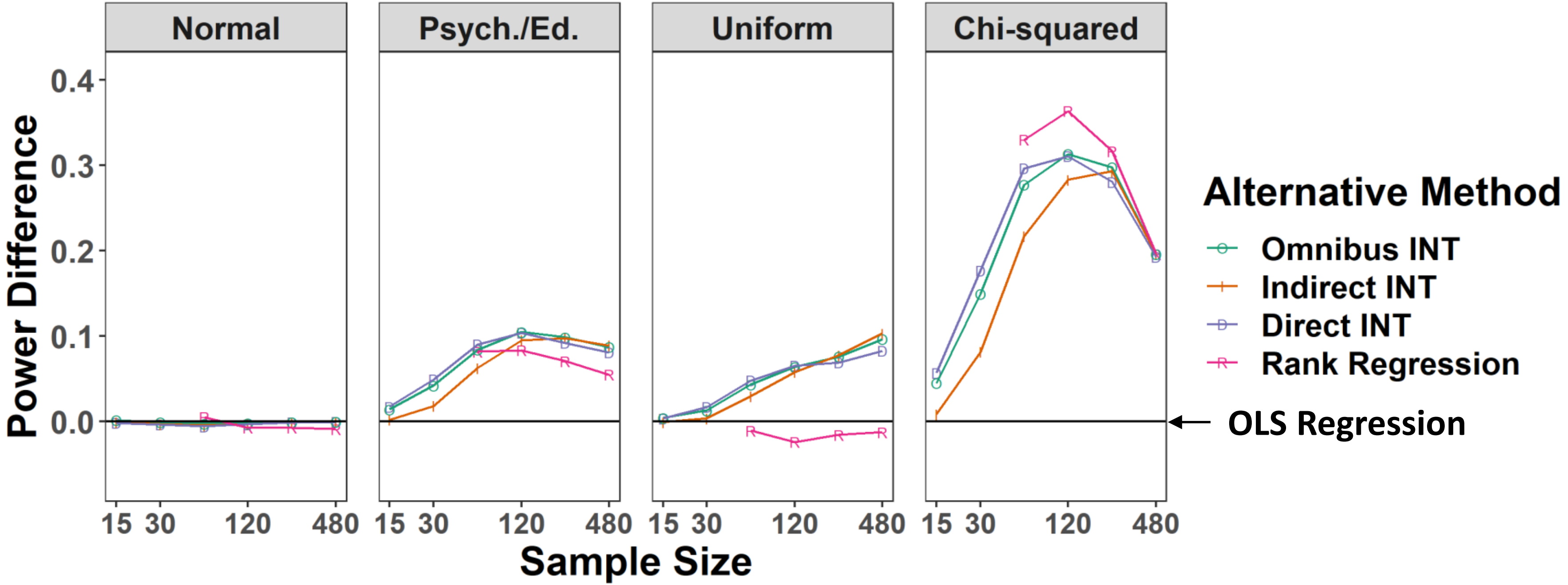
Indirect INT steps

Consider a model with <u>*c* covariates</u> and <u>*p* predictors of interest</u>

