Testing Many Restrictions Under Heteroskedasticity

Stanislav Anatolyev[1] and Mikkel Sølvsten[2]

[1]CERGE-EI (Czech Republic)

[2]University of Wisconsin-Madison (USA)

DAGStat 2022
Hamburg, Germany
March 2022

Overview

Regression tests (F, Wald) do not work properly when number of regressors and restrictions is large (e.g., Berndt and Savin, 1977)

Overview

Regression tests (F, Wald) do not work properly when number of regressors and restrictions is large (e.g., Berndt and Savin, 1977)

- Many regressors/controls/covariates at RHS
- Many restrictions to test (joint significance of controls, panel group effects)

Overview

Regression tests (F, Wald) do not work properly when number of regressors and restrictions is large (e.g., Berndt and Savin, 1977)

- Many regressors/controls/covariates at RHS
- Many restrictions to test (joint significance of controls, panel group effects)

Recent literature on inference with many regressors:

Overview

Regression tests (F, Wald) do not work properly when number of regressors and restrictions is large (e.g., Berndt and Savin, 1977)

– Many regressors/controls/covariates at RHS
– Many restrictions to test (joint significance of controls, panel group effects)

Recent literature on inference with many regressors:

♣ Corrected classical tests developed for many regressors and many restrictions (e.g., Calhoun, 2011; Anatolyev, 2012) fail under heteroskedasticity

♣ Corrected classical tests (t, Wald) developed for many regressors and heteroskedasticity (e.g., Cattaneo, Jansson, and Newey, 2018) fail with many restrictions

Feature presentation

♡ We propose new corrected classical test (based on F) that works with both many restrictions and heteroskedasticity

Feature presentation

♡ We propose new corrected classical test (based on F) that works with both many restrictions and heteroskedasticity

♡ Corrected F test: new critical value for good old F-statistic

Feature presentation

♡ We propose new corrected classical test (based on F) that works with both many restrictions and heteroskedasticity

♡ Corrected F test: new critical value for good old F-statistic

♡ We use new leave-two-out and leave-three-out technology as extension of leave-one-out (a.k.a. jackknife)

Framework

Linear regression:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad \mathbb{E}[\varepsilon_i \,|\, \boldsymbol{x}_i] = 0$$

Framework

Linear regression:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad \qquad \mathbb{E}[\varepsilon_i \,|\, \boldsymbol{x}_i] = 0$$

- n observations are independent across i
- full rank m of design matrix $\boldsymbol{S}_{xx} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i' = X'X$

Framework

Linear regression:

$$y_i = \boldsymbol{x}_i' \boldsymbol{\beta} + \varepsilon_i \qquad\qquad \mathbb{E}[\varepsilon_i \,|\, \boldsymbol{x}_i] = 0$$

- n observations are independent across i
- full rank m of design matrix $\boldsymbol{S}_{xx} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i' = X'X$

Hypothesis of interest:

$$H_0 \; : \; \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{q}$$

where matrix $\boldsymbol{R} \in \mathbb{R}^{r \times m}$ has full row rank r, and $\boldsymbol{q} \in \mathbb{R}^r$

Framework

Linear regression:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad\qquad \mathbb{E}[\varepsilon_i \,|\, \boldsymbol{x}_i] = 0$$

- n observations are independent across i
- full rank m of design matrix $\boldsymbol{S}_{xx} = \sum_{i=1}^{n} \boldsymbol{x}_i \boldsymbol{x}_i' = X'X$

Hypothesis of interest:

$$H_0 \;:\; \boldsymbol{R}\boldsymbol{\beta} = \boldsymbol{q}$$

where matrix $\boldsymbol{R} \in \mathbb{R}^{r\times m}$ has full row rank r, and $\boldsymbol{q} \in \mathbb{R}^r$

Key point: number of regressors and restrictions may be large relative to sample size and error term may be heteroskedastic

Outline

Repeating setup

Linear regression:

$$y_i = \boldsymbol{x}_i'\boldsymbol{\beta} + \varepsilon_i \qquad\qquad \mathbb{E}[\varepsilon_i|\boldsymbol{x}_i] = 0$$

- n observations are independent across i
- full rank m of design matrix $\boldsymbol{S}_{xx} = \sum_{i=1}^n \boldsymbol{x}_i\boldsymbol{x}_i' = X'X$
- heteroskedasticity in errors

$$\mathbb{E}[\varepsilon_i^2|\boldsymbol{x}_i] = \sigma^2(\boldsymbol{x}_i) \equiv \sigma_i^2$$

