



AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA W KRAKOWIE

Wydział Zarządzania

Projekt EFiD -Projekt 7B

Autorzy: *Patrycja Piła, Mateusz Strojek, Julia Szutka, Magdalena Wnuk*

Kierunek studiów: *Informatyka i Ekonometria*

Prowadzący: *dr hab. Tomasz Wójtowicz*

Kraków, 2025

Wstęp

Wersja 3:

Celem przeprowadzonego badania było sprawdzenie, jak nieuwzględnienie autokorelacji składników losowych wpływa na wyniki testu przyczynowości Grangera w modelu VAR(1):

$$X_t = \alpha_1 + \beta_1 X_{t-1} + \delta_1 Y_{t-1} + \varepsilon_t$$

$$Y_t = \alpha_2 + \beta_2 X_{t-1} + \delta_2 Y_{t-1} + \eta_t$$

Za pomocą symulacji zbadano, w jakim stopniu naruszenie założenia o braku korelacji serialnej reszt prowadzi do zniekształcenia rozmiaru testu (nadmiernego odrzucania hipotezy zerowej). W badaniu oceniono również skuteczność korekty macierzy kowariancji za pomocą dwóch wariantów estymatorów HAC. Analizę przeprowadzono w przekroju różnych długości próby ($N=100$ oraz $N=1000$), zróżnicowanej siły autokorelacji procesu błędu losowego ($\varepsilon_t \sim AR(1)$) oraz odmiennej dynamiki własnej badanych szeregów czasowych.

Metodyka badań

1. Generowanie danych

Podstawą symulacji był dwuwymiarowy model wektorowej autoregresji rzędu pierwszego -VAR(1). Proces generujący dane zdefiniowano za pomocą następującego układu równań:

$$\begin{aligned} X_t &= \alpha_2 + \beta_1 \cdot X_{t-1} + \delta_1 \cdot Y_{t-1} + \varepsilon_t \\ Y_t &= \alpha_2 + \beta_2 \cdot X_{t-1} + \delta_2 \cdot Y_{t-1} + \eta_t \end{aligned}$$

gdzie $t = 1, \dots, N$. Kluczowym elementem badania było wprowadzenie autokorelacji do składnika losowego ε_t w pierwszym równaniu. Został on zamodelowany jako proces autoregresyjny AR(1):

$$\varepsilon_t = \rho \cdot \varepsilon_{t-1} + u_t$$

gdzie u_t oraz η_t to niezależne zmienne losowe o rozkładzie normalnym (biały szum). Parametr ρ sterował siłą autokorelacji błędu.

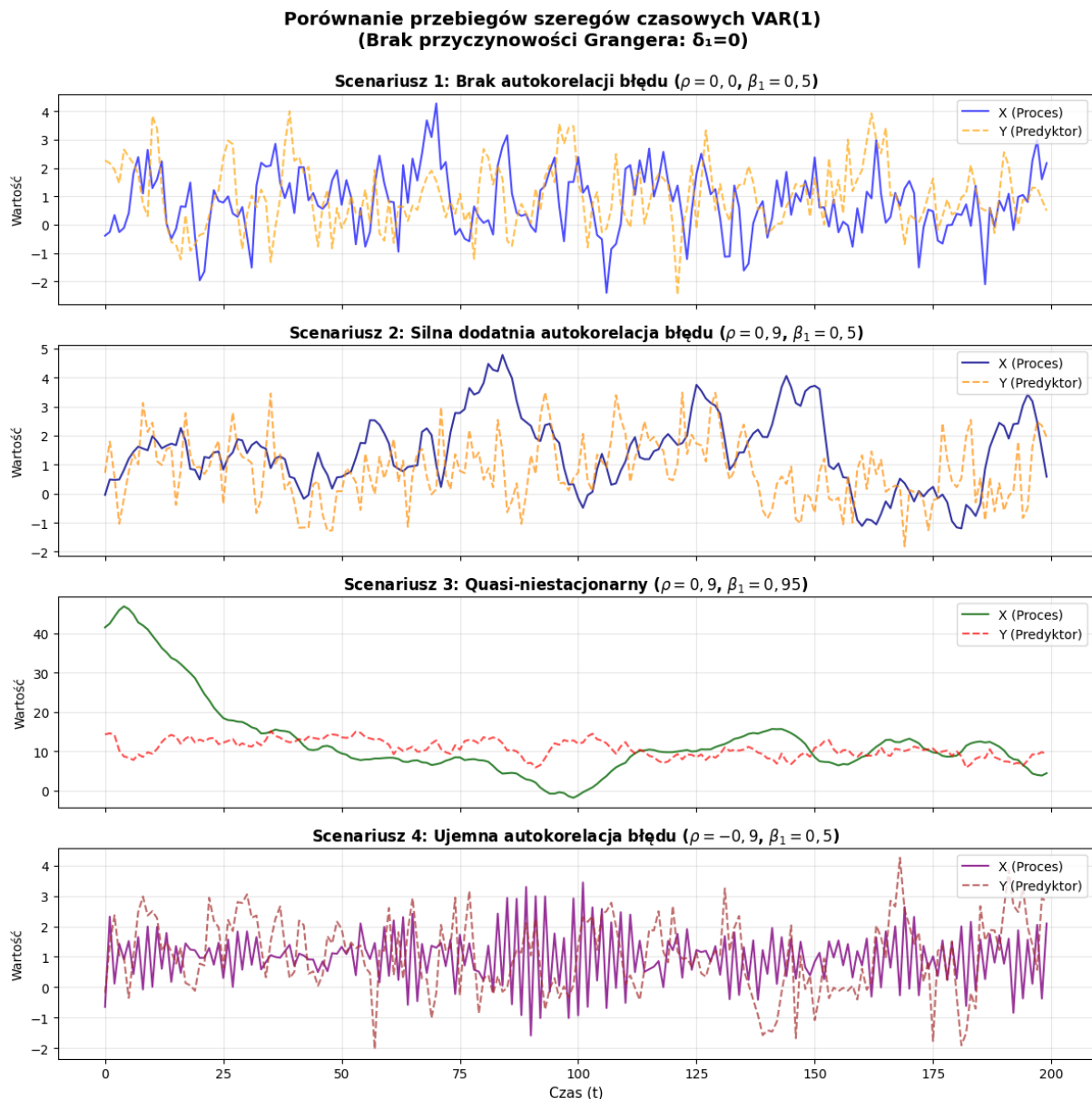
Aby zapewnić porównywalność wyników dla różnych wartości ρ , zastosowano skalowanie wariancji. Wariancja innowacji u_t była dobierana dynamicznie tak, aby bezwarunkowa wariancja procesu ε_t wynosiła zawsze 1, zgodnie ze wzorem:

$$Var(u_t) = 1 - \rho^2.$$

Proces symulacyjny uwzględniał:

