

Analiza danych panelowych cz. V

Stanisław Cichocki

Natalia Nehrebecka

Plan zajęć

▶ Diagnostyka

- a) Test RESET
- b) Test Jarque-Bera
- c) Testowanie heteroskedastyczności
 - a) groupwise heteroscedasticity
 - b) cross-sectional correlation
- d) Testowanie autokorelacji

Test RESET

- ▶ W celu sprawdzenia poprawności formy funkcyjnej modelu stosuje się test *RESET* (ang. *Regression Specification Error Test*), którego hipoteza zerowa zakłada liniową formę modelu.

```
=====
*** REgression Specification Error Tests (RESET) - Model= (xtre)
=====
Ho: Model is Specified - Ha: Model is Misspecified
-----
* Ramsey Specification ResetF Test
- Ramsey RESETF1 Test: Y= X Yh2          = 44.970 P-Value > F(1, 196) 0.0000
- Ramsey RESETF2 Test: Y= X Yh2 Yh3      = 26.572 P-Value > F(2, 195) 0.0000
- Ramsey RESETF3 Test: Y= X Yh2 Yh3 Yh4  = 17.846 P-Value > F(3, 194) 0.0000
-----
```

Test Jarque-Bera

- ▶ Test Jarque-Bera zakłada, że rozkład błędów losowych zachowuje się zgodnie z założeniami rozkładu normalnego. Jednocześnie hipoteza alternatywna mówi o braku normalności rozkładu.

	Observed	Bootstrap			Normal-based	
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
Skewness_e	1.310567	.6366598	2.06	0.040	.0627363	2.558397
Kurtosis_e	14.62481	6.854588	2.13	0.033	1.190066	28.05956
Skewness_u	-.2685204	.9683457	-0.28	0.782	-2.166443	1.629402
Kurtosis_u	1.419064	1.318737	1.08	0.282	-1.165612	4.00374
Joint test for Normality on e:				chi2(2) = 8.79	Prob > chi2 = 0.0123	
Joint test for Normality on u:				chi2(2) = 1.23	Prob > chi2 = 0.5393	

