

POLSKA AKADEMIA NAUK — ODDZIAŁ W KRAKOWIE
KOMISJA NAUK EKONOMICZNYCH I STATYSTYKI
KRAKOWSKA AKADEMIA
IM. ANDRZEJA FRYCZA MODRZEWSKIEGO

FOLIA OECONOMICA CRACOVIENSIA

Vol. LII
2011

WYDAWNICTWO ODDZIAŁU POLSKIEJ AKADEMII NAUK
KRAKÓW

REDAKTOR

prof. dr hab. Andrzej Iwasiewicz

KOMITET REDAKCYJNY

prof. dr hab. Anna Czubała

prof. dr hab. Henryk Gurgul

prof. dr hab. Jacek Osiewalski — Sekretarz Naukowy Komisji Nauk Ekonomicznych i Statystyki Oddziału PAN w Krakowie i sekretarz naukowy Komitetu Redakcyjnego

RECENZENCI ARTYKUŁÓW ZAMIESZCZONYCH W VOL. LI ORAZ VOL. LII

Recenzenci zewnętrzni:

dr hab. Jerzy Marzec — Uniwersytet Ekonomiczny w Krakowie

dr hab. Anna Pajor — Uniwersytet Ekonomiczny w Krakowie

dr hab. Mateusz Pipień — Uniwersytet Ekonomiczny w Krakowie

Recenzenci wewnętrzni:

prof. dr hab. Henryk Gurgul — Akademia Górniczo-Hutnicza
im. Stanisława Staszica w Krakowie

prof. dr hab. Andrzej Iwasiewicz — Krakowska Akademia
im. Andrzeja Frycza Modrzewskiego

prof. dr hab. Jacek Osiewalski — Uniwersytet Ekonomiczny w Krakowie

Adres redakcji

31-018 Kraków, ul. św. Jana 28

Wydanie publikacji finansowane przez
Ministra Nauki i Szkolnictwa Wyższego,
Krakowską Akademię im. Frycza Modrzewskiego

© *Copyright by Autorzy, Polska Akademia Nauk Oddział w Krakowie
i Krakowska Akademia im. Andrzeja Frycza Modrzewskiego
Kraków 2011*

ISSN 0071-674X



Realizacja wydawnicza: PAN Warszawska Drukarnia Naukowa
Skład i druk: PAN Warszawska Drukarnia Naukowa
ul. Śniadeckich 8, 00-656 Warszawa
tel./fax 22 628-76-14
e-mail: wdnpan@wdnpan.pl
www.wdnpan.pl

SPIS TREŚCI

Paweł Baster, Katarzyna Pochtowska: <i>Sieci neuronowe i polichotomiczne modele zmiennych jakościowych w analizie ryzyka kredytowego</i>	5
Krzysztof Malczyk: <i>Analiza wpływu zmian kursu walutowego na inflację w Polsce za pomocą modelu VECM</i>	19
Kamil Makiela: <i>The contemporary standards of national accounts — applicability and limitations in economic growth and productivity studies</i>	49
Jacek Osiewalski, Krzysztof Osiewalski: <i>Modele hybrydowe MSV-MGARCH z trzema procesami ukrytymi w badaniu zmienności cen na różnych rynkach</i>	71
Anna Prusak, Piotr Stefanów: <i>Badania nad właściwościami operacyjnymi metody AHP</i>	87
 Review	
Lech Czarnecki: <i>Review of the monograph entitled "POLAND-SERBIA: The Challenges of the Scientific Cooperation</i>	105

SIECI NEURONOWE I POLICHOTOMICZNE MODELE ZMIENNYCH JAKOŚCIOWYCH W ANALIZIE RYZYKA KREDYTOWEGO

PAWEŁ BASTER

e-mail: pabaster@gmail.com

KATARZYNA POCZTOWSKA

e-mail: k.pocztowska@gmail.com

Praca została przedstawiona przez autorów 9 lipca 2010 r. na sesji *Współczesna ekonometria w badaniach empirycznych młodych ekonomistów*, którą zorganizowały: Komisja Nauk Ekonomicznych Polskiej Akademii Umiejętności, Komisja Nauk Ekonomicznych i Statystyki Oddziału PAN w Krakowie, Katedra Ekonometrii i Badań Opreacyjnych Uniwersytetu Ekonomicznego w Krakowie.

ABSTRACT

P. Baster, K. Pocztońska. *Neural networks and models for polychotomous ordered data in credit risk analysis*. Folia Oeconomica Cracoviensia

Management of credit risk, one of the main bank activities, is currently a very important issue. This paper contains comparison of two instruments used in prediction of probability that consumer fails to fully repay a loan in agreed time: artificial neural networks and models for polychotomous ordered data. For the empirical research each client has been assigned to one of four categories reflecting his/her delay in payments. Estimation and validation of methods was performed on a 3000-item sample containing information about each loan agreement and repayment history originating from one of Polish banks, covering years 2000–2001. The dataset was repeatedly divided into train and validation sets. Multi-layer architecture of artificial neural network with logistic activation function was proposed. Ordered logit and probit models were estimated within maximum likelihood framework. Several alternative specifications were proposed differing in independent variable set (including their products and squares). Bank income was chosen as the main criterion of fitness. Problem of optimal decision and defining appropriate loss function was formulated on the basis of statistical decision theory. Furthermore, properties of estimated models related to inference about probability of repayment and credit risk factors were presented.

KEY WORDS — SŁOWA KLUCZOWE

credit risk, ordered logit, ordered probit, artificial neural networks

ryzyko kredytowe, wielomianowe modele zmiennych jakościowych, sieci neuronowe

1. RYZYKO KREDYTOWE — WPROWADZENIE

Ryzyko stanowi nieodłączny element działalności banku. W literaturze definiuje się je jako zagrożenie wynikające z nietrafnych decyzji bądź jako możliwość wystąpienia efektu niezgodnego z oczekiwaniami. W przypadku działalności kredytowej jest to niebezpieczeństwo niedotrzymania warunków umowy przez kredytobiorcę, który nie spłaca całości lub części kwoty kredytu i odsetek (Krysiak 2006).

Z punktu widzenia banku niespłacanie kredytów powoduje zmniejszenie potencjalnych zysków i konieczność tworzenia rezerw celowych zabezpieczających wkłady deponentów, co w ostateczności może skutkować utratą płynności i wiarygodności, a nawet bankructwem. Dlatego też zarządzanie ryzykiem jest tak ważną sferą funkcjonowania banku. Istotą działań w tym obszarze jest identyfikacja i pomiar ryzyka. Najczęściej przyjmowaną miarą są prawdopodobieństwa spłaty kredytu przez konkretnych klientów wyznaczane za pomocą odpowiednich metod. Do najpopularniejszych należą: analiza dyskryminacyjna, modele zmiennych jakościowych, systemy eksperckie, sieci neuronowe, data mining. Banki stosują różne procedury, a ich skuteczność może stanowić o sile i konkurencyjności na rynku. Wykorzystywane narzędzia powinny pozwalać też na wyznaczenie determinant ryzyka i podjęcie decyzji kredytowej (Gruszczynski 2001). Najistotniejszym dla banku kryterium porównawczym jest jednak wielkość zysku finansowego wynikającego z zastosowania poszczególnych metod na etapie udzielania kredytów.

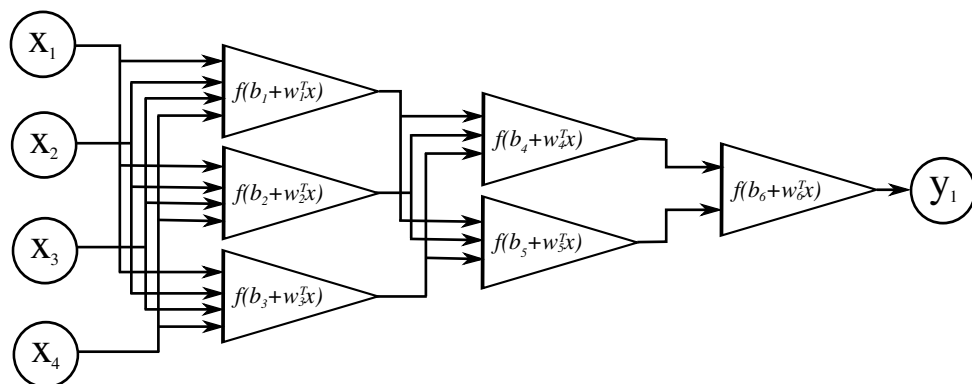
Przedmiotem rozważań w niniejszej pracy są dwa narzędzia stosowane w analizie ryzyka kredytowego: sieci neuronowe i polichotomiczne modele zmiennych jakościowych.

2. SIECI NEURONOWE

Przez *sztuczne sieci neuronowe* rozumiemy systemy przetwarzania danych symulujące uczenie się struktury mózgu (PWN 1996). Inna definicja głosi, że sieć neuronowa to uproszczony model mózgu, składający się z pewnej liczby elementów nazywanych sztucznymi neuronami, które przetwarzają informacje. *Sztuczne neurony* to uproszczone modele neuronów biologicznych, których parametry (nazywane wagami) decydują o własnościach sieci i podlegają optymalizacji w trakcie procesu nazywanego uczeniem (Tadeusiewicz 1993). Z matematycznego punktu widzenia, sieci neuronowe to modele regresyjne oparte na pewnej klasie sparametryzowanych funkcji nieliniowych, których parametry szacuje się zwykle metodami nieparametrycznymi.

W opisywanym badaniu wykorzystana została typowa architektura sieci wielowarstwowej jednokierunkowej z logistyczną funkcją aktywacji. Ta klasa sieci

jest w stanie odwzorować każdą funkcję ciągłą określoną na ograniczonym zbiorze z dowolną dokładnością (Cybenko 1989). Na rysunku 1 przedstawiony został schemat sieci tego typu. W sztucznym neuronie oblicza się ważoną sumę wejść, która jest następnie normowana przez funkcję logistyczną. Neurony ułożone są w warstwy (zwykle nie więcej niż trzy), które kolejno i równoległe przetwarzają zmienne wejściowe. Wartości generowane przez ostatnią warstwę sieci interpretowane są jako wartości wyjściowe — odpowiedź modelu.



Źródło: opracowanie własne.

Ryc. 1. Schemat sieci wielowarstwowej jednokierunkowej

Uczenie sieci to optymalizacja wag w sensie pewnej funkcji celu (najczęściej jest to minimalizacja błędu średniokwadratowego). Ocena efektywności sieci zwykle odbywa się przez testowanie jej działania na danych nie biorących udziału w optymalizacji — tzw. zbiorze testowym, stanowiącym zazwyczaj 10–20% całego zbioru obserwacji, podczas gdy dane biorące udział w uczeniu nazywane są zbiorem uczącym. Dopasowanie sieci do danych uczących rośnie wraz ze wzrostem liczby neuronów, lecz po początkowym wzroście zwykle spada dla danych testowych, gdyż sieć zaburzona przez przypadkowy szum w danych uczących będzie gorzej generalizować wyuczone zależności na nieobserwowane przypadki. Dla poprawy zdolności aproksymacyjnych sieci do funkcji celu dodaje się czasem wyraz penalizujący modele z wyższymi wartościami bezwzględnych wag. Do konstrukcji sieci, estymacji modelu i generowania prognoz opartych o sztuczne sieci neuronowe wykorzystuje się także czasem metody wnioskowania bayesowskiego (Lee 1999).

3. POLICHOTOMICZNE MODELE ZMIENNYCH JAKOŚCIOWYCH DLA KATEGORII UPORZĄDKOWANYCH

Polichotomiczne modele zmiennych jakościowych dla kategorii uporządkowanych (ang. multinomial response, polychotomous, multiple-choice models) to klasa modeli ekonometrycznych, w których zmienna endogeniczna ma dyskretny rozkład prawdopodobieństwa i przyjmuje więcej niż dwie wartości mierzone na skali porządkowej.

Ogromny wkład w rozwój tych modeli wnieśli D.L. McFadden i J. Heckman, którzy w 2000 r. za rozwój metodologiczny i zastosowanie modeli dyskretnego wyboru w praktyce otrzymali nagrodę Nobla w dziedzinie ekonomii. Konstrukcja modelu jest następująca (McKelvey, Zavoina 1975).

Niech y_1, \dots, y_T będzie ciągiem niezależnych zmiennych losowych o dyskretnym rozkładzie. Zakładamy, że przyjmują one z ustalonym prawdopodobieństwem p_{ij} wartości j , gdzie $j = 1, 2, \dots, M$ dla każdego t .

Zależność między $\{p_{ij}\}_{j=1}^M$ a wektorem zmiennych objaśniających $X = (x_1, \dots, x_K)$ wprowadza się przy pomocy nieobserwowalnej ciągłej zmiennej z_t , która według proponowanej w literaturze interpretacji reprezentuje stochastyczną funkcję użyteczności decydenta wybierającego spośród możliwych kategorii zmiennej y_t .

Wartości z_t determinują obserwowaną kategorię zmiennej y_t w następujący sposób:

$$\begin{cases} z_t = X_t B + e_t, \\ y_t = j \Leftrightarrow \alpha_{j-1} < z_t \leq \alpha_j, \end{cases}$$

dla $j = 1, \dots, M$ i $t = 1, \dots, T$.

W powyższym modelu zmienne losowe e_t są niezależne o znanym jednakowym rozkładzie (z zerową wartością oczekiwaną i znaną skończoną wariancją) i odpowiadają za wszelkie losowe zakłócenia. Jeśli jest to rozkład logistyczny to model nazywamy *logitowym*, a jeśli normalny to model określa się mianem *probitowego*.

(B, A) stanowi wektor parametrów podlegających estymacji za pomocą metody MNW. Składowe wektora $A = (\alpha_0, \alpha_1, \dots, \alpha_T)$ noszą nazwę *punktów ucięcia*. W celu zapewnienia identyfikowalności uwzględniamy w B wyraz wolny β_0 i ustalamy: $\alpha_0 = -\infty, \alpha_1 = 0, \alpha_M = +\infty, .$

Dla obserwacji charakteryzującej się wektorem zmiennych niezależnych X_T prawdopodobieństwo przyjęcia przez zmienną zależną y_t kategorii j (dla $j = 1, \dots, M$) można zapisać jako:

$$p_{tj} = \Pr(Y_t = j) = \Pr(\alpha_{j-1} < z_t < \alpha_j) = F(\alpha_j - X_t B) - F(\alpha_{j-1} - X_t B),$$

gdzie $A(\cdot)$ to dystrybuanta rozkładu zmiennych e_t .

Oznaczmy jako $W = (w_1, \dots, w_n)$ wektor wyjściowych zmiennych objaśniających reprezentujących określone interpretowalne cechy. Wtedy X_t może być wektorem dowolnych funkcji elementów W_t pod warunkiem, że są liniowe względem parametrów B . Jeżeli część regresyjna modelu ma postać $W_t B$ to taką specyfikację nazywamy *modelem I rzędu*. Jeżeli X_t jest wielomianem drugiego stopnia zmiennych z W_t to model jest *II rzędu*. Przyczyną wprowadzenia specyfikacji II rzędu jest jej zdolność do lepszej aproksymacji dowolnej funkcji zmiennych W_t i parametrów B , występującej potencjalnie w miejscu wyrażenia $X_t B$. Modele II rzędu wykazują tym samym lepsze własności w zakresie modelowania wpływu zmiennych wyjściowych na prawdopodobieństwa p_{ij} .

4. CHARAKTERYSTYKA DANYCH

Zbiór obserwacji składał się z 3000 rachunków kredytowych otwartych w okresie od 01.01.2000 do 30.09.2001. Dane pochodziły z polskiego banku komercyjnego i były już wcześniej wykorzystywane w badaniach empirycznych (Marzec 2008).

Zmienna objaśniana przyjmuje cztery wartości reprezentujące kategorię kredytu w zależności od przedawnienia w spłacie. Do podziału kredytu na takie grupy banki były zobowiązane treścią obowiązującej w tym czasie uchwały (Uchwała nr 8/1999 Komisji Nadzoru Bankowego z 22 grudnia 1999 r.), która obligowała również do tworzenia rezerwy zabezpieczającej depozyty liczonej jako procent wartości kredytów z danej kategorii. Tabela 1 zawiera informacje dotyczące konstrukcji zmiennej objaśnianej, okres opóźnienia w spłacie kredytu, a także procent wartości kredytu jaki bank musiał odłożyć w postaci rezerw.

Tabela 1

Konstrukcja zmiennej objaśnianej

j	Kategoria należności	Wysokość rezerw	Opóźnienie w spłacie
1	normalne	–	poniżej 1 m-ca
2	poniżej standardu	20%	od 1 do 3 m-cy
3	wątpliwe	50%	od 3 do 6 m-cy
4	stracone	100%	powyżej 6 m-cy

Źródło: opracowanie własne.

Zbiór charakterystyk obserwacji składał się ze zmiennych opisujących:
 — kredytobiorcę: płeć (1-mężczyzna, 0-kobieta), wiek w latach, źródło dochodu — reprezentowany przez trzy pomocnicze zmienne binarne: dochód1

(1-renta, emerytura), dochód2 (1-własna działalność gospodarcza, umowa zlecenie i o dzieło), dochód3 (1-inne źródło np. stypendium), referencyjna wartość to umowa o pracę;

- dotychczasowe kontakty klienta z bankiem: posiadanie kart kredytowych (1-posiada, 0-nie posiada) i rachunku ROR (1-posiada, 0-nie posiada), a także wielkość kwartalnych wpływów na ROR w tys. zł;
- kredyt: kwota w tys. zł, okres w latach, typ (1-konsumpcyjny, 0-hipoteczny), waluta (1-obca, 0-krajowa), sposób przyznania (1-przez pośrednika, 0-przez bank).

W tabeli 2 przedstawiono rozkład poszczególnych zmiennych w próbie. Określono też wartości wszystkich zmiennych (na poziomie mediany) dla dwóch wyróżnionych klientów — typowego klienta pozyskanego przez bank lub przez pośrednika.

Tabela 2

Podstawowe ilościowe informacje o rachunkach kredytowych i ich właścicielach

Zmienna	Struktura	Typowy klient	
		Pośrednik=0	Pośrednik=1
<i>Płeć</i>	54%	1	1
<i>Wiek</i>	39 lat	39 lat	39 lat
<i>Wpływy</i>	2,4 tys. zł	2,4 tys. zł	2,4 tys. zł
<i>ROR</i>	57%	1	1
<i>Karty</i>	33%	0	0
<i>Pośrednik</i>	38%	0	1
<i>Typ</i>	94%	1	1
<i>Okres</i>	1,5 roku	1,5 roku	1,5 roku
<i>Kwota</i>	5 tys. zł	5 tys. zł	5 tys. zł
<i>Waluta</i>	3%	0	0
<i>Dochód 1</i>	17%	0	0
<i>Dochód 2</i>	7%	0	0
<i>Dochód 3</i>	2%	0	0

Źródło: opracowanie własne.

5. MODEL FUNKCJONOWANIA BANKU W KONTEKŚCIE PODEJMOWANIA DECYZJI KREDYTOWEJ

W celu szacowania korzyści finansowych z zastosowanych narzędzi należy określić zysk banku z pojedynczego klienta w zależności od decyzji kredytowej i kategorii kredytu. W literaturze (Osiewalski 2007, Marzec 2008) rozważa się to zagadnienie na gruncie statystycznej teorii decyzji. Oznaczmy przez m marżę banku, czyli różnicę w oprocentowaniu kredytów i depozytów. Bank podpisując umowę z solidnym kredytobiorcą może się spodziewać zysku w wysokości m (dla jednostkowej kwoty kredytu). Natomiast odmawiając ponosi koszty utraconych możliwości równe $-m$. Decyzja o przyznaniu kredytu, który nie będzie spłacany wiąże się z utratą pewnej części kapitału i odsetek. Należy więc przyjąć odpowiednie wagi dla poszczególnych kategorii należności. Zasadne wydaje się odwołanie do wielkości rezerw, które bank musi tworzyć w celu zabezpieczenia środków pozyskanych od deponentów. Zakładamy więc, że kredyty poniżej standardu nie będą spłacone w 20% (strata banku to $0,2(1 + m)$), kredyty zagrożone w 50% (strata to $0,5(1 + m)$), a stracone nie zostaną spłacone w całości (strata to $(1 + m)$). Powoduje to również, że bank zyskuje $0,8m$ lub $0,5m$ z kredytów odpowiednio drugiej i trzeciej grupy. W tabeli 3 zaprezentowano jak w tej sytuacji będzie się kształtował zysk ekonomiczny banku.

Tabela 3

Tabela wypłat zysku banku w zależności od kategorii kredytu i decyzji kredytowej

Decyzja	Kategoria kredytu			
	$j = 1$ normalny	$j = 2$ poniżej standardu	$j = 3$ wątpliwy	$j = 4$ stracony
Udzielić	m	$0,6m - 0,2$	$-0,5$	$-(1 + m)$
Odmówić	$-m$	0	0	0
Prawdopodobieństwo	p_{t1}	p_{t2}	p_{t3}	p_{t4}

Źródło: opracowanie własne.

Na podstawie oszacowanych dla konkretnego klienta prawdopodobieństw p_{t1} , p_{t2} , p_{t3} , p_{t4} , i funkcji wypłaty z tabeli 3 określimy optymalną regułę wskazującą jaką decyzja kredytowa jest dla banku opłacalna. Zakładamy, że przyznanie kredytu powinno nastąpić, kiedy przy ustalonej marży m oczekiwany zysk z udzielenia kredytu będzie większy od kosztów odmowy, co zachodzi, gdy spełniona jest nierówność:

$$mp_{t1} + (0,6m - 0,2)p_{t2} - 0,5p_{t3} - (1 + m)p_{t4} \geq -mp_{t1}$$

6. WYNIKI EMPIRYCZNE

Dla potrzeb porównania modeli wielokrotnie dzielono zbiór obserwacji na zbiór uczący (85%) i testowy (15%) i uzyskane na nich wyniki dla każdego z modeli uśredniano. Za miarę jakości modelu przyjęto średni zysk oraz medianę zysku z kredytu o wartości jednostkowej (w sensie funkcji wypłaty zdefiniowanej w tabeli 3), przy ustalonej stopie procentowej $m = 10\%$. Miara ta przyjmuje wartości pomiędzy -1 a $0,1$. Do porównania wybrano model logitowy i probitowy dla kategorii uporządkowanych oraz sztuczną sieć neuronową z wyrazem penali-zującym modele z wyższą sumą kwadratów wag. Estymację przeprowadzono za pomocą pakietu R. Jako zmienne objaśniane przyjęto znormalizowane wartości wszystkich dostępnych cech oraz dwie nieliniowe kombinacje tych zmiennych: *kwota/okres* (odpowiadająca w przybliżeniu wysokości raty kredytowej) oraz *kwota/okres/wpływy* (lub 1 jeśli wpływy wynosiły zero). Wyniki porównania po ponad dwustu iteracjach zaprezentowano w tabeli 4.

Tabela 4

Zysk z kredytu o wartości jednostkowej

	logit	probit	sztuczna sieć neuronowa
średnia	0,0267	0,0266	0,0255
mediana	0,0267	0,0266	0,026

Źródło: opracowanie własne.

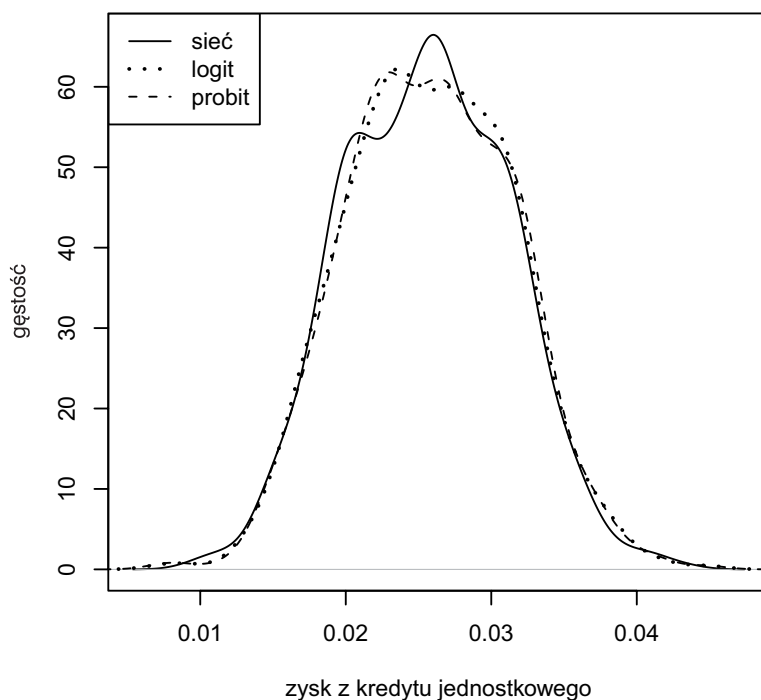
Wyniki porównano testem Wilcoxon dla obserwacji zależnych otrzymując rezultaty przedstawione w tabeli 5 oraz wyrysowano ich rozkład za pomocą ją-drowej estymacji gęstości (rysunek 2).

Tabela 5

P-value dla jednostronnych testów Wilcoxon

$H_0: Me(A) = Me(B)$ vs $H_1: Me(A) \neq Me(B)$		
<i>A</i>	<i>B</i>	<i>P-value</i>
logit	probit	0,0013
logit	sz. sieć neuronowa	$< 10^{-6}$
probit	sz. sieć neuronowa	$< 10^{-6}$

Źródło: opracowanie własne.



Źródło: opracowanie własne.

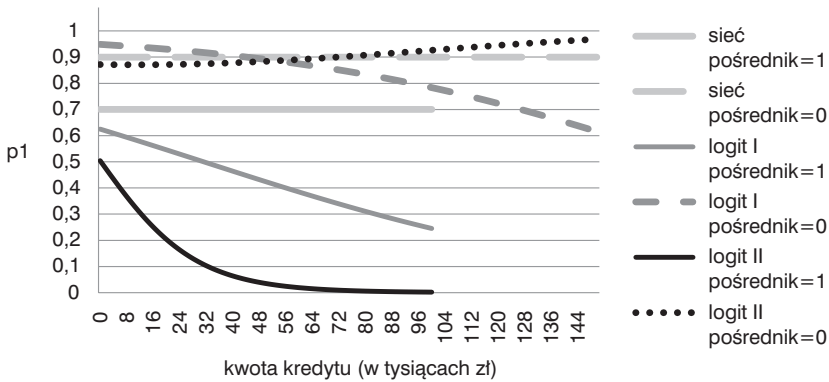
Ryc. 2. Jądrowa estymacja gęstości zysku

Wyniki te pokazują, że mimo minimalnych różnic w wartościach zysku, polichotomiczne modele zmiennych jakościowych dla kategorii uporządkowanych radziły sobie istotnie lepiej w ocenie ryzyka kredytowego niż sztuczne sieci neuronowe. Ponadto zaobserwowano, iż:

- jeżeli sztuczna sieć neuronowa korzysta ze wszystkich zmiennych objaśniających wówczas największy zysk generuje sieć składająca się tylko z jednego neuronu (co do postaci funkcyjnej równoważna modelowi logitowemu). Dodatkowe warstwy neuronów, będące źródłem elastyczności sztucznych sieci neuronowych i zwykle stanowiące o ich przewadze, w tym przypadku jedynie powodują pogarszanie wyników z powodu przeuczenia. Ograniczenie zbioru zmiennych objaśniających przez usuwanie zmiennych nieistotnych poprawia wyniki. Dla zredukowanego zbioru zmiennych wejściowych testy pokazują przewagę sieci z jedną warstwą ukrytą zawierającą 1–2 neuronów.
- Po ponad dwustu iteracjach nie można odrzucić hipotezy, że rozszerzenie modelu logitowego dwumianowego do modelu polichotomicznego dla kategorii uporządkowanych nie zmienia istotnie zysku (p -value dla testu Wilcoxon: 0,31). Test zdecydowanie odrzuca jednak analogiczną hipotezę dla modelu probitowego (p -value: $2 \cdot 10^{-22}$).

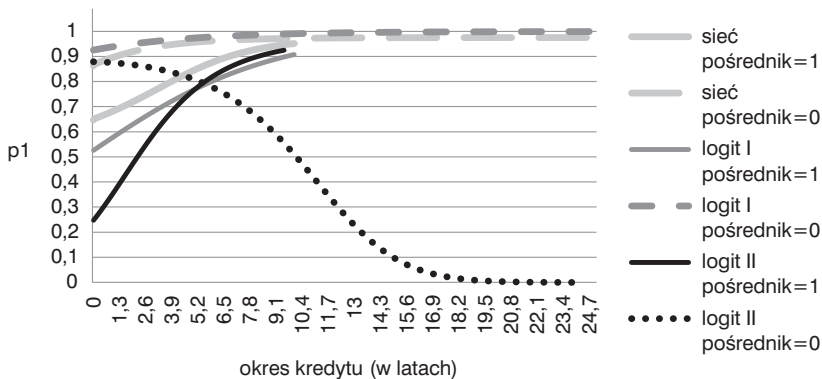
— Niewielka co do wartości zysku lecz istotna statystycznie przewaga modelu logitowego nad siecią neuronową (nawet o tylko jednym neuronie) sugeruje, że — wobec identycznych postaci funkcyjnych — różnica wynika z metod estymacji. Lepsze wyniki metod parametrycznych względem typowego dla sieci poszukiwania minimum błędu średniokwadratowego lub maksimum funkcji wypłaty sugerują zwrócenie się w kierunku prób wykorzystania metod parametrycznych do uczenia sieci neuronowych jako uogólnienia modelu logitowego. Stosowanie do tego celu metod wnioskowania bayesowskiego proponuje Herbert Lee (Lee 1999).

Na rysunkach 3–6 pokazano zależność między oszacowanym prawdopodobieństwem spłaty a wybranymi zmiennymi dla sieci neuronowych oraz modelu logitowego (I i II rzędu) oraz dla różnych wartości zmiennej *pośrednik*. Wartości pozostałych zmiennych objaśniających ustalone zostały na poziomie mediany ich wartości obserwowanych w próbie.



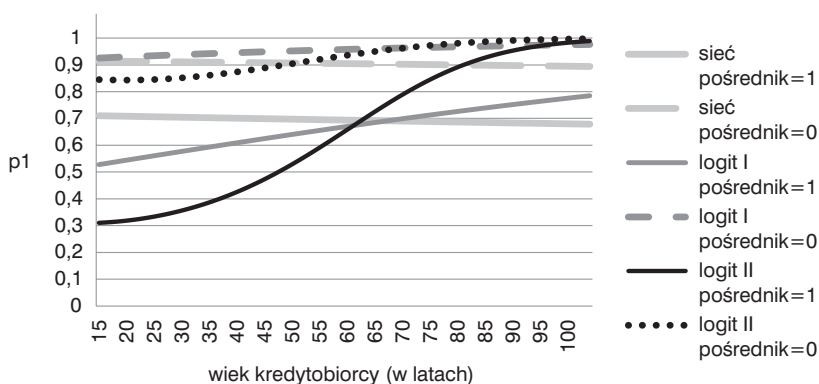
Źródło: opracowanie własne.

Ryc. 3. Ocena prawdopodobieństwa spłaty w zależności od kwoty kredytu



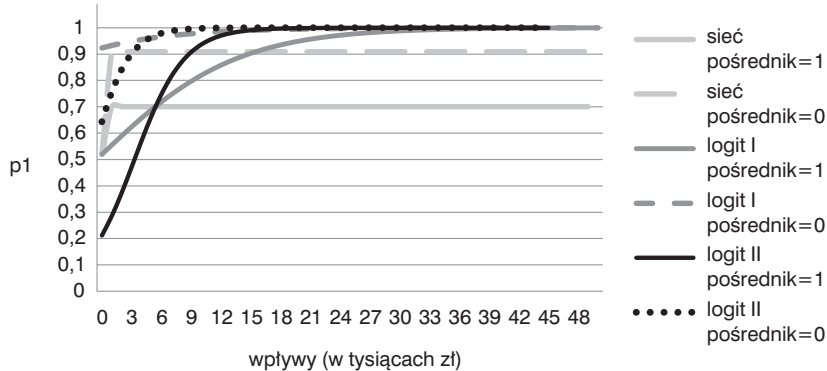
Źródło: opracowanie własne.

Ryc. 4. Ocena prawdopodobieństwa spłaty w zależności od okresu spłaty



Źródło: opracowanie własne.

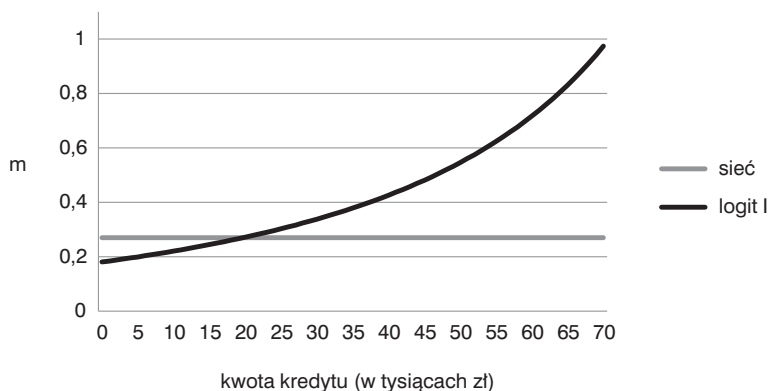
Ryc. 5. Ocena prawdopodobieństwa spłaty w zależności od wieku kredytobiorcy



Źródło: opracowanie własne.

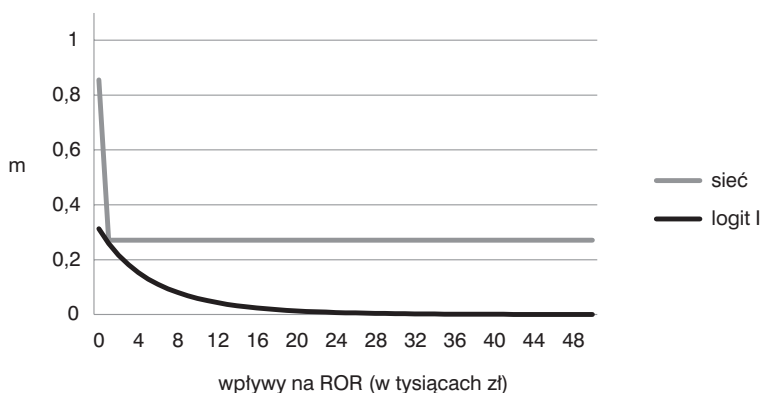
Ryc. 6. Ocena prawdopodobieństwa spłaty w zależności od kwartalnych wpływów na ROR klienta

Niektóre z tych zależności mogą służyć bankowi do zaproponowania klientom nowych warunków umowy, gdy ich podania zostały sklasyfikowane jako zbyt ryzykowne. W szczególności model może odpowiedzieć na pytanie o jaką kwotę lub o jaki okres spłaty może wnioskować klient, aby klasyfikator uznał jego szanse na terminową spłatę za wystarczające. Proponowane podejście może zostać wykorzystane w procedurze przyznawania kredytu z indywidualnym ustalaniem marży równoważącej oszacowane ryzyko umowy. Rysunek 7 i 8 pokazuje wysokości minimalnej marży zapewniającej bankowi zysk w długim okresie w zależności od kwoty kredytu oraz kwartalnych wpływów na rachunek ROR dla klienta, który stara się o kredyt przez pośrednika (pozostałe jego cechy przyjmują mediany wartości obserwowanych).



Źródło: opracowanie własne.

Ryc. 7. Minimalna marża w zależności od kwoty kredytu (w tys. zł)
(dla wniosków składanych przez pośrednika)



Źródło: opracowanie własne.

Ryc. 8. Minimalna marża w zależności od kwartalnych wpływów na ROR klienta
(dla wniosków składanych przez pośrednika)

7. PODSUMOWANIE

Formalne porównanie sieci neuronowych i polichotomicznych modeli zmiennych jakościowych pokazało, że — dla danych obserwacji — istotnie lepsze wyniki pod względem uzyskiwanego zysku dają modele zmiennych jakościowych. W pracy zaproponowano metodę wyznaczania zysku banku w zależności od decyzji kredytowej i kategorii analizowanego kredytu. Zaprezentowano też regułę, która przy danych prawdopodobieństwach spłaty kredytu (szacowanych

przez konkurencyjne modele) pozwala na podjęcie decyzji kredytowej i określenie gwarantujących przyznanie kredytu wartości zmiennych charakteryzujących konkretnego klienta bądź indywidualne wyznaczenie marży.

BIBLIOGRAFIA

- Cybenko G. 1989. *Approximations by superpositions of sigmoidal function*, Mathematics of Control, Signals and Systems 2, 303–314.
- Gruszczyński M. 2001. *Modele i prognozy zmiennych jakościowych w finansach i bankowości*, Monografie i Opracowania SGH 6, Warszawa.
- Krysiak Z. 2006. *Ryzyko kredytowe a wartość firmy. Pomiar i modelowanie*, Oficyna Wydawnicza, Kraków.
- Lee H. 1999. *Model Selection and Model Averaging for Neural Networks*, PhD thesis, Carnegie Mellon University, Pittsburgh, USA. <http://users.soe.ucsc.edu/~herbie/thesis.pdf>.
- Marzec J. 2008. *Bayesowskie modele zmiennych jakościowych i ograniczonych w badaniach niespłacalności kredytów*, Wydawnictwo UE w Krakowie, Kraków.
- Osiewalski J. 2007. *Bayesowska statystyka i teoria decyzji w analizie ryzyka kredytu detalicznego* [w:] *Finansowe warunkowania decyzji ekonomicznych*, Krakowska Szkoła Wyższa im. Andrzeja Frycza Modrzewskiego, Kraków.
- Petrozolin-Skowrońska B. 1996. *Neronowe sieci*. [w:] *Nowa Encyklopedia Powszechna PWN*, t. 4, Wydawnictwo Naukowe PWN, Warszawa, 439.
- McKelvey R.D., Zavoina W. 1975. *A Statistical Model for the Analysis of Ordinary Level Dependent Variables*, Journal of Mathematical Sociology 4, 103–120.
- Tadeusiewicz R. 1993. *Sieci Neuronowe*, Akademicka Oficyna Wydawnicza, Kraków.

