

# Prognozowanie i symulacje

## Wykład 4

---

# Testy stabilności parametrów strukturalnych

---

- Test Chowa
- Test CUSUM
- Test QLR

# Test stabilności parametrów strukturalnych – test Chowa

---

**Test Chowa** – do oceny stabilności parametrów strukturalnych / do oceny punktu zwrotnego **przy znanym** punkcie załamania

- $H_0$ : brak zmian strukturalnych  
 $H_A$ : występują zmiany strukturalne

- $$F = \frac{(SSE - SSE_1 - SSE_2) / k + 1}{(SSE_1 + SSE_2) / (n - 2(k + 1))}$$

SSE – suma kwadratów reszt w całej próbie,

SSE1 – suma kwadratów reszt w pierwszej podpróbie,

SSE2 – suma kwadratów reszt w drugiej podpróbie

- Statystyka F ma rozkład F z  $k+1$  i  $n-2(k+1)$  stopniami swobody

# Przykład – test Chowa

---

Roczne dane na temat pozycji czeskiego sektora usług (H46) w globalnych łańcuchach wartości w okresie 2000 – 2014.

Sprawdzamy stabilność parametrów strukturalnych pomiędzy dwoma przedziałami czasowymi: 2000-2011 i 2012-2014 (okres prognozowany)

Model 1: Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)

Zmienna zależna (Y): GVC\_position

	<i>Współczynnik</i>	<i>Błąd stand.</i>	<i>t-Studenta</i>	<i>wartość p</i>	
const	0,0868497	0,0100743	8,621	<0,0001	***
time	-0,00317799	0,00110803	-2,868	0,0132	**
Średn.aryt.zm.zależnej	0,061426		Odch.stand.zm.zależnej	0,02283	
Suma kwadratów reszt	0,004469		Błąd standardowy reszt	0,01854	
Wsp. determ. R-kwadrat	0,387550		Skorygowany R-kwadrat	0,34043	
F(1, 13)	8,226231		Wartość p dla testu F	0,01319	
Logarytm wiarygodności	39,60583		Kryt. inform. Akaike'a	-75,2116	
Kryt. bayes. Schwarza	-73,79556		Kryt. Hannana-Quinna	-75,2267	
Autokorel.reszt - rho1	0,695774		Stat. Durbina-Watsona	0,410073	

# Przykład – test Chowa

Pomocnicze równanie regresji dla testu Chowa

Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)

Zmienna zależna (Y): GVC\_position

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	0,0744820	0,00949764	7,842	7,89e-06	***
time	-0,000665199	0,00129048	-0,5155	0,6164	
splitdum	0,0135580	0,153321	0,08843	0,9311	
sd_time	-0,00373080	0,0109880	-0,3395	0,7406	

Średn. aryt. zm. zależnej	0,061426	Odch. stand. zm. zależnej	0,022830
Suma kwadratów reszt	0,002620	Błąd standardowy reszt	0,015432
Wsp. determ. R-kwadrat	0,641000	Skorygowany R-kwadrat	0,543091
F(3, 11)	6,546898	Wartość p dla testu F	0,008409
Logarytm wiarygodności	43,61192	Kryt. inform. Akaike'a	-79,22384
Kryt. bayes. Schwarza	-76,39164	Kryt. Hannana-Quinna	-79,25401
Autokorel. reszt - rho1	0,402544	Stat. Durbina-Watsona	1,002924

Test Chowa na zmiany strukturalne przy podziale próby w obserwacji 2012

$F(2, 11) = 3,88294$  z wartością p 0,0530

splitdum – zmienna 0-1, która przyjmuje wartość 1 dla wybranego okresu załamania i okresów dalszych, dla pozostałych okresów przyjmuje wartość 0

# Test stabilności parametrów strukturalnych – test QLR

---

Test QLR (ilorazu wiarygodności Quandta) – do oceny stabilności parametrów strukturalnych, gdy **nieznany jest moment załamania strukturalnego  $\tau$** .

