

Jan Acedański

Uniwersytet Ekonomiczny w Katowicach

KRYTERIA WYBORU DYNAMICZNYCH MODELI CZYNNIKOWYCH DLA CELÓW PROGNOSTYCZNYCH

Wprowadzenie

Dynamiczne modele czynnikowe (*dynamic factor models* – DFM) są ważnym narzędziem wykorzystywanym w krótkookresowym prognozowaniu makroekonomicznym. Podejście to często jest stosowane przez banki centralne jako narzędzie wspomagające krótkookresowe prognozowanie kluczowych zmiennych makroekonomicznych [5; 6; 7]. Istota metody sprowadza się do agregacji dużej liczby potencjalnych zmiennych objaśniających do kilku wzajemnie niezależnych czynników, które następnie wykorzystywane są do prognozowania wybranej zmiennej. Wyodrębnianie czynników najczęściej dokonywane jest za pomocą klasycznej metody składowych głównych lub jej modyfikacji. Równanie progностyczne opisujące zależność pomiędzy zmienną prognozowaną a czynnikami ma zwykle postać liniową. Oprócz czynników w równaniu tym mogą wystąpić ich opóźnienia, a także składniki o charakterze autoregresyjnym.

Jednym z ważniejszych problemów związanych ze stosowaniem omawianego podejścia jest specyfikacja modelu. W szczególności konieczne jest podjęcie decyzji dotyczących liczby czynników branych pod uwagę przy prognozowaniu, a także liczby ewentualnych opóźnień występujących w równaniu progностycznym. W literaturze przedmiotu opisywanych jest wiele podejść stosowanych w celu wyboru modelu dla celów progностycznych, np. wybór optymalnej liczby czynników dokonywany był często na podstawie zmodyfikowanych kryteriów informacyjnych zaproponowanych przez [3; 7], a wybór dokładnej specyfikacji równania progностycznego był oparty na standardowym kryterium informacyjnym BIC [2; 13]. Dość często także liczba czynników oraz opóźnień ustalana była *ad hoc* na arbitralnym, zwykle niewielkim poziomie [5; 12].

W ostatnich latach pojawiły się w literaturze podejścia, które pozwalają na dokonanie jednoczesnego wyboru liczby czynników oraz opóźnień w równaniu prognostycznym. Bai i Ng [3] oraz Groen i Kapetanios [11] zaproponowali zastosowanie w tym celu zmodyfikowanych kryteriów informacyjnych. Jeszcze inną możliwością jest wybór najlepszego modelu na podstawie analizy jakości prognoz wygasłych modeli z różną liczbą czynników oraz opóźnień.

Celem pracy jest porównanie omówionych trzech podejść do wyboru najlepszego modelu DFM dla celów prognostycznych: zmodyfikowanych kryteriów informacyjnych, wyłącznej analizy dokładności prognoz wygasłych oraz podejścia tradycyjnego łączącego kryteria informacyjne i analizę prognoz wygasłych. Badane jest, które z tych trzech podejść generuje najlepsze prognozy poza próbą. Podstawowym kryterium oceny jakości prognoz jest błąd średniokwadratowy RMSE. W badaniach wykorzystano zarówno szeregi generowane na podstawie symulacji Monte Carlo, jak i dane rzeczywiste dotyczące prognozowania miesięcznej stopy inflacji w Polsce.

1. Modele czynnikowe w prognozowaniu

Istota prognozowania za pomocą dynamicznych modeli czynnikowych opiera się na założeniu, że źródłem dynamiki bardzo wielu zmiennych ekonomicznych jest niewielka liczba niezależnych czynników [16]. Jeżeli \mathbf{X}_t oznacza wektor N zmiennych objaśniających obserwowanych w chwili t , wtedy powyższe założenie można sformalizować w postaci:

$$\mathbf{X}'_t = \mathbf{F}'_t \mathbf{L} + \boldsymbol{\varepsilon}'_t, \quad t = 1, 2, \dots, T \quad (1)$$

gdzie \mathbf{F}_t jest K -wymiarowym wektorem czynników, \mathbf{L} – $K \times N$ -wymiarową macierzą ładunków czynnikowych, natomiast $\boldsymbol{\varepsilon}_t$ reprezentuje N -wymiarowy wektor składników losowych.

O czynnikach zakłada się, że są one między sobą niezależne: $E(F_{it}, F_{jt}) = 0$ dla $i \neq j$ oraz niezależne od zaburzeń losowych: $E(F_{it}, \varepsilon_{it}) = 0$. Ponadto zakłada się, że składniki losowe mają zerową wartość oczekiwaną $E(\varepsilon_{it}) = 0$ oraz są niezależne między sobą $E(\varepsilon_{it}, \varepsilon_{jt}) = 0$, $i \neq j$. Oprócz powyższych założeń zwykle przyjmuje się także, że macierz ładunków czynnikowych jest stała oraz że nie występuje autokorelacja składników losowych, chociaż uchylenie tych założeń jest możliwe [9]. Liczba czynników K oraz ich realizacje nie są obserwowane i muszą być estymowane przez badacza. Fakt, że czynniki nie są obserwowane sprawia, że w omawianej sytuacji nie powinno się stosować standardowych kryteriów informacyjnych przy wyborze liczby czynników oraz opóźnień w równaniu prognostycznym.

Najpopularniejszą metodą estymacji czynników jest metoda składowych głównych (*principal components analysis* – PCA). Pozwala ona na znalezienie takich N -wymiarowych wektorów czynników $\hat{\mathbf{F}}_t$ oraz takiej $N \times N$ -wymiarowej macierzy ładunków czynnikowych $\hat{\mathbf{L}}$, że:

$$\mathbf{X}'_t = \hat{\mathbf{F}}'_t \hat{\mathbf{L}}, \quad t = 1, 2, \dots, T \quad (2)$$

Estymowane czynniki $\hat{\mathbf{F}}_t$ są wzajemnie niezależne oraz uporządkowane w takiej kolejności, że zmiany kolejnego czynnika wyjaśniają coraz mniejszy odsetek wahań zmiennych obserwowanych z wektorów \mathbf{X}_t .

Przy szacowaniu $\hat{\mathbf{F}}_t$ oraz $\hat{\mathbf{L}}$ metodą składowych głównych wykorzystuje się dekompozycję macierzy $\mathbf{X}\mathbf{X}'$, w której $\mathbf{X} = [\mathbf{X}_1 \ \mathbf{X}_2 \ \dots \ \mathbf{X}_T]'$. Kolejne kolumny macierzy $\hat{\mathbf{L}}$ są wektorami własnymi macierzy $\mathbf{X}\mathbf{X}'$ związanymi z kolejnymi, uporządkowanymi malejąco wartościami własnymi tej macierzy [16]. Przy danej macierzy ładunków czynnikowych realizacje czynników można łatwo obliczyć jako $\hat{\mathbf{F}}'_t = \mathbf{X}'_t \hat{\mathbf{L}}^{-1}$. Czasami zdarza się, że niektóre wektory \mathbf{X}_t są niekompletne, szczególnie na końcu lub początku próby. W takiej sytuacji przy estymacji czynników stosowany jest dodatkowo algorytm maksymalizacji oczekiwań (*expectations maximization* – EM) [16]. Algorytm ten ma charakter rekurencyjny, składający się z dwóch podstawowych kroków. W pierwszym kroku w oparciu o dane wartości $\hat{\mathbf{F}}_t$ oraz $\hat{\mathbf{L}}$ uzupełniane są brakujące elementy w niekompletnych wektorach \mathbf{X}_t . W drugim kroku korzystając już z kompletnych wektorów \mathbf{X}_t dokonuje się estymacji $\hat{\mathbf{F}}_t$ oraz $\hat{\mathbf{L}}$ zwykłą metodą składowych głównych. Startowe wartości $\hat{\mathbf{F}}_t$ oraz $\hat{\mathbf{L}}$ w procedurze wyznacza się najczęściej biorąc pod uwagę tylko kompletne wektory \mathbf{X}_t .

Metoda PCA zakłada statyczną zależność pomiędzy zmiennymi objaśnianymi a czynnikami. W przypadku przyjęcia założenia, że zależność ta nie ma charakteru statycznego stosowane są inne metody estymacji czynników, np. wykorzystujące analizę spektralną [9] lub filtry Kalmana [8].

W drugim kroku, po wyborze właściwej liczby czynników specyfikowane jest równanie prognostyczne, które można przedstawić w następującej postaci:

$$Y_{t+h} = \alpha_0 + \sum_{i=0}^{ny} \alpha_{i+1} Y_{t-i} + \sum_{j=0}^{nf} \beta_j \hat{\mathbf{F}}_t^{(k)} + \varepsilon_{Yt} \quad (3)$$

gdzie Y_t oznacza zmienną prognozowaną, $\hat{\mathbf{F}}_t^{(k)}$ reprezentuje wektor k pierwszych oszacowanych czynników, ε_{Y_t} jest składnikiem losowym, natomiast α_i oraz β_j są parametrami równania. Horyzont prognozy oznaczony został przez h , n_y jest nieznaną liczbą opóźnionych wartości zmiennej prognozowanej, natomiast n_f reprezentuje opóźnienia czynników. Parametry równania (3) szacowane są metodą najmniejszych kwadratów.