ANALIZA WPŁYWU ZMIAN KURSU WALUTOWEGO NA INFLACJĘ W POLSCE ZA POMOCĄ MODELU VECM

KRZYSZTOF MALCZYK

e-mail: krzysztof.malczyk@onet.eu

Praca została przedstawiona przez autora 9 lipca 2010 r. na sesji *Współczesna ekonometria w badaniach empirycznych młodych ekonomistów*, którą zorganizowały: Komisja Nauk Ekonomicznych Polskiej Akademii Umiejętności, Komisja Nauk Ekonomicznych i Statystyki Oddziału PAN w Krakowie, Katedra Ekonometrii i Badań Operacyjnych Uniwersytetu Ekonomicznego w Krakowie.

ABSTRACT

K. Malczyk. *Exchange Rate Pass-Through to Inflation in Poland: VECM Analysis*. Folia Oeconomica Cracoviensia

This paper applies Vector Error Correction (VEC) methodology to investigate effects of exchange rate shocks to inflation and price setting processes in Poland, since National Bank of Poland (NBP) have adopted inflation targeting policy and floating exchange rate regime in 1998. The size and the speed of the pass-through at different stage of pricing chain (import prices, producer prices, consumer prices) is measured by impulse response function (IRF). In addition, the relative importance of the exchange rate shocks using forecast errors variance decompositions from the estimated VECM is investigated. The results are interpreted in the context of NBP's conduct of monetary policy, in the prospect of euro area accession.

KEY WORDS — SŁOWA KLUCZOWE

exchange rate, inflation, monetary policy, Vector Error Correction Model

kurs walutowy, inflacja, polityka pieniężna, wektorowy model korekty błędów

1. WPROWADZENIE

Celem niniejszego artykułu jest przedstawienie wyników analizy wpływu zmian kursu walutowego oraz procesów je kształtujących na ceny i inflację w Polsce, od momentu wprowadzenia przez Narodowy Bank Polski strategii bezpośredniego celu inflacyjnego, to jest od roku 1998.

Na przestrzeni ostatnich dwóch dekad, zjawisko przenoszenia zmian kursu na ceny (ang. *exchange rate pass-through*) w gospodarkach krajów wysoko uprzemysłowionych stało się przedmiotem intensywnych analiz zarówno o charakterze teoretycznym, jak i empirycznym (m.in. McCarthy [1999], Hahn [2003], Bussiere [2007], Landau i Skudelny [2009]). Natomiast w przypadku polskiej gospodarki można stwierdzić, iż właściwe i kompletne poznanie mechanizmu przenoszenia zmian kursowych na ceny nie zostało jeszcze dokonane. Aspekty analizowanego zjawiska pojawiają się w pracach analitycznych Narodowego Banku Polskiego (m. in. Kokoszcyński i in. [2002], Cholewiński [2008], Grabek i in. [2008]). Teoretyczne zależności między kursem a cenami szeroko przedstawia Misztal [2008], [2009].

Zmiany polityki pieniężnej, perspektywa członkostwa Polski w trzecim etapie Unii Gospodarczej i Walutowej, a także efekty kryzysu gospodarczego ostatnich lat, stanowią przesłankę do ciągłej aktualizacji prowadzonych badań, zarówno poprzez wykorzystywanie możliwie najświeższych danych, jak i udoskonalanie wykorzystywanych metod ekonometrycznych. Niniejszy artykuł podejmuje próbę szerszej ilustracji efektu *pass-through* w rodzimej gospodarce.

W rozdziale drugim przedstawione zostaną teoretyczne założenia dotyczące istoty oraz determinant niepełnego przenoszenia zmian kursowych na ceny, ze szczególnym naciskiem na relacje łączące badane zjawisko z polityką pieniężną. Rozdział trzeci poświęcony zostanie teoretycznemu zaprezentowaniu modeli SVAR i SVECM oraz formalnemu przedstawieniu wskaźnika służącego do pomiaru efektu przenoszenia zmian kursowych na ceny. W rozdziale czwartym przedstawione zostaną wyniki wstępnych analiz rzeczywistych danych i zaprezentowane konkluzje analiz ich statystycznych własności. Rozdział piąty przedstawiać będzie ocenę siły i szybkości przenoszenia się zmian kursu walutowego na ceny i inflację, do czego wykorzystane zostaną funkcja reakcji na bodźce i dekompozycja wariancji błędu prognozy.

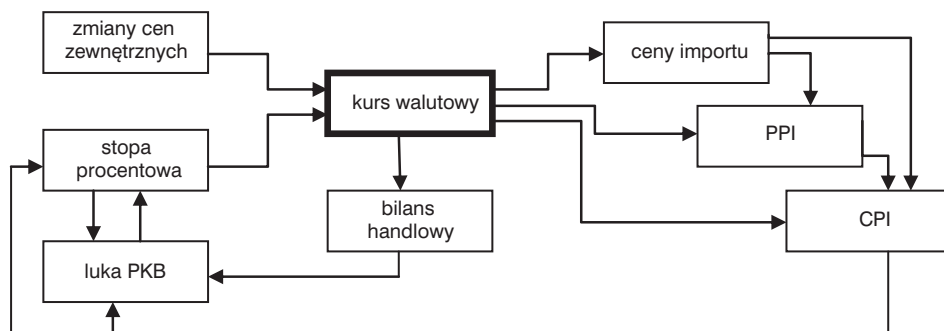
2. ISTOTA I PRZYCZYNY NIEPEŁNEGO PRZENOSZENIA ZMIAN KURSU WALUTOWEGO NA CENY

W literaturze ekonomicznej spotkać można dwa ujęcia efektu przenoszenia zmian kursu na ceny wewnętrzne. Ghosh i Rajan [2009] definiują to pojęcie jako stopień wrażliwości cen dóbr importowanych na jednoprocetową zmianę kursu walutowego wyrażonego w walucie importera. Jednak, jak zauważają Hall i Taylor [2007], w przypadku tzw. małych gospodarek otwartych z kursem walutowym związany jest ściśle również krajowy ogólny poziom cen. Towary importowane bezpośrednio konkurują z produktami krajowymi, natomiast dobra eksportowe sprzedawane są po cenach światowych. Stopień wpływu zmian kursu na ogólny poziom cen w kraju jest więc o wiele szerszy (por. Ihrig i in. [2007]). Dodatkowo

zmiany kursu poprzez wpływ na konkurencyjność eksportu determinują zmiany produkcji, co z kolei wpływa na ceny krajowe. Kanał kursowy jest też ogniwem transmisji szoków zewnętrznych.

W związku z powyższym można spotkać ujęcie szersze opisywanego zagadnienia, w którym to efekt *pass-through* rozumiany jest jako łączna reakcja zmian różnych czynników takich jak: szoki podażowe (zmiany cen zewnętrznych) i szoki popytowe (zmiany luki popytowej) na wskaźnik cen towarów i usług konsumpcyjnych (CPI) (Kokoszyczyński i in. [2002]). Transmisja w tym przypadku przebiega nie tylko bezpośrednio (poprzez ceny dóbr importowanych), ale też kanałem pośrednim poprzez wpływ cen importu zaopatrzeniowego wywołujący zmiany cen produkcji sprzedanej przemysłu (PPI). Dodatkowo wahania cen wewnętrznych wpływając na popyt krajowy zmieniają wielkość luki popytowej, co z kolei determinuje dalsze zmiany wskaźnika CPI.

Kurs walutowy jest również jednym z kanałów transmisji impulsów polityki pieniężnej. Spadek wewnętrznej stopy procentowej prowadzi do obniżenia się atrakcyjności inwestycyjnej kraju. Implikuje to odpływ kapitału krótkoterminowego pociągający za sobą deprecjację rodzimej waluty. Spadek kursu korzystnie wpływa na konkurencyjność eksportu, natomiast pogarsza warunki importu, co sprzyja poprawie bilansu handlowego. Rośnie krajowa produkcja i umacnia się wzrost gospodarczy. Odwrotny wpływ na gospodarkę będzie miał wzrost stopy procentowej (szerzej Kaźmierczak [2008] s. 124–125). Wszystkie wymienione zależności obrazuje schemat 1.



Źródło: opracowanie własne.

Schemat 1. Mechanizm wpływu zmian kursu walutowego na ceny

Szukając przyczyn niepełnego przekładania się zmian kursu na ceny, w jednej z pierwszych prac empirycznych analizujących opisywane zagadnienie, Dornbusch [1987] rozważa takie jego determinanty jak: udział importu w rynku kraju przeznaczenia, stopień homogeniczności i substytucyjności dóbr oraz skalę

rywalizacji eksporterów. Jeśli udział towarów zagranicznych w danym rynku jest wysoki, eksporterzy będą skłonni absorbować negatywne zmiany kursu walutowego w celu utrzymania konkurencyjności swoich produktów i zachowania udziałów w rynku. Podobnie będzie w sytuacji, gdy oferowane dobra konsumenci będą w stanie zastąpić produktami krajowymi. Zatem im rynek jest bardziej konkurencyjny, tym efekt *pass-through* jest mniejszy. Dodatkowo autor wymienia koszty transportu i dystrybucji zmniejszające skalę spadku cen w wyniku aprecjacji waluty importera.

Nie bez znaczenia są też tzw. koszty menu, czyli obciążenia jakie ponoszą sprzedawcy zmuszeni do częstych zmian cen. Dlatego wahania kursu uznawane za krótkotrwałe lub niewielkie mogą pozostać bez odzwierciedlenia w cenach sprzedaży dóbr. Niemniej jednak Pollard i Coughlin [2004] dowodzą, że gdy o skali przenoszenia decydują jedynie koszty menu, w długim okresie możemy mówić o pełnym efekcie *pass-through*.

Skala przenoszenia zmian kursu na ceny determinowana może być przez celowe zachowania producentów, w szczególności stosowaną przez nich strategię cenową i wybór waluty fakturowania sprzedaży (por. Fukuda i Ono [2006], Campa i Goldberg [2002]). Strategia ustalania cen w walucie lokalnej (*local currency pricing, LCP*) będzie przyczyniać się do nominalnej sztywności cen na rynku docelowym, czyli niskiego stopnia przenoszenia zmian kursowych (Devereux i Engel [2003]). Misztal [2008] dodaje, że konsekwencją wyboru takiej strategii będą dostosowania jedynie w wolumenie importu. Do odmiennych rezultatów prowadzi strategia ustalania cen w walucie kraju eksportera (*producer currency pricing, PCP*), w którym to przypadku mamy do czynienia z bezpośrednim przenoszeniem zmian kursowych na ceny importu.

Anderton [2003] analizując różnice w efekcie *pass-through* między Stanami Zjednoczonymi a krajami Unii Europejskiej także opiera się na czynnikach mikroekonomicznych. Autor konkluduje, że im wyższa elastyczność cenowa popytu na towary importowane tym mniejsza skłonność producentów do przenoszenia zmian kursu na ceny. Skłonność ta natomiast rośnie wraz ze wzrostem elastyczności cenowej podaży dóbr importowanych. Istotna wydaje się też inna konkluzja autora, mówiąca o względnie niskim stopniu przekładania się zmian kursowych w przypadku krajów członkowskich UE ubiegających się o członkostwo w strefie euro. Można zatem wysnuć wniosek, że perspektywa wejścia Polski do strefy euro może stabilizować ceny handlu międzynarodowego, gdyż wahania kursowe mogą być postrzegane jako tymczasowe i przemijające.

Wpływ zmian kursowych na ceny może również zależeć od prowadzonej przez władze monetarne danego kraju polityki pieniężnej, w tym polityki kursu walutowego. Taylor [2000] twierdzi, że skala przenoszenia zmian kursowych jest niższa w krajach o niższej inflacji, gdzie polityka monetarna ma bardziej stabilny charakter. W warunkach uporczywej inflacji firmy nie będą postrzegać wahań kursowych jako krótkotrwałych, stąd większa ich skłonność do dostosowań ce-

nowych. Założenie to znalazło potwierdzenie w analizach empirycznych (por. Choudhri i Hakura [2006], Ca' Zorzi, Hahn i Sanchez [2007]).

Warto zaznaczyć, że kanał kursowy polityki pieniężnej wydaje się mieć istotne znaczenie w przypadku możliwego przystąpienia Polski do strefy euro. Od stopnia podobieństwa pomiędzy mechanizmami transmisji polityki pieniężnej (w tym reakcji na szoki walutowe) zależeć będzie adekwatność i skuteczność wspólnej polityki monetarnej dla polskiej gospodarki. Wymienione wyżej prace dostarczają wniosków, że stopień przenoszenia zmian kursowych na ceny jest wyższy w gospodarkach wschodzących niż w krajach gospodarczo zaawansowanych. Ca' Zorzi, Hahn i Sanchez [2007] argumentują, że może to zależeć od wyboru reżimu kursowego i obawą przed rzeczywistym pełnym upłynnieniem kursu (*fear of floating*). Sławiński [2007] uzasadnia, że banki centralne zwłaszcza krajów rozwijających się często unikają prowadzenia swobodnej polityki pieniężnej (czego warunkiem jest kurs płynny), bojąc się właśnie wywoływania dużych wahań kursowych.

Wraz z rozwojem badań nad wpływem zmian kursowych na ceny większą wagę zaczęto przykładac do analiz efektu *pass-through* w kontekście polityki stabilizacyjnej banku centralnego oraz reakcji na szoki krajowe i zewnętrzne (m.in. Hahn [2003]; Ito, Sasaki i Sato [2005]). Pojawiły się pytania, jak zmienność kursu przekłada się na inflację konsumenta (CPI) oraz wzrost gospodarczy, tym samym jaka powinna optymalna polityka pieniężna przy wykazanym niepełnym przełożeniu zmian kursowych na ceny. Na przykładzie wysoce otwartych gospodarek Szwecji i Wielkiej Brytanii, Bernanke i in. [1999] pokazują, że niska inflacja może być rezultatem wprowadzenia przez banki centralne Szwecji i Wielkiej Brytanii strategii bezpośredniego celu inflacyjnego. Bardzo znikomy wpływ zmian kursu na inflację konsumenta potwierdza też szereg badań empirycznych (por. m.in. McCarthy [1999], Gagnon i Ihrig [2004]).

Mishkin [2008] zaznacza jednak, że współzależność między niewielkim efektem *pass-through* a niską inflacją utrzymywać będzie się tak długo, jak władze monetarne będą zdecydowanie reagować na szoki, mogące niekorzystnie wpłynąć na dynamikę cen. Autor dodaje również, że analizy stwierdzające niski wpływ zmian kursowych na inflację w rzeczywistości ukazują uśredniony efekt z różnych okresów, a rzeczywiste reakcje zależą od tego, jaki konkretnie szok dotknął gospodarke.

Dlatego też istotne wydają się badania prowadzone na aktualnych danych, ujmujące dynamicznie zmieniające się otoczenie gospodarcze. W przypadku Polski będą to czynniki wewnętrzne takie jak zmiany reżimu kursowego i całej strategii polityki pieniężnej, oraz zewnętrzne związane m. in. z akcesją do Unii Europejskiej czy skutkami kryzysu światowego ostatnich lat. Pytanie o zależność między kursem a inflacją wydaje się istotne szczególnie w perspektywie wejścia do strefy euro i związanej z tym konieczności spełnienia zarówno kryterium niskiej inflacji, jak i stabilnego kursu.

3. SPECYFIKACJA MODELU

Analiza opisanego we wprowadzeniu zagadnienia wymaga uporządkowania opisu powiązań między wymienianymi wielkościami ekonomicznymi. Powszechnymi narzędziami stosowanymi w analizie polityki ekonomicznej (a w szczególności polityki pieniężnej) są — zaproponowane przez Simsa [1980] — strukturalne modele wektorowej autoregresji (*Structural Vector Autoregressions*, SVAR) (Charemza i Deadman [1997] s. 155, Kokoszcyński [2004] rozdz. 7).

Niech punktem wyjścia będzie K -wymiarowy model VAR(p) zapisany w formie zredukowanej:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \quad (1)$$

gdzie y_t jest wektorem obserwowalnych zmiennych o wymiarach $(K \times 1)$, $A_j (j = 1 \dots p)$ to $(K \times K)$ wymiarowe macierze parametrów stojących przy opóźnionych zmiennych endogenicznych, natomiast u_t reprezentuje K -wymiarowy białoszumowy gaussowski składnik losowy $u_t \sim iin(\mathbf{0}, \Sigma_u)$.

Przy założeniu stacjonarności zmiennych równanie (1) można przedstawić w postaci reprezentacji wektorowej średniej ruchomej (*Vector Moving Average*, VMA) (Lütkepohl [2005]):

$$y_t = u_t + \Phi_1 u_{t-1} + \Phi_2 u_{t-2} + \dots \quad (2)$$

gdzie

$$\Phi_s = \sum_{j=1}^s \Phi_{s-1} A_j, \quad s = 1, 2, \dots \quad (3)$$

oraz $\Phi_0 = I_K$ i $A_j = 0$ dla $j > p$.

Reprezentacja w postaci VMA jest pomocna w wyprowadzeniu podstawowego narzędzia w badaniu czyli funkcji reakcji na bodziec (*impulse response function*, IRF). Określa ona dynamiczną reakcję zmiennej w modelu VAR na zaburzenie składnika losowego. Formalnie, (i, j) -ty element macierzy Φ_s obrazuje wpływ jednostkowego zaburzenia, które dotknęło zmienną j w momencie t , na i -tą zmienną w modelu, po upływie s okresów od momentu t .

Niestety, przy założeniach modelu (1), wartości funkcji reakcji mogą nie odzwierciedlać relacji między jedynie dwiema zmiennymi. Wynika to z możliwości wystąpienia jednoczesnych korelacji między składnikami losowymi poszczególnych równań (macierz Σ_u nie jest diagonalna). Sposobem na przezwycięzenie problemu jednoczesnych korelacji między składnikami losowymi modelu (1) jest przedstawienie go w postaci strukturalnej:

$$B^{-1} y_t = B^{-1} A_1 y_{t-1} + \dots + B^{-1} A_p y_{t-p} + \varepsilon_t \quad (4)$$

Należy dodać, że macierz B jest nieosobliwa wymiaru $(K \times K)$. Natomiast wektor ε_t nazywany jest wektorem strukturalnych innowacji. Modele (1) i (4) łączy zależność:

$$u_t = B\varepsilon_t \quad (5)$$

W związku z powyższym zachodzi relacja $\Sigma_u = B\Sigma_\varepsilon B'$, co przy założeniu braku korelacji i jednostkowych wariancji elementów ε_t , tj. $\Sigma_\varepsilon = I_K$, czyli $\varepsilon_t \sim iin(\mathbf{0}, 1_K)$ daje następującą postać macierzy kowariancji składnika losowego u_t :

$$\Sigma_u = BB'. \quad (6)$$

Widać więc, że wyjściowy model (1) można interpretować jako jedną z wersji modelu strukturalnego (4). Do identyfikacji modelu potrzebne jest nałożenie restrykcji na macierz B. Z własności symetrii macierzy kowariancji otrzymuje się $K(K+1)/2$ restrykcji. Do określenia identyfikowalności modelu strukturalnego niezbędne jest nałożenie dodatkowych $K(K-1)/2$ restrykcji. W niniejszym badaniu macierz B została otrzymana w wyniku dekompozycji Choleskiego macierzy kowariancji Σ_u . W tym przypadku macierz B jest macierzą trójkątną dolną, a cały model ma strukturę rekursywną. Każda kolumna macierzy B jest identyfikowalna z dokładnością do znaku.

Przy takim sposobie identyfikacji, postać (2) modelu można zapisać następująco:

$$y_t = \Lambda_0 \varepsilon_t + \Lambda_1 \varepsilon_{t-1} + \dots \quad (7)$$

gdzie $\Lambda_i = \Phi_i B$ ($i = 0, 1, 2, \dots$). Macierze zawierają wartości tzw. ortogonalnych funkcji reakcji impulsowej, tj. po wyeliminowaniu jednoczesnych korelacji między składnikami losowymi poszczególnych równań. Takie rozwiązanie pozwala wykorzystać kolejne narzędzie interpretacyjne w analizie modeli VAR jakim jest dekompozycja wariancji błędu prognozy.

Jeśli do oznaczenia (i, j) -tego elementu macierzy Λ_s użyje się oznaczenia $\lambda_{ij,s}$ to wariancję błędu prognozy $y_{k,T+h} - y_{k,T+h|T}$ można zapisać jako

$$\sigma_k^2(h) = \sum_{s=0}^{h-1} (\lambda_{k1,s}^2 + \dots + \lambda_{kK,s}^2) = \sum_{j=1}^K (\lambda_{kj,0}^2 + \dots + \lambda_{kj,h-1}^2) \quad (8)$$

Wyrażenie $(\lambda_{kj,0}^2 + \dots + \lambda_{kj,h-1}^2)$ interpretuje się jako udział zmiennej j w wariancji błędu prognozy zmiennej k dla horyzontu h . Po podzieleniu powyższego wyrażenia przez $\sigma_k^2(h)$ otrzymuje się procentowy udział zmiennej j w wariancji błędu prognozy zmiennej k dla h -tego okresu prognozy:

$$\omega_{kj}(h) = \frac{\lambda_{kj,0}^2 + \dots + \lambda_{kj,h-1}^2}{\sigma_k^2(h)} \quad (9)$$

Estymacja modelu SVAR jest zasadna przy założeniu kowariancyjnej stacjonarności zmiennych. Może być tak, że dane niestacjonarne kategorie ekonomiczne zmieniają się w zbliżony do siebie sposób, a łączące je relacje mają trwałe i długookresowy charakter, będący wynikiem wspólnego trendu stochastycznego. W takim przypadku kombinacje liniowe między zmiennymi mogą

być stacjonarne. Jeśli istnieje długookresowy związek między zmiennymi, a odchylenia od długookresowej ścieżki równowagi są stacjonarne to zmienne takie nazywa się skointegrowanymi (Charemza i Deadman [1997]). Pominięcie faktu, że zmienne mogą się ze sobą kointegrować, prowadzić będzie do utraty istotnej informacji w modelowaniu i pozbawienia modelu własności długookresowych (por. Charemza i Deadman [1997], Favero [2001]).

Narzędziem pozwalającym na analizę szeregów skointegrowanych, dekompozycję ich zmienności na relacje długookresowe, krótkookresowe wahania i losowe zaburzenia (a także eliminującym wspólne trendy stochastyczne) są wektorowe modele korekty błędu (*Vector Error Correction Model*, VECM) (Johansen [1988]).

Zredukowaną formę K -wymiarowego modelu VAR(p) postaci (1) można przekształcić do następującej postaci:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \quad (10)$$

gdzie:

$$\Pi = -(I_K - A_1 - \dots - A_p), \quad (11)$$

$$\text{zaś dla } p > 1 \quad \Gamma_i = -(A_{i+1} + \dots + A_p), \quad (12)$$

$$i = 1, \dots, p-1$$

Jeśli rząd macierzy Π równa się r i spełniony jest warunek $1 \leq rz(\Pi) \leq K - 1$, to można dokonać dekompozycji macierzy Π na iloczyn dwu macierzy pełnego rzędu kolumnowego, mianowicie

$$\Pi = \alpha \beta' \quad (13)$$

gdzie α i β mają wymiary $(K \times r)$. W takim przypadku zapisaną równaniem (10) formę modelu VAR nazywa się wektorowym modelem korekty błędu (VECM). Macierz β zwana jest macierzą kointegrującą, a każda kolejna jej kolumna odpowiada wektorowi kointegrującemu liniowo niezależnemu od pozostałych, wyznaczającemu równowagę długookresową. Iloczyn $\beta' y_t$ jest stacjonarny i stanowi jeden z r mechanizmów korekty błędu. Macierz α określa szybkość, z jaką dokonuje się korekta odchyłeń poszczególnych zmiennych od równowagi długookresowej, stąd zwana jest macierzą dostosowań (por. Charemza i Deadman [1997], Syczewska [2007]).

Model VECM można traktować jako zredukowaną postać strukturalnego wektorowego modelu korekty błędu (SVECM), który z kolei jest izomorficznym przekształceniem modelu SVAR. Dla zdefiniowanego równaniem (4) modelu SVAR odpowiadająca mu postać SVECM przedstawia się następująco:

$$B^{-1} y_t = B^{-1} \alpha \beta' y_{t-1} + B^{-1} \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t \quad (14)$$

gdzie wszystkie oznaczenia pozostają niezmiennie, tj. y_t jest K -wymiarowym wektorem obserwowalnych zmiennych, α i β są odpowiednio macierzami dostosowań i kointegrującą o wymiarach $(K \times r)$, $\Gamma_j (j = 1 \dots p - 1)$ to $(K \times K)$ wymiarowe macierze parametrów krótkookresowych, natomiast ε_t reprezentuje K -wymiarowy białoszumowy składnik losowy $\varepsilon_t \sim iin(0, I_K)$ i zachowana jest zależność (5).

Zgodnie z koncepcją wysuniętą przez McCarthy'ego [1999] siłę przełożenia zmian kursowych na ceny można mierzyć jako stosunek skumulowanej zmiany odpowiedniego agregatu cenowego, jaka nastąpiła od momentu zmiany kursu, do łącznej zmiany kursu walutowego, która tę zmianę cen wywołała. Powyższa zależność równa jest stosunkowi wartości funkcji reakcji danego agregatu cenowego w momencie $t + s$ na szok kursowy powstały w momencie t , do wartości funkcji reakcji kursu walutowego na innowację, która tę zmianę wywołała. Zależność tę dla cen importu można zapisać jako:

$$PT_{t,t+s} = \frac{IRF(imp, \mathcal{E}^{neer}, s)}{IRF(neer, \mathcal{E}^{neer}, s)} \quad (15)$$

gdzie $IRF(i, j, s) = \partial y_{i,t+s} / \partial \varepsilon_{jt}$ oznacza wartość odpowiedniej funkcji reakcji na impuls, a *imp* oraz *neer* oznaczają odpowiednio ceny importu oraz kurs walutowy.

Współczynnik zdefiniowany równaniem (15) można wykorzystać do analizy reakcji cen jak i pozostałych wielkości ekonomicznych na zaburzenia innych składników losowych (pamiętając, że szoki powinny być wzajemnie nieskorelowane, co przy założeniu łącznej normalności oznacza ich niezależność stochastyczną).

Zgodnie z zależnościami przedstawionymi na schemacie 1, przedmiotem zainteresowania jest analiza przenoszenia zmian kursowych nie tylko na ceny towarów importowanych, lecz także przekładanie się tych zmian na poszczególnych etapach łańcucha cenowego, tj. od cen importu, poprzez ceny produkcji sprzedanej przemysłu (PPI), aż do cen towarów i usług konsumpcyjnych (CPI) (por. Blanchard [1982]). Należy przy tym zaznaczyć, że zmiany kursu mogą oddziaływać na ceny na poszczególnych etapach w sposób bezpośredni lub być efektem przeniesienia zmian z poprzedniego ogniwa, co ze względu na wpływ na szybkość i siłę zjawiska *pass-through* uwzględniono podczas formalizacji modelu. Dodatkowo, w celu precyzyjnego opisu zmian kursu walutowego do analizy włączono te czynniki, które dotykając gospodarkę w istotny sposób wpływają na zmienność kursu, przede wszystkim takie jak zewnętrzne szoki popytowe i podażowe. Nie można też pominąć efektów wpływu polityki pieniężnej na cały modelowany system.

Bazując na pracach McCarthy'ego [1999] i Hahn [2003] zaproponowano łańcuch transmisji, mający oddawać przyczynowo-skutkowy charakter przenoszenia się zaburzeń między zmiennymi, w tej samej jednostce czasu (por. definicję

łańcucha dystrybucji zaproponowaną przez Cholewińskiego [2008]). Odpowiednie uporządkowanie zmiennych jest niezwykle istotne, gdyż może stanowić przesłankę teoretyczną do narzucenia odpowiedniej struktury na macierz B w opisanym uprzednio modelu SVAR.

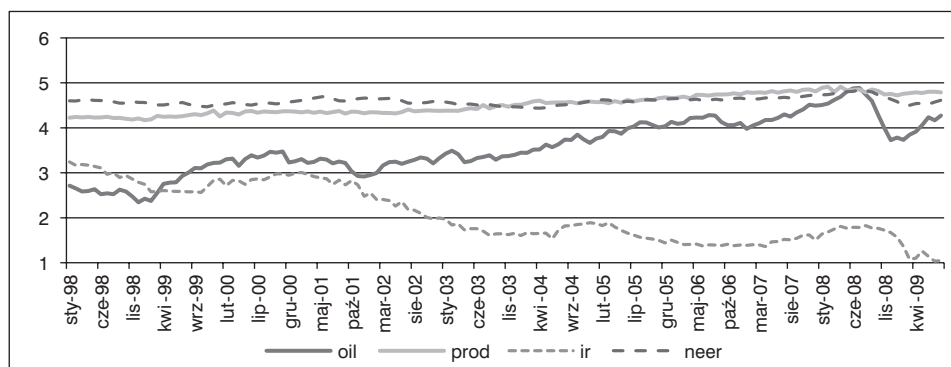
Na pierwszych dwóch miejscach łańcucha umiejscowione zostały zmienne oddające efekty szoków odpowiednio podażowych (*oil*) i popytowych (*prod*). Oznacza to, że szoki podażowe wpływają jednocześnie na wszystkie pozostałe zmienne w systemie, natomiast nie ma zależności odwrotnej. Szoki popytowe w tej samej jednostce czasu wpływać będą na wszystkie zmienne za wyjątkiem pierwszej, itd. Na trzecim miejscu umiejscowiono zmienną reprezentującą narzędzie banku centralnego (*ir*), co oznacza, że reakcje władz monetarnych natychmiastowo uwzględniają zmiany czynników umiejscowionych na pierwszym i drugim miejscu. Odwrotną kolejność między stopą procentową a zmienną obrazującą popytową stronę gospodarki zaobserwować można w pracy Hahn [2003], która argumentuje, że dane o PKB dostępne są z opóźnieniem, wobec czego bardziej zasadne wydaje się umiejscowienie tej zmiennej na trzeciej pozycji. Zdaniem autora ciężko utrzymać to założenie, ze względu na ewidentne opóźnione efekty decyzji władz monetarnych w sferze realnej. Kolejną zmienną w łańcuchu jest kurs walutowy (*neer*). Jego zmienność w danym okresie będzie determinowana przez szoki podażowe, popytowe oraz nieoczekiwane zmiany stopy procentowej, które miały w tym okresie miejsce. Ostatnie trzy elementy łańcucha to odpowiednio ceny importu (*imp*), producenta (*ppi*) i konsumenta (*cpi*), co odzwierciedla poszczególne etapy produkcji bądź dystrybucji dóbr.

Na podstawie opisu przedstawionego powyżej można przedstawić konkretny wektor zmiennych endogenicznych postaci: $y_t = [oil_t, prod_t, ir_t, neer_t, imp_t, ppi_t, cpi_t]'$ natomiast zapis równania (5) dla przyjętych oznaczeń zmiennych będzie miał postać:

$$\begin{bmatrix} u_t^{oil} \\ u_t^{prod} \\ u_t^{ir} \\ u_t^{neer} \\ u_t^{imp} \\ u_t^{ppi} \\ u_t^{cpi} \end{bmatrix} = \begin{bmatrix} b_{11} & & & & & & \\ b_{21} & b_{22} & & & & & \\ b_{31} & b_{32} & b_{33} & & & & \\ b_{41} & b_{33} & b_{34} & b_{44} & & & \\ b_{51} & b_{34} & b_{35} & b_{45} & b_{55} & & \\ b_{61} & b_{35} & b_{36} & b_{46} & b_{56} & b_{66} & \\ b_{71} & b_{36} & b_{37} & b_{47} & b_{57} & b_{67} & b_{77} \end{bmatrix} \begin{bmatrix} \mathcal{E}_t^{oil} \\ \mathcal{E}_t^{prod} \\ \mathcal{E}_t^{ir} \\ \mathcal{E}_t^{neer} \\ \mathcal{E}_t^{imp} \\ \mathcal{E}_t^{ppi} \\ \mathcal{E}_t^{cpi} \end{bmatrix} \quad (16)$$

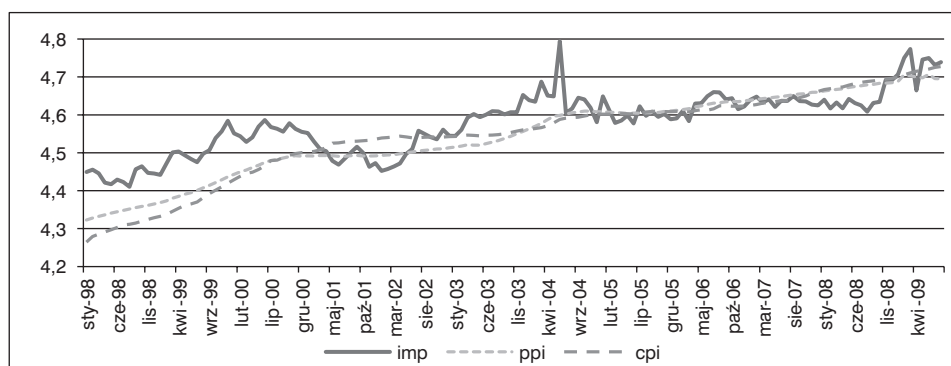
4. WSTĘPNE ANALIZY BADANYCH SZEREGÓW CZASOWYCH

Podstawę badania stanowi następujący zbiór zmiennych: *oil* — logarytm naturalny średniej światowej ceny baryłki ropy naftowej (w USD), *prod* — logarytm naturalny wyrównanego sezonowo indeksu produkcji przemysłowej w Polsce (2005=100), *ir* — logarytm naturalny średniej rocznej rynkowej stopy procentowej 1-miesięcznych lokat międzybankowych (WIBOR 1M, Warsaw Interbank Offer Rate), *neer* — logarytm naturalny indeksu nominalnego efektywnego kursu walutowego złotego (2005=100) oraz wyrównane sezonowo logarytmy naturalne indeksów cen: *imp* — importu do Polski (2005=100), *ppi* — produkcji sprzedanej przemysłu w Polsce (2005=100), *cpi* — koszyka towarów i usług konsumpcyjnych w Polsce (2005=100).



Źródło: opracowanie własne.

Wykres 1. Ceny ropy naftowej oil_t , produkcja przemysłowa $prod_t$, stopa procentowa ir_t , kurs walutowy $neer_t$ (logarytmy naturalne)



Źródło: opracowanie własne.

Wykres 2. Indeksy: cen importu imp_t , cen producenta ppi_t , cen konsumenta cpi_t (logarytmy naturalne)

Wykorzystane dane obejmują okres od stycznia 1998 roku do sierpnia 2009 roku, co stanowi 140 miesięcznych obserwacji. Aby zachować interpretacyjną spójność wyników, początek rozważanego okresu wiąże się z wprowadzeniem kluczowych zmian instytucjonalnych w NBP oraz zmianą strategii polityki pieniężnej na bezpośredni cel inflacyjny. Dodatkowo należy podkreślić, że okres od lutego 1998 uważany jest za faktyczny początek obowiązywania systemu kursu płynnego (Sławiński [1999]), pamiętając jednocześnie, że obowiązująca do kwietnia 2000 r. miesięczna administracyjna deprecjacja złotego, mogła determinować zachowania podmiotów gospodarczych ustalających ceny na rynku krajowym. Jest to istotne w analizie zjawiska *pass-through*, którego poziom może zależeć od reżimu kursowego. Koniec wybranego okresu (sierpień 2009) wiąże się z dostępnością danych w momencie rozpoczęcia analizy.

W celu analizy stopnia zintegrowania procesów generujących szeregi wykorzystane zostały testy pierwiastka jednostkowego Dickeya i Fullera, Phillipsa i Perrona oraz test stacjonarności KPSS (dokładne wyniki analizy przedstawia Malczyk [2010]). Wyniki testów wskazały na stacjonarność pierwszych przyrostów cen ropy naftowej, indeksu produkcji przemysłowej (czyli dynamiki produkcji) oraz kursu walutowego. W przypadku stopy procentowej testy ADF i PP bez zmiennych deterministycznych wskazały na stacjonarność na poziomach, lecz gdy zmniejszono poziom istotności do 0,01 nie było podstaw do odrzucenia hipotezy zerowej. Może to sugerować błąd II rodzaju, tym bardziej, że pozostałe testy sugerowały wynik odmienny. Z analizy wykresu (patrz wykres 1) wynika, że badany szereg nie jest generowany przez proces stacjonarny, wobec czego uznano, że jest zintegrowany w stopniu pierwszym. W przypadku szeregów tworzących łańcuch cenowy wątpliwości nie budziły jedynie rezultaty testów cen importu, wskazujące na stacjonarność ich pierwszych przyrostów czyli inflacji importu. Nie ma natomiast jednoznacznej odpowiedzi na pytanie czy stacjonarne są inflacja producenta oraz inflacja konsumenta. Wartości statystyk w testach ADF oscylują na granicy odrzucenia hipotezy zerowej. Co więcej, zaobserwować można wzrost ich wartości wraz ze zbliżaniem się obliczonej pomocniczo wartości statystyki DW do 2. W związku z tym wyniki sugerujące odrzucenie H_0 mogą być zafałszowane występowaniem autokorelacji składnika losowego w równaniu regresji. Stacjonarność inflacji sugerują natomiast testy PP. Testy KPSS nie dają jednoznacznych rozstrzygnięć.

