

Anna Plich

Uniwersytet Łódzki

ZASTOSOWANIE MODELU LOGITOWEGO DO PROGNOZOWANIA KRYZYSÓW GIEŁDOWYCH NA PRZYKŁADZIE NEW YORK STOCK EXCHANGE

Streszczenie: Kryzys giełdowy jest przykładem zmiennej jakościowej, która może zostać przekształcona w zmienną ilościową. W takiej sytuacji przyjmuje ona dwie wartości. Aby modelować taką zmienną, należy użyć modelu binarnego wyboru i dobrać odpowiednią transformację (np. przekształcenie logitowe), która pozwoli na estymację parametrów modelu. W artykule zaprezentowano zastosowanie modelu logitowego do opisu i prognozowania sytuacji na New York Stock Exchange. W przeprowadzonym badaniu porównano prognozy wykorzystujące dwa alternatywne podejścia.

Słowa kluczowe: modele logitowe, giełda papierów wartościowych, prognozy ostrzegawcze

1. Wstęp

Gwałtowne spadki obserwowane na amerykańskiej giełdzie w końcu 2007 r. i na początku 2008 r. wywołały falę paniki i spekulacje. Pesymiści przewidywali krach na miarę wielkiego kryzysu lat trzydziestych ubiegłego wieku. Powstało więc pytanie: czy ówczesne wydarzenia na giełdzie w USA w istocie zasługiwały na to miano?

Kryzysy finansowe, walutowe i giełdowe są przykładami zjawisk, w których badaniu uzasadnione staje się użycie modelu dla zmiennych przyjmujących przeliczalną liczbę wartości. Do takich należy model logitowy.

Celem niniejszej pracy jest próba diagnozy sytuacji na amerykańskiej giełdzie na przykładzie New York Stock Exchange poprzez budowę modelu logitowego i sformułowanie na jego podstawie prognozy do końca 2009 roku¹.

¹ Podobne badania dotyczące kryzysów walutowych przeprowadzono m.in. w ramach prac Katedry Ekonometrii na Wydziale Ekonomiczno-Socjologicznym Uniwersytetu Łódzkiego [Milo, Kozera 2004].

2. Metodologia badania – charakterystyka modelu dwumianowego i transformacji logitowej

Punktem wyjścia² jest model dwumianowy, w którym zmienna może przyjmować dwie wartości:

$$Y_i = \begin{cases} 0, & \text{gdy dany wariant nie występuje,} \\ 1, & \text{gdy dany wariant występuje.} \end{cases}$$

Wartości zero i 1 są przyjmowane z określonymi prawdopodobieństwami, które są warunkowe ze względu na zmienne objaśniające. Można założyć, że owe prawdopodobieństwa są funkcjami zmiennych objaśniających i wektora parametrów α , związanego ze zmiennymi objaśniającymi:

$$P(Y_i = 1 | \mathbf{x}_i, \alpha) = F(\mathbf{x}_i, \alpha), \text{ oraz } P(Y_i = 0 | \mathbf{x}_i) = 1 - F(\mathbf{x}_i, \alpha).$$

Można udowodnić, że przy przyjęciu liniowej postaci funkcji $F(\mathbf{x}_i, \alpha)$ i zdefiniowaniu modelu regresji mamy do czynienia z heteroskedastycznością składnika losowego. W związku z tym klasyczna metoda najmniejszych kwadratów nie jest właściwym sposobem szacowania parametrów modelu dwumianowego. Problem może zostać rozwiązany przy użyciu uogólnionej metody najmniejszych kwadratów. Takie rozwiązanie nie gwarantuje jednak uzyskania wyników interpretowalnych w kategoriach prawdopodobieństwa. W celu uniknięcia otrzymania większych od jedności lub ujemnych wartości oszacowań zmiennej objaśnianej należy sformułować model tak, aby spełnione były warunki: $\lim_{\mathbf{x}_i \cdot \alpha \rightarrow -\infty} P(Y_i = 1 | \mathbf{x}_i) = 0$ oraz

