

Sławomir Śmiech

Uniwersytet Ekonomiczny w Krakowie

ŹRÓDŁA ZMIENNOŚCI STÓP ZWROTU AKCJI NOTOWANYCH NA GPW

1. Wstęp

Modele wieloczynnikowe wykorzystywane do opisywania stóp zwrotu instrumentów finansowych są dzielone na trzy grupy, tj. czynnikowe modele makroekonomiczne, fundamentalne oraz statystyczne modele czynnikowe.

W statystycznym modelu czynnikowym zmienne objaśniające nie są bezpośrednio obserwowalne. Dlatego budowa tego typu modeli sprowadzać się musi do dwóch etapów. W pierwszym należy wyznaczyć (wyodrębnić) z pewnego zbioru danych odpowiednie (interesujące z jakiegoś punktu widzenia) zmienne. Zastosowanie w tym celu metody głównych składowych powinno skutkować tym, że otrzymane zmienne w znacznym stopniu będą kształtować wariancję zmiennych objaśnianych oraz będą ortogonalne. Dzięki temu mogą być rozumiane jako niezależne źródła informacji wpływające na modelowane zmienne.

W artykule przedstawione zostaną podstawy teoretyczne budowy statystycznego modelu czynnikowego dla stóp zwrotu. Następnie zaprezentowane zostaną wyniki analizy służącej wyodrębnieniu czynników w podzbiorach akcji będących składnikami indeksów WIG20 oraz WIG40.

2. Statystyczny model czynnikowy dla stóp zwrotu

Założmy, że mamy notowania N spółek w kolejnych T jednostkach czasu. Niech r_{it} oznacza logarytmiczną stopę zwrotu i -tej akcji w chwili t . Ogólna postać modelu czynnikowego przyjmuje postać [Tsey 2005, s. 406]:

$$r_{it} = \alpha_i + \beta_{i1}f_{1t} + \dots + \beta_{iK}f_{Kt} + \varepsilon_{it}, \quad t=1, \dots, T; \quad i=1, \dots, N, \quad (1)$$

gdzie α_i to wyraz wolny (przecięcie), f_{it} wartość i -tego czynnika w chwili t , β_{ij} to tzw. ładunek czynnikowy (*factor loading*) i -tej akcji dla j -tego czynnika, zaś ε_{it} to składnik losowy. Specyfikacja równania (1) wymaga przyjęcia określonych założeń, które można znaleźć w cytowanej wyżej pracy.

W statystycznym modelu czynnikowym zmienne objaśniające (równanie 1) nie są obserwowalne. Ich wyodrębnieniu służy metoda czynników głównych (w skrócie PCA – *principal components analysis*). Można ją stosować w przypadku, gdy liczba okresów T , w których obserwujemy stopy zwrotu, jest większa od liczby analizowanych walorów N ($T > N$). W przeciwnym razie macierz kowariancji stóp zwrotu obliczona dla próby jest osobliwa, czyli nieodwracalna. Punktem wyjścia do wyodrębnienia czynników w metodzie PCA jest analizowanie empirycznej macierzy kowariancji stóp zwrotu określonej jako

$$\hat{\Omega}_N = \frac{1}{T} RR' . \quad (2)$$

Szacowanie współczynników regresji (ładunków czynnikowych) przebiegać może następującymi metodami: MNK, UMNK lub metodą największej wiarygodności. Ponieważ wykorzystywane są tutaj dane przekrojowo-czasowe, wymaga się od danych, aby miały dostatecznie długą, ale również stabilną historię.

W statystycznym modelu czynnikowym dla stóp zwrotu zmienne ukryte (czynniki) definiuje się następująco:

$$f_i = \sum_{j=1}^N \gamma_{ji} R_j . \quad (3)$$

Czynniki są więc kombinacjami liniowymi wszystkich analizowanych stóp zwrotu i charakteryzują się następującymi własnościami:

- a) są ortogonalne,
- b) $\text{var}(f_1) + \dots + \text{var}(f_N) = \text{var}(R_1) + \dots + \text{var}(R_N)$,
- c) $\text{var}(f_1) \geq \dots \geq \text{var}(f_N)$.

Przydatność wyodrębnionych zmiennych ukrytych w kontekście analizy stóp zwrotu i ich ryzyka oparta jest na przedstawionych poniżej własnościach.

