

GRZEGORZ SZAFRAŃSKI

SYNTETYCZNE WSKAŹNIKI WYPRZEDZAJĄCE W PROGNOZOWANIU INFLACJI W POLSCE^{1,2}

1. WSTĘP

Prognozowanie inflacji jest jednym z podstawowych zadań banków centralnych, szczególnie tych, które realizując strategię bezpośredniego celu inflacyjnego deklarują utrzymywanie tempa zmian cen konsumenta w określonym paśmie wahań³. Takie zainteresowanie władz monetarnych jest konieczne w sytuacji, gdy nominalna i realna sfera gospodarki reagują z opóźnieniem na działanie instrumentów polityki pieniężnej. Tymczasem na inflację oddziałuje bardzo wiele czynników i celem tego badania jest sprawdzenie, które zmienne lub grupy zmiennych niosą z odpowiednim wyprzedzeniem użyteczną informację o zmianach cen koszyka dóbr konsumenta. Zmienne te nazywane są wskaźnikami wyprzedzającymi lub wiodącymi (ang. *leading indicators*).

W artykule sprawdzono, jakie korzyści daje zastosowanie metody wskaźników wyprzedzających do systematycznego, comiesięcznego prognozowania podstawowego wskaźnika inflacji w Polsce tj. indeksu cen towarów i usług konsumpcyjnych (dalej CPI). Postawiono roboczą hipotezę, że za pomocą metod statystycznych (analiza korelacji) można wskazać zmienne, których wahania poprzedzają zmiany inflacji w Polsce. W eksperymencie prognostycznym obejmującym okres od grudnia 2010 r. do listopada 2012 r. zastosowano proste (liniowe i jednorównaniowe) modele regresji, dalej zwane modelami wskaźników wyprzedzających⁴. Do prognoz wskaźnika inflacji w horyzoncie od 1 do 12 miesięcy naprzód wykorzystano jedynie informacje

¹ Artykuł jest znacznie zmienioną wersją artykułu Szafrąński (2011) opublikowanego w roboczej wersji w Materiałach i Studiach NBP. Zmiany dotyczą przede wszystkim części empirycznej tj. celu badania, zastosowanych metod i zakresu próby, a także motywacji i przeglądu literatury. Artykuł nie wyraża ani poglądów ani oficjalnego stanowiska Narodowego Banku Polskiego.

² Podziękowania dla współpracowników NBP za okazaną pomoc, a także dla uczestników seminariów naukowych (IE NBP i Katedry Ekonometrii UŁ) za motywującą dyskusję i dla anonimowego recenzenta za cenne uwagi do ostatecznej wersji tekstu.

³ W Polsce bank centralny stosuje strategię bezpośredniego celu inflacyjnego od 1999 roku.

⁴ W artykule Szafrąński (2011) sprawdzono, że modele te dostarczają na ogół równie dokładnych lub dokładniejszych punktowych prognoz CPI niż takie modele wzorcowe jak np. model oparty o auto-regresję.

dostępne w chwili sporządzania prognoz (tzw. dane w czasie rzeczywistym). Aby ocenić korzyści, jakie może przynieść redukcja informacji pochodzącej z dużego zbioru danych, jakoś prognoz na podstawie wskaźników indywidualnych porównano z prognozami opartymi na wskaźnikach syntetycznych. Wskaźniki syntetyczne zostały określone jako czynniki opisujące wspólną część zmienności grupy zmiennych i uzyskane metodą analizy głównych składowych. Elementem dodanym tego artykułu, w porównaniu do innych badań dotyczących prognozowania inflacji w Polsce na podstawie dużych zbiorów danych, jest sprawdzenie, jaki wpływ na jakość uzyskiwanych prognoz ma wstępny dobór zmiennych do analizy czynnikowej na podstawie kryteriów statystycznych. Celem pośrednim badania jest również ustalenie zawartości informacyjnej wybranych wskaźników z punktu widzenia przewidywania inflacji. Oczywiście ustalenie tego typu relacji ma charakter głównie symptomatyczny, a nie przyczynowo-skutkowy. Jednak poszukiwanie natury tej relacji może mieć duże znaczenie dla określenia źródeł i genezy zmian cen w myśl idei, że jednym z kryteriów potwierdzenia związku przyczynowo-skutkowego pomiędzy zmiennymi jest możliwość uzyskania na jego podstawie dokładnej prognozy.

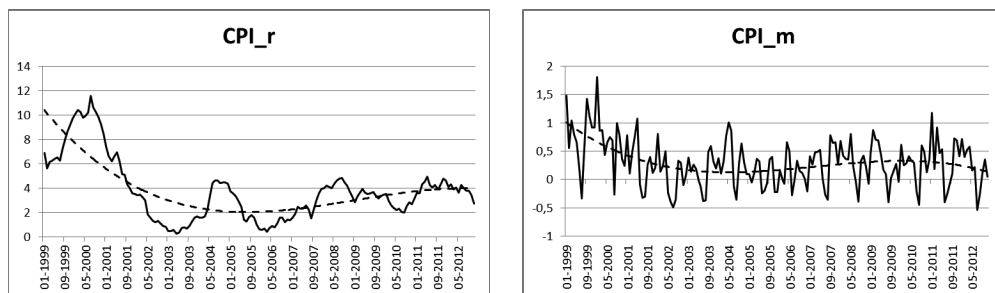
Struktura opracowania jest następująca. W drugiej części przedstawiono fakty dotyczące przebiegu inflacji w Polsce po okresie transformacji i postawiono szczegółowe hipotezy badawcze. Następnie wyjaśniono genezę koncepcji wskaźników wyprzedzających i skomentowano jej wykorzystanie w dotychczasowych badaniach nad prognozowaniem inflacji w Polsce i na świecie. W części trzeciej opisano konstrukcję modeli wskaźników wyprzedzających, a także szczegółowe założenia eksperymentu prognostycznego. Część czwarta zawiera opis konstrukcji indywidualnych i syntetycznych wskaźników wyprzedzających. W części piątej oceniono korzyści wynikające z ich stosowania oraz podjęto próbę oceny ich zawartości ekonomicznej.

2. INFLACJA W POLSCE PO OKRESIE TRANSFORMACJI – MOTYWACJA BADAWCZA

W latach 90. XX w. dwucyfrowa inflacja (w ujęciu rocznym) była w Polsce zjawiskiem powszechnym. Proces stopniowego jej spadku (dezinflacji) postępował wraz z wprowadzaniem nowych, systemowych rozwiązań w polityce pieniężnej. Trwały spadek inflacji poniżej 10% w skali roku i jej ustabilizowanie na średnim poziomie ok. 3% odnotowano dopiero po 2001 roku, a więc w dwa lata po wprowadzeniu strategii bezpośredniego celu inflacyjnego (por. rysunek 1).

Jednak po 2001 roku, wbrew temu, co pokazują statystyki opisowe wskaźnika CPI w ujęciu rok do roku (dalej CPI r/r) zmienność inflacji wcale nie uległa zmniejszeniu (por. tabela 1). Wskaźnik zmian cen w ujęciu miesięcznym (CPI m/m) ma bowiem zbliżoną wariancję w dwóch podpróbach obejmujących lata 2001–2006 i 2007–2012. Ponadto porównując wskaźniki publikowane ze wskaźnikami odsezonowanymi za pomocą procedury Tramo/Seats można spostrzec, że wskaźnik CPI m/m jest w dużym stopniu pod wpływem wahań sezonowych, które w latach 2001–2012 stanowiły

prawie 55% wariancji tego wskaźnika⁵. Zatem po wyłączeniu w miarę regularnych wahań sezonowych zmienność, która pozostaje do objaśnienia za pomocą modeli ekonometrycznych, mierzona odchyleniem standardowym, jest rzędu 0,25 pp. wobec ok. 0,36 pp. całkowitej zmienności inflacji w skali miesiąca (por. wyniki w tabeli 1). Wskazuje to, jak ważne, gdy badamy zmienność cen, jest poprawne uwzględnienie zmian o charakterze sezonowym.



Rysunek 1. Wskaźniki inflacji (CPI) w ujęciu rocznym (lewy panel) i miesięcznym (prawy panel) za okres 1999–2012 (w pp.) na tle nieliniowego trendu (linia przerywana)

Źródło: GUS, obliczenia własne.

