

GRZEGORZ PERCZAK, PIOTR FISZEDER

ESTYMACJA WARIANCJI ARYTMETYCZNEGO RUCHU BROWNA NA PODSTAWIE ZNANYCH WARTOŚCI MINIMUM, MAKSIMUM, KOŃCOWEJ ORAZ DRYFU¹

1. WPROWADZENIE

Zmienność jest jednym z ważniejszych pojęć współczesnych finansów. Jej znaczenie w teorii finansów jest fundamentalne. Wynika zarówno z teorii finansów, jak i zastosowań praktycznych. Zmienność ma zastosowanie zarówno w klasycznej teorii portfela, wycenie opcji czy analizach obejmujących problem pomiaru ryzyka. Wariancja logarytmicznej stopy zwrotu instrumentu finansowego może być miarą niepewności co do wielkości przyszłych zmian ceny tego instrumentu. Estymacja tego parametru dokonywana jest na podstawie historycznych notowań rynkowych. W praktyce występuje jednak zazwyczaj problem z dostępem do szczegółowych danych. W ogólnodostępnych publikowanych notowaniach cen instrumentów finansowych podawane są najczęściej ceny: otwarcia, minimalna, maksymalna oraz zamknięcia, jakie zostały zaobserwowane w danym dniu. Z jednej strony, tak niepełny zbiór danych utrudnia precyzyjne oszacowanie wariancji. Z drugiej strony, bardzo często w analizie danych finansowych wykorzystuje się szeregi czasowe tylko cen zamknięcia, co powoduje, że estymator wariancji konstruowany jest na podstawie jeszcze bardziej skąpego zbioru danych składającego się z dziennych stóp zwrotu. Celowe wydaje się zatem przedstawienie problemu wykorzystania dodatkowych informacji o minimalnej i maksymalnej dziennej cenie rynkowej w estymacji dziennej wariancji, którym to terminem określano w dalszej części pracy wariancję stopy zwrotu zaobserwowaną między zamknięciami notowań w kolejnych dniach.

Literatura naukowa dostarcza wielu metod estymacji dziennej wariancji. Często różnią się one przyjętymi założeniami, a także wykorzystują różnego typu dostępne notowania. Konstruowane w ten sposób estymatory różnią się również znacząco efektywnością. W niniejszym opracowaniu zostaną bliżej omówione tylko te, które wykorzystują znajomość co najwyżej cen otwarcia, minimalnej, maksymalnej i zamknięcia. Przedmiotem zainteresowania nie są estymatory konstruowane na podstawie śróddziennych stóp zwrotu (patrz np. Doman, 2011).

¹ Praca została sfinansowana ze środków Narodowego Centrum Nauki projekt numer 2012/05/B/HS4/00675 nt. „Modelowanie i prognozowanie zmienności – wykorzystanie dodatkowych informacji o cenach minimalnych i maksymalnych”.

Analiza publikacji zawierających metody estymacji dziennej wariancji: Parkinson (1980), Garman, Klass (1980), Rogers, Satchell (1991), Kunitomo (1992), Yang, Zhang (2000) prowadzi do wniosku, że w każdej z nich zastosowano inne podejście modelowe. Zostanie zatem podjęta próba kompleksowego spojrzenia na ten problem i dokonany zostanie przegląd najbardziej popularnych estymatorów, skonstruowanych na podstawie tego samego aparatu matematycznego. Prezentacja istniejącego stanu wiedzy z jednej wspólnej perspektywy pozwoli ocenić wady i zalety istniejących estymatorów, wyznaczyć analitycznie ich efektywność, ocenić zasadność założeń, na podstawie których są konstruowane, ocenić własności estymatorów w sytuacji, gdy nie są spełnione założenia modelowe, a także stać się inspiracją do poszukiwań innych sposobów estymacji wariancji.

Podstawą prowadzonych w pracy analiz będzie rozkład łączny trójwymiarowego wektora, którego współrzędnymi są zmienne losowe minimum, maksimum i wartości końcowej arytmetycznego ruchu Browna. Znajomość rozkładu tego wektora umożliwia obliczenie wartości oczekiwanych praktycznie dowolnych funkcji współrzędnych tego wektora. Zostanie pokazane, że niektóre z tych wartości oczekiwanych są równe wariancji procesu, uzyskany zostanie zatem pewien zbiór estymatorów tego parametru. To z kolei umożliwi zarówno sprawdzenie nieobciążoności znanych estymatorów, jak również obliczenie ich efektywności oraz konstrukcję nowego, nieobciążonego estymatora. Ponadto, zostanie pokazane, że jest on efektywniejszy od wielu innych dotychczas zaproponowanych estymatorów, pomimo że nie wymaga spełnienia restrykcyjnych założeń.