Przyglądając się problemowi od strony teoretycznej należy zauważyć, że okres transformacji w Polsce łączy się z prowadzeniem przez władze monetarne polityki dezinflacyjnej. Może to sugerować, że indeksy cenowe są szeregami zintegrowanymi w stopniu drugim (Majsterek [2008]). Z drugiej strony okres analizy nie obejmuje pierwszych lat przemian, kiedy władze monetarne walczyły z uporczywą inflacją. Natomiast późniejsze wprowadzenie strategii BCI przyczyniło się do stabilizacji zmienności inflacji.

Biorąc pod uwagę powyższe argumenty, ciężko jednoznacznie rozstrzygnąć o stopniu integracji indeksów PPI i CPI (konkluzje testów stopnia zintegrowania szeregów czasowych przedstawia tabela 1). W związku z tym, dalsza analiza przeprowadzona została w dwóch wariantach, uwzględniających przypadki: zintegrowania inflacji producenta i konsumenta w stopniu pierwszym oraz traktującym te szeregi jako stacjonarne. Rozważone zostały 2 warianty wektora zmiennych endogenicznych:

$$\text{Wariant I: } y_t^{(I)} = (oil_t, prod_t, ir_t, neer_t, imp_t, ppi_t, cpi_t)'$$

$$\text{Wariant II: } y_t^{(II)} = (oil_t, prod_t, ir_t, neer_t, imp_t, \Delta ppi_t, \Delta cpi_t)'$$

Tabela 1

Podsumowanie analizy stacjonarności szeregów czasowych

Zmienna	ADF	PP	KPSS	Decyzja
<i>oil</i>	I(1)	I(1)	I(1)	I(1)
<i>prod</i>	I(1)	I(1)	I(1)	I(1)
<i>neer</i>	I(1)	I(1)	I(1)	I(1)
<i>ir</i>	I(0)/I(1)	I(0)/I(1)	I(1)	I(1)
<i>imp</i>	I(1)	I(1)	I(1)	I(1)
<i>ppi</i>	I(1)/I(2)	I(1)	I(1)	I(1)/I(2)
<i>cpi</i>	I(1)/I(2)	I(1)	I(1)/I(2)	I(1)/I(2)

Źródło: opracowanie własne.

W badaniu założono, że każdy z wariantów zawiera zmienne zintegrowane w stopniu pierwszym, to jest: $y_t^{(I)} \sim I(1)$ jak i $y_t^{(II)} \sim I(1)$, tym samym analizie zostało poddane występowanie potencjalnej kointegracji typu CI(1,1). Należy przy tym zaznaczyć, że obniżenie stopnia integracji zmiennych I(2) do I(1) poprzez różnicowanie nie jest jedynym rozwiązaniem podczas analizy kointegracyjnej. Możliwe jest też przekształcenie modelu VECM do postaci pozwalającej analizować związki kointegracyjne typu CI (2,2), co rozszerza możliwość interpretacji ekonomicznych (por. Majsterek [2003], [2008]).

W celu odpowiedniego doboru rzędu opóźnień p , pomocniczo przeprowadzono analizę reszt modelu VAR. Ze względu na fakt, że analizowane poziomy zmiennych charakteryzują się trendem oraz niezerową średnią, wprowadzono do modelu odpowiednie zmienne deterministyczne. Dodatkowo włączona została zmienna sztuczna $impulse_t = 1$ dla $t = 78$, $impulse_t = -1$ dla $t = 79$ i $impulse_t = 0$ dla $t \neq 78$ i 79 . Okres 78 w szeregu oznacza czerwiec 2004 roku, miesiąc po wejściu polski do Unii Europejskiej. Dane cen importu wykazały w tamtym momencie chwilowy, silny wzrost, po czym sytuacja wróciła do równowagi, czyli inflacja importu w okresie 78 charakteryzuje się silnym skokiem oraz analogicznym

spadkiem w okresie następnym. W związku z tym zasadne wydaje się zniesienie wpływu tej zmiennej na oszacowania parametrów modelu (por. Cholewiński [2008]). Ostatecznie do dalszych analiz wykorzystano model postaci:

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + CD_t + u_t \quad (17)$$

gdzie $D_t = [1, t, \text{impulse}_t]'$ jest wektorem zmiennych deterministycznych, a C macierzą współczynników stojących przy tych zmiennych.

W celu wyboru odpowiedniego rzędu opóźnienia dokonano analizy autokorelacji reszt oszacowane zostały modele VAR w obu wariantach, wychodząc od najmniejszego rzędu opóźnień, tj. $p = 1$. Następnie analizie poddawano składniki losowe tych modeli. Gdy wartość funkcji autokorelacji dla reszt modelu nie była zadowalająca zwiększano rząd opóźnienia, aż do momentu uzyskania wyników pozwalających uznać reszty za niewykazujące istotnej autokorelacji. Dla obu wariantów przy rzędzie opóźnienia $p = 3$ uznano, że składniki losowe nie wykazują istotnej autokorelacji. Inną konkluzją analizy było stwierdzenie, że poszerzenie modelu VAR o dodatkowe roczne opóźnienie mogłoby korzystnie wpłynąć na rozkład składnika losowego w tym równaniu. Zwiększenie rzędu opóźnień modelu, przy tak krótkim szeregu czasowym, prowadzi jednak do spadku istotności oszacowanych parametrów. W związku z tym rząd równy 3 uznano za wystarczający (szerzej Malczyk [2010]).

Ponadto przeprowadzono test Jarque-Berra na normalność rozkładu składników losowych w każdym z równań. Przy rzędzie opóźnienia $p = 3$ składniki losowe pięciu pierwszych równań można uznać za podlegające rozkładowi normalnemu. Natomiast zwiększanie liczby opóźnień (podczas badania wykonano analizy do VAR(6)) nie pozwala na uznanie za normalne składników losowych w dwóch ostatnich równaniach, zarówno w wariancie pierwszym jak i drugim. Szczegółową analizę reszt dla modelu z trzema opóźnieniami przedstawia tabela 2. Wartości współczynnika skośności oraz kurtozy dla równań cen producenta i konsumenta (wariant I) oraz ich odpowiednich pierwszych przyrostów (wariant II) znacznie odbiegają od wartości teoretycznych tych momentów dla rozkładu normalnego. Ciężko zatem utrzymać wyartykułowane w rozdziale trzecim, robocze założenie o normalności składnika losowego modelu VAR. Dla celu analizy kointegracji założenie to zostało nadal podtrzymane, przy czym należy pamiętać, że do wniosków płynących z tej analizy należy podchodzić ostrożnie.

W celu ustalenia rzędu kointegracji badanych szeregów wykorzystane zostały zaproponowane przez Johansena [1988] testy śladu oraz test największej wartości własnej. Ostatecznie wykorzystany został model postaci:

$$\Delta y_t = \Upsilon D_t + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + CD_t + u_t \quad (18)$$

gdzie $D_t = [1, \text{impulse}_t]'$.

Podsumowanie testu JB dla modelu z trzema opóźnieniami

	Wariant I				Wariant II			
	JB	p-value	skośność	kurtoza	JB	p-value	skośność	kurtoza
u_1	1,470	0,479	-0,192	2,662	2,192	0,334	-0,289	2,766
u_2	1,178	0,554	0,228	2,989	1,190	0,551	0,227	2,936
u_3	4,259	0,118	0,125	3,833	4,777	0,091	0,021	3,920
u_4	2,647	0,266	-0,338	2,892	5,667	0,058	-0,496	3,141
u_5	0,270	0,873	0,019	3,216	2,362	0,306	-0,138	3,586
u_6	1214,130	0,000	1,231	17,483	1062,417	0,000	1,614	16,358
u_7	34,577	0,000	0,601	5,128	21,885	0,000	0,412	4,791

Źródło: opracowanie własne.

Brak restrykcji na stałą w modelu oznacza, że poziomy zmiennych obserwowalnych mogą podlegać trendowi liniowemu, lecz jest on ortogonalny w stosunku do przestrzeni kointegracyjnej. Innym rozwiązaniem byłoby włączenie trendu lub stałej do przestrzeni kointegracyjnej. Takie warianty modeli, zgodnie z sugestiami Lütkepohla [2005] (podrozdział 8.2.8), były testowane za pomocą testów logarytmu wiarygodności. Konkluzją tej procedury było stwierdzenie, że najlepszym modelem jest ten zaproponowany równaniem (18).

W celu doboru odpowiedniej liczby opóźnień w modelu, oprócz analizy reszt posłużono się dodatkowo sugerowanymi w literaturze kryteriami informacyjnymi: Akaike (AIC), Schwarz (SC), Hannana-Quinna (HQ), końcowym błędem predykcji wewnątrzpróbkowej (ang. *Final Prediction Error*, FPE) oraz testem logarytmu wiarygodności (LR). Ponieważ wykorzystane procedury dostarczyły różnych rezultatów, testy rzędu kointegracji wykonane zostały dla liczby opóźnień modelu VAR od 1 do 7. Wyniki przedstawia tabela 3. W ostatniej kolumnie każdego z wariantów zaznaczono, czy na dany rząd opóźnień wskazywało jakiegokolwiek kryterium.

Jeden rząd opóźnień, mimo wskazań kryteriów HQ i SC odrzucono, gdyż w modelu VECM oznaczałoby to brak opóźnionych przyrostów. Na podstawie analizy reszt modelu VAR podjęto również decyzję o wykluczeniu rzędu opóźnień równego dwa. W przypadku trzech opóźnień w wariancie pierwszym zarówno test śladu jak i największej wartości własnej wskazują na dwa wektory kointegrujące. W związku z tym dalszej analizie poddano model VECM o takiej specyfikacji. Identyczne postąpiono w wariancie drugim, gdzie na rząd opóźnień równy trzy wskazywały kryteria informacyjne FPE i AIC. Liczbę wektorów kointegrujących określono zgodnie z wynikiem testu największej wartości własnej.

Tabela 3

Podsumowanie testów rzędu kointegracji w zależności od liczby opóźnień modelu VAR

p	Wariant I			Wariant II		
	TRACE	MAX	Kryterium	TRACE	MAX	Kryterium
1	4	4	HQ, SC	5	3	HQ, SC
2	4	2		3	3	
3	2	2		5	2	FPE, AIC
4	2	0	FPE	3	0	
5	2	1		3	1	
6	2	1		2	2	LR
7	3	1	LR, AIC	3	2	

Źródło: opracowanie własne.

Należy zaznaczyć, że każda decyzja — zarówno co do rzędu opóźnień jak i liczby wektorów kointegrujących — ma charakter arbitralny. Problem ten można rozwiązać na gruncie bayesowskim. Przykład analizy kointegracyjnej z wykorzystaniem bayesowskiego porównywania modeli i łączenia wiedzy (którego podstawy omawia Osiewalski [2001], punkt 1.2) można odnaleźć m.in. w pracy Wróblewskiej [2009].

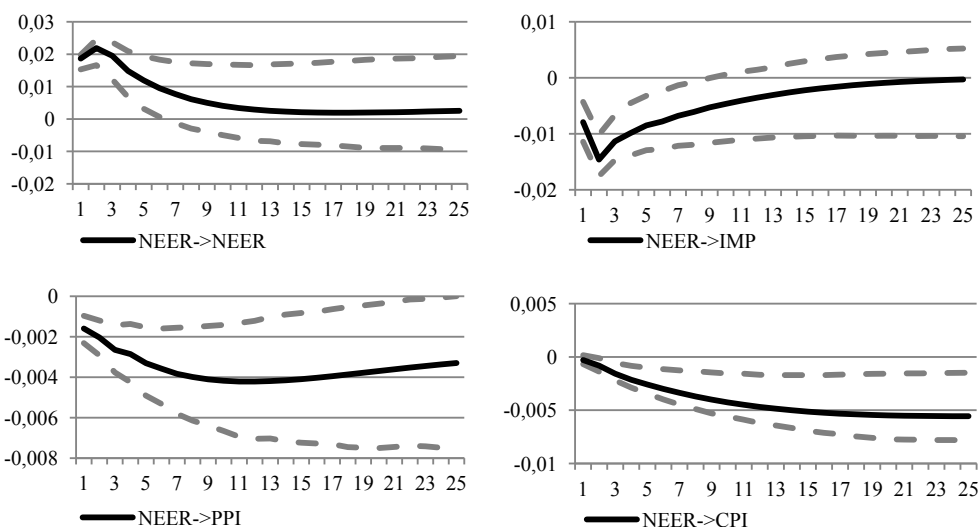
5. WYNIKI EMPIRYCZNE

Konkluzją analizy kointegracji zmiennych jest oszacowanie dwóch wariantów modeli, z których każdy zawiera dwa wektory kointegrujące. Zgodnie z podejściem Johansena [1988] w pierwszym etapie estymacji wykorzystano regresję niepełnego rzędu (*reduced rank regression*) do oszacowania wartości wektorów kointegrujących. Posiadając oceny metody największej wiarygodności dla macierzy β , oszacowano model VECM zwykłą metodą najmniejszych kwadratów (równanie po równaniu). Następnie, w celu wyeliminowania parametrów charakteryzujących się słabą statystyczną istotnością wykorzystano procedurę sekwencyjnej redukcji zmiennych (*System Sequential Elimination of Regressors, SER*). Polega ona na tym, że w każdym kroku procedury parametr z najniższym ilorazem t jest potencjalnie eliminowany z modelu, decyzja o wykluczeniu zmiennej podejmowana jest na podstawie wybranego kryterium informacyjnego. W niniejszym badaniu wykorzystano kryterium Akaike. Po nałożeniu zerowych restrykcji na statystycznie nieistotne parametry model VECM został oszacowany Uogólnioną Metodą Najmniejszych Kwadratów. Ostatnim etapem estymacji była dekompozycja Choleskiego macierzy kowariancji składnika losowego w modelu VECM

na iloczyn macierzy trójkątnych BB' , co po przemnożeniu macierzy parametrów przez odwrotność macierzy warunków strukturalnych B^{-1} doprowadziło do przekształcenia do modelu SVECM (zgodnie z równaniem (11)). Wartości oszacowań obu wariantów modelu zawiera praca Malczyka [2010].

Współczynniki reakcji na impuls

Wartości współczynników odpowiedzi na impuls bazują na ocenach punktowych funkcji pierwotnych parametrów modelu. Oby podjąć próbę oceny niepewności tych oszacowań, za pomocą metod bootstrapowych wyznaczono 95% przedziały ufności wykorzystując metodę percentyli (por. Efron i Tibshirani [1993]). W tym celu przeprowadzono 1000 prób bootstrapowych.



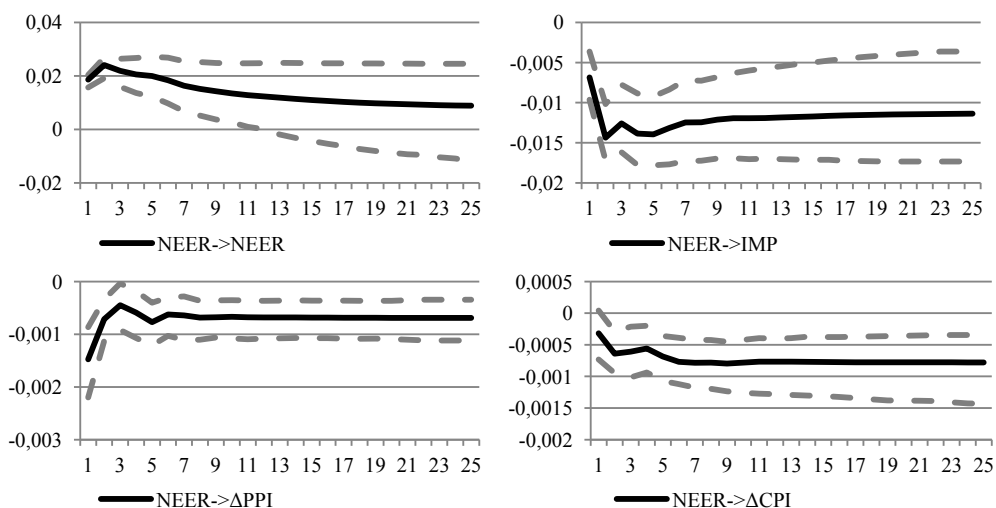
Źródło: opracowanie własne.

Wykresy 3–6. Funkcje reakcji kursu walutowego, cen importu, producenta i konsumenta na impuls kursu walutowego (wariant pierwszy)

W pierwszym kroku obliczono wartości funkcji reakcji cen importu, producenta i konsumenta na impuls kursu walutowego. Bezpośrednio wyniki te otrzymać można w wariancie pierwszym modelu, gdzie wszystkie zmienne są wyrażone w poziomach. Szokowy wzrost kursu oznacza aprecjację waluty krajowej. Jak widać na wykresach 3–6, w krótkim okresie najsilniejsza jest spadkowa reakcja cen importu, natomiast w perspektywie dwóch lat ceny wracają do poziomu wyjściowego. Inaczej sytuacja wygląda w przypadku cen producenta i konsumenta, których reakcje są o wiele słabsze, natomiast utrzymują się w dłuższym

okresie. Dodatkowo przedstawiono wykres autoodpowiedzi kursu walutowego. Jednorazowa zmiana kursu powoduje jeszcze większy wzrost w okresie pierwszego kwartału, po czym w perspektywie 14–15 miesięcy kurs stabilizuje się na poziomie nieznacznie wyższym od wyjściowego. Tym samym reakcja ta będzie miała bezpośredni wpływ na wielkość współczynników *pass-through*, gdyż równoległe dostosowania po stronie kursu będą obniżać skalę przenoszenia zmian kursowych na ceny.

Zbliżone wartości funkcji reakcji cen importu oraz autoodpowiedzi kursowej zaobserwować można w wariancie drugim (wykresy 7–10). Ceny importu nie powracają jednak do stanu pierwotnego, lecz w dłuższej perspektywie utrzymują się na niższym poziomie. Bezpośrednio z modelu odczytać można również reakcje inflacji PPI i CPI na zmiany kursu (wartości na prawej osi). W przypadku obu agregatów cenowych reakcja jest zbliżona, w długim okresie można zaobserwować, że dynamika cen konsumenta charakteryzuje się mniejszą wrażliwością na wahania kursu.

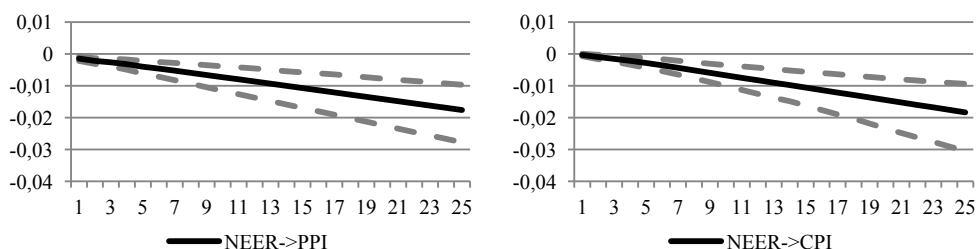


Źródło: opracowanie własne.

Wykresy 7–10. Funkcje reakcji kursu walutowego, cen importu, inflacji producenta i konsumenta na impuls kursu walutowego (wariant drugi)

Wykorzystując skumulowane wartości funkcji reakcji zmiennych wyrażonych w postaci pierwszych przyrostów na impuls walutowy, uzyskano wartości funkcji reakcji poziomów cen producenta i konsumenta w wariancie drugim. Wykresy 11–12 przedstawiają skumulowaną funkcję reakcji inflacji producenta i konsumenta na impuls kursu walutowego. Podobnie jak w przypadku wariantu pierwszego reakcja cen producenta jest w krótkim okresie silniejsza od CPI, na-

tomiast w dłuższej perspektywie zaobserwować można zależność odwrotną. W tym ostatnim przypadku reakcja poziomów cen producenta i konsumenta ma charakter silnie malejący, nie wykazujący ograniczeń w zadanej perspektywie dwóch lat. Jest to konsekwencją reakcji przyrostów cen, które w długim okresie stabilizują się na poziomie wyższym od zera, co prowadzi do narastającej reakcji poziomów zmiennych. Próbę ekonomicznego wyjaśnienia tego zjawiska podejmuje Cholewiński [2008], interpretując podobny rezultat jako efekt ciągłego wzrostu cen w gospodarce.



Źródło: opracowanie własne.

Wykresy 11–12. Funkcje skumulowanych odpowiedzi inflacji producenta i konsumenta na impuls kursu walutowego (wariant drugi)

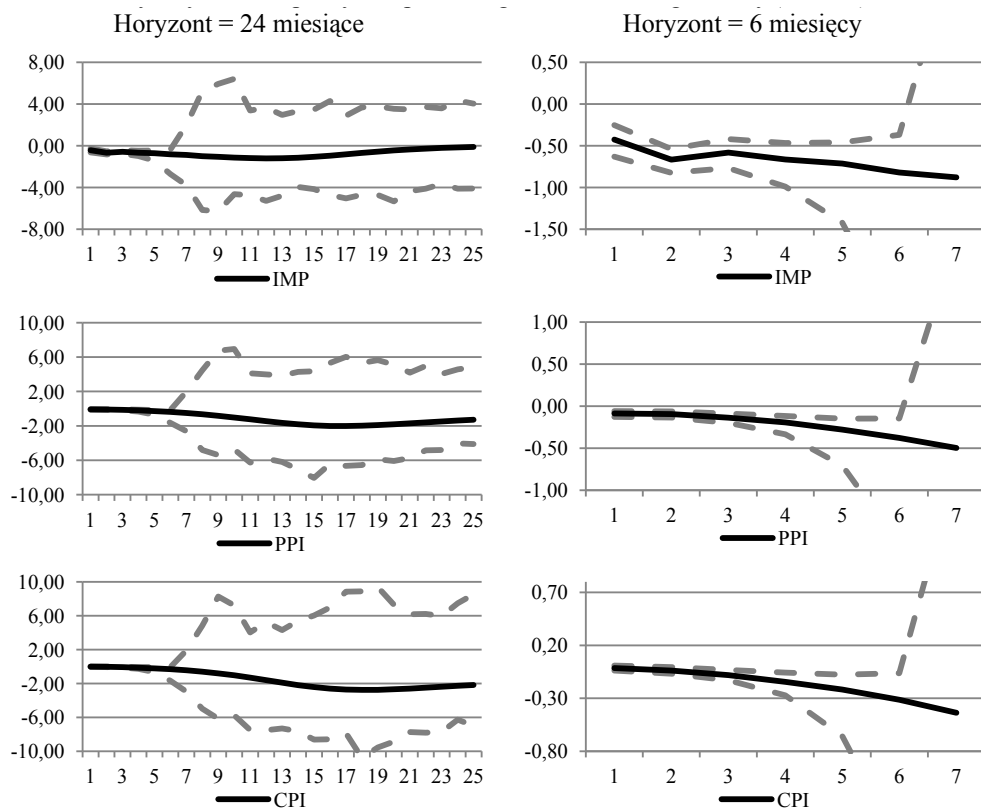
Współczynniki przełożenia zmian kursowych

Dysponując wartościami funkcji odpowiedzi indeksów cenowych na impuls kursu walutowego oszacowano wartości *pass-through*. Podobnie jak w przypadku funkcji reakcji dla ocen wskaźników *pass-through* wyznaczono bootstrapowe przedziały ufności. Oba warianty pozwalają na analizę szybkości i siły przekładania się zmian kursowych na ceny na każdym etapie łańcucha cenowego.

Na wykresach 13–18 zaprezentowane zostały wartości współczynników *pass-through* w wariacie pierwszym modelu. Pierwotnie wyznaczono wartości wskaźników dla okresu prognozy równego 24 miesiące (wykresy w pierwszej kolumnie). Można jednak zauważyć, że dla horyzontu prognozy powyżej 6 miesięcy drastycznie wzrasta niepewność oszacowań wskaźników, znacznie osłabiająca jakiegokolwiek wnioski ekonomiczne. W związku z tym ocenę efektu przekładania się zmian kursowych na ceny ograniczono do okresu krótkiego, zawierającego się w jednym półroczu (wykresy w drugiej kolumnie).

Zgodnie z oczekiwaniami najsilniejszą reakcję można zaobserwować po stronie cen importu. Z okresu na okres skala przełożenia zmian kursowych jest coraz wyższa, osiągając po trzech miesiącach wartość -0,66 (przedział ufności ograniczają wartości -0,99 i -0,47), postępującą w kolejnych miesiącach. Reakcje po stro-

nie cen producenta i konsumenta są również ujemne, lecz znacznie słabsze od cen dóbr importowanych. Jednocześnie charakteryzują się spadkiem niepewności oszacowań. Po upływie kwartału wskaźnik przełożenia dla cen producenta wynosi $-0,19$ (z przedziałem ufności pomiędzy $-0,33$ i $-0,12$), natomiast ocena reakcji cen konsumenta w takim okresie kształtuje się na poziomie $-0,14$ (z przedziałem ufności pomiędzy $-0,27$ i $-0,05$).

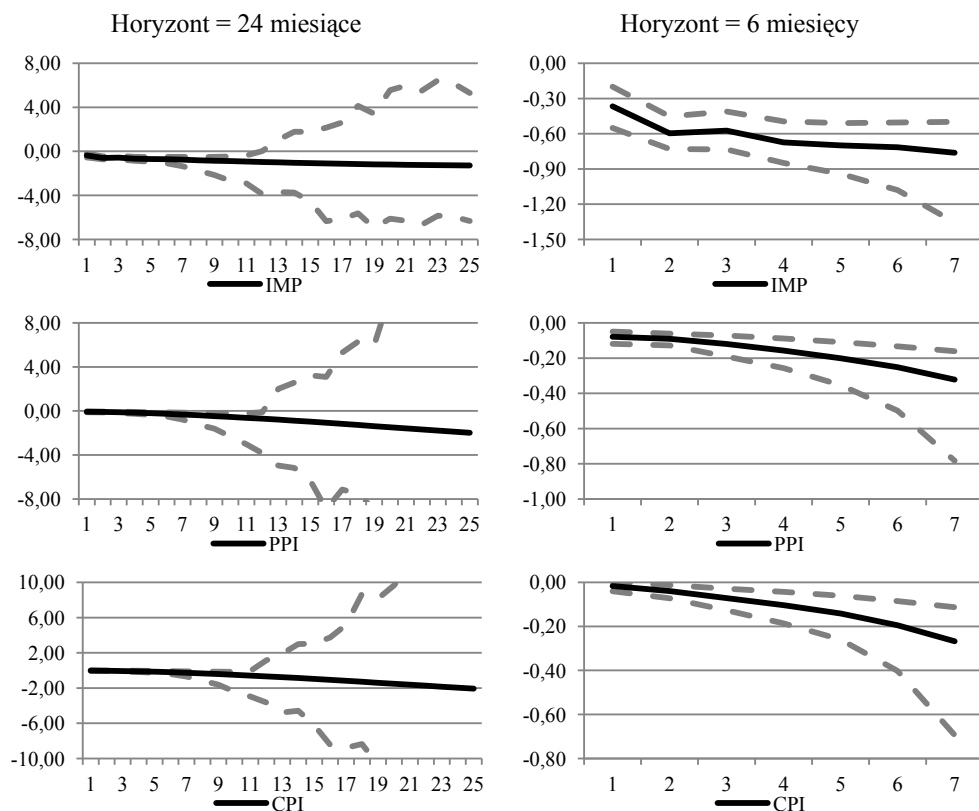


Źródło: opracowanie własne.

Wykresy 13–18. Współczynniki pass-through kursu walutowego na ceny (wariant I)

Skala przełożenia zmian kursu na ceny w wariacie drugim w krótkim okresie przebiega w sposób zbliżony. Istotne różnice w ocenach punktowych można zaobserwować po upływie roku. Jednak podobnie jak w wariacie pierwszym, ze względu na drastycznie szerokie przedziały ufności dla ocen parametrów, interpretację ograniczono do okresu krótkiego. Podobnie jak w wariacie pierwszym reakcja cen importu jest najsilniejsza i po upływie 3 miesięcy wynosi $-0,67$ z przedziałem ufności $(-0,83;-0,49)$. Reakcja cen producenta jest słabsza i wynosi $-0,15$

z przedziałem ufności $(-0,26; 0,09)$, natomiast w sposób najbardziej delikatny reagują ceny konsumenta, dla których ocena wskaźnika *pass-through* po upływie kwartału wynosi $-0,10$ z przedziałem ufności $(-0,19; 0,04)$. W perspektywie kolejnych okresów reakcje wymienianych agregatów są silniejsze, jednocześnie wzrasta niepewność oszacowań wskaźników przełożenia.

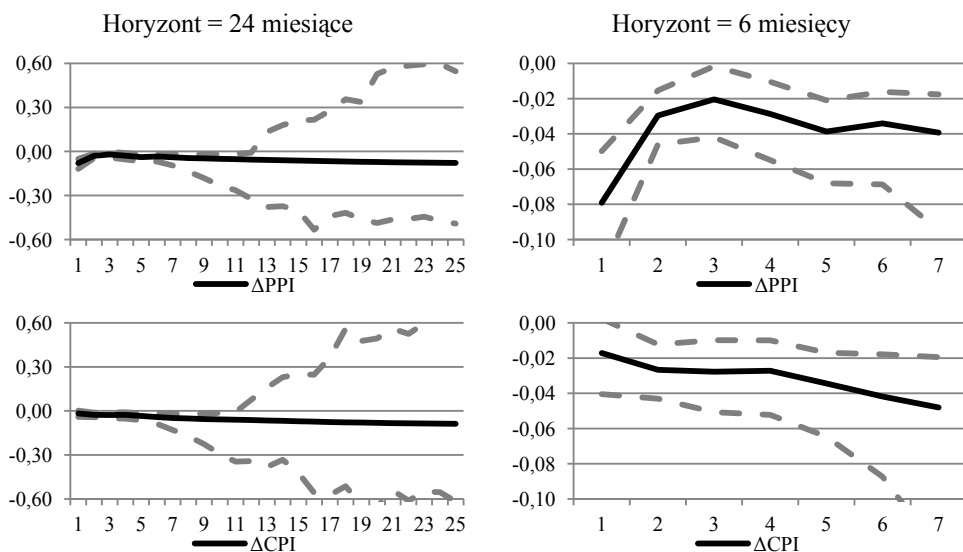


Źródło: opracowanie własne.

Wykresy 19–24. Współczynniki *pass-through* kursu walutowego na ceny (wariant II)

Podsumowując szacunki stopnia przekładania się zmian kursu walutowego na ceny należy stwierdzić, że malejące wartości wskaźników *pass-through* na kolejnych ogniwach łańcucha cenowego potwierdzają teoretyczne założenia o częściowej absorpcji wahań kursowych przez podmioty gospodarcze występujące na kolejnych etapach produkcji bądź dystrybucji towarów. Z drugiej strony zwiększające się wraz ze wzrostem horyzontu prognozy wartości wskaźników pokazują, że efekt *pass-through* charakteryzuje się kilkumiesięcznym opóźnieniem.

Z punktu widzenia polityki pieniężnej istotny wydaje się nie tyle wpływ zmian kursowych na ceny, co na stopę inflację. W związku z tym przeprowadzono analiza zjawiska *pass-through* w odniesieniu do dynamiki cen. Wpływ zmian kursu na stopę inflacji PPI i CPI można w sposób bezpośredni odczytać w wariancie drugim modelu, gdzie explicite umieszczone zostały pierwsze przyrosty logarytmów PPI i CPI. Na wykresach 25–28 przedstawiono wartości wskaźników *pass-through* na miesięczną stopę inflacji producenta i konsumenta. Ich wielkość jest zbliżona za wyjątkiem pierwszego okresu, kiedy to zaobserwować można silniejszą reakcję inflacji producenta. Ogólnie wpływ zmian kursu na inflację zmienia się w tempie malejącym, w półrocznej perspektywie oscylując wokół wielkości $-0,04$.



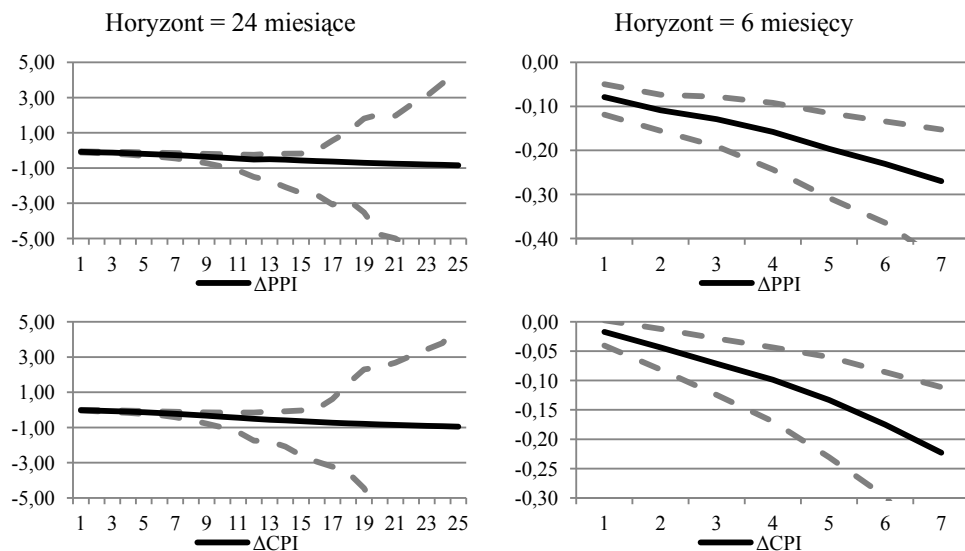
Źródło: opracowanie własne.

Wykresy 25–28. Współczynniki *pass-through* kursu walutowego na inflację PPI i CPI m/m

Praktycznie większe znaczenie ma analiza wpływu zmian kursowych na stopę inflacji w ujęciu rocznym. W tym celu można obliczyć wskaźnik *pass-through* posługując się krocząca sumą przyrostów logarytmów poziomów cen z okresu 12 miesięcy, co odpowiada rocznej stopie inflacji (Cholewiński [2008]). Jednak ze względu na okoliczność, że wyniki oszacowań wskaźników *pass-through* dla horyzontu wybiegającego ponad jedno półrocze okazały się obarczone dużą miarą niepewności, wpływ zmian kursu na roczną stopę inflacji sprowadzi się skumulowania odpowiednich wskaźników dla pierwszych 6 miesięcy.

Wyniki obliczeń przedstawiono na wykresach 29–32. Oba wskaźniki zmieniają się w sposób do siebie zbliżony, niemniej nieznacznie silniejszą reakcją cha-

rakteryzuje się inflacja producenta. Z okresu na okres wpływ zmian kursowych na inflację jest coraz wyższy, po pierwszym kwartale dla inflacji PPI wynosi $-0,16$ z przedziałem ufności $(-0,24;-0,09)$, a po upływie kolejnego kwartału wzrasta do $-0,27$ z przedziałem ufności $(-0,45;-0,15)$. Dla analogicznych okresów wartości wskaźnika dla inflacji CPI wynoszą odpowiednio $-0,10$ z przedziałem ufności $(-0,17;-0,04)$ oraz $-0,22$ z przedziałem ufności $(-0,41;-0,11)$. Tym samym można mówić o znacznym opóźnieniu zjawiska *pass-through*.



Źródło: opracowanie własne.

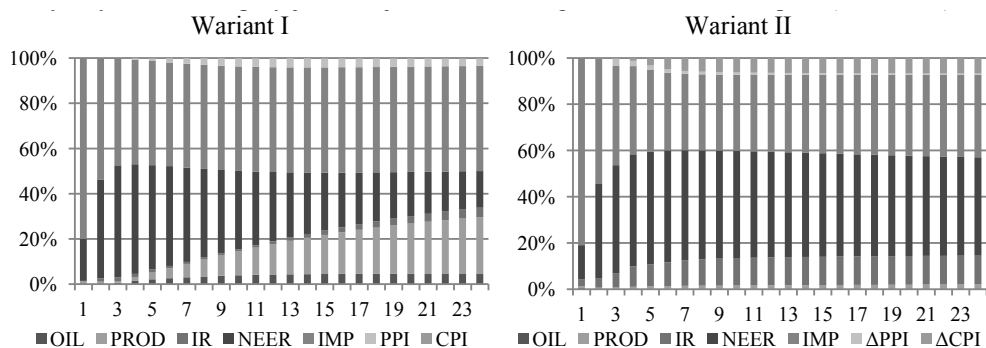
Wykresy 29–32. Współczynniki *pass-through* kursu walutowego na inflację PPI i CPI r/r

Dekompozycja wariancji

Ostatnim etapem analizy przenoszenia zmian kursu walutowego na ceny i inflację jest dekompozycja wariancji składnika losowego. O ile za pomocą funkcji reakcji można zbadać siłę i rozmiar przekładania się zmian kursowych na odpowiednie agregaty cenowe, to dzięki dekompozycji wariancji można odpowiedzieć na pytanie w jakim stopniu wariancja składnika resztowego na każdym etapie łańcucha cenowego jest pochodną odpowiednich szoków. Analizę przeprowadzono dla dwuletniego horyzontu prognozy, niemniej należy pamiętać, że wnioski na temat okresu dłuższego niż 6 miesięcy należy traktować ostrożnie.