- Inicjalizację - symulację rozpoczynano od teoretycznych wartości oczekiwanych procesów X i Y (stan stacjonarny).

- Okres rozbiegowy (Burn-in) - generowano dodatkowe 100 obserwacji początkowych, które następnie odrzucano w celu wyeliminowania wpływu warunków początkowych i ustabilizowania procesu.
- Hipotezę zerową - w każdym wariancie przyjęto $\delta_1 = 0$, co oznacza brak przyczynowości Grangera od zmiennej Y do X. Pozwoliło to na zbadanie rozmiaru testu (prawdopodobieństwa popełnienia błędu I rodzaju).



2. Zastosowane estymatory i procedury testowe

Badanie weryfikowało hipotezę zerową $H_0: \delta_1 = 0$ przy użyciu testu Walda. Estymację przeprowadzono metodą najmniejszych kwadratów dla równania zmiennej X_t . Ze względu na obecność autokorelacji w resztach porównano trzy podejścia do estymacji macierzy kowariancji parametrów:

- Estymator standardowy (OLS) -zakłada brak autokorelacji i homoskedastyczność reszt,
- Estymator HAC (Newey-West) -krótkie okno -estymator odporny na heteroskedastyczność i autokorelację, z ustaloną liczbą opóźnień równą 2,
- Estymator HAC (Newey-West) -długie okno -estymator z liczbą opóźnień dostosowaną do wielkości próby zgodnie z regułą kciuka: $L \approx N^{0,25}$.

Poziom istotności testu ustalono na $\alpha = 0,05$.

3. Procedura symulacyjna

Procedura badawcza obejmowała następujące kroki:

1. Definicja scenariuszy

Aby kompleksowo zbadać wpływ dynamiki szeregów czasowych na właściwości testu, zdefiniowano dziesięć scenariuszy parametrów autoregresyjnych (β_1 dla X_t oraz δ_2 dla Y_t). Scenariusze podzielono na cztery grupy analityczne:

- Symetryczna inercja dodatnia -analizowano procesy o rosnącej sile zależności od przeszłości. Uwzględniono trzy poziomy parametrów: (0,3; 0,3), (0,6; 0,6) oraz (0,9; 0,9).
- Granica niestacjonarności - włączono scenariusz (0,98; 0,98), reprezentujący procesy near-unit-root. Charakteryzują się one bardzo długą pamięcią i stanowią wyzwanie dla standardowych aproksymacji asymptotycznych.
- Ujemna autokorelacja (procesy oscylacyjne) -zbadano przypadki z ujemnymi parametrami autoregresji: (-0,5; -0,5) oraz (-0,8; -0,8).
- Asymetria i mieszane znaki parametrów -rozważono sytuacje, w których zmienne X i Y różnią się dynamiką, na przykład X wolny ($\beta_1 = 0,9$) i Y szybki ($\delta_2 = 0,2$), a także odwrotnie. Uwzględniono też przypadki o przeciwnych znakach parametrów, takie jak (-0,2; 0,2).

2. Parametryzacja autokorelacji składnika losowego

Rozszerzono zakres badania parametru ρ w procesie błędu ε_t . Analizie poddano pięć wartości:

$$\rho \in \{-0,9, -0,5, 0,0, 0,5, 0,9\}$$

Pozwoliło to ocenić zachowanie testu zarówno przy typowej autokorelacji dodatniej, jak i ujemnej, która może prowadzić do odmiennych zniekształceń, np. nadmiernie konserwatywnego działania testu.

3. Stabilizacja procesu (burn-in)

Ze względu na występowanie scenariusza o bardzo wysokiej inercji (0,98), okres rozbiegowy ustawiono na 100 obserwacji. W połączeniu z inicjalizacją od wartości

oczekiwanej procesu pozwoliło to zapewnić, że analizowane próbki pochodzą ze stanu stacjonarnego i są wolne od wpływu warunków początkowych.

4. Iteracja i ocena

Dla każdej kombinacji scenariusza strukturalnego, poziomu autokorelacji ρ oraz długości próby $N \in \{100, 1000\}$ przeprowadzono M replikacji symulacji. W każdej symulacji estymowano model i weryfikowano hipotezę zerową o braku przyczynowości. Zliczano częstość odrzuceń hipotezy dla estymatora OLS oraz estymatorów HAC, co pozwoliło na ocenę stabilności i poprawności testu w różnych warunkach dynamicznych. Wyniki oceniano w oparciu o rozkład Bernoulliego. Dla M symulacji i poziomu istotności α skonstruowano 95% przedział ufności dla frakcji odrzuceń. Wyniki wykraczające poza ten przedział interpretowano jako statystycznie istotne zniekształcenie rozmiaru testu (np. nadmierną skłonność do odrzucania prawdziwej hipotezy zerowej).

Każda pojedyncza symulacja to próba Bernoulliego $B(1, p)$, gdzie p to rzeczywisty rozmiar testu. Suma odrzuceń K w M symulacjach ma rozkład dwumianowy $K \sim B(M, p)$.