Tabela 1.

Statystyki opisowe wskaźników inflacji (CPI) w ujęciu rocznym (w pp.), miesięcznym i po odsezonowaniu w latach 2001–2012 i w dwóch podokresach

Okres	statystyki opisowe	Indeks CPI r/r	Indeks CPI m/m	Odsezonowany CPI m/m	Wskaźnik sezonowości
2001–2012	Średnia	2,967	0,230	0,230	-0,000
	Odchylenie standardowe	1,540	0,363	0,246	0,267
2001–2006	Średnia	2,479	0,171	0,173	-0,001
	Odchylenie standardowe	1,884	0,358	0,264	0,268
2007–2012	Średnia	3,468	0,289	0,221	0,000
	Odchylenie standardowe	0,864	0,361	0,211	0,267

Wskaźniki sezonowości uzyskane w pakiecie Demetra+ 4.3 przy użyciu Tramo/Seats. Źródło: obliczenia własne.

⁵ Przy założeniu, że wskaźnik sezonowości nie jest skorelowany ze wskaźnikiem CPI wyrównanym sezonowo.

Pomimo, że zmienność inflacji wyrównanej sezonowo jest stosunkowo mała, to jej źródła nie są jednorodne. Potwierdzeniem tego może być porównanie częstości występowania nietypowych obserwacji w CPI ogółem i jego komponentach wykrytych w automatycznej procedurze Tramo. O ile w przebiegu szeregu CPI ogółem obserwacje nietypowe zdarzają się stosunkowo rzadko, to w jego komponentach występują one o wiele częściej⁶. Pomimo, że wpływ odstających obserwacji w komponentach jest wygładzany w indeksie CPI ogółem za pomocą wag, to nadal te nietypowe obserwacje mogą mieć w pojedynczych okresach istotny wpływ na kształtowanie się całego agregatu. Dodatkowo skład koszyka dóbr i usług nabywanych przez gospodarstwa domowe podlega co roku zmianom ocenianym na podstawie badań budżetów gospodarstw domowych. Wszystkie wymienione tu elementy dotyczące pomiaru inflacji w Polsce składają się na obraz podstawowych problemów, jakie można napotkać w systematycznym prognozowaniu wskaźnika CPI.

Na podstawie tego krótkiego przeglądu zmienności wskaźnika cen konsumpcyjnych po 2001 roku postawiono następującą hipotezę. Ze względu na wiele potencjalnych źródeł zmienności cen koszyka konsumenta uwzględnienie dużego zbioru informacji jest konieczne, żeby móc systematycznie prognozować inflację w sposób wiarygodny i trafny. Jednak w klasycznej analizie regresji na podstawie liniowych modeli (np. modelu indywidualnych wskaźników wiodących) oczekiwany błąd prognozy jest proporcjonalny do ilorazu liczby zmiennych objaśniających (predyktorów) i liczby obserwacji (N/T) – por. Stock i Watson (2006). Gdy liczba wybranych predyktorów przekracza kilkanaście zmiennych i jest niewiele większa od liczby obserwacji, błąd estymatora MNK może mieć znaczący wpływ na oczekiwany błąd prognozy. Ponadto powiększanie liczby użytych w modelu wskaźników indywidualnych może poprawiać dopasowanie modelu w próbie kosztem pogorszenia jakości prognoz poza próbą. W tym artykule sprawdzono, czy praktycznym rozwiązaniem problemu zbyt dużej liczby regresorów nie może być utworzenie na podstawie tych zmiennych jednej lub kilku zmiennych syntetycznych opisujących jak największą część zmienności szeregów w danej grupie. W badaniu empirycznym syntezy informacji dokonano za pomocą metod analizy czynnikowej. Jeżeli podejście to jest zasadne, to wartość predykcyjna syntetycznych wskaźników wyprzedzających powinna okazać się relatywnie większa niż wskaźników indywidualnych.

⁶ W automatycznej procedurze TRAMO dla CPI ogółem znaleziono zaledwie dwie obserwacje odstające w okresie 2001–2012 (tj. sierpień 2000 roku i maj 2001 roku) wobec przynajmniej 9 tego typu wystąpień w takich komponentach CPI. Najwięcej w takich grupach dóbr jak: pieczywo i produkty zbożowe, nabiał, cukier, odzież, obuwie oraz w cenach usług edukacyjnych.

3. WSKAŹNIKI WYPRZEDZAJĄCE INFLACJI – PRZEGLĄD BADAŃ⁷

Metodę konstrukcji syntetycznych wskaźników wyprzedzających (ang. *leading indicators*) opracowano w amerykańskiej instytucji badawczej National Bureau of Economic Research (NBER) na potrzeby monitorowania i prognozowania stanu koniunktury ogólnogospodarczej (por. Burns i Mitchell, 1946). Wybierano w tym celu wskaźniki, które indywidualnie wykazywały największą zdolność do objaśnienia zmian przyszłej aktywności, a następnie nadawano im arbitralne wagi w mierniku syntetycznym. Usystematyzowaniem tej koncepcji zajęli się Stock i Watson (1989), którzy identyfikację wskaźników wyprzedzających oparli na zdolności prognozowania zmian szeregu referencyjnego (np. koniunktury czy inflacji). Przy czym przez szereg referencyjny rozumiano zmienną nieobserwowalną wyznaczaną na podstawie zmian wspólnych dla większości obserwowalnych aproksymant tego procesu⁸. Wspólne czynniki określane są „siłami sprawczymi” (ang. *driving forces*), a popularnym sposobem ich uzyskiwania są metody analizy czynnikowej wprowadzone do makroekonomii przez Geweke’a, Sargenta i Simsa w latach 70. XX wieku (za Stock i Watson, 2002a).

W latach 90. metody analizy czynnikowej przeprowadzanej na podstawie dużych zbiorów danych wykazały przydatność w prognozowaniu procesów makroekonomicznych. W 1998 roku Stock i Watson jako jedni z pierwszych⁹ do wyznaczenia syntetycznego wskaźnika wyprzedzającego inflacji użyli metod analizy czynnikowej, a w 1999 r. zastosowali je do modelowania krzywej Philipsa. W następnych latach liczba badań wykorzystujących metody analizy czynnikowej w prognozowaniu inflacji szybko rosła, a wśród nich warto wymienić artykułu Angelini i in. (2001), Stocka i Watsona (2003), Banerjee i in. (2005). W ramach tych badań pojawiają się dyskusje na temat optymalnych rozmiarów zbiorów danych do analizy czynnikowej (Boivin, Ng, 2006), zgodnych metod estymacji modeli czynnikowych (por. Stock i Watson, 2002a, Forni i in., 2005, Doz i in., 2006), sposobów wykorzystania dynamicznego modelu czynnikowego w prognozowaniu (por. Boivin i Ng, 2005) oraz sposobu postępowania w sytuacji asynchronicznego napływu danych ekonomicznych (w polskiej literaturze por. Łupiński, 2012).

W dotychczasowych badaniach statystycznych nad wskaźnikami wyprzedzającymi inflacji analizowano najczęściej następujące grupy zmiennych: komponenty inflacji (Hendry i Hubrich, 2006; Kapetanios, 2004), oraz konstruowane na ich podstawie syntetyczne miary inflacji bazowej określane jako *core inflation* (Giannone i Matheson, 2007; Cristadoro i in., 2005), czy *pure inflation* (Reis i Watson, 2010, a dla Polski artykuł Brzozy-Brzeziny i Kotłowskiego, 2009), oczekiwania inflacyjne

⁷ Ta część jest skróconą i uaktualnioną wersją przeglądu literatury zawartego w Szafranski (2011).

⁸ Oprócz części wspólnej każdy szereg aproksymant szeregu referencyjnego zawiera zmiany swoiste zwane idiosynkratycznymi.

⁹ Ostateczną wersję pracy opublikowano w 2002 roku (Stock i Watson, 2002b).

osób prywatnych i podmiotów gospodarczych (por. Ang i in., 2007, dla Polski Łyziak i Stanisławska, 2008), mierniki koniunktury stosowane w różnych wariantach krzywej Philipsa (por. ich przegląd w pracy Stocka i Watsona, 2008), miary podaży pieniądza (Hofmann, 2008), czy informacje z rynku finansowego o cenach i rentowności aktywów finansowych, a także czynnikach terminowej struktury stóp procentowych (Stock i Watson, 2003).