Na wykresach 33–34 zobrazowano dekompozycję wariancji dla równań indeksów cen importu w obu wariantach modelu. Zmienność błędu prognozy

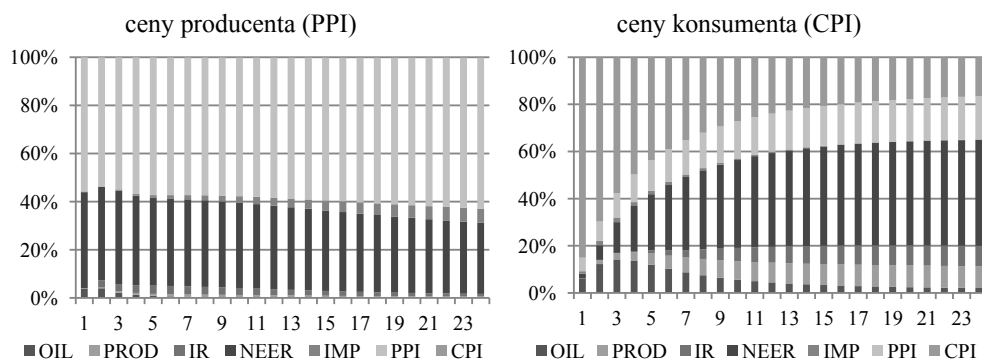
cen importu w wariancie pierwszym w krótkim okresie w prawie połowie wyjaśniana jest przez szoki kursu walutowego, po czym wraz ze wzrostem horyzontu prognozy zależność ta spada do około 16% po dwóch latach. Rośnie natomiast wpływ szoku popytowego, utożsamianego ze zmianami produkcji przemysłowej (do ok. 25% po dwóch latach). Zmienność błędu prognozy w wariancie drugim różni się dość istotnie. Wpływ szoku kursowego również wynosi w pierwszych miesiącach około 50% natomiast maleje w tempie zdecydowanie wolniejszym i po dwóch latach nadal wynosi ponad 40%. Po kilku miesiącach uwidacznia się także dość znaczący (ponad 12%) wpływ zmienności stopy procentowej, co może sugerować reakcję władz monetarnych na zagrożenie inflacją kosztową.



Źródło: opracowanie własne.

Wykresy 33–34. Dekompozycja wariancji składnika losowego w równaniach cen importu (wariant I i II)

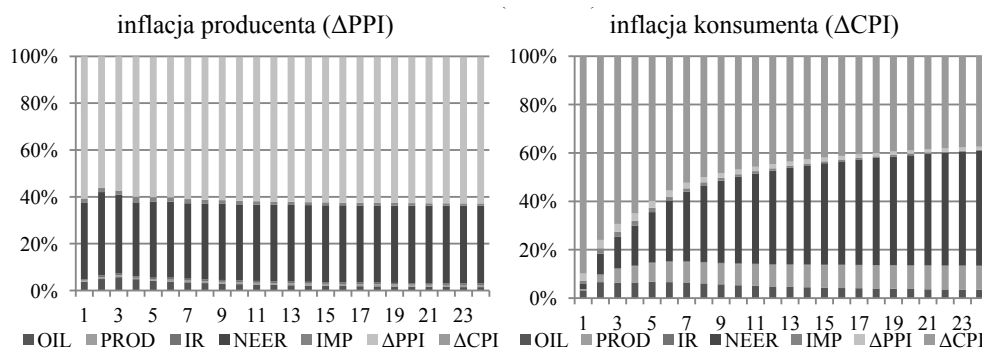
W przypadku cen producenta w wariancie pierwszym, w głównej mierze ich zmienność wyjaśniana jest sama przez się. Wpływ szoku kursowego na początku wynosi około 40% po czym spada, jednak w tempie wolniejszym niż w przypadku poprzedniego agregatu cenowego. Należy też zwrócić uwagę, że z okresu na okres zmienność cen PPI w coraz większym stopniu wyjaśniana jest zmiennością cen importu, co sugerować może odłożony wpływ zmian cen importu zaopatrzeniowego na ceny produkcji. Dla równania cen konsumenta w wariancie pierwszym udział szoku kursowego w wyjaśnianiu zmienności CPI jest początkowo znikomy, po czym rośnie w tempie malejącym do około 45% po dwóch latach, co potwierdza opóźniony efekt *pass-through*. Około 15–20% wariancji wyjaśniana jest zmiennością równania cen producenta. Na tym etapie względnie istotny okazuje również wpływ szoków monetarnych, co może być efektem tego, że głównym przedmiotem zainteresowania władz monetarnych jest zmienność cen koszyka towarów i usług konsumpcyjnych. Opisowaną sytuację zobrazowano na wykresach 35–36.



Źródło: opracowanie własne.

Wykresy 35–36. Dekompozycja wariancji składnika losowego w równaniach cen producenta i cen konsumenta (wariant I)

Wyniki dla wariantu drugiego przedstawiają wykresy 37–38. Wpływ zaburzeń kursowych na zmienność inflacji PPI jest praktycznie stała w całym okresie prognozy i wynosi około 33%. Reszta zmienności tego równania wyjaśniana jest głównie sama przez się. Jedynie w krótkim okresie zaobserwować można silniejszy (ok. 5%) wpływ szoków podaźowych, szybko wygasający wraz ze wzrostem horyzontu prognozy.



Źródło: opracowanie własne.

Wykresy 37–38 Dekompozycja wariancji składnika losowego w równaniach inflacji producenta i konsumenta (wariant II)

W przypadku równania inflacji konsumenta — podobnie jak w odniesieniu do poziomów cen — wpływ zmienności kursu walutowego jest opóźniony, lecz systematycznie z okresu na okres wzrasta. W pierwszym roku dość znaczny wydaje się również wpływ szoków podaźowych (ok. 5–7%), co odzwierciedla

wpływ zmian cen surowców energetycznych na inflację CPI (choć uwzględnione zostały jedynie ceny ropy naftowej, to zmiany cen pozostałych surowców zazwyczaj przebiegają w tym samym kierunku, co najwyżej z niewielkim opóźnieniem czasowym). W wariancie drugim zaskakująco niski okazuje się wpływ zmienności stopy procentowej, nie przekraczający 1%. Można na tej podstawie wnioskować, że inflacja nie jest wrażliwa na zmiany polityki pieniężnej, co przeczy intuicji ekonomicznej. Model sugeruje natomiast względnie znaczący (ok. 10%) wpływ szoków popytowych na zmienność inflacji CPI, co skłania do stwierdzenia, że istotny wpływ na wariację inflacji mogą mieć zmiany polityki fiskalnej.

6. PODSUMOWANIE

Niniejsza praca miała na celu przedstawienie wyników analizy wpływu zmian kursu walutowego na kształtowanie się cen i inflacji w Polsce w okresie od początku 1998 r., tj. od momentu prowadzenia przez Radę Polityki Pieniężnej strategii bezpośredniego celu inflacyjnego.

Elementem badania było wyspecyfikowanie dwóch wariantów modelu SVECM, z których każdy zawierał dwa wektory kointegrujące. Następnie zaprezentowane zostały wartości funkcji reakcji indeksów cenowych na impuls kursu walutowego, które posłużyły do obliczenia współczynników *pass-through*. Dla wymienionych funkcji parametrów modelu wyznaczono bootstrapowe przedziały ufności, pozwalające w sposób przybliżony wnioskować o skali niepewności oszacowań opisywanych wskaźników.

Oba warianty, zgodnie z oczekiwaniami, sugerują silniejszy i szybszy stopień przełożenia zmian kursowych na ceny importu, niż na ceny produkcji i konsumpcji. Korzystając z wariantu drugiego modelu, analizie poddano również reakcję dynamiki cen produkcji i konsumpcji. Inflacja zarówno w ujęciu miesięcznym, jak i rocznym, reagowała na zmiany kursu walutowego w sposób zbliżony. W krótkim okresie silniejszą reakcją charakteryzowała się inflacja PPI.

Zależności te zaobserwowano w rozpatrywanej perspektywie 6 miesięcy, do której ograniczono się w próbie ekonomicznej interpretacji opisywanych zagadnień. Wnioskowanie w dłuższym horyzoncie okazało się obarczone olbrzymią niepewnością i w kontekście przeprowadzonych analiz wydaje się nieuprawnione. Taki stan rzeczy poddaje w wątpliwość wyniki badań opierające się jedynie na ocenach punktowych parametrów (bądź ich funkcji) i staje się sugestią do uwzględniania w analizach także miar niepewności estymowanych parametrów.

Dekompozycja wariacji zaprezentowana w ostatniej części pracy dostarcza informacji o determinantach kształtowania się indeksów cenowych w perspektywie czasu. Wyniki analizy w obu wariantach potwierdzają opóźniony efekt *pass-through*.

Podsumowując całość wyników badania należy stwierdzić, że w półrocznej perspektywie ceny w gospodarce polskiej nie dostosowują się w pełni do zmian kursu walutowego. Kierunek zależności w krótkim okresie jest zgodny z intuicją ekonomiczną, natomiast wykorzystywane narzędzia nie dają podstaw do wnioskowania o zależnościach między kursem a cenami w długim okresie. Należy dodatkowo pamiętać, że wybór i specyfikacja modelu obarczona była dużą dozą arbitralności, co może mieć istotny wpływ na wyniki płynące z badania. Należy więc poszukiwać precyzyjniejszych metod ekonometrycznych pozwalających przewyższać tradycyjne ograniczenia makromodelowania (takie, jak krótkie szeregi czasowe) czy też zaobserwowaną w niniejszym badaniu niestabilność wariancji wielu z nich. Pomocne w przewyżczeniu tych ograniczeń może być wykorzystanie modeli klasy VAR-SV czy też wspomniane wcześniej przejście na grunt ekonometrii bayesowskiej.

Sposób modelowania powinien również odzwierciedlać możliwy asymetryczny charakter zjawiska *pass-through*. W ostatnich latach badania empiryczne uwzględniające to zagadnienie dla małych gospodarek otwartych prowadzili m.in. Alexius i Post [2008], a w Polsce Przystupa i Wróbel [2009]. Pogłębianie analiz mechanizmu transmisji polityki pieniężnej w świetle istnienia kanału kursowego wydaje się kluczowe dla prowadzenia skutecznej polityki monetarnej, szczególnie wobec możliwego ograniczenia swobody prowadzenia tej polityki koniecznością utrzymania niskiej zmienności kursu Złotego w stosunku do Euro, w ramach systemu ERM2. Autor pozostawia te zagadnienia przyszłym badaniom.

BIBLIOGRAFIA

- Alexius A., Post E. 2008. *Exchange rates and asymmetric shocks in small open economies*, Empirical Economics, Springer, vol. 35(3).
- Anderton B. 2003. *Extra-Euro Area Manufacturing Import Prices and Exchange Rate Pass-Through*, ECB Working Paper, nr 219.
- Bernanke B., Laubach T., Mishkin F., Posen A. 1999. *Inflation Targeting: Lessons from the International Experience*, Princeton University Press, Princeton.
- Blanchard O. 1982. *Price Desynchronisation and Price Level Inertia*, NBER Working Paper, nr 900.
- Bussier E. M. 2007. *Exchange rate pass-through to trade prices. The role of non-linearities*, ECB Working Paper, nr 822.
- Ca' Zorzi M., Hahn E., Sanchez M. 2007. *Exchange rate pass-through in emerging markets*, ECB Working Paper, nr 739.
- Campa J.M., Goldberg L. 2002. *Exchange Rate Pass-Through into Import Prices: A Macro and Micro Phenomenon?*, NBER Working Paper, nr 8934.
- Charemza W., Deadman D. 1997. *Nowa Ekonometria*, PWE, Warszawa.
- Cholewiński R. 2008. *Wpływ zmian kursu walutowego na dynamikę procesów inflacyjnych*, Materiały i Studia nr 226, NBP.
- Choudhri E., Hakura D. 2006. *Exchange rate pass-through to domestic prices: Does the inflationary environment matter?*, Journal of International Money and Finance, Vol. 25(4).

- Devereux M., Engel Ch. 2003. *Monetary Policy in the Open Economy Revisited: Exchange Rate Flexibility and Price Setting Behavior*, Review of Economic Studies, Vol. 70(4).
- Dornbusch R. 1987. *Exchange Rate and Prices*, American Economic Review, Vol. 77.
- Efron B., Tibshirani R.J., *An introduction to the Bootstrap*, Chapman & Hall, New York.
- Favero C. 2001. *Applied Macroeconomics*, Oxford University Press, Oxford.
- Fukuda S., Ono M. 2006. *On the Determinants of Exporters' Currency Pricing: History vs. Expectations*, NBER Working Paper, nr 12432.
- Ghosh A., Rajan R. 2009. *Exchange rate pass-through*, [w:] The Princeton Encyclopedia of the World Economy, Princeton University Press, Princeton.
- Grabek G., Kłos B., Kokoszcyński R., Łyziak T., Przystupa J., Wróbel E. 2002. *Porównanie podstawowych cech mechanizmu transmisji monetarnej w Polsce i strefie euro*, [w:] Raport na temat pełnego uczestnictwa Rzeczypospolitej Polskiej w trzecim etapie Unii Gospodarczej i Walutowej, NBP.
- Hahn E. 2003. *Pass-through of external shocks to euro area inflation*, ECB Working Paper, nr 243.
- Hall R., Taylor J. 2007. *Makroekonomia*, Wydawnictwo Naukowe PWN, Warszawa.
- Ihrig J., Kamin S., Lindner D., Marquez J. 2007. *Some Simple Tests of the Globalization and Inflation Hypothesis*, International Finance Discussion Papers nr 891, Board of Governors of the Federal Reserve System, Washington.
- Ito T., Sasaki Y., Sato K. 2005. *Pass-Through of Exchange Rate Changes and Macroeconomic Shocks to Domestic Inflation in East Asian Countries*, Discussion Paper nr 05-E-020, Research Institute of Economy, Trade and Industry.
- Johansen S. 1988. *Statistical Analysis of Cointegration Vectors*, Journal of Economic Dynamics and Control, Vol. 12.
- Każmierczak A. 2008. *Polityka pieniężna w gospodarce otwartej*, Wydawnictwo Naukowe PWN, Warszawa.
- Kokoszcyński R. 2004. *Współczesna polityka pieniężna w Polsce*, PWE, Warszawa.
- Kokoszcyński R., Łyziak T., Pawłowska M., Przystupa J., Wróbel E. 2002. *Mechanizm transmisji polityki pieniężnej — współczesne ramy teoretyczne, nowe wyniki empiryczne dla Polski*, Materiały i Studia nr 151, NBP.
- Landau B., Skudelny F. 2009. *Pass-through of external shocks along the pricing chain. A panel estimation for the euro area*, ECB Working Paper, nr 1104 .
- Lütkepohl H. 2005. *New Introduction to Multiple Time Series Analysis*, Springer, Berlin.
- Majsterek M. 2003. *Zmienne zintegrowane w stopniu drugim w modelowaniu ekonometrycznym*, Przegląd Statystyczny, Vol.50(2).
- Majsterek M. 2008. *Wielowymiarowa analiza kointegracyjna w ekonomii*, Rozprawy habilitacyjne Uniwersytetu Łódzkiego, Wydawnictwo UŁ, Łódź.
- Malczyk K. 2010. *Analiza przenoszenia zmian kursu walutowego na ceny w Polsce*, maszynopis pracy magisterskiej napisanej pod kierunkiem prof. dra hab. Jacka Osiewalskiego, Kraków.
- McCarthy J. 1999. *Pass-Through of Exchange Rates and Import Prices to Domestic Inflation in Some Industrialised Economies*, BIS Working Paper, nr 79.
- Mishkin F.S. 2008. *Exchange rate pass-through and monetary policy*, Conference on Monetary Policy "Jarle Berge Colloquium: Globalisation and Monetary Policy", Oslo, 7.03.2008.
- Miształ P. 2008. *Niepełne przenoszenie zmian kursu waluty na ceny w handlu zagranicznym*, Ekonomista, nr 5/08.
- Miształ P. 2009. *Zmiany kursu walutowego a dynamika cen w Polsce*, Ekonomista, nr 4/09.
- Osiewalski J. 2001. *Ekonometria Bayesowska w zastosowaniach*, Wydawnictwo Akademii Ekonomicznej w Krakowie, Kraków.
- Pollard P., Coughlin C. 2004. *Size Matters: Asymmetric Exchange Rate Pass-Through at the Industry Level*, Federal Reserve Bank of St. Louis Working Paper.

- Przystupa J., Wróbel E. 2009. *Asymmetry of the exchange rate pass-through: An exercise on the Polish data*, MPRA Paper 17660, University Library of Munich.
- Sims C. 1980. *Macroeconomics and Reality*, *Econometrica*, Vol. 48(1).
- Sławiński A. 1999. *Ewolucja mechanizmu kursowego w Polsce*, *Gospodarka Narodowa*, nr 7– 8/99.
- Sławiński A. 2007. *Znaczenie czynników ryzyka towarzyszących wchodzeniu Polski do ERM2 i do strefy euro*, Referat na Kongres Polskich Ekonomistów, Warszawa 29–30 listopada 2007r.
- Syczewska E.M. 2007. *Ekonometryczne modele kursów walutowych*, Monografie i opracowania nr 547, Oficyna Wydawnicza SGH, Warszawa.
- Taylor J. 2000. *Low Inflation, Pass-Through, and Pricing Power of Firms*, *European Economic Review*, Vol. 44(7).
- Wróblewska J. 2009. *Bayesian Model Selection in the Analysis of Cointegration*, *Central European Journal of Economic Modelling and Econometrics*, Vol. 1(1)

THE CONTEMPORARY STANDARDS OF NATIONAL ACCOUNTS – APPLICABILITY AND LIMITATIONS IN ECONOMIC GROWTH AND PRODUCTIVITY STUDIES

KAMIL MAKIEŁA

Kielce University of Technology,
e-mail: kamilmakiela@gmail.com

Praca została przedstawiona przez autora na posiedzeniu Komisji Nauk Ekonomicznych i Statystyki Oddziału PAN w Krakowie 24 października 2011 roku.

ABSTRACT

The purpose of this article is to survey the contemporary standards of modern national accounts, and to assess their applicability in tracing differences in economic growth across countries. In order to perform a growth accounting study one requires good quality and mutually comparable information about the three main macroeconomic indicators: i) production output, ii) capital input and iii) labour input. Thus, the author outlines the sources and reasons behind the creation of such statistics as well as the national accounts standards they comply with. Each of the above indicators is discussed in respect of these standards, limitations in applying and availability in international databases. The study also provides insights into the current stage of development of the System of National Accounts and how different measurement standards can augment inference on economic growth.

ABSTRAKT

Celem niniejszej pracy jest analiza współczesnych standardów dla rachunków księgowych budżetów państw oraz ich ocena pod kątem przydatności w pomiarze różnic we wzroście gospodarczym pomiędzy krajami. Badania typu *growth accounting* wymagają wysokiej jakości, porównywalnych między sobą informacji o trzech głównych wskaźnikach makroekonomicznych: i) wartości produkcji, ii) nakładu kapitału rzeczowego oraz iii) nakładu pracy. Dlatego też, autor niniejszej pracy zarysowuje źródła i przyczyny stojące za powstaniem takich danych oraz przedstawia standardy, którym podlegają. Każdy z wyżej wymienionych wskaźników omawiany jest po kątem tych standardów, ograniczeń w jego zastosowaniu oraz dostępności w międzynarodowych bazach danych. Artykuł zarysowuje również obecny poziom rozwoju *System of National Accounts* oraz pokazuje, jak różne standardy pomiaru mogą wpływać na wnioskowanie o wzroście gospodarczym.

KEYWORDS

macroeconomic data, system of national accounts, measurement standards, growth accounting

1. INTRODUCTION

Evaluating countries' productive capacities becomes a crucial element in today's globalized world. Benchmarking economies allows us to evaluate their industrial potential, stage of development and rate of economic growth. Growth accounting in particular has become a preferred framework for such research [Chen *et al.* 2010]. Though it does not, by itself, explain the underlying causes of each factor's contribution to the production output it can serve as a powerful policy review tool when complemented by historical and case study analyses [Schreyer 2004]. Such a comprehensive approach brings the essence of quantitative and qualitative research together, and allows us to fully understand the reasons of growth, innovation and productivity change. In doing so, growth accounting methods bring a fair share of knowledge that comes from the source of economic growth observation – meaning the data. Since every quantitative analysis of economic growth is dependent on the data accuracy and cross-country comparability, their methodologies and proper usage become an ongoing concern. It poses a considerable research challenge in the early stages of virtually every growth accounting study.

Because crunching the numbers can be so time consuming and complex, many economists spend little time considering the pedigree of the pre-crunched data. Compiling countries' productivity statistics is not only complex but also involves as much politics as science. What is more, once such data are made available throughout a set of countries, their comparability and across-nation applicability is still questionable. This issue was raised decades ago by international organizations¹ in the post-war era. Back then, such data were crucial to form the bases of policy recommendations and guidelines to efficiently allocate scarce resources needed to rebuild Europe after two world wars [Ward 2004].

At first, due to the world's division between American and Soviet spheres of influence, the earliest versions of international systems of accounts ("western" System of National Accounts and "eastern" Material Product System) were not entirely applicable or even comparable [Ward 2004]. However much changed after the USSR collapse and for some time now, we have been witnessing a gradual convergence towards mutual comparability of main macroeconomic indicators, that is, the production output and capital and labour inputs. The purpose of this article is to survey the contemporary state of knowledge about these international standards, discuss the outstanding issues, outline databases suitable for use in growth accounting studies and show the implications of using different measurements.

Section 2 of this work outlines principles of growth accounting where macroeconomic indicators are used to trace sources of economic growth. Section 3 focu-

¹ The League of Nations, precursor to the UN and OEEC later to be renamed as OECD .

ses on standards for macroeconomic accounts and their origins. These concepts set the methodological framework for compiling the main national statistics on which we focus in section 4. Section 5 provides insights into tools and methods necessary for bringing special-temporal observations to comparability. Finally, section 6 demonstrates what implications measurement standards may have on economic growth inference and section 7 concludes with a discussion.

2. PRINCIPLES OF GROWTH ACCOUNTING

Growth accounting procedures are largely based on macroeconomic production theory and their purpose is to trace each factor's contribution to economic growth. The underlying assumption is that a change in macroeconomic output, given as:

$$Y_{ti} = f(K_{ti}, L_{ti}; B_t) \cdot EF_{ti} \quad (1)$$

where Y_{ti} is the macroeconomic production output, K_{ti} is capital input and L_{ti} denotes labour input, is the result of a change in the i) quantity of inputs and ii) the way they are used in production. The latter is broadly referred to as the change in productivity and there are two ways to consider it. First, when the production technology is progressing (or regressing) it augments parameters (B_t) of the function that describes it. This way, more (or less) product can be made given the same quantity of inputs. Second, productivity may shift as the result of change in a country's technical efficiency (EF_{ti}). This may be due to a number of factors, like i) changes in work culture over years, ii) governmental policies, or the recently discussed iii) malicious practices of worldwide financial institutions. In short, the mainstream growth accounting framework can be summarised as²:

$$OC_{t+1,i} = IC_{t+1,i} \times TC_{t+1,i} \times EC_{t+1,i} \quad (2)$$

where IC is input change, TC is technical change, EC is efficiency change, i is country index and $t + 1$ denotes a change from t to $t + 1$ period. Suffice to say that increase in any of the three factors results in economic growth.

There have been several methodologies suggested to implement the growth accounting framework. In recent years, stochastic frontier analysis (SFA), independently developed by Aigner, Lovell, Schmidt [1977] and Meeusen and Van den Broeck [1977], seems to have become a preferred parametric approach (see,

² There have been many conceptual frameworks in this field. Very few, however, stood the test of time. One alternative idea recently mentioned in the literature, and introduced by Caselli and Coleman [2006], assumes that each country has its own unique technology and the aim of growth accounting is to trace differences among these different technologies. Considering the ongoing globalisation, however, this idea has already earned some critique.

e.g., Fried, Lovell and Schmidt [2008] for a lengthy list of applications in macroeconomics). A typical SFA model is denoted as:

$$Y_{ti} = f(K_{ti}, L_{ti}; B_t) \cdot \exp(v_{ti} - u_{ti}) \quad (3)$$

where $f(\cdot)$ is the production frontier³, B_t is a vector of technology parameters (in period t), v_{ti} reflects stochastic nature of the frontier (symmetric disturbance) and " $-u_{ti}$ " is the inefficiency term of country i in period t . Inefficiency is measured as the distance between the observed output and the world production frontier⁴. Having this, we can compute country i 's efficiency as:

$$EF_{ti} = \frac{\exp(y_{ti})}{\exp(f(k_{ti}, l_{ti}; B_t) + v_{ti})} = \exp(-u_{ti}) \quad (4)$$

where lower case letters (y, k, l) indicate natural logs of upper case letters (Y, K, L).

The above model can be easily re-arranged to accommodate growth accounting framework. Given any two corresponding periods t and $t + 1$, if we consider world frontiers as well as country i 's inputs and inefficiencies, the expected increase in the log of its macro-output is [Koop, Osiewalski and Steel 1999]:

$$\frac{1}{2}(x_{t+1,i} + x_{ti})'(\beta_{t+1} - \beta_t) + \frac{1}{2}(\beta_{t+1} + \beta_t)'(x_{t+1,i} - x_{ti}) + (u_{ti} - u_{t+1,i}) \quad (5)$$

where the first component captures technical progress (or regress), the second reflects input change and the third accounts for shifts in efficiency over time. Thus, if we define these components as input change: $IC_{t+1,i} = \exp(\frac{1}{2}(\beta_{t+1} + \beta_t)'(x_{t+1,i} - x_{ti}))$; technical change: $TC_{t+1,i} = \exp(\frac{1}{2}(x_{t+1,i} + x_{ti})'(\beta_{t+1} - \beta_t))$ and efficiency change: $EC_{t+1,i} = \exp(u_{ti} - u_{t+1,i})$, we can rewrite (5) as:

$$OC_{t+1,i} = IC_{t+1,i} \times TC_{t+1,i} \times EC_{t+1,i}$$

which is exactly the same formula as in (2) that we want to analyze.

³ The function here is assumed to be linear with respect to natural logs of Y, K and L

⁴ Meaning: the potentially obtainable output given inputs under the current level of technology. Also, here I assume that there is a common frontier for all investigated countries. See Makiela [2009] for a discussion.

3. SOURCES OF MODERN NATIONAL ACCOUNTS

Quite obviously, countries' economic and financial structures vary significantly. Therefore, when compiling datasets for a growth accounting study, it is important to use data that maintain comparability. Issues on calculating either a given country's output or the level of its capital stock are still left fairly open, so when analyzing growth one should remember that research conclusions also depend on calculation methodologies used in a given dataset. The data should be collected from databases that provide international comparability instead of directly from National Statistical Offices (NSOs hereafter). Slowly but steadily over the last 50 years, economists and statisticians have been unifying the national accounts methodology. Their efforts have been much appreciated and it is no accident that the two major creators of modern national accounts have been both awarded Nobel prizes for Economics – Simon Kuznets (USA) in 1971 and Richard Stone (UK) in 1984.

The System of National Accounts (SNA hereafter) had its origin in the policy monitoring and evaluation tools used during the rebuilding of postwar Europe. The SNA can be traced back to 1947 when, at its first meeting, the United Nations Statistical Commission (UNSC), chaired by Richard Stone, expressed the need to develop international statistical standards that would enable policy monitoring. This was especially crucial for the postwar Western Europe as, in order for the Marshal Plan to succeed, scarce resources had to be properly managed and efficiently allocated.

The first SNA was introduced in 1953 and adopted mainly by western economies. Though consisting of only six main tables, it enabled the basic policy reviews of the postwar reconstruction efforts [Bos 2008]. Several revisions to the 1953 version were issued (in 1960 and 1964), but it wasn't until 1968 when the first milestone achievement in unifying national accounts was made. The 1968 SNA comprised a set of balance sheets, input-output tables and, due to inter-industry sectoring, allowed policymakers and researchers to conduct more extensive macroeconomic analyses. Moreover, efforts have been made for 1968 SNA to be compatible with the Material Product System (MPS), a methodology which had been concurrently developed by the USSR and its satellite economies⁵.

The new standard, however, was not adopted as widely as its creators had anticipated. The Western economies did not have the slightest problem converting to it, as it directly responded to their policy and planning evaluation needs. Because the 1968 SNA was tailor-made for the West, NSOs and analysts encountered difficulties adopting it for non-Western economies. Furthermore, advancing complexity of financial and economic systems as well as technological progress quickly made it clear that SNA needs further development.

⁵ Countries like Cuba and North Korea still produce accounts in line with MPS methodology

The spread of intangible instruments such as intellectual property, electronic transfers and financial services were the main reasons of the 1993 revision. 1993 SNA was released jointly under the auspices of the United Nation (UN), the Organization for Economic Co-operation and Development (OECD), the European Commission (EC), International Monetary Fund (IMF) and the World Bank (WB). It was the most comprehensive issue of SNA compiled to date, revising the national accounts framework and bringing them up-to-date. Based on internationally agreed classifications, concepts and definitions, macroeconomic data could be gathered and presented in a format that is suitable for international comparative analysis. The 2008 revision of the System of National Accounts addressed issues left open in previous updates and provided advances in methodologies like the concept of *capital services* and *labour services*.

As far as international standards are concerned there is one more that should be mentioned. In 1995 the European Commission introduced the European System of National Accounts (ESA95) which was consistent with 1993 and 2008 SNA releases but provided strict guidelines to some issues that were deliberately left open in SNA. These, however, were necessary because national accounts in the EU are used by the European Commission to distribute development funds, calculate Members' contributions to the EU budget and, more recently, to monitor the sustainability of Members' public finances. Since ESA95 is part of the European Union legislation system, the international comparability of national accounts is a legal requirement for all Member States and for EU candidates. The SNA standard is designed to be flexible in order to be applicable for countries with different economic systems and at various stages of economic development. ESA95 is therefore more effective than SNA in ensuring international comparability. However, unlike SNA, not every country can adhere to its standards. Currently ESA is undergoing a five year revision plan, which is scheduled to conclude in 2012 [Gueye 2007].

4. MAIN PRODUCTIVITY INDICATORS IN SNA

Gross Domestic Product

The most frequently used macroeconomic production output indicator in the System of National Accounts is Gross Domestic Product (GDP). This single figure combines the production of all the companies, government bodies and non-profit institutions in a given country during a certain period. GDP is usually calculated annually, but in some countries also quarterly or even monthly. When aggregated from the microeconomic to the macroeconomic level, it follows three essential rules of SNA [OECD 2003]: i) avoid double counting, ii) relate to aggregates that are economically significant (i.e., which value is independent of non-economic factors) and iii) create indicators that are measurable in practice.

Apart from GDP, SNA defines one more output indicator – the Net Domestic Product, called NDP in short. It is used to assess the genuine level of newly created wealth during a given production period. Thus, subtraction from GDP must be made to account for the costs of using up capital assets. In 2008 SNA and ESA95 this is done through a figure called “consumption of fixed capital”. When this consumption is deducted, the result is Net Value Added, and the NDP is equal to all net values added summed across industries: $NDP = \Sigma \text{Net Values Added}$. Although less widely used than GDP, NDP in theory is a better measure of the wealth produced as it deducts the costs of machinery wear-off and other capital assets used-up in the production process. However, economists tend to prefer GDP for two reasons. First, methods and techniques for calculating consumption of fixed capital are rather complex and tend to differ between countries, making NDP comparability uncertain. Second of all, when ranking countries or analyzing their growth, the differences between GDP and NDP are small and do not change the conclusions.

When considering GDP (or NDP) as a production output, we should bear in mind that it does not account for i) home produced durables, ii) volunteer work, iii) wealth earned before, and most notably iv) makes no account of the “grey area” which may vary significantly across countries as well as in time. So, in principle GDP (or NDP) should be regarded more as a proxy rather than a good measure of countries’ production output (or welfare). It reflects output with no regard to its inputs that are used or even depleted in production. In some countries people may be working longer hours to maintain a comparable life standard, while in others they may be running down country’s natural resources for the same purpose. Furthermore, the fact that GDP does not reflect various kinds of economic activity, such as home production, may make a difference when GDPs of two nations are compared. If the first one is caring for its young and elder “for free” at home, while the other does it through market-based services, the latter will register higher GDP level. This does not mean, however, that the latter is actually better off [Gylfason 1999].

GDP can be calculated using *income*, *expenditure* or *output* approach (see, e.g., Chamberlin and Yeuh [2006]), but all the methods arrive at the same value only in theory. In practice the resulting estimates differ, since they are subjects to errors and omissions during aggregation process. The most significant discrepancy is between GDP acquired through *output* approach also known as *value added* approach. GDP should be obtained when intermediate consumption for total economy is deducted from its Gross Output. In practice, however, the Gross Value Added (GVA) calculated in this way does not equal Gross Domestic Product. To arrive at GDP level one needs to add the income from taxed goods and services, and deduct subsidies for them. From the output perspective, however, GDP is supposed to be a proxy for the total production output in a given economy. Hence, some countries, like the United States of America, define GDP from *output*

approach as GVA leaving discrepancies between the three methods behind. Also, differences between GVA and GDP are small and the most important thing is to use the same indicator for the whole dataset.

Being the most recognized macroeconomic output indicator, GDP is published by all significant international statistical institutions. United Nations Statistics Division and the World Bank provide the most comprehensive datasets of Gross Domestic Product. The statistics are gathered either directly from NSOs or, more often, through other international organizations such as OECD or Eurostat. Although they contain most numerous GDP dataset, variety of sources may make international comparability questionable.

IMF and OECD also provide estimates for GDP in IMF's *World Economic Outlook* and OECD's *Economic Outlook* databases. The two databases are called similarly not by a coincidence. By using the same data sources they usually provide the same estimates. However, when choosing between the two databases one should know that IMF's online database sometimes publishes rounded estimates directly from OECD's datasets⁶.

Capital input

Measuring capital input at the national level and assuring its international comparability is an ongoing problem for several reasons [OECD 2001]. Firstly, not all NSOs regularly publish data on physical capital stocks, which are the indicators needed to assess the level of capital input in an economy. Even if such data are made available their international comparability is vague. Secondly, there are several types of capital stock measures and each has its analytical applicability [Schreyer and Webb 2006]. Thirdly, we cannot measure capital stock directly. Most estimates mentioned by SNA are estimated by NSOs using available data according to local methodologies, although there is an increasing convergence towards international standards. This is mainly due to OECD's active involvement in recent years. The organization has issued numerous papers and handbooks on how to produce unified capital stock estimates.

Another reason for problems with obtaining capital stock estimates may be due to large data requirements. A given NSO needs to have at least data on i) all assets (by type), ii) investment volumes (by type of asset), iii) price deflators (by type of asset), iv) industry by asset-type investment matrices and v) a benchmark level of capital stock for no less than one year [OECD 2001]. Moreover, some types of capital measures, like *capital services*, require additional information like average service life (by asset) and depreciation rate of each asset type.

The first attempt to produce unified capital stock estimates for international comparisons was made at the *Center for International Comparisons* at the Univer-

⁶ Purchasing Power Parity estimates would be one such an example

sity of Pennsylvania (CICUP). Alan Heston, Robert Summers and Bettina Aten, developed a database called *Penn World Tables*. Version 5.6 contains *Physical Capital Stock per worker* estimates. Unfortunately they are based on an older version of SNA from 1968. The 1993 issue of SNA, however, dealt with several new concepts like i) how to allocate software and other intangible assets to investment (see Ahmad [2003] or Lequiller, Ahmad, Varjonen, Cave and Ahn [2003] for details), or ii) how to use quality-adjusted prices to deflate investment in information and communication technologies (ICT) assets. The new way of constructing national accounts changed significantly the way we now measure capital and proved the former capital estimates to be inconsistent [Schreyer 2007]. So far, the *Center for International Comparisons* has not published an update of their capital stock estimates.

OECD on the other hand, has been very active over the past years in developing new standards and ensuring capital stock comparability across its members. According to OECD, there are two main concepts of capital stock [Schreyer 2003]. The first type of capital stock is defined as a services provider in production. Hence, productivity of each asset is taken into consideration and the concept of *capital services* is introduced (see, e.g., OECD [2003], SNA [2008] or Timmer, O'Mahony and van Ark [2007] for details regarding methodology). In this case, not only the quantity of capital goods involved but also their physical characteristics play a role in assessing the total *capital services* level. Statisticians estimate it by weighting different types of stocks by their relative productivity. Unfortunately *capital services* hadn't been recognized by SNA until its 2008 edition, and thus only a few countries regularly publish data on their *productive stocks*. For now there are only three international databases that provide estimates on *capital services* at an international level. That is OECD's *Productivity database*, EU *KLEMS* project and *The Conference Board Total Economy Database* (only growth rates).

The second concept of capital stock measurement traces its role as an indicator of wealth. The *net* stock, also known as the *wealth* stock, represents the market value of all (fixed) capital goods. It is usually acquired from the *gross* capital stock by accounting for the decline in assets' value before they retire. The purpose here is to track capital's role as a sum of assets with their market values [OECD 2003]. This indicator, however, should be treated with caution in growth accounting or productivity studies, as the actual asset market value may not always reflect its productive potential. Thus, while *net* capital stock is more informative in terms of price value of the capital stock (wealth), *gross* capital stock or, if available, *productive stocks* are preferred measures of the capital's productive potential [OECD 2003].

