

## Wpływ ogłoszenia upadłości na złożoność strukturalną zmian cen na GPW

Paweł Fiedor\*, Artur Hołda\*\*

### Streszczenie

W artykule prezentujemy metodę badania efektywności rynku giełdowego za pomocą teorii informacji. Efektywność rynku badana jest przez stopień redundancji w szeregach czasowych opisujących zmiany cen, a konkretnym narzędziem jest stopa entropii Shannona, którą można także interpretować jako miarę przewidywalności zmian cen (w sensie granicy przewidywalności). Metody tej używamy do analizy szeregów czasowych opisujących logarytmiczne zmiany cen akcji wybranych spółek z Giełdy Papierów Wartościowych, które podległy procesowi upadłościowemu. Znane są badania efektywności całego rynku, natomiast nie jest zbadane dogłębnie, jak skrajnie negatywna sytuacja ekonomiczna samej spółki, a konkretnie informacja o niej, wpływa na efektywność procesów cenotwórczych dotyczących jej akcji oraz na przewidywalność zmian cen tych akcji. Przegląd prezentowany w niniejszym badaniu, oparty na 44 spółkach giełdowych, jest wstępem do szerszej gamy badań dotyczących wpływu zdarzeń pozagiełdowych na złożoność strukturalną procesów cenotwórczych.

**Słowa kluczowe:** ekonofizyka, złożoność, bankructwa, efektywność, finanse

**Kody JEL:** G14, G33

**DOI:** <http://dx.doi.org/10.17451/eko/41/2015/75>

---

\* Uniwersytet Ekonomiczny w Krakowie, Rakowicka 27, 31-510 Kraków,  
e-mail: Pawel.F.Fiedor@ieec.org

\*\* Uniwersytet Ekonomiczny w Krakowie, Rakowicka 27, 31-510 Kraków,  
e-mail: aholda@uek.krakow.pl

## 1. Wstęp

Rynki finansowe są dobrze zdefiniowanymi adaptacyjnymi systemami złożonymi (Mantegna, Stanley 2000; Tumminello *et al.* 2005), które są jednak rzadko traktowane *explicite* jako systemy złożone w badaniach głównego nurtu ekonomii. Z pomocą przychodzi częściowo ekonofizyka, w której to dyscyplinie, ze względu na duże zaangażowanie badaczy wywodzących się z teorii złożoności, kładziony jest szczególny nacisk na uwzględnienie i badanie złożoności (Mantegna, Stanley 2000; Rosser 2008; Fiedor, Hołda 2015), choć nie zawsze adaptacyjności (Fiedor 2014). W niniejszym opracowaniu przyglądamy się konkretnemu rodzajowi złożoności, a konkretnie złożoności strukturalnej, opartej na teorii informacji, która w przypadku badania szeregów czasowych opisujących zmiany cen giełdowych może być interpretowana jako przewidywalność (w sensie limitu mocy predykcyjnych najszerzej klasy modeli predykcyjnych). Złożoność strukturalna finansowych szeregów czasowych ma także związek z hipotezą efektywnego rynku (Samuelson 1965), gdyż większa złożoność strukturalna tych procesów oznacza mniejszą ich przewidywalność, a w szczególności w limicie sytuację postulowaną przez hipotezę efektywnego rynku. Badania złożoności strukturalnej finansowych szeregów czasowych można więc traktować jako test tej hipotezy (Fiedor 2014a; Navet, Chen 2008).

Przewidywalność systemu definiuje się zwykle, w badaniach systemów dynamicznych, jako możliwość znajomości przyszłego stanu tego systemu z uwagi na znajomość stanów historycznych. Najczęściej mierzy się przewidywalność systemów dynamicznych za pomocą takich narzędzi jak entropia Kołmogorowa-Sinaia lub wykładniki Lapunowa. Wykładniki Lapunowa mogą nie posiadać skończonych wartości dla procesów cenotwórczych na rynkach finansowych, dlatego też korzystamy z pierwszego z wymienionych narzędzi. W praktyce przewidywalność szeregów czasowych to ilość wzorców, które w tych szeregach występują (Fiedor 2014a).

Przewidywalność była początkowo badana w fizyce, natomiast obecnie staje się ona istotnym przedmiotem badań także w naukach społecznych, a niektórzy badacze twierdzą nawet, że jest jednym z kluczowych pytań w naukach społecznych (Navet, Chen 2008). Zachowania ludzkie na rynkach (nie tylko finansowych), jako zachowania wpływające na popyt i na podaż (jednocześnie zatem na ceny i ich zmienność), to tylko jeden z możliwych do badania aspektów przewidywalności zachowań ludzkich. W naukach społecznych niezwykle ważny jest stopień, w jakim zachowania te mogą być przez badaczy przewidywalne. Jest to niezwykle istotne przede wszystkim z powodu opierania na procesach stochastycznych większości modeli w naukach społecznych.

W badaniach rynków finansowych istnieje nacisk (choć nie jest to reguła, używane są także testy niezakładające *explicite* modelu, jak np. testy serii) na bada-

nie praktycznych możliwości predykcji opartych na konkretnych modelach, a nie na samej przewidywalności (limitu możliwości predykcyjnych). Badania takie pokazują jedynie, jakie są możliwości przewidywania przy użyciu konkretnych proponowanych przez badaczy narzędzi predykcyjnych, a nie pokazują przewidywalności danego procesu, która jest niezależna od używanych metod predykcji. Nie umniejsza to wartości merytorycznej tych badań, ale pokazuje potrzebę badania przewidywalności samej w sobie w celu poznania charakterystyki procesów cenotwórczych i innych procesów społecznych w zakresie ich złożoności. Analiza tego rodzaju pozwala m.in. na umieszczenie mocy predykcyjnej obecnie używanych modeli na tle teoretycznych możliwości przewidywania danego procesu. Analiza taka jest więc istotna zarówno z teoretycznego, jak i z praktycznego punktu widzenia dla zrozumienia charakterystyki procesów cenotwórczych, możliwości ich predykcji, oraz – będąc podbudową predykcji – ma też znaczenie ze względu na praktyczną działalność inwestorów i przedsiębiorstw.

Wracając do adaptacyjności systemów złożonych, w tym rynków giełdowych, należy zaznaczyć, że wiele badań w finansach, w tym te pokazujące potencjalną zyskowność strategii na danych historycznych, nie jest metodologicznie uzasadniona, gdyż te metody (przy ich użyciu w praktyce) zmieniałyby zachowanie rynków, zmieniając tym samym efektywność badanych metod. Traktuje się często systemy ekonomiczne jako systemy chaotyczne lub też systemy klasyczne, podczas gdy są one kompleksowymi systemami adaptacyjnymi, które wymagają innego podejścia (Fiedor 2014). Badania nad limitem predykcji oraz złożonością procesów nie napotykają tego problemu, jako że mierzenie tej charakterystyki rynku nie zmienia i nie może zmieniać zachowań graczy rynkowych, co czyni wyniki miarodajnymi.