- Wyodrębnione zmienne ukryte *a posteriori* mogą mieć ekonomiczną interpretację. Wtedy „wartością dodaną” przeprowadzonej procedury jest znalezienie nowych zmiennych, które mają udział w objaśnianiu wariacji stóp zwrotu.
- Wyodrębnione czynniki mogą być traktowane jako niezależne (są ortogonalne) źródła (przyczyny) zmienności stóp zwrotu.
- Ponieważ każda kolejna zmienna ukryta objaśnia coraz mniej, można zatem zrezygnować z części najmniej pożytecznych zmiennych ukrytych. W ten sposób można ograniczyć liczbę parametrów w docelowej analizie regresji.

Metoda PCA pozwala zastąpić zbiór zmiennych oryginalnych nowym zbiorem czynników głównych. Jeśli zostaną wykorzystane wszystkie czynniki główne, wtedy model regresji będzie wnosił dokładnie taką samą informację jak model, w którym zmienne objaśniające to oryginalne stopy zwrotu. Korzyść z zastosowania PCA polega na tym, że duży zbiór oryginalnych stóp zwrotu zostanie zastąpiony mniejszym zbiorem czynników głównych¹. Pozostaje kwestia, ile czynników (zmiennych ukrytych) powinno być wykorzystanych w analizie? Aby odpowiedzieć na to pytanie, można się posłużyć zestawionymi poniżej kryteriami:

- regułą „wartości własnej (λ) większej od jedności” (kryterium Kaisera),
- metodą procentu wariancji tłumaczonej przez czynniki główne,
- metodą testu ospiska,
- testem LR (*likelihood ratio test*), jeśli wykorzystuje się metodę MNW (por. [Ziwot 2002, s. 582]).

Interpretacja znaczenia czynników w modelu regresji jest możliwa dzięki współczynnikom regresji, które w metodzie PCA są nazywane ładunkami czynnikowymi. Wyrażają one stopień „nasycenia” ustalonej stopy zwrotu danymi czynnikami. Dodatkowo mogą być rozumiane jako współczynniki korelacji zmiennej pierwotnej z poszczególnymi czynnikami. Ortogonalne czynniki nie są wyznaczone w sposób jednoznaczny. Jeśli weźmiemy dowolną ortogonalną macierz H , taką że $H' = H^{-1}$, to możemy rozważany model regresji zapisać następująco:

$$R_t = \mu + BHH' f_t + \varepsilon_t = \mu + B^* f_t^* + \varepsilon_t. \quad (4)$$

Można zatem powiedzieć, że czynniki są interpretowane jednoznacznie z dokładnością do transformacji ortogonalnej. Ta własność pozwala obracać (rotować) czynniki tak, aby ułatwić ich identyfikację i interpretację. W kontekście statystycznych modeli czynnikowych dla stóp zwrotu najbardziej odpowiednia wydaje się rotacja *varimax* [Abdi 2007, s. 4 i dalsze].

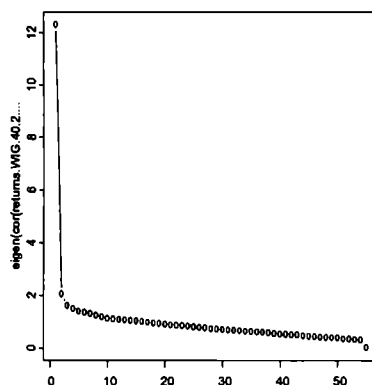
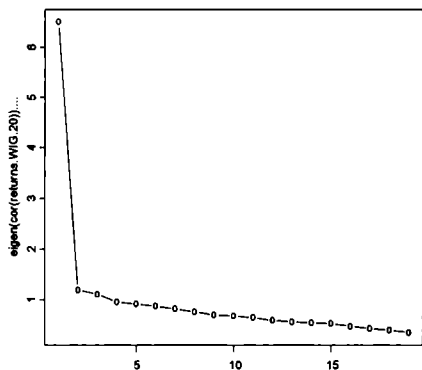
3. Wyodrębnianie czynników dla zbioru stóp zwrotu akcji WIG20 i WIG40

Niniejsza część poświęcona będzie przedstawieniu wyników budowania statystycznych modeli czynnikowych dla stóp zwrotu akcji notowanych na GPW. Przeprowadzana analiza ma umożliwić odpowiedzi na następujące pytania:

- Czy można wyodrębnić niezależne (ortogonalne) czynniki, które będzie można interpretować jako niezależne źródła przyczyny zmian stóp zwrotu notowanych akcji?
- Czy czynniki te zależą od pojemności zbioru rozważanych akcji?
- Jaka jest moc objaśniania wariancji stóp zwrotu przez wyznaczone czynniki?