Hypothesis of interest:

$$H_0 \ : \ \boldsymbol{R\beta} = \boldsymbol{q}$$

where matrix $\boldsymbol{R} \in \mathbb{R}^{r\times m}$ has full row rank r, and $\boldsymbol{q} \in \mathbb{R}^r$

F-statistic

Define

- OLS estimator of $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}} = \boldsymbol{S}_{xx}^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i y_i$$

- residual variance

$$\hat{\sigma}_{\varepsilon}^2 = \frac{1}{n-m} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i' \hat{\boldsymbol{\beta}})^2$$

F-statistic

Define

- OLS estimator of $\boldsymbol{\beta}$

$$\hat{\boldsymbol{\beta}} = \boldsymbol{S}_{xx}^{-1} \sum_{i=1}^{n} \boldsymbol{x}_i y_i$$

- residual variance

$$\hat{\sigma}_{\varepsilon}^2 = \frac{1}{n-m} \sum_{i=1}^{n} (y_i - \boldsymbol{x}_i' \hat{\boldsymbol{\beta}})^2$$

F-statistic is

$$F = \frac{\left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{q}\right)' \left(\boldsymbol{R}\boldsymbol{S}_{xx}^{-1}\boldsymbol{R}'\right)^{-1} \left(\boldsymbol{R}\hat{\boldsymbol{\beta}} - \boldsymbol{q}\right)}{r\hat{\sigma}_{\varepsilon}^2}$$

- let $\mathcal{F}$ denote numerator of F, i.e., $\mathcal{F} = r\hat{\sigma}_{\varepsilon}^2 F$

Conventional critical value

"Exact" F test of size $\alpha \in [0, 1]$ reads

$$\phi_{\mathsf{EF}}(\alpha) = \mathbf{1}\left\{F > q_{1-\alpha}(F_{r,n-m})\right\}$$

Conventional critical value

"Exact" F test of size $\alpha \in [0, 1]$ reads

$$\phi_{\mathsf{EF}}(\alpha) = \mathbf{1}\left\{F > q_{1-\alpha}(F_{r,n-m})\right\}$$

Centering and scaling yields

$$\phi_{\mathsf{EF}}(\alpha) = \mathbf{1}\left\{\frac{F-1}{\sqrt{2/r+2/(n-m)}} \qquad > \frac{q_{1-\alpha}(F_{r,n-m})-1}{\sqrt{2/r+2/(n-m)}}\right\}$$

Here,

Conventional critical value

"Exact" F test of size $\alpha \in [0,1]$ reads

$$\phi_{\mathsf{EF}}(\alpha) = \mathbf{1}\left\{F > q_{1-\alpha}(F_{r,n-m})\right\}$$

Centering and scaling yields

$$\phi_{\mathsf{EF}}(\alpha) = \mathbf{1}\left\{\frac{\mathcal{F} - r\hat{\sigma}_{\varepsilon}^{2}}{\sqrt{2r\hat{\sigma}_{\varepsilon}^{4} + 2r^{2}\hat{\sigma}_{\varepsilon}^{4}/(n-m)}} > \frac{q_{1-\alpha}(F_{r,n-m}) - 1}{\sqrt{2/r + 2/(n-m)}}\right\}$$

Here,

- $r\hat{\sigma}_{\varepsilon}^{2}$ estimates null mean of $\mathcal{F}$
- $2r\hat{\sigma}_{\varepsilon}^{4} + 2r^{2}\hat{\sigma}_{\varepsilon}^{4}/(n-m)$ estimates null variance of $\mathcal{F} - r\hat{\sigma}_{\varepsilon}^{2}$
- these estimators are inconsistent, in general

New critical value

Change red quantities to robust versions:

$$\phi_{\mathsf{LO}}(\alpha) = \mathbf{1}\left\{ \frac{\mathcal{F} - \hat{E}_{\mathcal{F}}}{\hat{V}_{\mathcal{F}}^{1/2}} > \frac{q_{1-\alpha}(F_{r,n-m}) - 1}{\sqrt{2/r + 2/(n-m)}} \right\}$$

where

- $\hat{E}_{\mathcal{F}}$ is (conditionally) unbiased for null mean of $\mathcal{F}$
- $\hat{V}_{\mathcal{F}}$ is unbiased for null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$

With many restrictions and heteroskedasticity, $\phi_{\mathsf{LO}}(\alpha)$ controls size