Należy zwrócić uwagę, że omówione tutaj podejście jest zbliżone do koncepcji prognozowania za pomocą wskaźników wyprzedzających. Równanie (3) wskazuje bowiem, że czynniki mają charakter wskaźników wyprzedzających. W tym miejscu ujawnia się pewna niespójność całego omawianego podejścia, gdyż przy estymacji czynników zakładano, że zależności pomiędzy czynnikami a zmiennymi \mathbf{X}_t mają charakter jednoczesny. Tymczasem w praktyce do zbioru zmiennych objaśniających zalicza się bardzo wiele zmiennych nie patrząc na to, czy mają one charakter wskaźników wyprzedzających względem zmiennej prognozowanej. Specyfikacja (3) ma więc charakter *ad hoc*.

Alternatywne podejście polega na specyfikacji równania opisującego zależność jednoczesne pomiędzy Y_t a $\hat{\mathbf{F}}_t^{(k)}$ wraz z ewentualnymi opóźnieniami [7], jednak w takiej sytuacji stawianie prognoz wymaga znajomości przyszłych wartości czynników, a więc konieczne jest skonstruowanie jeszcze jednego modelu służącego do prognozowania przyszłych wartości czynników. W tym celu najczęściej stosowane są modele wektorowej autoregresji.

2. Metody wyboru liczby czynników oraz opóźnień w równaniu prognostycznym

W literaturze opisywanych jest kilka metod wyboru optymalnej liczby czynników oraz specyfikacji równania prognostycznego. Można je podzielić na dwie grupy. W pierwszej obie decyzje są dokonywane oddzielnie. Do drugiej grupy należą metody jednoczesnego wyboru liczby czynników i opóźnień.

2.1. Kryteria Bai i Ng wyboru liczby czynników

Najpopularniejszymi kryteriami wyboru liczby czynników są modyfikacje standardowych kryteriów informacyjnych zaproponowane przez Bai i Ng [3].

Metoda składowych głównych generuje N czynników, a więc tyle samo, ile jest zmiennych obserwowanych tworzących wektory \mathbf{X}_t . Bai oraz Ng zaproponowali kryteria, które pozwalają na estymację nieznaną liczbę K czynników

rzeczywistych. Metoda ta daje właściwe oszacowania asymptotycznie, tzn., gdy badacz dysponuje jednocześnie długimi szeregami czasowymi – $T \rightarrow \infty$ oraz dużym zbiorem zmiennych objaśniających – $N \rightarrow \infty$. W omawianym podejściu rozważa się sekwencję modeli postaci (1), w których uwzględnianych jest k pierwszych czynników $k = 1, 2, \dots, k_{MAX}$ wraz z odpowiednimi postaciami macierzy ładunków czynnikowych uzyskanych metodą głównych składowych. Jeżeli przez $V(k)$ oznaczyć łączną wariancję wszystkich reszt w modelu z k czynnikami:

$$V(k) = \frac{1}{NT} \sum_{i=1}^N \sum_{t=1}^T \left(X_{it} - \sum_{j=1}^k \hat{F}_{jt} \hat{L}_{ji} \right)^2 \quad (4)$$

natomiast $\hat{\sigma}^2$ oznacza oszacowanie wariancji wszystkich składników losowych ε_i , wtedy optymalna liczba czynników \hat{k} powinna zostać ustalona na takim poziomie, aby wartości kryteriów:

$$PC(k) = V(k) + k\hat{\sigma}^2 \frac{N+T}{NT} \ln\left(\frac{NT}{N+T}\right) \quad (5)$$

$$IC(k) = \ln(V(k)) + k\hat{\sigma}^2 \frac{N+T}{NT} \ln(\min\{N, T\}) \quad (6)$$

były jak najmniejsze. W praktycznych zastosowaniach $\hat{\sigma}^2$ może być przybliżane przez $V(k_{MAX})$. W pracy Bai i Ng powyższe kryteria były oznaczane jako $PC_{p1}(k)$ oraz $IC_{p2}(k)$. W badaniach porównawczych kryteria te odznaczały się dobrymi własnościami w porównaniu do innych kryteriów rozważanych w cytowanym opracowaniu. Należy jeszcze zaznaczyć, że w ostatnich latach pojawiło się w literaturze kilka alternatywnych kryteriów o korzystnych własnościach w porównaniu do kryteriów przedstawionych powyżej, choć nie są one jeszcze tak popularne w zastosowaniach prognostycznych [1; 14; 15].

2.2. Metody analizy prognoz wygasłych w wyborze liczby opóźnień w równaniu prognostycznym

Wyboru właściwej liczby opóźnień n_y i n_f można także dokonywać poprzez porównanie jakości prognoz wygasłych modeli z różnymi liczbami opóźnień. W zbiorze T obserwacji wyodrębnia się r okien obserwacyjnych, począwszy od pierwszego obejmującego obserwacje o numerach 1, 2, ..., T_p . Kolejne okna są

albo przesuwane o jedną obserwację wprzód (*rolling window*), albo rozszerzane o kolejną obserwację (*expanding window*). W pierwszym przypadku okna liczą zawsze obserwacji T_p . W drugim liczba obserwacji wzrasta o jeden.

Następnie dla każdej metody m oraz okna i szacowane są parametry równania (3) oraz wyznaczane są prognozy na h okresów do przodu: $\hat{Y}_{T_p}(m, h, i)$, które są porównywane z rzeczywistymi wartościami zmiennej prognozowanej w danym okresie. Jako podstawową miarę jakości prognoz przyjęto w pracy błąd średniokwadratowy:

$$RMSE(m, h) = \sqrt{\frac{1}{r} \sum_{i=1}^r (\hat{Y}_{T_p}(m, h, i) - Y_{T_p+i+h})^2} \quad (7)$$

Jako najlepszy model do prognozowania przy danym horyzoncie prognozy wybierano ten model, który cechował się najmniejszym błędem średniokwadratowym.

W literaturze przedmiotu cały czas toczy się dyskusja dotycząca właściwego sposobu konstrukcji okien prognostycznych. Za oknami rozszerzonymi przemawia fakt, że pozwalają one uwzględnić wszystkie obserwacje. Zwolennicy okien przesuwanych podkreślają natomiast, że podejście to cechuje się większą odpornością w przypadku niestabilności procesu generującego dane. Badania empiryczne i symulacyjne nie są w stanie jednoznacznie rozstrzygnąć tej kwestii [10].

Biorąc pod uwagę dwa kryteria wyboru optymalnej liczby czynników oraz dwie metody konstrukcji okien przy wyborze najlepszej specyfikacji równania prognostycznego łącznie w pracy rozważane były cztery tradycyjne podejścia do wyboru najlepszego modelu DFM dla celów prognozowania.

2.3. Metody jednoczesnego wyboru liczby czynników i opóźnień

Na alternatywne podejścia do wyboru optymalnego modelu DFM, mające na celu jednoczesne podjęcie decyzji dotyczącej liczby czynników oraz opóźnień w równaniu prognostycznym można patrzeć jako na rozszerzenia metod opisanych w poprzednich dwóch rozdziałach. Można je podzielić na dwie grupy. W pierwszej znajdują się metody wykorzystujące zmodyfikowane wersje kryteriów informacyjnych. Druga bazuje na analizie prognoz wygasłych.

Odnośnie do kryteriów z pierwszej grupy Bai i Ng [4] zaproponowali kryterium finalnego błędu predykcji (*final prediction error* – FPE), które przy dużych próbach pozwala na wybór modelu postaci (3) minimalizującego średniokwadratowy błąd predykcji przy danym horyzoncie prognozy:

$$FPE = \ln(\hat{\sigma}^2) + (2 + n_y + k(1 + n_f)) \frac{\ln T}{T} + \frac{\hat{\gamma}' \hat{\Sigma} \hat{\gamma}}{\hat{\sigma}^2} \frac{\ln N}{N} \quad (8)$$

gdzie $\hat{\gamma}$ oznacza wektor ocen wszystkich parametrów strukturalnych α_i , β_j w równaniu (3), natomiast $\hat{\Sigma}$ reprezentuje oszacowanie macierzy kowariancji zmiennych wszystkich zmiennych niezależnych równania (3).

Do pierwszej grupy należą także kryteria podane przez Groena i Kapetaniosa [11]. Są one modyfikacją standardowych kryteriów informacyjnych uwzględniającą fakt, że zmienne objaśniające w równaniu prognostycznym nie są z góry ustalone, ale muszą być estymowane. W odniesieniu do równania (3) kryteria te przyjmują następującą formę:

$$BICM = \frac{T}{2} \ln(\hat{\sigma}^2) + (2 + n_y) \ln T + k(1 + n_f) \ln T \left(1 + \frac{T}{N}\right) \quad (9)$$

$$HQICM = \frac{T}{2} \ln(\hat{\sigma}^2) + 2(2 + n_y) \ln(\ln T) + 2k(1 + n_f) \ln(\ln T) \left(1 + \frac{T}{N}\right) \quad (10)$$

We wszystkich trzech przypadkach jako najlepszy model wybiera się ten, dla którego wartości kryteriów są najmniejsze.