Dla dużej liczby symulacji ($M=1000$), z Centralnego Twierdzenia Granicznego (de Moivre-Laplace), frakcja odrzuceń $\hat{p} = \frac{K}{M}$ ma w przybliżeniu rozkład normalny:

$$\hat{p} \sim N\left(\alpha, \frac{\alpha(1-\alpha)}{M}\right)$$

Dla $\alpha = 0,05$ i $M = 1000$, odchylenie standardowe wynosi:

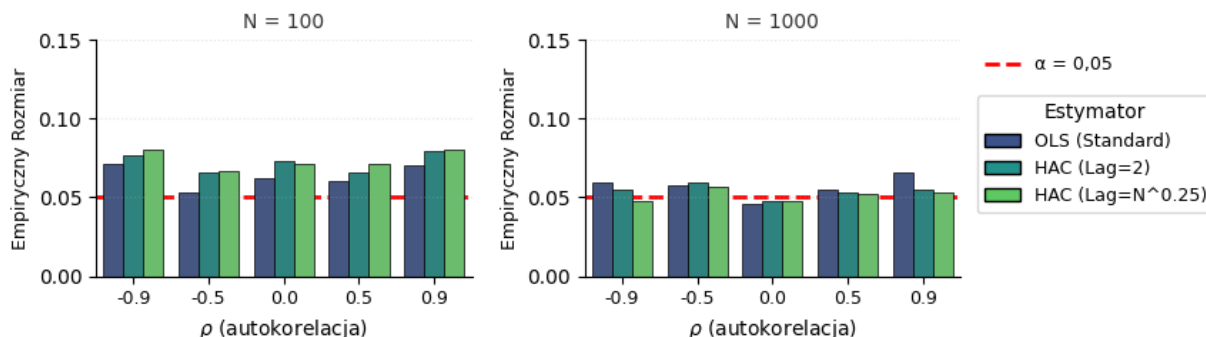
$$SE = \sqrt{\frac{0,05 * 0,95}{1000}} \approx 0,0069$$

Zatem 95% przedział ufności to $0,05 \pm 1,96 * 0,0069 \approx [0,036, 0,064]$

Wyniki symulacji

1. Scenariusz 1 ($\beta_1 = 0,3$; $\delta_2 = 0,3$)

1. $\beta=0,3$, $\delta_2=0,3$

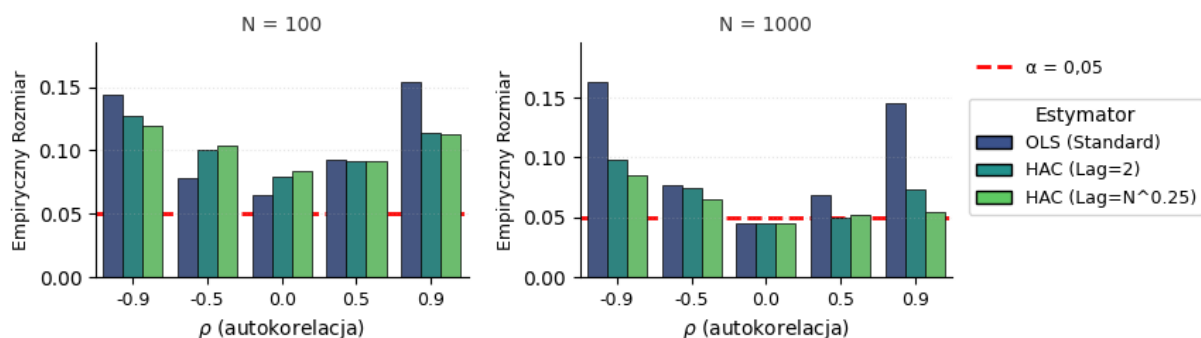


W krótkiej próbie wszystkie estymatory wykazują zawyżony rozmiar testu - częstość odrzucenia prawdziwej hipotezy zerowej przekracza poziom 5% za każdym razem. Efekt ten narasta wraz ze wzrostem autokorelacji błędu: im większe ρ , tym więcej test Grangera odrzuca H_0 . Zarówno standardowy OLS, jak i estymatory HAC nie korygują tego problemu.

W długiej próbie sytuacja stabilizuje się - rozmiar testu dla wszystkich estymatorów jest bliski ustalonemu poziomowi 5%, a wpływ autokorelacji błędu staje się znacznie słabszy. Zarówno OLS, jak i HAC pokazują zbliżone wartości. Przy dużej liczności próby test Grangera działa poprawnie i jest mniej podatny na zaburzenia wynikające z autokorelacji składnika losowego.