Both concepts have their disadvantages. On the one hand, it is logical to assume that different types of fixed assets will have different productivity capabilities. Countries with the same capital stock (capital wealth) may produce different output volumes only based on differences in their capital structure. On the other

hand, though *capital services* in theory provide much more exact productivity estimates of a given country's capital stock, they are always delivered in the form of a percentage change to the base year (e.g., 1995=100% in EU KLEMS database) and they cannot be used straightforward in a cross-country productivity analysis. One way to solve this issue would be first to define a given country's *capital services* for the benchmark year at the *gross* (or *net*) capital stock level. That way we take into consideration initial differences in capital input volumes between analysed countries. Then, *capital services* for remaining years can be easily calculated by adding the percentage change for the year of interest to the benchmark year estimate. Although intuitively this is the right course of action to acquire *capital services* at market prices that allow for a cross-sectional comparison, I have not encountered any study or growth accounting handbook that would provide justification to it.

Currently, there are six working repositories of internationally comparable capital stock estimates: four at OECD, one at EU KLEMS database (*capital input files*) and one at the Conference Board (*Total Economy Database*). OECD's *Economic Outlook* and *Productivity Database* contain annual aggregates, the latter measuring them in *productive stocks (capital services)*. OECD's *Structural Analysis (STAN)* and *Annual National Accounts (ANA)* on the other hand provide asset breakdown by industry. Eurostat's *National Accounts Team* is also planning to launch a web-based searchable database for its resources on capital stock. The launch date, however, is yet unknown.

Labour input

Usually, labour input in a given country is measured by the *average number of people employed* in a given year. According to many, however, this is not a good way of measuring economy's labour input, since it i) does not account for differences in work patterns across countries and ii) does not reflect the quality of labour (i.e., the level of human capital; see Gylfason [1999]).

In some countries the average number of hours worked per week by an employee may significantly differ from others, for example, as a result of discrepancies in the number of free days (holidays etc.). Moreover, the *average number of people employed* takes under consideration only those employed in enterprises and therefore leaving behind i) *self-employed workers* and ii) *family workers* [OECD 2009]. In order to account for such discrepancies in work patterns among countries and consider non-employed people who, nonetheless, are engaged in some productive activities, more detailed labour indicators are provided in 2008 SNA. By joint estimation of hours worked by employees and the two work groups mentioned above, NSOs can calculate the *total number of hours worked by persons engaged* in productive activities in the economy. This, however, is a complex figure and at least for now not many NSOs have the potential to produce it regularly.

The only datasets providing such indicators suitable for international comparisons are the ones available at the EU and OECD's databases.

Furthermore, contemporary standards also provide an internationally agreed on framework for considering the level of labour quality broadly referred to as human capital. By distinguishing between different types of labour in terms of gender, age and education attainment, NSOs can calculate a standardized measure of *labour services* (see SNA [2008], Van Ark, O'Mahony and Ypma [2007] or Timmer, O'Mahony and van Ark [2007] for details on methodology). Like *capital services* to *gross/net* capital stock, in theory *labour services* is a better estimate than the *total hours worked by persons engaged* because it grasps differences in labour quality, not only its quantity. Unfortunately also alike *capital services*, this figure is usually given as a percentage change to the base year, and due to its recent introduction into SNA only a handful of countries make such statistics available.

There are many repositories of labour statistics nowadays. Widely recommended and acknowledged statistics are available at OECD (*Employment database*, and for *labour services* OECD's *Productivity database*), European Commission (Eurostat's database, and for *labour services* EU KLEMS project) and the Conference Board (*Total Economy Database*).

5. BRINGING NATIONAL ACCOUNTS TO COMPARISON

When analyzing differences in economic growth across countries we should remember that apart from internal economic phenomenon like inflation, the data need to account for differences in currency values across countries and differences in their purchasing powers. Today, economists distinguish between the international market value of a given currency and its purchasing power. The first one, foreign exchange rate (*forex* rate), specifies how much one currency is worth in terms of the other. Price levels on the domestic markets are not taken into account and the exchange value is solely dependent on the currency attractiveness, which can be subjected to high volatility or speculation.

Shortcomings of this method have led to the creation of indices that base on the concept of "the law of one price", first introduced by Gustav Cassel (1921). According to Cassel exchange rate between two countries needs to be adjusted by their currencies' purchasing powers on their domestic markets so that a purchase in one currency would be equivalent to the other. Thus, Purchasing Power Parity (PPP) indices are a crucial element of data preparation for growth accounting studies as they bring the unifying element. The most acknowledged PPP indices convert countries currencies to a so-called "international dollar". They are compiled jointly by Eurostat and OECD and can be viewed in many databases, IMF and UNstats including. The methodology is based on "the basket of goods" concept, which is a complex and time-consuming study [Eurostat-OECD 2006].

Using PPP in a cross-section analysis is rather straightforward. However, when dealing with spatial and temporal observations we have two options to consider. One way is to use the *current international prices* base and apply benchmark PPPs from every year allowing the price structure to vary over time. Within the same year volumes are measured by the same price structure and are directly comparable. Comparison over time, however, carries effects of i) a relative change in volume and ii) changes in relative prices between countries [Schreyer and Koechlin 2002]. Moreover, benchmark PPP indices take time to compile and are usually made available after few years pass.

Another way is to set a base year and then extrapolate PPPs for the required period. This is done by applying countries' relative inflation rates to the chosen base year. Volumes measured in this way are at *constant international prices*. The underlying assumption of such measurement practise is that price structure is constant within the analyzed period. However, over time the relative price structure does change. By ignoring this, we may acquire a biased picture of economic development.

Which of the two methods should be used is dependent on the time scope of a particular analysis. The former one is advised for studies involving long periods of time (usually a decade or more; see, e.g., Schreyer and Koechlin [2002] for a discussion) while the latter for short.

6. MEASUREMENT STANDARDS AND THEIR IMPLICATIONS FOR GROWTH ACCOUNTING

In order to show what impact different measurement standards have on inference about economic growth let us consider two data sets: *A* and *B*. Both datasets contain information about sixteen countries over the period of eleven years (1995–2005). Both have the same output measure (Gross Value Added) and capital input measure (real fixed capital stock). The difference is only in the way labour input is defined. The first growth accounting estimation is conducted using dataset *A* which contains labour input defined as *total hours worked by persons engaged* (in millions in a given year). As mentioned in the previous section, using this indicator allows us to account for differences in countries' work structure, and to consider all people engaged in a productive activity (like self employed or family workers). Next, we take dataset *B* which uses *total number of employees* (in thousands in a given year), a more common but less precise labour input indicator. Then we run the growth accounting procedure again. All data come from the same database, EU KLEMS project, and can be accessed via its website. Purchasing Power Parities were obtained from OECD-Eurostat statistics and applied as described in section 5. Growth accounting is employed using the decomposition methodology first introduced by Koop, Osiewalski and Steel [1999], briefly outli-

ned in section 2. The estimation procedure is based on Bayesian approach to SFA (*Bayesian Frontier*), and follows Koop, Osiewalski and Steel [1999, 2000a, 2000b]. Since the full posterior distribution is too complex to derive marginal distributions analytically we solve the problem numerically using Gibbs sampler. The results are based on first 500 000 burnt draws and 120 000 retained to compute the characteristics of the posterior marginal distributions. Throughout the study we use posterior means as point estimates and posterior standard deviation as dispersion measures. The list of countries used for this comparison is similar to Makiela [2009], which provides a more in-depth analysis of their growth characteristics (based on dataset *A*). The purpose of this exercise, however, is merely to demonstrate what implications may the above mentioned measurement standards have on inference.

Tables from 1 to 3 and Figures 1 & 2 summarise the main results of such comparative analysis. Using different labour input indicators has a profound implication on economic growth inference. Economic regularity conditions imposed on the translog function based on dataset *A* have been significantly violated when dataset *B* was used (see Table 1 and Figure 1, 2). Though *Returns to Scale* (RTS) estimates in the two datasets are fairly close to each other⁷, estimated elasticities of capital and labour (grand averages) have shifted from a near 1:1 ratio to over 1:7 in favour of labour (in dataset *B*). This change is especially noticeable in Figure 1. All countries in dataset *B* are shifted relatively more to the right-bottom corner on the isoquant map, indicating generally much higher elasticities of labour than capital. Countries most influenced by such change are Denmark, Austria, Germany and Netherlands.

Furthermore, as reported in Table 2 structural decomposition using dataset *A* (which bases on a more detailed indicator of labour) shows an average decline of technical efficiency. Estimates based on dataset *B*, on the other hand, show efficiency growth over time⁸. Considering this as well as other discrepancies in decomposition results between the estimates from both samples, we can conclude that their posterior means are significantly away from each other.

As far as technical efficiency scores are concerned, Spearman's rank correlation coefficient between the two datasets is 0.7118. Sweden has lost its supremacy as the efficiency leader (Table 3). When dataset *B* is used in the analysis, Sweden's score falls below Italy's and is just slightly over Finland's, which jumped from 10th to 3rd third place. What is more, Germany has dropped 8 places, from 6th according to dataset *A* to 14th place in dataset *B*. Considering the underlying definitions of the two labour input indicators, such shifts may inform us of diffe-

⁷ Most importantly *RTS* order in the two datasets has largely remained the same. Spearman's rank correlation is 0.965, with USA and Japan as countries with the highest posterior means of average *Returns to Scale* in the analysed period.

⁸ This being said, I should point out that inference about efficiency change based on both datasets is very uncertain (high posterior standard deviations in respect to posterior means).

rences in work patterns among the countries. Work culture in countries such as Germany or Sweden may be less labour intensive, meaning that though many people are employed, they work relatively few hours per day in comparison to countries such as Italy or Finland.

7. CONCLUDING REMARKS

As indicated in section 6, empirical analysis of economic growth is conditioned upon the underlying data and their methodologies. Issues regarding data comparability over time, across countries, and even between different databases are an ongoing concern for statisticians and policymakers all over the world. Given the presented material, it is safe to say that today we have the means of producing unified, standardised macro-accounts, which present nations' economies in detail and are suitable for international comparisons. Unfortunately, as usual the practice is far from the theory. Developing countries often do not have the means and resources to adopt these standards, unless forced and subsidised by international institutions. Even some members of OECD or EU, organizations so active in bringing standardisation to national accounts, neglect their responsibilities in supplying the necessary statistics. For example Poland, since its accession to the EU, has not delivered a full dataset of fixed capital stock estimates to Eurostat⁹. The country, however, is not the only one and the previously mentioned international databases are missing data for many countries, which should be providing those statistics. Thus, in practice it is often impossible to obtain data for a predefined set of countries.

Currently we are not so much falling behind with setting the new accounting rules as much as with actually applying them in practice. As the world changes fast, new technologies constantly augment the way we think and make our livings. It is logical to assume that, due to the current pace of change, it will always be difficult to develop and apply standards that address our contemporary needs and account for all that is "new" in the economy. But today, even though we have increased the pace of SNA revision, we are still often missing the actual tools (meaning data) for international research and policymaking that this standard was supposed to deliver. Moreover, due to recent increase in SNA's complexity this problem is bound to become worse.

It seems that international organizations such as UN, OECD or EU should re-think their policy priorities regarding national accounts. More stress should be put on the "production" and "delivery" issues rather than SNA's "design" itself. This is because even the best and most up-to-date standard will fail, if its appli-

⁹ 29.07.2010 I obtained a spreadsheet from Eurostat. It contains all the data on (fixed) *gross/net* capital stock that EU members had provided Eurostat up to that date (contact person: Paul Allison). It does not contain data on capital stock for Poland after 2004.

cation across countries is neglected. Unfortunately, so far the pace with which modern national accounts are being implemented worldwide is falling sharply behind the rising quality of their standards, and there seems to be no particular interest in changing the situation.

ACKNOWLEDGEMENTS

This work has benefited greatly from the comments and suggestions provided by Mark Hoffman from Grand Valley State University and Jacek Osiewalski from Cracow University of Economics. All errors and omissions are mine.

REFERENCES

- Ahmad N. 2003. *Measuring Investment in Software*, "OECD Science, Technology and Industry Working Papers", 2003/6, OECD Publishing. doi: 10.1787/335303788330.
- Aigner D., Lovell C.K., Schmidt P. 1977. *Formulation and Estimation of Stochastic Frontier Production Function Models*, "Journal of Econometrics", Vol. 6, Iss. 1, 21–37.
- Bos F. 2008. *Uses of national accounts. History, international standardization and applications in the Netherlands*, "Eagle Economic & Statistics Working Paper", 2008/1, 7–63.
- Cassel G. 1921. *The World's Monetary Problems*, London: Constable.
- Caselli F, Coleman II WJ. 2006. *The World Technology Frontier*, "American Economic Review", Vol. 96, Iss. 3, 499–522.
- Chamberlin G., & Yeuh L., 2006. *Macroeconomics*, London: Thomson Learning.
- Chen V, Gupta A., Therrien A., Levanov G., & van Ark B. 2010. *Recent Productivity Developments in the World Economy: An Overview from The Conference Board Total Economy Database*, "International Productivity Monitor", Vol. 19, Iss. Spring 2010, 3–19.
- Eurostat-OECD. 2006. *Methodological Manual on Purchasing Power Parities*, Retrieved May 12, 2009, from OECD: <http://www.oecd.org/dataoecd/59/10/37984252.pdf>.
- Fried H.O., Lovell C.A.K., Schmidt S.S. 2008. *The Measurement of Productive Efficiency and Productivity Growth*, New York: Oxford University Press.
- Gueye G. 2007. *The revision of ESA 95. Implementation plans*, "Working Party on National Accounts", 2007/11, Luxembourg: Eurostat.
- Gylfason T. 1999. *Principles of Economic Growth*, Oxford: Oxford University Press.
- Intersecretariat Working Group on National Accounts (ISWGNA), 1993 *SNA*, Retrieved September 20, 2010, from United Nations Statistical Division: <http://unstats.un.org/unsd/sna1993/top.asp>.
- Koop G., Osiewalski J., Steel M.F. 1999. *The Components of Output Growth: A Stochastic Frontier Analysis*, "Oxford Bulletin of Economics and Statistics", Vol. 61, Iss. 4, 455–487.
- Koop G., Osiewalski J., Steel M.F. 2000a. *A Stochastic Frontier Analysis of Output Level and Growth in Poland and Western Economies*, "Economics of Planning", Vol. 33, Iss. 3, 185–202.
- Koop G., Osiewalski J., Steel M.F. 2000b. *Modelling the Sources of Output Growth in a Panel of Countries*, "Journal of Business & Economic Statistics", Vol. 18, Iss. 3, 284–299.
- Lequiller F, Ahmad N., Varjonen S., Cave W., & Ahn K.-H. 2003. *Report of the OECD Taskforce on Software Measurement in the National Accounts*, "OECD Statistics Working Papers", 2003/1, OECD Publishing. doi: 10.1787/334811030426.

- Makiela K. 2009. *Economic Growth Decomposition. An Empirical Analysis Using Bayesian Frontier Approach*, "Central European Journal of Economic Modelling and Econometrics", Vol. 1, Iss. 4, 333–369.
- Meeusen W., van den Broeck J. 1977. *Efficiency estimation from Cobb-Douglas Production Function with Composed Error*, "International Economic Review", Vol. 8, Iss. 2, 435–444.
- OECD. 2001. *Measuring Capital – OECD Manual*, Paris: OECD Publication Service.
- OECD. 2003. *Measuring Productivity. Measurement of Aggregate and Industry-Level Productivity Growth*, Paris: OECD Publication Service.
- OECD. 2009. *Glossary of Statistical Terms*, Retrieved May 10, 2009, from stats.oecd.org/glossary.
- Schreyer P. 2003. *Capital stocks, capital services and multi-factor productivity measures*, "OECD Economic Studies", Vol. 37, Iss. 2003/2, 163–184.
- Schreyer P. 2004. *Challenges for productivity measurement in OECD countries*, "8th OECD – NBS Workshop on National Accounts", 2004/12, Paris: OECD Headquarters.
- Schreyer P. 2007. *Old and New Asset Boundaries: A Review Article on Measuring Capital in the New Economy*, "International Productivity Monitor", Vol. 15, Iss. Fall 2007, 75–80.
- Schreyer P., & Koechlin F. 2002. *Purchasing power parities – measurement and uses*, "Statistics Brief", Vol. 3, Iss. 2002/3, 1–8.
- Schreyer P., & Webb C. 2006. *Capital stock data at the OECD: Status and Outlook*, "Mimeo", 2006/10, Paris: OECD.
- Timmer M., O'Mahony M., & van Ark, B. 2007. *The EU KLEMS Growth and Productivity Accounts: An Overview*, "International Productivity Monitor", Vol. 14, Iss. Spring 2007, 71–85.
- Van Ark B., O'Mahony M., & Ypma, G. 2007. *The EU KLEMS Productivity Report*, Groningen: EU KLEMS project.
- Ward M. 2004. *Some reflections on the 1968-93 SNA revision*, "Review of Income and Wealth", Vol. 50, Iss. 2, 299–313.

APPENDIX I. TABLES AND FIGURES

Table 1

Comparison of capital and labour elasticities estimates based on datasets *A* and *B*,
1995–2005 averages

Countries	Results based on dataset <i>A</i>			Results based on dataset <i>B</i>		
	el_K <i>D(el_K)</i>	el_L <i>D(el_L)</i>	RTS <i>D(RTS)</i>	el_K <i>D(el_K)</i>	el_L <i>D(el_L)</i>	RTS <i>D(RTS)</i>
Australia	0.3862 <i>0.0495</i>	0.5963 <i>0.0482</i>	0.9825 <i>0.0133</i>	0.0634 <i>0.1223</i>	0.8773 <i>0.1329</i>	0.9407 <i>0.0496</i>
Austria	0.2984 <i>0.0658</i>	0.6476 <i>0.0766</i>	0.9460 <i>0.0310</i>	-0.0129 <i>0.1557</i>	0.9185 <i>0.1849</i>	0.9056 <i>0.0669</i>
Czech Republic	0.5456 <i>0.0581</i>	0.4262 <i>0.0598</i>	0.9718 <i>0.0175</i>	0.1724 <i>0.1028</i>	0.7508 <i>0.1249</i>	0.9233 <i>0.0481</i>
Denmark	0.0238 <i>0.1415</i>	0.8927 <i>0.1653</i>	0.9166 <i>0.0606</i>	-0.1888 <i>0.2089</i>	1.0777 <i>0.2441</i>	0.8889 <i>0.0864</i>
Finland	0.4108 <i>0.0671</i>	0.5207 <i>0.0800</i>	0.9315 <i>0.0312</i>	0.1133 <i>0.1375</i>	0.7797 <i>0.1719</i>	0.8931 <i>0.0614</i>
Germany	0.2133 <i>0.1072</i>	0.8070 <i>0.0974</i>	1.0203 <i>0.0201</i>	-0.0021 <i>0.1483</i>	0.9946 <i>0.1360</i>	0.9925 <i>0.0430</i>
Italy	0.4200 <i>0.0734</i>	0.5990 <i>0.0777</i>	1.0190 <i>0.0094</i>	0.0249 <i>0.1296</i>	0.9429 <i>0.1252</i>	0.9679 <i>0.0445</i>
Japan	0.3465 <i>0.1212</i>	0.7059 <i>0.1253</i>	1.0524 <i>0.0182</i>	0.0094 <i>0.1521</i>	0.9995 <i>0.1367</i>	1.0089 <i>0.0396</i>
Korea	0.6468 <i>0.1113</i>	0.3894 <i>0.1311</i>	1.0362 <i>0.0283</i>	0.0715 <i>0.1166</i>	0.8902 <i>0.1151</i>	0.9617 <i>0.0429</i>
Netherlands	0.2513 <i>0.0714</i>	0.7116 <i>0.0732</i>	0.9629 <i>0.0284</i>	-0.0292 <i>0.1336</i>	0.9051 <i>0.1473</i>	0.9343 <i>0.0543</i>
Poland	0.5506 <i>0.0698</i>	0.4556 <i>0.0773</i>	1.0062 <i>0.0136</i>	0.1441 <i>0.0973</i>	0.8093 <i>0.1026</i>	0.9533 <i>0.0396</i>
Portugal	0.7446 <i>0.1040</i>	0.2325 <i>0.1053</i>	0.9770 <i>0.0284</i>	0.3808 <i>0.0725</i>	0.5491 <i>0.0874</i>	0.9299 <i>0.0350</i>
Slovenia	0.7832 <i>0.1406</i>	0.1330 <i>0.1411</i>	0.9162 <i>0.0435</i>	0.4360 <i>0.1335</i>	0.4352 <i>0.1622</i>	0.8712 <i>0.0573</i>
Sweden	0.5197 <i>0.0569</i>	0.4373 <i>0.0598</i>	0.9570 <i>0.0211</i>	0.2812 <i>0.0822</i>	0.6449 <i>0.1013</i>	0.9261 <i>0.0393</i>
United Kingdom	0.5080 <i>0.0846</i>	0.5186 <i>0.0961</i>	1.0266 <i>0.0151</i>	0.2173 <i>0.0838</i>	0.7717 <i>0.0763</i>	0.9890 <i>0.0233</i>
United States	0.3506 <i>0.1515</i>	0.7317 <i>0.1640</i>	1.0823 <i>0.0300</i>	0.0650 <i>0.1612</i>	0.9803 <i>0.1474</i>	1.0452 <i>0.0323</i>
Average	0.4375 <i>0.0921</i>	0.5503 <i>0.0986</i>	0.9878 <i>0.0256</i>	0.1128 <i>0.1274</i>	0.8329 <i>0.1373</i>	0.9457 <i>0.0477</i>

Note. “el_” labels denote posterior means of countries’ average elasticities in the analysed period (1995–2005); RTS stands for Returns to Scale; *D*(•) are the corresponding posterior standard deviations where *K* stands for capital and *L* for labour, written in italic; source: author’s calculations

Table 2

Comparison of growth decomposition results based on datasets A and B, 1995–2005 average growth rates

	Empirical GVA growth rate	Results based on dataset A					Results based on dataset B				
		Average technical growth	Average efficiency growth	Average input growth	Average productivity growth	Average output growth	Average technical growth	Average efficiency growth	Average input growth	Average productivity growth	Average output growth
Australia	5.7204	2.5009 <i>0.3261</i>	-0.0028 <i>0.4202</i>	3.3080 <i>0.1759</i>	2.4968 <i>0.2104</i>	5.8871 <i>0.1219</i>	1.3932 <i>0.2641</i>	0.2676 <i>0.2809</i>	4.1537 <i>0.0802</i>	1.6638 <i>0.1144</i>	5.8865 <i>0.0871</i>
Austria	3.9720	3.0096 <i>0.6035</i>	-0.7458 <i>0.6170</i>	1.7748 <i>0.0947</i>	2.2377 <i>0.1501</i>	4.0521 <i>0.1198</i>	2.1074 <i>0.3312</i>	0.2346 <i>0.3051</i>	1.6585 <i>0.0890</i>	2.3460 <i>0.1203</i>	4.0433 <i>0.0832</i>
Czech Republic	4.3918	1.9212 <i>0.2544</i>	0.6584 <i>0.2442</i>	1.8506 <i>0.1998</i>	2.5918 <i>0.2337</i>	4.4900 <i>0.1208</i>	1.9665 <i>0.2337</i>	1.4177 <i>0.2337</i>	1.0423 <i>0.0863</i>	3.4116 <i>0.1226</i>	4.4894 <i>0.0856</i>
Denmark	3.9079	6.0013 <i>3.0638</i>	-2.6315 <i>2.8729</i>	0.8366 <i>0.0981</i>	3.1239 <i>0.1556</i>	3.9865 <i>0.1211</i>	1.8159 <i>1.0832</i>	1.0401 <i>1.0222</i>	1.0908 <i>0.1782</i>	2.8641 <i>0.1996</i>	3.9857 <i>0.0854</i>
Finland	5.0423	2.2120 <i>0.4849</i>	0.6022 <i>0.5044</i>	2.2816 <i>0.1314</i>	2.8253 <i>0.1774</i>	5.1712 <i>0.1202</i>	2.7594 <i>0.2706</i>	-0.2481 <i>0.2811</i>	2.6099 <i>0.0687</i>	2.5037 <i>0.1071</i>	5.1790 <i>0.0846</i>
Germany	3.1007	3.7708 <i>0.8565</i>	-1.3109 <i>0.9590</i>	0.7323 <i>0.4264</i>	2.4031 <i>0.4497</i>	3.1512 <i>0.1159</i>	0.1404 <i>0.3931</i>	-0.2443 <i>0.4258</i>	3.2589 <i>0.2502</i>	-0.1056 <i>0.2553</i>	3.1492 <i>0.0843</i>
Italy	2.9829	2.2477 <i>0.2648</i>	-1.0578 <i>0.3027</i>	1.8557 <i>0.1460</i>	1.1655 <i>0.1766</i>	3.0426 <i>0.1138</i>	0.7334 <i>0.3023</i>	-0.2903 <i>0.3029</i>	2.5789 <i>0.0710</i>	0.4401 <i>0.1054</i>	3.0303 <i>0.0814</i>
Japan	3.1907	2.7018 <i>0.4141</i>	-0.3922 <i>0.5568</i>	0.9268 <i>0.6867</i>	2.2990 <i>0.7041</i>	3.2424 <i>0.1206</i>	0.0975 <i>0.4175</i>	-0.6793 <i>0.4320</i>	4.0520 <i>0.3456</i>	-0.7774 <i>0.3394</i>	3.2420 <i>0.0848</i>
Korea	5.6243	1.8677 <i>0.9452</i>	-0.6458 <i>1.1707</i>	4.5337 <i>0.4837</i>	1.1997 <i>0.4818</i>	5.7855 <i>0.1232</i>	0.9379 <i>0.2721</i>	0.2551 <i>0.3014</i>	4.5368 <i>0.1552</i>	1.1947 <i>0.1719</i>	5.7855 <i>0.0866</i>
Netherlands	5.1154	3.4260 <i>0.7159</i>	-0.1894 <i>0.8000</i>	1.9581 <i>0.1556</i>	3.2245 <i>0.1932</i>	5.2456 <i>0.1156</i>	1.4671 <i>0.2653</i>	0.5493 <i>0.2793</i>	3.1610 <i>0.0712</i>	2.0238 <i>0.1095</i>	5.2487 <i>0.0864</i>
Poland	5.8949	1.8974 <i>0.3612</i>	3.3341 <i>0.3763</i>	0.7391 <i>0.1321</i>	5.2936 <i>0.1850</i>	6.0717 <i>0.1234</i>	1.2623 <i>0.2746</i>	3.9951 <i>0.2974</i>	0.7263 <i>0.0655</i>	5.3071 <i>0.1102</i>	6.0719 <i>0.0866</i>
Portugal	4.5645	2.0038 <i>0.9225</i>	-3.0411 <i>0.9565</i>	5.8451 <i>0.4989</i>	-1.1059 <i>0.4794</i>	4.6723 <i>0.1221</i>	1.5845 <i>0.4549</i>	-1.6977 <i>0.4007</i>	4.8195 <i>0.2561</i>	-0.1416 <i>0.2570</i>	4.6705 <i>0.0853</i>
Slovenia	6.1409	1.7946 <i>1.1235</i>	-1.0311 <i>1.3375</i>	5.5677 <i>0.8552</i>	0.7331 <i>0.8311</i>	6.3346 <i>0.1210</i>	2.9740 <i>0.8530</i>	1.2246 <i>1.0075</i>	2.0202 <i>0.4294</i>	4.2273 <i>0.4481</i>	6.3310 <i>0.0866</i>
Sweden	4.2960	1.9206 <i>0.2386</i>	-0.0853 <i>0.1491</i>	2.5295 <i>0.2040</i>	1.8335 <i>0.2231</i>	4.4090 <i>0.1009</i>	1.8664 <i>0.2702</i>	-0.1323 <i>0.2307</i>	2.6221 <i>0.1167</i>	1.7310 <i>0.1410</i>	4.3984 <i>0.0825</i>
United Kingdom	5.0926	1.9800 <i>0.3250</i>	0.4242 <i>0.4027</i>	2.7456 <i>0.2447</i>	2.4116 <i>0.2688</i>	5.2228 <i>0.1120</i>	0.6625 <i>0.4394</i>	0.6410 <i>0.4449</i>	3.8674 <i>0.1401</i>	1.3059 <i>0.1596</i>	5.2236 <i>0.0849</i>
United States	5.2450	2.4414 <i>0.7813</i>	0.0682 <i>0.4211</i>	2.8077 <i>0.6686</i>	2.5096 <i>0.6745</i>	5.3833 <i>0.1131</i>	0.3216 <i>0.5393</i>	-0.1000 <i>0.4516</i>	5.8363 <i>0.4203</i>	-0.4230 <i>0.4034</i>	5.3870 <i>0.0849</i>
Average	4.6427	2.6061 <i>0.7301</i>	-0.3779 <i>0.7557</i>	2.5183 <i>0.3251</i>	2.2027 <i>0.3497</i>	4.7592 <i>0.1178</i>	1.3806 <i>0.4165</i>	0.3896 <i>0.4186</i>	3.0021 <i>0.1765</i>	1.7232 <i>0.1978</i>	4.7576 <i>0.0850</i>

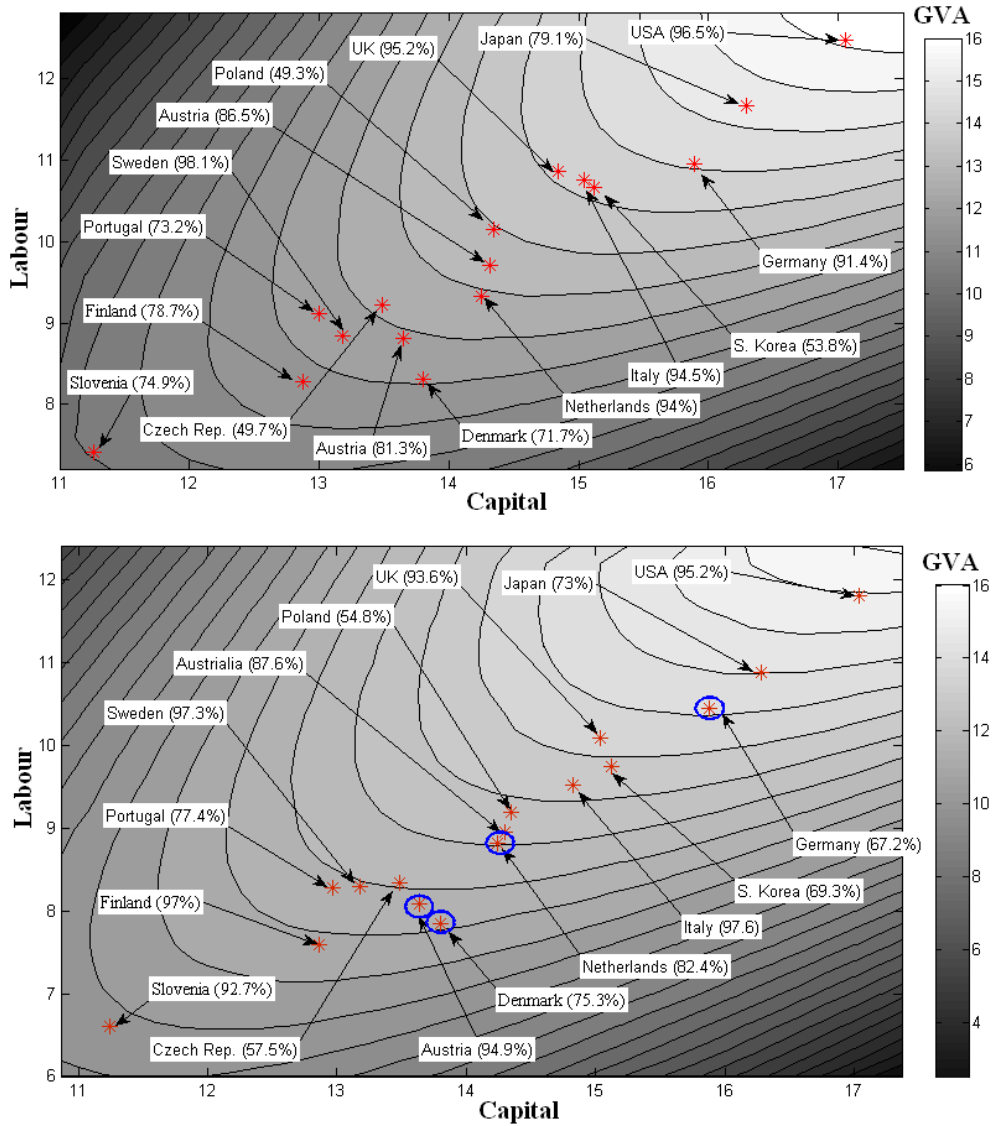
Note. Point estimates are posterior means of countries' average growth rates; posterior standard deviations are in italic; source: author's calculations.

Table 3

Estimation results for average technical efficiencies, datasets A and B

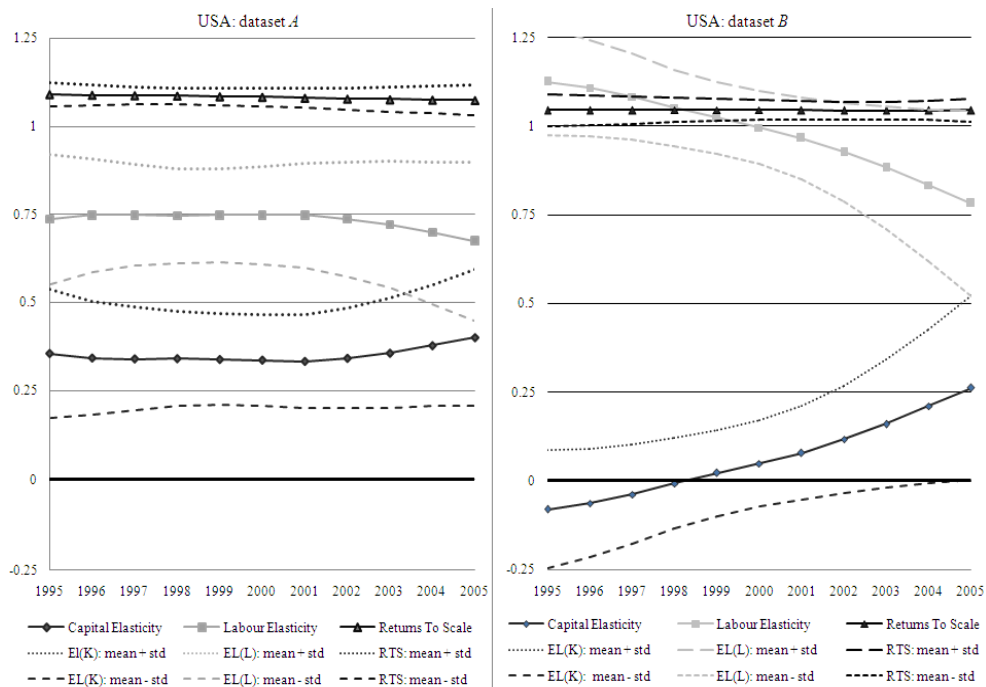
Countries	dataset A			dataset B		
	rank	AEF	<i>D(AEF)</i>	rank	AEF	<i>D(AEF)</i>
Sweden	1	0.9817	<i>0.0088</i>	2	0.9733	<i>0.0107</i>
United States	2	0.9657	<i>0.0245</i>	4	0.9524	<i>0.0260</i>
United Kingdom	3	0.9527	<i>0.0195</i>	6	0.9361	<i>0.0207</i>
Italy	4	0.9453	<i>0.0152</i>	1	0.9760	<i>0.0110</i>
Netherlands	5	0.9408	<i>0.0275</i>	9	0.8240	<i>0.0106</i>
Germany	6	0.9140	<i>0.0425</i>	14	0.6723	<i>0.0117</i>
Australia	7	0.8652	<i>0.0167</i>	8	0.8765	<i>0.0112</i>
Austria	8	0.8135	<i>0.0217</i>	5	0.9491	<i>0.0142</i>
Japan	9	0.7909	<i>0.0201</i>	12	0.7304	<i>0.0144</i>
Finland	10	0.7875	<i>0.0283</i>	3	0.9705	<i>0.0152</i>
Slovenia	11	0.7490	<i>0.0559</i>	7	0.9275	<i>0.0431</i>
Portugal	12	0.7323	<i>0.0317</i>	10	0.7738	<i>0.0145</i>
Denmark	13	0.7172	<i>0.0648</i>	11	0.7573	<i>0.0290</i>
Korea	14	0.5379	<i>0.0258</i>	13	0.6930	<i>0.0093</i>
Czech Rep	15	0.4967	<i>0.0073</i>	15	0.5752	<i>0.0064</i>
Poland	16	0.4936	<i>0.0099</i>	16	0.5480	<i>0.0073</i>
Average	–	0.7927	<i>0.0263</i>	–	0.8210	<i>0.0160</i>

Note. AEF's are posterior means of countries' average efficiency scores in the analysed period; *D(•)*'s denote posterior standard deviations, written in italic; source: author's calculations.



Note. Axes are in natural logs. Countries with negative elasticities are circled; percentages in brackets are based on estimated posterior average efficiencies in the analysed period; countries are placed on the maps according to their productive frontier and average inputs in 1995–2005; source: author’s calculations

Figure 1. Isoquant maps for datasets A and B, 1995–2005 averages



Note. Economic regularity conditions not met (at means) in dataset *B* between 1995 and 1998; source: author's calculations

Figure 2. Example of violated economic regularity conditions, USA, 1995–1998

MODELE HYBRYDOWE MSV-MGARCH Z TRZEMA PROCESAMI UKRYTYMI W BADANIU ZMIENNOŚCI CEN NA RÓŻNYCH RYNKACH

JACEK OSIEWALSKI

Katedra Ekonometrii i Badań Operacyjnych Uniwersytetu Ekonomicznego w Krakowie,
e-mail: eosiewa@cyf-kr.edu.pl

KRZYSZTOF OSIEWALSKI

studia doktoranckie Wydziału Finansów Uniwersytetu Ekonomicznego w Krakowie,
e-mail: krzysztof@osiewalski.eu

Praca została przedstawiona przez autorów na posiedzeniu Komisji Nauk Ekonomicznych i Statystyki Oddziału PAN w Krakowie 24 października 2011 roku.