New critical value

Change red quantities to robust versions:

$$\phi_{\mathsf{LO}}(\alpha) = \mathbf{1}\left\{ \frac{\mathcal{F} - \hat{E}_{\mathcal{F}}}{\hat{V}_{\mathcal{F}}^{\;1/2}} > \frac{q_{1-\alpha}(F_{r,n-m}) - 1}{\sqrt{2/r + 2/(n-m)}} \right\}$$

where

- $\hat{E}_{\mathcal{F}}$ is (conditionally) unbiased for null mean of $\mathcal{F}$
- $\hat{V}_{\mathcal{F}}$ is unbiased for null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$

With many restrictions and heteroskedasticity, $\phi_{\mathsf{LO}}(\alpha)$ controls size

Expressed in terms of F, we have $\phi_{\mathsf{LO}}(\alpha) = \mathbf{1}\{F > c_\alpha\}$ where

$$c_\alpha = \frac{1}{r\hat{\sigma}_\varepsilon^2} \left(\hat{E}_{\mathcal{F}} + \hat{V}_{\mathcal{F}}^{1/2} \frac{q_{1-\alpha}(F_{r,n-m}) - 1}{\sqrt{2/r + 2/(n-m)}} \right)$$

Outline

Leave-out estimability

Usually imposed assumption of leave-one-out full rank:

$\sum_{\ell \neq i} \boldsymbol{x}_\ell \boldsymbol{x}_\ell'$ is invertible for every $i \in \{1, \ldots, n\}$

equivalent to $P_{ii} = \boldsymbol{x}_i' \boldsymbol{S}_{xx}^{-1} \boldsymbol{x}_i < 1$ for all i

Here, we assume "leave-three-out full rank" of design matrix

Assumption 1

$\sum_{\ell \neq i,j,k} \boldsymbol{x}_\ell \boldsymbol{x}_\ell'$ is invertible for every $i, j, k \in \{1, \ldots, n\}$

- ♡ allows us to construct $\hat{\boldsymbol{\beta}}_{-ijk} = \left(\sum_{\ell \neq i,j,k} \boldsymbol{x}_\ell \boldsymbol{x}_\ell' \right)^{-1} \sum_{\ell \neq i,j,k} \boldsymbol{x}_\ell y_\ell$
- ♡ if not satisfied, we relax it to "leave-one-out full rank" at expense of smaller test size

Location problem

Null mean of $\mathcal{F}$ is

$$\sum_{i=1}^{n} B_{ii}\sigma_i^2$$

where $B_{ii} = B(\boldsymbol{x}_i, \boldsymbol{S}_{xx}, \boldsymbol{R})$ is observable

Location problem

Null mean of $\mathcal{F}$ is

$$\sum_{i=1}^{n} B_{ii}\sigma_i^2$$

where $B_{ii} = B(\boldsymbol{x}_i, \boldsymbol{S}_{xx}, \boldsymbol{R})$ is observable

Unbiased estimator of σ_i^2 from Kline, Saggio, and Sølvsten (2020) is

$$\hat{\sigma}_i^2 = y_i\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_{-i}\big)$$

Location problem

Null mean of $\mathcal{F}$ is

$$\sum_{i=1}^{n} B_{ii}\sigma_i^2$$

where $B_{ii} = B(\boldsymbol{x}_i, \boldsymbol{S}_{xx}, \boldsymbol{R})$ is observable

Unbiased estimator of σ_i^2 from Kline, Saggio, and Sølvsten (2020) is

$$\hat{\sigma}_i^2 = y_i\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_{-i}\big)$$

We therefore define

$$\hat{E}_{\mathcal{F}} = \sum_{i=1}^{n} B_{ii}\hat{\sigma}_i^2$$

Variance problem

Null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$ is

$$\sum_{i=1}^{n} \sum_{j \neq i} U_{ij} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{n} \left(\sum\nolimits_{j \neq i} V_{ij} \boldsymbol{x}_j' \boldsymbol{\beta} \right)^2 \sigma_i^2,$$

where $U_{ij} = U(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ and $V_{ij} = V(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ are observable

Variance problem

Null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$ is

$$\sum_{i=1}^{n} \sum_{j \neq i} U_{ij} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{n} \Big(\sum\nolimits_{j \neq i} V_{ij} \boldsymbol{x}_j' \boldsymbol{\beta} \Big)^2 \sigma_i^2,$$

where $U_{ij} = U(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ and $V_{ij} = V(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ are observable

