

Michał Masłowski, Marek Nowiński

Akademia Ekonomiczna we Wrocławiu

TEST *WHITE'S REALITY CHECK* JAKO METODA WERYFIKACJI AUTOMATYCZNYCH SYSTEMÓW TRANSAKCYJNYCH

1. Wstęp

Budowanie modeli inwestycyjnych, wykorzystujących w swej konstrukcji osiągnięcia analizy technicznej, generujących sygnały kupna-sprzedaży instrumentów finansowych na rynku kapitałowym nieodłącznie wiąże się z problemem określonym jako efekt *data-snooping*. Występuje on wtedy, gdy dobry wynik systemu inwestycyjnego powstaje dzięki szczęściu w doborze parametrów, a nie w wyniku faktycznej siły prognostycznej budowanego modelu. Efekt *data-snooping* najczęściej pojawia się w momencie, gdy zestaw danych używanych do opracowania modelu inwestycyjnego jest używany wielokrotnie, zarówno podczas konstruowania modelu, jak i na etapie jego weryfikacji.

Celem niniejszego artykułu jest przedstawienie wybranych metod weryfikacji występowania efektu *data-snooping*, a także szczegółowe omówienie jednej z nich, znanej jako *White's reality check*, na przykładzie z wykorzystaniem danych pochodzących z polskiego rynku kapitałowego.

2. Systemy transakcyjne i ich optymalizacja

W niniejszym artykule będą rozpatrywane przypadki powstania efektu *data-snooping* przy konstruowaniu automatycznych systemów transakcyjnych, stosowanych na rynku kapitałowym, bazujących na osiągnięciach analizy technicznej.

Systemem transakcyjnym¹ nazywamy zespół reguł wykorzystywanych do generowania sygnałów kupna oraz sprzedaży instrumentów finansowych.

Można także podać następującą definicję systemu transakcyjnego, pozwalającą na późniejszą formalizację zapisów konkretnych systemów transakcyjnych²:

¹ Opracowanie własne na podstawie [Schwager 2002].

² Opracowanie własne na podstawie [Skouras 1998].

$$s_t = D(\mathbf{I}_t, \bar{\mathbf{x}}) : R_+^N \times R^k \rightarrow \Omega \quad (1)$$

$$\mathbf{I}_t = (P_t, P_{t-1}, \dots, P_{t-N+1}), \bar{\mathbf{x}} \in B \subseteq R^k, \mathbf{I}_t \in R_+^N, \Omega = \{-1, 0, 1\} \quad (2)$$

gdzie: Ω – zbiór możliwych decyzji inwestycyjnych (-1 – krótka pozycja, 0 – brak pozycji (pozycja neutralna, poza rynkiem), 1 – długa pozycja),

B – przestrzeń parametrów,

$\bar{\mathbf{x}}$ – wektor parametrów,

$\mathbf{I}_t$ – wektor cen dostępnych w okresie t (N ostatnich cen),

P_t – cena instrumentu bazowego w okresie t ,

D – funkcja sygnału, która na podstawie dostępnych cen $\mathbf{I}_t$ i wartości parametrów $\bar{\mathbf{x}}$ generuje kierunek aktualnej pozycji rynkowej,

s_t – bieżąca wartość funkcji sygnału D w momencie t .

Inwestorzy w celu poszukiwania optymalnego zestawu parametrów dla konkretnego systemu transakcyjnego na danym rynku stosują procedurę zwaną optymalizacją automatycznych systemów transakcyjnych. Optymalizacja polega na przeszukaniu wszystkich możliwych kombinacji parametrów systemu transakcyjnego w poszukiwaniu takiego zestawu, którego zastosowanie da najlepszy wynik (największą stopę zwrotu). Optymalizację można sprowadzić do rozwiązania następującego zadania:

$$\max_{\bar{\mathbf{x}}} \left\{ \frac{1}{P} \sum_{t=R}^T \ln(1 + y_{t+1} s_t) \right\} \quad (3)$$

$$s_t = D(\mathbf{I}_t, \bar{\mathbf{x}}) : R_+^N \times R^k \rightarrow \Omega \quad (4)$$

$$\mathbf{I}_t = (P_t, P_{t-1}, \dots, P_{t-N+1}), \bar{\mathbf{x}} \in B \subseteq R^k, \mathbf{I}_t \in R_+^N, \Omega = \{-1; 0; 1\} \quad (5)$$

gdzie: P jest liczbą okresów (dni), w których testowany jest dany system inwestycyjny indeksowanych od R do T ,