2. Scenariusz 2 ($\beta_1 = 0,6$; $\delta_2 = 0,6$)

2. $\beta=0,6$, $\delta_2=0,6$

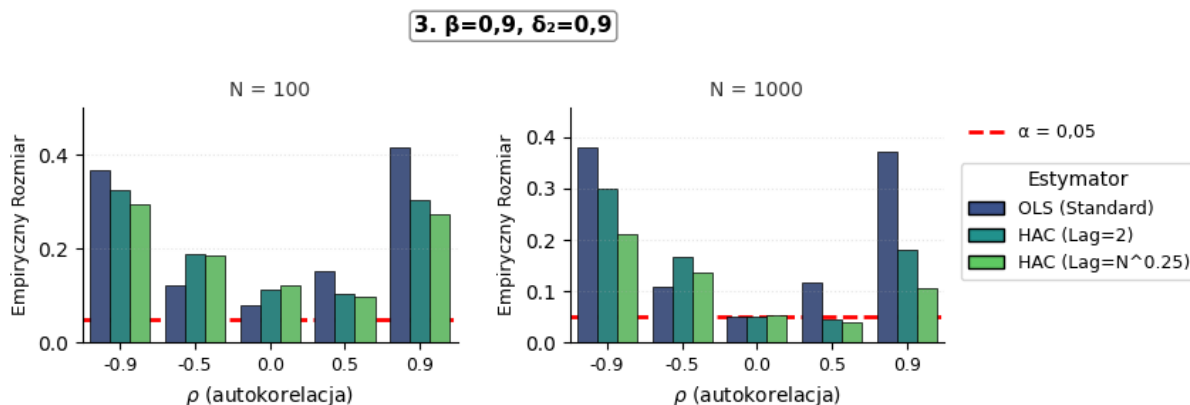


W krótkiej próbie ($N=100$) wszystkie estymatory wykazują silne zaburzenia rozmiaru testu. Standardowy OLS znacząco przekracza poziom 5% przy dużych wartościach autokorelacji błędu ($\rho=0,9$) - rozmiar testu jest trzykrotnie wyższy niż nominalny. Wraz z malejącą siłą autokorelacji OLS maleje. Estymatory HAC zachowują się lepiej, ale również wykazują zauważalne odchylenia od poziomu 5%, zwłaszcza w obszarze dodatniej autokorelacji. Słaba kontrola poziomu istotności wynika z połączenia umiarkowanie silnej dynamiki procesu ($\beta_1=0,6$ i $\delta_2=0,6$) i autokorelacji w błędzie.

W długiej próbie ($N=1000$) estymatory HAC kontrolują rozmiar testu poprawnie w całym zakresie wartości ρ , osiągając wyniki bliskie 0,05 (wyjątek stanowi ujemna autokorelacja). Standardowy OLS nadal wyraźnie zawyża liczbę odrzuceń przy silnej autokorelacji ($\rho=\pm 0,9$), co potwierdza jego wrażliwość na naruszenia założeń.

Większa liczność próby eliminuje większość problemów obserwowanych dla $N=100$, a HAC okazuje się skuteczny w stabilizowaniu rozmiaru testu.

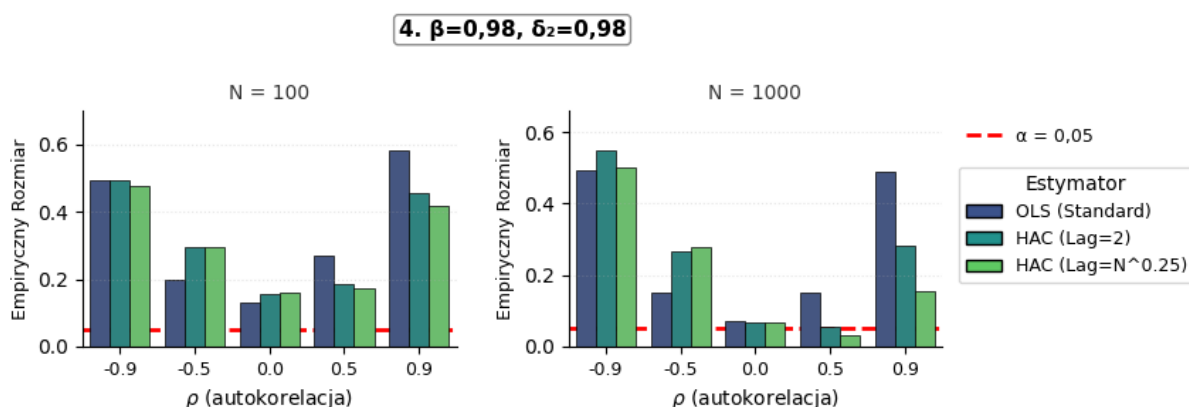
3. Scenariusz 3 ($\beta_1 = 0,9$; $\delta_2 = 0,9$)



W tym scenariuszu dynamika procesu jest bardzo silna, co wyraźnie przekłada się na zachowanie testu. W krótkiej próbie obserwuje się duże zniekształcenie rozmiaru testu - w skrajnych przypadkach wartości empiryczne przekraczają nawet 0,4, co oznacza ponad ośmiokrotne przekroczenie nominalnego poziomu istotności. Standardowy OLS zachowuje się „agresywnie”, natomiast estymatory HAC, mimo poprawy względem OLS, również przekraczają poziom 5% i nie są w stanie skutecznie kontrolować błędu I rodzaju przy tak silnej dynamice.

W długiej próbie ($N=1000$) sytuacja ulega częściowej poprawie, szczególnie przy braku autokorelacji błędów. Dla $\rho = 0$ empiryczny rozmiar testu znajduje się bardzo blisko wartości 0,05, co oznacza, że przyczynowość Grangera jest prawidłowo identyfikowana, gdy proces błędów spełnia klasyczne założenia. Jednak nawet duża liczność próby nie eliminuje problemów w przypadku wysokiej autokorelacji. Przy $\rho = \pm 0,9$ test nadal znacząco zawyża liczbę odrzuceń - zarówno OLS, jak i HAC - choć wartości są niższe niż w próbie $N=100$. Oznacza to, że silna inercja szeregów w połączeniu z autokorelacją błędów prowadzi do trwałego zaburzenia rozmiaru testu, a zwiększenie próby łagodzi problem, lecz go nie usuwa.