Niespełnienie założeń

- ▶ W modelu *Random Effects* zakładaliśmy, że

$$\text{Var}(\boldsymbol{\varepsilon}) = \sigma_{\varepsilon}^2 \mathbf{I}$$

$$\text{Var}(\mathbf{u}) = \sigma_u^2 \mathbf{I}$$

- ▶ W modelu *Fixed Effects* zakładaliśmy, że

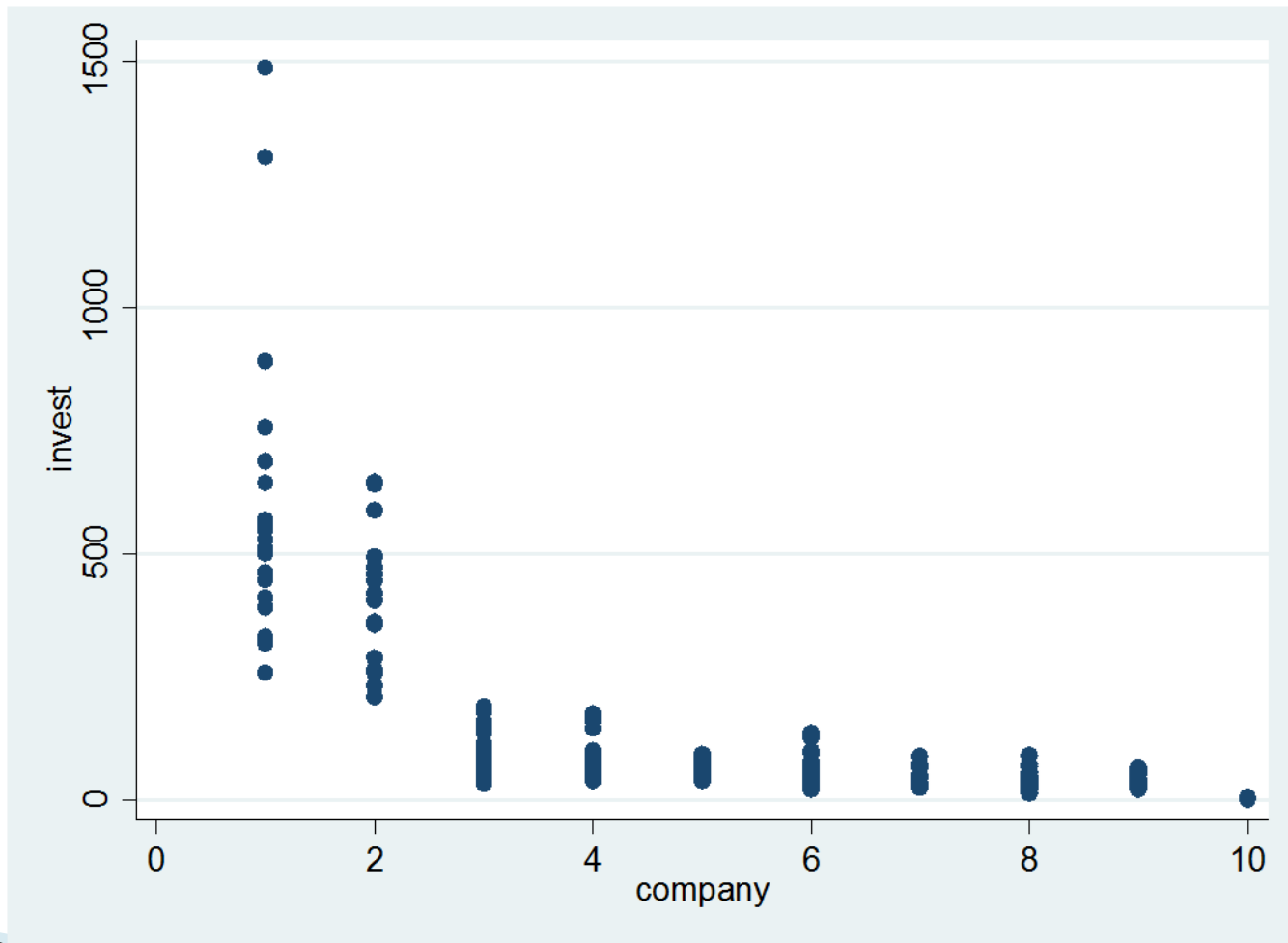
$$\text{Var}(\boldsymbol{\varepsilon}) = \sigma_{\varepsilon}^2 \mathbf{I}$$

Niespełnienie założeń – grupowa heteroskedastyczność

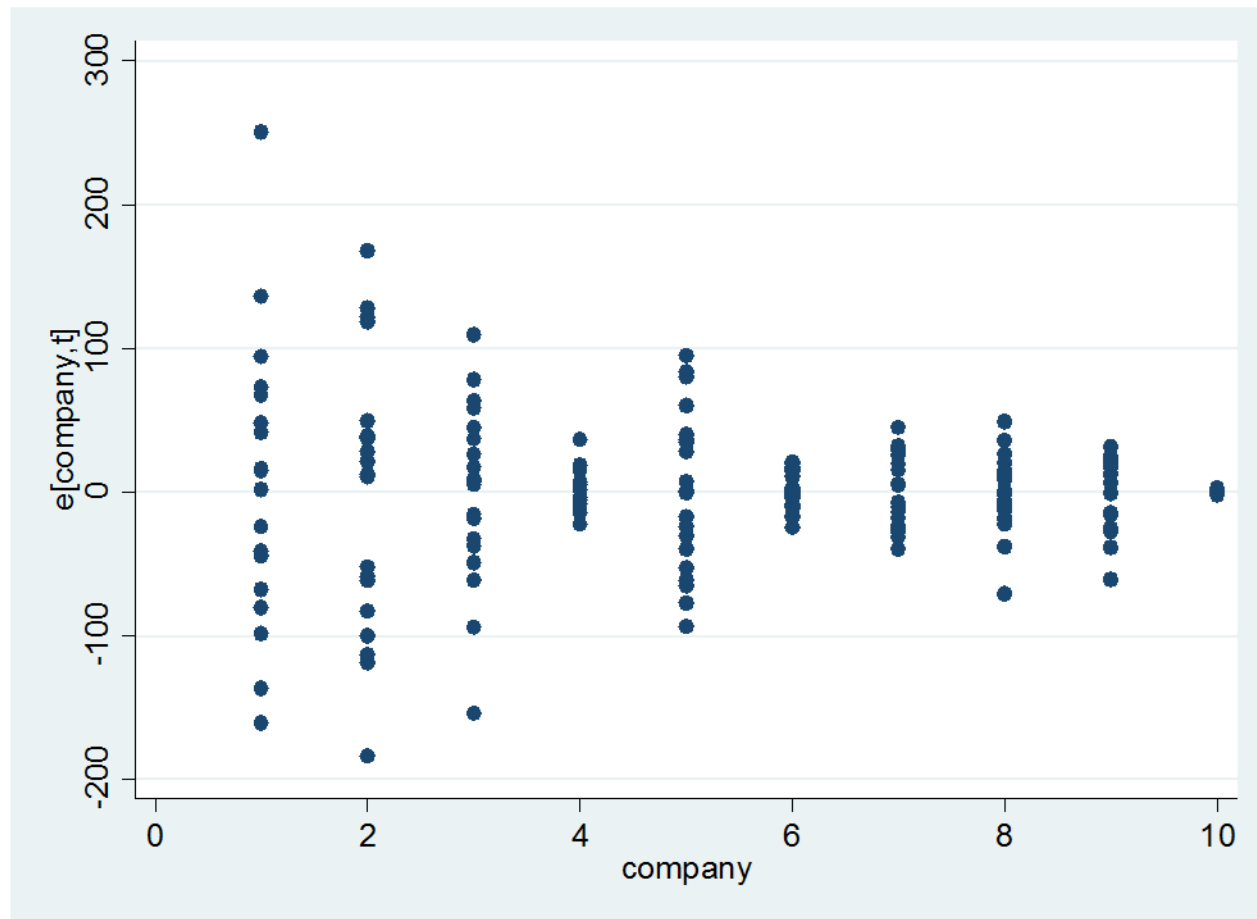
- ▶ **groupwise heteroscedasticity** – macierze Ω_i są różne dla **poszczególnych jednostek**, co może być spowodowane zmiennością wariancji czysto losowego składnika lub (w modelu *random effects*) zmiennością wariancji efektów indywidualnych.