Szukamy, momentu czasowego w którym wartość statystyki Chowa jest największa

Zał. Zmiany strukturalne mają charakter stochastyczny

Model podstawowy przyjmuje postać: 
$$y_t = \sum_{i=0}^k \beta_{it} x_{it} + \varepsilon_t$$

Parametry modelu są zmienne w czasie: 
$$\beta_{it} = \begin{cases} \beta_i, & t \leq \tau \\ \beta_i + \delta_i, & t > \tau \end{cases}$$

# Test stabilności parametrów strukturalnych – test QLR

---

Punkt zwrotny  $\tau$  jest nieznan. Weryfikujemy hipotezy:

$$H_0 : \delta_i = 0, H_A : \delta_i \neq 0$$

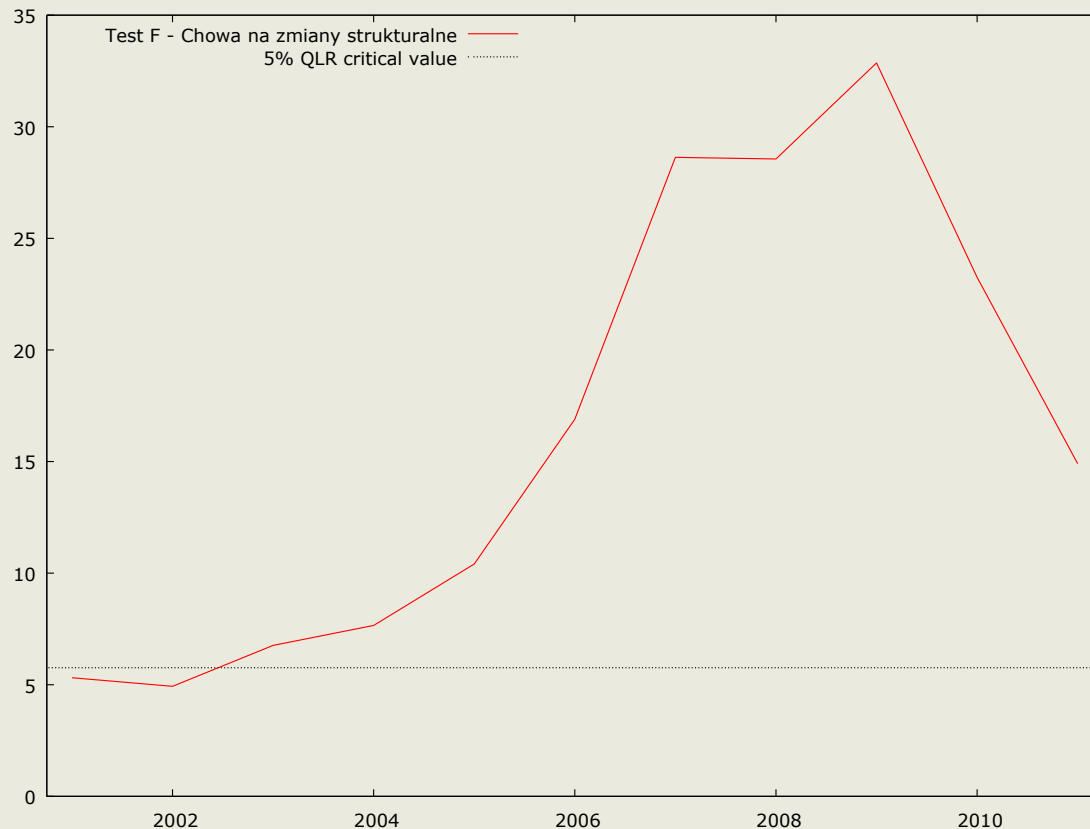
W modelu:

$$y_t = \sum_{i=0}^k \beta_{it} x_{it} + \sum_{i=0}^k \delta_i x_{it} z_t(\tau) + \varepsilon_t$$

Gdzie  $z_t(\tau)$  – zmienna 0-1 postaci:

$$z_t(\tau) = \begin{cases} 0, & t \leq \tau \\ 1, & t > \tau \end{cases}$$

# Test stabilności parametrów strukturalnych – test QLR



Test ilorazu wiarygodności Quandta na występowanie załamania strukturalnego

w nieokreślonym momencie, z 15 procentowym błędem oceny:

Maksymalne  $F(2, 11) = 32,8606$  dla obserwacji 2009

Asymptotyczna wartość  $p = 1,77661e-013$  dla  $\chi^2(2) = 65,7212$



# Test stabilności parametrów strukturalnych – test QLR vs. test Chowa

Test ilorazu wiarygodności Quandta na występowanie załamania strukturalnego

w nieokreślonym momencie, z 15 procentowym błędem oceny:

Maksymalne  $F(2, 11) = 32,8606$  dla obserwacji 2009

Asymptotyczna wartość  $p = 1,77661e-013$  dla  $\chi^2(2) = 65,7212$

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	0,0578790	0,00554460	10,44	4,81e-07	***
time	0,00343568	0,000985301	3,487	0,0051	***
splitdum	0,108700	0,0236756	4,591	0,0008	***
sd_time	-0,0134837	0,00207348	-6,503	4,41e-05	***

Średn. arytm. zm. zależnej	0,061426	Odch. stand. zm. zależnej	0,022830
Suma kwadratów reszt	0,000641	Błąd standardowy reszt	0,007632
Wsp. determ. R-kwadrat	0,912189	Skorygowany R-kwadrat	0,888241
$F(3, 11)$	38,08977	Wartość p dla testu F	4,20e-06
Logarytm wiarygodności	54,17295	Kryt. inform. Akaike'a	-100,3459
Kryt. bayes. Schwarz	-97,51369	Kryt. Hannana-Quinna	-100,3761
Autokorel. reszt - rho1	0,183221	Stat. Durbina-Watsona	1,468558

Test Chowa na zmiany strukturalne przy podziale próby w obserwacji 2009

$F(2, 11) = 32,8606$  z wartością  $p = 0,0000$

# Test stabilności parametrów strukturalnych – test CUSUM

---

Test CUSUM – weryfikacja stabilności modelu bez konieczności wskazywania punktu zwrotnego

Zał. Zmiany strukturalne mają charakter deterministyczny

Statystyka testowa t:

$$t = \frac{\bar{w}}{s} \sqrt{t - k - 1}$$

Gdzie wt:

$$w_t = \frac{\hat{\varepsilon}_t}{\sqrt{1 + x_t^T (X_{t-1}^T X_{t-1})^{-1} x_t}}$$