Wyboru najlepszego modelu prognostycznego w całości można również dokonać porównując jakość prognoz wygasłych, zgodnie z procedurą opisaną w poprzednim rozdziale, przy czym oprócz liczby opóźnień w równaniu (3) uwzględnia się teraz dodatkowo różne liczby czynników w modelu n_f .

3. Metodologia badań porównawczych omawianych metod

Łącznie w pracy porównywano 9 metod wyboru najlepszego modelu czynnikowego dla celów prognostycznych. Były to 4 podejścia klasyczne łączące kryteria Bai-Ng z badaniem jakości prognoz wygasłych – w dalszej części pracy oznaczane jako: *PC_rol* (kryterium *PC* i okna przesuwane), *PC_exp* (kryterium *PC* i okno rozszerzane), *IC_rol* (kryterium *IC* i okna przesuwane), *IC_exp* (kryterium *IC* i okno rozszerzane) – oraz 5 podejść nowych pozwalających na jednoczesny wybór liczby czynników i liczby opóźnień w równaniu prognostycznym: zmodyfikowane kryteria informacyjne – *FPE*, *BICM*, *HQICM* oraz kryteria bazujące na badaniu jakości prognoz wygasłych – *RF_rol* (z oknami przesuwanymi) i *RF_exp* (z oknami rozszerzanymi).

Badania przeprowadzono zarówno na symulowanych, jak i na rzeczywistych szeregach czasowych. W przypadku danych symulowanych w pierwszym kroku generowano szereg czasowy K czynników:

$$\mathbf{F}_t = \rho_F \mathbf{F}_{t-1} + \mathbf{v}_{Ft}, \quad \mathbf{v}_{Ft} \sim N(\mathbf{0}, \mathbf{I}_K) \quad (11)$$

gdzie ρ_F określa jednakowy dla wszystkich czynników poziom autokorelacji, \mathbf{I}_K symbolizuje natomiast K -wymiarową macierz jednostkową. Dla każdej symulacji liczbę czynników K wybierano losowo ze zbioru $\{1, 2, \dots, 5\}$. Następnie generowano szereg $(N+1) \times K$ -wymiarowych macierzy ładunków czynnikowych:

$$\begin{aligned} \mathbf{L}_t &= [1 + \psi_L (\boldsymbol{\eta}_{L_t} \otimes \mathbf{I}_k)] \mathbf{L}_{t-1} \\ \boldsymbol{\eta}_{L_t} &= \rho_L \boldsymbol{\eta}_{L_{t-1}} + \mathbf{v}_{L_t}, \quad \text{vec}(\mathbf{v}_{L_t}) \sim N(\mathbf{0}, \mathbf{I}_{(N+1)K}) \end{aligned} \quad (12)$$

przy czym ψ_L jest parametrem skalującym, $\boldsymbol{\eta}_{L_t}$ jest ciągiem $(N+1) \times K$ -wymiarowych realizacji procesu autoregresyjnego o współczynniku autokorelacji ρ_L , natomiast \otimes reprezentuje iloczyn Kroneckera, a więc oznacza zamianę macierzy w wektor kolumnowy. Z powyższego zapisu wynika, że w ogólnym przypadku dopuszczano zmienność macierzy ładunków czynnikowych. W większości symulacji zakładano jednak, że $\psi_L = 0$, co oznaczało stałość macierzy \mathbf{L}_t . Początkowe wartości składowych wektora czynników \mathbf{F}_0 oraz macierzy ładunków czynnikowych \mathbf{L}_0 losowane były z rozkładu $N(0, 1)$. Wreszcie na podstawie wektorów \mathbf{F}_t oraz macierzy \mathbf{L}_t tworzone były $(N+1)$ -wymiarowe wektory obserwacji:

$$\mathbf{Z}_t = \mathbf{F}_t \mathbf{L}_t + \boldsymbol{\eta}_{Z_t}, \quad \boldsymbol{\eta}_{Z_t} = \rho_Z \boldsymbol{\eta}_{Z_{t-1}} + \mathbf{v}_{Z_t}, \quad \mathbf{v}_{Z_t} \sim N(\mathbf{0}, \mathbf{I}_{N+1}) \quad (13)$$

Pierwsza kolumna macierzy \mathbf{Z}_t reprezentowała zmienną prognozowaną Y_t . Pozostałe N kolumn traktowano jako potencjalne zmienne objaśniające \mathbf{X}_t . Przed prowadzeniem dalszych obliczeń kolumny macierzy \mathbf{Z}_t zostały zstandardyzowane. Przyjmując powyższą specyfikację nie zakładano więc żadnego sprecyzowanego modelu parametrycznego dla zmiennej prognozowanej, ale traktowano ją jak jedną z wielu zmiennych ekonomicznych, których dynamika była kształtowana przez nieobserwowane czynniki.

Przeprowadzając badania przyjmowano, że prognozowanie nie ma charakteru jednorazowego, ale sekwencyjny. To znaczy zakładano, że prognozy stawiane są przez p kolejnych okresów zawsze na podstawie T ostatnich obserwacji. Takie postępowanie jest charakterystyczne dla omawianych modeli, które zwykle wykorzystywane są przez instytucje do regularnego comiesięcznego stawiania prognoz wybranych zmiennych. Symulowane szeregi czasowe liczyły więc $T + p + h_{MAX}$ obserwacji, gdzie h_{MAX} oznacza maksymalny horyzont prognozy. Na ich podstawie dla każdej metody m oraz dla każdego horyzontu prognozy h wybierano najlepszy model i generowano prognozy $\hat{Y}_{T+i-1}(m, h)$, $i = 1, 2, \dots, p$, które porównywano z rzeczywistymi realizacjami prognozowanej zmiennej. W ten sposób uzyskiwano ciąg p błędów prognoz:

$$u_i(m, h) = \hat{Y}_{T+i-1}(m, h) - Y_{T+i-1+h}, \quad i = 1, 2, \dots, p \quad (14)$$

Przy porównywaniu jakości prognoz uzyskanych różnymi metodami stosowano kryterium błędu średniokwadratowego prognoz wygasłych, przy czym aby ocenić, czy zaobserwowana w próbie p błędów różnica jakości była statystycznie istotna korzystano z testu warunkowej zdolności prognostycznej (*conditional prediction ability* – CPA) opracowanego przez Giacomini i White'a [10]. Hipoteza podstawowa w omawianym teście wskazuje, że jakość prognoz generowanych przez dwa konkurencyjne podejścia była w próbie taka sama. Hipoteza alternatywna jest prostym zaprzeczeniem hipotezy badanej.

Jeżeli przez $\Delta L_i(m_1, m_2, h)$ oznaczyć różnicę między błędami średniokwadratowymi prognoz wygasłych o horyzoncie h uzyskanych z dwóch różnych modeli m_1 i m_2 dla próby i , natomiast przez \mathbf{S}_i – s -elementowy wektor informacji dodatkowej dostępnej w próbie i , wtedy statystyka testowa testu CPA przyjmuje postać:

$$\chi^2 = p \bar{\mathbf{L}}' \hat{\mathbf{\Omega}}^{-1} \bar{\mathbf{L}} \quad (15)$$

gdzie $\bar{\mathbf{L}} = \frac{1}{p} \sum_{i=1}^p \Delta L_i(m_1, m_2, h) \mathbf{S}_i$, natomiast $\hat{\mathbf{\Omega}}$ oznacza oszacowanie macierzy kowariancji wektorów postaci $\Delta L_i(m_1, m_2, h) \mathbf{S}_i$ uzyskane metodą Neweya-West.

Asymptotycznie, przy dużej liczbie porównywanych prognoz ($p \rightarrow \infty$), statystyka testowa ma rozkład χ^2 z s stopniami swobody. W przypadku braku warunkowej informacji dodatkowej $S_i = 1$. Wtedy test ma charakter bezwarunkowy. Test bezwarunkowy wskazuje podejście, które generowało średnio dokładniejsze prognozy. Można więc uznać, że powinno ono także wskazywać, które podejście będzie lepsze w pewnym, bliżej nieokreślonym momencie w przyszłości. Test względny odpowiada natomiast na pytanie, czy na podstawie innych dostępnych w próbie informacji, poza średnią jakością prognoz wygasłych, można wskazać podejście, które będzie generowało prognozy lepsze w określonych warunkach. Omawiany test może być stosowany tylko w sytuacji, gdy kolejne prognozy obliczane są na podstawie takiej samej liczby obserwacji, co jest dość istotnym ograniczeniem. Z tej właśnie przyczyny długość prób T była zawsze stała.