4. Scenariusz 4 ($\beta_1 = 0,98$; $\delta_2 = 0,98$)

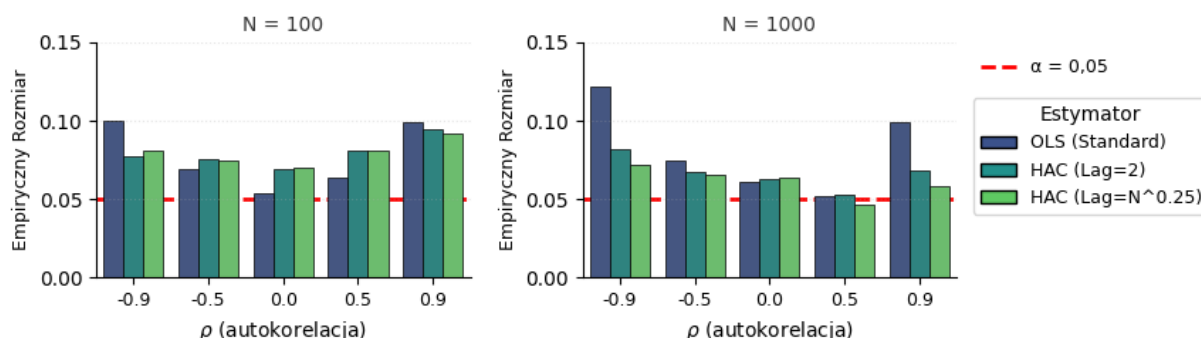


Scenariusz z parametrami β i δ bliskimi jednoŹci jest najbardziej wymagajĄcy, poniewaŹ proces ma charakter quasi-niestacjonarny i wykazuje ekstremalnie silnĄ pamieć. W krótkiej próbie ($N = 100$) prowadzi to do powaŹnych zniekształceń rozmiaru testu Grangera: przy $\rho = -0,9$ odsetek błędnych odrzuceń H_0 przekracza 0,4, a przy dodatniej autokorelacji błędu ($\rho = 0,9$) osiąga wartości od ok. 0,4 do 0,6. Wysoki odsetek błędów I rodzaju wynika z dwóch nakładających się problemów, których nie rozwiązuje zastosowanie estymatorów HAC:

- W modelu dynamicznym (gdzie opóŹniona zmienna X_{t-1} jest regresem), występowanie autokorelacji składnika losowego ε powoduje, Źe regresory stają się skorelowane z błędem losowym (problem endogeniczności). Prowadzi to do tego, Źe estymator OLS staje się obciążony i niezgodny. Estymatory HAC (Newey-West) korygują jedynie macierz kowariancji (błędy standardowe), zakładając, Źe same parametry są oszacowane poprawnie. W tym przypadku, przez złe wyspecyfikowanie modelu (brak odpowiednio wysokich opóŹnień w VAR przy autokorelacji reszt), same współczynniki są błędne, co czyni statystykę Walda bezużyteczną.
- Bliskość parametrów do jednoŹci (0,98) sprawia, Źe standardowe rozkłady asymptotyczne (t-Studenta, normalny) mogą nie być dobrym przybliŹeniem, a statystyki testowe zaczynają podlegać niestandardowym rozkładom (typu Dickeya-Fullera). Zwiększenie próby do $N=1000$ pomaga jedynie w ograniczonym stopniu, poniewaŹ przy tak silnej autokorelacji błędu $\rho = 0,9$ obciążenie estymatora zanika bardzo powoli.

5. Scenariusz 5 ($\beta_1 = -0,5$; $\delta_2 = -0,5$)

5. $\beta=-0,5$, $\delta_2=-0,5$



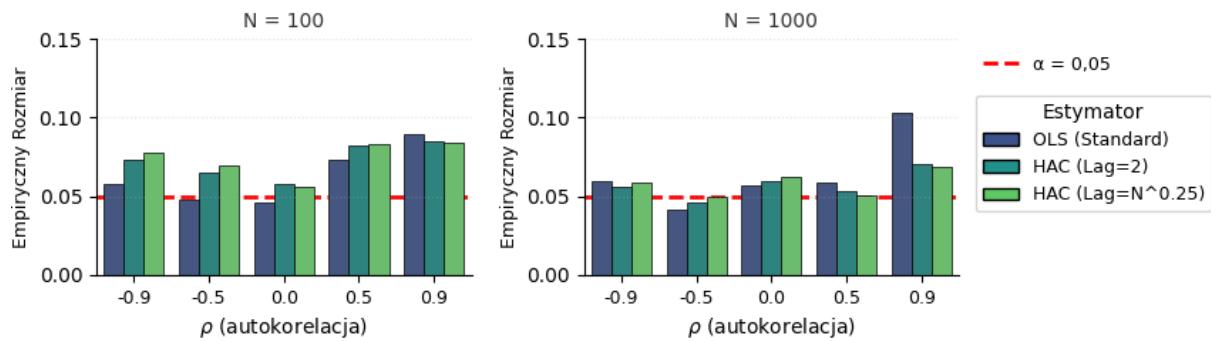
W scenariuszu z ujemnymi parametrami autoregresji szeregi zachowują się w sposób naprzemienny - po dodatniej wartości często pojawia się ujemna i odwrotnie. Proces ma więc skłonność do odbijania się od poprzedniego poziomu, zamiast kontynuować ruch w jednym kierunku. Taka oscylacyjność sprawia, że nie tworzą się długie, pozorne zależności, które mogłyby wprowadzać test w błąd.

W krótkiej próbie ($N = 100$) empiryczny rozmiar testu jest umiarkowanie zawyżony, ale zdecydowanie niższy niż w scenariuszach o dodatniej lub bliskiej jedności inercji. Wyniki są relatywnie stabilne w całym zakresie wartości ρ (od ok. 0,06 do 0,1). Różnice między OLS a HAC pozostają niewielkie. Oznacza to, że ujemna dynamika procesu częściowo „chroni” test Grangera przed nadmiernym odrzucaniem hipotezy zerowej, choć autokorelacja błędu nadal powoduje lekkie zniekształcenia.

W dużej próbie ($N = 1000$) rozmiar testu zbliża się do nominalnego poziomu 5% dla większości estymatorów i wartości ρ . Estymatory HAC zachowują się stabilnie i konsekwentnie kontrolują poziom istotności. OLS nieco przeszacowuje liczbę odrzuceń przy silnej autokorelacji błędu. Ogólnie jednak scenariusz z ujemną dynamiką okazuje się dużo mniej problematyczny niż te z dużą dodatnią autokorelacją (scenariusz 3 i 4) - zarówno rozmiar testu jest bliższy wartościom pożądanym, jak i różnice między estymatorami są mniejsze.

