

# Kointegracja

- korelacja a związek o charakterze przyczynowo-skutkowym

- korelacja a związek o charakterze przyczynowo-skutkowym

### Przyczynowość w sensie Grangera

Zmienna  $x$  jest przyczyną w sensie Grangera zmiennej  $y$  jeżeli bieżące wartości zmiennej  $y$  można dokładniej prognozować uwzględniając przeszłe wartości zmiennej  $x$  niż bez ich uwzględniania

- Mądrość ludowa głosi, że nisko latające jaskółki są przyczyną deszczu.

- Mądrość ludowa głosi, że nisko latające jaskółki są przyczyną deszczu.
- Czy są one przyczyna w sensie Grangera?

- Mądrość ludowa głosi, że nisko latające jaskółki są przyczyną deszczu.
- Czy są one przyczyna w sensie Grangera?
- Przed deszczem rośnie wilgotność powietrza

- Mądrość ludowa głosi, że nisko latające jaskółki są przyczyną deszczu.
- Czy są one przyczyna w sensie Grangera?
- Przed deszczem rośnie wilgotność powietrza
- Owady latają niżej

- Mądrość ludowa głosi, że nisko latające jaskółki są przyczyną deszczu.
- Czy są one przyczyna w sensie Grangera?
- Przed deszczem rośnie wilgotność powietrza
- Owady latają niżej
- Ptaki latają niżej

# Formalny test przyczynowości

- W modelu ARDL

$$y_t = a(t) + \sum_{i=1}^k \alpha_i y_{t-i} + \sum_{i=1}^l \beta_i x_{t-i} + \varepsilon_i$$

# Formalny test przyczynowości

- W modelu ARDL

$$y_t = a(t) + \sum_{i=1}^k \alpha_i y_{t-i} + \sum_{i=1}^l \beta_i x_{t-i} + \varepsilon_i$$

- Testem Grangera będzie weryfikacja hipotezy  $\beta = 0$ .

# Formalny test przyczynowości

- W modelu ARDL

$$y_t = a(t) + \sum_{i=1}^k \alpha_i y_{t-i} + \sum_{i=1}^l \beta_i x_{t-i} + \varepsilon_i$$

- Testem Grangera będzie weryfikacja hipotezy  $\beta = 0$ .
- Jeżeli nie można jej odrzucić, to mówimy, że  $x$  nie jest przyczyną w sensie Grangera zmiennej  $y$

# Formalny test przyczynowości

- W modelu ARDL

$$y_t = a(t) + \sum_{i=1}^k \alpha_i y_{t-i} + \sum_{i=1}^l \beta_i x_{t-i} + \varepsilon_i$$

- Testem Grangera będzie weryfikacja hipotezy  $\beta = 0$ .
- Jeżeli nie można jej odrzucić, to mówimy, że  $x$  nie jest przyczyną w sensie Grangera zmiennej  $y$
- Jej fałszywość świadczy o występowaniu przyczynowości

Source	SS	df	MS
Model	2096.93501	5	419.387001
Residual	11.6377918	124	.09385316
Total	2108.5728	129	16.3455256

Number of obs =	130
F( 5, 124) =	4468.54
Prob > F =	0.0000
R-squared =	0.9945
Adj R-squared =	0.9943
Root MSE =	.30635

inflacja	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
inflacja						
L1.	1.19583	.0831751	14.38	0.000	1.031203	1.360457
L2.	-.2241438	.0810825	-2.76	0.007	-.3846289	-.0636587
stopa						
L1.	-.1837957	.1194775	-1.54	0.127	-.4202751	.0526838
L2.	.1501882	.2230509	0.67	0.502	-.291292	.5916684
L3.	.0162387	.1260694	0.13	0.898	-.233288	.2657655
_cons	.3391173	.1753408	1.93	0.055	-.0079313	.6861659

Source	SS	df	MS
Model	2096.93501	5	419.387001
Residual	11.6377918	124	.09385316
Total	2108.5728	129	16.3455256