$$\bar{\Omega} = \begin{bmatrix} \Omega & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Omega \end{bmatrix} \quad \text{vs} \quad \bar{\Omega} = \begin{bmatrix} \Omega_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Omega_N \end{bmatrix}$$

Niespełnienie założeń – grupowa heteroskedastyczność



Niespełnienie założeń – grupowa heteroskedastyczność



Testowanie grupowej heteroskedastyczności

- ▶ **1.** Podstawowym testem wykrywającym występowanie heteroskedastyczności błędów losowych jest **test mnożników Lagrange'a**.
 - Dysponując resztami z modelu z ograniczeniami (*restricted*), czyli zakładającej homoskedastyczność

$$LM = \frac{T_i}{2} \sum_i \left(\frac{\hat{\sigma}_i^2}{\hat{\sigma}^2} - 1 \right)^2$$

- ▶ Statystyka testowa: $\chi^2(N - 1)$

Testowanie grupowej heteroskedastyczności

- ▶ 2. Wstawiając estymatory uzyskane analogicznie na podstawie reszt z modelu bez ograniczeń (*FGLS*), otrzymamy statystykę **Walda**:

$$W = \frac{T_i}{2} \sum_i \left(\frac{\hat{\sigma}^2}{\hat{\sigma}_i^2} - 1 \right)^2$$

- ▶ Statystyka testowa: $\chi^2(N)$

Testowanie grupowej heteroskedastyczności

- ▶ 3. I wreszcie, estymując oba modele za pomocą metody największej wiarygodności (lub iteracyjnej FGLS) i znajdując $\hat{\sigma}^2$ z *restricted* model oraz $\hat{\sigma}_i^2$ z *unrestricted* model, możemy obliczyć statystykę **testu ilorazu wiarygodności**:

$$LR = NT_i \ln \hat{\sigma}^2 - \sum_i T_i \ln \hat{\sigma}_i^2$$

- ▶ Statystyka testowa: $\chi^2(N - 1)$
- ▶ **Niestety, wszystkie trzy statystyki są wrażliwe na założenie o normalności reszt.**

Testowanie grupowej heteroskedastyczności

Likelihood-ratio test
(Assumption: homosk nested in hetero)

LR chi2(9) = 470.23
Prob > chi2 = 0.0000

Testowanie grupowej heteroskedastyczności

- ▶ W przypadku braku spełnienia założenia dotyczącego rozkładu normalnego reszt stosuje się **skorygowaną statystykę Walda**:

$$W = \sum_{i=1}^N \frac{(\hat{\sigma}_i^2 - \hat{\sigma}^2)^2}{V_i}$$

- ▶ gdzie:

$$V_i = \frac{1}{T_i(T_i - 1)} \sum_{t=1}^{T_i} (e_{it}^2 - \hat{\sigma}_i^2)^2$$