Badania dotyczące regularnego prognozowania inflacji w Polsce na podstawie metody wskaźników wyprzedzających nie są zbyt liczne. Na podstawie tradycyjnej metody NBER w ramach prac instytutu badawczego Bureau for Investment and Economic Cycles (BIEC) powstaje syntetyczny wskaźnik wyprzedzający dla inflacji w Polsce, tak zwany Wskaźnik Przyszłej Inflacji (WPI) – por. Białowolski i Żochowski (2006)¹⁰. Słabością tego podejścia jest opieranie się na ustalanych i zmienianych arbitralnie wagach. Z kolei podejścia, w których autorzy korzystają w sposób systematyczny z relatywnie dużych zbiorów danych można znaleźć w ramach badań porównawczych nad prognozowaniem zharmonizowanego indeksu inflacji (HICP) dla grupy 5 krajów (Polski, Czech, Węgier, Słowacji i Słowenii) wstępujących do Unii Europejskiej (UE) w 2004 roku – zob. Banerjee i in. (2004) i dla większej grupy nowych członków UE w artykule Arratibel i in. (2009). W tym pierwszym badaniu stosowano modele ze wspólnymi czynnikami uzyskanymi na podstawie kilkudziesięciu kwartalnych szeregów czasowych (56 dla Polski), a w tym drugim tradycyjne modele z jednym lub dwoma indywidualnymi wskaźnikami wyprzedzającymi wykazując ich wyższość nad prostymi metodami analizy szeregów czasowych.

W Polsce badania inflacji prowadzone na podstawie dynamicznych modeli czynnikowych i dużych zbiorów danych pojawiły się stosunkowo niedawno – por. artykuły Kotłowski (2008), Baranowski, i in. (2010). W pracach tych podobnie jak w tym opracowaniu do wyznaczenia wspólnych czynników wykorzystano metodę analizy głównych składowych (ang. *principal component analysis*, dalej PCA), której własności asymptotyczne (gdy liczba stacjonarnych szeregów czasowych oraz ich obserwacji po czasie rośnie do nieskończoności) opisano m.in. w Stock i Watson (2002a), Bai (2003). Warto wspomnieć, że w wyniku zastosowania metody PCA dochodzi do redukcji liczby zmiennych objaśniających w równaniu prognostycznym. Ze względu na potencjalną współliniowość niektórych regresorów redukcja ta powinna mieć wpływ na uzyskanie precyzyjniejszych ocen parametrów. Jednakże głównym celem zastosowania metody PCA w analizie czynnikowej dużych zbiorów danych jest wydobywanie w sposób syntetyczny informacji o wspólnych czynnikach, która może być użyteczna w prognozowaniu inflacji.

W przeprowadzonym badaniu wstępnie ograniczono liczbę zmiennych przeznaczonych do analizy czynnikowej metodami statystycznymi. Redukcja ta nie jest jedy-

¹⁰ Wskaźnik WPI zawiera jako komponenty następujące zmienne: kursy walutowe, indeksy cen importu i usług, zadłużenie sektora publicznego i osób prywatnych, jednostkowe koszty pracy, oraz dane ankietowe z testów koniunktury przemysłu przetwórczego IRG-SGH.

nie zabiegiem czysto praktycznym, ułatwiającym ekonomiczną interpretację wyników eksperymentu. Jest to również przyczynek do obecnej w literaturze dyskusji na temat statystycznych metod doboru zmiennych do analizy czynnikowej. Metody te, znane pod nazwą metod „najlepszych predyktorów” (tłum. własne ang. *targeted predictors*, termin użyty po raz pierwszy w artykule Bai i Ng, 2008), wskazują na korzyści z pominięcia w analizie czynnikowej szeregów o dużej zmienności idiosynkratycznej (swoistej, własnej). Ich obecność może bowiem znacząco obniżyć dokładność, z jaką wyznacza się nieobserwowalne wspólne czynniki (por. Boivin i Ng, 2006).

4. EKSPERYMENT PROGNOSTYCZNY

4.1. DANE I ICH PRZYGOTOWANIE

Do przeprowadzenia badania wykorzystano bazę danych zawierającą te informacje makroekonomiczne i finansowe, które w okresie od lipca 2009 r. do grudnia 2012 r. były co miesiąc publicznie dostępne w dniu publikacji wskaźnika CPI przez GUS¹¹. Daje to łącznie 42 generacje danych, w których najnowsze opublikowane dane często podlegały zmianom i rewizjom statystycznym, a niektóre z nich uzyskiwano na podstawie informacji za bieżący (niepełny) miesiąc¹². W skład bazy wchodzi najbardziej aktualne obserwacje dotyczące realizacji 189 miesięcznych szeregów czasowych, z których najstarsze sięgają stycznia 2000 roku. Niektóre z szeregów charakteryzują się późniejszym terminem publikacji niż termin ogłaszania wskaźnika CPI, czyli tzw. opóźnieniem publikacyjnym, które sięga od 1 do kilku miesięcy. Inne, w tym dane z rynków finansowych, paliwowych i rolnych, dotyczą miesiąca następnego w stosunku do opublikowanego wskaźnika CPI. Przy przyjętej metodzie prognozowania (por. 4.2) jako zasadę przyjęto, że dane z okresu t , to dane dostępne w okresie t , a zatem nie zawsze opisują one sytuację gospodarczą w okresie t . W skład bazy wchodzi zmienne, które przyporządkowano do następujących, jednorodnych tematycznie grup danych:

- komponenty wskaźnika CPI (grupa CPI i ZYWU),
- czynniki dotyczące kreacji pieniądza (PIEN),
- czynniki determinujące koszty produkcji i dystrybucji dóbr (z podziałem na krajowe KRAJ i zagraniczne ZAGR),
- wskaźniki (tzw. „twarde dane” GUS) dotyczące sytuacji ogólnogospodarczej (popytowe, GUS),

¹¹ Inflacja za miesiąc poprzedni jest publikowana co do zasady około 15 dnia bieżącego miesiąca. Ze względu na układ dni wolnych w danym roku termin ten nie jest stały i podawany jest co roku przez GUS w kalendarium publikacji danych.

¹² Dotyczy oszacowań wskaźników średniomiesięcznych, które uzyskiwano na podstawie danych dostępnych w chwili prognozowania jedynie za pierwszą połowę miesiąca, a w szczególności danych dziennych z rynków finansowych oraz danych tygodniowych z rynku paliw i surowców rolnych.

- informacje dotyczące oczekiwań (w tym ankietowe wskaźniki koniunktury KON, oczekiwania inflacyjne osób prywatnych CiH, oraz wskaźniki dochodowości na rynku finansowym RF).

Zbiornicze informacje o zawartości bazy danych zawarto w Załączniku. Wszystkie potencjalne wskaźniki indywidualne, poddano rekurencyjnie procedurze odsezonowania (o ile metoda Tramo/Seats wskazała na obecność wahań sezonowych). Następnie wyrównane sezonowo szeregi poddano takim transformacjom, żeby w ich wyniku otrzymać szeregi stacjonarne. Szczegóły dotyczące źródeł, odsezonowania i transformacji danych przedstawiono w Załączniku.

