Evaluating generic versus specific corrections for heteroskedasticity

Robert L. Kaufman

Department of Sociology, Ohio State University

Research questions

How well does the common "generic" heteroskedasticity correction known as HC3 (Long and Ervin 2000) perform in comparison to a correction for the actual specific form of heteroskedasticity?

I.e., how well do actual Type I errors in hypothesis tests correspond to the nominal .05 level and what is the power of the hypothesis test?

Past simulations have compared the use of *HC3* to other generic corrections or to the use of *OLS* without any correction. This research compares it to the use of *EGLS* (Estimated Generalized Least Squares) which corrects for an exact, specified form of heteroskedasticity.

How do the performances of OLS, HC3 and EGLS vary with

- **Sample size**
- **The assumed error distribution**
- **Common forms of heteroskedasticity**
- **The degree of heteroskedasticity**
- **The use of a screening test to decide whether to use a heteroskedasticity correction and/or which form to specify**

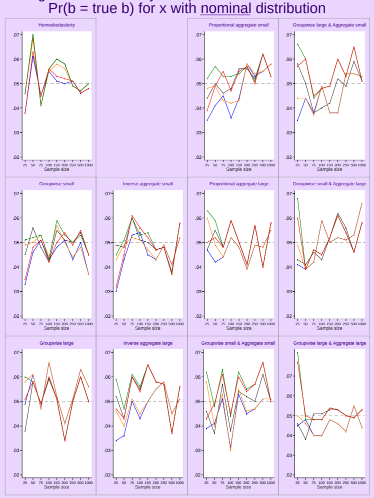
Monte carlo simulation design

Created population of 100,000 observations

- **1000 replications of 176 combinations of situations**
  - ***Sample size (8) by Error distribution (2) by Form and degree of heteroskedasticity (11)***
- **Varying sized random samples without replacement**
  - ***N = 25, 50 ,75, 100, 200, 250, 500, 1000***
- **Functional form**
  - $y_i = 1 + 1{*}x_{1i} + 1{*}x_{Di} + 1{*}x_{3i} + 1{*}x_{4i} + \tau{*}\varepsilon_i$
  - **τ used to set $R^2 = .4$**
  - ***Used varied distributions for the predictors (x)***
    - **x_1 uniform distribution**
    - **x_D nominal (dummy variable)**
    - **x_3 bell-shaped distribution**
    - **x_4 skewed distribution**
  - ***Two different error distributions***
    - **Normal: $\varepsilon \sim N(0,1)$**
    - **Skewed: $\varepsilon \sim \chi2(5)$**
- **Five common forms of heteroskedasticity with varying degree**
  1. ***None:* Homoskedastic**
  2. ***Groupwise:* varies across two groups corresponding to a nominal predictor in the analysis**
  3. ***Inverse aggregate*: varies inversely with the size of aggregate units of analysis**
  4. ***Proportional aggregate*: varies proportional to the size of aggregate units of analysis**
  5. ***Combination form*: Group-wise and aggregate inverse heteroskedasticity simultaneously**
- **Evaluation of Performance of Estimation Methods**
  - ***Size tests***
    - **% of times rejected $H_0 : b_k$ = true value (= 1)**
  - ***Power tests***
    - **% of times rejected H0 : $b_k = 0$ (Not presented)**
- **Diagnostic screening test used**
  - ***Breusch -Pagan (1979) test for a specified from of heteroskedasticity (not for all predictors)***

Significance level by sample size for normal errors Pr(b = true b) for x with uniform distribution

Significance level by sample size for normal errors Pr(b = true b) for x with nominal distribution



Key

- True model
- HC3 always
- HC3 but OLS if B-P not significant
- EGLS but OLS if B-P not significant
- EGLS but HC3 if B-P not significant

Definition of True Models

Homoskedasticity

$$\sigma_i^{\,2} = \text{constant}$$

Groupwise small

$$\sigma_{1i}^2 = 1.44\sigma_{2i}^2$$

Groupwise large

$$\sigma_{1i}^2 = 9.00\sigma_{2i}^2$$

Inverse aggregate small

$$\sigma_i^2 = \frac{\sigma^2}{n_i} \qquad n_i \sim \gamma(1000,2,1.41)$$

Inverse aggregate large

$$\sigma_i^2 = \frac{\sigma^2}{n_i} \qquad n_i \sim \gamma(1000,1,2)$$

Proportional aggregate small

$$\sigma_i^2 = \sigma^2 * n_i \qquad n_i \sim \gamma(1000,2,1.41)$$

Proportional aggregate large

$$\sigma_i^2 = \sigma^2 * n_i \qquad n_i \sim \gamma(1000,1,2)$$

Groupwise small & Inverse aggregate small

$$\sigma_{1i}^2 = \frac{\sigma_1^{\,2}}{n_i}, \sigma_{2i}^2 = \frac{\sigma_2^{\,2}}{n_i} \qquad \frac{\sigma_1^2}{\sigma_2^2} = 1.44 \quad n_i \sim \gamma(1000,2,1.41)$$