ABSTRACT

J. Osiewalski, K. Osiewalski *Hybrid MSV–MGARCH models with three latent processes in examining price volatility on different markets*, Folia Oeconomica Cracoviensia

J. Osiewalski and A.Pajor (2007, 2009) and J. Osiewalski (2009) introduced hybrid multivariate stochastic variance — GARCH (MSV–MGARCH) models, where the conditional covariance matrix is the product of a univariate latent process and a matrix with a simple MGARCH structure (Engle’s DCC or scalar BEKK). The aim was to parsimoniously describe volatility of a large group of assets. The proposed hybrid specifications, similarly as other models from the MSV class, require the Bayesian approach equipped with MCMC simulation tools. In order to jointly describe volatility on two different markets (or of two different groups of assets), J. Osiewalski and K.Osiewalski (2011) consider more complicated hybrid models with two latent processes. These new specifications seem very promising due to their good fit and moderate computational requirements.

This paper is devoted to hybrid specifications with three latent processes, even more complicated and located on the edge of possibilities of conducting exact Bayesian analysis. We present full Bayesian inference for such models and propose efficient MCMC simulation strategy. Our approach is used to jointly model volatility of six daily time series representing three different groups: two stock indices, prices of gold and silver, prices of oil and natural gas. We formally compare joint modelling to individual bivariate volatility modelling for each of three groups.

KEY WORDS — SŁOWA KLUCZOWE

Quantitative finance, Volatility analysis, multivariate SV processes, Bayesian inference

Ilościowe finanse, Analiza zmienności, wielowymiarowe procesy SV, wnioskowanie bayesowskie

1. WSTĘP

Osiewalski (2009) oraz Osiewalski i Pajor (2007, 2009) wprowadzili hybrydowe wielowymiarowe modele stochastycznej zmienności lub wariancji (*hybrid multivariate stochastic volatility or variance (MSV) models*). Ich macierz warunkowych kowariancji jest iloczynem pojedynczego procesu ukrytego i macierzy o strukturze zaczerpniętej z prostych procesów MGARCH (DCC, SBEKK). Zaproponowane modele hybrydowe (zwłaszcza model MSF–SBEKK typu I) są przydatne zarówno ze względu na wysoką zdolność opisu zmienności stóp zwrotu, jak i możliwość zastosowania do wielu cen łącznie; Osiewalski i Pajor (2010, 2011) wykorzystują tę specyfikację w analizie ryzyka portfela inwestycyjnego złożonego z akcji kilkunastu spółek.

Jednak pojedynczy proces ukryty wydaje się ograniczać zasadność stosowania modeli hybrydowych tylko do jednorodnych danych (np. zmian cen na jednym rynku). Dlatego Osiewalski i Osiewalski (2011) zaproponowali uogólniony model hybrydowy MSF–SBEKK wykorzystujący dwa procesy ukryte. Przedstawili jego pełną bayesowską analizę statystyczną oraz zastosowanie w przypadku czterowymiarowym: cen złota i srebra oraz dwóch indeksów giełdowych. Przykład ten uzasadnił użycie odrębnych procesów ukrytych w łącznym opisie zmienności na rynkach akcji i metali, a także ukazał powiązania między tymi odrębnymi rynkami zachodzące nawet przed ostatnim kryzysem (dane pochodzą z okresu 8.01.1999–1.02.2006).

W tej pracy proponujemy ogólniejszą strukturę hybrydową, wykorzystującą trzy procesy ukryte. Takie modele MSV są jeszcze dość oszczędne pod względem liczby parametrów i procesów ukrytych, a powinny lepiej (niż model MSF–SBEKK) opisywać zmienność cen na trzech odrębnych rynkach czy składowych portfela zawierającego trzy odrębne grupy aktywów. Aby empirycznie sprawdzić, czy łączne ujęcie zmienności jest wciąż zasadne, rozważamy ten sam okres, ale sześć szeregów czasowych. Przejście od modelu hybrydowego z dwoma procesami ukrytymi do modelu o trzech procesach wydaje się warte uwagi. Głównym problemem nie jest strona teoretyczna takiego uogólnienia, lecz konsekwencje obliczeniowe, gdyż ze wzrostem liczby zmiennych ukrytych (o wielkość równą liczbie obserwacji) znacząco rośnie czas niezbędny do przeprowadzenia symulacji MCMC z rozkładu *a posteriori*. Celem częściowym pracy jest więc sprawdzenie możliwości prowadzenia dokładnej analizy bayesowskiej w nowym przypadku oraz porównanie analizy dokładnej z przybliżoną, polegającą na zastąpieniu głównych macierzy parametrów wartościami opartymi na ocenach MNK.

Następna (druga) część pracy poświęcona jest nowemu modelowi i jego ujęciu bayesowskiemu (zarówno dokładnemu, jak i przybliżonemu). W części trzeciej zastosowano nowy model do łącznego opisu zmienności dwóch indeksów giełdowych, cen złota i srebra oraz ropy naftowej i gazu ziemnego. Podsumowanie wskazujące kierunki dalszych badań stanowi część czwartą (ostatnią).

2. POSTAĆ MODELU I JEGO ANALIZA BAYESOWSKA

Rozważamy n synchronicznych szeregów czasowych (po T obserwacji), które zawierają logarytmiczne stopy zwrotu (zmian cen) $n = n_1 + n_2 + n_3$ aktywów finansowych, należących do trzech odrębnych grup. Dane z okresu t , zawarte w wierszu $r_t = (r_{t1} \dots r_{tn})$ opisujemy (standardowo) procesem VAR(1) — wektorowej autoregresji rzędu 1:

$$r_t = \delta_0 + r_{t-1}\Delta + \varepsilon_t, \quad t = 1, \dots, T, \quad (1)$$

charakteryzowanym przez $n(n+1)$ -wymiarowy wiersz parametrów $\delta = (\delta_0 (\text{vec } \Delta)')$, który to zapis uzyskujemy dzięki wektoryzacji macierzy współczynników Δ . Dla ε_t , tj. wiersza n składników losowych w równaniu VAR(1), przyjmujemy (warunkowy względem przeszłości, oznaczanej ψ_{t-1} , i bieżących zmiennych ukrytych) n -wymiarowy rozkład normalny o zerowej wartości oczekiwanej i macierzy kowariancji

$$\Omega = \begin{bmatrix} g_{11}H_{t,11} & \sqrt{g_{11}g_{12}}H_{t,12} & \sqrt{g_{11}g_{13}}H_{t,13} \\ \sqrt{g_{12}g_{11}}H_{t,21} & g_{12}H_{t,22} & \sqrt{g_{12}g_{13}}H_{t,23} \\ \sqrt{g_{13}g_{11}}H_{t,31} & \sqrt{g_{13}g_{12}}H_{t,32} & g_{13}H_{t,33} \end{bmatrix}. \quad (2)$$

Skalary g_{ii} w (2) są dodatnimi zmiennymi nieobserwowalnymi (ukrytymi), zaś $H_{t,ij}$ ($i, j = 1, 2, 3$) są blokami ($n_i \times n_j$) macierzy kwadratowej (stopnia n) postaci

$$H_t = (1 - \beta - \gamma)A + \beta\varepsilon_{t-1}\varepsilon_{t-1}' + \gamma H_{t-1}, \quad (3)$$

gdzie β i γ są dodatnimi parametrami skalarnymi, spełniającymi warunek $\beta + \gamma < 1$, zaś A jest macierzą symetryczną i dodatnio określoną, zawierającą $n(n+1)/2$ swobodnych parametrów. Macierz H_t jest samoistną macierzą warunkowych kowariancji w tzw. skalarnym modelu BEKK (SBEKK), prostej specyfikacji z klasy wielowymiarowych modeli GARCH (MGARCH), które omawiają Bauwens, Laurent i Rombouts (2006). Gdyby przyjąć $g_{ii} = 1$ ($i = 1, 2, 3$), ε_t byłyby opisane modelem SBEKK(1,1) z warunkowym rozkładem normalnym. Jednak w naszej specyfikacji zakładamy, że 1° ciąg zmiennych ukrytych $\ln(g_{ii})$ ($t \in \mathbb{Z}$, ustalone i) tworzy gaussowski proces AR(1), a 2° ciąg warunkowych macierzy kowariancji Ω_t zależy od trzech procesów ukrytych tak, że wariancje n_i aktywów z grupy i zawierają czynnik g_{ii} . Warunkowe kowariancje zależą od g_{iiv} , jednak sposób wprowadzenia zmiennych ukrytych do macierzy Ω_t sprawia, że warunkowe współczynniki korelacji zależą nie od g_{iiv} lecz od elementów macierzy H_t i mają taką samą postać $h_{t,kl} / \sqrt{h_{t,kk}h_{t,ll}}$ ($k, l = 1, \dots, n$), jak w modelu SBEKK.

W podstawowej wersji naszego hybrydowego modelu MSV ukryte procesy AR(1) mają odrębne parametry i są niezależne:

$$\begin{aligned} \ln(g_{ii}) &= \phi_i \ln(g_{i-1,i}) + \eta_{ii} \quad (i = 1, 2, 3) \\ \eta_{ii} &\sim iN(0, \tau_i^{-1}), \quad \eta_{ii} \perp \eta_{ij} \quad (i, j = 1, 2, 3) \end{aligned} \quad (4)$$

przy czym $\tau_{it} > 0$ oraz $-1 < \varphi_{it} < 1$. Założenie, iż procesy ukryte są identyczne ($g_{t1} = g_{t2} = g_{t3}$) sprowadziłoby nasz nowy model do modelu MSF-SBEKK (typu I), czyli specyfikacji hybrydowej z jednym procesem ukrytym, którą zaproponował Osiewalski (2009). Jedynym teoretycznie uzasadnionym i praktycznie realizowalnym podejściem do wnioskowania statystycznego jest w przypadku modelu (1)–(4) podejście bayesowskie, umożliwiające intuicyjną, probabilistyczną interpretację wyników, stosunkową łatwą analizę dla modeli ze zmiennymi ukrytymi i formalne porównanie mocy objaśniającej konkurencyjnych specyfikacji niezagnieżdżonych.

Bayesowska wersja proponowanego modelu MSV-MGARCH zdefiniowana jest przez rozkład prawdopodobieństwa obserwacji, zmiennych ukrytych i parametrów. Jego funkcja gęstości ma postać

$$\begin{aligned} & p(r_1, \dots, r_T, g_{11}^{-1}, \dots, g_{T1}^{-1}, g_{12}^{-1}, \dots, g_{T2}^{-1}, g_{13}^{-1}, \dots, g_{T3}^{-1}, \theta | \psi_0) \\ &= p(\omega) \left(\prod_{j=1}^3 p(\varphi_j) p(\tau_j) \right) \prod_{t=1}^T p(r_t, g_{t1}^{-1}, g_{t2}^{-1}, g_{t3}^{-1} | \theta, \psi_{t-1}), \end{aligned} \quad (5)$$

gdzie ω grupuje parametry struktur VAR i SBEKK, θ — wszystkie parametry, zaś

$$\begin{aligned} p(r_t, g_{t1}^{-1}, g_{t2}^{-1}, g_{t3}^{-1} | \theta, \psi_{t-1}) &= (2\pi)^{-n/2} (\det H_t)^{-1/2} \exp\left(-\frac{1}{2} \varepsilon_t' \Omega_t^{-1} \varepsilon_t\right) \times \\ &\times \prod_{j=1}^3 (g_{tj}^{-1})^{(n_j/2)-1} f_N^1(\ln g_{tj}^{-1} | \varphi_j \ln g_{t-1,j}^{-1}, \tau_j^{-1}) \end{aligned} \quad (6)$$

jest gęstością warunkowego (względem przeszłości) rozkładu obserwacji i zmiennych ukrytych z okresu t . Gęstości *a priori* $p(\omega)$, $p(\varphi_j)$, $p(\tau_j)$ odzwierciedlają założenia o możliwych wartościach parametrów i są identyczne, jak w modelu MSF-SBEKK (Osiewalski i Pajor, 2009). Wektor parametrów warunkowej wartości oczekiwanej dla r_t , δ , ma $n(n+1)$ -wymiarowy standardowy rozkład normalny, ucięty przez warunek (zwykle niewiążący *a posteriori*), by wartości własne macierzy Δ były wewnątrz koła jednostkowego w przestrzeni liczb zespolonych. Parametr φ_i procesu ukrytego $\{g_{it}, t \in Z\}$ ma rozkład normalny ze średnią 0 i wariancją 100, ucięty do przedziału $(-1, 1)$; parametr τ_i ma rozkład wykładniczy o średniej 200. Wartości początkowe procesów ukrytych ustalamy na poziomie $g_{01} = g_{02} = g_{03} = 1$. Macierz odwrotna do A (w strukturze SBEKK) ma rozkład Wisharta o n stopniach swobody i średniej I_n ; β i γ mają łączny rozkład jednostajny na trójkącie zdefiniowanym warunkami: $\beta > 0$, $\gamma > 0$, $\beta + \gamma < 1$. Warunek początkowy dla H_t ma postać $H_0 = h_0 I_n$ gdzie h_0 jest parametrem o wykładniczym rozkładzie *a priori* ze średnią 1. Łącznie jest $1,5n(n+1) + 9$ parametrów.

Wiedza o wielkościach nieobserwowalnych, uzyskana po wglądzie w dane, zawarta jest w łącznym rozkładzie *a posteriori* o gęstości proporcjonalnej do (5):

$$\begin{aligned}
& p(\mathbf{g}_{11}^{-1}, \dots, \mathbf{g}_{T1}^{-1}, \mathbf{g}_{12}^{-1}, \dots, \mathbf{g}_{T2}^{-1}, \mathbf{g}_{13}^{-1}, \dots, \mathbf{g}_{T3}^{-1}, \boldsymbol{\theta} \mid r_1, \dots, r_T, \psi_0) \\
& \propto p(\omega) \left(\prod_{j=1}^3 p(\varphi_j) p(\tau_j) \right) \prod_{t=1}^T p(r_t, \mathbf{g}_{t1}^{-1}, \mathbf{g}_{t2}^{-1}, \mathbf{g}_{t3}^{-1} \mid \boldsymbol{\theta}, \psi_{t-1});
\end{aligned} \tag{7}$$

rozkład ten jest $[3T + 1,5n(n + 1) + 9]$ -wymiarowy, niestandardowy. Sposobem uzyskania jego charakterystyk jest metoda MCMC, tj. symulacja Monte Carlo za pomocą łańcucha Markowa o rozkładzie stacjonarnym z gęstością (7). Wykorzystujemy łańcuch Gibbsa, zbudowany przez sekwencyjne losowania z warunkowych rozkładów *a posteriori* o gęstościach postaci (notacja zakłada, że indeksy h, i, j są różnymi liczbami ze zbioru $\{1, 2, 3\}$):

$$\begin{aligned}
& p(\omega \mid \psi_0, r_1, \dots, r_T, \mathbf{g}_{11}, \dots, \mathbf{g}_{T1}, \mathbf{g}_{12}, \dots, \mathbf{g}_{T2}, \mathbf{g}_{13}, \dots, \mathbf{g}_{T3}, \varphi_1, \varphi_2, \varphi_3, \tau_1, \tau_2, \tau_3) \\
& \propto p(\omega) \prod_{t=1}^T (\det H_t)^{-1/2} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_t \boldsymbol{\Omega}_t^{-1} \boldsymbol{\varepsilon}_t'\right)
\end{aligned} \tag{8}$$

– gęstość rozkładu $[1,5n(n+1)+3]$ -wymiarowego, niestandardowego (losowania wymagają stosowania algorytmu Metropolis i Hastingsa);

$$\begin{aligned}
& p(\varphi_j, \tau_j \mid \psi_0, r_1, \dots, r_T, \mathbf{g}_{11}, \dots, \mathbf{g}_{T1}, \mathbf{g}_{12}, \dots, \mathbf{g}_{T2}, \mathbf{g}_{13}, \dots, \mathbf{g}_{T3}, \omega, \varphi_h, \tau_h, \varphi_i, \tau_i) \\
& \propto p(\varphi_j) p(\tau_j) \prod_{t=1}^T f_N^1(\ln g_{jt} \mid \varphi_j, \ln g_{t-1,j}, \tau_j^{-1})
\end{aligned} \tag{9}$$

– gęstość standardowa, normalna-gamma, ucięta przez restrykcje $-1 < \varphi_j < 1$;

$$\begin{aligned}
& p(\mathbf{g}_{ij}^{-1} \mid \psi_0, r_1, \dots, r_T, \mathbf{g}_{th}^{-1}, \mathbf{g}_{it}^{-1}, \mathbf{g}_{11}^{-1}, \dots, \mathbf{g}_{t-1,1}^{-1}, \mathbf{g}_{t+1,1}^{-1}, \dots, \mathbf{g}_{T1}^{-1}, \\
& \mathbf{g}_{12}^{-1}, \dots, \mathbf{g}_{t-1,2}^{-1}, \mathbf{g}_{t+1,2}^{-1}, \dots, \mathbf{g}_{T2}^{-1}, \mathbf{g}_{13}^{-1}, \dots, \mathbf{g}_{t-1,3}^{-1}, \mathbf{g}_{t+1,3}^{-1}, \dots, \mathbf{g}_{T3}^{-1}, \boldsymbol{\theta}) \\
& \propto (\mathbf{g}_{ij}^{-1})^{(n_j/2)-1} \exp\left(-\frac{1}{2} \boldsymbol{\varepsilon}_t \boldsymbol{\Omega}_t^{-1} \boldsymbol{\varepsilon}_t'\right) f_N^1(\ln g_{ij}^{-1} \mid \mu_{ij}, \sigma_{ij}^2)
\end{aligned} \tag{10}$$

– gęstość niestandardowa, gdzie dla $t = 1, 2, \dots, T-1$

$$\mu_{ij} = \frac{\varphi_j}{1 + \varphi_j^2} (\ln g_{t-1,j}^{-1} + \ln g_{t+1,j}^{-1}), \quad \sigma_{ij}^2 = \frac{1}{\tau_j (1 + \varphi_j^2)},$$

a dla $t = T$: $\mu_{ij} = \varphi_j \ln g_{t-1,j}^{-1}$, $\sigma_{ij}^2 = \tau_j^{-1}$;

dobry pomocniczy mechanizm losowania w algorytmie Metropolis i Hastingsa jest zapewniony przez rozkład o gęstości

$$q(\mathbf{g}_{ij}^{-1}) = f_G(\mathbf{g}_{ij}^{-1} \mid a_{ij} + \frac{n_j}{2}, b_{ij} + \frac{1}{2} \boldsymbol{\varepsilon}_t^{(j)} H_{t,jj}^{-1} (\boldsymbol{\varepsilon}_t^{(j)})'), \tag{11}$$

gdzie $a_{ij} = [\exp(\sigma_{ij}^2) - 1]^{-1}$, $b_{ij} = a_{ij} \exp(-\mu_{ij} - (\sigma_{ij}^2 / 2))$; gęstość (11) wykorzystuje informację o g_{ij} zawartą tylko w tej części wektora ε_t , która odpowiada aktywowi z grupy j ($j = 1, 2, 3$).

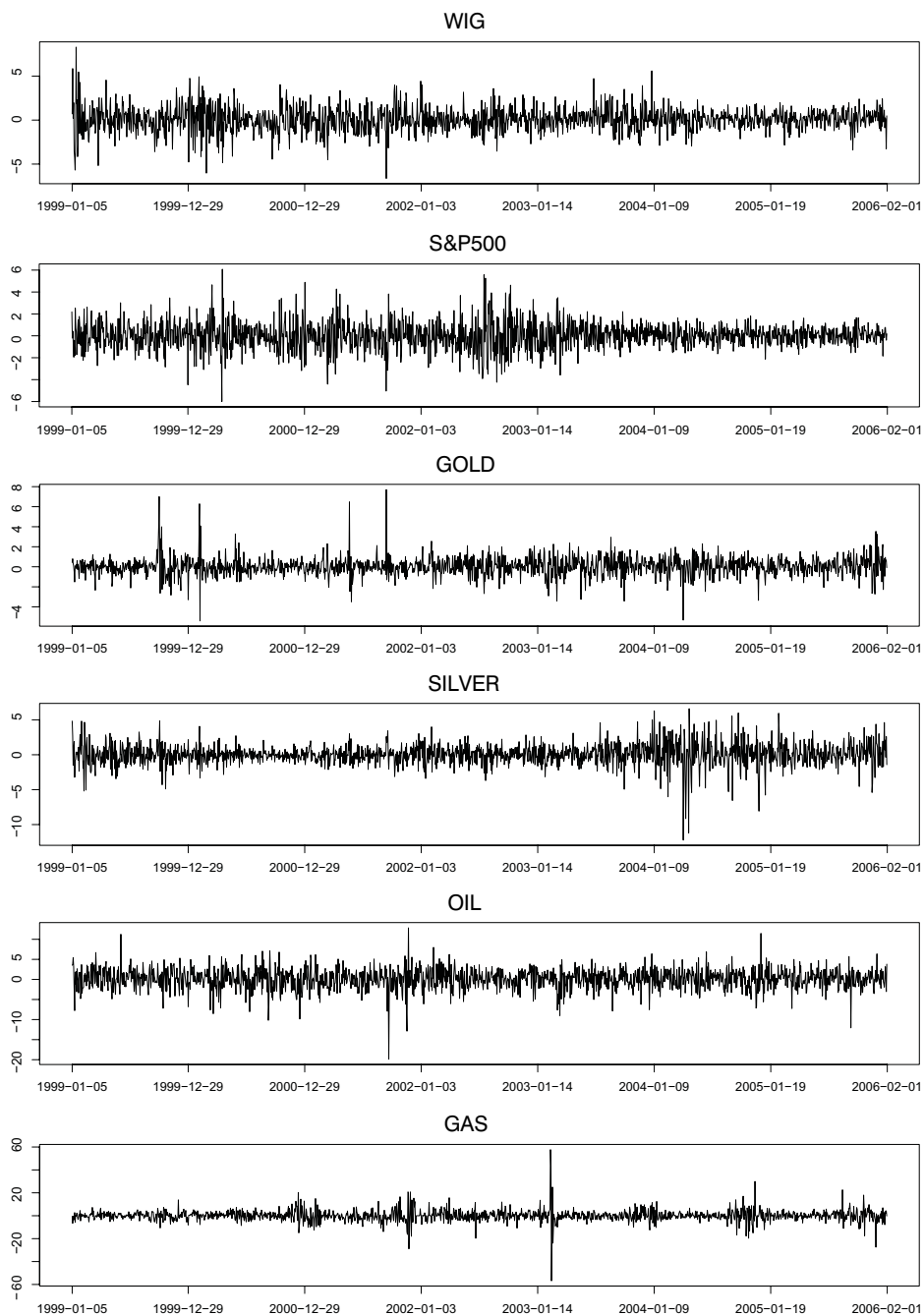
Sekwencyjne losowanie z powyższych rozkładów warunkowych (czyli próbnik Gibbsa) prowadzi — po dostatecznej liczbie losowań wstępnych — do próby zależnej z łącznego rozkładu *a posteriori* o niestandardowej gęstości (7). Na podstawie tej próby uzyskujemy dowolne charakterystyki tego rozkładu, czyli realizujemy wnioskowanie bayesowskie w sposób dokładny. Jednak wymiar wektora ω rośnie z kwadratem liczby szeregów czasowych, czyniąc losowanie Metropolisa i Hastingsa z rozkładu o gęstości (8) zadaniem zbyt trudnym już przy umiarkowanych wartościach n . Zatem Osiewalski (2009) proponuje, a Osiewalski i Pajor (2009) praktycznie stosują, zastąpienie estymacji parametrów macierzowych: δ — ocenami MNK z modelu VAR(1), A — empiryczną macierzą kowariancji reszt MNK. Ze względu na trudności z doбором właściwego mechanizmu pomocniczych losowań i szybko rosnący czas obliczeń, takie przybliżone podejście bayesowskie staje się niezbędne, gdy modelujemy łącznie więcej niż 5 lub 6 szeregów czasowych. Będzie ono w tej pracy porównane z dokładnym wnioskowaniem dla nowego modelu z trzema procesami ukrytymi.

3. ŁĄCZNE MODELOWANIE ZMIENNOŚCI SZEŚCIU SZEREGÓW CZASOWYCH

W tej części przedstawiamy wyniki bayesowskiego modelowania dziennych procentowych logarytmicznych stóp zwrotu (stóp zmian) sześciu szeregów czasowych łącznie. Dwa szeregi reprezentują indeksy giełdowe WIG i S&P500 (rynk akcji) i ten sam okres, co w poprzednich pracach (w celu zachowania porównywalności części wyników — zob. Osiewalski, Pajor, Pipień, 2007; Osiewalski i Pajor, 2009; Osiewalski i Osiewalski, 2011). Dwa szeregi pochodzą z rynków metali szlachetnych — dotyczą cen złota i srebra w tamtym okresie (*London Fix*, USD/oz); zob. Osiewalski i Osiewalski (2011). Dwa szeregi przedstawiają ceny ropy naftowej (*Brent Spot Price*, USD/Barrel) i gazu ziemnego (*Henry Hub Gulf Coast Natural Gas Spot Price*, USD/MMBTU).

Synchronizacja $n = 6$ szeregów pozostawiła $T = 1656$ obserwacji dla stóp zmian, które przedstawiono na Rys. 1. Przyjęto model hybrydowy (1)–(5) z 3 procesami ukrytymi ($n_1 = n_2 = n_3 = 2$). W Tabelach 1 i 2 zebrano statystyki opisowe danych.

Zmienność indeksów giełdowych, cen metali i cen surowców paliwowych jest wyraźnie inna. W stosunku do stóp zmian cen metali, stopy zmian indeksów z rynku akcji charakteryzują się zdecydowanie mniejszą (co do modułu) skośnością i kurtozą, są też słabo między sobą skorelowane. Słaba korelacja stóp zmian dotyczy też cen ropy i gazu, jednak rozproszenie obu tych stóp jest



Źródło: opracowanie własne.

Rysunek 1. Modelowane dzienne procentowe stopy zmian (8.01.1999–1.02.2006)

duże, a kurtoza w przypadku gazu jest ogromna. Warto wyjaśnić, że wielkość kurtozy w przypadku gazu ziemnego wynika ze skoku cenowego (z 8 do 22 USD/MMBTU), a następnie powrotu do wielkości go poprzedzającej, w dniach 23–27 lutego 2003. Według raportu *Staff Investigating Team (Federal Energy Regulatory Commission, 23 czerwca 2003)* było to podyktowane splotem wielu czynników destabilizujących równowagę popytu i podaży na rynku gazu. Z jednej strony zimny front okalający wschodnią część USA po bardzo mroźnej zimie spowodował znacznie większe zużycie gazu niż przewidywano na tę porę roku. Z drugiej, nie dość że rezerwy gazu zostały przez ową zimę w dużej mierze skonsumowane, to zimny front spowodował zamarznięcie części studni wydobywczych, przez co uniemożliwił wydobycie gazu w południowo-wschodniej części USA.

Tabela 1

Statystyki opisowe dziennych procentowych stóp zmian ($n = 6, T = 1656$)

szereg (stopy zmian)		średnia	odch. st.	skośność	kurtoza
WIG	8.01.1999–1.02.2006	0.0626	1.3686	0.0703	5.5488
S&P500		0.0018	1.2233	0.1160	5.0695
GOLD		0.0413	0.9540	0.6527	11.3989
SILVER		0.0415	1.5580	-0.6256	9.4357
OIL		0.1117	2.5474	-0.5465	6.9719
NAT.GAS		0.0873	5.1060	0.3954	28.3150

Źródło: opracowanie własne.

Tabela 2

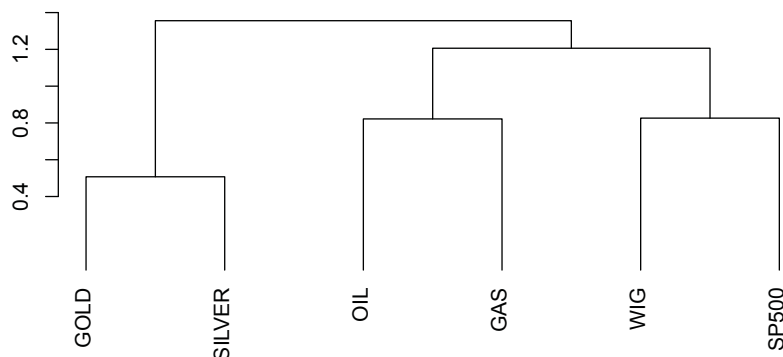
Współczynniki korelacji dziennych procentowych stóp zmian ($n = 6, T = 1656$)

	WIG	S&P500	GOLD	SILVER	OIL	NAT.GAS
WIG	1					
S&P500	0.1745	1				
GOLD	-0.0467	-0.0984	1			
SILVER	0.0546	-0.0417	0.4928	1		
OIL	0.0066	-0.0440	0.0743	0.0242	1	
NAT.GAS	-0.0124	-0.0149	0.0701	0.0273	0.1785	1

Źródło: opracowanie własne.

Aby zobrazować podobieństwo dynamiki stóp zmian posłużono się analizą skupień, dokonaną na podstawie współczynników korelacji podanych w Tabeli 2. Największe podobieństwo wykazywały w badanym okresie stopy zmian cen

złota i srebra (zob. Rys. 2); znacząco mniejsze cechowało stopy zmian indeksów giełdowych, tworzących drugie skupienie, oraz zmian cen ropy i gazu, tworzących skupienie trzecie. Występowanie tych trzech skupień w sposób nieformalny potwierdza zasadność użycia modelu z trzema procesami ukrytymi, po jednym dla pary o podobnej dynamice zmian.



Źródło: opracowanie własne.

Rysunek 2. Podobieństwo dziennych procentowych stóp zmian (metoda Warda)

W Tabeli 3 pokazujemy logarytmy dziesiętne czynnika Bayesa, uzyskane przez zastąpienie wartości brzegowych gęstości macierzy obserwacji ich symulacyjnymi ocenami Newtona i Raftery'ego. Wyniki wskazują, że pojedynczy proces ukryty to w rozważanym przypadku za mało, bo model MSF-SBEKK (na nim oparty) jest ok. 10^{118} razy mniej prawdopodobny *a posteriori* (przy równych szansach *a priori*) niż model obecnie proponowany. Jeśli do nowego modelu wprowadzamy *ad hoc* oceny MNK parametrów macierzowych (nie prowadzimy dla nich formalnego wnioskowania), to i tak nowy model — z taką silną (i niezbyt trafną) restrykcją na większość parametrów — jest ok. 10^{54} razy bardziej prawdopodobny *a posteriori* niż dokładnie szacowany model MSF-SBEKK. Cena pełnej bayesowskiej estymacji modelu z 3 procesami ukrytymi jest duża, bo czas obliczeń sięga granicy praktycznej stosowalności takiej specyfikacji, i to przy tylko 6 szeregach czasowych i liczbie obserwacji niewielkiej jak na dane dzienne. Nawet wstępne zastosowanie MNK dla większości parametrów nie skraca dostatecznie czasu obliczeń, co może stwarzać istotne ograniczenia dla wykorzystania modelu z 3 procesami ukrytymi. Trudno wyobrazić sobie jego użycie w analizie portfela 50 akcji, który rozważali Osiewalski i Pajor (2010, 2011), opisując zmienność za pomocą modelu MSF-SBEKK z ocenami MNK. Nasz nowy model może być jednak przydatny w badaniu efektu zarażania na niewielu rynkach, reprezentowanych przez główne ceny lub indeksy cenowe (jak w obecnym przykładzie).

Tabela 3

Bayesowskie porównanie hybrydowych modeli MSV dla 6 szeregów danych

Model VAR(1) + M_i	Liczba swobodnych parametrów (i zm. ukrytych)	$\log_{10}(B_{NOWY,i})$	czas obliczeń
NOWY (dokładnie)	72 (+3T)	0	26 h 3 min
NOWY (z MNK)	9 (+3T)	64	16 h 14 min
MSF-SBEKK typ I (dokładnie)	68 (+T)	118	7 h 53 min
MSF-SBEKK typ I (z MNK)	5 (+T)	155	2 h 32 min

Źródło: opracowanie własne.

Należy zwrócić uwagę na wartości oczekiwane (i — w nawiasach — odchylenia standardowe) *a posteriori* parametrów nowego modelu:

$$E(\delta') = \begin{bmatrix} \mathbf{0.073} \\ (0.027) \\ \mathbf{0.025} \\ (0.025) \\ \mathbf{0.010} \\ (0.018) \\ -\mathbf{0.015} \\ (0.027) \\ \mathbf{0.187} \\ (0.052) \\ \mathbf{0.079} \\ (0.072) \end{bmatrix}, \quad E(\Delta') = \begin{bmatrix} \mathbf{0.028} & \mathbf{0.282} & \mathbf{0.058} & \mathbf{0.023} & -\mathbf{0.002} & -\mathbf{0.010} \\ (0.024) & (0.029) & (0.035) & (0.019) & (0.011) & (0.005) \\ \mathbf{0.028} & -\mathbf{0.051} & \mathbf{0.040} & -\mathbf{0.024} & -\mathbf{0.003} & -\mathbf{0.002} \\ (0.021) & (0.026) & (0.032) & (0.017) & (0.010) & (0.005) \\ \mathbf{0.023} & -\mathbf{0.011} & -\mathbf{0.038} & -\mathbf{0.001} & \mathbf{0.008} & \mathbf{0.004} \\ (0.014) & (0.015) & (0.029) & (0.016) & (0.007) & (0.003) \\ \mathbf{0.035} & \mathbf{0.003} & \mathbf{0.254} & -\mathbf{0.202} & \mathbf{0.013} & -\mathbf{0.001} \\ (0.019) & (0.021) & (0.043) & (0.029) & (0.010) & (0.005) \\ \mathbf{0.031} & \mathbf{0.070} & -\mathbf{0.044} & \mathbf{0.019} & \mathbf{0.022} & -\mathbf{0.007} \\ (0.043) & (0.047) & (0.057) & (0.034) & (0.026) & (0.014) \\ \mathbf{0.037} & \mathbf{0.036} & -\mathbf{0.143} & \mathbf{0.002} & \mathbf{0.004} & \mathbf{0.013} \\ (0.060) & (0.069) & (0.092) & (0.053) & (0.038) & (0.026) \end{bmatrix},$$

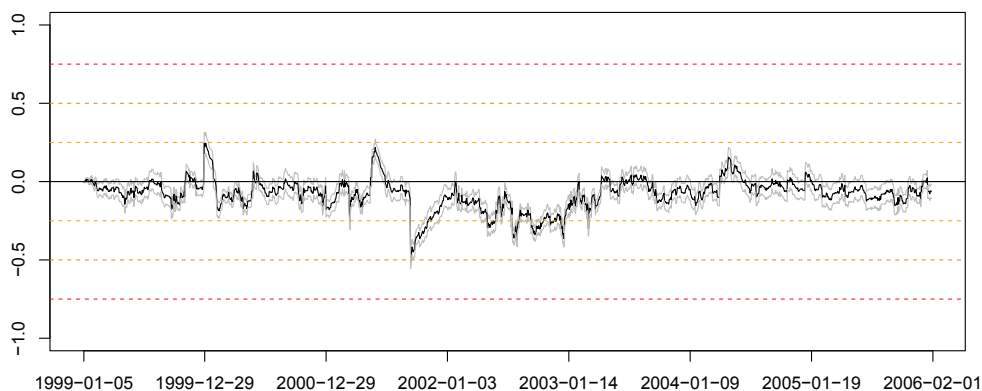
$$E(\varphi') = \begin{bmatrix} \mathbf{0.977} \\ (0.014) \\ \mathbf{0.503} \\ (0.071) \\ \mathbf{0.613} \\ (0.102) \end{bmatrix}, \quad E[(\tau^{-1})'] = E[(\sigma^2)'] = \begin{bmatrix} \mathbf{0.012} \\ (0.005) \\ \mathbf{0.482} \\ (0.072) \\ \mathbf{0.281} \\ (0.069) \end{bmatrix},$$

$$E(A) = \begin{bmatrix} \mathbf{1.913} & \mathbf{0.312} & -\mathbf{0.054} & \mathbf{0.027} & -\mathbf{0.027} & -\mathbf{0.022} \\ (0.655) & (0.127) & (0.041) & (0.060) & (0.126) & (0.188) \\ & \mathbf{1.454} & -\mathbf{0.064} & -\mathbf{0.031} & -\mathbf{0.123} & -\mathbf{0.095} \\ & (0.518) & (0.037) & (0.054) & (0.113) & (0.163) \\ & & \mathbf{0.469} & \mathbf{0.329} & \mathbf{0.119} & \mathbf{0.144} \\ & & (0.044) & (0.045) & (0.071) & (0.109) \\ & & & \mathbf{1.026} & -\mathbf{0.052} & \mathbf{0.060} \\ & & & (0.104) & (0.107) & (0.162) \\ & & & & \mathbf{5.192} & \mathbf{1.634} \\ & & & & (0.474) & (0.369) \\ & & & & & \mathbf{10.298} \\ & & & & & (1.034) \end{bmatrix},$$

$$E(\beta) = \mathbf{0.021} \quad (0.003), \quad E(\gamma) = \mathbf{0.943} \quad (0.008), \quad E(h_0) = \mathbf{2.381} \quad (0.710);$$

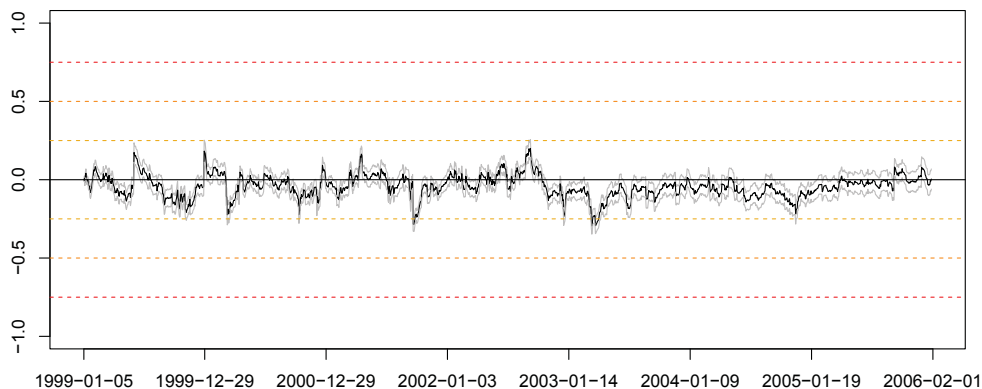
pogrubioną czcionką zapisano wartości oczekiwane tych parametrów, w przypadku których wartość zerowa nie mieści się w przedziale między kwantylami *a posteriori* na poziomie 0.025 i 0.975 (w tym sensie są to parametry istotnie różne od zera).