Issue: $\hat{\sigma}_i^2 \hat{\sigma}_j^2$ is systematically biased due to dependence between $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$

Variance problem

Null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$ is

$$\sum_{i=1}^{n} \sum_{j \neq i} U_{ij} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{n} \Big(\sum\nolimits_{j \neq i} V_{ij} \boldsymbol{x}_j' \boldsymbol{\beta} \Big)^2 \sigma_i^2,$$

where $U_{ij} = U(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ and $V_{ij} = V(\boldsymbol{x}_i, \boldsymbol{x}_j, \boldsymbol{S}_{xx}, \boldsymbol{R})$ are observable

Issue: $\hat{\sigma}_i^2 \hat{\sigma}_j^2$ is systematically biased due to dependence between $\hat{\sigma}_i^2$ and $\hat{\sigma}_j^2$

Solution: use leave-three-out estimates of individual variances

$$\hat{\sigma}_{i,-jk}^2 = y_i \big(y_i - \boldsymbol{x}_i' \hat{\boldsymbol{\beta}}_{-ijk} \big)$$

Leave-three-out

So we introduce leave-three-out estimators of individual variances

$$\hat{\sigma}^2_{i,-jk} = y_i \big(y_i - \boldsymbol{x}_i' \hat{\boldsymbol{\beta}}_{-ijk} \big)$$

Leave-three-out

So we introduce leave-three-out estimators of individual variances

$$\hat{\sigma}^2_{i,-jk} = y_i\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_{-ijk}\big)$$

and analogous leave-two-out estimator

$$\hat{\sigma}^2_{i,-j} := \hat{\sigma}^2_{i,-jj} = y_i \sum_{k \neq j} \check{M}_{ik,-ij} y_k$$

where $\check{M}_{ik,-ij} = M(\boldsymbol{x}_i, \boldsymbol{x}_k, \boldsymbol{x}_j, \boldsymbol{S}_{xx})$ is observable

Leave-three-out

So we introduce leave-three-out estimators of individual variances

$$\hat{\sigma}^2_{i,-jk} = y_i\big(y_i - \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_{-ijk}\big)$$

and analogous leave-two-out estimator

$$\hat{\sigma}^2_{i,-j} := \hat{\sigma}^2_{i,-jj} = y_i \sum_{k\neq j} \check{M}_{ik,-ij} y_k$$

where $\check{M}_{ik,-ij} = M(\boldsymbol{x}_i, \boldsymbol{x}_k, \boldsymbol{x}_j, \boldsymbol{S}_{xx})$ is observable

Combining these for different i and j leads to product estimator

$$\widehat{\sigma^2_i \sigma^2_j} = y_i \sum_{k\neq j} \check{M}_{ik,-ij} y_k \cdot \hat{\sigma}^2_{j,-ik}$$

Variance estimator

We use these leave-out estimators to construct

$$\hat{V}_{\mathcal{F}} = \sum_{i=1}^{n} \sum_{j \neq i} \left(U_{ij} - V_{ij}^2 \right) \cdot \widehat{\sigma_i^2 \sigma_j^2} + \sum_{i=1}^{n} \sum_{j \neq i} \sum_{k \neq i} V_{ij} y_j \cdot V_{ik} y_k \cdot \hat{\sigma}_{i,-jk}^2$$

Variance estimator

We use these leave-out estimators to construct

$$\hat{V}_{\mathcal{F}} = \sum_{i=1}^{n} \sum_{j \neq i} \left(U_{ij} - V_{ij}^2 \right) \cdot \widehat{\sigma_i^2 \sigma_j^2} + \sum_{i=1}^{n} \sum_{j \neq i} \sum_{k \neq i} V_{ij} y_j \cdot V_{ik} y_k \cdot \hat{\sigma}_{i,-jk}^2$$

Recall that null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$ is

$$\sum_{i=1}^{n} \sum_{j \neq i} U_{ij} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{n} \left(\sum\nolimits_{j \neq i} V_{ij} \boldsymbol{x}_j' \boldsymbol{\beta} \right)^2 \sigma_i^2$$

Variance estimator

We use these leave-out estimators to construct

$$\hat{V}_{\mathcal{F}} = \sum_{i=1}^{n} \sum_{j \neq i} \Big(U_{ij} - V_{ij}^2 \Big) \cdot \widehat{\sigma_i^2 \sigma_j^2} + \sum_{i=1}^{n} \sum_{j \neq i} \sum_{k \neq i} V_{ij} y_j \cdot V_{ik} y_k \cdot \hat{\sigma}_{i,-jk}^2$$

