

Porównanie trafności jednorocznych prognoz polskiej koniunktury sporządzanych przez krajowe i międzynarodowe instytucje ekonomiczne

Aleksander Grechuta*

Nadesłany: 17 marca 2017 r. Zaakceptowany: 3 stycznia 2018 r.

Streszczenie

Dobra znajomość przyszłych warunków działania jest niezbędna do osiągnięcia sukcesu w przedsięwzięciach gospodarczych. W makroekonomii warunki te nazywa się koniunkturą. W tym świetle znaczenie przewidywania koniunktury nie budzi wątpliwości, nie dziwi zatem, że wiele podmiotów zajmuje się tworzeniem prognoz ekonomicznych. Prognozy te w niejednakowym stopniu różnią się w zależności od instytucji, więc bardzo ważnym zagadnieniem jest ocena, które z nich charakteryzują się większą trafnością. W niniejszej pracy porównano trafność jednorocznych prognoz polskiej koniunktury sporządzanych przez Narodowy Bank Polski (NBP), Ministerstwo Finansów (MF), Komisję Europejską (KE), Organizację Współpracy Gospodarczej i Rozwoju (OECD) oraz Międzynarodowy Fundusz Walutowy (IMF). Badano trafność prognoz *ex post* w okresie 2004–2016, z uwzględnieniem korekt dla poszczególnych instytucji, związanych z dostępnością danych. Zmiany koniunktury opisywano za pomocą wybranych charakterystyk: rocznej dynamiki produktu krajowego brutto (PKB), inflacji mierzonej wskaźnikiem cen i towarów konsumpcyjnych (CPI) lub zharmonizowanym indeksem cen konsumpcyjnych (HICP) oraz stopy bezrobocia. Do badania trafności prognoz sporządzonych przez poszczególne instytucje wykorzystano wybrane miary jakości predykcji *ex post*, m.in.: średni błąd prognozy (ME), średni bezwzględny błąd prognozy (MAE), pierwiastek błędów średniokwadratowego (RMSE), współczynnik U-Theila oraz dekompozycje błędów średniokwadratowego.

Słowa kluczowe: PKB, inflacja, bezrobocie, trafność prognoz krótkookresowych, gospodarka Polski

JEL: E02, O1, F41

* e-mail: aleksander@grechuta.com.

1. Wstęp

Instytucje ekonomiczne publikują swoje prognozy z różną częstotliwością, przez co ich porównywanie jest utrudnione. Zarówno Narodowy Bank Polski (NBP), jak i Ministerstwo Finansów (MF) są instytucjami oddziałującymi na krajową gospodarkę, co oznacza, że sporządzane przez nie prognozy mają w pewnym sensie charakter planu. Inaczej jest w przypadku Komisji Europejskiej (KE), OECD i Międzynarodowego Funduszu Walutowego (IMF); ich prognozy pełnią jedynie funkcję informacyjną.

Na przełomie XX i XXI w. prowadzone były podobne badania nad trafnością prognoz opracowywanych przez instytucje ekonomiczne, dotyczących najważniejszych zmiennych dla wielu krajów. Pomijały one niestety polską gospodarkę. Można się domyślać, że nie miała ona w tamtym okresie dostatecznie dużego znaczenia. Zmieniło się to po przystąpieniu Polski do OECD oraz do Unii Europejskiej. Większość badań nie uwzględnia też wielu zmian, jakie zaszły w ostatnich 15 latach, oraz ważnych wydarzeń, jak np. kryzys finansowy zapoczątkowany w 2007 r.

Punktem wyjścia do rozważań w tej pracy jest przegląd wybranej literatury oraz najczęściej powtarzających się wniosków i spostrzeżeń. Oller i Barot (2000) zbadali trafność prognoz PKB i inflacji o horyzoncie jednorocznym, sporządzanych przez OECD, IMF i instytucje narodowe z 13 krajów (poświęcili także uwagę prognozom o horyzoncie 2-letnim). Autorzy poruszyli istotną kwestię, że oprócz samego badania trafności prognoz wyrażonej przy użyciu miar liczbowych istotne jest także badanie zdolności do przewidywania zmian kierunku (punktów zwrotnych) zjawisk ekonomicznych. Stwierdzili, że z punktu widzenia decyzji biznesowych trafność oceniana za pomocą tego typu miar (np. MAE i RMSE) nie ma takiego znaczenia, jak przewidywanie momentów odwrócenia trendu. Trafność ich prognozowania została przez autorów uznana za widoczny problem dla instytucji, które objęły badaniem. Zwrócili również uwagę na występowanie autokorelacji w błędach prognoz dla niektórych instytucji oraz sprawdzali występowanie obciążenia (czyli optymizmu lub pesymizmu). Uznali także, że prognozy badanych instytucji były trafniejsze niż prognozowanie naiwne oraz że inflację łatwiej prognozować niż PKB (na podstawie RMSE).

W pracy Blix i in. (2001) uwzględniono bardzo liczną grupę podmiotów prognozujących, których sporą część stanowiły instytucje komercyjne. Spośród międzynarodowych organizacji wybrano OECD i IMF. Dane, które wykorzystano w analizie, pochodziły z lat 1991–2000. Autorzy doszli m.in. do wniosku, że inflacja jest łatwiejsza do przewidywania niż PKB oraz że częściej jest przeszacowywana, podczas gdy wzrost gospodarczy ulegał raczej niedoszacowaniu. Ponadto zaobserwowali, że mimo silnej globalizacji gospodarek nie występują istotne związki między błędami prognoz dla poszczególnych państw. Stwierdzili też, podobnie jak autorzy poprzedniej pracy, że największą trudnością instytucjom prognozującym sprawiało prawidłowe przewidywanie punktów zwrotnych. Wszystkie analizowane przez nich prognozy były trafniejsze niż uzyskane metodą naiwną.

Kreini (2000) skupił się jedynie na prognozach IMF oraz OECD dla największych gospodarek. Poza PKB i inflacją analizował także bezrobocie i bilans handlowy. Stwierdził, że prognozy IMF i OECD były trafniejsze od prognoz naiwnych. Zwrócił również uwagę na problem z przewidywaniem punktów zwrotnych.

Po 2010 r. można zaobserwować powrót zainteresowania omawianym zagadnieniem – uznano, że ze względu na ostatnie wydarzenia warto zaktualizować posiadaną wiedzę. Gonzales-Cabanillas i Terzi (2012) podjęli się oceny trafności prognoz Komisji Europejskiej. Uważali to za potrzebne m.in. ze względu na kryzys gospodarczy, który dotknął Unię Europejską i strefę euro po 2007 r. Jednym z celów ich

pracy było sprawdzenie, jak wpłynęło to na trafność prognoz. Ponadto uwzględnili w swojej pracy kraje, które dołączyły do Unii Europejskiej w 2004 r. (a więc także Polskę) oraz w 2007 r. Autorzy zaznaczyli także, że ich praca jest pierwszym artykułem KE badającym jakość prognoz dla krajów, które weszły w skład wspólnoty w późniejszym czasie. Zostało odnotowane, że jakość prognoz dla tych krajów była wyraźnie gorsza niż dla pozostałych państw UE. Co ważne, było kilka wyjątków, wśród których znalazła się polska gospodarka (a dokładniej jej prognozy o horyzoncie jednorocznym). Błędy prognoz dla Polski (mierzone za pomocą MAE i RMSE) okazały się porównywalne z błędami dla krajów ze strefy euro, Wielkiej Brytanii i Skandynawii. Badano także występowanie systematycznego obciążenia związanego z tendencją prognostów do pesymizmu lub optymizmu oraz występowanie autokorelacji w błędach. Zaobserwowano, że w prognozach KE występuje ona tylko w odniesieniu do kilku państw i nie dotyczy Polski. Głównym wnioskiem pracy było jednak stwierdzenie, że kryzys gospodarczy ma duży wpływ na wyniki prognozowania; w latach recesji wartości błędów były zdecydowanie większe niż w pozostałych okresach. Podobnie jak w opracowaniach z początku XXI w. prognozowanie punktów zwrotnych okazało się mało skuteczne. Autorzy porównali również prognozy KE z prognozami OECD i IMF i uznali, że ich trafność jest podobna. Dane omawiane w pracy pochodzą z lat 1969–2011, porównano jednak wyniki dla okresu przed- i pokryzysowego (czyli osobno okres 1969–2007 i pełne dane). Warto nadmienić, że nie dla wszystkich omawianych krajów liczba obserwacji była taka sama.

Makarova (2014) przeprowadziła analizę prognoz makroekonomicznych Narodowego Banku Polskiego, praca została więc w całości poświęcona polskiej gospodarce. Autorka poruszyła kwestię niepewności makroekonomicznej, metod jej pomiaru oraz jej znaczenia dla skuteczności prognozowania. Zwróciła uwagę na potrzebę szukania konsensusu pomiędzy metodami opartymi na rozkładzie prognoz budowanych przy użyciu ankiet makroekonomicznych oraz metod wykorzystujących modele ekonometryczne (rozważali to także Blix i in. 2001). Podobnie jak autorzy wcześniej omówionych prac (Oller, Barot 2000; Blix i in. 2001; Gonzales-Cabanillas, Terzi 2012) poruszyła też problem występowania pewnego systematycznego obciążenia prognoz. Autorka przypuszcza, że jego przyczyną może być sugerowanie się innymi prognozami.

Z rodzimych badań trafności prognoz dla polskiej gospodarki warto wspomnieć pracę Acedańskiego, Bernais i Mastalerz-Kodzis (2013), która koncentruje się na rynku pracy (a dokładniej na wynagrodzeniach i liczbie pracujących). Uwzględniono w niej okres 1998–2013 i wykorzystano dane zarówno miesięczne, jak i kwartalne. Autorzy przeanalizowali skuteczność wybranych rodzajów modeli prognostycznych – VAR (ang. *vector autoregression*), DFM (ang. *dynamic factor model*), wskaźników wyprzedzających i metody prognoz łączonych. Podobne podejście zaprezentowali Baranowski, Leszczyńska i Szafranski (2010). Ich praca dotyczyła inflacji i skupiła się na skuteczności modeli DFM, z wykorzystaniem danych miesięcznych z lat 1999–2009. Autorzy wykazali wysoką skuteczność modeli DFM w porównaniu z pozostałymi badanymi metodami (zwłaszcza przy dłuższych horyzontach).

W wielu wymienionych pracach (zwłaszcza tych z początku XXI w. oraz w opracowaniu Gonzales-Cabanillas i Terzi (2012) do oceny trafności stosuje się miary ME, MAE i RMSE oraz statystykę U-Theila, które wykorzystano również w niniejszej pracy. W artykule Ollera i Barota (2000) zastosowano także miarę RMSE/SD, czyli RMSE podzielone przez odchylenie standardowe, która stanowi przybliżenie statystyki U-Theila.

Na podstawie przeglądu literatury można wskazać kilka najczęściej rozważanych problemów. Po pierwsze, prawie zawsze badano możliwość przewidywania punktów zwrotnych (w dalszej części pracy znajduje się przybliżenie tego terminu). Ze względu na dużą zgodność badaczy w tej kwestii problem

ten zostanie tutaj przebadany w kontekście polskiej gospodarki. Po drugie, analizowano występowanie autokorelacji błędów. W tym celu stosowano różnego rodzaju testy na autokorelację (np. test Ljunga-Boxa). Po trzecie, często sprawdzano występowanie obciążenia, które można rozumieć jako skłonność prognostów do pesymizmu lub optymizmu. Po czwarte, analizowano, czy prognozy były trafniejsze niż uzyskane metodą naiwną. W niniejszej pracy zostanie zweryfikowane, w jaki stopniu sprawdza się to w przypadku polskiej gospodarki.