$\lim_{\mathbf{x}_i \cdot \alpha \rightarrow +\infty} P(Y_i = 1 | \mathbf{x}_i) = 1$. Jednym z przekształceń to zapewniających jest transformacja stosowana w modelu logitowym: $P(Y_i = 1 | \mathbf{x}_i) = \frac{\exp(\mathbf{x}_i \cdot \alpha)}{1 + \exp(\mathbf{x}_i \cdot \alpha)}$. Zakłada się w nim, że

składnik losowy podlega rozkładowi logistycznemu. Przy założeniu, że zmienna y_i jest związana z pewną nieobserwowalną zmienną y_i^* przyjmującą wartości ze zbioru $(-\infty, +\infty)$ i że funkcja $F(\mathbf{x}_i, \alpha)$ jest liniowa, model dwumianowy można zapisać w postaci:

$$y_i^* = \mathbf{x}_i \cdot \alpha + \varepsilon_i, \quad t = 1, \dots, T, \quad y_i = 1 \{y_i^* > 0\}.$$

² Część teoretyczna opracowana na podstawie fragmentów nieopublikowanej jeszcze rozprawy doktorskiej, udostępnionych dzięki uprzejmości mgr. W. Grabowskiego z Katedry Modeli i Prognoz Ekonometrycznych Wydziału Ekonomiczno-Socjologicznego Uniwersytetu Łódzkiego [Grabowski (w przygotowaniu)].

Przyjmując pewne założenia co do rozkładu składnika losowego, można dokonać estymacji parametrów modelu metodą największej wiarygodności. Otrzymane oszacowania mogą być interpretowane co do kierunku, jednak nie posiadają bezpośredniej interpretacji co do wartości. Efekty krańcowe zależą bowiem od wartości zmiennych objaśniających. Dlatego w celu otrzymania interpretowalnych wyników konieczne jest przekształcenie pierwotnych oszacowań przy wykorzystaniu założeń dotyczących rozkładu składnika losowego.

Z uwagi na specyfikę modelu logitowego klasyczne miary dopasowania nie znajdują zastosowania w ocenie jego wyników. Konieczne staje się zastosowanie innego rodzaju miar, a wśród nich indeksu ilorazu wiarygodności McFaddena, którego interpretacja jest analogiczna do współczynnika determinacji w klasycznym modelu regresji liniowej, dlatego bywa nazywany pseudo R^2 .

3. Pomiar empiryczny ryzyka wystąpienia kryzysu giełdowego na New York Stock Exchange

3.1. Charakterystyka zmiennych

Podstawową trudnością w prowadzeniu rachunków ryzyka dla kryzysu giełdowego jest brak jednoznacznie sprecyzowanej definicji tego zjawiska. Jest to konieczne do ustalenia wartości obserwacji dla zmiennej objaśniającej. W przedstawionym badaniu zmienna została zdefiniowana na podstawie danych kwartalnych dotyczących wartości indeksu Dow Jones Industrial Average na zamknięciu notowań. W celu porównania przyjęto dwie definicje kryzysu:

- a) spadek wartości indeksu z kwartału na kwartał o 10% lub więcej,
- b) spadek wartości indeksu z kwartału na kwartał o 15% lub więcej.

i oszacowano dwa warianty modelu.

Oprócz problemów z definicją kryzysu trudności nastęrcza również ustalenie listy zmiennych, jednoznacznie i zawsze wpływających na zwiększenie prawdopodobieństwa jego wystąpienia. Wśród takich przyczyn często ważną rolę odgrywają czynniki niemierzalne, których nie można ująć w modelu. Co więcej, każdy z dotychczasowych kryzysów posiadał odrębną specyfikę [Zieliński 2008]. W odniesieniu do czynników mierzalnych istnieje więc konieczność rozważenia wielu zmiennych o charakterze makroekonomicznym, potencjalnie wpływających na zwiększenie ryzyka wystąpienia kryzysu, a następnie wybranie spośród nich na podstawie kryteriów statystycznych zmiennych istotnych i dobrze opisujących dane empiryczne. Poniższa lista przedstawia zmienne brane pod uwagę w konstrukcji modelu logitowego:

- Realny kurs walutowy i jego odchylenia od długookresowego trendu – w badaniu wykorzystano dane dla Dollar Broad Trade Weighted Exchange Rate, bę-

dącego ważonym kursem dolara względem walut największych partnerów handlowych USA.