# Test stabilności parametrów strukturalnych – test CUSUM

Test CUSUM (CUMulated SUM of residual) na stabilność parametrów modelu

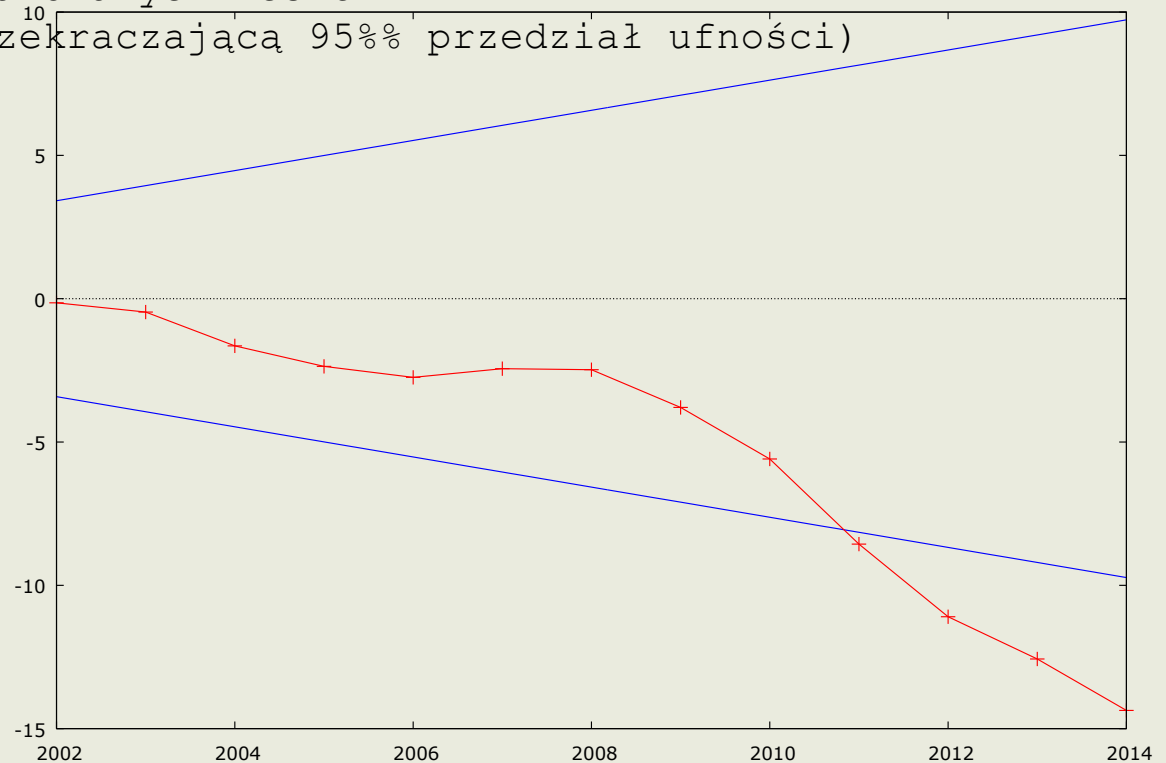
średnia arytmetyczna dla przeskalowanych reszt =  $-0,0139915$

odchylenie standardowe dla przeskalowanych reszt =  $0,0126624$

Skumulowana suma przeskalowanych reszt Wykres CUSUM z 95% przedziałem ufności

('\*' oznacza wartość przekraczającą 95% przedział ufności)

2002	-0,144	
2003	-0,466	
2004	-1,646	
2005	-2,361	
2006	-2,745	
2007	-2,442	
2008	-2,476	
2009	-3,794	
2010	-5,592	
2011	-8,563	*
2012	-11,095	*
2013	-12,569	*
2014	-14,364	*



Statystyka testu Harvey'a-Colliera  $t(12) = -3,98399$  z wartością  $p = 0,001813$

# Modele tendencji rozwojowej

---

W szeregu w którym:

$$y_t = T_t + \varepsilon_t$$

Trend może mieć postać:

— addytywną

$$y_t = f(t) + \varepsilon_t$$

— multiplikatywną

$$y_t = f(t) \cdot \varepsilon_t$$

— gdzie  $f(t)$  – funkcja czasu, charakteryzująca tendencję rozwojową szeregu – funkcja trendu

Wyznaczenie funkcji trendu  $f(t)$  – wygładzenie szeregu prognozowanego

# Modele tendencji rozwojowej

---

- Modele liniowe

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t$$

# Postać analityczna funkcji trendu – zjawiska rosnące coraz szybciej

---

- Modele wykładnicze – dla zjawisk o coraz szybszym tempie wzrostu

$$y_t = e^{\beta_0 + \beta_1 t + \varepsilon_t} \quad , \quad \beta_1 > 0$$

$$y_t = \beta_0 \beta_1^t \quad , \quad \beta_1 > 1$$

- Modele potęgowe – dla zjawisk które w układzie współrzędnych logarytmicznych wykazują przebieg liniowy

$$y_t = e^{\beta_0} t^{\beta_1} e^{\varepsilon_t} \quad , \quad \beta_1 > 1$$

- Wielomiany – elastyczne, posiadają kilka parametrów

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3 + \dots + \varepsilon_t$$

# Postać analityczna funkcji trendu – zjawiska rosnące coraz wolniej

---

- Model logarytmiczny

$$y_t = \beta_0 + \beta_1 \ln t + \varepsilon_t, \beta_1 > 0$$

- Model potęgowy

$$y_t = e^{\beta_0} t^{\beta_1} e^{\varepsilon_t}, 0 < \beta_1 < 1$$

- Wielomian (drugiego stopnia)