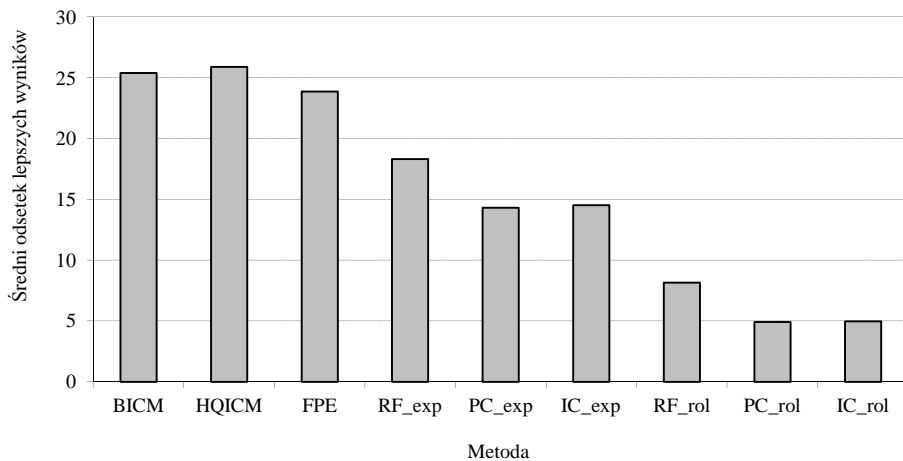
W badaniach symulacyjnych w zależności od liczby stawianych prognoz p , procedura generowania szeregów czasowych i testowania jakości prognoz powtarzana była 100 lub 1000 razy.

Własności różnych metod wyboru optymalnego modelu badano także na danych rzeczywistych. W tym celu wykorzystano szereg miesięcznej inflacji konsumenckiej w Polsce obejmujący okres 11.2000-08.2011. Liczył on więc 130 obserwacji. Zbiór zmiennych objaśniających składał się z 54 zmiennych, wśród których były m.in.: sektorowe indeksy cen, wskaźniki inflacji bazowej, indeksy produkcji sprzedanej przemysłu, produkcja ważniejszych wyrobów, zmienne reprezentujące dynamikę handlu zagranicznego, podstawowe zmienne rynku pracy, wyniki badań przewidywanej koniunktury gospodarczej, ceny surowców, poziom stóp procentowych oraz kursy walut. Wszystkie zmienne zostały pozbawione trendu oraz odsezonowane w programie Demetra. Tak przygotowane szeregi potraktowano tak, jak jeden z wysymulowanych szeregów opisanych powyżej, przy czym przy badaniu analizowano zarówno próby przesuwane, jak i rozszerzane. A więc na podstawie okien o długości co najmniej T_p wybierano najlepszy model według każdego kryterium, który następnie służył do obliczania prognoz. Podstawą oceny jakości prognoz były błędy średniokwadratowe oraz średnie błędy absolutne prognoz wygasłych oraz wyniki testów CPA.

4. Wyniki

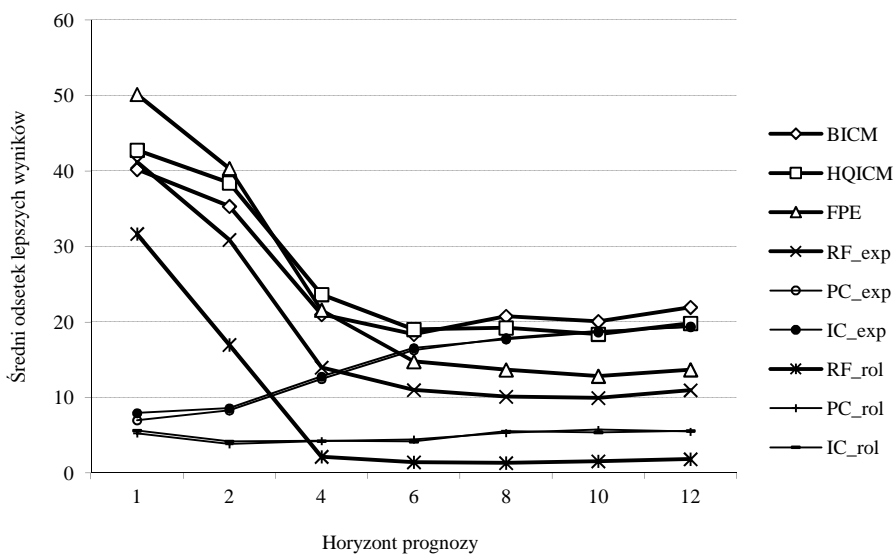
Dla danych symulowanych prezentowane są wyniki wskazujące, w jakim odsetku prób model wybrany określoną metodą okazał się lepszy z punktu widzenia bezwarunkowego testu CPA na poziomie istotności od innych modeli, przy czym dla zwiększenia czytelności prezentacji podawane są rezultaty uśrednione dla wszystkich konkurencyjnych modeli.

Wyniki uzyskano dla szeregów liczących 212 obserwacji, z czego szereg służący do wyboru najlepszego modelu liczył $T = 100$ obserwacji, a liczba przesuwanych okien była równa $p = 100$. Dokładności prognoz badano dla 7 różnych horyzontów: $h = 1, 2, 4, 6, 8, 10, 12$. Liczba zmiennych objaśniających została ustalona na poziomie $N = 100$. W przypadku metod wykorzystujących badanie dokładności prognoz wygasłych okna stosowane przy wyborze modelu liczyły $T_p = 50$ obserwacji. Liczba generowanych szeregów była równa $n = 1000$. We wszystkich przypadkach zbiór parametrów, względem którego dokonywano wyboru najlepszego modelu był postaci $K = \{1, 2, \dots, 6\}$, $n_y = \{0, 1\}$, $n_f = \{0, 1, 2\}$.



Rys. 1. Ranking metod – wariant podstawowy ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe \mathbf{L})

W podstawowym wariancie zakładano, że szereg czynników cechuje się autokorelacją na poziomie $\rho_F = 0,9$ (zob. równanie 11), wszystkie składniki losowe są homoskedastyczne, a macierz ładunków czynnikowych \mathbf{L} jest stała w czasie. Wyniki dla takiego wariantu zilustrowano na rys. 1. Przedstawia on średnią liczbę przypadków, w których dana metoda okazała się lepsza od metod konkurencyjnych, przy czym podane są wyniki średnie ze względu na metody konkurencyjne oraz rozważane horyzonty prognozy. Z rysunku wynika więc, że metody BICM i HQICM okazywały się średnio w 25% spośród 1000 przypadków statystycznie lepsze od innych metod. Niewiele niższy wynik uzyskuje kryterium FPE. Metody wykorzystujące analizę jakości prognoz wygasłych na podstawie okien rozszerzanych uzyskują już wyniki gorsze – około 17% dla metody jednoczesnej RF_ext oraz poniżej 15% dla metod łączonych PC_exp i IC_exp. Najgorsze rezultaty uzyskują metody korzystające z okien przesuwanych. Dla podejść łączonych jest to mniej więcej 5%.

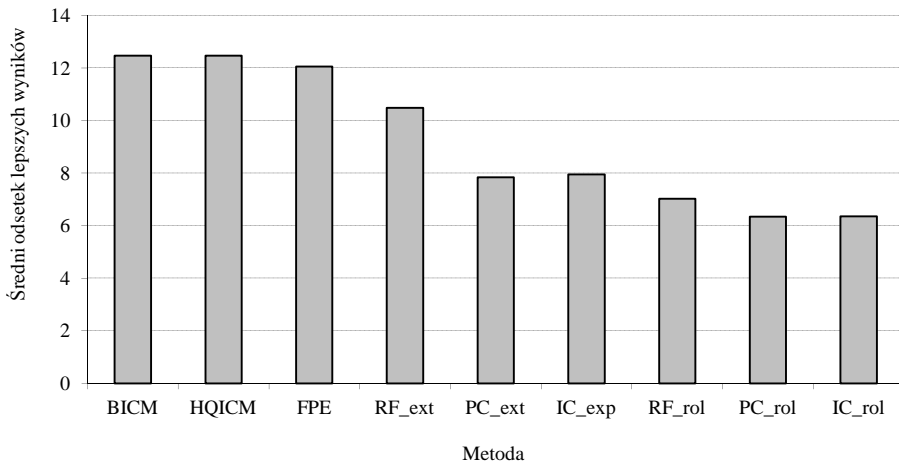


Rys. 2. Ranking metod według horyzontu prognoz – wariant podstawowy ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe \mathbf{L})

Rysunek 2 prezentuje bardziej szczegółowe wyniki z rozbiciem na różne horyzonty prognoz. Wynika z niego, że dla horyzontu $h = 1$ metody jednoczesne BICM, HQICM, FPE, RF_ext oraz RF_rol osiągają znacznie lepsze rezultaty od metod kombinowanych. Przykładowo kryterium FPE dawało lepsze wyniki od