Number of obs =	130
F( 5, 124) =	4468.54
Prob > F =	0.0000
R-squared =	0.9945
Adj R-squared =	0.9943
Root MSE =	.30635

inflacja	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
inflacja					
L1.	1.19583	.0831751	14.38	0.000	1.031203 1.360457
L2.	-.2241438	.0810825	-2.76	0.007	-.3846289 -.0636587
stopa					
L1.	-.1837957	.1194775	-1.54	0.127	-.4202751 .0526838
L2.	.1501882	.2230509	0.67	0.502	-.291292 .5916684
L3.	.0162387	.1260694	0.13	0.898	-.233288 .2657655
_cons	.3391173	.1753408	1.93	0.055	-.0079313 .6861659

- Czy rzeczywiście nie ma przyczynowości?

```
. test stopa l1.stopa l2.stopa
```

```
( 1) stopa = 0
```

```
( 2) L.stopa = 0
```

```
( 3) L2.stopa = 0
```

```
F( 3, 124) = 3.45  
Prob > F = 0.0186
```

```
. test stopa l1.stopa l2.stopa
```

```
( 1) stopa = 0
```

```
( 2) L.stopa = 0
```

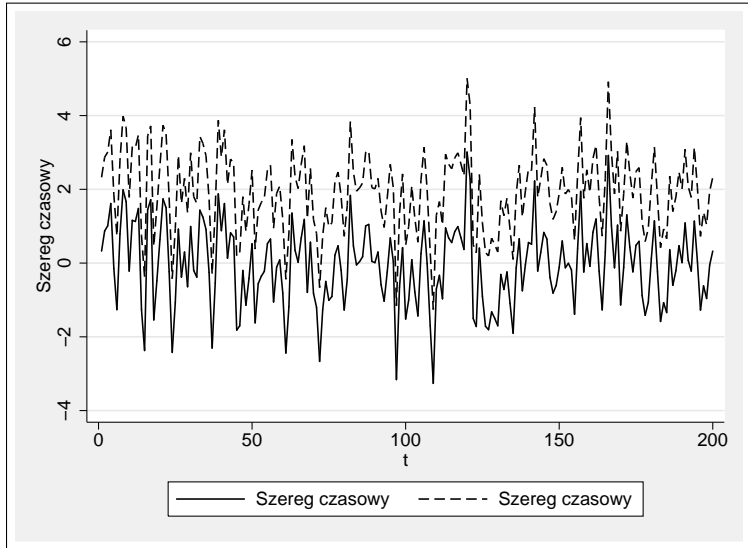
```
( 3) L2.stopa = 0
```

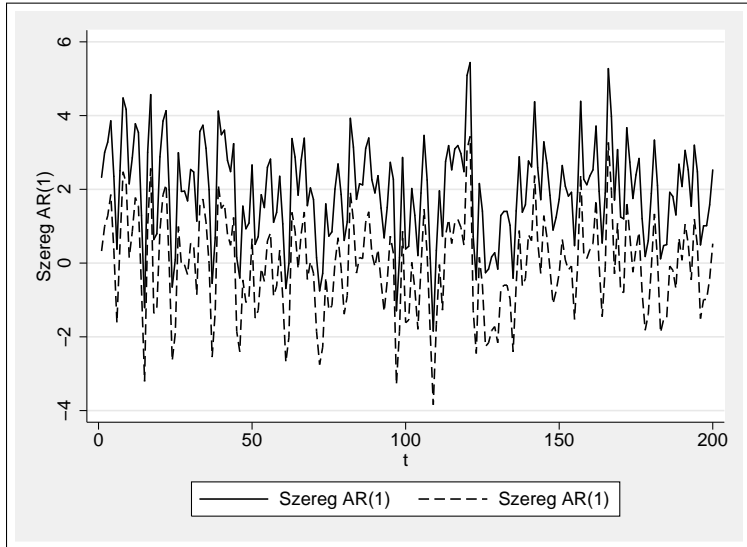
```
F( 3, 124) = 3.45  
Prob > F = 0.0186
```