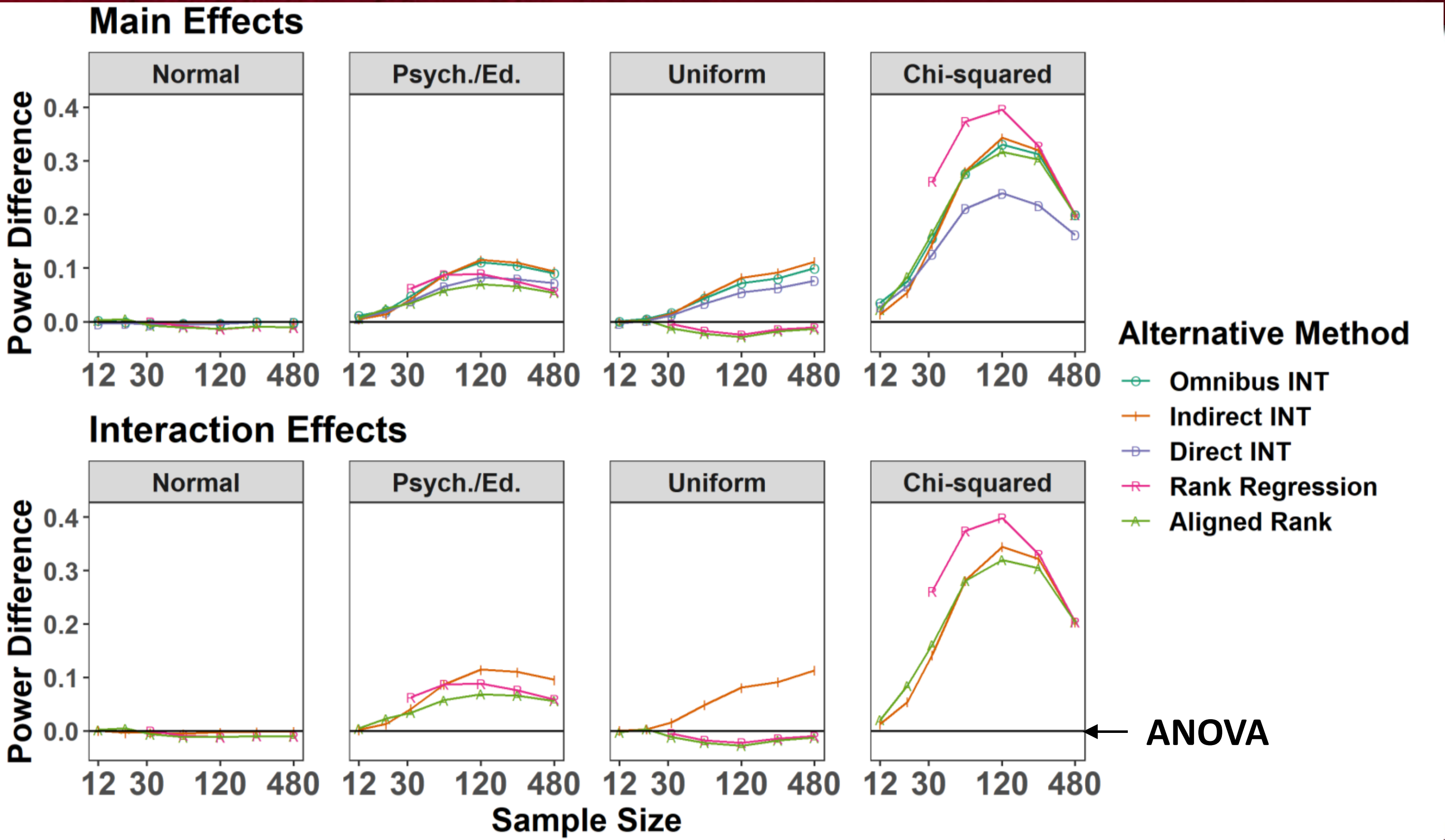
1) Estimate residuals in a restricted model using only the *c* covariates	$\hat{y}_i = \widehat{b_0} + \widehat{b_1} x_{1i} + \cdots + \widehat{b_c} x_{ci} + \hat{\epsilon}_{1i}$
2) Repeat with *y* replaced by the transformed residuals from step 1	$INT(\hat{\epsilon}_{1i}) = \widehat{b_0} + \widehat{b_1} x_{1i} + \cdots + \widehat{b_c} x_{ci} + \hat{\epsilon}_{2i}$
3) Estimate the full model with transformed residuals from step 1	$INT(\hat{\epsilon}_{1i}) =$ $\widehat{b}_0 + \widehat{b}_1 x_{1i} + \cdots + \widehat{b}_c x_{ci} +$ $\widehat{b}_{c+1} x_{(c+1)i} + \ldots + \widehat{b}_{(c+p)} x_{(c+p)i} + \hat{\epsilon}_{3i}$
4) Compare models 2 & 3 with a nested *F*-test	$F_{p,\, n-c-p} = \frac{\{\sum_{i=1}^{n} (\hat{\epsilon}_{2i})^2 - \sum_{i=1}^{n} (\hat{\epsilon}_{3i})^2\}/p}{\{\sum_{i=1}^{n} (\hat{\epsilon}_{3i})^2\}/(n-c-p)}$

Simulation 2: Designs with Covariates

Power Relative to OLS Regression



Simulation 3: Factorial ANOVA



Conclusions

Type I Error rates were inflated with:

- Rank-Based Regression for $n \leq 30$
- MM Estimation $n \leq 240$
  - with Fast Robust Bootstrap $n \leq 60$
- Direct & Omnibus INT Interactions

Power:

- INTs usually matched or exceeded other methods' power
- Relative power of Direct vs. Indirect varied, but Omnibus INT approached whichever was better
- Conditioning INTs on normality tests provided no benefit (not shown here)

Recommendation: ***When the residual distribution is in doubt, use the Omnibus INT for main effects and Indirect INT for interactions.***