Na podstawie wniosków płynących z literatury oraz własnych przemyśleń sformułowano kilka hipotez, które będą sprawdzane w tej pracy. Przypuszcza się, że trafność prognoz zależy od charakteru badanej zmiennej i sposobu jej pomiaru. Wyniki znane z przedstawionej literatury wskazują, że inflacja jest łatwiejsza do prognozowania niż PKB oraz że inflację przeliczowano, a PKB niedoszacowywano. Kontynuując tę myśl, założono, że trafność prognoz w dużym stopniu zależy od faktycznego przebiegu badanej zmiennej. W sytuacji, gdy jest on stabilny i obciążony niewielką zmiennością, trafność powinna być większa niż wówczas, gdy jest gwałtowny i ma większe odchylenia. Jeśli chodzi o punkty zwrotne, to przypuszcza się, że podobnie jak w innych badaniach ich trafne prognozowanie rzadko kończy się sukcesem. Na początku tego rozdziału poruszono problem braku zgodności w czasie w sporządzaniu prognoz. Dokumenty zawierające prognozy publikowane są w różnych miesiącach, momenty publikacji mogą być zaś różne dla każdej instytucji. Różnice te wzrastają wraz ze zwiększaniem puli analizowanych podmiotów. W pracy zostanie omówione, jak wpływa to na dokonywanie porównań, zaproponowano również sposób radzenia sobie z tą niedogodnością. Intuicja podpowiada, że prognozy oparte na świeższych zbiorach informacyjnych będą trafniejsze. Z punktu widzenia interpretacji najwygodniej byłoby, gdyby większą trafność prognoz jednej instytucji potwierdzały wszystkie miary jakości predykcji równocześnie. Prawdopodobnie jest to jednak rzadko spotykane, a największa trafność rozkłada się pomiędzy różne statystyki.

Struktura pracy przedstawia się następująco. W rozdziale drugim opisano szeregi czasowe, których prognozy zostaną poddane analizie. W rozdziale trzecim omówiono instytucje krajowe i zagraniczne sporządzające prognozy dla polskiej gospodarki. Zaprezentowano też ich metodykę oraz schemat przygotowywania prognoz. Rozdział czwarty zawiera objaśnienie i interpretacje miar zastosowanych do oceny trafności prognoz. W rozdziale piątym rozważa się zagadnienie porównywania prognoz, gdy występuje niezgodność momentów, w jakich były publikowane. Rozdział szósty zawiera analizę jakości prognoz za pomocą metod zaprezentowanych w rozdziale piątym. Rozdział siódmy to podsumowanie wniosków płynących z pracy.

2. Wybrane zmienne makroekonomiczne oraz znaczenie ich prognozowania

Do analizy wybrano następujące zmienne: średnioroczny wzrost PKB wyrażony w cenach stałych, inflację oraz stopę bezrobocia. Dane PKB pochodzą z Głównego Urzędu Statystycznego. Zastosowano procentowy wzrost w stosunku do roku poprzedniego, gdzie za rok bazowy przyjmuje się okres poprzedni ($t - 1$). Wykorzystane dane PKB (tabela 1) obliczono za pomocą metody ESA'2010, ustanowionej przez Parlament Europejski jako zintegrowana metoda rachunków narodowych dla wszystkich krajów członkowskich. W niektórych sytuacjach może to budzić wątpliwości co do porównywalności danych obliczanych za pomocą starszych metod, m.in. ze względu na liczne rewizje. Uznano jednak, że najświeższe dane najlepiej odzwierciedlają faktyczny stan gospodarki. Z tego powodu wszystkie prognozy są porównywane z PKB obliczonym na podstawie ESA'2010.

Kolejną analizowaną zmienną jest inflacja. Stosuje się różne metody jej pomiaru; obecnie dwiema najbardziej popularnymi są: wskaźnik zmian cen dóbr i usług konsumpcyjnych CPI (ang. *consumer price index*) oraz zharmonizowany indeks cen konsumpcyjnych HICP (ang. *harmonised index of consumer prices*) (Hałka, Leszczyńska 2011). Prognozowanie tego zjawiska ma tym większe znaczenie, że banki centralne prowadzą politykę celu inflacyjnego, a zatem zdolność przewidywania inflacji pozwala na wcześniejsze podjęcie odpowiednich działań i utrzymanie jej ustalonego poziomu. W przypadku Narodowego Banku Polskiego cel inflacyjny wynosi 2,5% z możliwością odchylenia o 1 pkt proc. (NBP 2016). Dla Polski inflacja CPI i HICP są zbliżone, ponieważ różnice w doborze koszyka dóbr i usług są nieznaczne (tabela 1). Należy jednak pamiętać, że nie dotyczy to wszystkich krajów.

Jako zmienną reprezentującą sytuację na polskim rynku pracy wybrano stopę bezrobocia. Według ustawy osobę bezrobotną definiuje się jako niezatrudnioną, nieprowadzącą własnej działalności gospodarczej oraz niewykonującą pracy zarobkowej, jednak zdolną i gotową do jej podjęcia w pełnym wymiarze czasu. W zależności od instytucji zajmującej się badaniami rynku pracy definicja bezrobocia oraz sposób jego pomiaru mogą się jednak w niewielkim stopniu różnić. Dlatego prognozy analizowanych zmiennych mogą dotyczyć innej zmiennej. Przykładowo, stopa bezrobocia prognozowana przez NBP to skorygowane dane BAEL, natomiast Ministerstwo Finansów w ustawie budżetowej opiera swoją prognozę na danych dla bezrobocia rejestrowanego. Komisja Europejska oraz OECD stosują z kolei stopę bezrobocia Eurostatu (stopę BAEL). W przypadku IMF do analiz prognoz stopy bezrobocia nie włączono tej instytucji, ponieważ dla grupy gospodarek, do której przypisano Polskę, prognozy tej zmiennej zaczęto publikować dopiero w październiku 2010 r. Uznano, że krótszy okres może negatywnie wpłynąć na porównywalność wyników.

3. Instytucje publikujące prognozy dla polskiej gospodarki i dobór okresów

Do sporządzania swojej prognozy Narodowy Bank Polski wykorzystuje model ekonometryczny NECMOD. Został on wprowadzony w 2008 r. i jest ulepszoną wersją wcześniej stosowanego modelu ECMOD (Budnik i in. 2009). Jest to model hybrydowy; składa się z 279 równań oraz 369 zmiennych. Dzięki takiemu rozbudowaniu umożliwia analizowanie elementów polskiej gospodarki zarówno w ujęciu całościowym, jak i w ujęciu cząstkowym. Ujęto w nim specyficzne cechy naszej gospodarki, m.in. dotacje unijne czy wpływ okresu transformacji. Równania modelu są podzielone na (Greszta i in. 2012):

- blok cen,
- rynek pracy,
- handel zagraniczny,
- blok gospodarstw domowych,
- stopy procentowe,
- akumulację.

Model był kilkakrotnie aktualizowany (Budnik i in. 2009; Greszta i in. 2010; Greszta i in. 2011; Greszta i in. 2012):

- 2009 r. – modyfikacja specyfikacji bloku handlu zagranicznego,
- 2010 r. – modyfikacja równania inwestycji przedsiębiorstw,
- 2011 r. – modyfikacja bloku cen, uwzględnienie zjawiska histerezy bezrobocia,
- 2012 r. – modyfikacja bloku aktywności zawodowej.

Projekcje opracowywane za pomocą modelu NECMOD są publikowane trzy razy w roku: do 2010 r. w lutym, czerwcu i październiku, od początku 2010 r. – w marcu, lipcu oraz w listopadzie. Horyzont obejmuje pozostałą część roku bieżącego, w którym sporządzono projekcję, oraz kolejne dwa lata.

Prognozy rządowe (Ministerstwa Finansów), w odróżnieniu od prognoz pozostałych omawianych instytucji, są publikowane raz w roku jako komplementarna część ustawy budżetowej. Jednoroczny horyzont sporządzania budżetu państwa wiąże się z potrzebą zachowania współpracy między władzą ustawodawczą, sporządzającą ustawę, oraz władzą wykonawczą, wprowadzającą ją w życie. Prognozy makroekonomiczne tworzone są przez Ministerstwo Finansów i przedstawiane do weryfikacji specjalistom – zarówno krajowym, jak i europejskim (Owsiak 2013). Przewidywanie wielkości makroekonomicznych określonych w ustawie jest wyjątkowo istotne, ponieważ są powiązane ze zmiennymi znajdującymi się w budżecie.

Kolejnym rozpatrywanym podmiotem prognozującym jest Komisja Europejska (KE) – organ wykonawczy Unii Europejskiej. Podstawowym zadaniem KE jest przedkładanie wniosków ustawodawczych Parlamentowi Europejskiemu; KE jest jedyną instytucją UE, która ma takie prawo. Wspólnie z Radą i Parlamentem zarządza środkami finansowymi wspólnoty, sporządza roczny budżet (który przedstawia tym organom) i wraz z Europejskim Trybunałem Obrachunkowym nadzoruje wydawanie środków. We współpracy z Trybunałem Sprawiedliwości egzekwuje przestrzeganie prawa unijnego przez kraje członkowskie oraz reprezentuje UE na arenie międzynarodowej i w jej imieniu negocjuje umowy międzynarodowe (EC 2012).

Za sporządzanie prognoz ekonomicznych w Komisji Europejskiej odpowiada Dyrekcja ds. Gospodarczych i Finansowych, bardziej znana pod skrótem DG ECFIN (Directorate General for Economic and Financial Affairs). Jej celem jest wspieranie wzrostu i dobrobytu gospodarczego obywateli Unii Europejskiej oraz krajów do niej nienależących przez promowanie polityki pozwalającej na stały wzrost ekonomiczny i stabilność finansową. Dostarczanie prognoz gospodarczych jest jednym ze sposobów realizowania tego celu (EC 2016). Opracowania KE, pod nazwą *Economic Forecast*, publikowane są trzy razy w roku: zimą, wiosną i jesienią, czyli zgodnie z semestrami europejskimi – okresami, w których odbywa się procedura koordynacji działań gospodarczych UE. Taka częstotliwość została wprowadzona dopiero w 2013 r.; wcześniej prognozy sporządzano dwa razy w roku, bez edycji zimowej, wraz z dwiema prognozami okresowymi (*interim forecast*). Nie są one jednak uwzględnione w analizie, ponieważ jedynie aktualizują prognozy obejmujące pozostałą część roku. Nie dotyczą horyzontu jednorocznego, któremu poświęcona jest niniejsza praca. Nie uwzględniają także wszystkich zmiennych szacowanych w *Economic Forecast*. DG ECFIN nie stosuje jednego modelu prognostycznego, ale modele specyficzne dla gospodarek konkretnych krajów, uzupełnione wiedzą ekspercką.