4.2. PRZEDMIOT PROGNOZY I METODA PROGNOZOWANIA

Zmienną objaśnianą we wszystkich analizowanych modelach wskaźników wyprzedzających są wskaźniki inflacji CPI w ujęciu miesiąc do miesiąca (m/m). W badaniu wykorzystano metodę prognozowania bezpośredniego¹³ (ang. *direct forecasting*), czyli dla prognozy wyznaczanej na h miesięcy naprzód ($h = 1, \dots, 12$) konstruowany jest model prognostyczny wykorzystujący jedynie ostatnie dostępne obserwacje o zmiennych objaśniających (tj. sprzed h miesięcy w stosunku do ostatnio opublikowanej inflacji). Prognozy oparte na liniowych modelach wskaźników wyprzedzających (LI) przyjmują następującą postać ogólną:

$$\hat{\pi}_{t+h|t} = \sum_{i=1}^I \alpha_i^{(h)} LI_{it}^{(h)} + sez_t, \quad (1)$$

gdzie:

$\hat{\pi}_{t+h|t}$ – prognoza wskaźnika inflacji CPI m/m (okres poprzedni=1) na okres $t + h$ uzyskana na podstawie informacji dostępnej do okresu t ,

$LI_{i,t}^{(h)}$ – „ i ”-ty wiodący (wyprzedzający) wskaźnik inflacji wyznaczany oddzielnie dla każdego z horyzontów prognoz ($h = 1, \dots, 12$) wg metod opisanych w pkt 4.3 i 4.4,

$\alpha_i^{(h)}$ – ocena parametru uzyskana metodą najmniejszych kwadratów (MNK) dla wskaźnika indywidualnego lub wspólnego czynnika ($i = 1, \dots, I$),

sez_t – oznacza estymowane łącznie ze wskaźnikami wyprzedzającymi deterministyczne efekty sezonowe, co *de facto* sprowadza się do uzmiennienia wyrazu wolnego w każdym miesiącu.

¹³ Z przyjęcia metody bezpośredniej w prognozowaniu wynika określenie osobnych wskaźników i osobnych modeli na każdy horyzont prognozy (zob. Stock i Watson, 2008).

4.3. INDYWIDUALNE WSKAŹNIKI WIODĄCE

W modelu wskaźników indywidualnych wszystkie odsezonowane zmienne z bazy danych potraktowano jako potencjalnych kandydatów na wskaźniki wiodące. Przyjęto oznaczenia $LI_{i,t}^{(h)} = \tilde{x}_{j(h),t}$, gdzie $j(h)$ oznacza kolejność uporządkowania indeksów zmiennych $j = 1, 2, \dots, 189$ dla horyzontu prognozy h ($h = 1, \dots, 12$). Rangowania zmiennych w $j(h)$ dokonano „w próbie” na podstawie kryterium jakości dopasowania równania (1) z jedną zmienną objaśniającą ($I = 1$), oddzielnie dla każdego z horyzontów prognozy. Ze względu na potencjalną niestabilność wyboru najlepszych predyktorów dokonywanego na podstawie kolejnych generacji danych ranking utworzono na podstawie okresu testowego obejmującego 17 generacji danych, przy czym pierwsza generacja zawiera dane dostępne w lipcu 2009 r., a ostatnia w listopadzie 2010 r. W ten sposób zmiennym przypisano średnie pozycje w rankingu, które następnie uporządkowano do postaci rankingu wskaźników $j(h)$ wyznaczonego oddzielnie na każdy horyzont prognozy. W tabeli 2 widoczne są trzy pierwsze pozycje każdego z tych rankingów, a syntetyczne informacje o pełnym rankingu 20 najlepszych wskaźników zaprezentowano w podziale na grupy zmiennych w tabeli 3. W celu sprawdzenia zdolności prognostycznej wybranych modeli „poza próbą” do modelu prognostycznego (1) włączane są kolejne zmienne od pozycji 1 do pozycji 30 określonej na podstawie tego rankingu.

Tabela 2.

Rankingi 3 „najlepszych w próbie” indywidualnych wskaźników wyprzedzających na horyzonty od 1 do 12 miesięcy (na podstawie generacji danych: lipiec 2009 – grudzień 2010 r.)

Rankingi	Pozycja 1		Pozycja 2		Pozycja 3	
	Nazwa grupy / symbol zmiennej	średnia pozycja	Nazwa grupy / symbol zmiennej	średnia pozycja	Nazwa grupy / symbol zmiennej	średnia pozycja
h=1	CPI bazowa 15%	1,0	CPI Wyposaż. mieszkań	2,4	CPI ogółem (SA)	3,8
h=2	CPI bazowa 15%	1,4	CiH IPSOS3	2,2	KRAJ Mleko MRol	5,1
h=3	CiH IPSOS4	2	CiH IPSOS1	4,0	CiH IPSOS3	4,1
h=4	CiH IPSOS1	1,6	CiH IPSOS4	2,0	CiH IPSOS3	4,7
h=5	CiH IPSOS3	1,4	CiH IPSOS1	3,3	CPI Opał	5,6
h=6	CiH IPSOS3	1,8	CiH IPSOS4	4,8	CPI Usługi medyczne	6,9

h=7	ZYWU Nabiał	1,5	RF Stopa 12M	2,3	RF Stopa 3M	4,3
h=8	ZYWU Nabiał	1,0	RF Stopa 12M	2,0	RF Stopa 2Y	3,1
h=9	RF Stopa 12M	1,1	RF Stopa 3M	2,5	RF Stopa 2Y	3,2
h=10	RF Stopa 10Y	1,1	ZYWU Napoje bezalk.	2,5	RF Stopa 2Y	2,6
h=11	ZYWU Napoje alk.	1,1	ZAGR MFW oliwa	2,8	RF WIG Telekom	2,9
h=12	WWG Bezrob. rejestr.	1,3	RF WIG Telekom	4,2	RF WIG 20	4,9

Średni ranking ustalono na podstawie rankingów dla 18 generacji danych w okresie testowym utworzonych w oparciu o jakość dopasowania modelu (1) dla wskaźnika indywidualnego (λ). Oznaczenia niektórych zmiennych; w grupie wskaźników CPI i ZYWU (zawierającej wszystkie odsezonowane i w ujęciu m/m); „bazowa 15%” – inflacja bazowa „15% średnia obciąża”, ogółem (SA) – wyrównany sezonowo wskaźnik CPI ogółem, pozostałe grupy CPI zgodne z opisem tworzone na podstawie jednej lub kilku grup COICOP, w grupie CiH; IPSOSx – udziały odpowiedzi na pytanie „x” w ankiecie dot. oczekiwań inflacyjnych osób prywatnych IPSOS, w grupie KRAJ; Mleko MRol– mleko spożywcze pasteryzowane (źródło: tygodniowy monitoring rynków rolnych Ministerstwa Rolnictwa i Rozwoju Wsi, dalej MRol), ZAGR; MFW – ceny surowców na rynkach światowych, RF; stopy dochodowości na rynku finansowym (na wybrane terminy), oraz dochodowości indeksów giełdowych (WIG). Oznaczenia nazw grup zmiennych i źródło danych zob. Załącznik. Źródło: obliczenia własne.

Tabela 3.

Udział procentowy zmiennych pochodzących z grup prognostycznych w rankingu 20 wskaźników indywidualnych najsilniej skorelowanych z przyszłą inflacją w horyzoncie od 1 do 12 miesięcy

Grupy	h=1	h=2	h=3	h=4	h=5	h=6	h=7	h=8	h=9	h=10	h=11	h=12
CiH	15	15	15	20	20	15	15	0	5	5	5	0
PiB	0	15	5	0	0	0	0	5	0	5	5	10
RF	0	5	0	0	10	10	40	35	35	25	35	30
KRAJ	30	20	30	20	5	5	0	15	15	0	10	10
ZAGR	0	0	0	5	5	25	0	0	0	20	30	15
WWG	0	0	0	0	0	0	10	5	5	10	5	10
ZWG	0	0	0	0	0	5	0	0	0	0	0	0
PIEN	0	5	20	20	30	20	15	0	5	10	5	0
CPI	40	15	10	15	20	15	10	15	15	15	0	20
ZYWN	15	25	20	20	10	5	10	25	20	10	5	5

Źródło: obliczenia własne.

4.4. SYNTETYCZNE WSKAŹNIKI WIODĄCE

Wskaźniki syntetyczne gromadzą informację zawartą w zmiennych występujących w rankingu na pozycji 1 do R ($R = 5, 10, 15, 20, 25, 30$) skondensowano do jednego lub kilku wspólnych czynników za pomocą metody PCA:

$$LI_{i,t}^{(h)} = \sum_{j=1}^R \hat{\lambda}_j^{(h)} \tilde{x}_{j(h),t}, \quad (2)$$

gdzie $\hat{\lambda}_j$ jest współczynnikiem stojącym przy zmiennej o numerze i w wektorze własnym odpowiadającym j -tej największej wartości własnej macierzy kowariancji wystandaryzowanych zmiennych $\tilde{x}_{j(h),t}$. Wspólne czynniki wyznaczone na podstawie formuły (2), oddzielnie na każdy horyzont $h = 1, 2, \dots, 12$, wykorzystywane są następnie jako wskaźniki wiodące w równaniu (1).