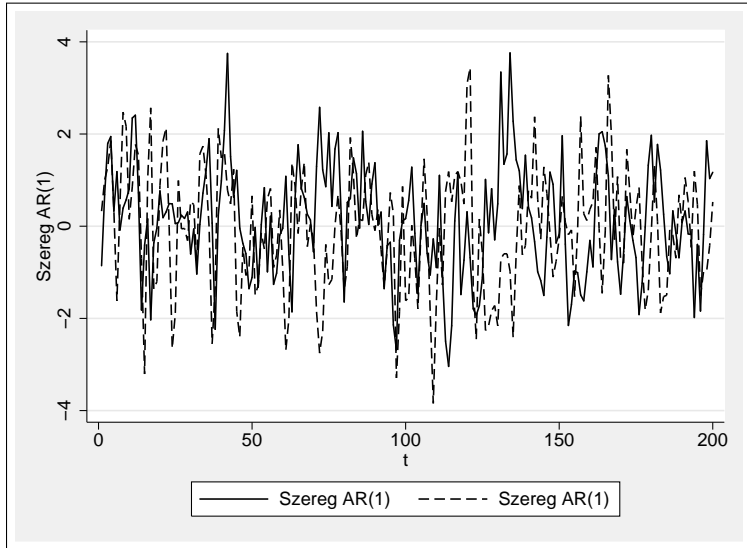
- Zatem należy uznać stopę bezrobocia za przyczynę inflacji w sensie Grangera

- Kointegracja dotyczy zmiennych zintegrowanych, o stopniu integracji większym niż 0.

- Kointegracja dotyczy zmiennych zintegrowanych, o stopniu integracji większym niż 0.
- jeżeli dwie zmienne niestacjonarne kształtują się w kolejnych okresach w sposób podobny, tzn. że zmiany ich wartości są w stałej relacji to **mogą być** skointegrowane.







## Kointegracja

O procesach stochastycznych mówimy, że są skointegrowane rzędu  $(d, b)$  jeżeli oba są zintegrowane rzędu  $d$ , oraz istnieje ich kombinacja liniowa rzędu  $d - b$

## Kointegracja

O procesach stochastycznych mówimy, że są skointegrowane rzędu  $(d, b)$  jeżeli oba są zintegrowane rzędu  $d$ , oraz istnieje ich kombinacja liniowa rzędu  $d - b$

- Kointegracja oznacza istnienie długookresowego związku

## Kointegracja

O procesach stochastycznych mówimy, że są skointegrowane rzędu  $(d, b)$  jeżeli oba są zintegrowane rzędu  $d$ , oraz istnieje ich kombinacja liniowa rzędu  $d - b$

- Kointegracja oznacza istnienie długookresowego związku

## Wektor kointegrujący

Współczynniki stacjonarnej kombinacji liniowej procesów niestacjonarnych nazywamy wektorem kointegrującym

## Kointegracja

O procesach stochastycznych mówimy, że są skointegrowane rzędu  $(d, b)$  jeżeli oba są zintegrowane rzędu  $d$ , oraz istnieje ich kombinacja liniowa rzędu  $d - b$

- Kointegracja oznacza istnienie długookresowego związku

## Wektor kointegrujący

Współczynniki stacjonarnej kombinacji liniowej procesów niestacjonarnych nazywamy wektorem kointegrującym

- W praktyce dwie zmienne zawierające trend mogą być skointegrowane

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- $x_t$  jest  $I(1)$  z definicji

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- $x_t$  jest  $I(1)$  z definicji
- $y_t$  jako funkcja  $x_t$  jest niestacjonarna, ale

$$y_t - \beta x_t = \alpha + \varepsilon_t \tag{1}$$

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID (0, \sigma_u^2)$$

- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID (0, \sigma_\varepsilon^2)$$