¹ Jak dowodzi Maddala [Maddala 2006, s. 320], wymaga to użycia pewnej wiedzy *a priori*.

Analiza została przeprowadzona dla dwóch zbiorów walorów. Pierwszy obejmował akcje wchodzące w skład WIG20. Pominięta została spółka CEZ (była zbyt krótko notowana). Wymiar czasowy analizy obejmował okresy wspólnych notowań walorów z całego badanego zbioru (481 dni roboczych pomiędzy 16.05.2005 a 27.08.2007). Drugi zbiór zawierał akcje wchodzących w skład indeksów WIG20 oraz WIG40. W tym przypadku analizowano stopy zwrotu 56 spółek, a okres analizy był identyczny jak dla pierwszego zbioru akcji.



Rys. 1a. Test osypiska dla akcji WIG20

Rys. 1b. Test osypiska dla akcji WIG20 i WIG40

Źródło: opracowanie własne.

W obu przypadkach analiza przeprowadzana była w ramach następujących etapów:

1. Określenie odpowiedniej liczby czynników.
2. Próba interpretacji.
3. Ewentualne zrotowanie czynników i ponowna próba interpretacji

W celu określenia dostatecznej liczby czynników do opisu macierzy kowariancji stop zwrotu akcji ze zbioru WIG20 wykorzystano test osypiska, regułą wartości własnej, a także obliczono wartość p -value dla testu LR .

Rys 1a przedstawia test osypiska dla walorów ze zbioru WIG20. Nie widać na wykresie wyraźnych załamania krzywej, które mają wskazywać, jaką liczbę czynników wyodrębnić. Przedstawiony wykres pozwala za to zauważyć, że wartości własne 4 pierwszych czynników są większe od jedynki. Test LR , którego wartości p -value przedstawia tab. 1, również sugeruje, że należy wziąć do analizy 4 czynniki.

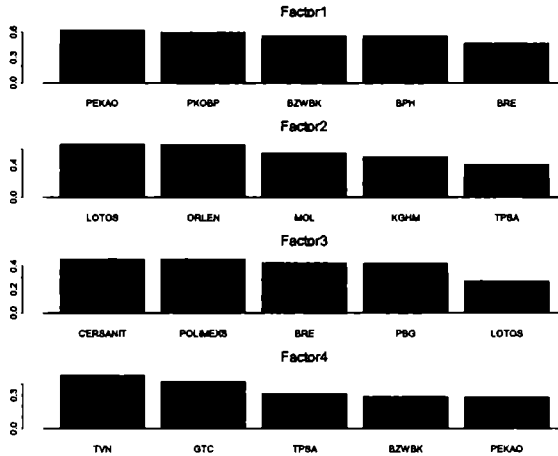
Tabela 1. Wartości własne oraz wartość p -value dla testu LR odpowiadające rosnącej liczbie czynników. Wyniki dla zbioru stop zwrotu akcji WIG20

Liczba czynników	2	3	4	5	6
p -value	0	0,05	0,124	0,311	0,605

Źródło: opracowanie własne.

Analiza ładunków czynników (ich wartości nie zostały tutaj zestawione ze względu na ograniczoną objętość tekstu) miała umożliwić interpretację wyodrębnionych czynników. Interpretacja ta okazała się możliwa tylko dla pierwszego czynnika, który charakteryzował się tym, że miał dodatnią korelację ze wszystkimi stopami zwrotu. Można go więc interpretować jako ryzyko rynkowe (w tym przypadku utożsamiane z indeksem WIG20).

Z powodu trudności z interpretacją analizowane czynniki zostały „obrócone” przy użyciu rotacji *varimax*.



Rys. 2. Zmienne najbardziej skorelowane ze zrotowanymi czynnikami (WIG20)
Źródło: opracowanie własne.

Zbiór akcji najsilniej związany z poszczególnymi czynnikami został przedstawiony na rys. 2. Pierwszy czynnik jest najsilniej związany ze stopami zwrotu akcji sektora bankowego. Największe wartości ładunków czynnikowych (współczynniki korelacji) ma pierwszy czynnik dla stóp zwrotu wszystkich występujących w indeksie WIG20 banków (PEKAO, PKOBP, BZWBK, BPH, BRE). Drugi czynnik to z kolei spółki sektora paliwowego (Lotos, Orlen, Mol). Trzeci silnie zależy od spółek branży budowlanej lub deweloperskiej (CERSANIT, POLIEXS, PBG). Możemy założyć, że czwarty składnik odpowiada za informacje z pozostałych branż.