Badania nad złożonością strukturalną i przewidywalnością cen mają długą historię, głównie w kontekście obalania wspomnianej hipotezy efektywnego rynku. Badania takie były także przeprowadzane zgodnie z metodyką używaną w niniejszym opracowaniu (Fiedor 2014a; Navet, Chen 2008). Badania te dotyczyły jednak efektywności całego rynku. Można oczywiście spojrzeć głębiej i analizować złożoność strukturalną procesów cenotwórczych danych spółek. Można w szczególności użyć metody przesuwnego okna, tak aby zbadać, jak ta złożoność zmieniała się w czasie, co czynimy w niniejszym opracowaniu. Takie badania dotyczące pojedynczych spółek muszą w sposób oczywisty mieć jakieś tło nadające im sens merytoryczny. W tym przypadku chcemy badać, jak zdarzenia zewnętrzne w stosunku do samej giełdy wpływają na procesy cenotwórcze wewnątrz danego rynku finansowego. Podobne badania, dotyczące np. wpływu wiadomości z portali takich jak Wikipedia i Twitter na procesy cenotwórcze, są obecnie niezwykle popularne (Alanyali *et al.* 2013; Kristoufek 2013; Mestyán *et al.* 2013; Moat *et al.* 2013; Preis *et al.* 2013). W niniejszym opracowaniu badamy, czy fakt skrajnie złej sytuacji ekonomicznej (a konkretnie informacji o tej sytu-

acji wpływającej na rynek) spółek wpływa na złożoność strukturalną i efektywność ich procesów cenotwórczych. Dla znalezienia spółek w skrajnie złej sytuacji wystarczy wybrać spółki, które uległy upadłości (ogłoszenie publiczne informacji o upadłości jest też dosyć łatwe do odnalezienia w czasie). Zatem interesującym nas pytaniem jest to, czy informacje płynące ze spółki, które trafiają do uczestników rynku w sposób asymetryczny, wpływają na przewidywalność zmian cen tej spółki. Wydawać się może, że tak być powinno, jednak ogromna złożoność rynków finansowych powodować może to, że efektywność tych procesów nie zostanie zakłócona ze względu na licznosc inwestorów i ich celów oraz metod inwestycyjnych. Hipoteza efektywnego rynku została, w swojej mocnej wersji, w dużej mierze odrzucona przez ekonomistów, niemniej jednak należy zanotować, że hipoteza ta wskazywałaby na to, że złożoność strukturalna procesów cenotwórczych nie będzie się zmieniać ze względu na problemy zewnętrzne, co jest naszą hipotezą w niniejszym badaniu.

Struktura tej pracy jest następująca. W kolejnej części prezentujemy metody estymacji złożoności strukturalnej szeregów czasowych. W części trzeciej prezentujemy dane eksperymentalne oraz wyniki naszych badań wraz z dyskusją. Niniejszą pracę kończymy wnioskami i podsumowaniem wraz z propozycjami dalszych badań.

## 2. Metodologia

W tej części niniejszego opracowania przedstawiamy metodę pomiaru złożoności strukturalnej (jednocześnie przewidywalności) szeregów czasowych, którą będziemy używać do badań praktycznych na danych giełdowych. Wpierw należy jednak wprowadzić badania złożoności strukturalnej procesów stochastycznych na tle innych metod badania przewidywalności procesów stochastycznych, w tym także metod testowania efektywności rynku.

Historia modelowania szeregów czasowych na potrzeby predykcji sięga co najmniej wynalezienia autoregresji przez Yule'a w 1927 roku. Od tej pory liczba strategii używanych do tego celu rośnie w ogromnym tempie, w tym w ostatnich 30 latach szczególnie mocno jest to widoczne w badaniach nad rynkami finansowymi. Zajmujemy się problemem kwantyfikowania struktury predykccyjnej obecnej w szeregach czasowych opisujących zwroty z akcji na warszawskiej giełdzie. Metody służące do takiej kwantyfikacji mogą być podzielone na dwie grupy: analiza błędów oparta na modelach i analiza informacji niezakładająca modelu. Większość metod używanych w tym celu należy do pierwszej grupy. Tymczasem metody z grupy pierwszej nie mogą powiedzieć nic o przewidywalności finansowych szeregów czasowych, jeśli przewidywalność rozumiemy tak, jak jest ona definiowana w matematyce systemów dynamicznych. Metody te mogą jedynie

opisywać moc predykcyjną danego modelu czy strategii dla danej klasy szeregów czasowych. A zatem badacze używający tych metod nie mogą powiedzieć, że danych szeregów czasowych nie da się przewidywać, a jedynie, że ich modele tego nie potrafią.

Bardziej formalnie, pierwsza z prezentowanych grup analizuje rozkłady błędów badanych modeli. Metody te opisują *de facto* lokalną moc predykcyjną, aproksymując rozkład błędów dla różnych regionów szeregów czasowych, używając lokalnej predykcji w próbie. Te rozkłady są następnie używane jako estymatory błędów poza próbą w tych samych regionach. Takie podejście pokazuje, że różne części szeregów czasowych mogą mieć różną lokalną moc predykcyjną dla badanego modelu. Analiza lokalnych błędów opartych na modelu działa całkiem dobrze w analizowaniu lokalnej niepewności predykcyjnej w odniesieniu do ustalonego modelu. Metody te nie mogą natomiast kwantyfikować przewidywalności samych szeregów czasowych, a więc w szczególności nie mogą być używane do wyciągania wniosków co do struktury tych szeregów w odniesieniu do przewidywania przy użyciu innych modeli.

Analiza globalnych błędów opartych na modelu podąża w tym kierunku. Używa rozkładów błędów poza próbą, obliczonych dla klasy modeli, ustalając, który z modeli jest najlepszy. Często bada się także, czy błędy modeli mają rozkład normalny. Jeśli nie mają, sugeruje to, że w badanych szeregach czasowych istnieją struktury, które nie są odzwierciedlane i używane przez badany model. Głównym problemem tego podejścia jest brak ogólności. Normalność rozkładu błędów oznacza jedynie, że model odzwierciedla strukturę danych najlepiej jak to możliwe w ramach danego paradygmatu (np. najlepsze dopasowanie liniowe dla danych nieliniowych). Taka procedura nie wskazuje, czy inne podejście mogłoby przynieść lepsze rezultaty. Proponowane są także metody pośrednie, które przechodzą od lokalnych do globalnych rozwiązań.

Wszystkie te metody oparte są na konkretnych założeniach dotyczących procesu generującego dane i wiedzy o tym, co dzieje się z błędami w przypadku, gdy te założenia są lub nie są spełnione. Analiza przedstawiona w prezentowanej rozprawie dotyczy drugiej grupy: analizy informacji niezakładającej żadnego modelu. Analiza taka odsuwa się od powyższych ograniczeń. A zatem podejściem takim mierzy się właściwą danym szeregom czasowym złożoność strukturalną, która jest wyznacznikiem ogólnej przewidywalności tych szeregów czasowych. Po ustaleniu złożoności strukturalnej szeregów czasowych ich przewidywalność może być korelowana z mocą predykcyjną konkretnych modeli. Innymi słowy, tak obliczona przewidywalność jest niejako wyznacznikiem rzeczywistej mocy danych modeli, a więc odniesioną do tego, co jest możliwe ze względu na strukturę badanych szeregów czasowych.

Metody badania złożoności strukturalnej szeregów czasowych oparte są na pojęciu redundancji, które formalnie określa, jak informacje propagowane są w czasie

w szeregach czasowych: ile informacji o przyszłych wartościach szeregów czasowych zawarty jest w ich przeszłych stanach. Innymi słowy jest to wzajemna informacja pomiędzy przeszłymi stanami a stanem obecnym. Redundancja procesów stochastycznych charakteryzujących się własnością *iid* jest równa 0, gdyż wszystkie obserwacje w takim procesie są od siebie niezależne. Z drugiej strony dla systemów deterministycznych (nawet chaotycznych), redundancja jest wysoka (z maksimum w limicie). Oznacza to, że mogą one być idealnie przewidywane, jeśli posiadamy wystarczająco długą historię. W praktyce trudno jest estymować redundancję szeregów czasowych. Aby tego dokonać należy znać albo entropię Kołmogorowa-Sinaj, albo wartości wszystkich dodatnich wykładników Lapunowa danego systemu. Obydwie procedury są dosyć trudne, a druga jest praktycznie niemożliwa dla danych finansowych, dla których wykładniki Lapunowa nie są skończone.