Recall that null variance of $\mathcal{F} - \hat{E}_{\mathcal{F}}$ is

$$\sum_{i=1}^{n} \sum_{j \neq i} U_{ij} \sigma_i^2 \sigma_j^2 + \sum_{i=1}^{n} \Big(\sum\nolimits_{j \neq i} V_{ij} \boldsymbol{x}_j' \boldsymbol{\beta} \Big)^2 \sigma_i^2$$

Note: product of $(j = k)$-th terms in second component of $\hat{V}_{\mathcal{F}}$ generate bias of $V_{ij}^2 \sigma_i^2 \sigma_j^2 \Rightarrow U_{ij} - V_{ij}^2$ instead of U_{ij} in first component of $\hat{V}_{\mathcal{F}}$

Recap

We propose test

$$\phi_{\mathsf{LO}}(\alpha) = \mathbf{1}\left\{F > c_\alpha\right\}$$

where

$$c_\alpha = \frac{1}{r\hat{\sigma}_\varepsilon^2}\left(\hat{E}_{\mathcal{F}} + \hat{V}_{\mathcal{F}}^{1/2}\frac{q_{1-\alpha}(F_{r,n-m}) - 1}{\sqrt{2/r + 2/(n-m)}}\right)$$

$$\begin{aligned}
\hat{E}_{\mathcal{F}} &= \sum_{i=1}^{n} B_{ii}\hat{\sigma}_i^2 \\
\hat{V}_{\mathcal{F}} &= \sum_{i=1}^{n}\sum_{j\neq i}\Big(U_{ij} - V_{ij}^2\Big)\cdot\widehat{\sigma_i^2\sigma_j^2} + \sum_{i=1}^{n}\sum_{j\neq i}\sum_{k\neq i} V_{ij}y_j \cdot V_{ik}y_k \cdot \hat{\sigma}_{i,-jk}^2
\end{aligned}$$

Outline

1. Framework and hypothesis test
2. Leave-out-estimation
3. Asymptotic theory
4. Simulations

Regularity conditions

Assumption 2

$(i)\ \max_i \left(\mathbb{E}[\varepsilon_i^4|\boldsymbol{x}_i] + \sigma_i^{-2}\right) = O_p(1)$
$(ii)\ \max_i \left(\boldsymbol{x}_i'\boldsymbol{\beta}\right)^2 = O_p(1)$
$(iii)\ \max_i \left(\sum_{j\neq i} V_{ij}\boldsymbol{x}_j'\boldsymbol{\beta}\right)^2/r = o_p(1)$
$(iv)\ \max_{i\neq j\neq k\neq i} D_{ijk}^{-1} = O_p(1)$

- (i) and (ii) place tail restrictions on data
- (iii) is high-level condition on weight each observation receives
- (iv) is slight strengthening of Assumption 1, as construction of $\hat{V}_{\mathcal{F}}$ involves division by

$$D_{ijk} = \begin{vmatrix} M_{ii} & M_{ij} & M_{ik} \\ M_{ij} & M_{jj} & M_{jk} \\ M_{ik} & M_{jk} & M_{kk} \end{vmatrix} \quad \text{where } M_{ij} = \mathbf{1}\{i=j\} - \boldsymbol{x}_i'\boldsymbol{S}_{xx}^{-1}\boldsymbol{x}_j$$

Size and power

Theorem 1

If Assumptions 1 and 2 hold, then under H_0

$$\lim_{n,r\to\infty} \Pr\left(F > c_\alpha\right) = \alpha$$

Theorem 2

Consider sequence of local alternatives

$$H_\delta \;:\; \boldsymbol{R\beta} - \boldsymbol{q} = \left(\boldsymbol{R}\boldsymbol{S}_{xx}^{-1}\boldsymbol{R}'\right)^{1/2} \cdot \boldsymbol{\delta}$$

If Assumptions 1 and 2 hold, then

$$\lim_{n,r\to\infty} \Pr\left(F > c_\alpha\right) - \Phi\left(\Phi^{-1}\left(\alpha\right) + \mathbb{V}_0\left[\mathcal{F} - \hat{E}_{\mathcal{F}}\right]^{-1/2} ||\boldsymbol{\delta}||^2\right) = 0$$