Uwaga: w modelu indywidualnych wskaźników wyprzedzających uwzględniono następujące warianty dotyczące liczby wskaźników wyprzedzających $I = 1, 2, 3, 5, 10, 15, 20, 25, 30$. Stąd dla tych modeli na rysunku 2 i rysunku 3 oraz w tabeli 4 przyjęto oznaczenia: 1 zm, 2 zm, ..., 30 zm. Natomiast w modelu syntetycznych wskaźników wyprzedzających parametr $I = 1, 2, 3$ oznacza liczbę uwzględnionych w modelu wspólnych czynników uzyskanych metodą analizy głównych składowych spośród 5, 10, 15 lub 20 zm (np. 2 PCA-5 oznacza model z dwoma wspólnymi czynnikami wyznaczony na podstawie 5 najlepszych predyktorów).

4.5. ZAŁOŻENIA EKSPERYMENTU PROGNOSTYCZNEGO

Eksperyment polegał na wykonaniu prognoz inflacji na h miesięcy naprzód¹⁴ ($h = 1, 2, \dots, 12$) każdorazowo na podstawie oddzielnej generacji danych dostępnych co miesiąc w okresie od grudnia 2010 – do listopada 2012 r. Ze względu na to, że przedmiotem zainteresowania podmiotów gospodarczych jest tempo zmian cen w ujęciu rok do roku (r/r) wyniki prognoz punktowych zaprezentowano po ich przeliczeniu na wskaźniki r/r . Prognozy inflacji r/r są następnie porównane z realizacjami za pomocą syntetycznych miar błędów prognoz RMSE, ME i MAE, tj. odpowiednio pierwiastka błędu średniokwadratowego, błędu średniego i błędu absolutnego. Parametry modeli progностycznych (1) reestymowano metodą najmniejszych kwadratów z wykorzystaniem jedynie danych dostępnych do okresu t oraz stałego okna estymacji (ang. *rolling*

¹⁴ Uwaga: prognozowanie na jeden okres naprzód ($h = 1$), zwane *nowcastingiem*, dotyczy prognozy inflacji na bieżący miesiąc. Określenia okresów naprzód użyto jedynie w kontekście dostępności danych, a nie w odniesieniu do bieżącego okresu, w którym te dane są publikowane.

window)¹⁵. Przyjęty w eksperymencie dla wszystkich metod układ badania jakości prognoz i dostępna baza danych pozwala na weryfikację prognoz *ex post* inflacji rok do roku kolejno na podstawie: 24 obserwacji dla prognoz na $h = 1$, dwudziestu dwóch – na $h = 2$, ..., i dwunastu – na $h = 12$.

5. WYNIKI EMPIRYCZNE

Jako podstawowe kryterium dokładności prognozy punktowej, w opisanym w części 4.5 okresie weryfikacji obejmującym lata 2011–2012, przyjęto pierwiastek błędu średniokwadratowego (RMSE) liczony oddzielnie dla każdego horyzontu prognoz. W celu porównania prognoz z różnych modeli liczony jest również relatywny spadek błędu RMSE (oznaczający poprawę jakości prognoz) w stosunku do prognozy bazowej wyrażony w %. Pomocniczo analizowano również obciążenie prognoz za pomocą miary ME i relacji ME/MAE (dalej określanej jako stopień obciążenia prognoz¹⁶). Jakość prognoz sprawdzono pod kątem wpływu na nią liczby wiodących wskaźników indywidualnych i liczby wspólnych czynników użytych w równaniu (1), oraz obydwu tych elementów jednocześnie.

5.1. MODELE WSKAŹNIKÓW INDYWIDUALNYCH W PRÓBIE

Ranking zmiennych najsilniej skorelowanych z przyszłą inflacją w horyzoncie od 1 do 12 miesięcy naprzód okazał się być relatywnie stabilny w kolejnych 17 generacjach danych poprzedzających okres weryfikacji prognoz. Zmienne najsilniej skorelowane pozostawały na najwyższych pozycjach rankingu przez kilkanaście miesięcy¹⁷, na co wskazują średnie pozycje z kolejnych miesięcy dla 3 pierwszych zmiennych – por. tabela 2. Analizując zawartość ekonomiczną zmiennych dla poszczególnych horyzontów prognoz na pozycjach 1–3 w rankingu (zob. tabela 2) i grup zmiennych w rankingu 20 zmiennych najsilniej skorelowanych z przyszłą inflacją (zob. tabela 3) można wyciągnąć następujące wnioski:

1. W prognozie wskaźnika inflacji za bieżący miesiąc tj. dla $h=1$ w modelu (1) najbardziej przydatne są ostatnie informacje pochodzące ze wskaźników inflacji bazowej, inflacji CPI i jej nieżywnościowych komponentów, co może być oznaką

¹⁵ Stałe okno estymacji zastosowano między innymi w celu złagodzenia konsekwencji potencjalnie zmiennego wzorca sezonowości i wpływu na wyniki estymacji obserwacji pochodzących z początku próby.

¹⁶ Relacja tych miar błędów jest zawsze z przedziału $(-1,1)$. Jej graniczne wartości świadczą o systematycznym przeszacowaniu (dolna granica przedziału) lub systematycznym niedoszacowaniu (górną granicą) zmiennej prognozowanej.

¹⁷ Również na kolejnych pozycjach wybór zmiennych wykazywał się dużą stabilnością. Przykładowo wszystkie wskaźniki indywidualne do 7 pozycji w rankingu uzyskiwały w kolejnych miesiącach pozycję nie gorszą niż 10, a do 10 pozycji – nie gorszą niż 15.

- uporczywości (ang. *persistence*) inflacji rozumianej jako tendencja do utrzymywania się efektów impulsowego szoku dotyczącego cen w kolejnym miesiącu (por. Hertel, Leszczyńska 2013). W rankingu 20 najsilniej skorelowanych wskaźników indywidualnych dla *nowcastingu* indeksu CPI i jego komponenty stanowią ponad 55% wybranych zmiennych, a 40% po wyłączeniu żywnościowych komponentów CPI – grupa ZYWU. Ważną grupą prognostyczną (30% udział) są również wskaźniki związane ze zmianą kosztów dóbr konsumpcyjnych na rynku krajowym (KRAJ).
2. Dla prognoz od 2 do 6 miesięcy naprzód na najwyższych pozycjach w rankingu dominują informacje pochodzące z badań ankietowych konsumentów dotyczące ich oczekiwań inflacyjnych (IPSOS). Wśród wskaźników na 20 najwyższych pozycjach stanowią one (wraz z badaniami ankietowymi cen i handlu – grupa CiH) od 15% do 20%. W rankingu 20 wskaźników występują ponadto komponenty CPI, choć ich udział maleje wraz z horyzontem prognoz (od 40% dla $h=2$ do 20% dla $h=5$ i $h=6$). Podobnie wraz z wydłużeniem horyzontu prognozy tracą na znaczeniu informacje o kosztach krajowych (KRAJ np. informacje o cenach skupu produktów rolnych), a wzrasta znaczenie danych dotyczących podaży pieniądza (PIEN).
 3. Dla prognoz od 7 do 12 miesięcy naprzód najliczniej występującą grupą wskaźników indywidualnych (od 25% do 40% wskaźników) są dane pochodzące z rynku finansowego (RF, w tym przede wszystkim stopy procentowe z rynku finansowego, a na 11 i 12 miesięcy również stopy dochodowości rynku akcji). Kolejne grupy zmiennych w rankingu dotyczą kosztów dóbr konsumpcyjnych wytwarzanych w kraju (KRAJ) i zagranicą (ZAGR). W tej grupie prognoz koszty dóbr importowanych mają znaczenie tylko dla prognoz na 10–12 miesięcy naprzód. Wśród pozostałych wskaźników wyprzedzających w rankingu występują też komponenty CPI nieżywnościowe (CPI) i te związane z żywnością i napojami (ZYWU). Udział tej drugiej grupy wyraźnie maleje wraz z wydłużeniem horyzontu prognozy.
 4. O wiele mniejsze znaczenie prognostyczne w prognozach na okresy od 7 do 12 miesięcy wykazują „twarde” dane GUS na temat sytuacji gospodarczej branż i rynku pracy (WWG) i prognozowanej koniunktury gospodarczej (PiB). Wartość predykcyjna informacji o kondycji handlu zagranicznego (ZWG) w konstrukcji wskaźników wyprzedzających jest najmniejsza spośród wszystkich analizowanych grup zmiennych.