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \varepsilon_t, \beta_2 < 0$$

- Wielomian (drugiego stopnia) „odwrotnościowy”

$$y_t = \beta_0 + \beta_1 \frac{1}{t} + \beta_2 \frac{1}{t^2} + \varepsilon_t, \beta_2 < 0$$

- Prognozowanie polega na ekstrapolacji funkcji trendu



$$GVC\_position = e^{\beta_0 + \beta_1 t + \varepsilon_t}$$

Model 8: Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)

Zmienna zależna (Y): l\_GVC\_position

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	-3,14512	0,134803	-23,33	2,29e-011	***
time	0,225022	0,0387698	5,804	8,42e-05	***
time2	-0,0185494	0,00235626	-7,872	4,43e-06	***

$$\ln GVC\_position = -3,145 + 0,225t - 0,0185t^2 + \hat{\varepsilon}_t$$

Dla 95% przedziału ufności,  $t(9, 0,025) = 2,262$

	l_GVC_position	prognoza	błąd ex ante	95% przedział ufności
2012	-3,569285	-3,392239	0,209911	-3,867092 - -2,917387
2013	-3,444272	-3,686440	0,250669	-4,253493 - -3,119386
2014	-3,943307	-4,020363	0,303235	-4,706328 - -3,334398

# Przykład

---

$$\ln GVC\_position^P_{2012} = -3,392239$$

$$\ln GVC\_position^P_{2013} = -3,686440$$

$$\ln GVC\_position^P_{2014} = -4,020363$$



$$GVC\_position^P_{2012} = e^{-3,392239} = 0.033633$$

$$GVC\_position^P_{2013} = e^{-3,686440} = 0.025061$$

$$GVC\_position^P_{2014} = e^{-4,020363} = 0.017946$$

Uwaga na - Średni błąd predykcji ex-ante

Empiryczne i wyrównane wartości zmiennej: GVC\_position



Empiryczne i wyrównane wartości zmiennej: GVC\_position



# Test dla dwóch wariancji

---

Test F – do oceny istotności spadku wariancji przy zmianie stopnia wielomianu

$$H_0 : \sigma_1^2 = \sigma_2^2,$$

$$H_A : \sigma_1^2 > \sigma_2^2$$

gdzie:  $\sigma_1^2$  - wariancja wielomianu niższego stopnia,  $\sigma_2^2$  - wariancja wielomianu wyższego stopnia

$$F = \frac{\sigma_1^2}{\sigma_2^2}$$

Statystyka F ma rozkład F z liczbą stopni swobody odpowiednio  $n - (k_1 + 1)$  z modelu pierwszego i  $n - (k_2 + 1)$  z modelu drugiego

$$n - (k_1 + 1) > n - (k_2 + 1)$$

Model 4: Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)

Zmienna zależna (Y): GVC\_position

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	0,0868497	0,0100743	8,621	9,77e-07	***
time	-0,00317799	0,00110803	-2,868	0,0132	**

Średn.aryt.zm.zależnej 0,061426    Odch.stand.zm.zależnej 0,022830  
Suma kwadratów reszt 0,004469    Błąd standardowy reszt **0,018541**

Model 5: Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)

Zmienna zależna (Y): GVC\_position

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	0,0454347	0,00823596	5,517	0,0001	***
time	0,0114391	0,00236869	4,829	0,0004	***
time2	-0,000913566	0,000143958	-6,346	3,69e-05	***

Średn.aryt.zm.zależnej 0,061426    Odch.stand.zm.zależnej 0,022830  
Suma kwadratów reszt 0,001026    Błąd standardowy reszt **0,009246**

$$F = \frac{0.018541^2}{0.009246^2} = 4.021$$

$$F_{\alpha}(13, 12, 0.05) = 2.660$$

# Współczynnik Theila

---

Współczynnik Theila służy do obliczenia całkowitego względnego błędu prognozy w przedziale weryfikacji prognoz

$$I^2 = \frac{\sum_{t=n+1}^T (y_t - y_t^P)^2}{\sum_{t=n+1}^T y_t^2}$$

Współczynnik Theila przyjmuje wartość 0 dla prognoz idealnie trafnych, im większe rozbieżności pomiędzy realizacją a prognozą tym większa wartość współczynnika