6. Scenariusz 6 (Asymetria: X wolny, Y szybki)

6. Asymetria: X wolny, Y szybki



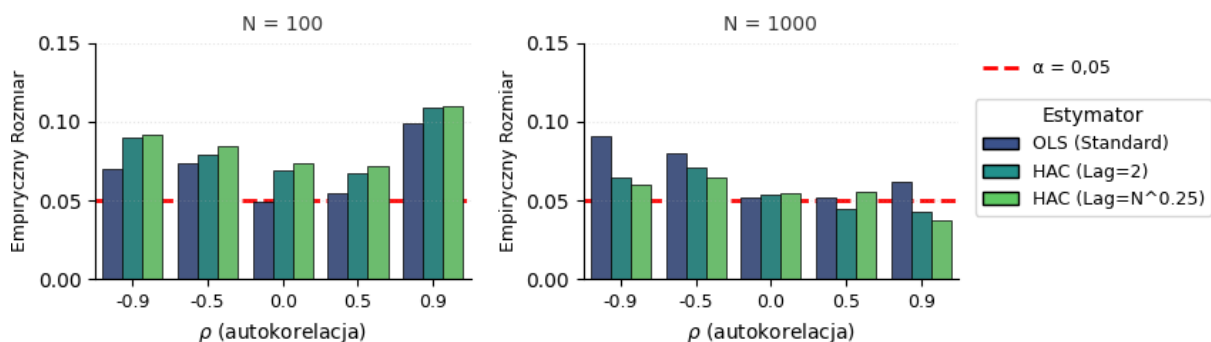
W scenariuszu asymetrycznej dynamiki, w którym X charakteryzuje się wysoką inercją ($\beta_1 = 0,9$), a Y reaguje znacznie szybciej ($\delta_2 = 0,2$), rozmiar testu pozostaje dość bliski wartości nominalnej dla większości estymatorów. Zarówno w próbie $N = 100$, jak i $N = 1000$ widoczna jest ogólna stabilność wyników - empiryczny rozmiar testu utrzymuje się w przedziale około 0,04-0,08 i nie wykazuje istotnych odchyśleń. Oznacza to, że asymetryczna struktura procesów nie generuje dużo fałszywych zależności między X a Y.

Natomiast pewne zaburzenie pojawia się jedynie w długiej próbie ($N = 1000$) przy silnej dodatniej autokorelacji błędu ($\rho = 0,9$), gdzie estymator OLS zawyża rozmiar testu do około 0,11. Jest to jednak odosobniony przypadek, a obie wersje HAC skutecznie utrzymują poziom istotności blisko wartości nominalnej.

Co istotne, dla ujemnej autokorelacji błędu ($\rho = -0,5$) wszystkie estymatory mieszczą się niemal idealnie w okolicach poziomu 0,05, co potwierdza, że asymetryczna struktura procesów nie generuje fałszywych zależności między X a Y. Scenariusz asymetryczny okazuje się mało problematyczny - przyczynowość Grangera jest prawidłowo identyfikowana, a autokorelacja błędu wywołuje jedynie niewielkie efekty uboczne.

7. Scenariusz 7 (Asymetria: X szybki, Y wolny)

7. Asymetria: X szybki, Y wolny

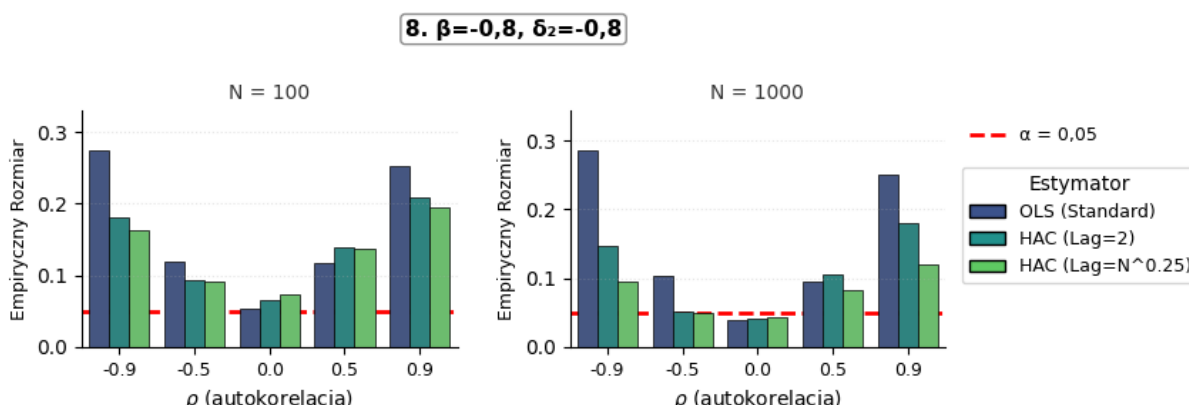


W scenariuszu, w którym X reaguje szybko ($\beta_1 = 0,2$), a Y charakteryzuje się wyższą inercją ($\delta_2 = 0,9$), rozmiar testu pozostaje kontrolowany w większości przypadków. W krótkiej próbie ($N = 100$) empiryczne wartości mieszczą się zazwyczaj w przedziale 0,05-0,10, co

wskazuje na umiarkowane zawyżenie poziomu istotności. Odstępstwo stanowi ponownie wysoka dodatnia autokorelacja, gdzie empiryczne wartości są powyżej 0,10.

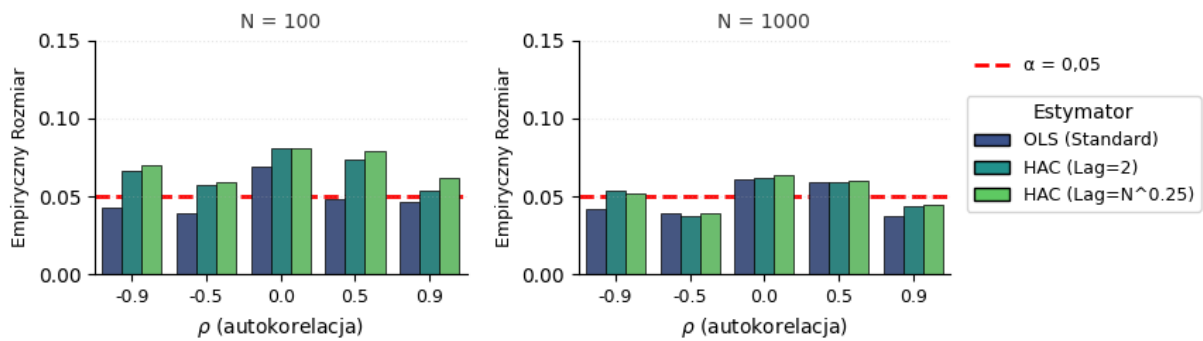
W długiej próbie ($N = 1000$) rozmiar testu jest zbliżony do poziomu nominalnego, zwłaszcza dla estymatorów HAC, które utrzymują wartości w okolicach 0,05 a czasami poniżej niego. Jedynie OLS przy silnie dodatniej autokorelacji błędu ponownie wykazuje skłonność do zawyżania liczby odrzuceń, choć efekt ten jest wyraźnie słabszy niż w scenariuszach o wysokiej dodatniej inercji. Test Grangera działa stabilnie, a korekty HAC skutecznie zabezpieczają przed wpływem autokorelacji błędu.