Wspólną cechą podstawowych, znanych estymatorów dziennej wariancji stopy zwrotu jest fakt, że albo zakładają one uproszczony model, gdy dryf ma wartość zerową, albo ich konstrukcja nie wykorzystuje znajomości wartości dryfu. Wydaje się natomiast, że wykorzystanie informacji o wartości dodatkowego parametru powinno zwiększyć efektywność estymatora. W opracowaniu zostaną przedstawione wyniki dotyczące wykorzystania takiego estymatora do estymacji wariancji procesu obserwowanego na rynku kapitałowym. Zostanie pokazane, że szacunki zmienności konstruowane na podstawie estymatorów wykorzystujących dodatkowo znajomość dziennych cen minimalnych i maksymalnych są bardziej trafne niż te, które konstruuje się na podstawie modelu GARCH. Estymator konstruowany na podstawie dodatkowych informacji o cenach minimalnych i maksymalnych może być w przyszłości wykorzystany do budowy modeli GARCH, dzięki czemu będzie możliwe uzyskanie jeszcze trafniejszych szacunków zmienności.

Układ artykułu jest następujący. W części drugiej zostanie przedstawiony rozkład łączny zmiennych losowych minimum, maksimum i wartości końcowej arytmetycznego ruchu Browna, a następnie zostanie wyznaczona jego gęstość i funkcja charakterystyczna oraz wartości oczekiwane iloczynów potęg tych zmiennych losowych. Dodatkowo obliczone zostaną wartości oczekiwane funkcji zmiennych losowych minimum, maksimum i wartości końcowej dla niezerowej wartości dryfu. Część trzecia zawierać będzie omówienie znanych estymatorów dziennej wariancji. Wyznaczona

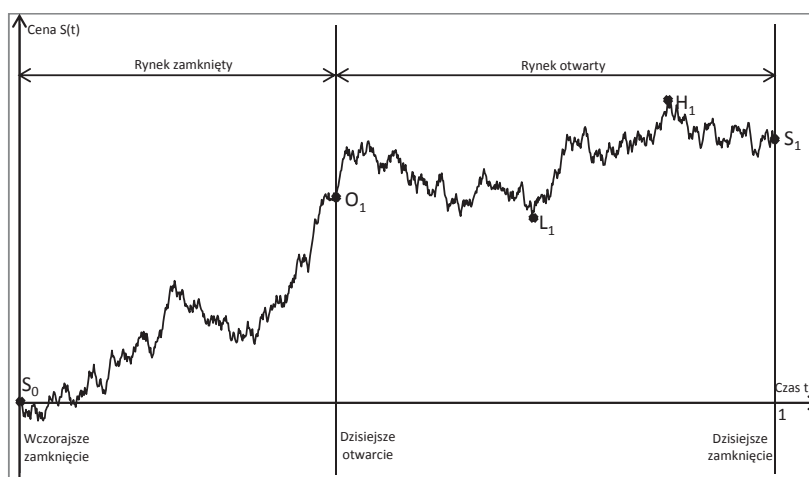
będzie nie tylko ich efektywność, lecz także wartość obciążenia, w sytuacji gdy nie są spełnione założenia modelowe, na podstawie których były one konstruowane. Informacje przedstawione w części drugiej zostaną wykorzystane do propozycji nowego estymatora dziennej wariancji. Estymator ten zostanie następnie zastosowany w badaniu dotyczącym szacowania zmienności stóp zwrotu indeksu WIG20 notowanego na GPW w Warszawie. Opracowanie kończy podsumowanie. Artykuł powstał na podstawie fragmentów pracy Perczaka (2013).

2. TRÓJWYMIAROWY WEKTOR LOSOWY MINIMUM, MAKSIMUM I WARTOŚCI KOŃCOWEJ RUCHU BROWNA

2.1. ZAŁOŻENIA ORAZ OZNACZENIA

W opracowaniu przyjęto, że: t jest ustaloną liczbą rzeczywistą dodatnią, jednostką czasu t jest jeden dzień (czyli wartość $t = 1$ odpowiada jednej dobie), zmienna τ spełnia warunek $0 \leq \tau \leq t$, $\tau = 0$ oznacza moment rozpoczęcia notowań w bieżącym okresie, który jest jednocześnie momentem otwarcia rynku w bieżącym okresie, B_τ jest standardowym procesem Wienera, S_τ jest ceną instrumentu finansowego odnotowaną w chwili τ . Stopa zwrotu, zdefiniowana jako $x_\tau = \ln(S_\tau/S_0)$ jest realizacją arytmetycznego ruchu Browna: $X_\tau = \mu\tau + \sigma B_\tau$ w punkcie $\tau = t$. Ponadto przyjęto: $C_t := \max_{0 \leq \tau \leq t} X_\tau$, $A_t := \min_{0 \leq \tau \leq t} X_\tau$.