Współczynnik można zdekomponować na 3 elementy wskazujące na źródła braku trafności prognozy

# Współczynnik Theila – z czego wynikają błędy

---

- Obciążoność prognozy – nieodgadnięcie średniej wartości zmiennej prognozowanej

$$I_1^2 = \frac{(\bar{y} - \bar{y}^P)^2}{\frac{1}{T-n} \sum_{t=n+1}^T y_t^2}$$

- Niedostateczna elastyczność – nieodgadnięcie wahań zmiennej prognozowanej

$$I_2^2 = \frac{(s - s^P)^2}{\frac{1}{T-n} \sum_{t=n+1}^T y_t^2}$$

- Niedostatecznej zgodności prognoz z rzeczywistym kierunkiem zmian zmiennej prognozowanej – nieodgadnięcie kierunków tendencji rozwojowej

$$I_3^2 = \frac{2ss^P(1-r)}{\frac{1}{T-n} \sum_{t=n+1}^T y_t^2}$$

Model 5: Estymacja KMNK, wykorzystane obserwacje 2000-2014 (N = 15)  
 Zmienna zależna (Y): GVC\_position

	współczynnik	błąd standardowy	t-Studenta	wartość p	
const	0,0454347	0,00823596	5,517	0,0001	***
time	0,0114391	0,00236869	4,829	0,0004	***
time2	-0,000913566	0,000143958	-6,346	3,69e-05	***

Program liczący współczynnik Theila

	Dane	Prognozy	$(t-y_t)^2$	$y_t^2$	$(t-\hat{y}_t)^2$	$(t-\hat{y}_t)^2$	$(t-\hat{y}_t)^2 \cdot y_t^2$
▶	0,028176	0,029802	2,64387...	0,00079...	2,82239...	0,00032...	3,03900...
	0,031928	0,012506	0,00037...	0,00101...	2,95066...	6,29377...	4,30938...
	0,019384	-0,00717	0,00070...	0,00037...	5,05805...	0,00035...	0,00013...
			0,00108...	0,00218...	8,29095...	0,00068...	0,00016...

Zenuj    Oblicz    S : 0,00525704505085002  
 Zachowaj "Dane"    S<sup>2</sup>: 0,0151041766703415  
 r: 0,709428587150518  
 Średnia prognoz wygasłych: 0,0117126666666667  
 Średnia prognozowanej: 0,026496  
 Ilość elementów: 3    Współczynnik Theila (I<sup>2</sup>): 0,495642380497735  
 Ilość prognoz wygasłych: 3    60,43 %    1.) I1<sup>2</sup>: 0,29951291002122  
 26,81 %    2.) I2<sup>2</sup>: 0,132889385600291  
 12,76 %    3.) I3<sup>2</sup>: 0,063240084876224  
 Suma składników 1.)+2.)+3.) = 0,495642380497735  
 Pierwiastek ze współczynnika Theila wynosi: 0,704    Przeciętny błąd prognozy wynosi: 70,4 %

vv Zapisz vv

Model	Theil (I)	Przeciętny błąd %	I1	I2	I3
1	0,495642380497735	70,4 %	60,43 %	26,81 %	12,76 %



# Współczynnik Theila – z czego wynikają błędy

---

- Przeciętny błąd prognoz wynosi 70,4% (pierwiastek ze współczynnika Theila)
- Z tego błędu
  - 60,43 % był to błąd wywołany złą oceną średniej zmiennej prognozowanej
  - 26,81% był to błąd wywołany złą oceną wahań zmiennej prognozowanej
  - 12,76 % był to błąd wywołany złą oceną trendu zmiennej prognozowanej

---

Dziękuję za uwagę!  
Aleksandra.Kordalska@pg.edu.pl