- $x_t$  jest  $I(1)$  z definicji
- $y_t$  jako funkcja  $x_t$  jest niestacjonarna, ale

$$y_t - \beta x_t = \alpha + \varepsilon_t \tag{1}$$

- Prawa strona (1) jest stacjonarna

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- $x_t$  jest  $I(1)$  z definicji
- $y_t$  jako funkcja  $x_t$  jest niestacjonarna, ale

$$y_t - \beta x_t = \alpha + \varepsilon_t \tag{1}$$

- Prawa strona (1) jest stacjonarna
- Zatem lewa strona (1) musi być stacjonarna

- Niech  $x_t$  będzie błędzeniem przypadkowym

$$x_t = x_{t-1} + u_t \quad u_t \sim IID(0, \sigma_u^2)$$

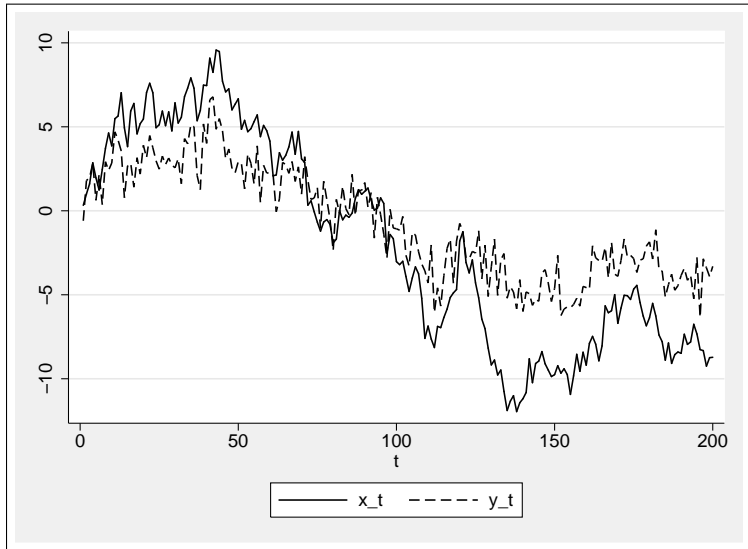
- Niech  $y_t$  będzie zdefiniowane następująco

$$y_t = \alpha + \beta x_t + \varepsilon_t \quad \varepsilon_t \sim IID(0, \sigma_\varepsilon^2)$$

- $x_t$  jest  $I(1)$  z definicji
- $y_t$  jako funkcja  $x_t$  jest niestacjonarna, ale

$$y_t - \beta x_t = \alpha + \varepsilon_t \tag{1}$$

- Prawa strona (1) jest stacjonarna
- Zatem lewa strona (1) musi być stacjonarna
- Więc  $[1, -\beta]$  jest wektorem kointegrującym



- Procedura testowa jest analogiczna do sposobu testowania stopnia integracji procesu stochastycznego

- Procedura testowa jest analogiczna do sposobu testowania stopnia integracji procesu stochastycznego
- Z reguły wykorzystuje się dwustopniową procedurę Engla-Grangera

- Procedura testowa jest analogiczna do sposobu testowania stopnia integracji procesu stochastycznego
- Z reguły wykorzystuje się dwustopniową procedurę Engla-Grangera
- W pierwszym kroku badamy stopień integracji zmiennych

- Procedura testowa jest analogiczna do sposobu testowania stopnia integracji procesu stochastycznego
- Z reguły wykorzystuje się dwustopniową procedurę Engla-Grangera
- W pierwszym kroku badamy stopień integracji zmiennych
- Stopień integracji zmiennej zależnej nie może być wyższy niż stopień integracji którejkolwiek zmiennej objaśniającej

- Procedura testowa jest analogiczna do sposobu testowania stopnia integracji procesu stochastycznego
- Z regułą wykorzystuje się dwustopniową procedurę Engla-Grangera
- W pierwszym kroku badamy stopień integracji zmiennych
- Stopień integracji zmiennej zależnej nie może być wyższy niż stopień integracji którejkolwiek zmiennej objaśniającej
- Procedura wykorzystuje fakt, że nawet jeśli zmienne są zintegrowane to estymator MNK jest zgodny.