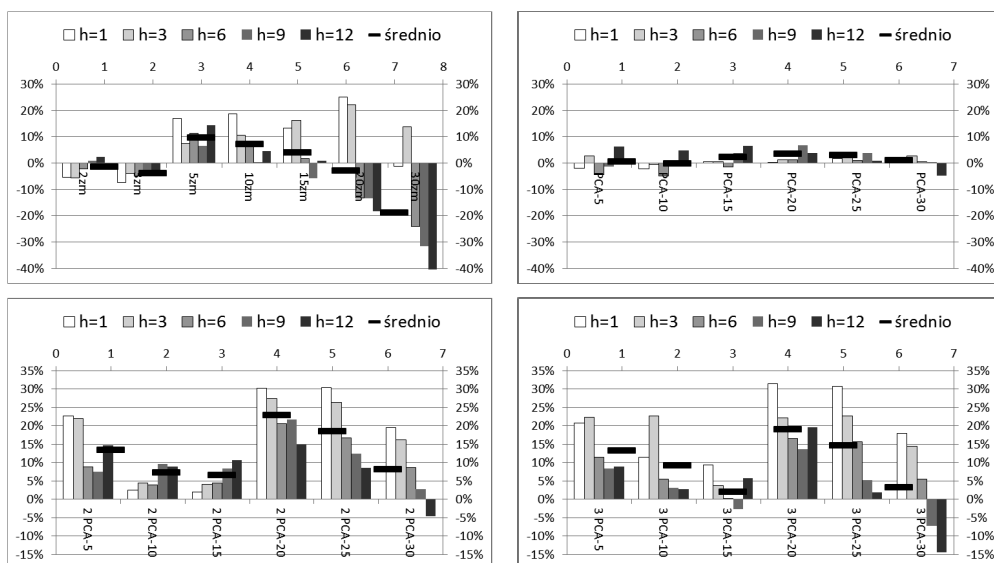
5.2. MODELE WSKAŹNIKÓW INDYWIDUALNYCH A MODELE WSKAŹNIKÓW SYNTETYCZNYCH – WŁASNOŚCI PROGNOSTYCZNE

Modelem bazowym do porównań modeli wskaźników indywidualnych i syntetycznych uczyniono model (1) z jednym wskaźnikiem wiodącym (oddzielnie na każdy horyzont). Na rysunku 2 przedstawiono poprawę jakości prognoz (spadek RMSE) dla modeli wskaźników indywidualnych i syntetycznych w wybranych (reprezentatywnych) horyzontach prognoz ($h=1,3,6,9,12$) oraz średnią poprawę jakości prognoz dla wszystkich 12 horyzontów w stosunku do modelu z jedną zmienną. W okresie weryfikacji skonstruowanym na podstawie generacji danych od grudnia 2010 r. do listopada 2011 r. można sformułować następujące wnioski:

1. Jakość prognoz dla wskaźników indywidualnych początkowo rośnie wraz ze wzrostem liczby wskaźników wiodących (choć nie monotonicznie) osiągając dla 5 najlepszych w próbie wskaźników wiodących (5 zm) RMSE niższe w stosunku do modelu bazowego od 2% dla $h = 10$ do 17% dla $h = 1$ (tj. średnio o mniej niż 10%). Gdy liczba zmiennych w modelu wskaźników indywidualnych przekracza 20, to jakość prognoz stopniowo się pogarsza w stosunku do prognoz z jednym wskaźnikiem, szczególnie dla horyzontów prognoz pow. 5 miesięcy.
2. Kondensowanie informacji ze wskaźników indywidualnych do jednego wspólnego czynnika przynosi skromną poprawę jakości prognoz średnio poniżej 5% w analizie czynnikowej dla 20 zmiennych (1-PCA-20), chociaż wraz ze wzrostem liczby zmiennych relatywne pogorszenie jakości prognoz nie jest tak znaczne jak dla wskaźników indywidualnych.
3. Włączanie do modelu wskaźników wyprzedzających drugiego (2-PCA) i trzeciego (3-PCA) wspólnego czynnika (tj. kolejnych głównych składowych macierzy korelacji uzyskanych metodą PCA) przynosi znaczną poprawę jakości prognoz, największą dla około 20 wskaźników wyprzedzających użytych w analizie czynnikowej – średnio o 23% dla 2 czynników (od 11% na $h = 11$ do 37% na $h = 2$), oraz o 19% dla 3 wspólnych czynników. Dalsze zwiększanie liczby zmiennych do analizy czynnikowej skutkuje spadkiem korzyści ze stosowania analizy czynnikowej, szczególnie dla horyzontów prognoz powyżej 6 miesięcy.
4. Porównanie jakości prognoz wybranych modeli wskaźników indywidualnych i wybranego modelu wskaźników syntetycznych dla 20 zmiennych i 2 wspólnych czynników wskazuje na korzyści ze stosowania analizy czynnikowej dla prognoz od 4 do 11 miesięcy wpród (por. tabela 4 i rysunek 3). Na podstawie prezentowanej metody wskaźników syntetycznych uzyskano w dwuletnim okresie weryfikacji prognozy przynajmniej o porównywalnej precyzji i obciążeniu w stosunku do pro-

gnoz korzystających z dowolnego zestawu optymalnie dobieranych wskaźników indywidualnych.

5. Kondensowanie informacji w postaci 2–3 wspólnych czynników zwiększa odporność prognoz opartych na wskaźnikach wyprzedzających na zmianę liczby zmiennych w modelu w porównaniu z metodami opartymi na wskaźnikach indywidualnych. Ma to praktyczne znaczenie, gdyż na ogół nie znamy optymalnej liczby zmiennych, których należy użyć do konstrukcji wskaźników wyprzedzających. Wówczas synteza informacji w postaci metod analizy czynnikowej jest bezpiecznym, choć nie zawsze najbardziej efektywnym sposobem uzyskiwania prognoz na podstawie dużych zbiorów danych.



Rysunek 2. Relatywna jakość prognoz w okresie weryfikacji od grudnia 2010 r. do listopada 2012 r. Jakość prognoz policzono dla modeli indywidualnych i syntetycznych wskaźników wyprzedzających w porównaniu z wariantem bazowym (tj. modelem z jednym wskaźnikiem indywidualnym na każdy horyzont) i wyrażono jako procentowy spadek RMSE. Określenie „średnia” dotyczy średniej jakości prognoz dla wszystkich horyzontów.

Źródło: obliczenia własne.

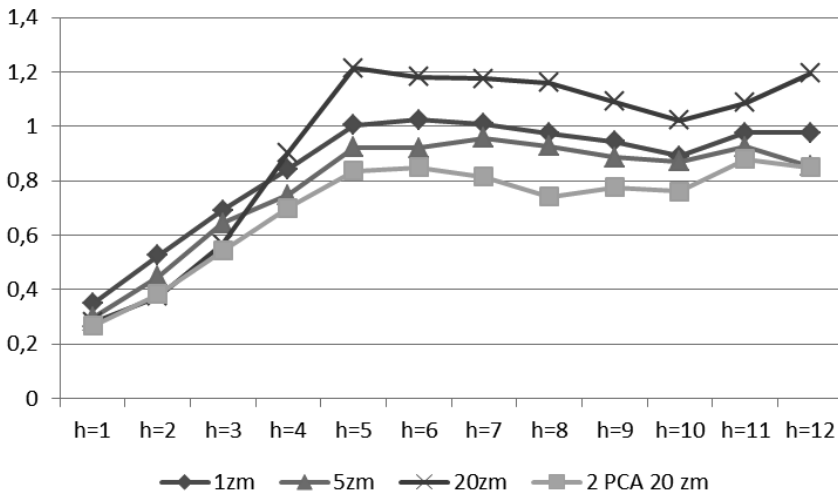
Tabela 4.