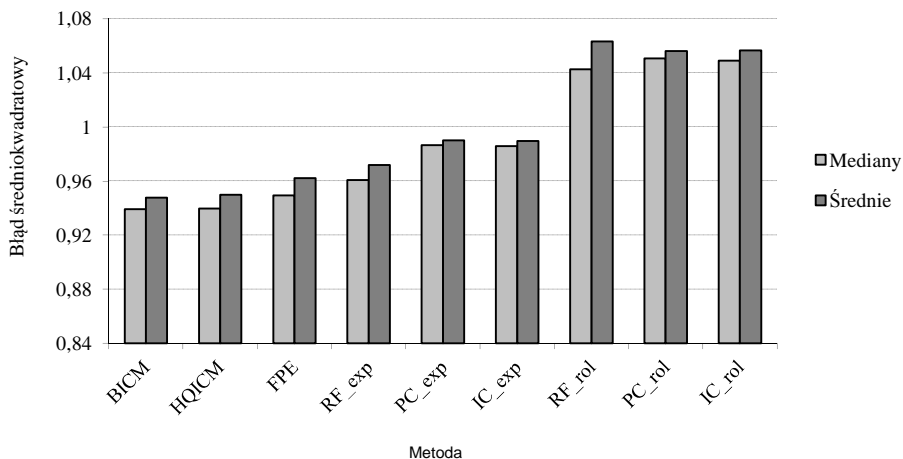
innych modeli średnio w 50% przypadków. Kryterium RF_rol było lepsze mniej więcej w 30% przypadków, a dla pozostałych było to w granicach 40%. Kryteria łączone dla tego horyzontu dawały lepsze wyniki w mniej niż 10% analizowanych przypadków, jednak w miarę wzrostu horyzontu przewaga metod jednoczesnych wyraźnie spada. Dla horyzontu $h = 12$ metody PC_exp oraz IC_exp dają praktycznie takie same rezultaty (około 20%), jak najlepsze kryteria jednoczesne BICM oraz HQICM. Zdecydowanie najsłabiej prezentuje się podejście RF_rol. Na omawianym rysunku widać także wyraźnie, że oba podejścia łączone wykorzystujące okna rozszerzane (PC_exp i IC_exp) dają praktycznie identyczne rezultaty. Podobne spostrzeżenie dotyczy metod PC_rol i IC_rol. Taka sytuacja ma miejsce we wszystkich wynikach prezentowanych w pracy.

Rysunek 3 ilustruje zbiorcze wyniki w przypadku stosowania wersji warunkowej testu CPA, przyjmując jako warunek ostatnią dostępną obserwację zmiennej prognozowanej: $S_i = Y_{T+i-1}$. Z takiej perspektywy różnice między podejściami są dużo mniejsze, choć pozycja metod w rankingu nie zmieniła się istotnie. Najlepsze kryteria jednoczesne okazywały się lepsze od innych metod średnio w ponad 12% przypadków. Najsłabsze wyniki osiągały podejścia wykorzystujące okna przesuwane – niewiele ponad 6%.



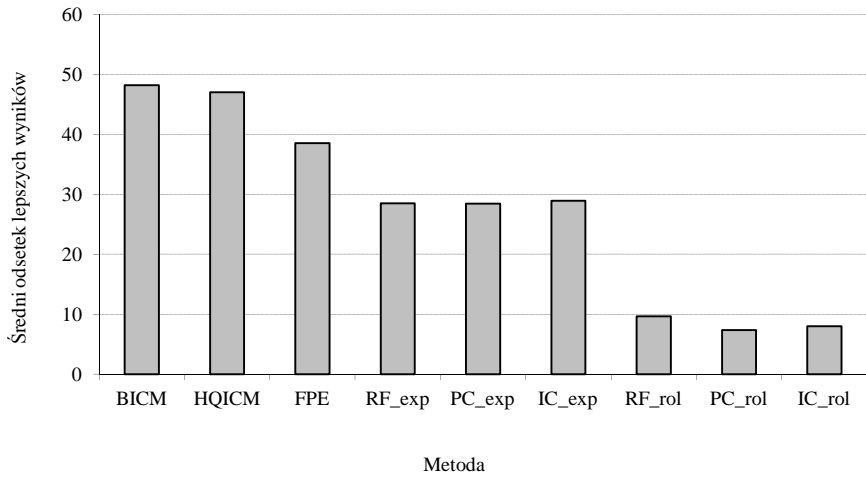
Rys. 3. Ranking metod – testy warunkowe, wariant podstawowy ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe L)

Na rys. 4 przedstawiono charakterystyki błędów średniokwadratowych prognoz wygasłych – średnie wartości median i średnich dla różnych horyzontów. Wyniki te są dość zgodne z rankingiem przedstawionym na rys. 1. Najniższe błędy uzyskiwano dla jednoczesnych kryteriów informacyjnych, nieco wyższe w przypadku metod korzystających z okien rozszerzanych. Zdecydowanie najwyższymi błędami cechowały się podejścia związane z oknami przesuwanymi. Należy też zwrócić uwagę, że relatywne różnice pomiędzy błędami nie są duże. Rozpiętość wyników nie przekracza 10%.

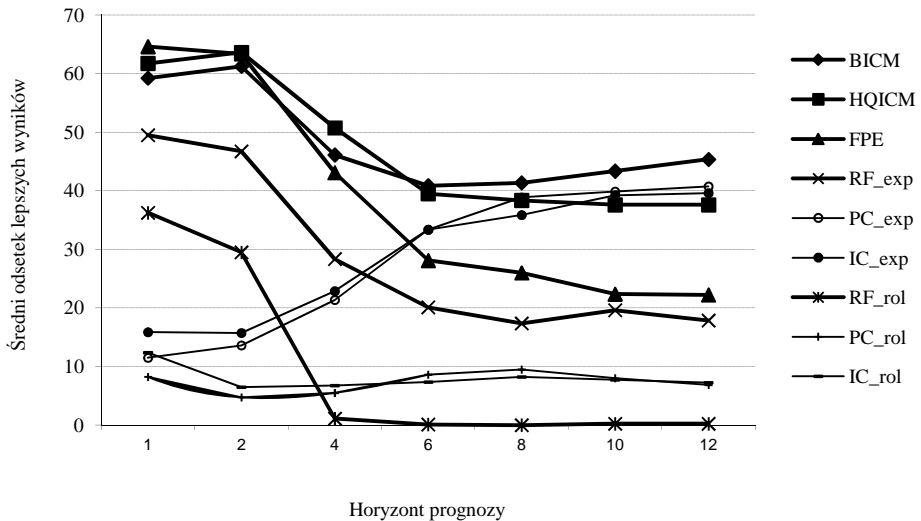


Rys. 4. Błędy średniokwadratowe dla wariantu podstawowego ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe L)

Dla zweryfikowania rezultatów przedstawionych na rys. 1 i 2 przeprowadzono podobne badanie, ale przy większej liczbie prognoz $p = 500$ stawianej dla każdego symulowanego szeregu. Większa liczba obserwacji sprawia, że wyniki testów CPA są bardziej wiarygodne i częściej wskazują na odrzucenie hipotezy o braku różnic w jakości prognoz dla dwóch modeli, jednak ze względu na czasochłonność obliczeń dla takiej liczby prognoz przeprowadzono tylko 100 powtórzeń procedury. Wyniki badań prezentują rys. 5 i 6. Potwierdzają one rezultaty zaprezentowane dla wariantu podstawowego na rys. 1 i 2. Najlepsze wyniki dają kryteria BICM oraz HQICM. Nieco gorsze FPE i metody korzystające z prognoz wygasłych i okien rozszerzanych. Zdecydowanie najslabiej wypadają metody wykorzystujące okna przesuwane. Również rozłożenie wyników w zależności od horyzontu prognoz jest podobne jak na rys. 2.