- W drugim kroku szacujemy wektor kointegrujący

$$y_t = x_t\beta + u_t \quad (2)$$

- W drugim kroku szacujemy wektor kointegrujący

$$y_t = x_t\beta + u_t \quad (2)$$

- Jeżeli  $y_t - x_t\beta$  jest stacjonarne to  $u_t$  musi być stacjonarne

- W drugim kroku szacujemy wektor kointegrujący

$$y_t = x_t\beta + u_t \quad (2)$$

- Jeżeli  $y_t - x_t\beta$  jest stacjonarne to  $u_t$  musi być stacjonarne
- Test kointegracji jest testem stacjonarności reszt  $u_t$

- W drugim kroku szacujemy wektor kointegrujący

$$y_t = x_t\beta + u_t \quad (2)$$

- Jeżeli  $y_t - x_t\beta$  jest stacjonarne to  $u_t$  musi być stacjonarne
- Test kointegracji jest testem stacjonarności reszt  $u_t$
- Regresja pomocnicza testu ADF ma postać

$$\Delta u_t = \rho u_{t-1} + \xi_t$$

- $H_0 : \rho = 0$  odpowiada niestacjonarności reszt co oznacza **brak kointegracji** między zmiennymi

- $H_0 : \rho = 0$  odpowiada niestacjonarności reszt co oznacza **brak kointegracji** między zmiennymi
- Rozkład statystyki testowej zależy od liczby zmiennych w równaniu (2)

- $H_0 : \rho = 0$  odpowiada niestacjonarności reszt co oznacza **brak kointegracji** między zmiennymi
- Rozkład statystyki testowej zależy od liczby zmiennych w równaniu (2)
- Problem autokorelacji

- $H_0 : \rho = 0$  odpowiada niestacjonarności reszt co oznacza **brak kointegracji** między zmiennymi
- Rozkład statystyki testowej zależy od liczby zmiennych w równaniu (2)
- Problem autokorelacji
- Problem trendu

- Kointegracja jest sposobem opisu równowagi długookresowej

- Kointegracja jest sposobem opisu równowagi długookresowej

### Twierdzenie Grangera o reprezentacji

Jeżeli zmienne losowe  $x_t, y_t$  są  $I(1)$  oraz są skointegrowane z wektorem kointegrującym  $[1, -\beta]$  wówczas zależność można przedstawić za pomocą Mechanizmu Korekty Błędem (ECM)

$$\Delta y_t = \sum_{i=0}^{K-1} \gamma_i \Delta x_{t-i} + \sum_{i=0}^{L-1} \theta_i \Delta y_{t-i} + \alpha(y_{t-1} - \beta x_{t-1}) + \varepsilon_t$$

gdzie  $\varepsilon_t \sim I(0)$

- Twierdzenie odwrotne jest również prawdziwe:

*Jeżeli zmienne są opisane mechanizmem korekty błędem to są skointegrowane*

- Twierdzenie odwrotne jest również prawdziwe:  
*Jeżeli zmienne są opisane mechanizmem korekty błędem to są skointegrowane*
- Twierdzenie Grangera o reprezentacji pozwala na ekonomiczną interpretację parametrów

- W równowadze  $y^* = E(y_t) = E(y_{t-1}) = \dots = E(y_{t-k})$ .  
Wobec tego

$$E(\Delta y_t) = E(y_t - y_{t-1}) = E(y_t) - E(y_{t-1}) = E(y_t) - E(y_t) = 0$$

- W równowadze  $y^* = E(y_t) = E(y_{t-1}) = \dots = E(y_{t-k})$ .  
Wobec tego

$$E(\Delta y_t) = E(y_t - y_{t-1}) = E(y_t) - E(y_{t-1}) == E(y_t) - E(y_t) = 0$$

- Podobna zależność zachodzi dla zmiennych egzogenicznych

$$x^* = E(x_t) = E(x_{t-1}) = \dots = E(x_{t-k}), \text{ więc } E(\Delta x_t) = 0$$

- W równowadze  $y^* = E(y_t) = E(y_{t-1}) = \dots = E(y_{t-k})$ .  
Wobec tego

$$E(\Delta y_t) = E(y_t - y_{t-1}) = E(y_t) - E(y_{t-1}) = E(y_t) - E(y_t) = 0$$

- Podobna zależność zachodzi dla zmiennych egzogenicznych

$$x^* = E(x_t) = E(x_{t-1}) = \dots = E(x_{t-k}), \text{ więc } E(\Delta x_t) = 0$$