8. Scenariusz 8 ($\beta_1 = -0,8$; $\delta_2 = -0,8$)



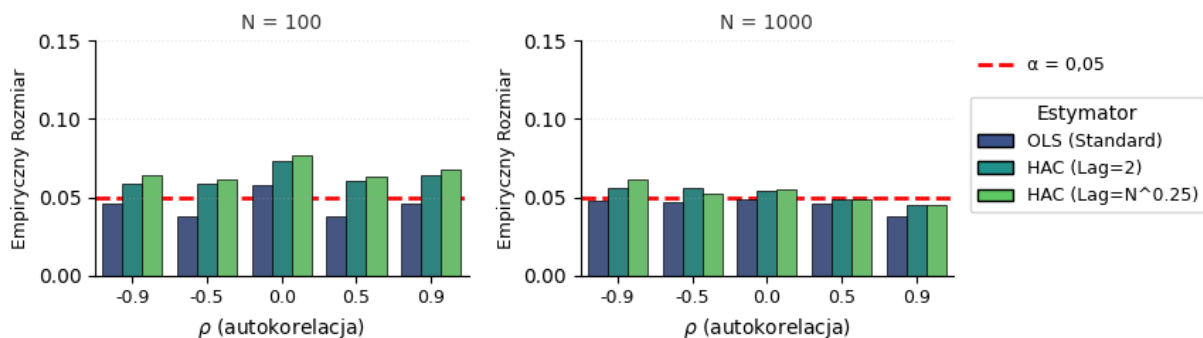
Wyniki dla tego scenariusza, w którym parametry wykazują silną ujemną autokorelację, pokazują, że test Grangera jest bardzo wrażliwy na sytuację, gdy zmienne mają oscylacyjny przebieg jednocześnie z ujemnie skorelowanym błędem losowym. Wykresy prezentują wyraźny kształt litery „U”, gdzie empiryczny rozmiar testu znacząco rośnie dla skrajnych wartości parametru ρ , osiągając dla estymatora OLS alarmujące poziomy rzędu 0,25-0,30. Oznacza to, że w obecności silnej autokorelacji reszt, standardowy test myli się niemal w 30% przypadków. Szczególnie istotnym wnioskiem płynącym z tego scenariusza jest ograniczona skuteczność estymatorów HAC w obliczu tak silnej dynamiki. Nawet w dużej próbie korekta Newey-West nie jest w stanie w pełni przywrócić nominalnego rozmiaru testu. Choć estymator z długim oknem opóźnień redukuje błąd I rodzaju z poziomu ok. 0,28 do ok. 0,10-0,12, jest to nadal wartość dwukrotnie przekraczająca dopuszczalny poziom istotności. Sugeruje to, że przy silnych procesach oscylacyjnych połączonych z dużą autokorelacją błędu, standardowe procedury odporne mogą być niewystarczające, a struktura macierzy wariancji-kowariancji jest zbyt złożona, by mogła zostać poprawnie oszacowana przy użyciu standardowych reguł wyboru opóźnień.

9. Scenariusz 9 ($\beta_1 = -0,2$; $\delta_2 = 0,2$)

9. $\beta = -0,2$, $\delta_2 = 0,2$ 

Analiza tego scenariusza ujawnia interesujące zachowanie testu Walda. W przypadku małej próby zaobserwowano wyraźne występowanie obciążenia wynikającego z krótkiej długości szeregu. Co istotne, nawet przy braku autokorelacji składnika losowego ($\rho = 0$) empiryczny rozmiar testu dla wszystkich estymatorów przewyższa nominalny poziom 0,05. Sugeruje to, że specyficzna struktura procesu VAR generuje w krótkich szeregach zniekształcenia statystyk testowych niezależne od samej autokorelacji reszt. Warto również odnotować zachowanie estymatorów przy silnej autokorelacji błędów. W małej próbie standardowy estymator OLS staje się nieco konserwatywny, czyli niedoszacowuje rozmiaru testu. Estymatory HAC (szczególnie wariant z długim okresem) korygują ten błąd, lecz mają tendencję do „przestrzeliwania”, prowadząc do nieznacznego nadmiernego odrzucania hipotezy zerowej.

Zwiększenie liczebności próby do $N = 1000$ skutkuje stabilizacją wyników i redukcją wariancji estymatorów. Zaobserwowano wyraźną zbieżność rozmiaru empirycznego do nominalnego poziomu alfa. Wyniki te prowadzą do wniosku, że przy słabej i niejednorodnej dynamice parametrów strukturalnych wpływ autokorelacji składnika losowego jest drugorzędny, a głównym źródłem zniekształceń w małych próbach pozostaje skończona liczba stopni swobody oraz specyfika samej estymacji parametrów VAR.