Kolejną instytucją sporządzającą prognozy jest OECD – Organizacja Współpracy Gospodarczej i Rozwoju (Organization for Economic Co-operation and Development). Założona w 1960 r. OECD jest międzynarodową organizacją ekonomiczną zrzeszającą obecnie 29 krajów, w tym od 1996 r. Polskę. Głównym celem istnienia OECD jest wspieranie państw członkowskich w zwiększaniu wzrostu gospodarczego i poziomu życia ich obywateli. Zajmuje się także gromadzeniem i publikowaniem danych ekonomicznych oraz sporządzaniem prognoz gospodarczych, obejmujących państwa w niej zrzeszone, ugrupowania ponadnarodowe, jak Unia Europejska czy strefa euro, oraz gospodarkę światową i wybrane kraje niebędące członkami UE (np. Chiny, Brazylia). Prognozy OECD mają charakter informacyjny (nie mają charakteru planu), w przeciwieństwie do prognoz rządowych, które traktuje się jako integralną część strategii rozwoju kraju (Nowak 2004). Są publikowane w raportach *OECD Economic Outlook*

dwa razy w roku: od 1967 do 2013 r. w czerwcu i w grudniu, obecnie w maju i listopadzie. Dotyczą wielu zmiennych makroekonomicznych, których liczba zwiększa się wraz z kolejnymi edycjami. W zależności od opracowania horyzont prognozy obejmuje dwa lub trzy lata. Okresowo sporządzane są również prognozy długoterminowe. Metodyka OECD ze względu na szeroki zakres podmiotów objętych prognozą składa się z wielu elementów, które zależą od tego, dla jakiego kraju jest sporządzana prognoza. Sama organizacja swoje podejście określa jako połączenie metody opinii eksperckich z narzędziami statystycznymi oraz modelami ekonometrycznymi. Na proces sporządzania prognoz składają się:

- ocena obecnej sytuacji makroekonomicznej,
- zastosowanie modeli wskaźnikowych,
- analiza zależności między kluczowymi zmiennymi (ang. *key variables*),
- wykorzystanie opracowanego przez OECD *forecast entry system*,
- wykorzystanie modeli makroekonomicznych,
- analiza budżetowa.

Punktem wyjścia do sporządzenia nowej prognozy jest analiza bieżącej sytuacji ekonomicznej poszczególnych krajów odrębnie oraz na tle gospodarki światowej, głównie za pomocą metod ekonometrycznych oraz wskaźników. Istotą początkowego etapu jest porównanie poprzednich założeń co do przyszłego stanu koniunktury z realizacją. Modele wskaźnikowe mają dwojakie zastosowanie: do krótkoterminowych analiz (głównie kwartalnych) PKB i handlu światowego oraz takich, które skupiają się na najważniejszych zmiennych ekonomicznych dla każdego państwa objętego prognozą. Przez analizę zależności między kluczowymi zmiennymi w metodyce OECD rozumie się analizę współzależności między: wydatkami gospodarstw domowych, zatrudnieniem, wynagrodzeniami oraz cenami, luką popytową i handlem zagranicznym. System *forecast entry* nie jest elementem metodologii, lecz scentralizowanym systemem zarządzania prognozami. Pozwala ekspertom na zapoznanie się z najnowszymi danymi oraz ich ewentualne aktualizacje. Jak w większości przypadków jednym z filarów prognoz ekonomicznych OECD jest podejście ekonometryczne. Organizacja stosuje model NiGEM, opracowany przez brytyjski Narodowy Instytut Badań Ekonomicznych i Społecznych (National Institute of Economic and Social Research). Model ma wiele cech modeli DSGE, jednak opiera się na danych historycznych, co pozwala na zachowanie równowagi między założeniami teoretycznymi a danymi, podobnie jak w przypadku opracowanego przez NBP modelu NECMOD. W odniesieniu do Polski model wykorzystuje 137 zmiennych (informacje o modelu pochodzą z portalu NIESR¹). Ostatnim elementem metodyki OECD jest analiza sytuacji budżetowej, oparta głównie na planie budżetowym państw, dla których sporządza prognozy (OECD 2011).

Ze względu na znaczną liczbę analizowanych podmiotów poszczególne metody zostały tu jedynie przybliżone. Można jednak założyć, że wobec każdego kraju stosuje się podejście indywidualne, a prognoza w dużej mierze opiera się na pracach i opiniach ekspertów mających specjalistyczną wiedzę na temat każdej z gospodarek.

Ostatnia wybrana instytucja opracowująca prognozy to Międzynarodowy Fundusz Walutowy, oznaczony tutaj skrótem IMF (International Monetary Fund). Powstał w 1944 r. na mocy postanowień z konferencji w Bretton Woods, gdzie jego statut podpisały 44 państwa. IMF rozpoczął działalność rok później, obecnie ma zaś 189 członków. Polska była obecna przy powstawaniu Funduszu, ale na skutek ówczesnej sytuacji politycznej opuściła jego struktury w 1950 r. Do grona członków powróciła w 1986 r. Głównym celem Funduszu było utworzenie systemu walutowego, który umożliwiłby rozwój handlu

¹ National Institute's Global Econometric Model, <https://nimodel.niesr.ac.uk/>.

i zatrudnienia bez nakładania ograniczeń w wymianie handlowej. Trzy podstawowe funkcje, jakie pełni IMF, to nadzorowanie przestrzegania przez państwa członkowskie ustanowionych reguł kursowych, dostarczanie państwom środków na utrzymanie bilansu płatniczego oraz tworzenie warunków do optymalnego funkcjonowania międzynarodowego systemu walutowego. Jednym z narzędzi realizowania tych celów jest sporządzanie prognoz ekonomicznych. Są one publikowane w dokumencie *World Economic Outlook*, dwa razy w roku (Nowak 2004; IMF 2017a). IMF określa swe metody sporządzania prognoz jako „podejście od dołu” (ang. *bottom-up*). Zespoły prognostów zajmujące się poszczególnymi krajami przygotowują scenariusze dla ich gospodarek, a Fundusz dostarcza prognozy dotyczące globalnej sytuacji ekonomicznej, kursów walut i cen dóbr. Następnie prognoza jest łączona i agregowana, po czym trafia do *World Economic Outlook* i bazy danych IMF (IMF 2017b). W swojej metodologii dzieli gospodarki krajowe na dwie podstawowe grupy: zaawansowane (ang. *advanced economies*) i rozwijające się (ang. *emerging market and developing economies*). Podział ten pomaga w ocenie znaczenia poszczególnych gospodarek podczas sporządzania prognozy globalnej sytuacji ekonomicznej (IMF 2017c).

W analizie jakości prognoz opisanych instytucji wykorzystano dane z lat 2004–2016. Wyjątkiem jest NBP, w przypadku którego datą początkową jest 2009 r. Wynika to z faktu, że instytucja ta nie udostępnia prognoz poza tymi, które znajdują się na jej stronie internetowej. Wybór roku 2004 został z kolei podyktowany tym, że od tego momentu Komisja Europejska rozpoczęła publikowanie pełnych prognoz dla polskiej gospodarki. Należy zaznaczyć, że miało to miejsce dopiero w drugiej, jesiennej edycji z 2003 r. (*Autumn Economic Forecast 2003*). Uznaje się, że jest to minimalne obciążenie; w przypadku pozostałych instytucji badanie obejmuje wszystkie prognozy sporządzone w 2003 r.

4. Miary oceny jakości predykcji *ex post*

Do oceny jakości predykcji zastosowano następujące statystyki opisowe: ME, MAE, RMSE (zależne od skali), statystykę U-Theila (niezależna od skali) oraz dekompozycję błędu kwadratowego. Podstawą do obliczania pierwszych trzech miar są błędy prognoz, czyli różnice pomiędzy wartością faktyczną a prognozą wygasłą. Błąd pojedynczej prognozy określa się następująco:

$$E_t = y_t - f_t \quad (1)$$

gdzie y_t – wartość faktyczna w okresie t , f_t – wartość prognozy na okres t .

W przypadku, gdy instytucja publikuje więcej niż jedną prognozę w roku:

$$E_{t(1)} = y_t - f_{t(1)}, E_t = y_t - f_{t(2)}, E_{t(n)} = y_t - f_{t(n)} \quad (2)$$

gdzie $E_{1,2,\dots,n}$ są błędami prognoz pochodzących z tego samego roku, $f_{1,2,\dots,n}$ to wartości prognoz.

W przypadku miary ME wynik może przyjmować wartości zarówno ujemne, jak i dodatnie. Na tej podstawie można wnioskować, czy prognozy były regularnie zaniżane lub zawyżane oraz z jakim natężeniem. Na skutek różnicy znaków błędy prognoz mogą się nawzajem znosić, więc miara ta nie

informuje o faktycznej wartości rozminięcia. Z tego powodu stosuje się także miarę MAE, informującą o bezwzględnej wartości błędu (Franek 2013). Należy pamiętać, że w przypadku ME znak dodatni interpretuje się jako zaniżanie wyniku przez prognozę, a ujemny jako zawyżanie.

W ogólnym rozrachunku miara MAE² wydaje się użyteczniejsza, ponieważ badacze z reguły wolą raczej wiedzieć, jakim ogólnym błędem jest obarczone podejście prognostyczne. Jednak w przypadku porównywania różnych instytucji interesujące jest, które z nich mają tendencję do zawyżania lub zaniżania swoich wyników (co można niekiedy interpretować jako optymizm lub pesymizm), co w prosty sposób pozwala śledzić miara ME. Należy jednak poczynić pewne założenia. Można sobie wyobrazić sytuację, w której szacowane wartości w większości przypadków (np. w 17 na 20) są dodatnie, oraz takie, gdy ich mniejsza część (np. 3 spośród 20) jest ujemna. Nie musi to oznaczać, że znak przed miarą ME świadczy o zawyżaniu, ponieważ w tym konkretnym przypadku jej wartości mogą być na tyle niskie, że dadzą wynik odwrotny (lub chociaż mocno zbliżony do zera), co może zatuszować występowanie bardzo częstych wartości dodatnich. Można jednak założyć, że prognozy instytucji międzynarodowych nie dają tak rozbieżnych wyników. Liczba wyższych oraz liczba niższych oszacowań powinny w nich być zbliżone.

Następnie zostanie zastosowana miara wyliczona na podstawie kwadratów błędów – pierwiastek błędu średniokwadratowego RMSE (ang. *root mean square error*). Statystyka ta w literaturze przedmiotu jest często przedstawiana jako alternatywa MAE. Główna różnica polega na tym, że RMSE wykazuje większą wrażliwość na odchylenia, co jest efektem podnoszenia ich do kwadratu. Można zaobserwować, że między tymi miarami występuje zależność: $MAE \leq RMSE$. W przypadku, gdy wszystkie błędy prognoz mają identyczne wartości ($E_1 = E_2 = \dots = E_n$, gdzie liczba prognoz wynosi n), MAE i RMSE są sobie równe. W literaturze różnie ocenia się zasadność stosowania jednej lub drugiej miary. Przykładowo Wilmott i Mastuura (2005) są zwolennikami stosowania MAE, a Chai i Draxler (2014) – RMSE. W tym artykule wartość RMSE będzie służyła do wnioskowania na temat występowania najbardziej nietrafionych prognoz.

Opisane miary charakteryzują się tym, że są zależne od skali (ang. *scale-dependent*). Oznacza to, że jeśli dane zjawisko jest mierzone w jakiejś jednostce, to wyniki tych miar również są wyrażone w tej jednostce (Hyndman 2014). Wobec PKB i inflacji jako jednostkę przyjmuje się średnioroczny procentowy wzrost, a wobec bezrobocia – jego stopę. Powoduje to, że nie można porównywać trafności prognoz poszczególnych zmiennych makroekonomicznych. Co więcej, nawet średniorocznego wzrostu nie można tutaj traktować jako tej samej jednostki. Taki sam wynik liczbowy ME lub MAE dla PKB oraz inflacji nie oznacza, że skuteczność prognozowania jest identyczna dla obu tych zmiennych (Hyndman 2006). W omawianej literaturze przedmiotu można zauważyć, że autorzy nie stosują się do tej zasady. Na przykład Oller i Barot (2000), a także Blix i in. (2001) na podstawie statystyki RMSE uznali, że inflacja jest łatwiejsza w prognozowaniu niż PKB. Niniejsza praca nie postuluje uznania takiego podejścia za błędne, jednak powstrzymuje się od wyciągania takich wniosków, dlatego do niektórych porównań stosowane będą tylko miary niezależne od skali.