- Stosując operator wartości oczekiwanej do postaci ECM otrzymujemy

$$0 = \alpha(y^* - x^*\beta)$$

- W równowadze  $y^* = E(y_t) = E(y_{t-1}) = \dots = E(y_{t-k})$ .  
Wobec tego

$$E(\Delta y_t) = E(y_t - y_{t-1}) = E(y_t) - E(y_{t-1}) == E(y_t) - E(y_t) = 0$$

- Podobna zależność zachodzi dla zmiennych egzogenicznych

$$x^* = E(x_t) = E(x_{t-1}) = \dots = E(x_{t-k}), \text{ więc } E(\Delta x_t) = 0$$

- Stosując operator wartości oczekiwanej do postaci ECM otrzymujemy

$$0 = \alpha(y^* - x^*\beta)$$

- Zatem nawias po prawej stronie zawiera warunek równowagi długookresowej

- W model o postaci ECM jest wpisana długookresowa równowaga

- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie  $(y - x\beta)$  to odchylenie od równowagi długookresowej

- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie  $(y - x\beta)$  to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją

- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie  $(y - x\beta)$  to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją
  - kształtowania się zjawiska w przeszłości

- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie  $(y - x\beta)$  to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją
  - kształtowania się zjawiska w przeszłości
  - wartości zmiennych egzogenicznych

- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie ( $y - x\beta$ ) to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją
  - kształtowania się zjawiska w przeszłości
  - wartości zmiennych egzogenicznych
  - odchylenia od równowagi długookresowej

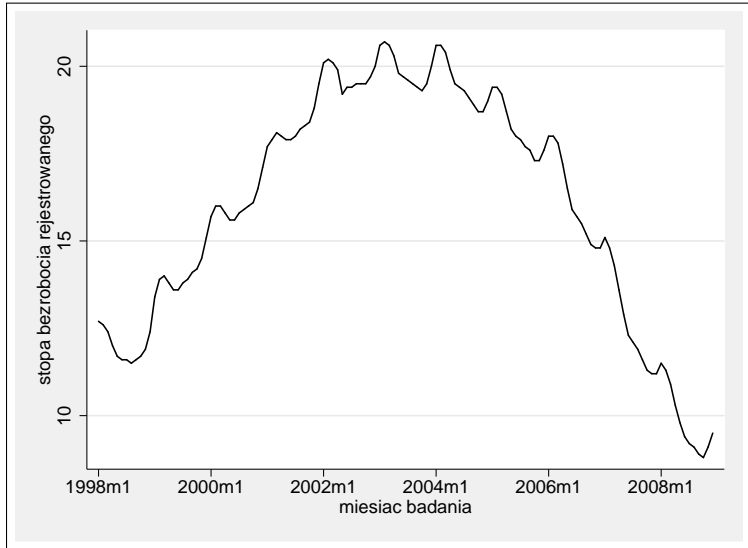
- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie ( $y - x\beta$ ) to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją
  - kształtowania się zjawiska w przeszłości
  - wartości zmiennych egzogenicznych
  - odchylenia od równowagi długookresowej
- Znak parametru  $\alpha$  informuje o kierunku odchylenia od równowagi długookresowej