Skupianie się istotnych parametrów macierzy A w osobne bloki, odpowiadające rynkom akcji, metali i surowców paliwowych, może nasuwać przypuszczenie, że łączne modelowanie wybranych sześciu szeregów jest zbędne. Z drugiej strony, elementy macierzy Δ nie potwierdzają tego przypuszczenia — wpływ opóźnionej stopy zmian ceny ropy naftowej na bieżącą stopę zmian indeksu S&P500 wydaje się istotny. W celu formalnej weryfikacji tego przypuszczenia oszacowano trzy odrębne modele MSF–SBEKK; w tym przypadku zakłada się całkowitą niezależność. Okazuje się, że łączny model hybrydowy (z trzema procesami ukrytymi dla sześciu szeregów) jest ok. $10^{5.5}$ razy bardziej prawdopodobny *a posteriori* niż trzy odrębne modele. Porównując to z wynikami w Tabeli 3 stwierdzamy z kolei, że trzy odrębne modele są ok. $10^{112.5}$ razy bardziej prawdopodobne niż łączny model z jednym procesem ukrytym. Okazuje się zatem, że odrębne procesy ukryte są ważniejsze dla wyjaśnienia zmienności tych szeregów niż łączne modelowanie z jednym tylko procesem ukrytym. Model najlepszy wśród rozważanych zapewnia i łączny opis, i pewną odrębność. Jego adekwatność wyjaśniają przebiegi warunkowych współczynników korelacji między dziennymi stopami zmian indeksu S&P500 i cen złota lub ropy naftowej; na Rys. 3 i 4 podano wartości oczekiwane (\pm dwa odchylenia standardowe) *a posteriori*. Warunkowe korelacje pozostają niezerowe (dodatnie lub, częściej, ujemne) przez długie okresy, więc zmienność na różnych rynkach nie jest niezależna. Natomiast średnia wartość oczekiwana *a posteriori* warunkowych współczynników korelacji (z S&P500) jest równa dla cen złota -0.0806 (przy średnim odchyleniu



Źródło: opracowanie własne.

Rys. 3. Wartości oczekiwane *a posteriori* warunkowego współczynnika korelacji stóp zmian indeksu S&P500 oraz ceny złota

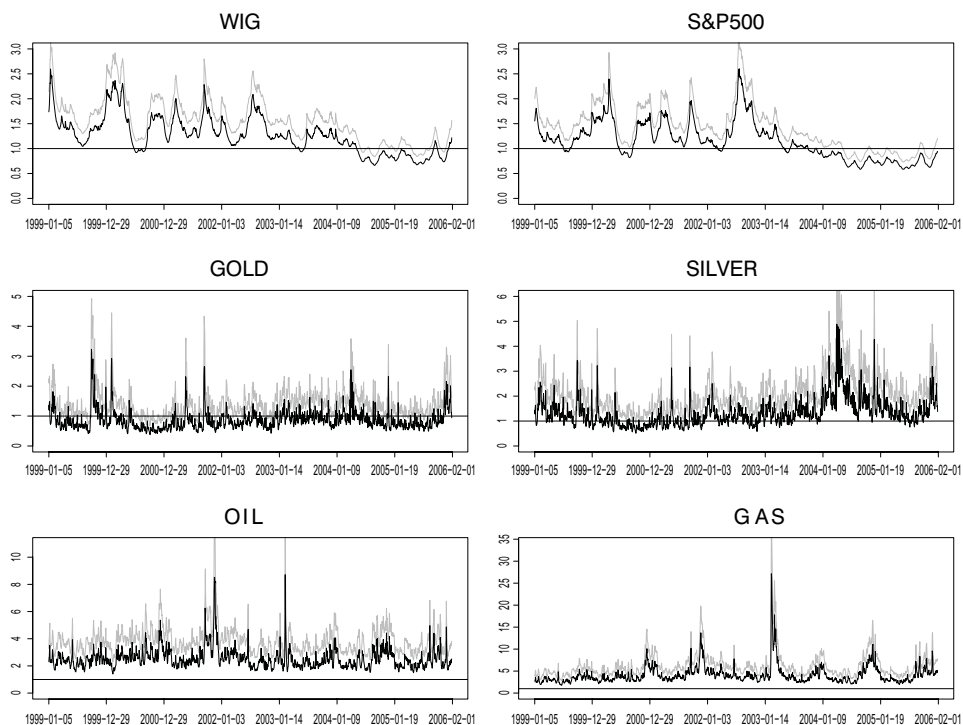


Źródło: opracowanie własne.

Rys. 4. Wartości oczekiwane *a posteriori* warunkowego współczynnika korelacji stóp zmian: indeksu S&P500 oraz ceny ropy naftowej

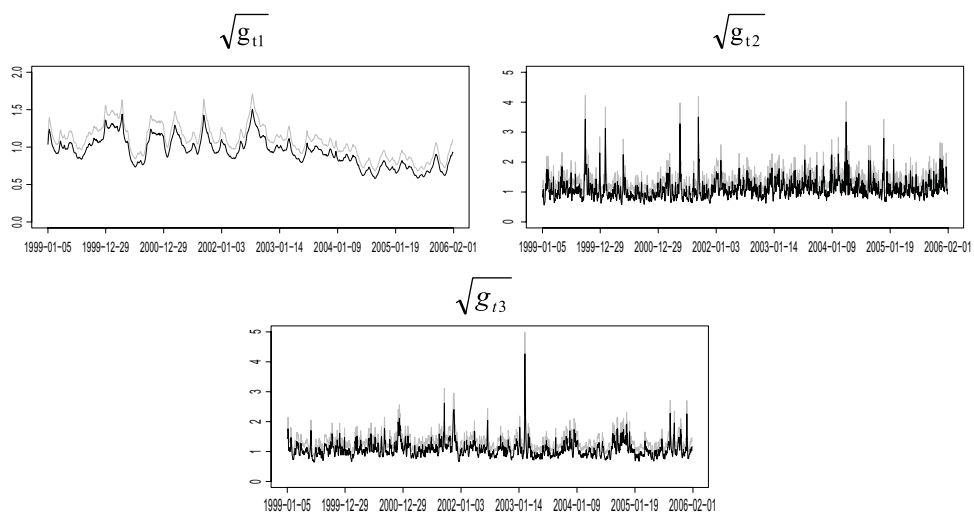
standardowym *a posteriori* równym 0.1001), więc nie różni się znacząco od empirycznej korelacji obu szeregów (-0.0984); w przypadku ropy mamy -0.0438 (0.0755) wobec bardzo podobnej, bliskiej zeru empirycznej korelacji (-0.0440). Wnioskowanie o warunkowych współczynnikach korelacji między stopami zmian cen na rozważanych rynkach potwierdza, że są między nimi zwykle dość słabe, ale jednak istotne zależności. Znaczące ujemne skorelowanie między indeksem S&P500 oraz cenami złota i ropy zachodziło w II połowie 2001 roku, świadcząc o istniejącej wtedy możliwości dywersyfikacji ryzyka inwestycji.

Na Rys. 5 pokazano wartości oczekiwane *a posteriori* warunkowych odchyłeń standardowych, a na Rys. 6 ich składowej ukrytej $\sqrt{g_{ii}}$ ($i=1,2,3$). Zmienność indeksów giełdowych (mierzona warunkowym odchyleniem standardowym) jest znacznie mniej wahliva niż zmienność cen metali. Bardziej nieregularny jest przebieg drugiego procesu ukrytego, z gwałtownymi wybuchami powyżej wyjściowego poziomu 1. Duże różnice w przebiegu procesów ukrytych są zgodne z wartościami oczekiwanymi *a posteriori* ich parametrów, bardzo od siebie odległymi. Dla φ_1 mamy 0.977, tj. silną autokorelację procesu $\ln(g_{t1})$, o bezwarunkowej wariancji ok. $0.012/(1-0.977^2) = 0.264$. Dla φ_2 mamy 0.503, co oznacza słabą autokorelację $\ln(g_{t2})$, przy wariancji ok. 0.645, czyli relatywnie dużej. Dla φ_3 mamy 0.613, umiarkowaną autokorelację $\ln(g_{t3})$, przy wariancji ok. 0.450, czyli też umiarkowanej. Procesy ukryte mają istotny udział w opisie zmienności sześciu szeregów czasowych, a odmienność ich przebiegu wyjaśnia ogromną przewagę nowego modelu hybrydowego nad modelem MSF-SBEKK (z jednym procesem ukrytym dla wszystkich szeregów).



Źródło: opracowanie własne.

Rys. 5. Wartości oczekiwane (+ dwa odchylenia standardowe) *a posteriori* warunkowych odchyłeń standardowych



Źródło: opracowanie własne.

Rys. 6. Wartości oczekiwane (+ odchylenie standardowe) *a posteriori* pierwiastków zmiennych ukrytych

4. UWAGI KOŃCOWE

Celem pracy była prezentacja nowego wielowymiarowego hybrydowego modelu stochastycznej zmienności (lub wariancji), jego bayesowskiej analizy statystycznej i praktycznej przydatności w łącznym opisie zmienności cen na różnych rynkach. Nowy model, wykorzystujący warunkową strukturę korelacyjną SBEKK i trzy procesy ukryte, pozwala na adekwatne i oszczędne ujęcie dynamiki stóp zwrotu dla wielu aktywów, które można podzielić w sposób naturalny na trzy grupy.

Dalsze badania powinny mieć na celu uogólnienie modelu na dowolną (choć w praktyce niewielką) liczbę odrębnych grup aktywów, a także do sprawdzenia jego przydatności w przypadku dużej łącznej liczby aktywów. W szczególności ważny jest wstępny podział portfela na kilka grup aktywów o podobnej dynamice cen, a następnie porównanie analizy ryzyka dużego portfela, przeprowadzanej za pomocą modelu z jednym procesem ukrytym (zob. Osiewalski i Pajor, 2010, 2011) oraz za pomocą specyfikacji o kilku procesach ukrytych. Innym kierunkiem możliwych zastosowań jest analiza powiązań zmienności na wielu rynkach (finansowych, metali, surowców energetycznych i innych) w okresie ostatniego kryzysu o charakterze globalnym. Odmiennym polem badań metodycznych i empirycznych jest bayesowskie porównanie najnowszych specyfikacji MSV: naszych hybrydowych (zawierających elementy struktur MGARCH) oraz tych, które zaproponowali Philipov i Glickman (2006) oraz Gouriéroux, Jasiak i Sufana (2009).

PODZIĘKOWANIA

Autorzy wyrażają szczególną wdzięczność Annie Pajor (za wprowadzanie w niełatwą tematykę modeli MSV) oraz Błażejowi Mazurowi (za zwrócenie uwagi na ograniczenia modelu MSF-SBEKK w opisie zmienności niejednorodnych portfeli).

LITERATURA

- Bauwens L., Laurent S., Rombouts J.V.K. 2006. *Multivariate GARCH models: A survey*, "Journal of Applied Econometrics" 21, 79–109.
- Gouriéroux C., Jasiak J., Sufana R. 2009. *The Wishart Autoregressive process of multivariate stochastic volatility*, "Journal of Econometrics" 150, 167–181.
- Osiewalski J. 2009. *New hybrid models of multivariate volatility (a Bayesian perspective)*, "Przegląd Statystyczny (Statistical Review)" 56, 15–22.
- Osiewalski J., Osiewalski K. 2011. *Modele hybrydowe MSV-MGARCH z dwoma procesami ukrytymi*, „Zeszyty Naukowe Uniwersytetu Ekonomicznego w Krakowie — seria Finanse” (w druku).

- Osiewalski J., Pajor A. 2007. *Flexibility and parsimony in multivariate financial modelling: a hybrid bivariate DCC–SV model* [w:] *Financial Markets. Principles of Modeling, Forecasting and Decision-Making* (FindEcon Monograph Series No. 3, edited by W.Milo and P.Wdowiński), Łódź University Press, Łódź, 11–26.
- Osiewalski J., Pajor A. 2009. *Bayesian analysis for hybrid MSF–SBEKK models of multivariate volatility*, “Central European Journal of Economic Modelling and Econometrics” 1, 179–202.
- Osiewalski J., Pajor A. 2010. *Bayesian Value-at-Risk for a portfolio: multi- and univariate approaches using MSF–SBEKK models*, “Central European Journal of Economic Modelling and Econometrics” 2, 253–277.
- Osiewalski J., Pajor A. 2011. *Bayesian Value-at-Risk and Expected Shortfall for a large portfolio (multi- and univariate approaches)*, “Acta Physica Polonica A” (w druku).
- Osiewalski J., Pajor A., Pipień M. 2007. *Bayesian comparison of bivariate GARCH, SV and hybrid models* [w:] “MACROMODELS’2006, Proceedings of the 33rd International Conference” (edited by W.Welfe and A.Welfe), Absolwent, Łódź, 247–277.
- Pajor A. 2010. *Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej — ujęcie bayesowskie*, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie, Kraków.
- Philipov A., Glickman M. 2006. *Multivariate Stochastic Volatility Via Wishart Processes*, “Journal of Business and Economic Statistics” 24(3), 313–328.

BADANIA NAD WŁAŚCIWOŚCIAMI OPERACYJNYMI METODY AHP

ANNA PRUSAK

Uniwersytet Ekonomiczny w Krakowie
Katedra Zarządzania Jakością

PIOTR STEFANÓW

Krakowska Akademia im A.F. Modrzewskiego
Katedra Metod Statystycznych
e-mail: piotr@stefanow.net

ABSTRACT

An ability to generate consistent results is the most critical operational feature of every method. The objective of the present paper was to demonstrate the results of the study conducted within the Ministry of Science and Higher Education Grant NN111345138 (*Methodology of multicriteria comparative analysis of elements*), regarding the relationship between the level of inconsistency of the AHP results, measured by Consistency Ratio (CR), and number of the compared criteria. Experiment involved analysis of 3-8 characteristics of selected objects (laptop, printer, public transport, mobile phone, flat and food store). The results showed that the analysis of as few as 3 characteristics resulted in a very high fraction of inconsistent opinions (more than 50%). In case of 6,7 and 8 characteristics this fraction exceeded 80%.

KEY WORDS — SŁOWA KLUCZOWE

analytic hierarchy process (AHP), analytic network process (ANP), consistency ratio (CR),
Multiple criteria decision analysis

analityczny proces hierarchiczny (AHP), analityczny proces sieciowy (ANP),
zgodność wyników (CR), wielowymiarowa analiza porównawcza (WAP)

1. WPROWADZENIE

Metoda Analitycznego Procesu Hierarchicznego (AHP — *Analytic Hierarchy Process*) oraz Analitycznego Procesu Sieciowego (ANP — *Analytic Network Process*) to jedne z najpopularniejszych na świecie wielokryterialnych narzędzi wspomagających podejmowanie decyzji. Zostały one zaproponowane w latach 70 XX w.

przez amerykańskiego matematyka, T.L. Saaty'ego. U podstaw metody AHP leży założenie, że większość złożonych problemów decyzyjnych można rozłożyć na czynniki pierwsze i przedstawić w postaci drzewa hierarchicznego (struktury hierarchicznej). Struktura hierarchiczna podlega następnie analizie poprzez dokonanie porównań parami poszczególnych jej elementów oraz obliczeniu odpowiednich współczynników wagowych. W ten sposób metoda AHP porządkuje warianty decyzyjne od optymalnego do najmniej pożądanego, tym samym ułatwiając podjęcie decyzji co do wyboru któregoś z nich. Poszczególne etapy tej metody zostały w sposób syntetyczny omówione w dalszej części pracy.

Problem decyzyjny nie zawsze można zaprezentować w postaci drzewa hierarchicznego, które nie uwzględnia pewnych złożonych relacji, jak np. sprzężenia zwrotnego pomiędzy poszczególnymi elementami. Dlatego też zostało opracowane inne narzędzie — ANP, które stanowi rozszerzenie metody AHP, uwzględniające różnego rodzaju powiązania pomiędzy oraz wewnątrz poszczególnych grup kryteriów i subkryteriów. Metoda ANP nie jest przedmiotem rozważań niniejszego artykułu, ale należy podkreślić, że analiza danych w przypadku obydwu metod oparta jest na tej samej, uniwersalnej procedurze postępowania, polegającej na dokonywaniu porównań parami między sobą wszystkich elementów modelu przy użyciu fundamentalnej skali Saaty'ego. Procedury te generują jednak szereg problemów, zarówno na poziomie budowy modelu jak i analizy, które w praktyce prowadzą do niskiej wiarygodności wyników. Są to m.in. nadmierne uproszczenie rzeczywistości w modelu hierarchicznym, agregacja wyników AHP w grupowym podejmowaniu decyzji, zmiana wyników na skutek pojawienia się dodatkowego elementu w modelu oraz trudność w uzyskaniu spójnych (zgodnych) wyników. Przykłady wybranych problemów zostały omówione np. w [Stefanów & Prusak 2011].

Problem dotyczący trudności w uzyskaniu zgodnych (spójnych) wyników, szczególnie gdy dokonywana jest duża liczba porównań, jest najpoważniejszym zarzutem wobec AHP/ANP. Tymczasem zdolność do generowania niesprzecznych wyników jest krytyczną właściwością operacyjną każdej metody. Dostrzegając ten problem, twórca narzędzi AHP/ANP zaproponował specjalny wskaźnik zgodności (współczynnik *CR* — *Consistency Ratio*), za pomocą którego na każdym etapie analizy kontrolowana jest spójność (zgodność) otrzymanych rezultatów. Jeśli są one niespójne, badanie należy powtórzyć bądź odrzucić. Nasuwa się w związku z tym pytanie, jak często takie niezgodności się zdarzają i od czego zależą. Wyniki badań przedstawione w niniejszym artykule, prowadzone w ramach Grantu MNiSW NN111345138: *Metodologia wielokryterialnej analizy porównawczej obiektów*, są jedną z nielicznych prób udzielenia odpowiedzi na to pytanie. Istnieje bowiem bardzo wiele publikacji dotyczących tego problemu, ale większość z nich próbuje rozwiązać go od strony analitycznej. Nie są znane autorom żadne próby badania od strony empirycznej poziomu niezgodności wyników AHP w zależności od takich czynników jak np. sposób prezentacji danych

i administrowania ankiety, liczba porównywanych elementów oraz liczba stopni na skali porównań¹.

Prowadzenie badań dotyczących kluczowych właściwości operacyjnych AHP/ANP jest szczególnie ważne ze względu na rosnącą w świecie popularność tych metod, zarówno w nauce, jak i w praktyce. Mimo, iż mają one wielu krytyków, którzy zwracają uwagę na ich poważne niedociągnięcia, to w bazach czasopism wysokopunktowanych można znaleźć ogromną liczbę artykułów na temat aplikacji AHP/ANP w różnych obszarach. Żadna inna wielokryterialna metoda wspomagająca decyzje (np. VDA, Electre, Promethe) nie doczekała się takiej liczby publikacji. Po metody AHP/ANP coraz częściej sięgają też praktycy. Duża popularność AHP/ANP w porównaniu do innych wielokryterialnych metod podejmowania decyzji wynika m.in. z tego, że metody te pozwalają w prosty sposób zdekomponować złożony problem decyzyjny i zanalizować go, w wyniku czego otrzymuje się konkretne wyniki liczbowe wskazujące optymalny wariant decyzyjny. Ponadto, posiadają one łatwo dostępne dla użytkownika oprogramowanie komputerów². Duża liczba zastosowań wymaga zwrócenia szerszej uwagi na wiarygodność obliczeń, które mogą mieć potem wpływ na realne decyzje.

Celem niniejszego artykułu jest przedstawienie rezultatów badań dotyczących zależności pomiędzy poziomem niezgodności wyników AHP, mierzonym za pomocą powszechnie stosowanego współczynnika CR, a liczbą porównywanych elementów. Dokonano tego za pomocą eksperymentu, obejmującego analizę od 5 do 8 właściwości różnych obiektów. Przedstawiono również cząstkowe badania dotyczące 3 kryteriów.

W pierwszej części pracy krótko przedstawiono zasady działania metody AHP. Druga część artykułu koncentruje się na prezentacji uproszczonej metody szacowania współczynnika zgodności porównań CR. Kolejna część pracy stanowi syntetyczny przegląd światowej literatury dotyczącej AHP/ANP, ze szczególnym naciskiem na metodologiczne aspekty tych metod. Pracę kończy opis, omówienie wyników eksperymentu oraz wnioski i implikacje dla przyszłych badań.

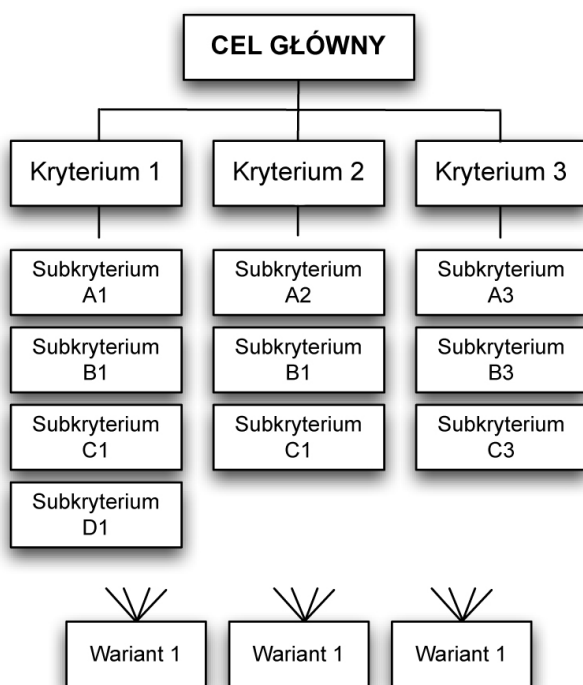
2. ZASADY DZIAŁANIA METODY AHP

Etapy analizy AHP oraz ich podstawy matematyczne można znaleźć praktycznie w każdym podręczniku i artykule T. Saaty'ego oraz każdej publikacji dotyczącej metod AHP/ANP. Wiele z nich odwołuje się także do tzw. 7 filarów AHP [Saaty

¹ Wpływ tych czynników na zgodność wyników AHP uzyskanych na podstawie porównań dokonywanych na skali wielostopniowej stanowi przedmiot badań ww. grantu MNiSW NN111345138: *Metodologia wielokryterialnej analizy porównawczej obiektów* oraz grantu NCN 2011/01/D/HS4/04006: *Czynniki wpływające na zgodność porównań parami w metodach analitycznego procesu hierarchicznego i sieciowego (AHP/ANP)*.

² *ExpertChoice* przeznaczone tylko dla AHP oraz *SuperDecisions*, przeznaczone dla ANP, ale można go również stosować dla modeli AHP.

2001]. Informacje te są również w przystępny sposób przekazane w artykułach publikowanych w języku polskim [np. Adamus 2006]. Przykład najprostszej i najczęściej konstruowanej struktury hierarchicznej został zaprezentowany na rys. 1. Składa się ona z kilku poziomów, z których najwyższy to cel główny problemu decyzyjnego. Na realizację tego celu wpływają położone o poziom niżej kryteria, a te z kolei są zależne od odpowiadających im subkryteriów (czynników). Na samym dole znajdują się warianty decyzyjne (zwane również „alternatywami”). W tak zbudowanym modelu ostateczna decyzja polega na wyborze tego wariantu decyzyjnego, który w najwyższym stopniu spełnia najważniejsze subkryteria oraz kryteria.



Źródło: Opracowanie własne na podstawie np. [Saaty 1994].

Rys. 1. Przykład struktury hierarchicznej w modelu AHP

Ważność poszczególnych elementów względem siebie wyznaczana jest przez dokonanie ich porównań parami. Porównania te dokonywane są z użyciem tzw. fundamentalnej skali porównań Saaty’ego, gdzie przewaga jednego elementu nad drugim oznaczana jest od „takiego samego znaczenia” (=1), poprzez „słabą przewagę” (=3), „mocną przewagę” (=5), „bardzo mocną przewagę” (=7) aż

do „absolutnej/ekstremalnej przewagi” (=9). W przypadku, gdy osoba oceniająca nie jest zdecydowana do końca, można do skali wprowadzić liczby pośrednie (2,4,6,8). Porównania dokonywane są na zasadzie odwrotności: jeśli element A jest 3 razy lepszy od B, to B stanowi 1/3 elementu A. W sumie skala daje aż 17 możliwości. Tak wyrażone opinie wprowadzane są do kwadratowej macierzy A porównań parami o wymiarach ($n \times n$), gdzie n jest liczbą porównywalnych kryteriów. Macierz ta stanowi główne narzędzie analizy metody AHP [np. Saaty 1996]:

$$\mathbf{A} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ 1/a_{12} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 1/a_{1n} & 1/a_{2n} & \cdots & 1 \end{bmatrix} \quad (1)$$

Taka macierz konstruowana jest na każdym poziomie modelu AHP. W modelu zaprezentowanym na rys. 1, w pierwszej kolejności budowana jest macierz służąca analizie ważności kryteriów względem celu głównego. Następnie budowane są macierze dla analizy znaczenia subkryteriów w stosunku do właściwych im kryteriów. Macierzy tych jest tyle, ile w modelu jest „klastrow” zależności kryterium-subkryteria. Dla każdej macierzy obliczane są tzw. wektory priorytetów $W(w_1, w_2, w_3)$, które określane są także jako współczynniki wagowe (wagi), i które wskazują relatywną ważność porównywanych elementów. Ponieważ uzyskanie

Tabela 1

Obliczanie współczynników wagowych (priorytetów)

			Średnia geometryczna r_i	Priorytety (wagi) w_i
1	a_{12}	a_{13}	$\sqrt[3]{1 \times a_{12} \times a_{13}}$	$r_1 / \sum_{i=1}^3 r_i$
$1/a_{12}$	1	a_{23}	$\sqrt[3]{1/a_{12} \times 1 \times a_{23}}$	$r_2 / \sum_{i=1}^3 r_i$
$1/a_{13}$	$1/a_{23}$	1	$\sqrt[3]{1/a_{13} \times 1/a_{23} \times 1}$	$r_3 / \sum_{i=1}^3 r_i$
			$\sum_{i=1}^3 r_i$	$\sum_{i=1}^3 w_i = 1$

Źródło: na przykład [Adamus & Łasak 2009].

dokładnych wyników wymaga wykonania szeregu skomplikowanych procedur matematycznych dla każdej macierzy³, w literaturze dotyczącej AHP pokazane są uproszczone metody obliczania, za pomocą których uzyskuje się zbliżone wyniki. Do najczęściej stosowanych należy obliczanie za pomocą średniej geometrycznej, gdzie stopień pierwiastka równy jest liczbie porównywanych elementów [np. Adamus & Łasak 2009]. Poszczególne wartości tak obliczonego wektora macierzy **A** są normalizowane, dając wektory priorytetów w_1, w_2, w_3 . Procedura ta została przedstawiona w tabeli 1.

Powyższa procedura musi być przeprowadzona dla każdej macierzy. Współczynniki wagowe (priorytety) określają względną ważność elementów decyzyjnych na każdym poziomie modelu hierarchicznego. Są to tzw. *priorytety lokalne*, czyli wartości obejmujące relację względem elementu położonego jedynie o jeden poziom wyżej w strukturze hierarchicznej. *Priorytety globalne* określają stopień ważności elementów względem elementu zlokalizowanego w hierarchii o więcej niż jeden poziom wyżej, np. znaczenie subkryteriów w odniesieniu do celu. Ich wartości globalne oblicza się poprzez przemnożenie wartości priorytetu dla relacji subkryterium-kryterium przez wartość właściwego priorytetu kryterium-cel. Tak przeprowadzona procedura nie gwarantuje jednak wiarygodności ostatecznych wyników. Metody AHP/ANP wymagają bowiem na każdym etapie analizy obliczenia tzw. współczynnika niezgodności *CR* (*Consistency Ratio*), który wskazuje na zgodność uzyskanych porównań. Procedura ta została w skrócie omówiona poniżej.

3. WŁAŚCIWOŚCI OPERACYJNE: ZDOLNOŚĆ GENEROWANIA ZGODNYCH WYNIKÓW

Podstawową właściwością operacyjną każdej metody jest zdolność do generowania niesprzecznych, czyli zgodnych, logicznych rezultatów. Wyniki analizy struktury hierarchicznej, według procedury przedstawionej w poprzednim rozdziale, tylko wtedy są wiarygodne, a więc dające rzetelną wskazówkę do podjęcia decyzji, kiedy zostaje spełniony warunek zgodności ocen. Współczynnik zgodności nie może przekraczać 10% ($CR \leq 0,10$)⁴. Podobnie jak w przypadku priorytetów, obliczanie współczynnika *CR* następuje automatycznie przy korzystaniu z odpowiedniego oprogramowania AHP/ANP (ExpertChoice, SuperDecisios). Podstawą do obliczenia *CR* jest wyznaczenie największej wartości własnej (λ_{max}) macierzy **A**. W uproszczony sposób oblicza się ją jako sumę iloczynów sumy wartości porównań w każdej kolumnie macierzy **A** oraz odpowiedniego dla danego elementu współczynnika wagowego. Korzystając z oznaczeń zapisanych w tabeli 1, uproszczoną formułę można przedstawić w następujący sposób:

³ Oprogramowanie ExpertChoice i SuperDecisions wykonują te obliczenia automatycznie.

⁴ Szerzej na temat wartości „granicznej” można znaleźć np. w [Saaty & Vargas 1982].

$$\lambda_{\max} = (1 + 1/a_{21} + 1/a_{31}) \cdot w_1 + (a_{12} + 1 + 1/a_{32}) \cdot w_2 + (a_{13} + a_{23} + 1) \cdot w_3 \quad (2)$$

Kolejnym krokiem jest obliczenie tzw. indeksu niezgodności (*Inconsistency Index* — *IC*) jako ilorazu różnicy największej wartości własnej macierzy i liczby porównywanych elementów⁵ (n) przez różnicę ($n - 1$):

$$IC = (\lambda_{\max} - n)/(n - 1) \quad (3)$$

Na samym końcu obliczamy współczynnik *CR* poprzez podzielenie *IC* przez stabilizowaną wartość *RI*:

$$CR = IC/RI \quad (4)$$

Stabilizowane wartości *RI* zostały wygenerowane przez Saaty'ego na podstawie symulacji dla 500 000 macierzy, i w zależności od liczby porównywanych elementów wynoszą np.:

dla $n = 3$, $RI = 0,52$;

dla $n = 4$, $RI = 0,89$;

dla $n = 5$, $RI = 1,11$;

dla $n = 6$, $RI = 1,25$;

Wartości *RI* dla większej liczby n , jak i dokładniejszy opis sposobu ich generowania, można znaleźć np. w [Alonso & Lamata 2006]. Jeśli *CR* przekroczy wartość 0,10, porównania należy uznać za niezgodne, a analizę należy powtórzyć [Saaty 2008]. Ponieważ powtarzanie analizy jest bardzo czasochłonne, wielu badaczy proponowało stosowanie różnych algorytmów mających na celu zredukowanie wartości *CR* przy minimalnej ingerencji (zmianie) danych (porównań) pierwotnych. Ponieważ w każdym przypadku „sztucznej” redukcji współczynnika zgodności zawsze następuje zaburzenie pierwotnych porównań, najlepiej jest uzyskać maksymalnie zgodne wyniki już na etapie gromadzenia danych, czyli dokonywania porównań parami. Aby to się udało, zaleca się porównywanie nie więcej niż 7 ± 2 elementów w jednej macierzy, gdyż tyle właśnie porównań jest w stanie dokonać przeciętny człowiek zachowując wymagany poziom spójności [Saaty & Ozdemir 2003, Miller 1956].

4. PRZEGLĄD LITERATURY DOTYCZĄCEJ AHP/ANP

Istnieje bardzo wiele publikacji na temat wielokryterialnych metod wspomagających procesy podejmowania decyzji w różnych obszarach nauki i praktyki. Oprócz metod AHP/ANP, należą do nich m.in. Electre, Promethe, VDA. Jak zo-

⁵ Im bardziej największa wartość własna macierzy zbliżona do n , tym mniejszy współczynnik *CR* [Saaty 2008].

stało wcześniej wspomniane, metody AHP/ANP cieszą się zdecydowanie największą popularnością w porównaniu do innych metod, o czym świadczy duża (i wciąż szybko rosnąca) liczba publikacji dotyczących tych metod w największych bazach czasopism naukowych, takich jak np. *ScienceDirect*, *Wiley*, *Scopus*. Korzystając tylko z wyszukiwarki bazy *ScienceDirect*, wpisano następujące słowa kluczowe: „analytic hierarchy process”, „analytic network process”, „ahp”, „anp”. Hasło „analytic hierarchy process” wygenerowało 20290⁶ artykułów, natomiast jego skrót „ahp” 10881⁷. Publikacji na temat ANP jest znacznie więcej, należy jednak wziąć pod uwagę fakt, iż prawie w każdym artykule dotyczącym ANP znajduje się odniesienie do metody AHP. Jak bowiem wcześniej wspomniano, „fundament” tych metod, 9-stopniowa skala porównań Saaty’ego, jest uniwersalny dla obydwu metod. Słowo kluczowe „analytic network process” spowodowało znalezienie aż 45890⁸ publikacji, natomiast „anp” dało 18882⁹ rezultatów. Jeszcze pół roku temu liczby te były średnio o kilkaset pozycji niższe, na przykład dla hasła „analytic network process” liczba znalezionych publikacji wynosiła 41000¹⁰. Porównując częstość występowania artykułów na temat AHP/ANP z innymi metodami wspomagania decyzji, „vda” jako słowo kluczowe w tej samej bazie daje 2093¹¹ rezultatów, „electre” 1241¹², natomiast „promethe” zaledwie 23¹³. Co więcej, wartości te nie zmieniły się znacząco w porównaniu ze stanem sprzed pół roku¹⁴. Należy ponadto wziąć pod uwagę fakt, że część z tych artykułów zawiera także odniesienie do metod AHP/ANP.

Artykuły dotyczące AHP/ANP publikowane są głównie w takich czasopismach jak *European Journal of Operational Research*, *International Journal of Production Economics*, *Information and Management*, *Computers and Operations Research*, *Decision Support Systems*, *Socio-Economic Planning Science*, *Journal of Operational Research Society*, *Expert Systems with Applications*, *Journal of Mathematical Psychology*. Publikacje te można generalnie podzielić na dwie grupy: o charakterze „aplikacyjnym”, skoncentrowanym na rozwiązywaniu za pomocą AHP/ANP problemie, oraz „metodologicznym”, gdzie „środek ciężkości” przypada na właściwości samej metody. Ogólne studium kilkuset abstraktów artykułów w obszarze AHP/ANP wykazało, że zdecydowana większość badań w tym zakresie ma charakter „aplikacyjny”. Dotyczą one zastosowania AHP/ANP w różnych obszarach i w odniesieniu do różnych celów decyzyjnych. Takimi celami decyzyjnymi mogą być na przykład planowanie, zakupy, alokacja zasobów, rozwiązywanie konfliktu,

⁶ Stan na 5.11.2011.

⁷ Stan na 5.11.2011.

⁸ Stan na 5.11.2011.

⁹ Stan na 5.11.2011.

¹⁰ Stan na 5.06.2011.

¹¹ Stan na 5.11.2011.

¹² Stan na 5.11.2011.

¹³ Stan na 5.11.2011.

¹⁴ Na przykład, w porównaniu do 5.06.2011 liczba artykułów dt. VDA wzrosła tylko o 2.

ewaluacja, optymalizacja, itd. W obszarze aplikacji, nie ma chyba dziedziny, w której by nie próbowano stosować narzędzi AHP/ANP. Obszary te, w powiązaniu z wyżej wspomnianymi typami celów decyzyjnych, zostały w sposób syntetyczny omówione przez [Vaidya & Kumar 2006]. Przykłady różnorodności zastosowań AHP/ANP to rozwiązywanie takich problemów jak ewaluacja dostawców [Akarte *et al.* 2001], wybór najlepszego oprogramowania [Lai *et al.* 2002], zdefiniowanie czynników sukcesu i opracowanie strategii wdrażania systemu ISO-14001 na podstawie analizy korzyści i kosztów [Chin *et al.* 1999], ewaluacja różnych strategii fuzji w bankowości [Arbel & Orger, 1990], wybór lokalizacji rafinerii w kilku krajach Bliskiego i Środkowego Wschodu [Badri 1999], poprawa jakości komunikacji pacjent-lekarz [Singpurwalla *et al.* 1999], analiza strategii USA wobec Iranu [Saaty 2008] i wiele innych.