Kolejną zastosowaną miarą będzie współczynnik U-Theila (1965), którego wartość może być uproszczona do następującej postaci (Makridakis, Wheelwright, Hyndman 1998):

² W literaturze można także spotkać alternatywną nazwę MAD – *mean absolute deviation* (por. Shim 2009; Hyndman 2005).

$$U = \sqrt{\frac{\sum_{t=n+1}^{n+k+1} \left(\frac{y_{t+1} - f_{t+1}}{y_t} \right)^2}{\sum_{t=n+1}^{n+k+1} \left(\frac{y_{t+1} - y_t}{y_t} \right)^2}} \quad (3)$$

Należy podkreślić, że dla modelu naiwnego prognoza wynosi $f_{t+1} = y_t$, a statystyka $U = 1$. Ponadto współczynnik U-Theila oparty jest na użyciu modelu naiwnego jako benchmarku, co pozwala na porównywanie jakości prognozowania różnych zmiennych (Hyndman 2006).

W niniejszej pracy zdecydowano się na rozszerzenie analiz jakości predykcji o niestosowaną w omawianej literaturze dekompozycję błędu średniokwadratowego MSE, zaproponowaną przez Theila (1966):

$$MSE = \left(\bar{y} - \bar{f} \right)^2 + \left(s_y - s_f \right)^2 + 2s_y s_f (1 - r_{yf}) \quad (4)$$

Na jej podstawie można wyodrębnić trzy elementy:

$$U_M = \frac{\left(\bar{y} - \bar{f} \right)^2}{MSE} \quad (5)$$

$$U_S = \frac{\left(s_y - s_f \right)^2}{MSE} \quad (6)$$

$$U_C = \frac{2s_y s_f (1 - r_{yf})}{MSE} \quad (7)$$

Muszą się one sumować do jedności:

$$U_M + U_S + U_C = 1 \quad (8)$$

Ich interpretacja wygląda następująco (Patrick, Teelucksingh 2002):

U_M – błąd wynikający z obciążenia (ang. *bias proportion*),

U_S – błąd wynikający z niedoszacowania wariancji (ang. *variance proportion*),

U_C – błąd wynikający z kowariancji (ang. *covariance proportion*).

U_M to błąd systematyczny; w tym przypadku mierzy on odległość między średnią prognoz wygłoszonych oraz realizacji. Oczekuje się, że wartości U_M będą możliwie bliskie zera (zero oznacza, że średnie są identyczne). U_S oznacza różnice wariancji i określa, w jakim stopniu zmienność badanego zjawiska została uchwycona przez prognozy. Wartości odległe od zera mogą świadczyć o silnym zaniżeniu lub zawyżeniu odchylenia standardowego badanej zmiennej. Składnik U_C , określony tutaj jako błąd wynikający z kowariancji, jest także nazywany składnikiem „szumu” (Polasek 2013). Reprezentuje bowiem pozostały błąd przypadkowy, czyli taki, który nie wynika z metod pomiaru. Uznaje się, że dla „dobrej” serii prognoz pierwsze dwa składniki powinny być równe zero ($U_M + U_S = 0$), a ostatni powinien wynosić jeden ($U_C = 1$). W takiej sytuacji cały błąd wynikałby z czysto przypadkowych wahań, niezwiązanych z metodą prognozowania.

5. Metoda porównywania prognoz przy braku pełnej zgodności zbiorów informacji

Ze względu na kolejność publikowania prognozy instytucji ekonomicznych można podzielić na pierwszorzeczne, drugorzeczne i trzeczorzeczne (dla NBP i KE). W przypadku porównań między instytucjami podział ten nie zawsze zapewnia zgodność tzw. zbiorów informacji (taki termin wprowadzili Blix i in. 2001). Można go rozumieć jako obecny stan informacji, na podstawie których formułuje się prognozę. Instytucje publikują swoje opracowania w różnych terminach (miesiącach), które nie zawsze są sztywno określone i bywają swobodnie zmieniane (zob. tabela 2). Z tego powodu prognoza na okres $t + 1$ sporządzona w grudniu roku t będzie miała przewagę nad tą, która ukazała się w np. styczniu czy maju roku t (nie oznacza to jednak, że zawsze będzie trafniejsza). Trzeba więc znaleźć sposób, by porównywać ze sobą z prognozy pochodzące z różnych momentów. Na przykład KE publikuje prognozy zgodnie z porami roku: wiosną, jesienią i od 2013 r. zimą. Według takiego podziału klasyfikowano tutaj prognozy badanych instytucji. Należy podkreślić, że jest on dosyć arbitralny, np. prognozy z września w przypadku MF traktuje się jako wiosenne, a w przypadku IMF³ jako jesienne. Przyczyna takiej decyzji jest złożona – stosowanie lub konstruowanie metod klasyfikowania prognoz do poszczególnych serii może powodować powstawanie szeregów o różnej długości lub nowych, bardzo krótkich, których porównywanie będzie mało zasadne. Dobrym przykładem może być klasyfikacja według kwartałów: jeśli instytucja przesunie prognozę np. z października (czwarty kwartał) na wrzesień (trzeci kwartał), jedna z serii traci obserwacje na rzecz innej. Tak było w przypadku IMF, który większość prognoz drugorzecznych publikuje w październiku (przed 2004 r. pojawiały się we wrześniu). W przypadku MF przez większość czasu prognozy ukazywały się we wrześniu, a dopiero od 2010 r. w przedziale od kwietnia do lipca i, jak wspomniano, wszystkie traktuje się jako serię wiosenną. Dodatkowo, podejście to pokrywa się w dużej mierze z porównywaniem prognoz według kolejności ich sporządzania, jednak zapewnia nieco większą zgodność pomiędzy zbiorami informacji. Dlatego serie prognoz będą porównywane w obrębie poszczególnych edycji (czyli ustalonych pór roku). Mogą się pojawić wnioski na temat różnic pomiędzy edycjami, jeśli zostaną uznane za wartość wspomnienia.

6. Ocena jakości prognoz

Prognozy dla każdej zmiennej ekonomicznej będą analizowane dla dwóch okresów: 2004–2016 i 2009–2016. Wartości prognoz zostały przedstawione na wykresach 1–4. Zastosowanie podejścia wizualnego pozwoliło zobrazować zjawiska, których nie da się dostrzec, analizując tylko miary liczbowe, zwłaszcza nietrafione przewidywanie punktów zwrotnych – problem bardzo często wspominany w literaturze przedmiotu. Punkty zwrotne są momentami, które oddzielają okresy rozwoju i spowolnienia ekonomicznego. Górny punkt zwrotny, zwany również szczytem, kończy okres wzrostu i analogicznie – dolny, zwany dnem, kończy okres spadków (Bry, Boschan 1971). Ze względu na krótki okres i charakter analiz każda zmiana kierunku jest w niniejszej pracy traktowana jako punkt zwrotny. W celu zaprezentowania, czym są omawiane załamania szeregu, stworzono wykres 5 (tylko realizacja PKB na podstawie wiosny, ponieważ dla inflacji i stopy bezrobocia realizacją dla każdej instytucji może być inna zmienna) służący jako przykład tych zjawisk.

³ Nie można ich znaleźć w tabeli 2, ponieważ taki system funkcjonował przed wybranym okresem. Uznano jednak, że jest to bardzo dobry przykład omawianego tu problemu.

6.1. Produkt krajowy brutto

Oceniona zostanie teraz jakość prognoz za pomocą miar ME, MAE, RMSE i współczynnika U-Theila. W przypadku PKB i prognoz rozpoczynających się w 2004 r. (tabela 3) w edycjach wiosennych najtrafniejsze były prognozy MF (najtrafniejsze ME, MAE, RMSE i U). Instytucje międzynarodowe osiągnęły bardzo zbliżone wartości, co dobrze widać na wykresie 1. Można zaobserwować bardzo podobne obserwacje odstające, zwłaszcza w latach kryzysu (2009 i 2010 r.) oraz przed kryzysem (w 2006 i 2007 r.), co mogło się przyczynić do uzyskania tak podobnych wartości RMSE. W przypadku jesiennych prognoz PKB wyniki uzyskane przez OECD były najdokładniejsze; potwierdzają to wszystkie miary, zwłaszcza MAE i RMSE. Statystyka U-Theila była bardzo podobna w wiosennych edycjach, a aktualniejsze prognozy miały – zgodnie z oczekiwaniami – niższe wartości. Wszystkie międzynarodowe instytucje wykazały się pesymizmem co do polskiego wzrostu gospodarczego, natomiast w MF był on minimalny. Co ciekawe, w KE i OECD obciążenie prognozy wiosennej było niższe niż prognozy jesiennej. W edycjach jesiennych można zaobserwować nieco większe trudności z przewidywaniem zmienności (wyższe U_s), podczas gdy w prognozach wiosennych ten problem nie występował (większość błędów znajduje się w składniku odpowiedzialnym za kowariancję U_c). W okresie 2009–2016 najniższym bezwzględnym błędem charakteryzują się jesienne prognozy OECD. W przypadku miary ME najlepiej wypadły późniejsze prognozy IMF. W tym przedziale czasowym widać ponaddwukrotną poprawę pomiędzy edycjami (w miarach ME i RMSE). Mimo krótszego okresu w prognozach wciąż można zaobserwować podobny problem z dekompozycją błędów średniokwadratowego; jesienne edycje charakteryzowały się znacznie wyższą wartością U_s . Uwagę zwraca także niepewność co do polepszenia się lub pogorszenia się polskiego PKB; w krótszym okresie (od 2009 r.) instytucje raz zawyżały (w wiosennych prognozach), a raz zaniżały jego wartość (w jesiennych). Tymczasem przyjmując rok 2004 jako początkowy, przeciętnie spodziewano się niższych wartości PKB. Zimowe prognozy NBP, mimo najmniejszego zbioru informacji, nie wypadły wyjątkowo słabo w stosunku do edycji wiosennych (lepsze MAE, RMSE i wskaźnik U-Theila). Warto dodać, że wyraźnie poprawia się trafność prognoz jesiennych. W badanych okresach (od 2004 r.) na podstawie wykresu 1 można zaobserwować występowanie korelacji pomiędzy prognozami a poprzednimi realizacjami (lata 2008 i 2009 oraz 2012 i 2013). Widać, że instytucje przewidywały wartości zbliżone do tych z poprzednich okresów. Mniej więcej od lat 2009–2010 zwiększyła się trafność przewidywania poziomu i zmian kierunku tempa wzrostu PKB, zwłaszcza w przypadku NBP, co widać na wykresie 4. Można uznać, że prognozowanie punktów zwrotnych przed kryzysem i podczas niego było o wiele trudniejsze niż po jego ustąpieniu (wciąż jednak bez stuprocentowej skuteczności). Ciekawy jest także fakt, że w badanym okresie wszystkim instytucjom udało się przewidzieć zwrot kierunku PKB w 2011 r. przynajmniej jedną z serii swoich prognoz. Prognozy ze wszystkich edycji (w obu przedziałach czasowych) były trafniejsze niż prognozy naiwne.