R jest minimalną liczbą obserwacji niezbędną do wygenerowania sygnału inwestycyjnego ($T=R+P-1$).

y_t jest dzienną stopą zwrotu liczoną $(P_{t+1}-P_t)/P_t$.

Z praktycznego punktu widzenia należałoby także podać średnią roczną stopę zwrotu (procentową). W odniesieniu do systemów transakcyjnych, w których bazowymi są instrumenty pochodne, należałoby także podać wynik systemu liczony w punktach. Wynika to z trudności przy ustaleniu wielkości kapitału początkowego (depozyt zabezpieczający i dodatkowe środki), a ten sposób podawania wyników systemów transakcyjnych ułatwia ich porównywanie.

3. Efekt *data-snooping*

Z jednej strony dokonuje się optymalizacji, aby uniknąć przypadkowego doboru parametrów, wykorzystuje się przy tym założenie [Schwager 2002], że zestaw parametrów, który cechował się wysoką skutecznością w przeszłości, cechuje wysokie prawdopodobieństwo skutecznego inwestowania w przyszłości. Z drugiej jednak strony przeszukiwanie całej przestrzeni parametrów niesie z sobą niebezpieczeństwo zaistnienia sytuacji, w której, mimo braku zdolności prognostycznych badanego modelu, jedna z rozpatrywanych kombinacji parametrów da przez przypadek dobre wyniki (wysoką stopę zwrotu). W praktyce często także dochodzi do sytuacji opisanej przez V. Tharpa [2000] jako „pokusa mnożenia stopni swobody”, w której inwestorzy, dodając kolejne ograniczenia i parametry do systemu transakcyjnego, dążą do osiągnięcia możliwie najwyższej stopy zwrotu, wskutek czego często otrzymują modele, które w doskonały sposób „dopasowują się” do danych historycznych, jednak całkowicie zawodzą na danych testowych. Efekt ten nazywamy efektem *data-snooping*.

4. Wybrane metody weryfikacji

W razie podejrzenia, iż osiągnięty dobry wynik inwestycyjny jest wynikiem działania efektu *data-snooping*, a nie faktycznej siły prognostycznej opracowanego modelu, zaleca się m.in. następujące sposoby weryfikacji [Sullivan, Timmermann, White 2003]. Do pierwszej grupy należą metody, w których unika się używania tych samych danych na etapie konstrukcji i testowania modeli inwestycyjnych.

- **Oczekiwanie, aż nowe dane staną się dostępne** – metoda, mimo że powszechnie możliwa do zastosowania, jest metodą niezwykle czasochłonną. Odmianą tej metody jest podział posiadanych danych na dane, na podstawie których buduje się model, i dane testowe, na których następuje jedynie weryfikacja skonstruowanego modelu. Jest to jednak możliwe do zastosowania jedynie przy dostępnej bardzo dużej liczbie danych. Na przykład w razie testowania jakichkolwiek z modeli inwestycyjnych, na danych miesięcznych, w odniesieniu do polskiego rynku kapitałowego mamy do dyspozycji jedynie około 180 notowań, co w razie budowy modeli długoterminowych wyklucza niniejszą metodę. Dodatkowo podział danych na dane służące do budowy modelu i do jego testowania jest zawsze arbitralny, a więc nie pozbawiony całkowitej obiektywności. Innym zagrożeniem jest efekt okresowej utraty skuteczności automatycznych systemów transakcyjnych. Modele, które w jednym okresie cechowały się wysoką skutecznością, w innym, np. na skutek zmiany dynamiki rynku, mogą tę skuteczność zatracić.
- **Zastosowanie opracowanego modelu na danych „podobnych”** – jest to powszechna praktyka stosowana na rynkach finansowych; z uwagi na brak danych