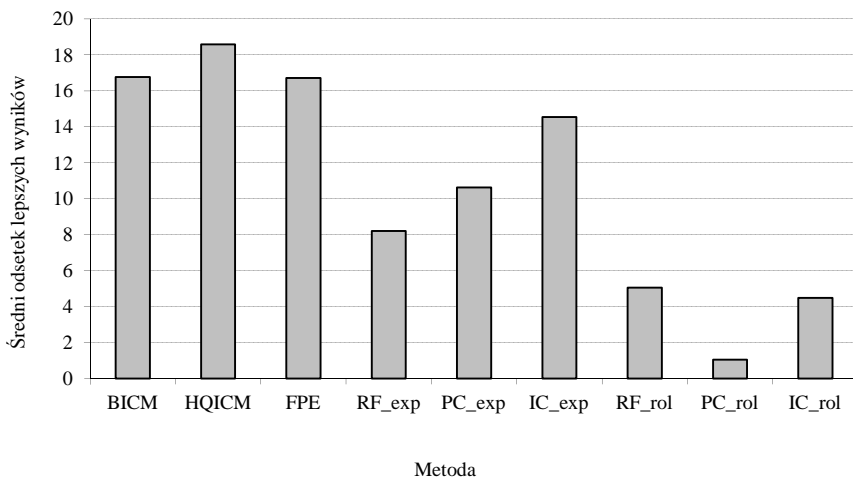
Rys. 5. Ranking metod – wariant podstawowy, większa liczba prognoz ($p = 500$, $n = 100$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe L)



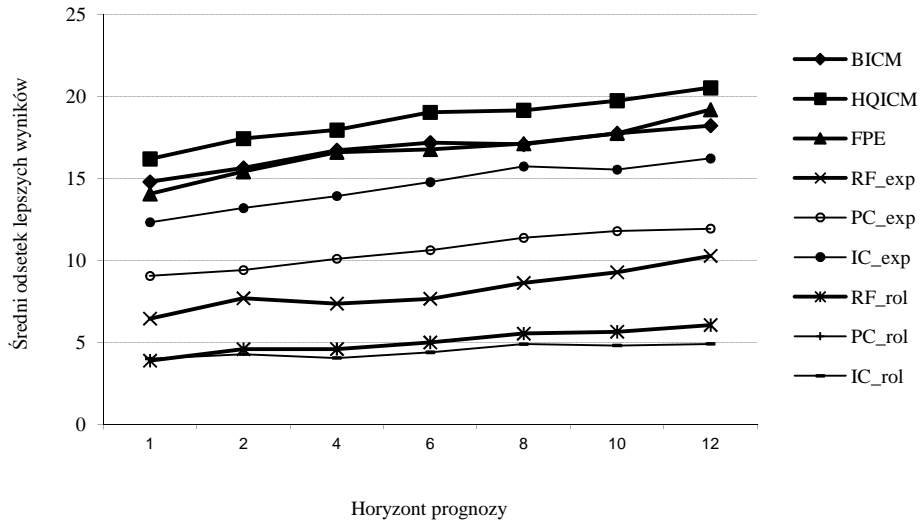
Rys. 6. Ranking metod według horyzontu prognoz – wariant podstawowy, większa liczba prognoz ($p = 500$, $n = 100$, $\rho_F = 0,9$, $\rho_Z = 0$, stałe L)

Należy zwrócić uwagę, że przy dużej liczbie prognoz testy zdecydowanie częściej wskazują na odrzucenie hipotezy o równych zdolnościach prognostycznych. W przypadku najlepszych kryteriów średnio w prawie 50% prób okazywały się one statystycznie lepsze od innych podejść, a dla krótkich horyzontów odsetek ten przekraczał 60%. Jest to więc mniej więcej dwa razy więcej niż w wariancie podstawowym.

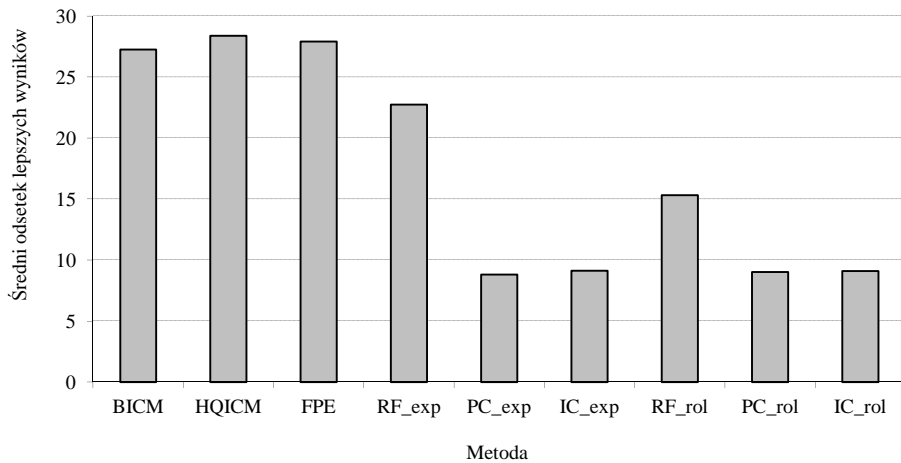
Rysunki 7 i 8 prezentują wyniki dla modelu czysto losowego, w którym nie występowała żadna autokorelacja $\rho_F = 0$, $\rho_Z = 0$. Także w tym przypadku najlepsze wyniki dają kryteria BICM, HQICM oraz FPE. Dobrze wypada także metoda IC_exp. Podobnie jak w poprzednich przypadkach, najgorsze prognozy otrzymuje się z modeli bazujących na prognozach wygasłych o oknach przesuwanych. W przeciwieństwie do poprzednich przypadków ranking ten jest stabilny ze względu na horyzont prognozy i prawie zawsze wygląda w taki sam sposób, co ilustruje rys. 8.



Rys. 7. Ranking metod – model czysto losowy ($p = 100$, $n = 1000$, $\rho_F = 0$, $\rho_Z = 0$, stałe L)



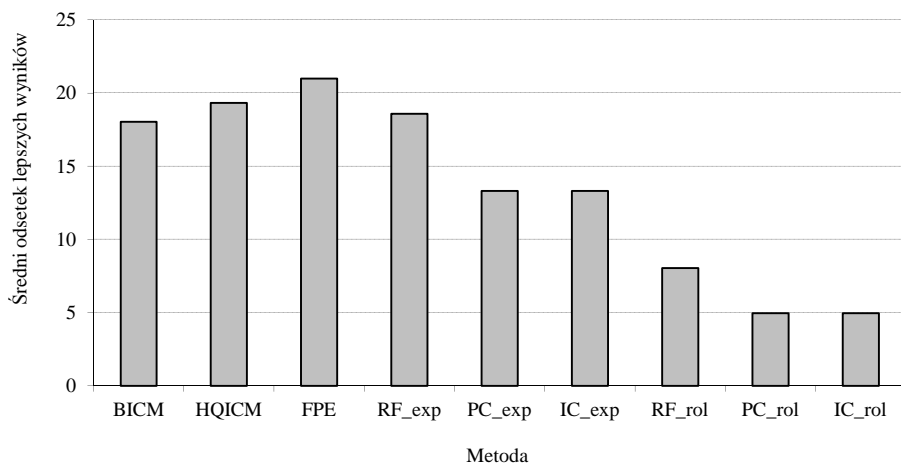
Rys. 8. Ranking metod według horyzontu prognoz – model czysto losowy ($p = 100$, $n = 1000$, $\rho_F = 0$, $\rho_Z = 0$, stałe \mathbf{L})



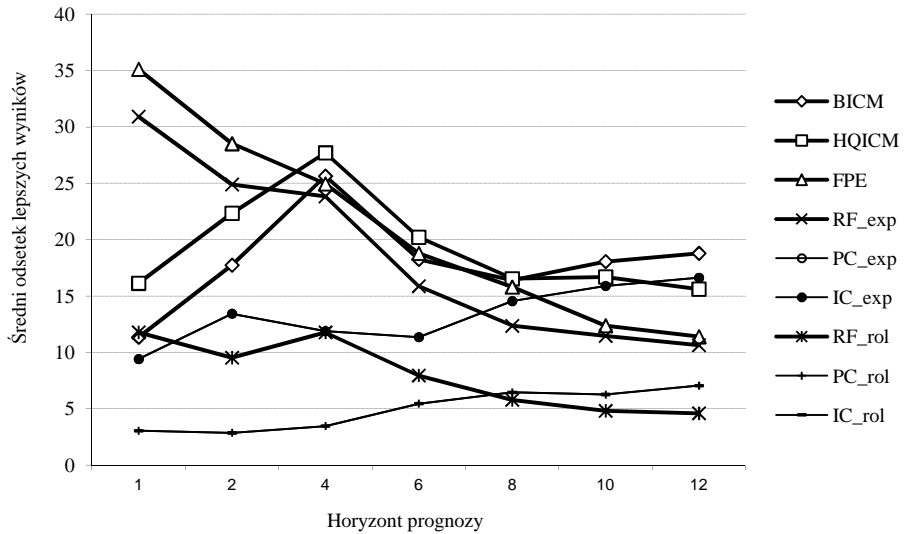
Rys. 9. Ranking metod – błędzenie losowe dla czynników ($p = 100$, $n = 1000$, $\rho_F = 1$, $\rho_Z = 0$, stałe \mathbf{L})

Na rys. 9 przedstawiono ranking metod w sytuacji, gdy dynamika czynników opisana jest procesem błędzenia losowego. Wyniki zbliżone są do uzyskanych w wariancie podstawowym, przy czym teraz najslabiej wypadają kryteria mieszane, zarówno korzystające z okien przesuwanych, jak i rozszerzanych. Były one lepsze od innych podejść średnio w mniej niż 10% przypadków. Wyniki z rozbiem na różne horyzonty nie są prezentowane, gdyż są one zbliżone do tych z rys. 2.