Błędy prognoz poza próbę dla wybranych modeli wskaźników wyprzedzających i horyzontów w okresie weryfikacji od stycznia 2011 r. do grudnia 2012 r.

Model	Miara błędów					
1 zm	RMSE	0,35	0,69	1,02	0,94	0,98
	MAE	0,28	0,61	0,85	0,77	0,85
	ME	-0,07	-0,08	-0,19	-0,36	-0,68
5 zm	RMSE	0,30	0,64	0,92	0,89	0,85
	MAE	0,24	0,50	0,76	0,74	0,75
	ME	-0,04	-0,13	-0,13	-0,06	-0,37
20 zm	RMSE	0,28	0,57	1,18	1,09	1,19
	MAE	0,22	0,44	0,97	0,89	1,09
	ME	0,03	-0,04	-0,03	0,07	-0,23
2 PCA 20 zm	RMSE	0,27	0,54	0,85	0,78	0,85
	MAE	0,22	0,46	0,68	0,62	0,72
	ME	0,01	-0,01	-0,06	-0,10	-0,33

Oznaczenia modeli dotyczą modeli z jedną (1 zm), pięcioma (5 zm) i dwudziestoma (20 zm) wskaźnikami indywidualnymi, oraz modelu wskaźników syntetycznych na podstawie 20 zmiennych i dwóch wspólnych czynników (2 PCA-20 zm).

Źródło: obliczenia własne.



Rysunek 3. Błędy RMSE (w pp.) dla wybranych modeli wskaźników wyprzedzających w okresie od stycznia 2011 do grudnia 2012 r. Oznaczenia modeli jak w Tabeli 4.

Źródło: obliczenia własne.

6. PODSUMOWANIE

Wnioski, jakie płyną z analizy modeli wskaźników wiodących dla inflacji w Polsce można wstępnie podsumować w następujący sposób. Po pierwsze, metoda wyboru zmiennych na podstawie analizy siły ich korelacji z przyszłą inflacją w modelu wskaźników wyprzedzających opartym o zasadę prognozowania bezpośredniego i uwzględniającym sezonowość daje (w okresie testowym) stosunkowo stabilne rezultaty z punktu widzenia wyboru predyktorów. Spośród zmiennych najsilniej skorelowanych z przyszłą inflacją dla horyzontów prognoz do 6 miesięcy występują zmienne niosące informację o krótkookresowej uporczywości inflacji (tj. miary agregatów CPI i jego komponentów), oraz oczekiwania inflacyjne gospodarstw domowych. Dla dłuższych horyzontów prognoz zyskują na znaczeniu informacje pochodzące z rynku finansowego, a także komponenty CPI dotyczące żywności. Po drugie, początkowo wraz ze wzrostem liczby indywidualnych wskaźników wyprzedzających w modelu prognostycznym rośnie jakość prognoz modelowych, co może świadczyć o tym, że wiele czynników kształtuje przyszłą inflację. Po trzecie, potwierdzony został wniosek z artykułu Stocka i Watsona (2006) o pogarszaniu się własności prognostycznych modelu wskaźników indywidualnych, gdy liczba zmiennych objaśniających w modelu prognostycznym przekroczy pewien pułap. Wówczas, jak pokazują wyniki wielu analiz, remedium na utratę dokładności estymatora może być syntetyzowanie informacji pochodzących z wielu źródeł za pomocą metod analizy czynnikowej i statystycznego doboru zmiennych do tej analizy. Dla prognozowania inflacji w Polsce w okresie weryfikacji obejmującym lata 2011–2012 daje ona szczególnie dobre rezultaty w postaci redukcji błędów *ex post* dla liczby wspólnych czynników większej niż jeden. Podobnie jak w modelu dla wskaźników indywidualnych również w modelu czynnikowym dołączanie do modelu kolejnych zmiennych (coraz słabiej skorelowanych w próbie) ma wpływ na pogorszenie jakości prognoz, ale skutki nie są tak znaczące jak w modelach wskaźników indywidualnych.

Uniwersytet Łódzki oraz Instytut Ekonomiczny NBP

LITERATURA

- [1] Ang, A., Bekaert, G., Wei, M., (2007), Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better? *Journal of Monetary Economics*, 54 (4), 1163–1212.
- [2] Angelini, E., Henry, J., Mestre, R., (2001), Diffusion Index-based Inflation Forecasts for the Euro Area. Working Paper Series 061, European Central Bank.
- [3] Arratibel, O., Kamps, C., Leiner-Killinger, N., (2009), Inflation Forecasting in the New EU Member States. Working Paper Series 1015, European Central Bank.
- [4] Bai, J., (2003), Inferential Theory for Factor Models of Large Dimensions, *Econometrica*, Elsevier, 71 (1), 135–172.

- [5] Bai, J., Ng, S., (2008), Forecasting Economic Time Series Using Targeted Predictors, *Journal of Econometrics*, Elsevier, 146 (2), 304–317.
- [6] Banerjee, A., Marcellino, M., Masten, I., (2004), Forecasting Macroeconomic Variables for the Acceding Countries. Working Papers 260, IGIER (Innocenzo Gasparini Institute for Economic Research), Bocconi University.
- [7] Banerjee, A., Marcellino, M., Masten, I., (2005), Leading Indicators for Euro-Area Inflation and GDP Growth. *Oxford Bulletin of Economics and Statistics*, 67 (s1), 785–813.
- [8] Baranowski, P., Leszczyńska, A., Szafranski, G., (2010), Krótkookresowe prognozowanie inflacji z użyciem modeli czynnikowych, *Bank i Kredyt*, 41 (4), 23–44.
- [9] Białowolski, P., Żochowski, D., (2006), Analiza własności prognostycznych komponentów WPI, w: Drozdowicz-Bieć M., (red.), *Wskaźniki wyprzedzające*. Prace i Materiały Instytutu Rozwoju Gospodarczego SGH. Warszawa.
- [10] Białowolski, P., Żochowski, D., (2006), *Wskaźniki wyprzedzające*. Prace i Materiały Instytutu Rozwoju Gospodarczego SGH, Rozdz. w: *Analiza własności prognostycznych komponentów WPI*, 167–197.
- [11] Boivin, J., Ng, S., (2005), Understanding and Comparing Factor-Based Forecasts, *International Journal of Central Banking*, 1 (3), 117–151.
- [12] Boivin, J., Ng, S., (2006), Are More Data Always Better for Factor Analysis?, *Journal of Econometrics*, Elsevier, 132(1), 169–194.
- [13] Burns, A. F., Mitchell, W. C., (1946), *Measuring Business Cycles*, Nowy Jork, NBER.
- [14] Brzoza-Brzezina, M., Kotłowski, J., (2009), Bezwzględna stopa inflacji w gospodarce polskiej, *Gospodarka Narodowa*, 9 (37), 1–21.
- [15] Cristadoro, R., Forni, M., Reichlin, L., Veronese, G., (2005), A Core Inflation Indicator for the Euro Area, *Journal of Money, Credit and Banking*, 37 (3), 539–560.
- [16] Doz, C., Giannone, D., Reichlin, L., (2006), A Quasi Maximum Likelihood Approach for Large Approximate Dynamic Factor Models, Working Paper Series 674, European Central Bank.
- [17] Forni, M., Hallin, M., Lippi, M., Reichlin, L., (2005), The Generalized Dynamic Factor Model: One-Sided Estimation and Forecasting, *Journal of the American Statistical Association*, 100, 830–840.
- [18] Giannone, D., Matheson, T., (2007), A New Core Inflation Indicator for New Zealand, CEPR Discussion Papers 6469.
- [19] Hendry, D. F., Hubrich, K., (2006), Forecasting Economic Aggregates by Disaggregates, Working Paper Series 589, European Central Bank.
- [20] Hertel, K., Leszczyńska, A., (2013), Uporczywość inflacji i jej komponentów – badanie empiryczne dla Polski, *Przegląd Statystyczny*, 60 (2), 187–210.
- [21] Hofmann, B., (2008), Do Monetary Indicators Lead Euro Area inflation?, Working Paper Series 867, European Central Bank.
- [22] Kapetanios, G., (2004), A Note on Modelling Core Inflation for the UK Using a New Dynamic Factor Estimation Method and a Large Disaggregated Price Index Dataset, *Economics Letters*, 85 (1), 63–69.
- [23] Kotłowski, J., (2008), Forecasting Inflation with Dynamic Factor Model – the Case of Poland, Working Papers 24, Department of Applied Econometrics, Warsaw School of Economics.
- [24] Łupiński M., (2012), Short-term Forecasting and Composite Indicators Construction with Help of Dynamic Factor Models Handling Mixed Frequencies Data with Ragged Edges, *Przegląd Statystyczny*, 59 (1), 48–73.
- [25] Łyziak, T., Stanisławska, E., (2008), Consumer Inflation Expectations in Europe: Some Cross-Country Comparisons, *Bank i Kredyt*, 39 (9), 14–28.
- [26] Reis, R., Watson, M. W., (2010), Relative Goods’ Prices, Pure Inflation, and the Phillips Correlation, *American Economic Journal: Macroeconomics*, 2 (3), 128–157.