Outline

Simulation setup

Inspired by MacKinnon (2013), outcome equation is

$$y_i = \beta_1 + \sum_{k=2}^{m} \beta_k x_{ik} + \varepsilon_i$$

where

- data is drawn *i.i.d.* across i
- sample sizes take values $80,\ 160,\ 320,\ 640$ and 1280
- number of unknown coefficients is $m = 0.8n$
- coefficient of determination R^2 is equal to 0.16
- homo-/heteroskedastic errors are $\varepsilon_i \,|\, \boldsymbol{x}_i \sim N(0, \sigma_i^2)$ with

$$\sigma_i = z_\zeta \, (1 + s_i)^\zeta$$

  for complex $s_i = f(x_{i2}, \ldots, x_{im})$ and $\zeta \in \{0, 2\}$ and $\mathbb{E}[\sigma_i^2] = 1$

Two designs and alternatives

Regressor design: $x_{i2}, \ldots, x_{im}$ are products of *i.i.d.* log-normals and $0.5 + u_i$ where u_i is standard uniform

The hypothesis of interest restricts values of last $r = 0.6n$ coefficients

Two types of alternatives:

– Sparse: a single coefficient deviates

– Dense: all coefficients deviate in equal proportions

Size

Nominal size			5%		
Test		LO	EF	W	$\%\hat{V}_{\mathcal{F}} < 0$
Homoskedasticity					
$n = 80$	$r = 48$	3	5	9	19.9
$n = 160$	$r = 96$	5	5	98	6.7
$n = 320$	$r = 192$	6	5	69	1.8
$n = 640$	$r = 384$	5	5	100	0.2
$n = 1280$	$r = 768$	5	5	100	0.0
Heteroskedasticity					
$n = 80$	$r = 48$	4	47	18	13.4
$n = 160$	$r = 96$	5	62	89	4.4
$n = 320$	$r = 192$	5	82	100	0.9
$n = 640$	$r = 384$	5	97	100	0.2
$n = 1280$	$r = 768$	5	100	100	0.0

Size

Nominal size		1%			5%			10%			
Test		LO	EF	W	LO	EF	W	LO	EF	W	$\%\hat{V}_{\mathcal{F}} < 0$
Homoskedasticity											
$n = 80$	$r = 48$	1	1	7	3	5	9	6	10	11	19.9
$n = 160$	$r = 96$	2	1	97	5	5	98	10	10	98	6.7
$n = 320$	$r = 192$	1	1	68	6	5	69	11	10	69	1.8
$n = 640$	$r = 384$	1	1	100	5	5	100	10	10	100	0.2
$n = 1280$	$r = 768$	1	1	100	5	5	100	10	10	100	0.0
Heteroskedasticity											
$n = 80$	$r = 48$	1	22	13	4	47	18	7	62	23	13.4
$n = 160$	$r = 96$	1	33	84	5	62	89	9	76	92	4.4
$n = 320$	$r = 192$	1	57	100	5	82	100	10	90	100	0.9
$n = 640$	$r = 384$	1	87	100	5	97	100	11	99	100	0.2
$n = 1280$	$r = 768$	1	99	100	5	100	100	10	100	100	0.0

Power

Deviation				Dense			
Nominal size				5%		10%	
Test				LO	EF	LO	EF
Homoskedasticity							
$n = 80$	$r = 48$			4	15	9	25
$n = 160$	$r = 96$			13	22	21	35
$n = 320$	$r = 192$			24	34	37	48
$n = 640$	$r = 384$			44	56	58	71
$n = 1280$	$r = 768$			68	84	80	92

Power

Deviation		Sparse				Dense			
Nominal size		5%		10%		5%		10%	
Test		LO	EF	LO	EF	LO	EF	LO	EF
Homoskedasticity									
$n = 80$	$r = 48$	6	15	12	25	4	15	9	25
$n = 160$	$r = 96$	15	23	25	35	13	22	21	35
$n = 320$	$r = 192$	28	35	42	48	24	34	37	48
$n = 640$	$r = 384$	48	54	62	67	44	56	58	71
$n = 1280$	$r = 768$	74	80	85	89	68	84	80	92

Summary and prospects

We propose new test (or new critical value for standard F statistic) with heteroskedasticity and many restrictions

- Size is controlled asymptotically and also in quite small samples
- There is acceptable loss in power from robustness to heteroskedasticity
- We extend to discrete regressors when invertability of $\sum_{\ell \neq i,j,k} \boldsymbol{x}_\ell \boldsymbol{x}_\ell'$ fails for some i, j, k by intentionally increasing $\hat{V}_{\mathcal{F}}$ resulting in smaller test size
- We have also developed extension robust to few regressors and restrictions