W dalszej kolejności przeprowadzono analizę dla stóp zwrotu akcji pochodzących z indeksów WIG20 oraz WIG40. Podobnie jak poprzednio posłużono się metodami „testu ospyska” oraz testu LR w celu wyznaczenia „odpowiedniej” liczby czynników. Otrzymane wyniki zostały przedstawione na rys. 1b oraz w tab. 2.

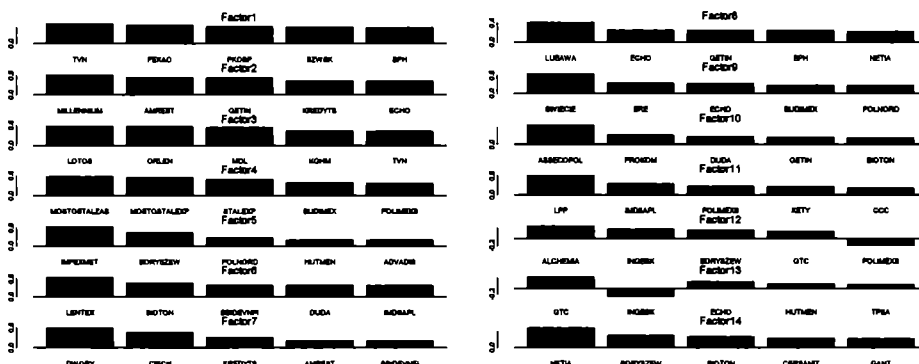
Analogiczne jak poprzednio rozumowanie pozwoliło wyodrębnić 14 pierwszych głównych składowych – 14 czynników. W takim wypadku (podobnie zresztą jak poprzednio) interpretacja kolejnych czynników okazała się niemożliwa. Jedyńm wyjątkiem był pierwszy czynnik, który tak jak uprzednio odzwierciedlał ryzyko rynkowe. Nie udało się znaleźć klucza, według którego można by nadać sens pozostałym czynnikom. Dlatego też dokonano rotacji czynników metodą *varimax*.

Tabela 2. Wartości p -value dla testu LR odpowiadające rosnącej liczbie czynników. Wyniki dla zbioru stóp zwrotu akcji z WIG20 i WIG40

Liczba czynników	1	...	11	12	13	14
p -value	0	...	0,005	0,018	0,047	0,123

Źródło: opracowanie własne.

Dzięki temu potencjalne interpretacje stają się łatwiejsze, gdyż można wyróżnić niewiele akcji dość silnie skorelowanych z danym czynnikiem i jednocześnie sporą część pozostałych walorów, których stopy zwrotu są od niego niezależne. Akcje, których stopy zwrotu są najsilniej skorelowane z poszczególnymi czynnikami, zostały przedstawione na rys 3a i 3b. Dwa pierwsze czynniki są zbudowane w dużej części na akcjach sektora bankowego (choć trzeba przyznać, że zależą one również od pewnych dość przypadkowych w tym kontekście akcji, jak np. TVN czy AMREST).

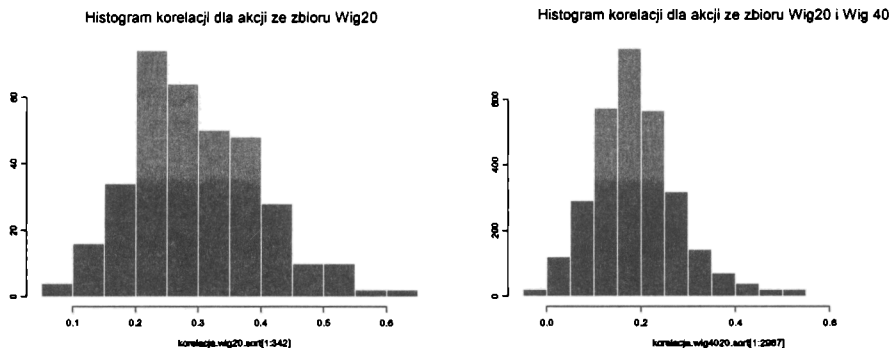


Rys. 3. Zmienne najsilniej skorelowane z czynnikami 1-14 (WIG20 i WIG40)

Źródło: opracowanie własne.