inwestorzy, w przypadku znalezienia modelu inwestycyjnego, który daje dobre wyniki na jednym rynku finansowym, testują jego skuteczność na innych rynkach. Dobra skuteczność badanego systemu transakcyjnego na porównywalnych rynkach jest interpretowana jako argument za przyjęciem hipotezy o dobrej skuteczności tego systemu na rynku podstawowym, z kolei złe wyniki na rynku finansowym służącym do testowania są przesłanką do przyjęcia hipotezy o złej skuteczności badanego modelu również na rynku podstawowym. Takie podejście ma jednak dwie wady. Zachowania obu rynków finansowych mogą być ze sobą silnie skorelowane, zatem dane z nich pochodzące mogą być silnie zależne, co będzie po części powodowało zagrożenie, że zarówno opracowanie, jak i testowanie systemu transakcyjnego będą się odbywały niemalże na tych samych danych. Z kolei całkowicie odmiennym zagrożeniem dla tego typu metod testowania systemów transakcyjnych są strukturalne różnice obydwu rynków, które mogą powodować, że strategie inwestycyjne, które dobrze funkcjonują na rynku podstawowym, mogą całkowicie zawieść na rynku testowym, co jednak nie wyklucza ich skutecznego stosowania.

Stosowanie powyższych metod jest znacznie utrudnione, gdyż porównywalne zbiory danych, służące wyłącznie do testowania opracowanych systemów transakcyjnych, mogą nie istnieć lub mogą być bardzo trudno osiągalne. Dla odmiany proponuje się bardziej formalne podejście, z użyciem danych, na których był budowany rozpatrywany system transakcyjny. Jedną z najbardziej znanych i popularnych jest bootstrapowa metoda zaproponowana w pracy [Brock, Lakonishok, LeBaron 1992], w której porównuje się skuteczność opracowanego systemu transakcyjnego zastosowanego na danych oryginalnych z wynikami tego systemu zastosowanego na danych zastępczych modelowanych jako błądzenie przypadkowe, szeregi autoregresyjne pierwszego rzędu i szeregi z warunkową wariancją. Metoda ta jest szeroko znana i opisywana w literaturze m.in. [Isakov, Hollistein 1998; Fernandez-Rodriguez, Sosvilla-Rivero, Andrada-Felix 1999; Ratner, Leal 1998; Kwon, Kish 2002], jednak do tej pory nie znalazła ona zastosowań praktycznych (np. oprogramowania w znanych pakietach do analizy technicznej).

Kolejna metoda to zaproponowany przez H. White'a [2000] test znany jako *White's reality check*. W. White zaproponował test weryfikujący hipotezę, czy najlepszy model inwestycyjny na danym rynku posiada siłę prognostyczną zdolną do wygenerowania w dłuższym horyzoncie czasu wyników lepszych od benchmarku, z uwzględnieniem efektu *data-snooping*. Jako benchmark najczęściej przyjmuje się dwie strategie: cały czas utrzymywanie pozycji długiej – strategia „kup i trzymaj” (ang. *buy and hold*) lub też pozostawanie poza rynkiem – strategia *hold cash*. Z uwagi na popularność hipotezy rynku efektywnego najczęściej spotykaną strategią zerową jest strategia „kup i trzymaj”.

5. White's reality check

Procedura zaproponowana przez H. White'a, na podstawie prac F. Diebolda i R. Mariano [1995] oraz K. Westa [1996], jest następująca:

Należy znaleźć empiryczną dystrybuantę rozkładu stóp zwrotu, generowanych przez wszystkie modele z założonego wcześniej „uniwersum” (zbioru wszystkich możliwych modeli inwestycyjnych – w omawianym przypadku systemów transakcyjnych). Oczywiście w praktyce niemożliwe jest wzięcie pod uwagę wszystkich systemów transakcyjnych z uwagi na to, że są one tworzone przez inwestorów w sposób całkowicie dowolny. W praktyce tworzy się sztucznie uniwersum jako stosunkowo duży zbiór systemów transakcyjnych³.

