Measurement Error- and Heteroskedasticity-Robust Analysis of Covariance

UNIVERSITY OF CALGARY

Andrew F. Hayes
University of Calgary

Jacob J. Coutts
University of Maryland

Using PROCESS, your test of group mean differences can be invulnerable to both random measurement error (RME) in covariates and heteroskedasticity

ANCOVA is *not* "robust"

Analysis of covariance (and its ordinary least squares (OLS) regression equivalent) is routinely used as a means of comparing groups that may (or may not) differ on one or more variables ("covariates") also related to the dependent variable Y.

Assumption	Robust?
Normality of within-group errors in estimation	YES
Homoskedasticity (equality of within-group variance)	NO
Covariate(s) free of random measurement error (**RME**)	NO

- Homoskedasticity should not be assumed, as heteroskedasticity is a common and more realistic situation. In it's presence, false rejections (claiming there are differences when there aren't) are more likely than you think. Fortunately, there are inferential techniques (e.g., **HC0** and **HC3** standard error estimators) that work well with heteroskedasticity, though are not widely known or used.
- Covariates are almost never free of RME (i.e., reliability less than 1). This can invalidate standard ANCOVA/regression-based tests of group differences. Structural equation modeling can be used to deal with RME. **Errors-in-variables (EIV) regression** is easier though not widely known in behavioural science or implemented (until now).
- **EIV regression?** Originally introduced in econometrics, it involves a simple modification of the data prior to OLS estimation, subtracting out random measurement error expected given an estimate of reliability of measurement. Standard errors required for inference are modified accordingly.

We have developed a new EIV-based standard error estimator for ANCOVA/regression that includes a heteroskedasticity robustification better than alternatives. It works and is now available in the PROCESS macro for SPSS, SAS, and R.

Simulation and Results

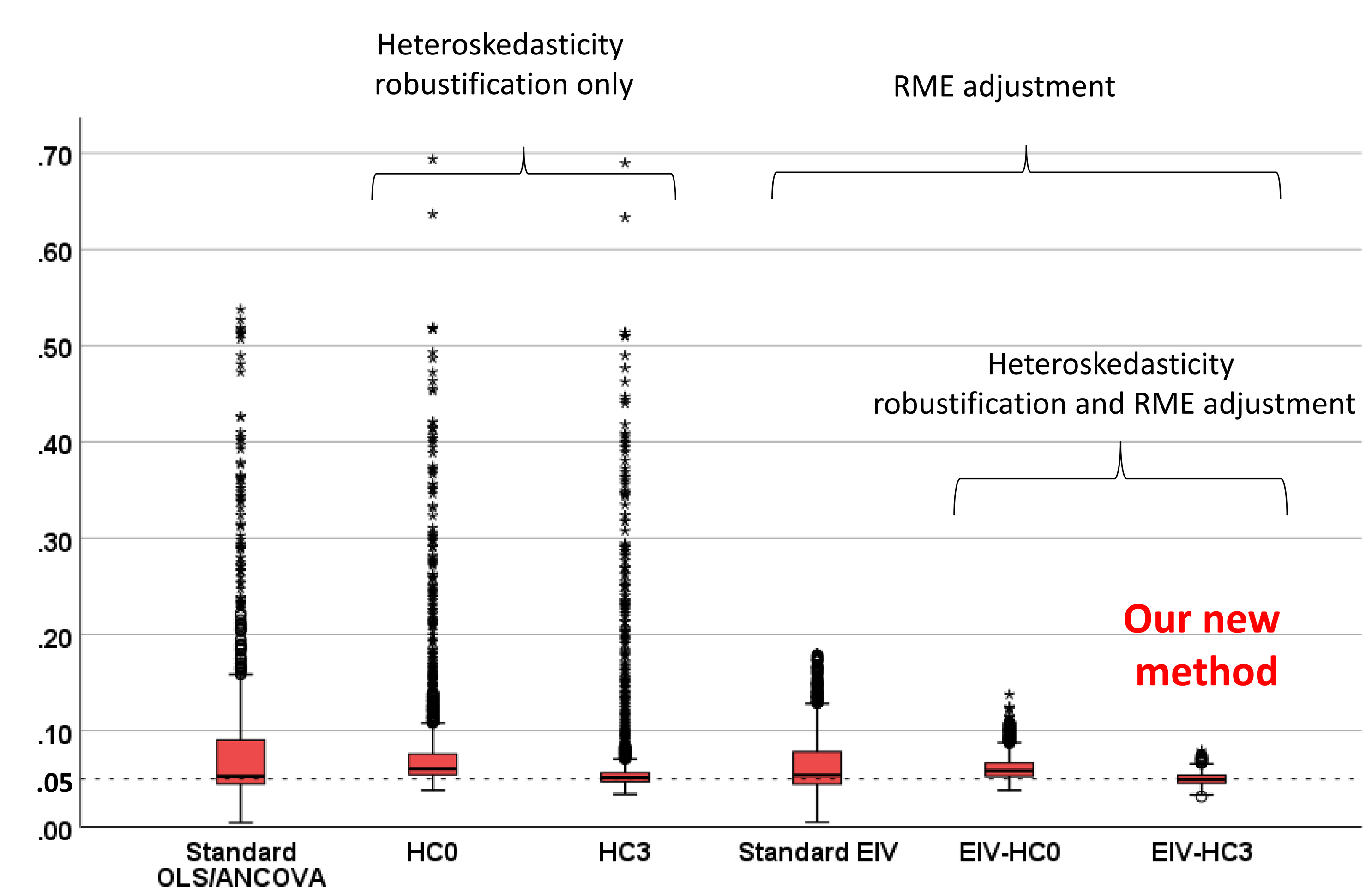
We examined Type I error rates for a test of equality of group means, comparing standard ANCOVA/regression to EIV regression with and without a heteroskedasticity-robustification, including a new one we introduce (**EIV-HC3**)