Takie podejście do analizy efektywności procesów cenotwórczych bliższe jest testom hipotezy błędzenia losowego takim, jak: testy serii, testy ilorazów wariancji, testy korelacji rang Spearmana, itd. Testy te często koncentrują się jednak na testowaniu hipotez, a nie na kwantyfikacji przewidywalności jako takiej, o czym jeszcze poniżej. Najbliżej naszego podejścia są te elementy analizy technicznej odwołujące się do poszukiwania wzorców w danych giełdowych, które są właściwie tożsame z poniżej przedstawioną metodą badania złożoności strukturalnej procesów cenotwórczych. Siłą naszego podejścia jest mocne zakorzenienie w matematycznej definicji przewidywalności. Zauważmy, że przewidywalność systemu definiuje się w matematyce jako możliwość znajomości przyszłego stanu tego systemu w chwili obecnej. Jak okaże się poniżej – wzór (3) – definicja stopy entropii Shannona może być wprost interpretowana jako niepewność co do ostatniego stanu systemu, gdy znana jest cała jego historia. W niniejszej pracy, estymując złożoność strukturalną procesów cenotwórczych, badamy zatem bezpośrednio ich przewidywalność w sensie matematyki systemów dynamicznych.

Co więcej, stopa entropii jest łatwa w interpretacji. Zawiera się ona zawsze w dobrze zdefiniowanym przedziale (w naszym przypadku  $[0;2]$ ). Stopę entropii można także łatwo znormalizować, tak aby przedział ten miał postać  $[0;1]$ . Stopa entropii bliska zera oznacza maksymalną przewidywalność i najmniejszą złożoność strukturalną (cały proces stochastycznych jest wtedy w zasadzie jednym wzorcem), a zatem następny stan systemu może być w takim przypadku przewidywany z dobrym skutkiem przy stosunkowo krótkiej historii. Natomiast stopa entropii bliska maksimum oznacza minimalną przewidywalność procesu cenotwórczego i zarazem jego największą złożoność strukturalną (bogactwo, wielość wzorców w szeregach czasowych opisujących zmiany cen), i w takim przypadku skuteczna predykcja następnego stanu systemu wymaga znajomości bardzo długiej historii stanów tego systemu (do nieskończoności w limicie). Co ważne, wartości stopy entropii dla procesów cenotwórczych pokazują także rzeczywiste możliwości predykcji zmian cen. Pokazano, że precyzja predykcji następnej zmia-

ny ceny opartej o regułę maksymalizacji entropii jest silnie skorelowana ze stopą entropii danego procesu, tak że precyzja tej metody wzrasta wraz ze spadkiem stopy entropii (Fiedor 2014c). Ta silna korelacja pokazuje, że sensowne interpretacyjnie nie są tylko skrajne wartości stopy entropii, ale także całe spektrum (gdyż przyrosty przewidywalności są tożsame z predykcją powyżej wspomnianą metodą), co jest mocną stroną tej metodologii w porównaniu do metod takich jak test ilorazów wariancji, które w zasadzie koncentrują się na testowaniu słabej wersji hipotezy efektywnego rynku a nie na kwantyfikowaniu złożoności strukturalnej procesów cenotwórczych.

Miarą niepewności, a zatem także przewidywalności, jest entropia w sensie teorii informacji. Niska entropia wskazuje na wysoką pewność przyszłych stanów (przewidywalność), a wysoka entropia odwrotnie (na niską pewność i ilość dostępnych do badań informacji). Entropia, w teorii informacji, jest miarą niepewności zmiennej losowej. Entropia pojedynczej zmiennej losowej  $X$  dana jest wzorem:

$$H(X) = - \sum_i p(x_i) \log_2 p(x_i) \quad (1)$$

w którym suma dotyczy wszystkich możliwych realizacji zmiennej losowej  $\{x_i\}$  i ich prawdopodobieństw  $p(x_i)$  (Shannon 1948).

Stopa entropii, która jest główną miarą używaną w niniejszej pracy, została wprowadzona do literatury przez Claude'a Shannona. Stopa entropii uogólnia pojęcie entropii dla sekwencji zależnych zmiennych losowych. Dla stacjonarnego procesu stochastycznego  $X = \{X_i\}$  stopa entropii dana jest wzorami:

$$H(X) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, X_2, \dots, X_n) \quad (2)$$

$$H(X) = \lim_{n \rightarrow \infty} H(X_n | X_1, X_2, \dots, X_{n-1}) \quad (3)$$

Pierwszy z powyższych jest prawdziwy dla dowolnego procesu stochastycznego, natomiast drugi wzór wymaga stacjonarności procesu stochastycznego, z czym nie ma problemu w przypadku badanych logarytmicznych zmian cen. Prawa strona pierwszego z powyższych równań może być interpretowana w następujący sposób: stopa entropii mierzy niepewność w czasie, zakładając, że zaobserwowano kompletną historię aż do momentu  $n$ . Stopa entropii oznacza zatem średnią entropię każdej zmiennej losowej w procesie stochastycznym.

Estymacja entropii jest problemem o niezwyklej wadze, gdyż jedynie w niewielu zastosowaniach znana jest rzeczywista entropia. Zainteresowanie naukowców estymacją entropii to przede wszystkim ostatnie dwie dekady, głównie ze względu na istotność tej procedury w neurobiologii (przydatność entropii i innych technik opartych na teorii informacji w badaniach funkcjonowania mózgu, w szczególności analizy EEG; zob. Maciejewski *et al.* 2008).

Metody estymacji entropii mogą być podzielone na dwie odrębne grupy (Gao *et al.* 2006):

- Estymatory największej wiarygodności, badające rozkład empiryczny wszystkich fraz danej długości w analizowanych szeregach czasowych (na przykład za pomocą obliczania entropii rozkładu łańcucha Markowa  $n$ -tego stopnia). Wadą tego podejścia są wykładniczo rosnące wymagania co do długości próbki wraz z liniowo rosnącą długością fraz, dzięki czemu estymatory te nie są przydatne w badaniu długookresowych zależności. Takie zależności nie mogą być ignorowane w badaniach ekonomicznych, gdzie nie można wykluczyć istnienia zależności średniego i długiego zasięgu. W związku z tym metody te nie są popularne w badaniach ekonomicznych.
- Estymatory oparte na algorytmach kompresji danych, w szczególności algorytmach Lempela-Ziva (Farah *et al.* 1995; Kontoyiannis *et al.* 1998; Lempel, Ziv 1977) i Wążeń Drzew Kontekstów (Kennel *et al.* 2005; Willems *et al.* 1995). Obydwie wymienione techniki charakteryzują dobrą precyzją dla ograniczonej liczby obserwacji, dzięki czemu są lepiej przystosowane do badania długookresowych zależności w danych empirycznych. W niniejszym artykule używać będziemy estymatora opartego na algorytmie Wążeń Drzew Kontekstów, zdefiniujemy jednak obydwie metody.