Test opiera się na l statystykach:

$$\bar{f}_k = P^{-1} \sum_{t=R}^T f_{k,t+1}(\hat{\beta}_t) \quad k=1, \dots, l \quad (6)$$

gdzie l jest liczbą rozpatrywanych systemów transakcyjnych, P jest liczbą okresów, w których dokonuje się prognozy, indeksowanych od R do T , tak że $T = R + P - l$, $\hat{\beta}_t$ jest wektorem estymowanych parametrów. W przypadku badania skuteczności systemów transakcyjnych zdefiniowanych wzorami (1), (2), zamiast estymowanych w każdym kroku wielkości, mamy parametryzację systemów transakcyjnych – kolejne zestawy parametrów β_k ($k = 1, \dots, l$), dzięki czemu otrzymujemy natychmiast wyniki stóp zwrotu dla poszczególnych systemów. Przy zastosowaniu niniejszej procedury do badania tylko systemów transakcyjnych bazujących na analizie technicznej przyjęto następujące oznaczenia [Sullivan, Timmermann, White 1999]:

$$f_{k,t+1}(\beta) = \ln[1 + y_{t+1}s_k(\mathbf{I}_t, \beta_k)] - \ln[1 + y_{t+1}s_0(\mathbf{I}_t, \beta_0)], \quad k=1, \dots, l \quad (7)$$

$$\mathbf{I}_t = (P_t, P_{t-1}, \dots, P_{t-N+1}) \quad (8)$$

gdzie: P_t jest oryginalnym szeregiem cenowym,

$$y_{t+1} = (P_{t+1} - P_t) / P_t,$$

s_k i s_0 są wartościami funkcji sygnału odpowiednio dla k -tego systemu transakcyjnego oraz dla benchmarku. Z reguły przyjmuje się, że $s_0=1$ (strategia *buy and hold*). Hipotezą zerową będzie hipoteza, w której zakłada się, że najlepszy znaleziony system transakcyjny nie daje lepszych wyników niż przyjęta strategia zerowa:

$$H_0 : \max_{k=1, \dots, l} \{\bar{f}_k\} \leq 0 \quad (9)$$

³ Na przykład w pracy [Sullivan, Timmermann, White 1999] wybrano uniwersum składające się z 7846 systemów transakcyjnych, w pracy [Qi, Wu 2001] – 2127, w pracy [Hsu, Kuan 2005] – 39 832.

Odrzucenie hipotezy zerowej będzie oznaczało, że istnieje strategia inwestycyjna dająca lepsze rezultaty od strategii kup i trzymaj (lub innej przyjętej, np. „poza rynkiem”). White pokazał, że hipoteza zerowa może być weryfikowana poprzez zastosowanie metodologii „stacjonarnego bootstrapu” (ang. *stationary bootstrap*) [Politis, Romano 1994] do obserwowanych wartości f_k , w celu uzyskania bootstrapowych szeregów zastępczych oznaczonych $f_{k,i}^*$, gdzie i jest numerem zastępczego szeregu ($i = 1, \dots, B$). Dla $B = 500$ skonstruowano następujące statystyki:

$$\bar{V} = \max_{k=1, \dots, J} \left\{ \sqrt{P}(\bar{f}_k) \right\} \quad (10)$$

$$\bar{V}_i^* = \max_{k=1, \dots, J} \left\{ \sqrt{P}(\bar{f}_{k,i}^* - \bar{f}_k) \right\}, \quad i = 1, \dots, B \quad (11)$$