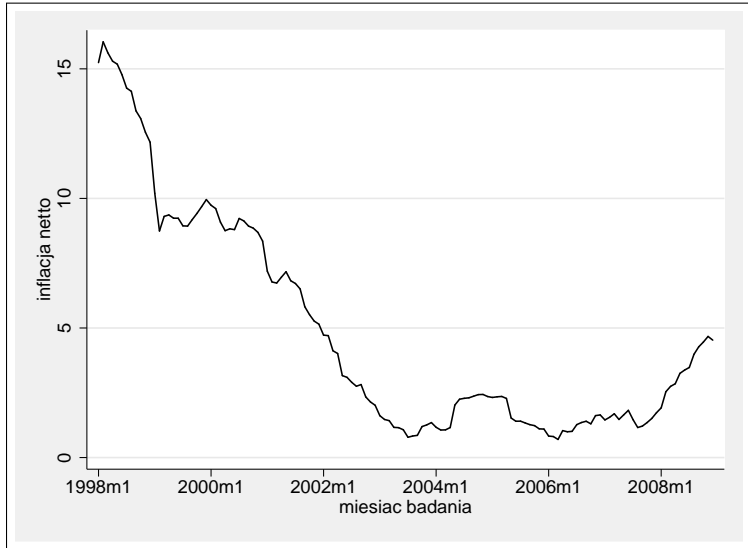
- W model o postaci ECM jest wpisana długookresowa równowaga
- W krótkim okresie ( $y - x\beta$ ) to odchylenie od równowagi długookresowej
- Zmiana poziomu zjawiska  $\Delta y_t$  jest funkcją
  - kształtowania się zjawiska w przeszłości
  - wartości zmiennych egzogenicznych
  - odchylenia od równowagi długookresowej
- Znak parametru  $\alpha$  informuje o kierunku odchylenia od równowagi długookresowej
- Wielkość parametru  $\alpha$  informuje o części odchylenia od równowagi długookresowej korygowanej podczas pierwszego okresu

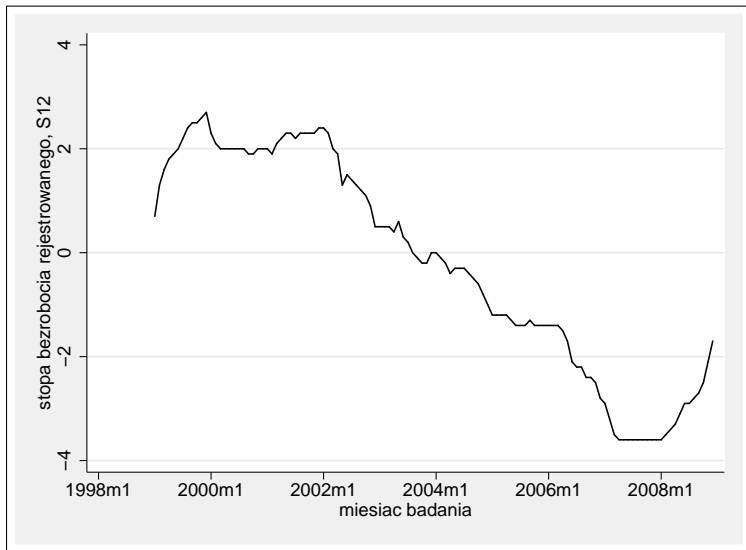
- Badamy zależność między inflacją a stopą bezrobocia

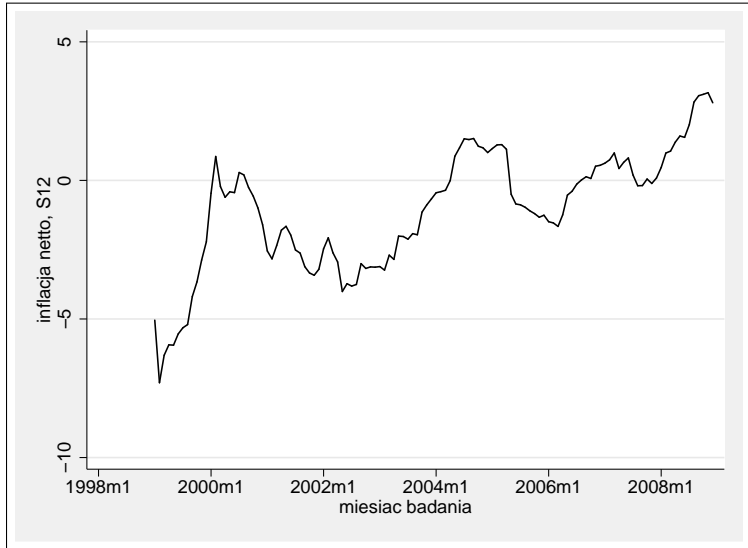
- Badamy zależność między inflacją a stopą bezrobocia
- W pierwszym kroku sprawdzimy stacjonarność zmiennych