Pierwsza z tych metod oparta jest na pracach Kołmogorowa. Złożoność w sensie Kołmogorowa (zdefiniowana jako rozmiar najmniejszego programu komputerowego, który zdolny jest wyprodukować taką sekwencję; zob. Cover, Thomas 1991) jest używana do estymowania stopy entropii. Algorytm Lempela-Ziva jest jedną z kilku praktycznych miar złożoności Kołmogorowa. Algorytm ten mierzy liniową złożoność. Został on zaproponowany przez Jacoba Ziva i Abrahama Lempela w 1976 roku (Lempel, Ziv 1976). W praktyce miara ta zlicza ilość wzorców występujących w badanym szeregu od lewej do prawej, a więc dla przykładu złożoność Lempela-Ziva szeregu:

$$s = 101001010010111110 \quad (4)$$

jest równa 8, jako że znajdujemy 8 różnych wzorców (patrzac od lewej do prawej) (Doganaksoy, Gologlu 2006):

$$1|0|10|01|010|0101|11|110| \quad (5)$$

Na podstawie tej miary badacze stworzyli wiele estymatorów stopy entropii Shannona. W niniejszym artykule używamy estymatora stworzonego przez Kontoyiannisa w roku 1998 (estymator  $\alpha$ ) (Kontoyiannis 1998). Estymator ten, dla szeregu czasowego  $s$  o długości  $n$  (gdzie  $s_i$  jest symbolem w szeregu  $s$  na pozycji  $i$ ) dany jest wzorem:



$$\hat{h}_{SM} = \left( \frac{1}{n} \sum_{i=1}^n A_i \right)^{-1} \log_2 n \quad (6)$$

gdzie  $A_i$  jest długością najkrótszego podszeregu zaczynającego się na pozycji  $s_i$ , który nie pojawia się jako podszereg wśród poprzednich i symboli, tj.  $s_0, \dots, s_{i-1}$ . Estymator ten jest szeroko używany w literaturze (Kennel *et al.* 2005) i pokazano, że ma lepsze własności statystyczne niż wcześniejsze estymatory oparte na algorytmie Lempela- Ziva (Kontoyiannis *et al.* 1998).

Efektywność estymatora zbadaliśmy poprzez obliczenie stopy entropii dla próbki złożonej z niezależnych realizacji zmiennej losowej o rozkładzie jednolitym przybierającej wartości z zakresu  $\{0,1,2,3\}$ . Teoretyczna entropia dla tej zmiennej jest równa:

$$H(P) = - \sum_{i=0}^3 \left( \frac{1}{4} \right) \log_2 \left( \frac{1}{4} \right) = 2 \quad (7)$$

Otrzymany wynik zależy od wielkości próbki, jakości generatora zmiennych losowych i efektywności estymatora. Na próbce wielkości 10000 realizacji zmiennej otrzymanych z generatora losowych liczb całkowitych, z losowością opartą na szumie atmosferycznym<sup>1</sup>, uzyskaliśmy wynik równy 1,95, co pokazuje, że estymator jest efektywny, jako że 2 jest maksimum możliwym do uzyskania przy założeniu perfekcyjnego generatora zmiennych losowych (Fiedor 2014a). Podobne wyniki efektywności  $\hat{h}_{SM}$  przy użyciu innych generatorów zmiennych losowych otrzymali inni badacze (Navet, Chen 2008).

Inną metodą estymacji stopy entropii Shannona jest tak zwany estymator Ważenia Drzew Kontekstów, o którym poniżej. Dla każdego dyskretnego, ergodycznego procesu stochastycznego  $X$  asymptotyczna zasada ekwipartycji (udowodniona dla skończonych, ergodycznych źródeł przez twierdzenie Shannona-McMillana-Breimana) stwierdza, że:

$$-\frac{1}{n} \log p(X_1^n) \rightarrow H(X) \text{ gdy } n \rightarrow \infty \quad (8)$$

gdzie  $p(X_1^n)$  jest prawdopodobieństwem procesu  $X_1^n$  ograniczonym do okresu  $\{1, \dots, n\}$ , natomiast  $H(X)$  to stopa entropii  $X$ , której istnienie zostało udowodnione dla wszystkich dyskretnych i stacjonarnych procesów (Fiedor 2014a). Powyższa zbieżność została udowodniona dla wszystkich przypadków z prawdopodobieństwem równym 1 (Cover, Thomas 1991). Dlatego też możliwa jest estymacja pośrednio, poprzez estymację prawdopodobieństwa długiej realizacji  $X$ .

<sup>1</sup> <http://www.random.org>

Algorytm Ważenia Drzew Kontekstów (Context Tree Weighting, CTW) jest algorytmem kompresji danych (Willems 1998; Willems *et al.* 1996, 1995), który może być także interpretowany jako Bayesowska procedura estymacji prawdopodobieństwa szeregu generowanego przez proces oparty na drzewie binarnym (Gao *et al.* 2008). Proces drzewa binarnego o głębokości  $D$  to stochastyczny proces binarny  $X$  o rozkładzie zdefiniowanym przez zbiór przyrostków  $S$  składający się z szeregów binarnych długości  $\leq D$  i wektor parametrów  $\Theta = (\Theta_s; s \in S)$ , gdzie każdy  $\Theta_s \in [0; 1]$ .

Jeśli dany szereg  $x_1^n$  został wygenerowany przez proces drzewa binarnego o głębokości  $\leq D$ , z nieznanym zbiorem przyrostków  $S^*$  oraz nieznanym wektorem parametrów  $\Theta^*$ , to możemy ustalić prawdopodobieństwo a priori  $\pi(S)$  dla każdego zbioru przyrostków  $S$  o głębokości  $\leq D$  oraz, znając  $S$ , możemy ustalić prawdopodobieństwo a priori  $\pi(\Theta|S)$  dla każdego wektora parametrów  $\Theta$ . Bayesowska aproksymacja rzeczywistego prawdopodobieństwa  $x_1^n$  (zakładając  $S^*$  i  $\Theta^*$ ) to prawdopodobieństwo mieszaniny:

$$\hat{P}_{D,mix}(x_1^n) = \sum_S \pi(S) \int P_{S,\theta}(x_1^n) \pi(\theta|S) d\theta \quad (9)$$

gdzie  $P_{S,\theta}(x_1^n)$  jest prawdopodobieństwem ( $x_1^n$ ), zakładając rozkład procesu drzewa binarnego ze zbiorem przyrostków  $S$  oraz wektorem parametrów  $\Theta$ . Wyrażenie zaprezentowane w równaniu (9) nie jest możliwe do bezpośredniego obliczenia, gdyż liczba zbiorów przyrostków głębokości  $\Theta$  jest rzędu  $2^D$ . Jest to nadmierna wartość dla zastosowań praktycznych przy obecnej technologii dla dowolnego  $D > 20$ .

Algorytm Ważenia Drzew Kontekstów jest efektywnym sposobem dla obliczenia prawdopodobieństwa mieszaniny z powyższego równania, przy konkretnym wyborze prawdopodobieństw a priori  $\pi(S)$ ,  $\pi(\Theta|S)$ . To prawdopodobieństwo dla  $S$  to

$$\pi(S) = 2^{-|S|-N(S)+1} \quad (10)$$

gdzie  $|S|$  to liczba elementów  $S$ , natomiast  $N(S)$  to liczba szeregów w  $S$  o długości mniejszej niż  $D$ . Jest to standardowe płaskie prawdopodobieństwo a priori na drzewie binarnym, używane rutynowo w teorii informacji. Przy założeniu zbioru przyrostków  $S$ , prawdopodobieństwo a priori  $\Theta$  jest wytworem rozkładu Dirichleta z parametrami  $(\frac{1}{2}, \frac{1}{2})$ .

Algorytm Ważenia Drzew Kontekstów pozwala obliczyć prawdopodobieństwo zdefiniowane powyżej z dużą dokładnością. Obliczenia te mogą być wykonane w czasie (oraz z pamięcią) rosnącym liniowo wraz z długością szeregu  $n$ . Dlatego też możliwe jest badanie przy znacznie większym  $D$ , niż jest to możliwe przy zastosowaniu wspomnianych estymatorów największej wiarygodności.