gdzie

$$\bar{f}_k = P^{-1} \sum_{t=R}^T f_{k,t} \quad (12)$$

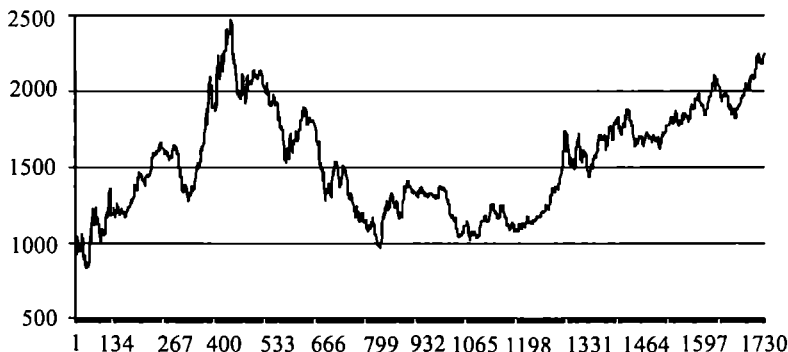
$$\bar{f}_{k,i}^* = P^{-1} \sum_{t=R}^T f_{k,t}^* \quad (13)$$

Porównując otrzymaną wartość $\bar{V}$ z kolejnymi wartościami $\bar{V}_i^*$ ($i = 1, \dots, 500$), otrzymujemy wartość- p (ang. *p-value*). Miara ta znana jest jako *White's reality check p-value* i pozwala decydować o przyjęciu lub odrzuceniu hipotezy zerowej przy pewnym założonym poziomie granicznym.

Badania, z wykorzystaniem powyższej metodologii, przeprowadzone przez R. Sullivana [Sullivan, Timmermann, White 1999] wykazały istotną przewagę niektórych stosowanych systemów transakcyjnych nad strategią „kup i trzymaj” dla indeksu DJIA w okresie do 1987 r. (na tych samych danych, co zastosowanych w pracy [Brock, Lakonishok, LeBaron 1992], około 100 lat notowań), z kolei pokazano brak przewagi analizy technicznej w kolejnych 10 latach na tym samym rynku dla tych samych reguł analizy technicznej. W pracy [Qi, Wu 2001] za pomocą testu *reality check* wykazano skuteczność analizy technicznej na wybranych parach walut, w końcu w jednej z nowszych prac [Hsu, Kuan 2005] podjęto próbę oceny skuteczności na dojrzałych rynkach (indeksy S&P500, DJIA) w porównaniu z rynkami stosunkowo młodymi (indeksy NASDAQ Composite i Russell 2000). Pokazano istotną skuteczność strategii technicznych na rynkach młodych oraz jej brak na rynkach dojrzałych, co w pewien sposób potwierdzają badania [Sullivan, Timmermann, White 1999].

6. Badania empiryczne

Rozpatrzony został przykład, w którym badano skuteczność strategii inwestycyjnych na kontraktach terminowych na WIG20. Wybór takiego instrumentu bazowego wynika z praktycznych możliwości zastosowań, zwłaszcza z możliwości zajmowania krótkiej pozycji. Stosowanie indeksów giełdowych do badań nad skutecznością strategii analizy technicznej jako instrumentów bazowych wiąże się z kilkoma problemami, m.in. są to: problemy w ustaleniu pojęć „kupno i sprzedaż indeksu”, problemy z zajmowaniem krótkiej pozycji, a także okresowe zmiany składu indeksów, co utrudnia porównywanie tego instrumentu w różnych okresach. Zastosowanie kontraktów terminowych jako instrumentu bazowego zamiast indeksów giełdowych lub kursów akcji powoduje odmienny sposób liczenia prowizji maklerskich: dla kontraktów są one liczone kwotowo, natomiast dla pozostałych, tradycyjnych instrumentów są liczone procentowo. Badaniu poddano dane z okresu od 7 kwietnia 1998 r. do 31 sierpnia 2005 r. (1856 obserwacji). Wedle przyjętych oznaczeń $R=101$, co oznacza, że potrzeba 100 obserwacji do wygenerowania sygnału inwestycyjnego przez „najwolniejszy” z rozpatrywanych systemów transakcyjnych, zatem $T=1856$ i $P=1755$. Wyniki rozpatrywanych systemów transakcyjnych były porównywane do benchmarku (w tym przypadku strategii „kup i trzymaj”) w okresie od 1 września 1998 r. do 31 sierpnia 2005 r. (1755 obserwacji – 7 lat).