- Saldo na rachunku obrotów bieżących w relacji do PKB – deficyt bilansu obrotów bieżących oznacza sytuację, gdy inwestycje przekraczają oszczędności, co zwiększa podatność giełdy na kryzysy.
- Wielkość rezerw walutowych w stosunku do wielkości agregatu pieniężnego M2 – miara zewnętrznej wrażliwości kraju.
- Zmiana stanu kredytów bankowych w relacji do PKB – zwiększanie wartości tej zmiennej zwiększa ryzyko wystąpienia kryzysu.
- Dynamika eksportu – rozwijający się eksport wpływa pozytywnie na nastroje inwestorów.
- Inflacja – duże wartości indeksu cen konsumpcyjnych są odzwierciedleniem nierównowagi gospodarczej, wpływającej na zwiększenie ryzyka wystąpienia kryzysu.
- Poziom stóp procentowych – w badaniu wykorzystano dane dotyczące Bank Prime Loan Rate. Wysokie wartości zwiększają ryzyko wystąpienia kryzysu.
- Udział deficytu budżetowego w PKB – w razie wystąpienia deficytu często następuje emisja papierów dłużnych Skarbu Państwa, co wpływa na stopę oprocentowania innych instrumentów finansowych. Ponadto następuje zmniejszenie strumienia środków pieniężnych napływającego na giełdę. Sytuacja taka wpływa niekorzystnie na nastroje inwestorów.
- Dynamika realnego PKB – niska dynamika może ochłodzić nastroje inwestorów i zwiększyć podatność giełdy na kryzys.
- Opóźnione odchylenia wartości indeksu giełdowego od długookresowego trendu – wysokie wartości powinny zwiększać ryzyko wystąpienia kryzysu. Wynika to z obserwacji, że wielkie kryzysy giełdowe były poprzedzone okresami przewartościowania akcji, po których następowało wymykające się spod kontroli odreagowanie.

3.2. Wyniki estymacji

Parametry były szacowane na podstawie danych o częstotliwości kwartalnej dla okresu 1973.1-2007.2. Z uwagi na konieczność uwzględnienia opóźnień efektywna liczebność próby wyniosła 134 obserwacje.

Dla wariantu A definicji kryzysu (spadek wartości indeksu z kwartału na kwartał o 10% lub więcej) najlepszy wynik pod względem zarówno dopasowania do danych rzeczywistych, jak i istotności zmiennych osiągnięto przy uwzględnieniu następujących zmiennych:

- E_dyn , $E_dyn(-4)$ – bieżąca i opóźniona o rok dynamika realnego eksportu,
- $BD_GDP(-1)$ – opóźniony o kwartał udział realnego deficytu budżetowego w realnym produkcie krajowym brutto,
- $GDP_dyn(-1)$ – opóźniona o kwartał dynamika realnego PKB,

- DJ_res(-1), DJ_res(-2) – opóźnione o jeden i dwa kwartały odchylenia wartości indeksu Dow Jones Industrial Average od długookresowego trendu. Tabele 1-4 zawierają szczegółowe wyniki estymacji dla tego wariantu:

Tabela 1. Wyniki estymacji dla wariantu A

Zmienna	Współczynnik	<i>p-value</i>	Istotność	Efekt krańcowy
const	142,552	0,07196	*	
E_dyn	-12,227	0,05937	*	-0,0593
E_dyn(-4)	-7,838	0,15388		-0,0380
BD_GDP(-1)	158,689	0,00652	***	0,7701
GDP_dyn(-1)	-124,345	0,13145		-0,6035
DJ_res(-1)	0,334	0,04090	**	0,0016
DJ_res(-2)	0,409	0,01587	**	0,0020
Pseudo R ² :	0,53034	LR (6df):	35,0045	

Źródło: opracowanie własne z wykorzystaniem programu komputerowego Gretl.