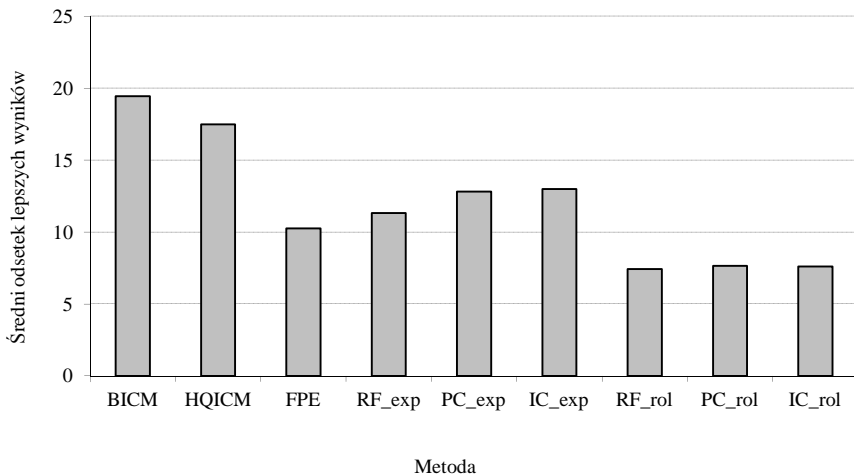
Rysunki 10 i 11 dotyczą modeli, w których zamiast autokorelacji czynników występuje autokorelacja składników losowych w równaniu (13). Należy zwrócić uwagę, że w takiej sytuacji wyniki różnią się w porównaniu do wariantu podstawowego. Najlepsze rezultaty przynosi stosowanie kryterium FPE (ponad 20%). Nieco mniej dają kryteria BICM, HQICM oraz RF_exp. Kryteria, w których stosowane są okna przesuwane dają najgorsze rezultaty. Dokładniejszy wgląd w wyniki daje rys. 11. Widać, że kryteria FPE oraz RF_exp mają zdecydowaną przewagę dla horyzontu $h = 1$. Dla $h = 4$ podobne wyniki dają również kryteria BICM oraz HQICM. Podobnie jak w wariancie podstawowym, dla dłuższych horyzontów poprawia się relatywna jakość prognoz dla modeli korzystających z rozszerzanych okien.



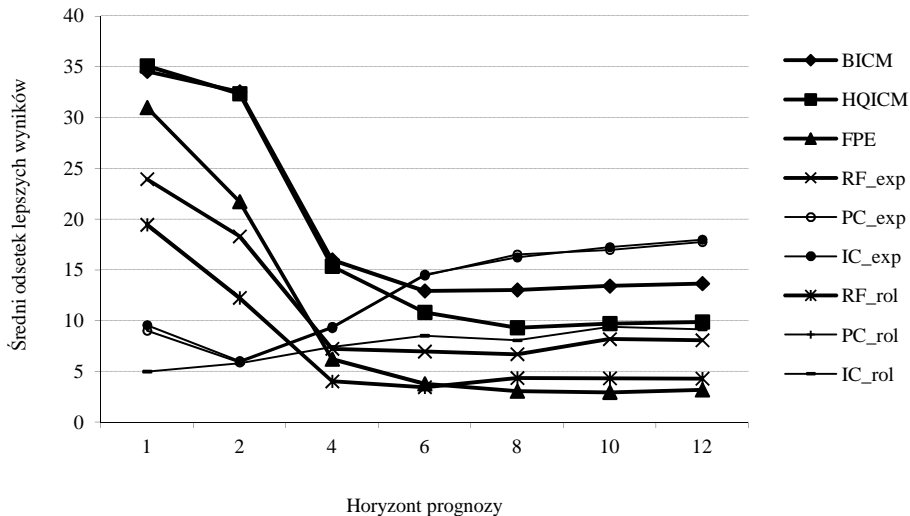
Rys. 10. Ranking metod – autokorelacja zaburzeń losowych ($p = 100$, $n = 1000$, $\rho_F = 0$, $\rho_Z = 0,9$, stałe \mathbf{L})



Rys. 11. Ranking metod według horyzontu prognoz – autokorelacja zaburzeń losowych ($p = 100$, $n = 1000$, $\rho_F = 0$, $\rho_Z = 0,9$, stałe L)



Rys. 12. Ranking metod – zmienna macierz ładunków czynnikowych ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, zmienne L)



Rys. 13. Ranking metod według horyzontu prognoz – zmienna macierz ładunków czynnikowych ($p = 100$, $n = 1000$, $\rho_F = 0,9$, $\rho_Z = 0$, zmienne \mathbf{L})

Na rys. 12 i 13 zaprezentowano wyniki dla zmieniającej się macierzy ładunków czynnikowych zgodnie z równaniem (12), w którym przyjęto $\psi_L = 0,015$ oraz $\rho_L = 0,9$. Widać, że nawet w sytuacji niestabilności procesu generującego dane, metody wykorzystujące analizę prognoz wygasłych z oknami przesuwanymi dają najgorsze wyniki. Ogólnie, rezultaty w takim ujęciu w dalszym ciągu są podobne do wyników uzyskanych dla wariantu podstawowego.

W dalszej części badano własności omawianych metod w odniesieniu do rzeczywistego szeregu inflacji konsumenckiej w Polsce. Szereg ten liczył 130 obserwacji. Na ich podstawie generowano $p = 48$ szeregów czasowych liczących $T = 70$ obserwacji lub więcej w przypadku okien rozszerzanych. Na ich podstawie wybierano najlepsze modele i obliczano wartości prognoz wygasłych oceniając ich dokładność. Dla metody wykorzystujących analizę prognoz wygasłych podokna miały długość co najmniej $T_p = 40$ obserwacji. Wyniki badań zestawiono w tab. 1-4, przy czym nie podawano rezultatów testów CPA, ponieważ w przeważającej liczbie przypadków nie dawały one podstaw do rozróżniania między zdolnościami prognostycznymi analizowanych podejść.

Tabela 1 zawiera błędy średniokwadratowe uzyskane dla prognoz przesuwanych w omawianym podejściu. Dwa najlepsze wyniki dla każdego horyzontu pogrubiono. Dwa najgorsze rezultaty oznaczono kolorem szarym. Podano także wyniki średnie, a w ostatniej kolumnie ranking metod ze względu na

średnie rezultaty. Najmniejsze błędy dawało stosowanie kryteriów BICM oraz HQICM, dobre wyniki także kryterium PC_exp. Najslabiej wypadły kryteria stosujące tylko analizę prognoz wygasłych RF_exp i RF_rol. Widać też dość wyraźnie, że kryteria łączone wykorzystujące okna przesuwane dawały gorsze wyniki niż te, które bazują na oknach rozszerzanych. Należy również podkreślić, że choć kryterium RF_exp daje najlepsze prognozy ze wszystkich metod dla horyzontu $h = 1$, to dla większych horyzontów daje często najgorsze wyniki.

Tabela 1

Błędy średniokwadratowe prognoz inflacji dla prognoz przesuwanych

| Model | $h = 1$ | $h = 2$ | $h = 4$ | $h = 6$ | $h = 8$ | $h = 10$ | $h = 12$ | Średnio | Rank |
|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------|----------|
| BICM | 0,236 | 0,228 | 0,234 | 0,246 | 0,257 | 0,273 | 0,268 | 0,249 | 1 |
| HQICM | 0,234 | 0,234 | 0,235 | 0,246 | 0,257 | 0,272 | 0,268 | 0,249 | 1 |
| FPE | 0,232 | 0,261 | 0,237 | 0,246 | 0,255 | 0,276 | 0,278 | 0,255 | 5 |
| RF_exp | 0,226 | 0,271 | 0,245 | 0,255 | 0,293 | 0,288 | 0,288 | 0,267 | 8 |
| PC_exp | 0,232 | 0,239 | 0,235 | 0,249 | 0,256 | 0,272 | 0,267 | 0,250 | 3 |
| IC_exp | 0,239 | 0,239 | 0,236 | 0,249 | 0,266 | 0,273 | 0,271 | 0,253 | 4 |
| RF_rol | 0,238 | 0,268 | 0,238 | 0,240 | 0,324 | 0,295 | 0,363 | 0,281 | 9 |
| PC_rol | 0,242 | 0,248 | 0,237 | 0,242 | 0,264 | 0,294 | 0,290 | 0,260 | 6 |
| IC_rol | 0,248 | 0,247 | 0,239 | 0,242 | 0,268 | 0,298 | 0,288 | 0,261 | 7 |

Dosyć podobne rezultaty uzyskuje się przy zastosowaniu okien rozszerzanych przy wyznaczaniu kolejnych prognoz, co ilustruje tab. 2. Choć względne różnice między metodami są mniejsze niż w poprzednim przypadku, to wyraźniejszy jest ranking metod. Najlepsze są kryteria BICM, HQICM oraz FPE, nieco gorzej wypadają kryteria korzystające z okien rozszerzanych. Ponownie ostatnie pozycje zajmują metody stosujące okna przesuwane.