- $\hat{\sigma}_i^2 = \frac{1}{T_i} \sum_{t=1}^{T_i} e_{it}^2$
- $\hat{\sigma}^2 = \frac{1}{N} \sum_i \hat{\sigma}_i^2$

- ▶ Statystyka krytyczna: $\chi^2(N)$

Testowanie grupowej heteroskedastyczności

Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: $\sigma(i)^2 = \sigma^2$ for all i

```
chi2 (10) = 1.7e+07  
Prob>chi2 = 0.0000
```

Testowanie grupowej heteroskedastyczności

Fixed-effects (within) regression
Group variable: company

Number of obs = 200
Number of groups = 10

R-sq:

within = 0.7668
between = 0.8194
overall = 0.8060

Obs per group:

min = 20
avg = 20.0
max = 20

F(2, 9) = 28.31

Prob > F = 0.0001

corr(u_i, X_b) = -0.1517

(Std. Err. adjusted for 10 clusters in company)

	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
invest						
mvalue	.1101238	.0151945	7.25	0.000	.0757515	.1444961
kstock	.3100653	.0527518	5.88	0.000	.1907325	.4293981
_cons	-58.74393	27.60286	-2.13	0.062	-121.1859	3.698079
sigma_u	85.732501					
sigma_e	52.767964					
rho	.72525012	(fraction of variance due to u _i)				

Niespełnienie założeń – korelacja międzygrupowa

- ▶ **cross-sectional correlation** – w macierzy blokowej $\bar{\Omega}$ poza przekątną występują macierze niezerowe; sytuację taką najczęściej opisuje się poprzez współczynnik korelacji równoczesnych realizacji „czystego” zaburzenia losowego

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{jt}) = \rho_{ij}$$

$$\bar{\Omega} = \begin{bmatrix} \Omega_1 & \cdots & \mathbf{0} \\ \vdots & \ddots & \vdots \\ \mathbf{0} & \cdots & \Omega_N \end{bmatrix}$$

Testowanie korelacji międzygrupowej

- ▶ 1. Test mnożnika Lagrange'a Breucha-Pagana będzie miał postać:

$$LM = T \sum_{i=2}^N \sum_{j=1}^{i-1} r_{ij}^2$$

- gdzie r_{ij} , to współczynniki korelacji reszt między obiektami i -tym i j -tym uzyskanymi z oszacowania *UMNK* (ze względu na założenie stałości parametrów strukturalnych względem obiektów).
- Statystyka krytyczna: $\chi^2 \left(\frac{N(N-1)}{2} \right)$
 - H_0 : brak korelacji międzygrupowej
 - UWAGA! stosować gdy *duże T i małe N*

Testowanie korelacji międzygrupowej

Correlation matrix of residuals:

	__e1	__e2	__e3	__e4	__e5	__e6	__e7	__e8	__e9
__e1	1.0000								
__e2	-0.0533	1.0000							
__e3	-0.2173	-0.0935	1.0000						
__e4	-0.1912	0.1540	0.2084	1.0000					
__e5	-0.4555	-0.3997	0.6075	0.0490	1.0000				
__e6	-0.2489	-0.1310	0.6277	0.1334	0.8320	1.0000			
__e7	-0.2234	-0.5074	0.5516	0.0291	0.9318	0.7896	1.0000		
__e8	-0.4390	-0.1120	0.8832	0.2432	0.8081	0.7542	0.7451	1.0000	
__e9	-0.3939	-0.3290	0.7952	0.0675	0.8772	0.7229	0.8399	0.9051	1.0000
__e10	0.4196	0.3937	-0.1534	-0.0030	-0.6206	-0.3208	-0.5452	-0.4038	-0.5352
__e10									
__e10	1.0000								

Breusch-Pagan LM test of independence: $\chi^2(45) = 246.329$, **Pr = 0.0000**
Based on 20 complete observations over panel units

Testowanie korelacji międzygrupowej

▶ 2. Testy Pesarana, Friedmana i Freesa

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{jt}) = \rho_{ij}$$

▶ Test Pesarana

$$\text{Statystyka testowa: } CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right)$$