Publikacji o charakterze „metodologicznym” zidentyfikowano znacznie mniej. Są to zarówno krytyczne analizy AHP/ANP, jak i propozycje rozwiązań problemów wynikających z matematycznych/algebraicznych własności tych metod. Do najczęściej dyskutowanych problemów związanych z praktycznym użytkowaniem AHP/ANP należą m.in.: agregacja indywidualnych priorytetów w grupowym podejmowaniu decyzji [np. Solms 2009, Forman & Peniwati 1998, Beynon 2005], zmiana kolejności wyników (*rank reversal*) na skutek wprowadzenia do modelu dodatkowego obojętnego elementu [np. Watson & Freeling 1982, Belton & Gear 1983, Weber 1997, Holder 1990, Perez *et al.* 2001] oraz trudności w uzyskaniu spójnych (zgodnych) wyników AHP/ANP [np. Vargas 1982, Finan & Hurley 1997, Apostolou & Hassel 1993; 2002, Alonso & Lamata 2005; 2006, Costa 2006, Ergu *et al.* 2011]. Dużo mniej informacji można znaleźć na temat „operacyjnych” (praktycznych) aspektów AHP/ANP, dotyczących m.in. sposobu gromadzenia danych. Tymczasem na podstawie badań empirycznych zaobserwowano, że końcowe rezultaty są zależne m.in. od sposobu prezentacji kwestionariusza [np. Webber *et al.* 1996] oraz rodzaju użytej skali [Weathers *et al.* 2005].

Przyczyny niezgodności rezultatów AHP/ANP już od dawna stanowiły przedmiot badań na całym świecie. Głównym obszarem, na którym skoncentrowane są badania dotyczące niezgodności, jest współczynnik CR. Istniejąca literatura w tym zakresie próbowała odpowiedzieć na trzy główne pytania badawcze. Po pierwsze, czy ustalona przez Saaty’ego wartość współczynnika niezgodności (CR) na poziomie $\leq 0,10$ powinna być tylko wskazówką, czy też sztywną regułą, odnoszącą się do każdego przypadku. Po drugie, czy współczynnik CR jest odpowiednią miarą poziomu niezgodności. Po trzecie, czy i jak można zredukować współczynnik CR na poziomie macierzy. Pytaniem, do którego praktycznie brak jest w literaturze odniesienia, to jakie czynniki na etapie gromadzenia danych (np. liczba porównywanych elementów) wpływają na poziom zgodności wyników.

W odniesieniu do pierwszej kwestii, część autorów sugeruje, że poziom współczynnika CR ustalony na stałym poziomie $CR \leq 0,10$ jest zbyt restrykcyjny.

[Apostolou & Hassel 1993] próbowali udowodnić za pomocą eksperymentu, że wartość współczynnika CR przekraczająca 0,10 nie wpływa w sposób istotny na końcowe wyniki badań. Eksperyment ten został zakwestionowany przez [Chu & Liu 2001], którzy z kolei znaleźli istotną różnicę pomiędzy wynikami z $CR \leq 0,10$, a tymi gdzie $CR > 0,10$. Inny problem był rozważany m.in. przez [Alonso & Lamata 2006], którzy zaproponowali odmienny sposób pomiaru poziomu niezgodności wyników prowadzący do akceptacji, bądź odrzucenia macierzy. To, czy dana macierz jest spójna bądź nie, zależy od dwóch czynników:

- a) indeksu zgodności (IC), za który autorzy uważają największą wartość własną macierzy λ_{max} ;
- b) wymaganego poziomu zgodności (α), $0 < \alpha \leq 1$, przy czym α jest adaptowalny w zależności od badanego problemu; użytkownik w różnych sytuacjach może potrzebować różnego poziomu zgodności i ten poziom określa procentowo.

Autorzy ponadto kwestionują stabilizowane wartości RI, wygenerowane przez Saaty'ego na podstawie symulacji 500 000 macierzy. Te bowiem są różne dla każdej podobnej symulacji.

Trzecie zagadnienie dotyczy redukcji współczynnika CR na poziomie macierzy. Na przykład, [Costa *et al.* 2006] zastosował w tym celu tzw. algorytm genetyczny (Genetic Algorithm — GA). Polega on na wygenerowaniu w macierzy, gdzie niezgodność wynosi powyżej 10% oraz rozwiązań z CR mniejszym lub równym 10%. Program najpierw analizuje zgodność, jeśli $CR > 10\%$, algorytm losowo generuje serię macryc z wartością niższą niż 10%, zbliżonych do macierzy pierwotnej. Z kolei [Ergu *et al.* 2011] zaproponował trzystopniową procedurę redukcji CR opartą na identyfikacji najbardziej niezgodnych elementów macierzy, [Finan & Hurley 1997] za pomocą symulacji Monte Carlo udowodnili, że matematyczna redukcja wartości CR zwiększa wiarygodność rezultatów analizy. Wielu badaczy proponowało i proponuje różne procedury i algorytmy mające na celu zmniejszenie niezgodności, ale wątpliwość budzi fakt, iż w wielu przypadkach tego typu operacje prowadzą do zmiany oryginalnych informacji uzyskanych podczas prowadzenia badań [np. Tung & Tang 1998].

Ponieważ $CR \leq 0,10$ nadal jest jedyną akceptowalną i powszechnie stosowaną miarą poziomu niezgodności wyników AHP/ANP, istnieje silna potrzeba wypracowania określonych zaleceń dotyczących redukcji możliwości uzyskania zgodnych wyników już na etapie gromadzenia danych. Do takich „działań prewencyjnych” należy przede wszystkim upewnienie się m.in. że istnieje dostateczna ilość informacji na temat badanego problemu, a model decyzyjny został odpowiednio ustrukturyzowany. Dużą rolę odgrywa także wiedza ekspertów, która w przypadku otrzymania niezgodnych wyników, powinna być poddana dodatkowej ocenie [Szczypińska & Piotrowski 2009]. Inną przyczyną generującą sprzeczne rezultaty może być fakt stosowania w AHP/ANP 9-stopniowej skali liniowej, podczas gdy większość problemów nie ma charakteru liniowego. Dla-

tego też jako bardziej wiarygodną sugerowano skalę multiplikatywną (geometryczną) [Holder 1990].

5. EKSPERYMENT

5.1. Opis eksperymentu

Celem przeprowadzonego eksperymentu było zbadanie, jaka jest empiryczna frakcja niezgodnych ocen, dokonywanych przez ekspertów według kryterium oceny zgodności zaproponowanej przez Saaty'ego (współczynnik CR), w zależności od liczby porównywanych parami kryteriów. Określono hipotezę badawczą, która brzmiała, że dla liczby kryteriów większej od 5 liczba sprzecznych (niezgodnych) ocen będzie bardzo wysoka (większa od 30%). Badanie polegało na obliczeniu wskaźnika CR uzyskanego na podstawie zastosowania przez ekspertów metody AHP. Zostali oni poproszeni o porównanie parami kryteriów dotyczących pięciu obiektów. Zostały one dobrane w taki sposób, aby ankietowani (ze względów organizacyjnych byli to najczęściej studenci krakowskich uczelni) mogli być traktowani jako eksperci.

1. Telefon komórkowy (T)
2. Mieszkanie (M)
3. Komunikacja miejska (K)
4. Drukarka (D)
5. Laptop (L).

Ze względu na możliwość popełnienia podczas badania błędu systematycznego polegającego na tym, że porównywanie pierwszego obiektu „zdecyduje” kolejne porównania zdecydowano w każdym zestawie — w sposób losowy — ustalić inną kolejność porównywanych obiektów.

Każdy obiekt był charakteryzowany maksymalnie przez osiem kryteriów, które zostały przedstawione w tabeli 2.

Ze względu na możliwość popełnienia podczas badania błędu systematycznego polegającego na tym, że początkowe kryterium „zdecyduje” kolejne porównania zdecydowano podczas każdego badania — w sposób losowy — ustalić inną kolejność kryteriów. Każdy z obiektów był oceniany ze względu na:

1. Pięć kryteriów
2. Sześć kryteriów
3. Siedem kryteriów
4. Osiem kryteriów

Każdy z ankietowanych został poproszony o ocenę czterech z pięciu obiektów (T, M, K, D, L). Liczba kryteriów dla każdego z obiektów była różna (ułożona w sposób losowy), co oznacza, że jeśli ankietowany oceniał np. (T) ze względu na 5 kryteriów, to pozostałe obiekty mógł ocenić ze względu na 6, 7 lub 8 kryteriów.

Jako kryterium zgodności przyjęto omówiony wcześniej współczynnik CR , którego wartość nie powinna przekraczać 0,10.

Tabela 2

Kryteria charakteryzujące porównywane obiekty

Telefon komórkowy (T)	Mieszkanie (M)	Komunikacja miejska (K)
<ul style="list-style-type: none"> • Wielkość wyświetlacza • Radio • Aparat fotograficzny • Wymiary (długość, szerokość) • Ciężar telefonu • Czas pracy (bateria) • Sygnał wibracyjny • Wielkość klawiatury 	<ul style="list-style-type: none"> • Wielkość łazienki • Odpowiednie piętro • Właściwa technologia • Piwnica • Wielkość mieszkania • Czas dojazdu do centrum • Wielkość kuchni • Lokalizacja 	<ul style="list-style-type: none"> • Punktualność • Wolne miejsca siedzące • Czas oczekiwania na pojazd • Cena biletu • Brak tłoku • Czystość • Bezpieczeństwo • Szybkość poruszania się
Drukarka (D)	Laptop (L)	
<ul style="list-style-type: none"> • Rozdzielczość • Czas wydruku strony • Możliwość druku w kolorze • Możliwość druku obustronnego • Głośność pracy • Wymiary • Koszty eksploatacji • Czas bezawaryjnej pracy 	<ul style="list-style-type: none"> • Pojemność dysku • Szybkość pracy (procesor) • Wielkość pamięci operacyjnej • Kamera • Odpowiednia wielkość ekranu • Odpowiedni kolor obudowy • Liczba miesięcy gwarancji • Ciężar 	

Źródło: Opracowanie własne.

5.2. Wyniki eksperymentu

Dla każdej liczby kryteriów obliczono frakcję ankiet, dla których $CR > 0,1$ według wzoru:

$$W_n = \frac{N_{n,f}}{N_n} \quad (5)$$

gdzie:

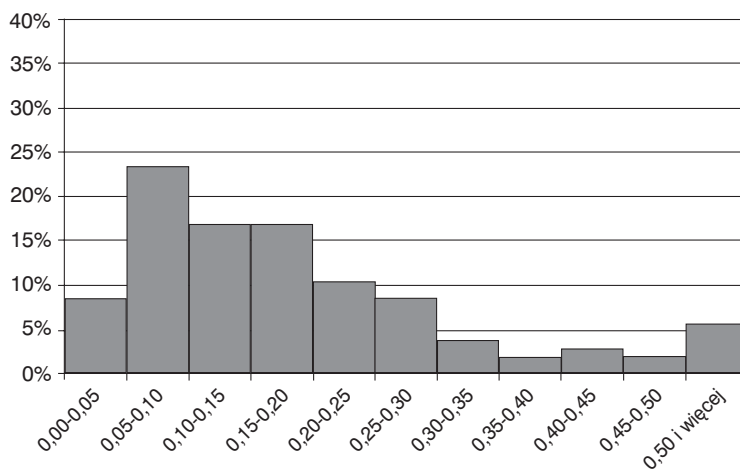
n — liczba kryteriów

W_n — frakcja ankiet w których $CR > 0,10$

$N_{n,f}$ — liczba ankiet, dla których $CR > 0,10$

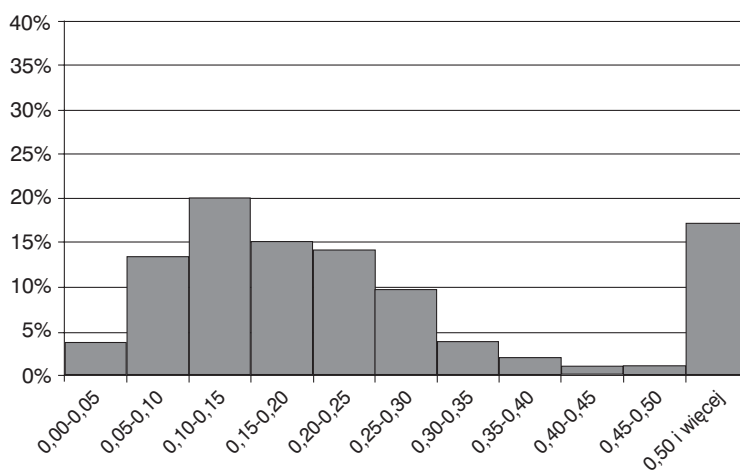
N_n — liczba wszystkich ankiet.

Na rysunkach rys. 2 — rys. 5 zostały przedstawione empiryczne rozkłady współczynnika zgodności dla różnej liczby kryteriów (5, 6, 7, 8).



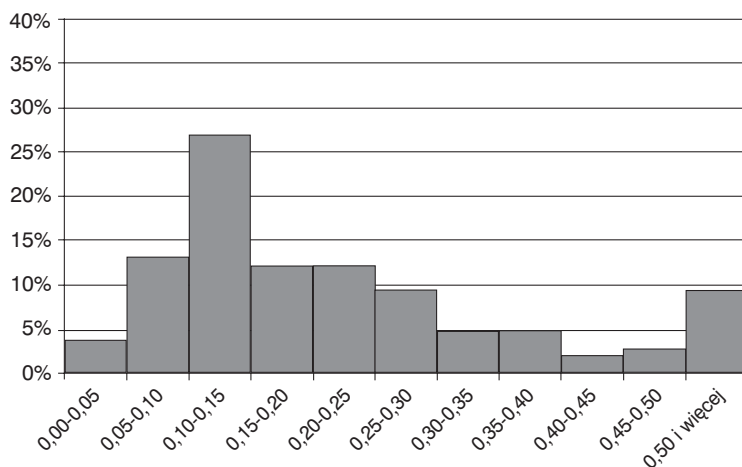
Źródło: Opracowanie własne.

Rys. 2. Rozkład współczynnika CR dla 5 kryteriów (przedział 0,05)



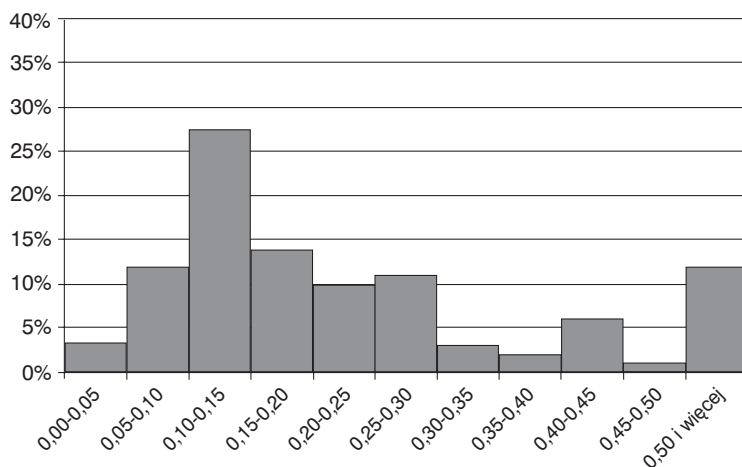
Źródło: Opracowanie własne.

Rys. 3. Rozkład współczynnika CR dla 6 kryteriów (przedział 0,05)



Źródło: Opracowanie własne.

Rys. 4. Rozkład współczynnika CR dla 7 kryteriów (przedział 0,05)



Źródło: Opracowanie własne.

Rys. 5. Rozkład współczynnika CR dla 8 kryteriów (przedział 0,05)

Przedstawione powyżej wykresy wymagają komentarza. Przede wszystkim można zauważyć, że odsetek niezgodnych porównań rośnie wraz ze wzrostem liczby kryteriów. Dla 6, 7 oraz 8 kryteriów (rys. 3, rys. 4 oraz rys. 5) najczęściej wartość wskaźnika niezgodności znajduje się w przedziale pomiędzy 0,1 a 0,15.

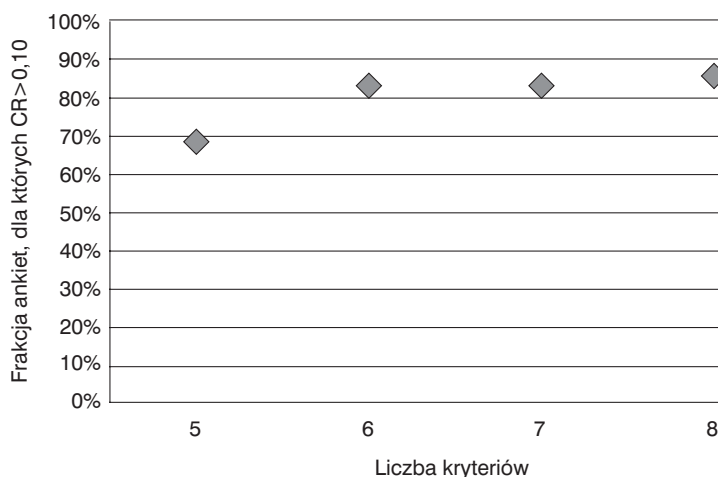
Dla wszystkich badań charakterystyczną właściwością jest występowanie wartości CR większych od wartości 0,5.

Wyniki eksperymentu dla 5, 6, 7 i 8 kryteriów zestawiono w sposób zbiorczy w tabeli 3 jako frakcje ankiet, dla których wartość CR jest większa od zakładanego przez Saaty'ego poziomu 0,10. Rys. 6 przedstawia wykres frakcji niezgodnych ocen w zależności od liczby porównywanych kryteriów. Jak się spodziewano, im wyższa liczba kryteriów, tym wyższa frakcja niezgodnych ankiet. Dla 5 kryteriów poziom ten wynosi 68,22%, co budzi wątpliwości dotyczące założenia metody AHP mówiącego, że człowiek jest zdolny do efektywnego dokonania porównań 7 ± 2 elementów. Jak ponadto wykazano, w przypadku większej liczby kryteriów (od 6 wzwyż), frakcja ta sięga ponad 80%, jednakże różnice w liczbie niezgodnych ankiet dla 6, 7, i 8 kryteriów są bardzo niewielkie ($W_6 = 83,02\%$, $W_7 = 83,33\%$, $W_8 = 85,29\%$).

Tabela 3

Wyniki badań dla poszczególnej liczby kryteriów jako frakcja ankiet $CR > 0,10$

	5 kryteriów	6 kryteriów	7 kryteriów	8 kryteriów
Liczba ankiet poprawnie wypełnionych (N_n)	107	106	108	102
Liczba ankiet z $CR > 0,10$ ($N_{n,f}$)	73	88	90	87
Frakcja ankiet z $CR > 0,10$ (W_n)	68,22%	83,02%	83,33%	85,29%



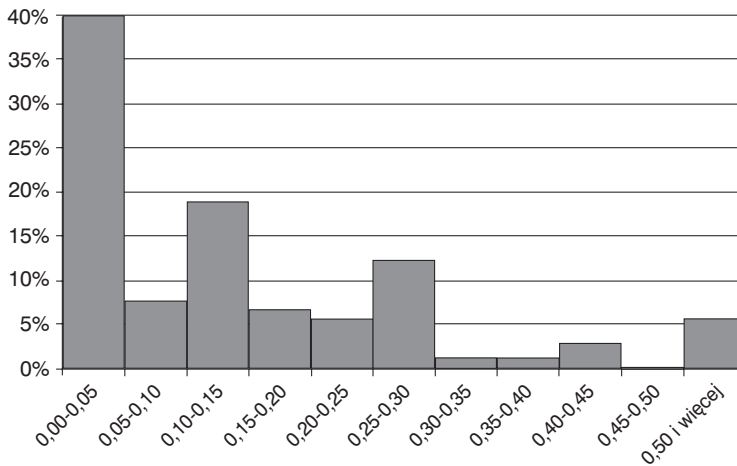
Źródło: Opracowanie własne.

Rys. 6. Frakcja niezgodnych ankiet w zależności od liczby porównywanych kryteriów

Wyniki powyższego eksperymentu skłoniły autorów do zadania sobie pytania, jakie wyniki otrzymamy przy niższej liczbie kryteriów. Dlatego też przeprowadzono eksperyment uzupełniający, którego celem było zbadanie, jak wysoka jest frakcja niezgodnych porównań w przypadku 3 kryteriów. Jako badany obiekt wybrano sklep spożywczy, który był analizowany ze względu na trzy kryteria:

- Szybkość obsługi,
- Łatwość dojazdu,
- Wybór produktów.

Wyniki pokazały, że dla 108 poprawnie wypełnionych ankiet frakcja niezgodnych ocen, obliczona na podstawie wyżej przedstawionego wzoru, wyniosła aż 52,78%. Rozkład współczynnika CR został natomiast przedstawiony na rys. 7.



Źródło: Opracowanie własne.

Rys. 7. Rozkład współczynnika CR dla 3 kryteriów (przedział 0,05)

6. WNIOSKI

W niniejszej pracy przedstawiono, na podstawie badań empirycznych, skalę występowania jednego z głównych problemów związanych ze stosowaniem metod AHP/ANP, a mianowicie trudność w uzyskaniu zgodnych (spójnych) wyników, nawet przy niewielkiej liczbie porównywanych elementów. Początkowe założenie, że dla liczby kryteriów większej od 5 liczba sprzecznych (niezgodnych) ocen będzie bardzo wysoka (większa od 30%), zostało potwierdzone. Co więcej, badania wykazały, że bardzo wysoka frakcja niespójnych odpowiedzi (ponad 50%) występuje już przy analizie 3 kryteriów, natomiast w przypadku 8 kryteriów przekracza 80%. Rezultaty eksperymentu pokazały również, iż w przypadku

większej liczby kryteriów różnice we frakcji liczbie niezgodnych ankiet dla 6, 7, i 8 kryteriów są bardzo niewielkie. Konieczne są zatem dalsze badania w tym zakresie, w szczególności biorąc pod uwagę takie czynniki, jak:

- a) wielkość próby; celem uzyskania miarodajnych wyników, badania należy powtórzyć na większej próbie,
- b) dobór ankietowanych do badania; w przedstawionych eksperymentach, na podstawie badania nielosowej próby krakowskich studentów wnioskowano o całą populację,
- c) wielkość ankiety; przeprowadzana ankieta była zbyt obszerna, gdyż oceniający otrzymywał jednorazowo ponad 20 stron (na każdej stronie były trzy porównania), co mogło powodować zmęczenie i w konsekwencji „losowe” odpowiedzi.

Biorąc pod uwagę podstawową właściwość operacyjną metody AHP, jaką jest generowanie niesprzecznych wyników, problem ten wymaga dalszych badań i precyzyjnego zdefiniowania warunków, przy których metodę tę należy stosować.

BIBLIOGRAFIA

- Adamus W. 2006. *Wejście Polski do strefy euro w ujęciu korzyści, kosztów, szans i ryzyka*, [w:] Nowe wyzwania w zarządzaniu organizacjami (red. W. Adamus), *Acta Academiae Modrewniana*, Kraków 2006.
- Adamus W., Łasak P. 2010. *Zastosowanie metody AHP do wyboru umiejscowienia nadzoru nad rynkiem finansowym*, *Bank i Kredyt*, 41(4): 73–100.
- Akarte M.M. et al. 2001. *Web based casting supplier evaluation using Analytic hierarchy process*, *Journal of the Operational Research Society*, 52(5): 511–522.
- Alonso J.A., Lamata M.T. 2005. *A Statistical Criterion of Consistency in the Analytic Hierarchy Process*, [w:] V. Torra et al. (Eds.), *MDAI 2005*, pp. 67–76, Springer-Verlag Berlin Heidelberg.
- Alonso J.A., Lamata M.T. 2006. *Consistency in the analytic hierarchy process: a new approach*, *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, 14(4): 445–459.
- Apostolou B., Hassel J.M. 1993. *An empirical examination of the sensitivity of the analytic hierarchy process to departures from recommended consistency ratios*, *Mathematical and Computer Modeling*, (4/5): 163–170.
- Apostolou B., Hassel J.M. 2002. *Note on Consistency Ratio: A Reply*, *Mathematical and Computer Modeling*, (35): 1081–1083.
- Arbel A., Orger Y.E. 1990. *An application of AHP to bank strategic planning: The merger and acquisitions process*, *European Journal of Operational Research*, 48(1): 27–37.
- Badri M. 1999. *Combining the AHP and GP for global facility location-allocation problem*, *International Journal of Production Economics*, 62(3): 237–248.
- Belton V., Gear T. 1983. *On a shortcoming of Saaty's method of analytic hierarchies*, *Omega*, 11(3): 228–230.
- Beynon M.J. 2005. *A method of aggregation in DS/AHP for group decision-making with the non-equivalent importance of individuals in the group*, *Computers and Operations Research*, 32: 1881–1896.
- Costa J.E., Wanderley A.J.M., Cosenza C.A.N. 2006. *A proposition to solve inconsistency problem in decision matrices using genetic algorithms*, third International Conference on Production Research — Americas' Region 2006 (ICPR-AM06).

- Ergu D., Kou G., Peng Y., Shi Y. 2001. *A Simple Method to Improve the Consistency Ratio of the Pair-wise Comparison Matrix in ANP*, European Journal of Operational Research, In press.
- Finan J.S., Hurley W.J. 1997. *The Analytic Hierarchy Process: Does adjusting a pairwise comparison matrix to improve the consistency ratio help?*, Computers and Operations Research, 24(8): 749–755.
- Forman E., Peniwati K. 1998. *Aggregating individual judgments and priorities with the Analytic Hierarchy Process*, European Journal of Operational Research, 108: 165–169.
- Holder R.D. 1990. *Some Comments on the Analytic Hierarchy Process*, The Journal of the Operational Research Society, 41(11): 1073–1076.
- Chin K.S., Chiu S., Tummala V.M.R. 1999. *An evaluation of success factors using AHP to implement ISO 14001 based EMS*, International Journal of Quality & Reliability Management, 16(4): 341–361.
- Chu P., Liu J.K.-H. 2002. *P. Chu and Note on consistency ratio*, Mathematical and Computer Modeling, 35: 1077–1080.
- Lai V., Wong B.K., Cheung W. 2002. *Group decision making in a multiple criteria environment: A case using the AHP in the software selection*, European Journal of Operational Research, 137(1): 134–144.
- Miller D. 1956. *The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information*, The Psychological Review, 63: 81–97.
- Perez J., Jimeno J.L., Mokotoff E. 2001. *Another potential strong shortcoming of AHP* (<http://ideas.repec.org/p/alc/alcdtd/8-02.html>).
- Saaty T. 1994. *Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process*, Pittsburgh, PA: RWS Publications.
- Saaty T.L. 1996. *Decision Making for Leaders: The Analytical Hierarchy Process for Decisions in a Complex World*, The Analytical Hierarchy Process Series, 2: 71–74.
- Saaty T. 2001. *The seven pillars of the analytic hierarchy process*, (www.creativedecisions.net).
- Saaty T. Ozdemir M. 2003. *Why the magic number seven plus or minus two*, Mathematical and Computer Modeling, 38: 233–244.
- Saaty T.L. 2008. *Relative Measurement and Its Generalization in Decision Making Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors The Analytic Hierarchy/Network Process*, RACSAM Rev. R. Acad. Cien. Serie A. Mat., 102(2): 251–318.
- Singpurwalla N., Forman, E., Zalkind D. 1999. *Promoting shared health care decision making using the analytic hierarchy process*, Socio-Economic Planning Sciences, 33(4): 277–299.
- Solms, S. 2009. *Homogeneity and choice aggregation in the analytic hierarchy process*, ISAHIP, Pittsburgh, Pennsylvania, USA, 29 July-1 August 2009.
- Stefanów P., Prusak A. 2011. *Badanie wiarygodności i skuteczności skali porównań Saaty'ego w metodzie AHP i ANP, Przedsiębiorcze aspekty organizacji i biznesu*, pod red. A. Chodyńskiego.
- Szczypińska A., Piotrowski E.W. 2009. *Inconsistency of the judgment matrix in the AHP method and the decision maker's knowledge*, Physica A, 388: 907–915.
- Tung S.L., Tang S.L. 1998. *A comparison of the Saaty's AHP and modified AHP for right and left eigenvector inconsistency*, European Journal of Operational Research, 106: 123–128.
- Vaidya O.S., Kumar S. 2006. *Invited Review Analytic hierarchy process: An overview of applications*, European Journal of Operational Research, 169: 1–29.
- Saaty T.L., Vargas L.G. 1982. *The Logic of Priorities*, Kluwer Nijhoff Publishing, Massachusetts.
- Watson S.R., Freeling A.N.S. 1982. *Assessing Attribute Weights*, Omega, 10: 582–583.
- Weathers D., Sharmab S., Niedricha R.W. 2005. *The impact of the number of scale points, dispositional factors, and the status quo decision heuristic on scale reliability and response accuracy*, Journal of Business Research, 58: 1516–1524.
- Webber S.A., Apostolou B., Hassel J.M. 1996. *The sensitivity of the analytic hierarchy process to alternative scale and cue presentations*, European Journal of Operational Research, 96: 351–362.
- Weber M. 1997. *Remarks on the paper "On the Measurement of Preferences in the Analytic Hierarchy Process"*, Journal of Multicriteria Decision Analysis, 6: 320–321.

REVIEW OF THE MONOGRAPH ENTITLED
“POLAND-SERBIA: The Challenges of the Scientific Cooperation”

Prof. LECH CZARNECKI, PhD, D.Sc.
Warsaw University of Technology

It is my pleasure to make remarks on the monograph entitled “POLAND-SERBIA: The Challenges of the Scientific Cooperation” published by Acta Academiae Modrevianae (Kraków, 2010) under scientific editorial supervision of Zbigniew Paszek. The book which lies in front of me has an elegant three-coloured cover – white, blue and red. Those are Serbian national colours. Therefore, there is also included the Polish “badge” in a symbolic way. The book contains the collection of 21 academic articles written by 28 authors — 20 from Serbia and 8 from Poland. The book has more than 300 pages. The title of the monograph (suggested by Slobodan Malinić) is adequate to the content because there are *challenges* but not only and not solely. The papers are arranged alphabetically according to the authors` names. However, behind the alphabetic courtesy someone could find several leading ideas in the book which organise the content. There are:

the fundamental civilisation idea:

the Paradigm of Sustainable Development has been presented by Biljana Jovanović Gavrilović. The history of sustainable concept since the famous report (1987) of the Brundtland Commission entitled “Our Common Future” has been described. It has been characterized as a long debate, over the last fifty years, which reflected the evolution of thinking development; from economic growth to sustainable development. The author considers various implications of the classic definition: *a development that meets the needs of the present without compromising the ability of future generations to meet their own needs*. The author’s statement that *the realisation of sustainable development is a true challenge for each country* could be the motto of the monograph. Not only has the essence of the concept been

presented but also the interdependence of the three key dimensions: economic, social and environmental. The author stressed that the measurements of sustainable development are a task of particular complexity and significance. *Sustainable development requires an integrated set of policies to jointly work towards improvement of human welfare. This way conditions will be created for the interaction between the key dimensions of sustainable development to be taken into account in the phase of designing different political interventions. Measures undertaken with the aim to affect one of the three dimensions (economic, social or environmental), can cause unforeseen consequences and costs without a good study of their effect on the other two. Responsibility to achieve sustainable development lies to a great extent in the hands of state, but other institutions, from the domain of private sector and civil society, should also become active partners in that process. Each of these management holders has a unique and important role in promotion of sustainable development.*

The Biljana Jovanović Gavrilović's words mentioned above made the wide spread framework for other topics presented in the monograph on various levels of generalisation and complexity.

the diagnosis of the current status:

Objective Dependence Cooperation Between the Economy and the State by Vlastimir Leković,

On Serbian Economic Competitiveness in Transitional Conditions by Ljiljana Maksimović,

Economic Crisis and Crisis in Serbian Enterprises – New Challenges for the Government, Managers and Accountants by Slobodan D. Malinić,

Results Achieved in Meeting Convergence Criteria in Serbia as a Necessary Condition for Joining EMU by Dragana Marković and Srdjan Furtula,

Internal Audit Analysis on the Example of Joint Stock Company Variant from Cracow by Vesna Milovanović,

Marketing in Poland in the Twentieth Year of Political Transformation by Roman Niestrój,

The Analysis of Business Success in Tourism by Gordana Radosavljević,

Export Performances of National Economy in the Conditions of Global Economic Crisis by Petar Veselinović.

quest and basic questions:

Can education be the key to welfare? What can a person educated in new times do? Is the intelligentsia heading towards the liberal or etatistic orientation? (Jerzy Mikułowski Pomorski).

Quest in the right way: Polish science and higher education against international background. *There is the long list of sins which also included underestimating the importance and role of competition both in the fields of science and education. Finally, the conclusion: The European Union, and Poland with it, will not overtake the USA*

or Asian countries if the foundation of the basic solutions both in economy and in education and science is not formed by healthy competition, immanent competition system in accessing R&D financing sources and the conviction that knowledge-based economy is an economy based on freedom, creativity and equality (Tadeusz Pomianek, Andrzej Rozmus).

How to make the two-way effective leader communication successful? The answer is that it is not exclusively *the matter of intellect and logic, but of trust, emotions and accepting difference of other people and the efforts to comprehend the intention of the communication, with no prejudice and priori refusing the message content. The ability of keeping such communication is of an essential importance in the process of organizational changes which request changes at personal level too, with efforts to understand and appreciate individual distinctions* (Vesna Stojanović Aleksić, Srdjan Šapić).

What is the goal of Serbian policy makers in the conditions of global economic crisis? How to create preconditions for increasing export possibilities? Final conclusion: *Without creating competitive advantages and technology development Serbian economy will not be involved in world trade with opportunity to obtain significant export results* (Petar Veselinović).

models and means to aim:

Innovation Processes Through SME's Networking by Verica Babić and Jelena Erić,
Regional Development Incentives in Serbia by Stevan Devetaković

The Concept of Corporate Social Responsibilities – the Impact on Practice by Momčilo Djordjević and Srdjan Šapić,

An Organic Model of Strategic Management by Marek Lisiński,

Alternative Sources of Financing Economic Development by Nikola Makojević,

Financial Liquidity Management by Predrag Stančić and Vladimir Stančić.

Possibilities of the Application of Pricing Strategies in the Development of Trade between Poland and Serbia (Klemens Budzowski) is one of the most valuable articles in the book. The paper presents the results of analyses according to which the trade exchange between Poland and Serbia is not satisfactory, especially in the context of historical, geographical backgrounds as well as objective possibilities. According to the Author the choice of appropriate price strategies could actuate further development and growth in the trade exchange between both countries. Advantages and disadvantages of high and low price strategy, concentration-based pricing strategy as well as imitation strategy have been well exemplified. A very interesting problem of price paradoxes such as: Giffen's, Veblen's, imitation and snob effect, income effect of price reduction, shock and timing effect, speculative effect and others were also presented.

Very peculiar place among other articles has been devoted to Benford's Law (Tadeusz Grabiński) which could be treated even as a general law of nature. From another point of view the Benford Law is of practical meaning and is utilized in

different science and real life domains: economic studies, quantitative research, earth sciences, technical studies and information technology studies.

It is worth emphasizing that published articles are particularly addressed to one of the title countries: Poland (J. Mikułowski-Pomorski, V. Milovanović, R. Niestrój, T. Pomianek, A. Rozmus), Serbia (Stevan Devetaković, N. Makojević, L. Maksimović, S.D. Malinić, D. Marković, S. Furtula, P. Veselinović, E. Vuksanović, R. Stefanović) or addressed to both countries (K. Budzowski). There are also more universal considerations (V. Babić, J. Erić, M. Djordjević, S. Šapić, T. Grabiński, B. Jovanović Gavrilović, V. Leković, M. Lisiński, G. Radosavljević, P. Stančić, V. Stančić, V. Stojanović Aleksić).

The monograph is dedicated to Professor Ilija Rosić, Rector of Kragujevac University (1990–1992) and Dean of the Faculty of Economics (1987–1990; 2000–2002). Who is the man to whom 28 authors from two countries addressed their papers. The answer can be found in two papers on the first pages of the book: *Life and Work of Ilija Rosić. Education and Scientific Achievements* by Slobodan D. Malinić and *Ilija Rosić — A Man Dedicated to Science, Education and Academic Cooperation with Polish Universities* by Zbigniew Paszek. The essence of Ilija Rosić personality is described by Zbigniew Paszek in the following words: *professor Ilija Rosić was always full of energy and enthusiasm in any work he undertook; very well known for being good-natured, for his optimism and cheerfulness as well as his witty comments.*

I am a little bit embarrassed that this review is so uniformly positive. However, I should mention that I am also emotionally involved because in 2003, as the Vice-Rector for Academic Affairs of Warsaw University of Technology, I had an opportunity to meet prof. Ilija Rosić as I organized his lecture for students of our Business School. Right now I would like to make very sincere appreciation to Zbigniew Paszek — doctor honoris causa of the Kragujevac University in the name of all who knew Ilija Rosić for his great job as the scientific editor of the book. The book is invaluable as is the tribute to the fascinating Person but firstly it defines the modern civilisation challenges in an interesting and precious manner.