W praktyce najczęściej zakłada się, że rozpatrywany okres t jest równy jednej dobie. Problem wyznaczenia minimum i maksimum dziennej logarytmicznej stopy zwrotu jest nieco bardziej złożony. Jego ilustrację przedstawia rysunek 1.



Rysunek 1. Przykładowa realizacja ceny instrumentu finansowego w ciągu dnia
Źródło: opracowanie własne na podstawie pracy Garmana, Klassa (1980, str. 70).

Z notowań w ciągu dnia zazwyczaj ogólnodostępnych jest tylko pięć wielkości:

S_0 – wczorajsza cena zamknięcia,

O_1 – dzisiejsza cena otwarcia,

H_1 – odnotowane w dniu dzisiejszym maksimum ceny,

L_1 – odnotowane w dniu dzisiejszym minimum ceny,

S_1 – dzisiejsza cena zamknięcia.

Wykorzystanie informacji dotyczących cen, jakie zaobserwowano tylko w okresie od momentu rozpoczęcia do zakończenia notowań rynkowych, nie pozawala na prawidłowe wyznaczenie zmienności dobowej. W takim wypadku można co najwyżej oszacować zmienność jaka miała miejsce w czasie funkcjonowania rynku. Wartości notowań L_1 i H_1 zostały ustalone dla okresu, w którym rynek był otwarty, a nie dla całej doby, co ilustruje rysunek 1. W związku z tym konieczne jest przededefiniowanie wartości minimalnej i maksymalnej tak, aby obejmowały one okres pomiędzy dwoma kolejnymi zamknięciami notowań (czyli pełnej doby). Ostatecznie zostaną zatem przyjęte następujące definicje dziennych stóp zwrotu minimum i maksimum $A_t := \min(\ln L_1, \ln S_0) - \ln S_0$, $C_t := \max(\ln H_1, \ln S_0) - \ln S_0$. Zmienna $X_t := \ln(S_t/S_0)$ będzie określana jako stopa zwrotu wartości końcowej. Ponadto, przyjęto założenie, że wariancja procesu jest stała w ciągu dnia, a jednocześnie dopuszczalna jest jej zmienność w kolejnych dniach. Z tego powodu ustalony został warunek: $0 < t \leq 1$.

2.2. ROZKŁAD ŁĄCZNY ZMIENNYCH LOSOWYCH STÓP ZWROTU MAKSYMUM, MINIMUM I WARTOŚCI KOŃCOWEJ RUCHU BROWNA

Znajomość rozkładu łącznego wektora trzech zmiennych losowych (A_t, C_t, X_t) umożliwi wyznaczenie wartości oczekiwanych wybranych funkcji tych zmiennych. Ponieważ rozkład ten ma skomplikowaną postać, zatem w pierwszej kolejności zaprezentowane zostaną rozkłady łączne wektorów (C_t, X_t) oraz (A_t, X_t) .

Istnieje wiele opracowań, w których omawiany jest problem rozkładu łącznego wektorów (C_t, X_t) , np. Cox, Miller (1965), Harrisson (1985). Zagadnienie to jest również bardzo często omawiane przy wycenie opcji barierowych (patrz np. Li, 1998; Jakubowski i inni, 2006).

Rozważane jest zdarzenie losowe $\{C_t \leq c, X_t \leq x\}$ gdy $x \leq c$ i $0 \leq c$. Dystrybuanta rozkładu łącznego wektora zmiennych losowych (C_t, X_t) opisana jest formułą (patrz Harrisson, 1985, str. 13, wzór (8)):

$$P(C_t \leq c, X_t \leq x; \mu, \sigma^2, t) = \Phi\left(\frac{x - \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu c}{\sigma^2}} \Phi\left(\frac{x - 2c - \mu t}{\sigma\sqrt{t}}\right), \quad (1)$$

gdzie Φ oznacza dystrybuantę standaryzowanego rozkładu normalnego.