Wreszcie, mając szereg binarny  $(x_1^n)$ , estymator stopy entropii Shannona, oparty ma algorytmie Ważenia Drzew Kontekstów  $\hat{H}_{ctw}$ , dany jest wzorem:

$$\hat{H}_{ctw} = -\frac{1}{n} \log \hat{P}_{D,mix}(x_1^n) \quad (11)$$

gdzie  $\hat{P}_{D,mix}(x_1^n)$  jest prawdopodobieństwem mieszaniny zdefiniowanym powyżej (Kennel, Mees 2002; London *et al.* 2002). Także ten estymator został testom na próbkach losowych w celu testowania jego własności, podobnie jak algorytm Lempela-Ziva. Różnią się one zasadniczo jedynie innym obciążeniem, a zatem są przesunięte względem siebie liniowo, co nie zmienia w żaden sposób poniższej analizy. Algorytm Wążenia Drzew Kontekstów jest znacznie szybszy, dlatego też będzie używany w niniejszym opracowaniu.

### 3. Wyniki empiryczne i dyskusja

Ogólne konkluzje dotyczące złożoności strukturalnej finansowych szeregów czasowych przedstawione zostały w naszych wcześniejszych badaniach (Fiedor 2014a), a zatem mamy punkt odniesienia dla badań dotyczących konkretnych spółek. W szczególności dzienne zmiany cen są bardzo złożone (w okolicach wartości równych 2), natomiast wewnątrzdzienne zmiany cen są mniej złożone (bardziej przewidywalne, z wartościami często odbiegającymi znacznie od 2) (Fiedor 2014a). Chcąc badać wpływ negatywnej kondycji finansowej spółek (a właściwie informacji o takiej kondycji napływającej na rynek) na złożoność strukturalną zmian cen ich akcji, musimy badać szeregi czasowe opisujące zmiany cen akcji spółek, które w pewnym momencie ogłosiły publicznie informacje o upadłości. Wybraliśmy zatem 44 spółki, które w pewnym momencie ogłosiły upadłość, tak aby badać złożoność strukturalną zmian cen ich akcji w czasie, obserwując zmiany w okresie tuż przed ogłoszeniem upadłości. Wybrane spółki to: A.pl Internet SA, ABM Solid SA, Advadis SA, Alterco SA, BGE SA, Bomi SA, BUDOPOL-WROCŁAW SA, Budostal-5 SA, BUDUS SA, Call2Action SA, Cash Flow SA, Cool Marketing SA, D&D SA, Direct eServices SA, Dolnośląskie Surowce Skalne SA, Drewex SA, ENERGOMONTAŻ-POŁUDNIE SA, Euromark Polska SA, Europejski Fundusz Hipoteczny SA, Firma Handlowa Jago SA, Fota SA, Gant Development SA, GREENECO SA (Anti), Ideon SA, Intakus SA, Internetowy Dom Zdrowia SA, InwazjaPC SA, KCSP SA, Mediatel SA, Mew SA, Motor Trade Company SA, Nicolas Entertainment Group SA, Partex SA, PBG SA, Polskie Jadło SA, Positive Advisory SA, Promet SA, R&C Union, Richter Med SA, Sobet SA, SSI SA, Synkret SA, Waspol SA, Wilbo SA.

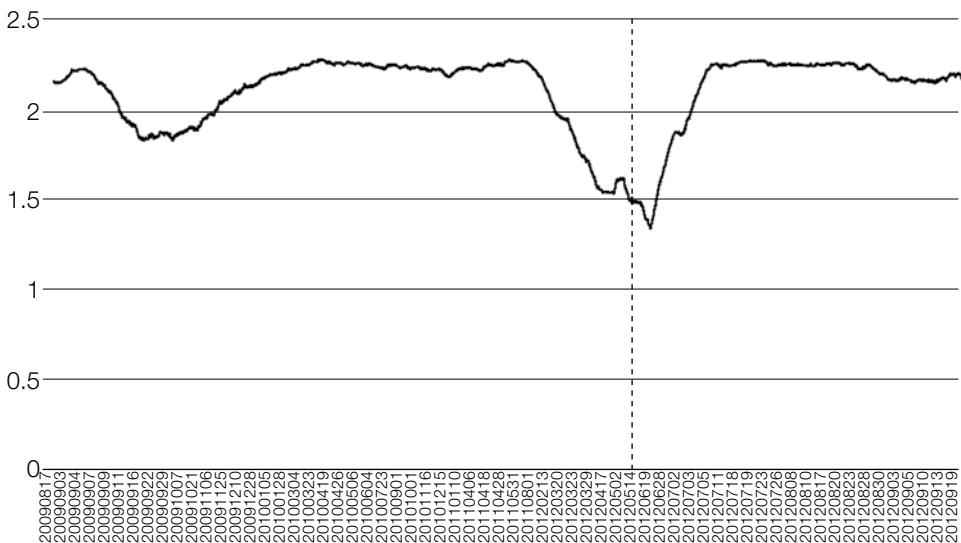
W tym miejscu należy wspomnieć powód, dla którego badamy zmiany cen, a nie same ceny akcji powyższych spółek. Szeregi czasowe opisujące ceny nie są stacjonarne, co wywołuje wiele problemów z analizą statystyczną. Dlatego też rutynowo bada się logarytmiczne zmiany cen. W niniejszym opracowaniu używamy bazy danych DM BOŚ zawierające historię cen próbkowanych co każdą zmianę ceny w całej historii

notowań danych spółek, od początku notowań aż do 5 lipca 2013 roku. Ceny po każdej transakcji ( $p$ ) zamieniamy na logarytmiczne zmiany cen, gdzie dla zmiany ceny  $t$ :

$$r_t = \ln(p_t - p_{t-1}) \quad (12)$$

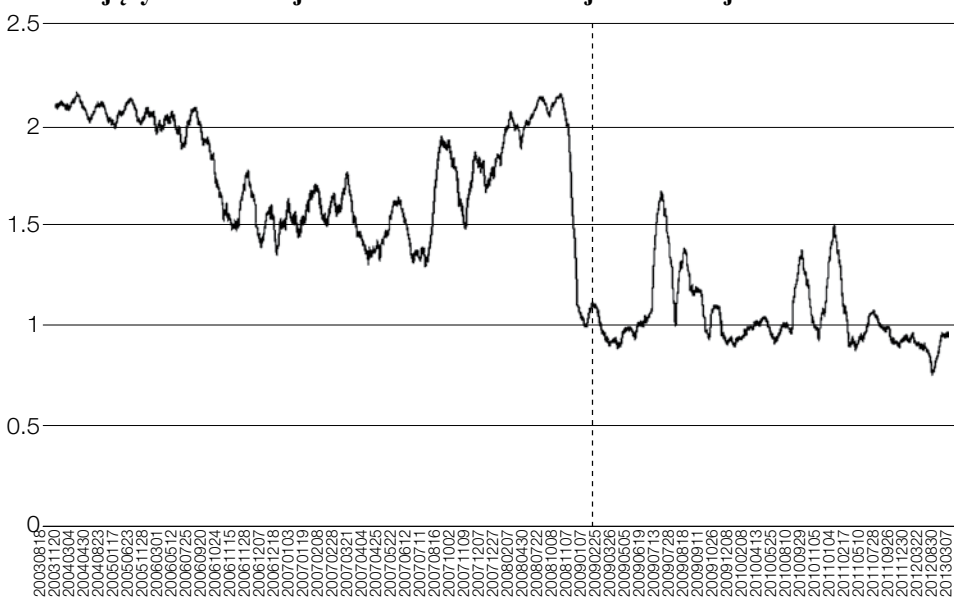
Te zmiany cen są na potrzeby opisanych algorytmów zamieniane na postać dyskretną. Dokonujemy tego, dzieląc logarytmiczne zmiany cen dla danego szeregu na cztery równe części (kwartyle). Taki podział nie wprowadza dodatkowych parametrów, które mogłyby zmieniać wyniki. Sama liczba części, na które dzielimy dane, nie ma wielkiego znaczenia (Fiedor 2014a), choć powinna być większa niż dwa (wtedy tracilibyśmy informacje o zmienności i jej grupowaniu, które są kluczowe; zob. Fiedor 2014b), i nie powinna być zbyt duża (taki wybór rozmywałby wzorce w danych). Taki sposób dyskretyzacji danych był używany w wielu badaniach z pozytywnym skutkiem (Navet, Chen 2008). Przy danych o czterech możliwych stanach stopa entropii może przybierać wartości od 0 dla skrajnie przewidywalnych procesów do około 2 dla procesów o największej złożoności strukturalnej (jak w hipotezie efektywnego rynku). Dla badania zmian złożoności strukturalnej w czasie używamy metody przesuwnego okna o długości 30 zmian cen. Długość taka jest stosunkowo niska dla używanego estymatora (Fiedor 2014a), ale zważywszy, że wyniki prezentowane dalej są z reguły uśredniane dla jakiegoś okresu, nie powinno to stanowić problemu. W przypadku Wykresów 1 i 2 prezentujemy wyniki dla przesuwnego okna o długości 1000 zmian cen, co pozwoli na bardziej stabilny wykres (alternatywą byłoby sztuczne łagodzenie wykresu przez uśrednianie po okresach). We wszelkich wynikach używamy końca okna jako daty, do której przypisujemy wynik estymacji złożoności strukturalnej dla tego okna, tak aby wyniki nie zawierały danych przyszłych w stosunku do prezentowanej daty. W tym miejscu należy zaznaczyć, że używamy czasu transakcyjnego, a nie czasu kalendarzowego czy zegarowego, a zatem szerokość okna w czasie rzeczywistym zależy od gęstości zmian cen w ciągu dnia. Nie powinno to jednak zmieniać znacząco analizy. Za datę publicznego ogłoszenia upadłości uznajemy datę złożenia wniosku o ogłoszenie upadłości.