6.2. Inflacja CPI

Tym, co rzuca się w oczy podczas analizy prognoz inflacji od 2004 r. (tabela 4), są wyjątkowo wysokie wartości statystyki U-Theila; w edycjach wiosennych wartość statystyki zazwyczaj przekracza 3,0. Można, co prawda, zaobserwować pewną poprawę w przypadku prognoz jesiennych, jednak wciąż nie są one trafniejsze niż prognozy naiwne. Autor przypuszcza, że częściowo wynika to z wystąpienia punk-

tu zwrotnego po 2012 r., w którym nastąpił silny i niespodziewany spadek inflacji. Dobrze widać to na wykresie 2 – żadna instytucja prognozująca nie spodziewała się ujemnej wartości inflacji. Analizując pozostałe miary prognoz z edycji wiosennych, można zauważyć, że najlepszą wartość statystyki ME uzyskało polskie MF. Należy jednak zwrócić uwagę na stosunkowo wysokie wartości statystyki RMSE w prognozach MF. Można się domyślać, że w dużej mierze przyczyniły się do tego obserwacje po sytuacji z 2012 r. Prognozy podmiotów międzynarodowych nieco bardziej zbliżyły się do faktycznego poziomu, jednak wciąż miały ograniczoną skuteczność. W prognozach jesiennych najlepiej wypadało OECD – miało widoczną przewagę, zwłaszcza jeśli chodzi o miarę ME. Jedynie w przypadku MAE prognozy OECD były minimalnie mniej trafne niż KE (różnica niewidoczna przy zaokrągleniu do drugiej cyfry po przecinku). Na wykresie 2 (wykres 4 dla NBP) można dostrzec, że prognozy sprzed 2010 r. były podobne do prognozy naiwnej, po czym wróciły na właściwy tor i nieco trafniej przewidywały poziomy inflacji (jest to widoczne zwłaszcza w prognozach jesiennych), aż do wspomnianego spadku inflacji w 2012 r. Ciekawe wydaje się także to, że wszystkie instytucje w każdej edycji spodziewały się wyższej inflacji niż faktycznie miała miejsce. Nasiliło się to po zmianie roku analizy na 2009, więc można wnioskować, że okres pokryzysowy był główną przyczyną zawyżania prognoz. Interesujący jest także fakt, że prognoza wiosenna NBP była trafniejsza nawet od edycji jesiennych. Można też zaobserwować, że większa część błędu średniokwadratowego znajduje się w składniku odpowiedzialnym za błąd przypadkowy (U_c) niż w późniejszej edycji. Sytuacja ta jest mniej jednoznaczna dla pozostałych instytucji. Dla nich nie można zaobserwować wyraźnych prawidłowości w zmianie kompozycji błędu średniokwadratowego.

6.3. Stopa bezrobocia

W przypadku prognozowania stopy bezrobocia przez MF (tabela 5) można zaobserwować, że w edycjach wiosennych od 2004 r. wartość ME była bardzo bliska zera. Spośród tych edycji prognozy MF były najtrafniejsze we wszystkich statystykach niezależnych od miary; wszystkie analizowane instytucje miały bardzo podobną wartość współczynnika U-Theila, około 1,3 (można zaobserwować wyraźną poprawę tej statystyki pomiędzy edycjami). W przypadku prognoz jesiennych należy zwrócić uwagę, że tylko KE spodziewała się niższej stopy bezrobocia niż ta, która faktycznie miała miejsce (dodatnia wartość statystyki ME). Prognozowała także stosunkowo wysokie wartości RMSE, zwłaszcza w porównaniu z prognozami OECD. Były one ponad dwa razy większe, co sugeruje występowanie kilku stosunkowo dużych rozminięć u KE. Na wykresie 3 można się ich doszukiwać w początkowych okresach (przed 2009 r., zwłaszcza w 2004 r.). W okresie 2009–2016 prognozy NBP były najtrafniejsze spośród wiosennych edycji dla każdej z miar. Z kolei prognozy MF cechowały się najmniejszą trafnością (z wyjątkiem ME). Można tutaj zaobserwować pewną rozbieżność co do oczekiwanej stopy bezrobocia wśród polskich instytucji – NBP spodziewał się wyższej, a MF niższej wartości, podczas gdy wszystkie instytucje międzynarodowe zgodnie ją zawyżały. Stosunek MAE do RMSE wydaje się wysoki w zimowej i wiosennej prognozie NBP, co jest dobrze widoczne na wykresie 3 – jesiennie prognozy wierniej podążają za faktycznymi poziomami stopy bezrobocia.

W jesiennych edycjach prognozy instytucji międzynarodowych okazały się lepsze niż NBP i osiągnęły wartość statystyki ME bliższą zera. Były nieco bardziej zbliżone do siebie w przypadku MAE, a jeszcze bardziej RMSE. Statystyka U-Theila w wiosennych i zimowych prognozach NBP przekroczyła

wartość 1. W jesiennych edycjach miary trafności prognoz wybranych instytucji przyjęły wartość poniżej jedności; najtrafniejsze były prognozy NBP i zbliżone do niej OECD. Dekompozycja statystyki U w prognozach wiosennych wydaje się stosunkowo dobra, zwłaszcza MF, w której większość błędu koncentruje się w składniku odpowiedzialnym za błąd przypadkowy U_c , podczas gdy w pozostałych prognozach niewielka część błędu rozkłada się między dwa pozostałe składniki (U_m , U_s). Podobna jest sytuacja w jesiennych edycjach instytucji międzynarodowych – można w nich zaobserwować, że wkład do wartości U jeszcze bardziej koncentruje się w ostatnim składniku. Jedyne w przypadku NBP znajduje się w nietrafionej średniej wartości prognoz (U_m). Zimowa prognoza NBP okazała się najmniej trafna spośród wszystkich edycji (z wyjątkiem MAE). Podobna sytuacja miała miejsce w prognozach jesiennych mierzonych za pomocą dekompozycji błędu średniokwadratowego, jednak tutaj część błędu znalazła się w składniku odpowiedzialnym za wariancję U_s . Po przeanalizowaniu obu badanych okresów można także stwierdzić, że edycje wiosenne i zimowa były mniej trafne niż prognozowanie naiwne (U-Theila powyżej jedności), w przeciwieństwie do jesiennych, które były skuteczniejsze.

Analizując wykresy 1–4, można potwierdzić, że trafne przewidywanie punktów zwrotnych sprawia trudności instytucjom prognozującym. Potwierdziło się sformułowane na początku przypuszczenie, że sytuacja będzie podobna, jak opisywano w literaturze przedmiotu. Za najbardziej szczególny punkt zwrotny można uznać spadek inflacji z 2012 r., którego nie przewidziała żadna z instytucji. Co więcej, można w tej sytuacji doszukiwać się także występowania autokorelacji błędów – mimo utrzymywania się ujemnej inflacji do końca badanego okresu żadna instytucja z wyjątkiem NBP nie prognozowała ujemnej wartości. Wszystkie instytucje pominęły szczyt dla PKB w 2007 r. oraz podwójne dno z lat 2008–2009, a także kolejne podwójne dno w okresie 2012–2013. W tym ostatnim przypadku jedynie prognozy NBP w dużym stopniu pokrywały się z faktycznymi wartościami. W odniesieniu do stopy bezrobocia sytuacja była podobna, jednak w przebiegu tej zmiennej można odnotować mniej punktów zwrotnych niż u poprzednich. Tak jak w przypadku inflacji z 2012 r. można tu doszukiwać się autokorelacji – mimo zmiany kierunku i wyraźnej zmiany poziomów prognozy często mają tendencję do utrzymywania podobnych wartości.

7. Podsumowanie

W pracy przedstawiono wyniki analizy trafności prognoz dla polskiej gospodarki w kontekście problemów poruszanych w wybranej literaturze przedmiotu. Badano także zagadnienie ustalania zbiorów informacji, według których klasyfikowano prognozy w sytuacji, gdy występuje niezgodność momentów ich sporządzania.

Analizę trafności prognoz przeprowadzono dla pięciu instytucji: NBP, MF, KE, OECD i IMF. Podzielono ją na dwa okresy ze względu na dostępność danych. Tabela 6 podsumowuje wyniki analizy. W okresie 2004–2016 można zaobserwować dużą trafność prognoz OECD, zwłaszcza jesiennych edycji, która osiem razy okazała się najcelniejsza. Prognozy polskiego MF były dwa razy najtrafniejsze dla miary ME, z wynikami bardzo bliskimi zera, co oznacza, że nie miało wyraźnych tendencji do pesymizmu czy optymizmu. KE osiągnęła wyniki podobne jak OECD, które okazały się trafniejsze tylko dwa razy. W krótszym okresie (2009–2016) prognozy OECD również były najczęściej trafne (sześć razy), z niewielką przewagą nad NBP, które było najtrafniejsze cztery razy. W tym okresie w odniesieniu do inflacji

został przełamany wcześniejszy schemat i aż trzy razy najtrafniejsza była prognoza wiosenna. Raz najtrafniejsza była prognoza KE i raz IMF, jednak w przypadku tej drugiej instytucji nie brano pod uwagę prognoz stopy bezrobocia.

Dla przeprowadzonych analiz bardzo ważne jest to, że ocenę sporządzono na podstawie skończonej liczby obserwacji. Korzystając z takiej próby, należy ostrożnie weryfikować hipotezy dotyczące różnic między trafnością prognoz – nie można bowiem badać mocy statystycznej zaobserwowanych wyników. Należy także pamiętać, że na ich kształt silnie wpływa metoda klasyfikowania prognoz do porównywanych serii. W pracy przedstawiono jedno z bardziej intuicyjnych podejść, pozwalające ominąć problem z potencjalną niezgodnością obserwacji, z którym może się wiązać stosowanie bardziej rozbudowanych metod.

Tak jak przypuszczano, prognozy oparte na nowszych zbiorach informacji są z reguły trafniejsze. We wszystkich przypadkach sprawdza się to dla statystyki U-Theila, ale dla miar zależnych od skali występują pewne wyjątki. Świadczą o tym chociażby bardzo dobre wyniki wiosennych prognoz inflacji opracowywanych w NBP. W przeciwieństwie do Ollera i Barota (2012) oraz Blix i in. (2001) nie można powiedzieć, że wszystkie analizowane prognozy są trafniejsze niż prognozy naiwne. Tylko w przypadku PKB wartości statystyki U-Theila okazały się mniejsze od jedności, a w odniesieniu do inflacji były zdecydowanie powyżej tej wartości. Można podejrzewać, że przyczynił się do tego jej nagły w spadek w 2012 r. oraz krótka seria danych. Można zatem stwierdzić, że wniosek Ollera i Barota (2000) oraz Blix i in. (2001), że prognozowanie PKB jest trudniejsze niż prognozowanie inflacji, nie potwierdza się w odniesieniu do polskiej gospodarki w badanym okresie.

Analiza wartości statystyki ME pozwala na sformułowanie pewnych wniosków. Inflacja była wyraźnie zawyżana przez wszystkie instytucje w obu uwzględnionych okresach. Zawyżane były również prognozy stopy bezrobocia, z niewielkimi wyjątkami – KE w okresie 2004–2016 i MF w latach 2009–2016. W okresie rozpoczynającym się w 2004 r. można zaobserwować tendencję do zaniżania prognoz PKB, która zanika po skróceniu okresu do lat 2009–2016.

Na początku pracy sformułowano założenie, że najwygodniej byłoby, gdyby większa trafność instytucji przejawiała się we wszystkich miarach jakości predykcji. Podejrzewano, że w praktyce jest to raczej mało prawdopodobne. Nie sprawdziło się jednak to w takim stopniu, jak zakładano. Na pewno nie występuje wyraźne uszeregowanie się instytucji według trafności miar jakości predykcji *ex post*, w niektórych przypadkach można jednak zauważyć, że prognozy którejs z instytucji są najtrafniejsze lub najmniej trafne dla wszystkich analizowanych statystyk. W latach 2004–2016 dotyczy to wiosennych prognoz PKB i stopy bezrobocia sporządzonych w MF, a w okresie 2009–2016 – inflacji i stopy bezrobocia w wiosennych prognozach NBP.

We wstępie wspomniano, że prognozy badanych instytucji krajowych mają charakter planu, a zatem powinny być trafniejsze. Nie można tego w pełni potwierdzić. Były one trafniejsze tylko w określonych przypadkach – w okresie uwzględniającym NBP (2009–2016) wiosenna prognoza tej instytucji osiągnęła bardzo dobre wyniki, a w dwóch przypadkach (PKB i stopa bezrobocia od 2004) prognoza MF miała ME bardzo bliskie zera.