Groupwise large & Inverse aggregate small

$$\sigma_{1i}^2 = \frac{\sigma_1^{\,2}}{n_i}, \sigma_{2i}^2 = \frac{\sigma_2^{\,2}}{n_i} \qquad \frac{\sigma_1^2}{\sigma_2^2} = 9.00 \quad n_i \sim \gamma(1000,2,1.41)$$

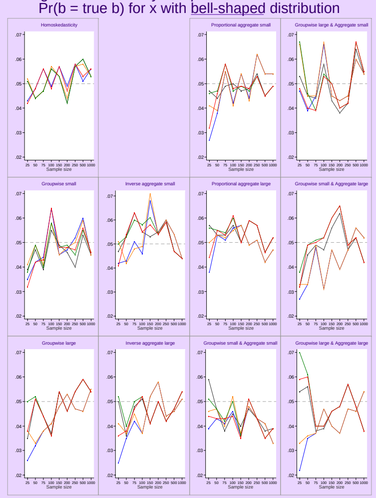
Groupwise small & Inverse aggregate large

$$\sigma_{1i}^2 = \frac{\sigma_1^{\,2}}{n_i}, \sigma_{2i}^2 = \frac{\sigma_2^{\,2}}{n_i} \qquad \frac{\sigma_1^2}{\sigma_2^2} = 1.44 \quad n_i \sim \gamma(1000,1,2)$$

Groupwise large & Inverse aggregate large

$$\sigma_{1i}^2 = \frac{\sigma_1^{\,2}}{n_i}, \sigma_{2i}^2 = \frac{\sigma_2^{\,2}}{n_i} \qquad \frac{\sigma_1^2}{\sigma_2^2} = 9.00 \quad n_i \sim \gamma(1000,1,2)$$

Significance level by sample size for normal errors Pr(b = true b) for x with bell-shaped distribution



Significance level by sample size for normal errors Pr(b = true b) for x with skewed distribution

HC3 vs. OLS vs. EGLS

- **When there is no heteroskedasticity, using either HC3 or OLS with a screening test typically worked well, while the use of EGLS with a screening test is almost as good.**

  Except HC3-based corrections fare poorly for the combination of a skewed predictor and skewed ($\chi2$) errors.

- **HC3 performs variably to consistently poorly when the heteroskedasticity is either Groupwise or the combination of Groupwise and Inverse Aggregate.**

  Note that both these models specify that heteroskedasticity is a function of a predictor in the model.

- **HC3 generally performs well (but not usually best) when the heteroskedasticity is either the inverse aggregate form or the proportional aggregate form but is often poor at smaller sample sizes; how small depends on the distribution of the predictor.**

  Note that neither model specifies that heteroskedasticity is a function of a predictor in the model.

- **EGLS for the specific form of heteroskedasticity performs consistently well across the range of forms of heteroskedasticity.**

  If the screening test for heteroskedasticity was not significant, using OLS vs. HC3 most often made little difference.

  But HC3 sometimes worked poorly for smaller sample sizes (how small depended on the distribution of the predictor).

Other findings

- **When the degree of heteroskedasticity is SMALLER, there is somewhat less distinction among the performances of the various estimation methods.**

  GREATER heteroskedasticity favors the use of EGLS, especially for forms including a Groupwise specification.

- **HC3 appears to be the most adversely affected by the distributional assumptions about the predictors.**

  Especially for a skewed distribution.

- **Note that in several models it is sensible to test for both Groupwise and Aggregate size forms of heteroskedasticity.**

  Use of a screening test for multiple forms consistently worked nearly as well as a screening test for only the true form.

  So only results for multiple screening are presented.

- **Except in a few limited instances (e.g., skewed predictor when the true model is OLS), the type of error distribution mattered little.**

  So only results for Normal errors are presented.

For further information

Please contact me at *kaufman.3@sbs.osu.edu*. You can find a PDF version of this poster on my website at http://www.sociology.ohio-state.edu/rlk/ASA_200_poster.pdf