Tabela 2. Dopasowanie do danych empirycznych dla wariantu A

		Przewidywane	
		0	1
Empiryczne	0	124	1
	1	3	6

Źródło: opracowanie własne z wykorzystaniem programu komputerowego Gretl.

Symbole „*”, „**” oraz „***” w tab. 1 odpowiadają kolejno poziomom istotności: 0,1; 0,05; 0,01. Jak łatwo zauważyć, zmiennych E_dyn(-4) i GDP_dyn(-1) nie można uznać za istotne na żadnym z tych poziomów, jednak pozostały w modelu, gdyż poprawiają trafność prognoz *ex post*. Wartość statyki LR testu dla łącznej istotności zmiennych potwierdza ich znaczenie w objaśnieniu prawdopodobieństwa wystąpienia kryzysu. Analiza dopasowania do danych empirycznych pozwala stwierdzić, że w 130 na 134 przypadków model poprawnie przewidział wystąpienie kryzysu lub jego brak, co stanowi 97% wszystkich obserwacji.

W wariancie B definicji kryzysu (spadek wartości indeksu z kwartału na kwartał o 15% lub więcej) w celu otrzymania jak najlepszych wyników zaistniała konieczność modyfikacji poprzedniego zestawu zmiennych. Usunięto opóźnioną o rok dynamikę eksportu i zastąpiono opóźnione o kwartał odchylenie wartości indeksu od długookresowego trendu opóźnieniem o trzy kwartały.

Tym razem mamy do czynienia z trzema zmiennymi nieistotnymi na standardowych poziomach ufności. Związane z nimi wartości *p-value* nie przekraczają jednak wartości 0,2, a zmienne te znacznie poprawiają trafność prognozy *ex post*, dlatego zostały pozostawione w modelu. Również w tym przypadku wysoka wartość

Tabela 3. Wyniki estymacji dla wariantu B

Zmienna	Współczynnik	p-value	Istotność	Efekt krańcowy
const	257,419	0,01563	**	
E_dyn	-11,742	0,17993		-0,0111
BD_GDP(-1)	174,855	0,02537	**	0,1658
GDP_dyn(-1)	-248,677	0,01864	**	-0,2357
DJ_res(-1)	0,367	0,13748		0,0003
DJ_res(-3)	0,338	0,16414		0,0003
Pseudo R2:	0,583673	LR (5df):	24,9199	

Źródło: opracowanie własne z wykorzystaniem programu komputerowego Gretl.

Tabela 4. Dopasowanie do danych empirycznych dla wariantu B

		Przewidywane	
		0	1
Empiryczne	0	129	0
	1	2	3

Źródło: opracowanie własne z wykorzystaniem programu komputerowego Gretl.

statystyki LR pozwala na odrzucenie hipotezy o braku łącznej istotności zmiennych. Na podstawie danych z tab. 4 można stwierdzić, że model poprawnie przewidział wystąpienie kryzysu lub jego brak 132 razy, co stanowi 98,5% przypadków.

4. Prognozy na podstawie dwóch wariantów modelu

Prognozy obejmują okres 2007.3-2009.4. Wydłużenie szeregów obserwacji dla zmiennych objaśniających odbyło się przy użyciu metod mechanicznych oraz ocen ekspertów. W drugim przypadku głównym źródłem informacji były strony internetowe The Financial Forecast Center.

W tabeli 5 zaprezentowano wyniki prognoz dla obu wariantów definicji kryzysu. Zmienne zdefiniowane wariantami A i B są nazwane odpowiednio: KR10 i KR15. Otrzymane wartości należy interpretować jako prawdopodobieństwo wystąpienia kryzysu giełdowego w danym okresie.