Na podstawie równania (1) wyprowadzono formułę na gęstość rozkładu łącznego wektora zmiennych losowych (C_t, X_t) :

$$\begin{aligned}
f_{C_t, X_t}(c, x; \mu, \sigma^2, t) &= \frac{\partial^2 P(C_t \leq c, X_t \leq x; \mu, \sigma^2, t)}{\partial c \partial x} = \\
&= \frac{\sqrt{2}(2c - x)e^{-\frac{(x-\mu t)^2 - 4c(x-c)}{2\sigma^2 t}}}{\sqrt{\pi}\sigma^3 t^{\frac{3}{2}}} = \frac{\sqrt{2}(2c - x)e^{-\frac{(x-2c-\mu t)^2 - 4\mu ct}{2\sigma^2 t}}}{\sqrt{\pi}\sigma^3 t^{\frac{3}{2}}}. \quad (2)
\end{aligned}$$

Wyznaczono również funkcję charakterystyczną rozkładu łącznego wektora zmiennych losowych (C_t, X_t) :

$$\begin{aligned}
\varphi_{C_t, X_t}(q, s; \mu, \sigma^2, t) &= \int_0^\infty \int_{-\infty}^c e^{iqc + isx} f_{C_t, X_t}(c, x; \mu, \sigma^2, t) dx dc = \\
&= \frac{((q + s)\sigma^2 - i\mu)e^{\frac{1}{2}(q+s)(-(q+s)\sigma^2 + 2i\mu)t} \operatorname{erfc}\left(\frac{(-\mu - i(q + s)\sigma^2)\sqrt{t}}{\sqrt{2}\sigma}\right)}{(q + 2s)\sigma^2 - 2i\mu} + \\
&\quad + \frac{(s\sigma^2 - i\mu)e^{\frac{1}{2}s(-s\sigma^2 + 2i\mu)t} \operatorname{erfc}\left(\frac{(\mu + is\sigma^2)\sqrt{t}}{\sqrt{2}\sigma}\right)}{(q + 2s)\sigma^2 - 2i\mu}, \quad (3)
\end{aligned}$$

gdzie $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$.

Z kolei rozważmy zdarzenie $\{A_t > a, X_t > x\}$ gdy $a \leq 0$ i $a \leq x$. Także w tym wypadku można wyznaczyć jego prawdopodobieństwo stosując analogię do poprzednio rozpatrywanego zdarzenia:

$$P(A_t > a, X_t > x; \mu, \sigma^2, t) = \Phi\left(\frac{-x + \mu t}{\sigma\sqrt{t}}\right) - e^{\frac{2\mu a}{\sigma^2}} \Phi\left(\frac{-x + 2a + \mu t}{\sigma\sqrt{t}}\right). \quad (4)$$

Gęstość rozkładu łącznego wektora zmiennych losowych (A_t, X_t) dana jest formułą:

$$\begin{aligned}
f_{A_t, X_t}(a, x; \mu, \sigma^2, t) &= \frac{\partial^2 P(A_t > a, X_t > x; \mu, \sigma^2, t)}{\partial a \partial x} = \\
&= \frac{\sqrt{2}(x - 2a)e^{-\frac{(x-\mu t)^2 - 4a(x-a)}{2\sigma^2 t}}}{\sqrt{\pi}\sigma^3 t^{\frac{3}{2}}} = \frac{\sqrt{2}(x - 2a)e^{-\frac{(x-2a-\mu t)^2 - 4\mu at}{2\sigma^2 t}}}{\sqrt{\pi}\sigma^3 t^{\frac{3}{2}}}. \quad (5)
\end{aligned}$$

Funkcja charakterystyczna rozkładu łącznego wektora zmiennych losowych (A_t, X_t) wygląda następująco:

$$\begin{aligned}
\varphi_{A_t, X_t}(p, s; \mu, \sigma^2, t) &= \int_{-\infty}^0 \int_a^{\infty} e^{ipa+isx} f_{A_t, X_t}(a, x; \mu, \sigma^2, t) dx da = \\
&= \frac{((p+s)\sigma^2 - i\mu)e^{\frac{1}{2}(p+s)(-(p+s)\sigma^2+2i\mu)t} \operatorname{erfc}\left(\frac{(\mu + i(p+s)\sigma^2)\sqrt{t}}{\sqrt{2}\sigma}\right)}{(p+2s)\sigma^2 - 2i\mu} + \\
&\quad + \frac{(s\sigma^2 - i\mu)e^{\frac{1}{2}s(-s\sigma^2+2i\mu)t} \operatorname{erfc}\left(\frac{(-\mu - is\sigma^2)\sqrt{t}}{\sqrt{2}\sigma}\right)}{(p+2s)\sigma^2 - 2i\mu}. \tag{6}
\end{aligned}$$