Rys. 1. Wykres kontynuacyjny kontraktów terminowych na indeks WIG20 w okresie od 1 września 1998 r. do 31 sierpnia 2005 roku

Źródło: opracowanie własne.

W tym okresie, zakładając realizowanie strategii „kup i trzymaj” oraz brak dźwigni finansowej, inwestor osiągnąłby stopę zwrotu w wysokości 142% zainwestowanego kapitału (roczna stopa zwrotu 20%).

Dla porównania wybrano jako uniwersum strategii inwestycyjnych do testów grupę systemów opartych na pojedynczych średnich ruchomych, dwóch średnich ruchomych, wybiciach, wskaźniku momentum oraz na wskaźniku MACD – są to systemy wykorzystujące istnienie trendów; systemy „grające z trendem”. Rozpa-

trazono też systemy transakcyjne oparte na wskaźniku ROC (system przeciwtrendowy) z dodatkową regułą stopu. Do systemów „grających z trendem” zastosowano również kolejne parametry, takie jak procentowe progi potwierdzające wystąpienie sygnału. W sumie rozpatrzono uniwersum składające się z 1655 systemów transakcyjnych.

Systemem transakcyjnym, który dał największą stopę zwrotu spośród rozpatrywanych, okazał się bardzo prosty jednoparametrowy system oparty na wskaźniku Momentum o wielkości parametru: 10 sesji giełdowych. Zakładając, podobnie jak w przypadku strategii „kup i trzymaj”, brak dźwigni finansowej, prowizje w wysokości jednego punktu od zawarcia transakcji jednym kontraktem terminowym, inwestor stosujący się do tej strategii osiągnąłby stopę zwrotu w wysokości 320%. Oznacza to przewagę nad strategią zerową w wysokości 178 pp.

Wartość *reality check p-value* dla tej strategii wynosi 0,01. Wartość ta może być interpretowana jako prawdopodobieństwo, że efekty uzyskane ze strategii „kup i trzymaj” będą lepsze od wyników uzyskanych po zastosowaniu najlepszego znalezionej systemu transakcyjnego lub też w tym przypadku jako minimalna wartość poziomu istotności, przy którym hipoteza zerowa będzie odrzucona. Przyjmując poziom graniczny 0,05, taki wynik oznacza odrzucenie hipotezy zerowej. H. White podaje, że stosując postępowanie *reality check* i rozpatrując pełną gamę strategii inwestycyjnych, uznaje się, że znaleziona strategia inwestycyjna pozwala uzyskiwać istotnie lepsze wyniki od strategii „kup i trzymaj”, z jednoczesną kontrolą efektu *data-snooping*.

W tabeli 1 przedstawiono wartości *reality check p-value* dla wybranych popularnych systemów transakcyjnych z zastosowanymi typowymi wartościami parametrów. Instrumentem bazowym w każdym przypadku jest, tak jak w prezentowanym przykładzie, kontrakt terminowy na indeks WIG20.

Jako najważniejszą wadę testu *reality check* wymienia się to, że daje znacznie zawyżone wyniki, tj. pokazuje skuteczność reguł analizy technicznej mimo ich częstego braku w sytuacji, gdy do modelu zostanie włączona znaczna liczba systemów transakcyjnych o bardzo niskiej stopie zwrotu. Jest to sytuacja bardzo częsta, gdyż na

Tabela 1. *White's reality check p-value* dla wybranych systemów transakcyjnych

System transakcyjny	Wynik [pkt.]	Stopa zwrotu [%]	Stopa zwrotu w stosunku rocznym [%]	<i>White's reality check p-value</i>
System oparty na średniej ruchomej 14-sesyjnej	2525	273	39	0,012
System oparty na wskaźniku Momentum o wartości parametru 10 sesji	2955	319	45	0,01
System oparty na wskaźniku Momentum o wartości parametru 15 sesji	1808	195	28	0,024
System oparty na wskaźniku MACD o długości średniej ruchomej SIG 5 sesji	2068	223	32	0,022