Jak widać, w obu wariantach największe prawdopodobieństwo wystąpienia kryzysu przypada na pierwszy kwartał 2008 r., kiedy mieliśmy do czynienia z dużymi spadkami na Wall Street, które stały się przyczyną dyskusji na temat powtórki wielkiego kryzysu lat trzydziestych. Wartości prognoz na kolejne kwartały nie potwierdzają jednak tej pesymistycznej wizji. W obu wariantach od trzeciego kwartału 2008 r. aż do końca 2009 r. wartości prawdopodobieństw są bliskie zeru. Zastanawiające jest to, że w przypadku drugiego kwartału roku 2008 r. prawdopodo-

Tabela 5. Wyniki prognozy

Okres	Prognoza KR10	Prognoza KR15
2007.3	0,671884	0,103404
2007.4	0,889235	0,190674
2008.1	0,978089	0,966319
2008.2	0,000596	0,295677
2008.3	0,000000	0,000000
2008.4	0,000007	0,000000
2009.1	0,000066	0,000019
2009.2	0,000008	0,000005
2009.3	0,000002	0,000002
2009.4	0,000000	0,000001

Źródło: opracowanie własne z wykorzystaniem programu komputerowego Gretl.

bieństwo wystąpienia kryzysu zdefiniowanego wariantem B jest większe niż w przypadku wariantu A, podczas gdy logiczniejsza byłaby sytuacja odwrotna. Należy jednak pamiętać, że zestawy zmiennych objaśniających różniły się od siebie w obu przypadkach, co uzasadnia otrzymanie pozornie niespójnego wyniku.

5. Wnioski

Mimo trudności wynikających z braku teorii dotyczącej definicji kryzysu giełdowego i jego przyczyn, przy przyjętych założeniach udało się zbudować model ekonometryczny opisujący to zjawisko. Jego wyniki wydają się satysfakcjonujące. Trzeba mieć jednak na względzie jego podstawową wadę: niejednoznaczność definicji kryzysu i ograniczoną liczbę uwzględnionych determinant. Pominięte zostały m.in. wszystkie czynniki jakościowe, mające wpływ na zachowania inwestorów. Z wymienionych względów wyniki modelu należy interpretować wyłącznie jako ewentualne ostrzeżenie, a nie ostateczną diagnozę kryzysu, co pozostaje w zgodzie z interpretacją w kategoriach prawdopodobieństwa. Nie umniejsza to jednak roli modelu logitowego jako atrakcyjnego narzędzia analizy sytuacji na giełdzie, które może stanowić wartościowe uzupełnienie krótkookresowych modeli szeregów czasowych.

Literatura

- Cieślak M., *Prognozowanie gospodarcze, metody i zastosowania*, wyd. 4, Wydawnictwo Naukowe PWN, Warszawa, 2005.
- Grabowski W., praca doktorska, Wydział Ekonomiczno-Socjologiczny Uniwersytetu Łódzkiego, Łódź (w przygotowaniu).

Milo W., Kozera Z., *O ryzyku kryzysu walutowego*, Acta Universitatis Lodziensis, Folia Oeconomica nr 177, UŁ, Łódź 2004.

Zieliński M., *Pięć i pół kryzysu*, www.rp.pl/artukul/87276.html (kwiecień 2008).

Żabińska J., *Kryzysy finansowe, ich przyczyny, skutki i sposoby przezwyciężania*, Wydawnictwo AE, Katowice 2000.

www.economagic.com (kwiecień 2008).

www.forecasts.org (kwiecień 2008).

THE APPLICATION OF LOGIT MODEL IN STOCK EXCHANGE CRISIS FORECASTING ON THE EXAMPLE OF THE NEW YORK STOCK EXCHANGE

Summary: Stock exchange crisis is an example of qualitative variable which may be expressed in quantitative terms. In this case it takes only two values. In order to model such a variable one should use binary choice model and choose appropriate transformation which enables to estimate its parameters, e.g. logit transformation. The paper presents the application of logit model to describe and forecast a situation on the New York Stock Exchange. The study presents and compares forecasts based on two alternative definitions of crisis.