-
- [27] Stock, J. H., Watson, M. W., (1989), New Indexes of Coincident and Leading Economic Indicators, w: Blanchard, O. J., Fischer, S., (red.), NBER Macroeconomics Annual 1989, 4, NBER Chapters, National Bureau of Economic Research, Inc, 351–409.
- [28] Stock, J. H., Watson, M. W., (2002a), Forecasting Using Principal Components From a Large Number of Predictors, *Journal of the American Statistical Association*, 97 (460), 1167–1179.
- [29] Stock, J. H., Watson, M. W., (2002b), Macroeconomic Forecasting Using Diffusion Indexes, *Journal of Business and Economic Statistics*, 20 (2), 147–162.
- [30] Stock, J. H., Watson, M. W., (2003), Forecasting Output and Inflation: The Role of Asset Prices, *Journal of Economic Literature*, 41 (3), 788–829.
- [31] Stock, J. H., Watson, M. W., (2006), Forecasting with Many Predictors, w: Elliot G., Granger C.W.J., Timmermann A., (red.), *Handbook of Economic Forecasting*, Elsevier, Roz. 10 (1), 515–554.
- [32] Stock, J. H., Watson, M. W., (2008), Phillips Curve Inflation Forecasts, NBER Working Papers 14322, National Bureau of Economic Research, Inc.
- [33] Szafranski G., (2011), Krótkoterminowe prognozy polskiej inflacji w oparciu o wskaźniki wyprzedzające, Materiały i Studia, Zeszyt nr 263, NBP, Warszawa.

ZAŁĄCZNIK

ZAWARTOŚĆ MERYTORYCZNA I STATYSTYCZNA BAZY DANYCH

Tabela 5.

Informacje na temat zawartości zbioru danych

Symbol grupy	Zawartość grupy	Źródło danych	Liczba zmiennych	Opóźnienie publikacyjne*	Transformacje/ sezonowość (SA)
Komponenty wskaźnika CPI					
CPI	Wskaźnik CPI ogółem, inflacja bazowa i komponenty CPI bez grupy ZYWU	GUS	24	1 mies.	Indeksy m/m (SA)
ZYWU	Komponenty CPI grupy: żywność, napoje, używki	GUS	13	1 mies.	Indeksy m/m (SA)
Podaż pieniądza i czynniki jego kreacji					
PIEN	Podaż pieniądza, rezerwa obowiązkowa i składniki bilansu MIF	NBP	20	2–3 mies.	Tempa zmian m/m (SA)
Koszty produkcji i surowców w podziale na źródła krajowe i zagraniczne					
KRAJ	Koszty produkcji (PPI)	GUS	9	2 mies.	Indeksy m/m (SA)
	Koszty surowców rolnych	MRol	6	0,5 mies.	Tempa zmian m/m (SA)
	Ceny paliw i ich koszty (w tym akcyza)	e-petrol.pl	6	0,5 mies.	
ZAGR	Kursy walutowe PLN	NBP	4	0,5 mies.	Zmiany m/m
	Ceny dóbr i surowców (w PLN)	MFW	15	1 mies.	Tempa zmian m/m (SA)
Wskaźniki makroekonomiczne dot. bieżącej sytuacji ogólnogospodarczej					
MAKR	Wskaźniki aktywności w branżach	GUS	6	2 mies.	Indeksy m/m (SA)
	Budżet państwa	GUS	4	2 mies.	Tempa zmian m/m (SA)
	Dane z rynku pracy	GUS	10	2 mies.	Zmiany m/m (SA)
	Handel zagraniczny	GUS	10	3 mies.	Tempa zmian m/m (SA)
	Dane dotyczące oczekiwań (w tym dane ankietowe i dane z rynku finansowego)				

CiH	Wskaźniki koniunktury (ceny i handel)	GUS	17	1 mies.	Zmiany m/m (SA)
	Oczekiwania inflacyjne	IPSOS	5	1 mies.	
KON	Wskaźniki koniunktury (ogółem, przemysł i budownictwo)	GUS	25	1 mies.	Zmiany m/m (SA)
RF	Stopy procentowe instrumentów dłużnych (lokaty mbank., obligacje SP)	Reuters	7	0,5 mies.	Tempa zmian m/m
	Stopy zwrotu (indeksy giełdowe)	Reuters	8	0,5 mies.	

* Opóźnienie publikacyjne liczone jest oddzielnie dla danych pochodzących z generacji danych. Jest to różnica między końcem miesiąca sporządzenia prognozy a okresem, którego dotyczą najnowsze dane dostępne w danej generacji danych, np. opóźnienie publikacyjne 1 miesiąc oznacza, że w miesiącu sporządzenia prognozy dostępne są dane za poprzedni miesiąc, „0,5” miesiąca oznacza, że do obliczenia średniej za bieżący miesiąc wykorzystywane są dane dzienne lub tygodniowe dot. pierwszej połowy bieżącego miesiąca.

Źródło: opracowanie własne.

SYNTETYCZNE WSKAŹNIKI WYPRZEDZAJĄCE W PROGNOZOWANIU INFLACJI
W POLSCE

Streszczenie

W artykule przedstawiono wyniki procedury doboru zmiennych pochodzących z dużego zbioru danych ekonomicznych i finansowych do systematycznego, krótkookresowego prognozowania wskaźnika inflacji (CPI) w Polsce. Doboru zmiennych najsilniej skorelowanych z przyszłą inflacją w horyzoncie od 1 do 12 miesięcy naprzód, zwanych wskaźnikami wyprzedzającymi inflacji, dokonano na podstawie generacji danych rzeczywiście dostępnych co miesiąc od lipca 2009 do listopada 2010 roku (okres testowy). Następnie jakość prognoz w oparciu o modele wskaźników indywidualnych porównano ze wskaźnikami syntetycznymi w okresie od grudnia 2010 do listopada 2012 roku (okres weryfikacji) stosując bezpośrednią metodę prognozowania. W porównaniu do modeli wskaźników indywidualnych ustalonych na podstawie rankingu w okresie testowym wskaźniki syntetyczne uzyskane za pomocą metod analizy czynnikowej okazują się być dobrymi predyktorami przyszłej inflacji pod względem precyzji, nieobciążoności i odporności prognoz na zmiany liczby wskaźników i wspólnych czynników.

Słowa kluczowe: wskaźniki wyprzedzające, analiza czynnikowa, inflacja w Polsce

COMPOSITE LEADING INDICATORS IN FORECASTING INFLATION IN POLAND

Abstract

We present the procedure for a systematic selection of short-term leading indicators for Polish consumer inflation (CPI) from large number of predictors. Using data set of 189 economic and financial time series (available from July 2009 to November 2012) we select the indicators highly (in-sample) correlated with future inflation from 1 to 12 months ahead. Then we perform a direct forecasting exercise on a real-time data base from December 2010 to November 2012. Compared to individual leading indicators we find that few common factors extracted from a principal component analysis (named composite leading indicators) are successful tools for predicting future inflation being precise, unbiased, and more robust to changes in a number of predictors and common factors.

Keywords: leading indicators, factor analysis, inflation in Poland