**Wykres 1. Złożoność strukturalna procesu cenotwórczego w czasie dla spółki Anti.** Linią wertykalną oznaczony jest moment ogłoszenia upadłości. Widać wyraźne odejście od efektywności procesu cenotwórczego dla akcji tej spółki przed ogłoszeniem jej upadłości, co może sugerować, że znaczna część inwestujących w te akcje miała takie informacje wcześniej.



**Wykres 2. Złożoność strukturalna procesu cenotwórczego w czasie dla spółki Duda.**

Linią wertykalną oznaczony jest moment ogłoszenia upadłości. Widać wyraźne odejście od efektywności procesu cenotwórczego dla akcji tej spółki przed ogłoszeniem jej upadłości, co sugerować może, że znaczna część inwestujących w te akcje miała takie informacje wcześniej.



Najpierw chcemy ustalić, czy złożoność strukturalna procesów cenotwórczych zmienia się po publicznej informacji o upadłości. W tym celu porównujemy złożoność strukturalną w dwóch okresach, a zatem w okresie przed (I) i po (II) dacie ogłoszenia upadłości. Prezentowana w Tabeli 1 złożoność strukturalna jest uśredniana w tychże okresach. Dodatkowo prezentowany jest także średni wolumen w tych okresach. Aż w 41 z 44 badanych spółek przewidywalność (odwrotność złożoności strukturalnej) procesów cenotwórczych dotyczących ich akcji wzrosła po ogłoszeniu upadłości, a zatem wydaje się, że ogłoszenie upadłości wpływa na efektywność rynku akcji danej spółki. Przewidywalność jest mocno skorelowana z wolumenem, zarówno w badanym zbiorze spółek, jak i w badaniach szerokiego rynku (dla szerokiego rynku zbadano korelację między średnim wolumenem i średnią złożonością strukturalną dla 707 spółek z warszawskiej giełdy) korelacja ta mieści się w granicach 0,7–0,8 (należy w tym miejscu nadmienić, że sytuacja taka ma miejsce jedynie dla danych wewnątrzdziennej, nie istnieje istotna statystycznie korelacja między średnim wolumenem a przewidywalnością dla cen zamknięcia dnia). Dlatego też w Tabeli 1 pokazano także średnie wolumeny w tych okresach. Rzeczywiście w 41 z 44 spółek wzrósł średni wolumen, choć sytuacja nie jest jednoznaczna, gdyż nie są to te same spółki, dla których wzrosła przewidywalność. A zatem zmiany wolumenu wydają się jedynie częścią powodu, dla którego wzrasta przewidywalność po ogłoszeniu upadłości. Wydaje się, że może to mieć związek z ujednoliceniem oczekiwań inwestorów w związku z informacją o upadłości. Podobną sytuacją jest kryzys finansowy, który powoduje większe grupowanie spółek (mniejszą modularność) w sieciach finansowych.

**Tabela 1. Średnia złożoność strukturalna i średni wolumen transakcji w okresach przed (I) i po (II) złożeniu wniosku o ogłoszenie upadłości spółek giełdowych. W 41 z 44 badanych spółek ich procesy cenotwórcze są bardziej przewidywalne po tym ogłoszeniu w stosunku do okresu przed ogłoszeniem. Także w 41 z 44 spółek wzrósł średni wolumen, choć nie są to zawsze te same spółki. Dodatkowo istnieje silna (-0.8) dodatnia korelacja między średnim wolumenem a przewidywalnością procesów cenotwórczych, co ma jednak miejsce także na szerokim rynku.**

Nazwa	I	II	wol I	wol II
A.pl Internet SA	1.897	2.073	255	615
ABM Solid SA	2.067	1.514	1 026	4 146
Advadis SA	1.510	1.802	22 842	77 218
Alterco SA	2.072	2.030	555	2 116
BGE SA	2.034	1.719	2 656	3 278
Bomi SA	1.952	1.455	3 681	30 198
BUDOPOL-WROCŁAW SA	1.837	1.445	4 043	20 329
Budostal-5 SA	2.057	1.885	1 064	2 535
BUDUS SA	1.953	1.934	124	4 471

Call2Action SA	1.901	1.519	10 804	45 283
Cash Flow SA	2.026	1.739	941	4 080
Cool Marketing SA	1.862	1.784	3 912	12 643
D&D SA	2.058	1.612	2 812	8 220
Direct eServices SA	2.014	1.517	1 532	10 027
Dolnośląskie Surowce Skalne SA	1.978	1.794	1 528	5 025
Drewex SA	1.992	1.619	2 003	10 663
ENERGOMONTAŻ-POŁUDNIE SA	2.016	1.448	2 021	15 963
Euromark Polska SA	2.026	1.484	1 223	23 544
Europejski Fundusz Hipoteczny SA	1.701	1.781	14 070	5 850
Firma Handlowa Jago SA	1.957	1.423	4 057	31 102
Fota SA	2.017	1.997	591	3 734
Gant Development SA	1.953	1.436	1 647	20 806
GREENECO SA (Anti)	2.024	1.558	4 119	2 657
Ideon SA	1.566	1.550	53 205	71 431
Intakus SA	1.939	1.617	4 266	28 235
Internetowy Dom Zdrowia SA	2.091	1.853	290	1 073
InwazjaPC SA	1.858	1.344	6 906	7 110
KCSP SA	2.026	1.430	2 094	11 526
Mediatel SA	2.034	1.776	856	5 808
Mew SA	2.119	1.999	395	2 360
Motor Trade Company SA	1.689	1.480	9 952	28 040
Nicolas Entertainment Group SA	1.613	1.602	30 493	73 859
Partex SA	1.774	1.443	9 596	6 600
PBG SA	1.997	1.900	455	2 678
Polskie Jadło SA	1.774	1.465	10 411	32 591
Positive Advisory SA	2.063	2.017	2 660	4 714
Promet SA	1.999	1.630	1 444	6 966
R&C Union	1.992	1.542	1 974	10 948
Richter Med SA	2.114	1.897	419	727
Sobet SA	1.873	1.536	4 497	37 114
SSI SA	1.564	1.311	12 016	58 686
Synkret SA	2.168	2.135	73	4 453
Waspół SA	2.005	1.459	636	5 643
Wilbo SA	2.021	1.463	2 308	3 614

Powyższa analiza dotyczy całkiem szerokiego czasowo horyzontu, a zatem należy także przeanalizować okresy bliższe (krótsze) samej daty ogłoszenia upadłości. A zatem chcemy pokazać także, że przewidywalność zmian cen spółek ogłaszających upadłość zmienia się właśnie wokół daty ogłoszenia upadłości. Zbadaliśmy zatem także średnią złożoność strukturalną dla okresów miesiąca przed i po ogłoszeniu upadłości. Dla ponad 68% badanych spółek złożoność strukturalna procesów cenotwórczych dotyczących ich akcji spadła w takim krótszym horyzontie czasowym, co wzmacnia powyższe ustalenia.