W pewnym stopniu potwierdza się także wniosek z pracy Gonzales-Cabanallis i Terzi (2012), że czas kryzysu pogarsza trafność prognoz. W okresie, który objął kryzys (2004–2016), można zaobserwować wyższe wartości miar jakości predykcji w przypadku PKB niż po 2009 r. Nie dotyczy to wszystkich zmiennych – w przypadku inflacji niektóre instytucje w dłuższym okresie wykazywały się większą trafnością (KE, OECD i IMF). Sytuacja jest jeszcze bardziej złożona w odniesieniu do stopy bezrobocia. W krótszym okresie objętym analizą prognozy MF okazały się gorsze, a prognozy instytucji międzynarodowych – lepsze.

Gonzalez-Cabanillas i Terzi (2012) porównywali także prognozy KE z pozostałymi uwzględnionymi tutaj instytucjami międzynarodowymi i stwierdzili, że ich trafność jest zbliżona do trafności prognoz OECD i IMF. W przypadku polskiej gospodarki również można doszukiwać się pewnych podobieństw. Dobrze obrazuje to m.in. przebieg prognoz PKB (wykres 1), w badanych instytucjach międzynarodowych był on zbliżony. Można zaobserwować, że w niektórych przypadkach miary jakości prognoz przyjmowały bardzo podobne wartości, co przejawia się najsilniej poprzez ME i RMSE. Jest to widoczne zwłaszcza w wiosennych i jesiennych prognozach PKB od 2004 r. (w tych drugich jedynie prognozy KE i IMF miały zbliżone wartości, OECD była od nich trafniejsza). Podobieństwo występuje także w przypadku MAE w prognozach inflacji pochodzących z jesiennych edycji (2004) sporządzonych przez KE i OECD. Kolejne można zaobserwować w krótszym okresie 2009–2016 dla stopy bezrobocia (tylko dla KE i OECD, IMF nie bierze udziału w analizie tej zmiennej). Co ciekawe, gdy spojrzeć się na lata 2004–2016, widać już znaczną rozbieżność w wartościach miar jakości predykcji w tych dwóch instytucjach.

Bibliografia

- Acedański J., Bernais J., Mastalerz-Kodzis A. (2014), Dokładność wybranych metod prognozowania wynagrodzeń i liczby pracujących w Polsce, *Bank i Kredyt*, 45(2), 163–196.
- Baranowski P., Leszczyńska A., Szafranski G. (2010), Krótkookresowe prognozowanie inflacji z użyciem modeli czynnikowych, *Bank i Kredyt*, 41(4), 28–44.
- Blix M., Wadefjord J., Wienecke U., Adahl M. (2001), How good is the forecasting performance of major institutions, *Economic Review (Penning – OchValutapolitk)*, 3/2001, 37–68.
- Bry G., Boschan C. (1971), Programmed selection of cyclical turning points, *Cyclical Analysis of Time Series: Selected Procedures and Computer Programs*, National Bureau of Economic Research.
- Budnik K. i in. (2009), *The new macroeconometric model of the Polish economy*, National Bank of Poland Working Paper, 62.
- Chai T., Draxler R. (2014), Root mean square error (RMSE) or mean absolute error (MAE)? Arguments against avoiding RMSE in the literature, *Geoscientific Model Development*, 7, 1247–1250.
- Franek S. (2013), *Wieloletnie planowanie budżetowe*, Difin.
- Gonzales-Cabanillas L., Terzi A. (2012), The accuracy of the European Commission's forecasts re-examined, *Economic Papers*, 476.
- Greszta M., Hulej M., Krzesicki O., Lewińska R., Pońsko P., Rybaczyk B., Tarnicka M. (2011), *Reestymacja kwartalnego modelu gospodarki polskiej NECMOD 2011*, Narodowy Bank Polski.
- Greszta M., Hulej M., Lewińska R., Michałek A., Pońsko P., Rybaczyk B., Schulz B. (2010), *Reestymacja kwartalnego modelu gospodarki polskiej NECMOD 2010*, Narodowy Bank Polski.
- Greszta M., Hulej M., Lewińska R., Michałek A., Pońsko P., Rybaczyk B., Schulz B. (2012), *Reestymacja kwartalnego modelu gospodarki polskiej NECMOD 2012*, Narodowy Bank Polski.
- Hałka A., Leszczyńska A. (2011), Wady i zalety wskaźnika cen towarów i usług konsumpcyjnych – szacunki obciążenia dla Polski, *Gospodarka Narodowa*, 9(241), 51–57.
- Hyndman R. (2006), Another look at measures of forecast accuracy, *Foresight*, 4, 43–46.
- Hyndman R. (2014), Measuring forecast accuracy, w: M. Gilliland, L. Tashman, U. Sglavo, *Business Forecasting: Problems and Solutions*, Wiley & Sons.

- IMF (2017a), *The IMF at a glance*, <http://www.imf.org/external/np/exr/facts/pdf/glance.pdf>.
- IMF (2017b), *World Economic Outlook (WEO)*, <https://www.imf.org/external/pubs/ft/weo/faq.htm>.
- IMF (2017c), *Database – WEO groups and aggregates information*, <https://www.imf.org/external/pubs/ft/weo/2017/01/weodata/groups.htm#w>.
- EC (2012), *The European Union explained – How the European Union works*, European Commission, Publications Office of the European Union.
- EC (2016), *Management Plan 2016 ECFIN*, Planning and Management Documents, European Commission.
- Kreini M. (2000), Accuracy of OECD and IMF projection, *Journal of Policy Modelling*, 22(1), 61–79.
- Makarova S. (2014), *Risk and uncertainty: macroeconomic perspective*, UCL School of Slavonic and East European Studies Economics and Business Working Paper, 129.
- Makridakis S., Wheelwright S., Hyndman R. (1998), *Forecasting Methods and Applications*, Willey & Sons.
- NBP (2016), *Założenia polityki pieniężnej na rok 2017*, Narodowy Bank Polski.
- Nowak A. (2004), Organizacje integracji gospodarczej państw wysoko uprzemysłowionych, w: T. Łoś-Nowak (red.), *Organizacje w stosunkach międzynarodowych*, Wydawnictwo Uniwersytetu Wrocławskiego.
- OECD (2011), *Forecasting methods and analytical tools*, http://www.oecd.org/eco/outlook/forecasting_methodsandanalyticaltools.htm.
- Oller L., Barot B. (2000), The accuracy of European growth and inflation forecast, *International Journal of Forecasting*, 16, 293–315.
- Owsiak S. (2013), Wieloletnie planowanie finansowe a funkcjonowanie budżetu zadaniowego, *Studia BAS*, 1(33), 37–59.
- Patrick W., Teelucksingh S. (2002), *A Practical Introduction to Econometric Methods: Classical and Modern*, University of West Indies Press.
- Polasek W. (2013), *Forecast evaluation for multiple time series: a generalized Theil decomposition*, Working Papers of the Rimini Centre for Economic Analysis, 23(2013).
- Theil H. (1965), *Economic Forecast and Policy*, North Holland Publishing Company.
- Theil H. (1966), *Applied Economic Forecasting*, North Holland Publishing Company.
- Wilmott C., Mastuura K. (2005), Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance, *Climate Research*, 30, 79–82.

Aneks

Tabela 1

Realizacje badanych zmiennych makroekonomicznych

Rok	PKB	CPI	HICP	Skorygowane bezrobocie BAEL (NBP)	Bezrobocie rejestrowane (MF)	Stopa bezrobocia – Eurostat (KE i OECD)
2004	5,1	3,4	3,6	–	19	19,0
2005	3,5	2,2	2,2	–	17,6	17,8
2006	6,2	1,3	1,3	–	14,8	13,8
2007	7,2	2,5	2,6	–	11,2	9,6
2008	3,9	4,2	4,2	–	9,5	7,1
2009	2,6	3,8	4	8,2	12,1	8,2
2010	3,7	2,6	2,6	9,6	12,4	9,6
2011	5,0	4,2	3,9	9,6	12,5	9,6
2012	1,6	3,6	3,7	10,1	13,4	10,1
2013	1,3	1,0	0,8	10,3	13,4	10,3
2014	3,3	0,1	0,1	9,0	11,5	9,0
2015	3,6	-0,9	-0,7	7,5	9,7	7,5
2016	2,7	-0,7	-0,2	6,2	8,3	6,2

Uwaga: stopa bezrobocia stosowana przez IMF nie została uwzględniona w analizie.

Źródło: opracowanie na podstawie danych GUS, NBP, OECD i Eurostatu.

Tabela 2
Miesiące publikowania prognoz przez poszczególne instytucje

Rok publikacji	NBP			MF		KE*		OECD		IMF	
	1 zima	2 wiosna	3 jesień	1 wiosna	0,5 zima	1 wiosna	2 jesień	1 wiosna	2 jesień	1 wiosna	2 jesień
2015	marzec	lipiec	listopad	czerwiec	luty	maj	listopad	czerwiec	listopad	październik	kwiecień
2014	marzec	lipiec	listopad	czerwiec	luty	maj	listopad	maj	listopad	październik	kwiecień
2013	marzec	lipiec	listopad	czerwiec	styczeń	maj	listopad	maj	listopad	październik	kwiecień
2012	marzec	lipiec	listopad	czerwiec	-	maj	listopad	maj	listopad	październik	kwiecień
2011	luty	czerwiec	październik	kwiecień	-	maj	listopad	maj	listopad	październik	kwiecień
2010	luty	czerwiec	październik	lipiec	-	maj	listopad	maj	listopad	październik	kwiecień
2009	luty	czerwiec	październik	wrzesień	-	maj	listopad	czerwiec	listopad	październik	kwiecień
2008	-	-	-	wrzesień	-	kwiecień	listopad	czerwiec	grudzień	październik	kwiecień
2007	-	-	-	wrzesień	-	maj	listopad	czerwiec	grudzień	październik	kwiecień
2006	-	-	-	wrzesień	-	maj	listopad	czerwiec	grudzień	październik	kwiecień
2005	-	-	-	wrzesień	-	kwiecień	listopad	czerwiec	grudzień	październik	kwiecień
2004	-	-	-	wrzesień	-	kwiecień	listopad	czerwiec	grudzień	październik	kwiecień
2003	-	-	-	wrzesień	-	kwiecień**	listopad	czerwiec	grudzień	październik	kwiecień

* Po przejściu na system prognozowania trzy razy w roku dodano edycję zimową jako pierwszą w roku, edycja wiosenna (do tej pory pierwsza) stała się drugą serią prognoz w roku, a jesienna trzecioroczną (do tej pory druga). Aby ułatwić interpretację, prognozę zimową traktuje się tutaj jako prognozę nr 0,5, w przeciwieństwie do NBP, gdzie w całym badanym okresie używano systemu trzech prognoz w roku.

** W tej edycji dokumentu nie sporządzono prognozy wszystkich badanych zmiennych dla polskiej gospodarki i z tego względu nie uwzględniono jej w analizie.

Źródło: opracowanie na podstawie danych NBP, GUS, Ministerstwa Finansów, Eurostatu, OECD i IMF.