In our simulation, we orthogonally manipulated (1728 conditions in all):

- Total sample size (60, 120, 240, 480)
- Number of groups (2, 3, 4)
- Inequality in group sample sizes (none, moderate, large)
- Heteroscedasticity in outcome *Y* (none, larger groups with 4× larger variance in *Y*, smaller groups with 4× larger variance in *Y*)
- Reliability of covariate (0.7, 1)
- Reliability of outcome (0.7, 1)
- Correlation between groups and covariate (0, 0.5)
- Correlation between covariate and outcome (0, 0.5)

Figure 1: False rejection (Type I error) rate using α = 0.05 criterion

2000 replications. Boxplots visualize performance across the 1728 conditions in the simulation.



Summary and Implementation

Use the new **EIV-HC3** *implementation in PROCESS and don't worry about heteroskedasticity or random measurement error in your ANCOVA/regression. As shown in* Figure 1 *(left bottom):*

- In many situations you are likely to encounter, you can't trust a test from standard ANCOVA/OLS regression
- Heteroskedasticity-robust methods (HC0 and HC3) reduce the vulnerability to false rejections but don't eliminate it when covariates are measured imperfectly and this isn't accounted for in the analysis.
- Accounting only for random measurement error using standard EIV regression (EIV only) helps too, but doesn't eliminate problems due to heteroskedasticity.
- Accounting for both RME and heteroskedasticity (EIV-HC0 or EIV-HC3) is the ticket, with our improvement (**EIV-HC3**) more uniformly keeping false rejections in control regardless of the presence of heteroskedasticity and/or random measurement error in a covariate.

Implementation in the PROCESS macro

As of version 5, PROCESS for SPSS, SAS, and R can do EIV-based ANCOVA/regression with heteroskedasticity robustification and while accounting for random measurement error in covariates. Only one line of code and a reliability estimate for your covariate(s) is needed.

Example PROCESS command comparing groups ("cond") on "perform" controlling for "support" measured with reliability 0.8

SPSS
```
process y=perform/x=cond/mcx=1/cov=support/relcov=0.8.
```

SAS:
```
%process(data=test,y=perform,x=cond,mcx=1,cov=support,relcov=0.8)
```

R
```
process(test,y="perform",x="cond",mcx=1,cov="support",relcov=0.8)
```

Presented at the annual convention of the Canadian Psychological Association, June 2025, in St. John's Newfoundland