Źródło: opracowanie własne.

wielu rynkach całe grupy systemów transakcyjnych cechują się niską skutecznością. Kolejną wadą jest to, że test *reality check* nie uwzględnia zmienności stóp zwrotu generowanych przez rozpatrywane systemy transakcyjne. Metoda ta, podobnie jak ta zaproponowana przez [Brock, Lakonishok, LeBaron 1992], nie znalazła praktycznego zastosowania w szeroko stosowanych aplikacjach komputerowych.

Literatura

- Brock W., Lakonishok J., LeBaron B., *Simple Technical Trading Rules and the Stochastic Properties of Stock Returns*, „The Journal of Finance” 1992 vol. 47, Issue 5 (December), s. 1731-1764.
- Diebold F., Mariano R., *Comparing Predictive Accuracy*, „Journal of Business and Economic Statistics” 1995 nr 13, s. 253-265.
- Fernandez-Rodriguez F., Sosvilla-Rivero S., Andrada-Felix J., *Technical Analysis in the Madrid Stock Exchange*, Fundacion Esudios de Economia Aplicada, Documento de Trabajo 99-05, 1999.
- Hsu P., Kuan C., *Re-examining the Profitability of Technical Analysis with White's Reality Check and Hansen's SPA Test*, 2005.
- Isakov D., Hollistein M., *Application of Simple Technical Trading Rules to Swiss Stock Prices: Is it Profitable?*, International Center for Financial Asset Management and Engineering, Research Paper nr 2, 1998.
- Kwon K., Kish R.J., *Technical Trading Strategies and Return Predictability: NYSE*, „Applied Financial Economics” 2002 nr 12, s. 639-653.
- Politis D.N., Romano J.P., *The Stationary Bootstrap*, „Journal of the American Statistical Association” 1994 vol. 89 nr 428, s. 1303-1313.
- Ratner M., Leal R., *Test of Technical Trading Strategies in the Emerging Equity Markets of Latin America and Asia*, „Journal of Banking and Finance” 1998.
- Schwager J., *Analiza techniczna rynków terminowych*. Wig-press, Warszawa 2002.
- Skouras S., *Financial Returns and Efficiency as seen by an Artificial Technical Analyst*, 1998.
- Sullivan R., Timmermann A., White H., *Data-Snooping, Technical Trading Rule Performance, and the Bootstrap*, „Journal of Finance” 1999 vol. 54, s. 1647-1691.
- Sullivan R., Timmermann A., White H., *Scientific Progress with Data Sharing*, „International Journal of Forecasting” 2003 nr 19, s. 217-228.
- Tharp V., *Gielda, wolność i pieniądze*, Wig-press, Warszawa 2000.
- Qi M., Wu Y., *Technical Trading-rule Profitability, Data-Snooping and Reality Check: Evidence from the Foreign Exchange Market*. 2001
- West K., *Asymptotic Inference about Predictive Ability*, „Econometrica” 1996 vol. 64, s. 1067-1084.
- White H., *A Reality Check for Data Snooping*, „Econometrica” 2000 vol. 68 nr 5, s. 1097-1126.

EFFECTIVENESS' VERIFICATION OF MECHANICAL TRANSACTIONS SYSTEMS USING WHITES'S REALITY CHECK

Summary

The article presents a few methods of effectiveness' verification of mechanical transactions systems. These methods are divided into two parts. The first class of methods is based on assumption that different data are used to build and test system and different to verify its effectiveness. The second class of methods are based on bootstrap approach proposed by Brock [1992]. One of the variants of that method was proposed by White [2000] and is known as White's Reality Check test. This test is presented in details. The paper is exemplified by data from the Polish Stock Exchange.