Tabela 3

Miary jakości predykcji *ex post* PKB dla okresów 2004–2016 i 2009–2016

Statystyka	NBP	NBP	MF	KE	OECD	IMF	NBP	KE	OECD	IMF	Najtrafniejsza prognoza ogółem
	Nr prog.	zima	wiosna				jesień				
2004–2016											
ME	–	–	0,03%	0,07%	0,13%	0,34%	–	0,39%	0,38%	0,22%	MF(W)
MAE	–	–	1,42%	1,64%	1,64%	1,60%	–	1,32%	1,11%	1,32%	OECD(J)
RMSE	–	–	1,61%	1,80%	1,81%	1,81%	–	1,46%	1,29%	1,45%	OECD(J)
U	–	–	0,9	1,0	1,0	1,0	–	0,66	0,54	0,70	OECD(J)
U_M	–	–	0,00	0,00	0,01	0,03	–	0,07	0,08	0,02	
U_S	–	–	0,07	0,04	0,05	0,16	–	0,11	0,20	0,21	
U_C	–	–	0,93	0,95	0,94	0,80	–	0,82	0,71	0,77	
2009–2016											
ME	-0,25%	-0,31%	-0,15%	-0,07%	-0,21%	-0,12%	0,18%	0,20%	0,14%	-0,02%	IMF(J)
MAE	1,33%	1,01%	1,35%	1,57%	1,60%	1,54%	0,74%	0,97%	0,72%	1,04%	OECD(J)
RMSE	1,50%	1,32%	1,53%	1,78%	1,83%	1,68%	0,96%	1,06%	0,79%	1,10%	OECD(J)
U	0,69	0,60	0,89	0,96	1,04	1,02	0,49	0,60	0,43	0,63	OECD(J)
U_M	0,03	0,06	0,01	0,00	0,01	0,01	0,03	0,04	0,03	0,00	
U_S	0,05	0,00	0,01	0,00	0,00	0,02	0,13	0,14	0,39	0,20	
U_C	0,93	0,94	0,99	1,00	0,98	0,97	0,84	0,83	0,58	0,80	

Tabela 4

Miary jakości predykcji *ex post* inflacji w latach 2004–2016 i 2009–2016

Statystyka	NBP		MF	KE	OECD	IMF	NBP		KE	OECD	IMF	Najtrafniejsza prognoza ogółem
	Nr prog.	zima	wiosna				jesień					
2004–2016												
ME	–	–	-0,13%	-0,36%	-0,33%	-0,41%	–	-0,35%	-0,10%	-0,32%	OECD(J)	
MAE	–	–	1,50%	1,26%	1,32%	1,17%	–	1,07%	1,07%	1,18%	KE(J)	
RMSE	–	–	1,71%	1,42%	1,49%	1,49%	–	1,23%	1,18%	1,32%	OECD(J)	
U	–	–	3,4	3,2	3,0	3,5	–	2,21	1,56	1,80	OECD(J)	
U_M	–	–	0,01	0,06	0,05	0,08	–	0,08	0,01	0,06		
U_S	–	–	0,47	0,54	0,19	0,61	–	0,13	0,62	0,50		
U_C	–	–	0,52	0,40	0,76	0,32	–	0,79	0,38	0,45		
2009–2016												
ME	-0,38%	-0,20%	-0,56%	-0,60%	-0,82%	-0,79%	-0,41%	-0,29%	-0,30%	-0,51%	NBP(W)	
MAE	1,22%	0,96%	1,83%	1,37%	1,50%	1,41%	1,11%	1,09%	1,24%	1,21%	NBP(W)	
RMSE	1,49%	1,22%	1,98%	1,55%	1,65%	1,76%	1,26%	1,27%	1,34%	1,37%	NBP(W)	
U	2,87	2,44	3,42	3,22	3,04	3,52	2,12	2,23	1,57	1,81	OECD(J)	
U_M	0,06	0,03	0,08	0,15	0,25	0,20	0,11	0,05	0,05	0,14		
U_S	0,30	0,16	0,49	0,53	0,14	0,51	0,14	0,37	0,75	0,65		
U_C	0,64	0,81	0,43	0,32	0,61	0,29	0,75	0,58	0,20	0,21		

Tabela 5

Miary jakości predykcji *ex post* stopy bezrobocia w latach 2004–2016 i 2009–2016

Statystyka	NBP		MF	KE	OECD	IMF	NBP		OECD	IMF	Najtrafniejsza prognoza ogółem
	NBP	zima					KE	jesień			
Nr prog.	zima		wiosna			jesień					
2004–2016											
ME	–	–	-0,10%	-1,37%	-1,40%	–	–	0,34%	-0,78%	–	MF(W)
MAE	–	–	1,85%	2,19%	2,01%	–	–	1,98%	1,08%	–	OECD(J)
RMSE	–	–	2,12%	2,54%	2,45%	–	–	3,83%	1,43%	–	OECD(J)
U	–	–	1,3	1,3	1,3	–	–	0,69	0,72	–	KE(J)
U_M	–	–	0,00	0,29	0,33	–	–	0,01	0,29	–	
U_S	–	–	0,01	0,12	0,05	–	–	0,00	0,02	–	
U_C	–	–	0,99	0,59	0,62	–	–	0,99	0,68	–	
2009–2016											
ME	-1,18%	-0,35%	0,44%	-0,46%	-0,47%	–	-0,53%	-0,23%	-0,25%	–	KE(J)
MAE	1,57%	1,04%	2,11%	1,69%	1,46%	–	0,85%	0,78%	0,74%	–	OECD(J)
RMSE	2,02%	1,44%	2,34%	1,80%	1,58%	–	1,03%	1,02%	0,91%	–	OECD(J)
U	1,94	1,39	1,71	1,59	1,43	–	0,75	0,87	0,77	–	NBP(J)
U_M	0,34	0,06	0,04	0,06	0,09	–	0,27	0,05	0,07	–	
U_S	0,14	0,12	0,00	0,06	0,04	–	0,00	0,02	0,01	–	
U_C	0,52	0,82	0,96	0,88	0,87	–	0,73	0,93	0,92	–	

■ wartość najbliższej 0, ■ wartość najdalej 0

Źródło: opracowanie na podstawie danych NBP, GUS, Ministerstwa Finansów, Eurostatu, OECD i IMF.

Tabela 6

Podsumowanie analiz: najtrafniejsza prognoza dla każdej statystyki (i zmiennej) oraz poszeregowana częstotliwość trafności dla każdej edycji i ogółem* w podziale na okresy

2004–2016			
Statystyka	PKB	Inflacja	Stopa bezrobocia
ME	MF(W)	OECD(J)	MF(W)
MAE	OECD(J)	KE(J)	OECD(J)
RMSE	OECD(J)	OECD(J)	OECD(J)
U2-Theila	OECD(J)	OECD(J)	KE(J)
Institucja	Wiosna	Jesień	Ogółem (wiosna i jesień)
OECD	0	8	8
MF	2	0	2
KE	0	2	2
IMF**	0	0	0
2009–2016			
Statystyka	PKB	Inflacja	Stopa bezrobocia
ME	IMF(J)	NBP(W)	KE(J)
MAE	OECD(J)	NBP(W)	OECD(J)
RMSE	OECD(J)	NBP(W)	OECD(J)
U2-Theila	OECD(J)	OECD(J)	NBP(J)
Institucja	Wiosna	Jesień	Ogółem (wiosna i jesień)
OECD	0	6	6
NBP***	3	1	4
KE	0	1	1
IMF**	0	1	1
MF	0	0	0

* Częstotliwość trafności (w tabeli: wiosna, jesień i ogółem) obliczana jest na zasadzie, jak często edycja danej instytucji trafiła do ostatniej kolumny tabel 3–5 (Najtrafniejsza prognoza ogółem).

** Nie uwzględniono w analizie stopy bezrobocia IMF.

*** Nie brano pod uwagę zimowych prognoz NBP.

Źródło: opracowanie na podstawie danych NBP, GUS, Ministerstwa Finansów, Eurostatu, OECD i IMF.