Dodatkowo zbadaliśmy korelacje Pearsona dla badanych spółek pomiędzy złożonością strukturalną a logarytmicznymi zmianami cen, poziomem cen i wolumenem. Współczynniki te prezentujemy w Tabeli 2. Nie istnieje prawidłowość dotycząca korelacji między złożonością strukturalną procesów cenotwórczych, a poziomem logarytmicznych zmian cen. Należy zauważyć jednak silną negatywną korelację pomiędzy złożonością strukturalną a wolumenem. Mniej przewidywalne są zatem te okresy, w których większy jest wolumen sprzedaży, jest to naturalne, gdyż większa ilość zainteresowanych powoduje większą efektywność rynku. Ciekawa jest natomiast silna dodatnia korelacja między złożonością strukturalną a poziomem cen dla wielu badanych spółek. Bardziej przewidywalne są zatem okresy, gdy ceny są niższe.

**Tabela 2. Współczynniki korelacji liniowej Pearsona pomiędzy złożonością strukturalną a logarytmicznymi zmianami cen (A), poziomem cen (B) i wolumenem (C) w czasie.**

**Należy zauważyć znaczącą negatywną korelację pomiędzy złożonością strukturalną a wolumenem. Mniej przewidywalne są zatem te okresy, w których większy jest wolumen sprzedaży; jest to naturalne, gdyż większa liczba zainteresowanych powoduje większą efektywność rynku. Ciekawa jest natomiast silna dodatnia korelacja między złożonością strukturalną a poziomem cen dla wielu badanych spółek. Wtedy bardziej przewidywalne są okresy, w których ceny są niższe.**

Nazwa	A	B	C
A.pl Internet SA	0.069	- 0.292	0.080
ABM Solid SA	0.019	0.092	- 0.210
Advadis SA	- 0.057	0.029	0.076
Alterco SA	- 0.026	- 0.009	0.007
BGE SA	- 0.041	0.268	- 0.251
Bomi SA	0.093	0.328	- 0.305
BUDOPOL-WROCŁAW SA	0.033	0.334	- 0.164
Budostal-5 SA	0.187	0.410	- 0.417
BUDUS SA	- 0.010	- 0.020	- 0.269
Call2Action SA	0.174	0.461	- 0.465
Cash Flow SA	0.043	0.136	- 0.197
Cool Marketing SA	0.006	- 0.387	- 0.027
D&D SA	0.052	0.407	- 0.421
Direct eServices SA	0.144	0.170	- 0.269
Dolnośląskie Surowce Skalne SA	- 0.006	0.250	- 0.064
Drewex SA	0.006	0.528	- 0.363
ENERGOMONTAŻ-POLUDNIE SA	0.072	0.092	- 0.268
Euromark Polska SA	0.129	- 0.158	- 0.201
Europejski Fundusz Hipoteczny SA	0.009	0.342	- 0.308
Firma Handlowa Jago SA	0.066	0.361	- 0.321
Fota SA	0.060	- 0.093	- 0.028



Gant Development SA	0.010	0.183	- 0.351
GREENECO SA (Anti)	0.083	0.323	0.046
Ideon SA	0.073	0.093	- 0.118
Intakus SA	0.078	0.291	0.159
Internetowy Dom Zdrowia SA	0.062	0.149	0.153
InwazjaPC SA	0.301	0.599	0.026
KCSP SA	0.206	0.123	- 0.498
Mediatel SA	- 0.003	0.124	- 0.129
Mew SA	0.034	0.087	- 0.071
Motor Trade Company SA	0.004	0.483	- 0.289
Nicolas Entertainment Group SA	- 0.039	0.251	- 0.120
Partex SA	0.115	0.257	- 0.232
PBG SA	- 0.008	0.099	- 0.078
Polskie Jadło SA	0.046	0.398	- 0.308
Positive Advisory SA	0.032	0.115	0.051
Promet SA	0.057	0.341	- 0.380
R&C Union	0.151	0.231	- 0.203
Richter Med SA	0.220	0.197	- 0.313
Sobet SA	- 0.048	0.367	- 0.567
SSI SA	0.013	0.136	- 0.007
Synkret SA	0.128	0.088	- 0.242
Wspol SA	0.250	0.187	- 0.309
Wilbo SA	0.015	0.319	- 0.103

Mając już obraz ogólny wpływu upadłości i jej ogłaszania na procesy cenotwórcze, pokażemy jeszcze w większym zakresie najciekawsze przypadki spośród analizowanych 44 spółek. Na Wykresie 1 prezentujemy złożoność strukturalną w czasie dla spółki Anti, wraz z linią pionową oznaczającą moment ogłoszenia o upadłości. Analogiczne dane dla spółki Duda zostały przedstawione na Wykresie 2. Należy nadmienić, że wykresy te są liniowe względem zmian cen, a nie względem czasu kalendarzowego.

W obydwu przypadkach widać, że ogólny obraz przekazany powyżej nie jest zawsze prawdziwy. Procesy cenotwórcze obydwu spółek stały się zdecydowanie bardziej przewidywalne jeszcze przed publicznym ogłoszeniem informacji o upadłości. A zatem istnieje podejrzenie, że informacje te były dostępne znaczącej części inwestorów zaangażowanym w handel akcjami tych spółek. Obraz ten wpisuje się także w ogólny trend spadku złożoności strukturalnej w okolicach informacji o upadłości pokazany w Tabeli 1.

#### 4. Podsumowanie

W niniejszym artykule zaprezentowaliśmy metodę badania złożoności strukturalnej zmian cen akcji giełdowych. Złożoność taką można traktować jako przeciwieństwo przewidywalności (w sensie limitu zdolności predykcyjnych dla danego procesu cenotwórczego). Zbadaliśmy, na przykładzie 44 spółek z GPW i New-Connect, które podległy procesowi upadłości, jak negatywna sytuacja finansowa spółek, a przede wszystkim informacje o niej, wpływają na efektywność procesów cenotwórczych dotyczących tych spółek, mierzoną ich złożonością strukturalną. Pokazaliśmy, że złożoność strukturalna po takim ogłoszeniu częściej spada, niż wzrasta, a zatem informacje o upadłości wpływają raczej na zwiększenie się przewidywalności zmian cen, co powodowane być może większą synchronizacją inwestorów w obliczu takich informacji. Pokazaliśmy także, że nie można mówić o ogólnej odczuwalnej zmianie złożoności przed samym ogłoszeniem, tak aby można je było przewidywać. Natomiast na przykładzie dwóch spółek: Anti i Duda, pokazaliśmy, że w niektórych przypadkach procesy cenotwórcze zmieniają się co do ich przewidywalności jeszcze przed ogłoszeniem upadłości, a zatem pokazuje to, że być może inwestorzy handlujący tymi spółkami mieli takie informacje, które były w posiadaniu większości inwestorów, lub też w posiadaniu istotnych inwestorów, do których działań dostosowali się następnie pozostali inwestorzy. Dalsze badania powinny spoglądać na wpływ innych zewnętrznych zdarzeń na złożoność strukturalną procesów cenotwórczych na warszawskiej giełdzie. Do dalszych badań mogą także zostać także użyte inne miary złożoności strukturalnej, takie jak chociażby entropia permutacyjna.