Kolejne dwa czynniki oparte są odpowiednio na spółkach sektora paliwowego oraz budowlanego. Dalsze czynniki charakteryzują się tym, że są silnie związane z jedną tylko spółką i są równocześnie niezależne od zmian stóp zwrotu pozostałych spółek. Taka charakterystyka czynników pokazuje, że nie można wyróżnić zbyt wielu grup spółek, których zmienność zależałaby od napływających wspólnych dla nich informacji.

Reasumując, wyniki przeprowadzonego badania skłaniają do następujących wniosków. Jeśli skupimy się na zbiorze stóp zwrotu największych spółek notowanych na GPW (WIG20), to możemy powiedzieć, że ich zmienność zależy od niezależnych informacji charakterystycznych dla sektorów, do których poszczególne spółki przynależą. Jeśli zbiór akcji rozszerzymy o walory uwzględnione w indeksie WIG40, to oprócz trzech pierwszych czynników (które możemy interpretować tak jak w mniejszym zbiorze) kolejne w zasadzie reprezentują poszczególne walory.



Rys. 4. Histogram współczynników korelacji dla stóp zwrotu akcji z indeksu WIG20 oraz WIG20 i WIG40

Źródło: opracowanie własne.

Na koniec można zadać pytanie: dlaczego w wypadku mniejszego zbioru danych nie udało się wyodrębnić grup spółek, które zależałyby od tych samych źródeł informacji? Prawdopodobna przyczyna tkwi we własnościach macierzy kowariancji dla stóp zwrotu obu zbiorów akcji.

Przeprowadzenie analizy głównych składowych jest zasadne wówczas, gdy poszczególne zmienne są ze sobą dość silnie skorelowane. Aby ocenić zależność pomiędzy spółkami, wyznaczono współczynniki korelacji i zestawiono je w postaci histogramów na rys. 4.

Przedstawione wykresy pokazują, że w wypadku zbioru akcji WIG20 dominują zależności korelacyjne, które są dość wysokie (powyżej 0,2). Przeważają tam pary, dla których współczynnik korelacji zawiera się w przedziale (0,2; 0,4). Inaczej jest w wypadku większego zbioru akcji. Tutaj dominują pary, dla których współczynnik korelacji wynosi ok. 0,2. W takim przypadku trudno jest wymagać, aby wyodrębnione składowe główne uwzględniały zmienność różnych (jak się okazuje – niezależnych) walorów. Na koniec należy dodać, że w przypadku WIG20 cztery czynniki objaśniły 41% całkowitej wariancji stóp zwrotu. W wypadku akcji ze zbioru WIG20 i WIG40 14 czynników objaśnia ok. 40% całkowitej wariancji.

4. Podsumowanie

Celem badania było ustalenie wspólnych dla różnych walorów, ale niezależnych źródeł zmienności stóp zwrotu spółek notowanych na GPW. Otrzymane wyniki dowodzą, że większość spółek najsilniej reaguje na specyficzne dla nich informacje. Wyjątek stanowią największe na giełdzie spółki. Udało się je pogrupować ze względu na ich przynależność do poszczególnych sektorów, tj. sektora bankowego, paliwowego oraz budowlanego. Można więc powiedzieć, że pojawiające się informacje albo wpływają w podobny sposób na wszystkie spółki (uwzględnio-

ne w indeksie WIG20) z jednego z wymienionych sektorów, albo oddziałują w sposób specyficzny na poszczególne walory.

Literatura

Abdi H., *Factor Rotations in Factor Analyses*, preprint, 2007.

Maddala G.S., *Ekonometria*, PWN, Warszawa 2006.

Tsay R.S., *Analysis of Financial Time Series*, Wiley&Sons, New Jersey 2005.

Ziwot E., Wang J., *Modeling Financial Time Series with S-Plus*, Springer-Verlag, New York 2002.

SOURCES OF VOLATILITY OF RETURNS IN WARSAW STOCK EXCHANGE

Summary

Multifactor models can be used to explain returns of financial instruments. There are three main groups of multifactor models i.e. fundamental, macroeconomic and statistical factor models. In statistical factor models factors are not directly observable. If one use principal component analysis to extract the factor realizations then it can be interpret as interdependent source of volatility of returns. This article presents results of extracting factors explaining volatility of returns of stocks on the Warsaw Stock Exchange.