W przypadku paneli niezbilansowanych: $CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \sqrt{T_{ij}} \hat{\rho}_{ij} \right)$

▶ Test Friedmana

$$\text{Statystyka testowa: } R_{ave} = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{r}_{ij}$$

◦ $r_{ij} = r_{ji} = \frac{\sum_{t=1}^T \{r_{i,t-(T+1/2)}\} \{r_{j,t-(T+1/2)}\}}{\sum_{t=1}^T \{r_{i,t-(T+1/2)}\}^2}$

▶ Test Freesa

$$\text{Statystyka testowa: } R_{ave}^2 = \frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{r}_{ij}^2$$

Testowanie korelacji międzygrupowej

Pesaran's test of cross sectional independence =
4.661, **Pr = 0.0000**

Testowanie korelacji międzygrupowej

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares

Panels: **heteroskedastic with cross-sectional correlation**

Correlation: no autocorrelation

Estimated covariances	=	55	Number of obs	=	200
Estimated autocorrelations	=	0	Number of groups	=	10
Estimated coefficients	=	3	Time periods	=	20
			Wald chi2(2)	=	3738.07
			Prob > chi2	=	0.0000

invest	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mvalue	.1127515	.0022364	50.42	0.000	.1083683	.1171347
kstock	.2231176	.0057363	38.90	0.000	.2118746	.2343605
_cons	-39.84382	1.717563	-23.20	0.000	-43.21018	-36.47746

Niespełnienie założeń - autokorelacja

- ▶ mogą wystąpić np. procesy autoregresyjne i/lub średniej ruchomej.

$$\text{Cov}(\varepsilon_{it}, \varepsilon_{is}) = \rho_{ts}$$

Niespełnienie założeń - autokorelacja

- ▶ Kolejnym ważnym zagadnieniem jest możliwość wystąpienia autokorelacji składnika losowego.
- ▶ Dla modelu *RE* stosuje się *Test* wykorzystujący test mnożników Lagrange'a.
 - **Hipoteza zerowa** zakłada brak występowania efektów losowych oraz brak autokorelacji pierwszego rzędu składników czysto losowych:
 - ρ – współczynnik autokorelacji składników czysto losowych

Niespełnienie założeń - autokorelacja

Tests for the error component model:

$$\text{invest}[\text{company}, t] = Xb + u[\text{company}] + v[\text{company}, t]$$

$$v[\text{company}, t] = \lambda v[\text{company}, (t-1)] + e[\text{company}, t]$$

Estimated results:

	Var	sd = sqrt(Var)
invest	47034.89	216.8753
e	2784.458	52.767964
u	7089.8	84.20095

Tests:

Serial Correlation:

$$\text{ALM}(\lambda=0) = 10.31 \quad \text{Pr}>\chi^2(1) = 0.0013$$

Joint Test:

$$\text{LM}(\text{Var}(u)=0, \lambda=0) = 808.47 \quad \text{Pr}>\chi^2(2) = 0.0000$$

Niespełnienie założeń - autokorelacja

- ▶ Testujemy autokorelację testem **Woolridge'a** (krótkie panele)
 - **H_0** : brak autokorelacji pierwszego rzędu
 - **H_1** : występuje autokorelacja pierwszego rzędu

- ▶ Rozwiązanie (gdy autokorelacja typu AR(1) jedynym problemem): estymatory uwzględniające taką strukturę składnika losowego
 - uogólnione modele z zakłóceniem AR(1):

Niespełnienie założeń - autokorelacja

Wooldridge test for autocorrelation in panel data

H0: no first-order autocorrelation

F(1, 9) = 263.592

Prob > F = 0.0000

Niespełnienie założeń - autokorelacja

Cross-sectional time-series FGLS regression

Coefficients: generalized least squares

Panels: heteroskedastic with cross-sectional correlation

Correlation: common AR(1) coefficient for all panels (0.9261)

Estimated covariances	=	55	Number of obs	=	200
Estimated autocorrelations	=	1	Number of groups	=	10
Estimated coefficients	=	3	Time periods	=	20
			Wald chi2(2)	=	220.77
			Prob > chi2	=	0.0000

invest	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
mvalue	.0826837	.0063148	13.09	0.000	.0703068	.0950605
kstock	.1942327	.0303987	6.39	0.000	.1346524	.253813
_cons	-18.6185	4.953675	-3.76	0.000	-28.32752	-8.909473

Dziękuję za uwagę