- Badamy zależność między inflacją a stopą bezrobocia
- W pierwszym kroku sprawdzimy stacjonarność zmiennych
- Jeżeli okażą się one być niestacjonarne oszacujemy parametry modelu równowagi długookresowej









```
. dfuller s12.stopa
Dickey-Fuller test for unit root                Number of obs   =           119
```

```
----- Interpolated Dickey-Fuller -----
          Test          1% Critical      5% Critical      10% Critical
          Statistic      Value           Value           Value
-----
Z(t)          -0.246          -3.504          -2.889          -2.579
-----
```

MacKinnon approximate p-value for Z(t) = 0.9329

```
. dfuller s12.inflacja
Dickey-Fuller test for unit root                Number of obs   =           119
```

```
----- Interpolated Dickey-Fuller -----
          Test          1% Critical      5% Critical      10% Critical
          Statistic      Value           Value           Value
-----
Z(t)          -1.336          -3.504          -2.889          -2.579
-----
```

MacKinnon approximate p-value for Z(t) = 0.6128

```
. reg s12.inflacja s12.stopa
```

Source	SS	df	MS	Number of obs	=	120
Model	240.973663	1	240.973663	F( 1, 118)	=	91.65
Residual	310.271352	118	2.62941824	Prob > F	=	0.0000
				R-squared	=	0.4371
				Adj R-squared	=	0.4324
Total	551.245015	119	4.63231105	Root MSE	=	1.6215

S12.inflacja	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
stopa					
S12.	-.686349	.0716952	-9.57	0.000	-.828325    -.544373
_cons	-1.228575	.148833	-8.25	0.000	-1.523305    -.9338454

```
predict u, resid
```

Source	SS	df	MS	Number of obs =	119
Model	1.01987416	1	1.01987416	F( 1, 117) =	3.83
Residual	31.1406747	117	.266159613	Prob > F =	0.0527
Total	32.1605489	118	.272547024	R-squared =	0.0317
				Adj R-squared =	0.0234
				Root MSE =	.51591

D.u	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
u						
L1.	-.0581122	.0296869	-1.96	0.053	-.1169057	.0006812
_cons	.05068	.0472985	1.07	0.286	-.0429921	.1443522

## Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
1	6.694	1	0.0097
2	7.785	2	0.0204
3	8.073	3	0.0445
4	8.138	4	0.0866
5	10.536	5	0.0614

Source	SS	df	MS
Model	4.52212342	4	1.13053085
Residual	22.5209061	111	.202891046
Total	27.0430295	115	.235156778

Number of obs	=	116
F( 4, 111)	=	5.57
Prob > F	=	0.0004
R-squared	=	0.1672
Adj R-squared	=	0.1372
Root MSE	=	.45043

D.u	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
u					
L1.	-.0873195	.0299708	-2.91	0.004	-.1467087 - .0279302
LD.	.3159623	.0916903	3.45	0.001	.1342718 .4976528
L2D.	.0728535	.0917579	0.79	0.429	-.1089708 .2546779
L3D.	.0593006	.0843714	0.70	0.484	-.1078869 .2264881
_cons	.0351022	.0424318	0.83	0.410	-.0489792 .1191836

```
. reg s12.inflacja s12.stopa l.u l(1/2).s12.inflacja
```

Source	SS	df	MS	Number of obs =	118
Model	474.938384	4	118.734596	F( 4, 113) =	637.14
Residual	21.0582306	113	.186356023	Prob > F	= 0.0000
Total	495.996614	117	4.2392873	R-squared	= 0.9575
				Adj R-squared	= 0.9560
				Root MSE	= .43169

S12.inflacja	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
stopa					
S12.	.4754522	.259508	1.83	0.070	-.0386799 .9895844
u					
L1.	-.7665232	.3771054	-2.03	0.044	-1.513637 -.0194093
inflacja					
LS12.	1.956976	.3942372	4.96	0.000	1.175921 2.738031
L2S12.	-.2792319	.0796877	-3.50	0.001	-.4371076 -.1213562
_cons	.9134467	.4721089	1.93	0.056	-.0218862 1.84878