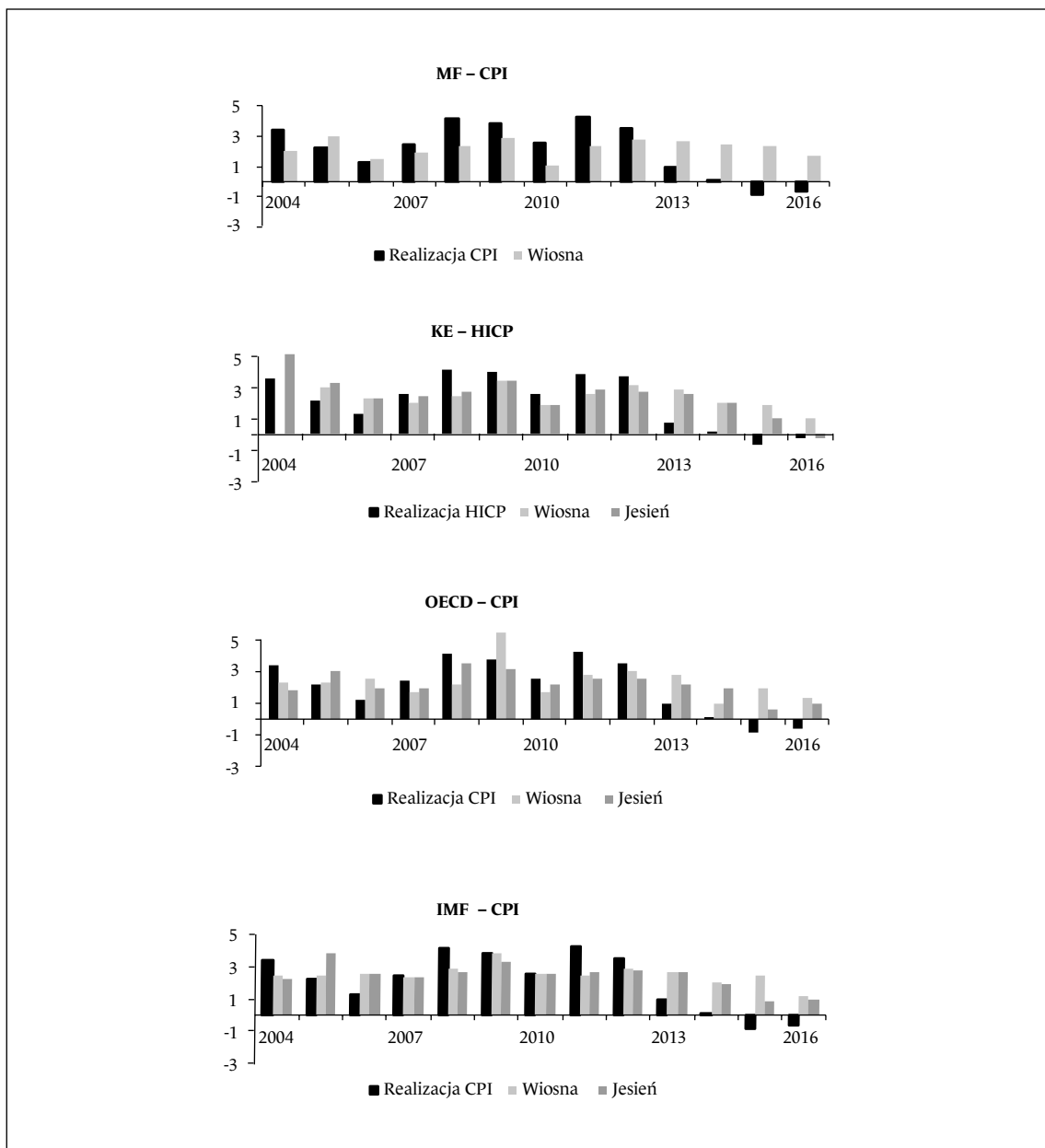
Wykres 1

Realizacja oraz prognozy PKB dla MF, KE OECD i IMF



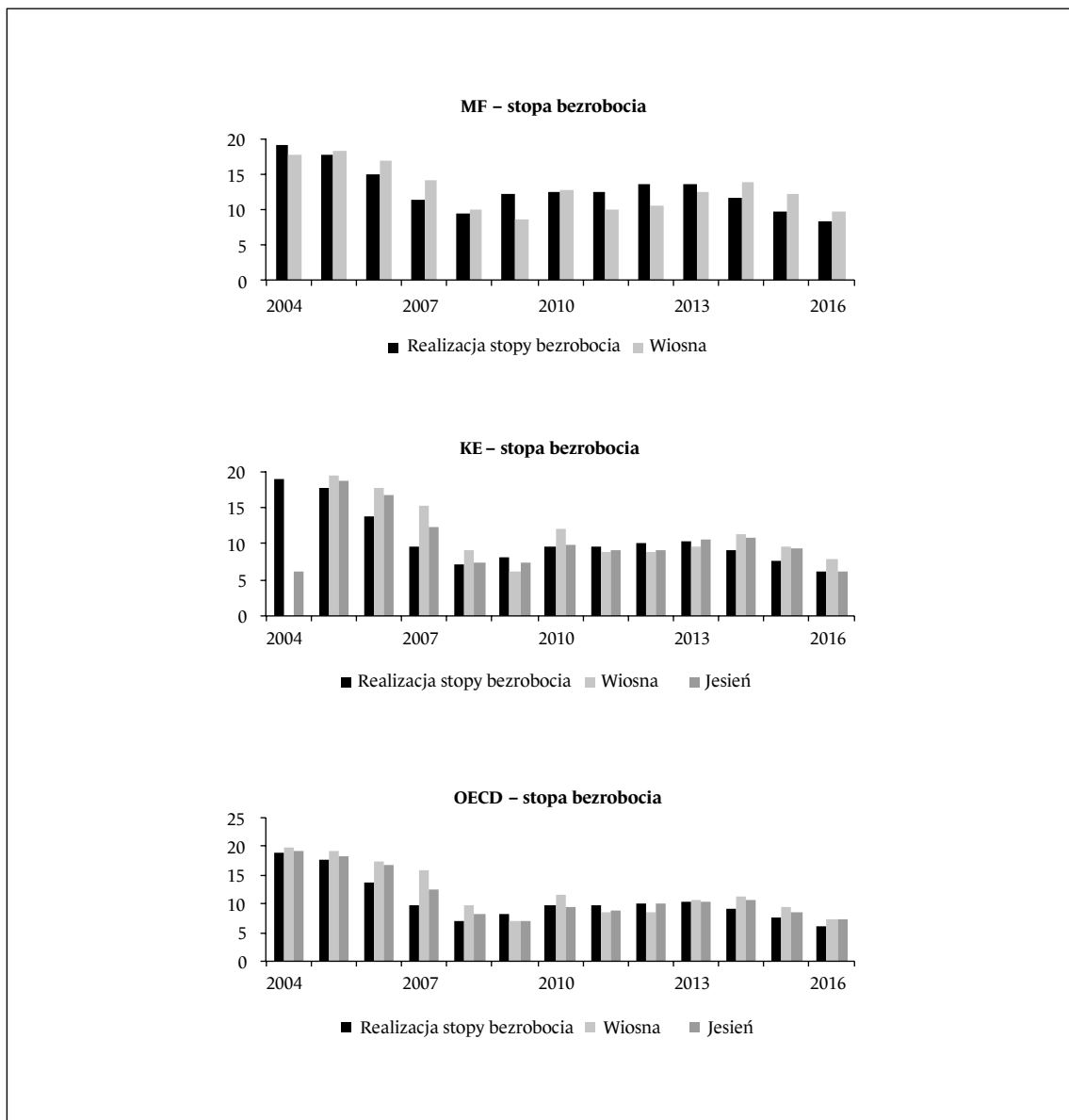
Wykres 2

Realizacja oraz prognozy inflacji dla MF, KE OECD i IMF



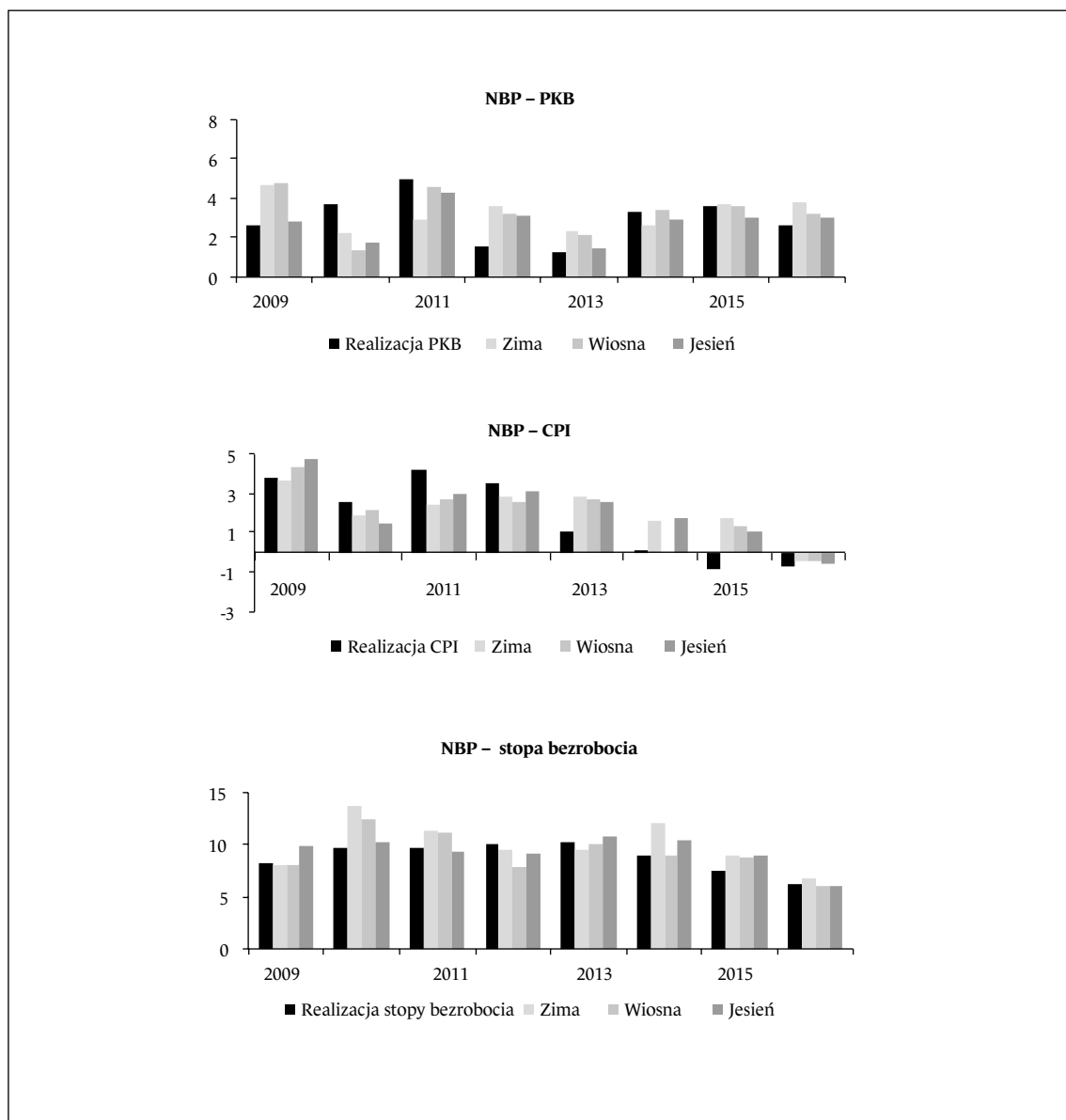
Wykres 3

Realizacja oraz prognozy stopy bezrobocia dla MF, KE i OECD



Wykres 4

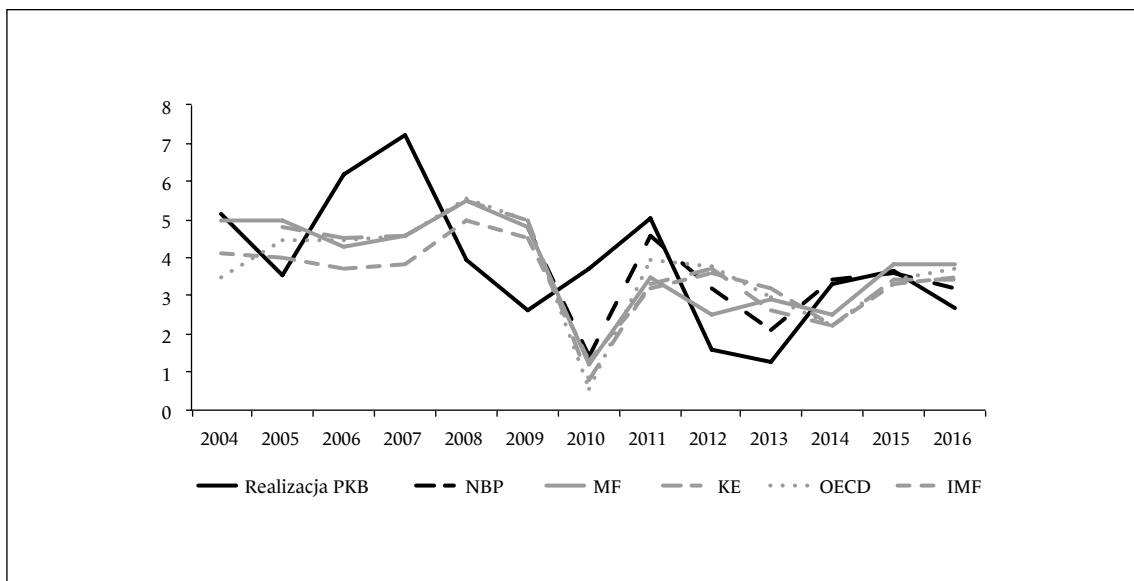
Realizacja oraz prognozy PKB, inflacji CPI i stopy bezrobocia dla NBP (2009–2016)



Źródło: opracowanie na podstawie danych NBP, GUS, Ministerstwa Finansów, Eurostatu, OECD i IMF.

Wykres 5

Przebieg prognoz wiosennych edycji PKB oraz realizacji



Źródło: opracowanie na podstawie danych NBP, GUS, Ministerstwa Finansów, Eurostatu, OECD i IMF.

A comparison of annual forecasts accuracy of Polish economic conditions compiled by national and international institutions

Abstract

A proper knowledge of future operating conditions is almost always considered a key factor of most economic projects. On a macroeconomic level, the quoted operating conditions are described as economic conditions. It is apparent that being able to correctly anticipate future conditions is of utmost importance. Therefore, it is no surprise that many economic entities focus on compiling economic forecasts. The forecasts of the different forecasting institutions almost never have the same results, which creates the need to evaluate their accuracy in order to establish which of them are the most reliable. This paper deals with a comparison of the accuracy of annual forecasts of the Polish economic conditions compiled by Narodowy Bank Polski (NBP), the Ministry of Finance of Poland (MoF), the European Commission, the Organisation of Economic Co-operation and Development (OECD) and the International Monetary Fund (IMF). The accuracy of the forecasts was examined using *ex post* methods. The selected time interval covers the years from 2004 to 2016, while taking it into account that some institutions take a different time span depending on data availability. The changes of economic conditions were represented by a selection of factors: changes in the average annual growth rate of Gross Domestic Product (GDP) in percentage, inflation described by the Consumer Price Index and in one specific case – the Harmonized Index of Consumer Prices, and the unemployment rate. Forecast accuracy was analysed using various *ex post* statistical methods, including mean error (ME), mean absolute error (MAE), root mean squared error (RMSE), Theil coefficient (Theil's U or U-Theil) and decomposition of mean squared error.

Keywords: GDP, inflation, unemployment, accuracy of short-term forecasts, Polish economic situation