Tabela 2

Błędy średniokwadratowe prognoz inflacji dla prognoz rozszerzanych

| Model | $h = 1$ | $h = 2$ | $h = 4$ | $h = 6$ | $h = 8$ | $h = 10$ | $h = 12$ | Średnio | Rank |
|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------|----------|
| BICM | 0,233 | 0,235 | 0,237 | 0,244 | 0,241 | 0,255 | 0,248 | 0,242 | 1 |
| HQICM | 0,237 | 0,240 | 0,237 | 0,244 | 0,241 | 0,255 | 0,248 | 0,243 | 2 |
| FPE | 0,218 | 0,263 | 0,237 | 0,242 | 0,241 | 0,255 | 0,244 | 0,243 | 2 |
| RF_exp | 0,227 | 0,277 | 0,238 | 0,247 | 0,247 | 0,254 | 0,247 | 0,248 | 6 |
| PC_exp | 0,237 | 0,241 | 0,237 | 0,245 | 0,248 | 0,260 | 0,247 | 0,245 | 4 |
| IC_exp | 0,237 | 0,241 | 0,237 | 0,245 | 0,248 | 0,260 | 0,247 | 0,245 | 4 |
| RF_rol | 0,270 | 0,287 | 0,240 | 0,258 | 0,298 | 0,287 | 0,276 | 0,274 | 9 |
| PC_rol | 0,245 | 0,251 | 0,243 | 0,253 | 0,258 | 0,274 | 0,280 | 0,258 | 7 |
| IC_rol | 0,245 | 0,251 | 0,243 | 0,253 | 0,258 | 0,274 | 0,280 | 0,258 | 7 |

W tab. 3 i 4 przedstawiono średnie błędy absolutne prognoz wygasłych. Z tego punktu widzenia najlepsze wyniki dawały kryteria BICM oraz QHICM, a także PC_exp oraz IC_exp. Ponownie zdecydowanie najslabsze wyniki były związane ze stosowaniem podejścia korzystającego z okien przesuwanych. Wyniki były więc ogólnie podobne do rezultatów uzyskanych przy rozważaniu błędów średniokwadratowych.

Tabela 3

Średnie błędy absolutne prognoz inflacji dla prognoz przesuwanych

| Model | $h = 1$ | $h = 2$ | $h = 4$ | $h = 6$ | $h = 8$ | $h = 10$ | $h = 12$ | Średnio | Rank |
|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------|----------|
| BICM | 0,170 | 0,172 | 0,177 | 0,187 | 0,192 | 0,206 | 0,203 | 0,187 | 1 |
| QHICM | 0,173 | 0,175 | 0,178 | 0,188 | 0,192 | 0,206 | 0,203 | 0,188 | 3 |
| FPE | 0,181 | 0,203 | 0,179 | 0,189 | 0,190 | 0,211 | 0,208 | 0,194 | 5 |
| RF_exp | 0,172 | 0,209 | 0,185 | 0,193 | 0,218 | 0,222 | 0,219 | 0,203 | 6 |
| PC_exp | 0,164 | 0,181 | 0,177 | 0,189 | 0,191 | 0,203 | 0,202 | 0,187 | 1 |
| IC_exp | 0,176 | 0,180 | 0,179 | 0,189 | 0,199 | 0,203 | 0,209 | 0,191 | 4 |
| RF_rol | 0,185 | 0,210 | 0,184 | 0,187 | 0,236 | 0,241 | 0,290 | 0,219 | 9 |
| PC_rol | 0,180 | 0,190 | 0,182 | 0,190 | 0,203 | 0,242 | 0,232 | 0,203 | 6 |
| IC_rol | 0,192 | 0,189 | 0,184 | 0,189 | 0,208 | 0,243 | 0,231 | 0,205 | 8 |

Tabela 4

Średnie błędy absolutne prognoz inflacji dla okien rozszerzanych

| Model | $h = 1$ | $h = 2$ | $h = 4$ | $h = 6$ | $h = 8$ | $h = 10$ | $h = 12$ | Średnio | Rank |
|--------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|---------|----------|
| BICM | 0,173 | 0,179 | 0,180 | 0,188 | 0,182 | 0,197 | 0,191 | 0,184 | 1 |
| QHICM | 0,176 | 0,182 | 0,180 | 0,188 | 0,182 | 0,197 | 0,191 | 0,185 | 2 |
| FPE | 0,164 | 0,212 | 0,180 | 0,184 | 0,182 | 0,197 | 0,188 | 0,187 | 5 |
| RF_exp | 0,183 | 0,225 | 0,181 | 0,191 | 0,189 | 0,195 | 0,190 | 0,193 | 6 |
| PC_exp | 0,176 | 0,185 | 0,180 | 0,188 | 0,187 | 0,199 | 0,190 | 0,186 | 3 |
| IC_exp | 0,176 | 0,185 | 0,180 | 0,188 | 0,187 | 0,199 | 0,190 | 0,186 | 3 |
| RF_rol | 0,203 | 0,229 | 0,184 | 0,194 | 0,227 | 0,233 | 0,223 | 0,213 | 9 |
| PC_rol | 0,184 | 0,188 | 0,184 | 0,193 | 0,201 | 0,222 | 0,227 | 0,200 | 7 |
| IC_rol | 0,184 | 0,188 | 0,184 | 0,193 | 0,201 | 0,222 | 0,227 | 0,200 | 7 |

Podsumowanie

Przedstawione w pracy wyniki w dość jednoznaczny sposób wskazują, że najlepsze modele DFM dla celów prognostycznych najczęściej były uzyskiwane na podstawie kryteriów jednoczesnych BICM oraz HQICM zaproponowanych przez Groena i Kapetaniosa. Tylko nieznacznie gorsze rezultaty były związane ze stosowaniem kryterium FPE Bai i Ng. Kolejne miejsca w rankingu przypadły kryteriom bazującym na oknach rozszerzanych przy wyborze najlepszego modelu. Natomiast najslabiej wypadały z reguły metody stosujące okna przesuwane. Spośród metod z dwóch ostatnich grup lepiej wypadały metody, w których proces wyboru najlepszego modelu przebiegał w całości w oparciu o badanie własności prognoz wygasłych. Wyniki wskazywały także, że w przypadku niewielkich horyzontów prognozy różnice pomiędzy metodami identyfikowane za pomocą testów zdolności prognostycznych CPA były z reguły znacznie wyższe niż w przypadku dłuższych horyzontów.

Literatura

1. Alessi L., Barigozzi M., Capasso., *A robust criterion for determining the number of factors in approximate factor models*, ECARES Working Paper 2009/023.
2. Artis M., Banerjee A., Marcellino M., *Factor forecasts for the UK*, „Journal of Forecasting” 2005, Vol. 24.
3. Bai J., Ng S., *Determining the number of factors in approximate factor models*, „Econometrica” 2002, Vol. 70 (1).
4. Bai J., Ng S., *Boosting diffusion indexes*, „Journal of Applied Econometrics” 2009, Vol. 24 (4).
5. Baranowski P., Leszczyńska A., Szafranski G., *Krótkookresowe prognozowanie inflacji z użyciem modeli czynnikowych*, „Bank i Kredyt” 2010, nr 41 (4).
6. Bernanke B., Boivin J., *Monetary policy in data-rich environment*, „Journal of Monetary Economics” 2003, Vol. 50 (3).
7. Breitung J., Eickmeier S., *Dynamic factor models*, Deutsche Bundesbank Discussion Paper Series 1: Economic Studies No 38/2005.
8. Doz C., Giannone D., Reichlin L., *A quasi maximum likelihood approach for large approximate dynamic factor models*, ECB Working Paper 674.
9. Forni M., Hallin M., Lippi M., Reichlin L., *The generalized dynamic factor model: Identification and estimation*, „The Review of Economics and Statistics” 2000, Vol. 82 (4).
10. Giacomini R., White H., *Tests of conditional predictive ability*, „Econometrica” 2006, Vol. 74 (6).
11. Groen J., Kapetanios G., *Model selection criteria for factor-augmented regressions*, Federal Reserve Bank of New York Staff Report No. 363, February 2009.

12. Marcellino M., Stock J., Watson M., *Macroeconomic forecasting in the Euro area: country specific versus Euro wide information*, „European Economic Review” 2003, Vol. 47.
13. Matheson T., *Factor model forecasts for New Zealand*, Reserve Bank of New Zealand DP2005/01.
14. Onatski A., *Testing hypotheses about the number of factors in large factor models*, „Econometrica” 2009, Vol. 77.
15. Otter P., Jacobs J., den Reijer A., *A criterion for the number of factors in a data-rich environment*, Mimeo 2011.
16. Stock J., Watson M., *Diffusion indexes*, NBER Working Paper 6702.

SELECTION CRITERIA FOR FORECASTING DYNAMIC FACTOR MODELS

Summary

The paper compares three groups of methods used for best dynamic factor model selection for forecasting: modified information criteria, methods exclusively based on ex post forecasts analysis and mixed algorithms. It searches for the approach that delivers best out-of-sample forecasts according to mean square error measure. The analysis utilizes both Monte Carlo generated samples as well as real time series used for forecasting consumer inflation in Poland. Results show that best forecasts are obtained from the modified information criteria proposed by Groen and Kapetanios, whereas the methods that employ ex post forecasts from rolling windows usually give the worst predictions.