10. Scenariusz 10 ($\beta_1 = 0,2$; $\delta_2 = -0,2$)**10. $\beta = 0,2$, $\delta_2 = -0,2$** 

Przedstawione wyniki dla Scenariusza 10, charakteryzującego się słabą dynamiką i mieszanymi znakami parametrów ($\beta_1 = 0,2$, $\delta_2 = -0,2$), wskazują na relatywnie wysoką

stabilność testu Grangera, nawet w obecności autokorelacji składnika losowego. W małej próbie zaobserwowano ciekawe zjawisko: standardowy estymator OLS wykazuje tendencję do niedoszacowania rozmiaru testu poniżej nominalnego poziomu 0,05, szczególnie przy silnej autokorelacji zarówno ujemnej, jak i dodatniej.

Z kolei estymatory HAC, mimo że teoretycznie poprawne, w małej próbie obarczone są obciążeniem skutkującym lekkim nadmiernym odrzucaniem hipotezy zerowej.

Sytuacja ulega wyraźnej poprawie w dużej próbie, co potwierdza zgodność asymptotyczną stosowanych estymatorów. Na prawym panelu widać, że wraz ze wzrostem liczebności próby rozmiar empiryczny dla wszystkich metod - zarówno OLS, jak i HAC - zbiega do teoretycznego poziomu istotności $\alpha = 0,05$, niezależnie od siły parametru ρ .

Oznacza to, że przy słabej inercji zmiennych procesowych (X_t , Y_t) negatywny wpływ autokorelacji reszt jest ograniczony i zanika asymptotycznie.

Wnioski

Tabela 1: Średnie wartości rozmiar testu dla $N = 100$

ρ	OLS (Standard)		HAC (Lag=2)		HAC (Lag= $N^{0,25}$)	
	Średnia	Maksimum	Średnia	Maksimum	Średnia	Maksimum
-0,9	0,1668	0,494	0,1567	0,493	0,1518	0,477
-0,5	0,0839	0,197	0,1078	0,293	0,1092	0,295
0	0,0666	0,13	0,0839	0,157	0,0867	0,16
0,5	0,0973	0,27	0,0952	0,186	0,095	0,174
0,9	0,1855	0,582	0,1566	0,455	0,1497	0,419

Dla próby $N = 100$ żadna z analizowanych kombinacji (ρ , estymator) nie mieści się w teoretycznym przedziale ufności $[0,036; 0,064]$, co wskazuje na systematyczne zawyżanie rozmiaru testu we wszystkich badanych wariantach. Choć średni wynik dla estymatora OLS przy braku autokorelacji ($\rho = 0$) wynosi 0,0666 - co mogłoby sugerować jedynie niewielką inflację błędu I rodzaju - to jednak analiza wartości maksymalnych ujawnia istotną niestabilność testu, maskowaną przez uśrednione wyniki.

W najbardziej wymagających scenariuszach dynamicznych, szczególnie przy silnej dodatniej autokorelacji ($\rho = 0,9$), empiryczny rozmiar testu drastycznie rośnie, osiągając w skrajnych przypadkach 0,582 dla OLS oraz 0,419 dla estymatora HAC z długim oknem. Oznacza to, że w specyficznych warunkach test odrzuca prawdziwą hipotezę zerową nawet w blisko 60% przypadków, co czyni wnioskowanie całkowicie niewiarygodnym. Ponadto nawet zastosowanie estymatorów odpornych (HAC) nie eliminuje problemu w małej próbie - dla ujemnej autokorelacji ($\rho = -0,9$) maksymalne odchylenia wciąż sięgają wartości zbliżonych do 0,50.

Co szczególnie istotne, nawet w teoretycznie bezpiecznych warunkach braku autokorelacji ($\rho = 0$), maksymalny odnotowany rozmiar testu mieści się w przedziale 0,13-0,16. Dowodzi to, że w krótkich szeregach czasowych test Grangera jest wysoce wrażliwy nie tylko na własności składnika losowego, lecz również na samą strukturę parametrów modelu VAR.

Tabela 2: Średnie wartości rozmiar testu dla $N = 1000$

ρ	OLS (Standard)		HAC (Lag=2)		HAC (Lag= $N^{0,25}$)	
	Średnia	Maksimum	Średnia	Maksimum	Średnia	Maksimum
-0,9	0,1747	0,495	0,146	0,549	0,1243	0,501
-0,5	0,0782	0,152	0,0894	0,265	0,0858	0,278
0	0,0535	0,073	0,0544	0,067	0,0556	0,066
0,5	0,0757	0,153	0,0567	0,105	0,0519	0,082
0,9	0,1662	0,491	0,1042	0,283	0,0739	0,154

Dla próby $N = 1000$ działanie testu Grangera ulega wyraźnej poprawie w ujęciu średnim, a w kilku przypadkach przeciętny rozmiar testu mieści się w teoretycznym przedziale ufności $[0,036; 0,064]$. Najbardziej stabilne wyniki uzyskuje się przy braku autokorelacji błędu ($\rho = 0$), gdzie nie tylko wartości średnie, lecz także maksymalne dla wszystkich estymatorów (ok. 0,066-0,073) pozostają blisko poziomu nominalnego $\alpha = 0,05$. Świadczy to o prawidłowej kalibracji testu w warunkach pełnej zgodności z założeniami modelu.

Jednak analiza wartości maksymalnych dla silnej autokorelacji ($\rho = \pm 0,9$) pokazuje, że nawet tak duża próba nie eliminuje problemów w najtrudniejszych scenariuszach. Choć średni rozmiar testu dla estymatora HAC (długie okno) przy $\rho = 0,9$ spada do poziomu 0,0739, to w najgorszym przypadku frakcja błędnych odrzuceń nadal osiąga 0,154, a dla ujemnej autokorelacji $\rho = -0,9$ wzrasta do wysokiego poziomu 0,50.

Wskazuje to, że w modelu z opóźnioną zmienną zależną silna autokorelacja składnika losowego prowadzi do niezgodności estymatorów OLS, której zastosowanie korekty HAC nie jest w stanie w pełni usunąć. Estymator HAC poprawia jedynie oszacowanie wariancji, nie eliminując obciążenia parametrów strukturalnych. W skrajnych przypadkach skutkuje to trwałym zniekształceniem wyników testu, nawet przy bardzo dużej liczbie obserwacji.