#### Bibliografia

- Alanyali, Merve, Helen Susannah Moat i Tobias Preis. 2013. „Quantifying the relationship between financial news and the stock market”. *Scientific Reports* 3, 3578.
- Cover, Thomas i Joy A. Thomas, 1991. *Elements of Information Theory*. New York: Wiley.
- Doğanaksoy Ali i Faruk Göloğlu. 2006. „On Lempel-Ziv Complexity of Sequences”, *Lecture Notes in Computer Science* 4086: 180–189.
- Farah, Martin, Michiel Noordewier, Serap Savari, Larry Shepp, Abraham Wyner i Jacoby Ziv. 1995. „On the entropy of DNA: algorithms and measurements based on memory and rapid convergence”. W: *SODA'95: Proceedings of the Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*, 48–57.
- Fiedor, Paweł. 2014. „The Social Dynamics of the Peter Principle”. *Journal of Engineering Science & Technology Review* 8: 56–60.

- Fiedor, Paweł. 2014a. „Frequency Effects on Predictability of Stock Returns”. W: *Proceedings of the IEEE Computational Intelligence for Financial Engineering & Economics 2014*, 247–254.
- Fiedor, Paweł. 2014b. „Information-theoretic approach to lead-lag effect on financial markets”, *European Physical Journal B* 87, 168.
- Fiedor, Paweł. 2014c. „Maximum Entropy Production Principle for Stock Returns”, arXiv: 1408.3728.
- Fiedor, Paweł i Artur Hołda A. 2015. „Istota ekonofizyki jako dziedziny naukowej”. *Studia Ekonomiczne*, w druku.
- Gao, Yun, Ioannis Kontoyiannis i Elie Bienenstock. 2006. „From the entropy to the statistical structure of spike trains”. W: *2006 IEEE International Symposium on Information Theory*, 645–649.
- Gao, Yun, Ioannis Kontoyiannis i Elie Bienenstock. 2008. „Estimating the Entropy of Binary Time Series: Methodology, Some Theory and a Simulation Study”, *Entropy* 10: 71–99.
- Kennel, Matthew B., Jonathon B. Shlens, Henry D.L. Abarbanel, E.J. Chichilnisky. 2005. „Estimating entropy rates with Bayesian confidence intervals”. *Neural Computations* 17: 1531–1576.
- Kennel, Matthew B. i Alistair Mees. 2002. „Context-tree modeling of observed symbolic dynamics”. *Physical Review E* 66: 056209.
- Kontoyiannis, Ioannis. 1998. „Asymptotically Optimal Lossy Lempel-Ziv Coding”. W: *Proceedings of the 1998 IEEE International Symposium on Information Theory*, 273.
- Kontoyiannis, Ioannis, P.H. Algoet., Y.M. Suhov i A.J. Wyner. 1998. „Nonparametric entropy estimation for stationary processes and random fields, with applications to English text”. *IEEE Transactions on Information Theory* 44(3): 1319–1327.
- Kristoufek, Ladislav. 2013. „Can Google Trends search queries contribute to risk diversification? ” *Scientific Reports* 3, 2713.
- Lempel, Abraham i Jacob Ziv. 1976. „On the complexity of finite sequences”, *IEEE Transactions on Information Theory* 22(2): 75–81.
- Lempel, Abraham i Jacob Ziv. 1977. „A Universal Algorithm for Sequential Data Compression”. *IEEE Transactions on Information Theory* 23(3): 337–343.
- London, Michael, Adi Schreiber, Michael Häusser, Matthew E. Larkum i Idan Segev. 2002. „The information efficacy of a synapse”. *Nature Neuroscience* 5: 332–340.
- Maciejewski, Artur, Mirosław Łątka i W. Jernajczyk. 2008. „Zastosowanie metody empirycznej dekompozycji modalnej i złożoności Lempel’a-Ziv’a do analizy EEG chorych na schizofrenię”. *Medycyna Dydaktyka Wychowanie* 60: 90–94.
- Mantegna, Rosario N. i Eugene H. Stanley. 2000. *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge: Cambridge University Press.
- Mestyán, Márton, Taha Yasseri i János Kertész. 2013. „Early prediction of movie box office success based on Wikipedia activity big data”, *PloS ONE* 8: e71226.

- Moat, Helen S., Chester Curme, Adam Avakian A., Dror Y. Kenett, Eugene Stanley, Tobias Preis. 2013. „Quantifying Wikipedia usage patterns before stock market moves”. *Scientific Reports* 3: 1801.
- Navet, Nicolas i Shu-Heng Chen. 2008. „On Predictability and Profitability: Would GP Induced Trading Rules be Sensitive to the Observed Entropy of Time Series? ”. *Natural Computing in Computational Finance* 100: 197–210.
- Preis, Tobias., Helen S. Moat i Eugen Stanley. 2013. „Quantifying trading behavior in financial markets using Google Trends”. *Scientific Reports* 3: 1684.
- Rosser, Barkley J. 2008 „Econophysics and Economic Complexity”. *Advances in Complex Systems* 11 (745): 745–760.
- Samuelson, Paul A. 1965. „Proof That Properly Anticipated Prices Fluctuate Randomly”. *Industrial Management Review* 6(2): 41–49.
- Shannon, Claude E. 1948. „A Mathematical Theory of Communication”. *Bell Systems Technical Journal* 27: 379–423, 623–656.
- Tumminello, M., T. Aste, T. Di Matteo i R.N. Mantegna. 2005. „A tool for filtering information in complex systems”. *Proceedings of the National Academy of Science of the United States of America* 102(3): 10421–10426.
- Willems, Frank J.M. 1998. „The context-tree weighting method: Extensions”. *IEEE Transactions on Information Theory* 44(2): 792–798.
- Willems, Frank J.M., Yuri M. Shtarkov, Tjalling Tjalkens. 1995. „The Context-Tree Weighting Method: Basic Properties”. *IEEE Transactions on Information Theory* 41(3): 653–664.
- Willems, Frank J.M., Yuri M. Shtarkov, Tjalling Tjalkens. 1996. „Context weighting for general finite-context sources”. *IEEE Transactions on Information Theory* 42(5): 1514–1520.

# The Effects of Bankruptcy on the Structural Complexity of the Price Changes on WSE

## Abstract

In this study we present a method of analysing market efficiency using information theory. The efficiency of a given market is studied by the degree to which redundancy is present in the time series describing stock returns, while the particular tool used is called Shannon's entropy rate, and can be interpreted as a measure of the predictability of stock returns (understood as the limits of prediction). We use this method to analyse time series describing logarithmic returns of chosen companies listed at Warsaw Stock Exchange, which have undergone bankruptcy. There exists a body of research analysing the efficiency of the whole market, but there are no detailed studies analysing how strongly negative economic situation of a company (and particularly information about this situation) affects the efficiency of price formation processes with regards to the shares of this company, and how it affects the predictability of the changes in the prices of these shares. The review presented in this study, based on 44 stocks, is meant to be a prelude to many detailed studies of the influence of effects of events outside of the stock market on the structural complexity of the price formation processes themselves.

**Keywords:** econophysics, complexity, bankruptcy, efficiency, finance

**JEL Codes:** G14, G33

**DOI:** <http://dx.doi.org/10.17451/eko/41/2015/75>