Problem wyznaczenia rozkładu wektora (A_t, C_t, X_t) rozważany był np. w opracowaniach Coxa, Millera (1965), Li (1998). Gęstość rozkładu X_t z górną i dolną barierą absorbującą równymi odpowiednio c i a opisuje następujący wzór (patrz Cox, Miller, 1965, str. 222, formuła (78)):

$$\begin{aligned}
f_{X_t}(x; C_t \leq c, A_t > a; \mu, \sigma^2, t) &= \\
&= \frac{1}{\sqrt{2\pi t}\sigma} \sum_{k=-\infty}^{\infty} e^{\frac{2k(c-a)\mu}{\sigma^2}} \left(e^{-\frac{(x-2k(c-a)-\mu t)^2}{2\sigma^2 t}} - e^{-\frac{2c\mu t - (x-2c-2k(c-a)-\mu t)^2}{2\sigma^2 t}} \right), \tag{7}
\end{aligned}$$

gdzie $a \leq 0 \leq c$, $a \leq x \leq c$.

Korzystając ze wzoru (7) wyznaczono gęstość rozkładu łącznego wektora zmiennych losowych (A_t, C_t, X_t) :

$$\begin{aligned}
f_{A_t, C_t, X_t}(a, c, x; \mu, \sigma^2, t) &= -\frac{\partial^2 f_{X_t}(x; C_t \leq c, A_t > a; \mu, \sigma^2, t)}{\partial a \partial c} = \\
&= \sum_{k=-\infty}^{\infty} (g(a, c, x; k, k, \mu, \sigma, t) - g(a, c, x; k, k+1, \mu, \sigma, t)), \tag{8}
\end{aligned}$$

gdzie funkcja g opisana jest zależnością:

$$\begin{aligned}
g(a, c, x; m, n, \mu, \sigma, t) &= \\
&= \frac{4mn}{\sqrt{2\pi}\sigma^5 t^{5/2}} \left[(x - 2(nc - ma))^2 - \sigma^2 t \right] e^{\frac{2(nc-ma)\mu t - (x-2(nc-ma)-\mu t)^2}{2\sigma^2 t}}. \tag{9}
\end{aligned}$$

Funkcję charakterystyczną rozkładu łącznego wektora zmiennych losowych (A_t, C_t, X_t) opisuje formuła:

$$\begin{aligned} \varphi_{A_t, C_t, X_t}(p, q, s; \mu, \sigma^2, t) &= \\ &= \int_{-\infty}^0 \int_0^{\infty} \int_a^c e^{ipa+iqc+isx} f_{A_t, C_t, X_t}(a, c, x; \mu, \sigma^2, t) dx dc da. \end{aligned} \quad (10)$$

Z uwagi na bardzo złożoną postać tej funkcji jej szczegółowe przedstawienie zostanie w tym przypadku pominięte².

Znajomość postaci funkcji charakterystycznych umożliwia obliczenie wartości oczekiwanych wybranych zmiennych losowych. Przykładowo, dla całkowitych wartości u, v, w uzyskujemy:

$$E[A_t^u C_t^v X_t^w] = \frac{1}{i^n} \frac{\partial^n \varphi_{A_t, C_t, X_t}(p, q, s; \mu, \sigma^2, t)}{\partial p^u \partial q^v \partial s^w}, \quad (11)$$

gdzie $n = u + v + w$.

W sytuacji, gdy $u = 0$ lub $v = 0$, wygodniej jest skorzystać odpowiednio z formuł:

$$E[C_t^v X_t^w] = \frac{1}{i^n} \frac{\partial^n \varphi_{C_t, X_t}(q, s; \mu, \sigma^2, t)}{\partial q^v \partial s^w} \quad (12)$$

oraz

$$E[A_t^u X_t^w] = \frac{1}{i^n} \frac{\partial^n \varphi_{A_t, X_t}(p, s; \mu, \sigma^2, t)}{\partial p^u \partial s^w}. \quad (13)$$

2.3. MOMENTY ZWYKŁE ZMIENNYCH LOSOWYCH STÓP ZWROTU MINIMUM, MAKSIMUM I WARTOŚCI KOŃCOWEJ RUCHU BROWNA

Przyjęto następujące oznaczenia i definicje:

$$v := \frac{\mu\sqrt{t}}{\sigma}, \quad (14)$$

$$Z(n, y, v) := \sum_{k=1}^{\infty} \frac{1}{(k+y)^n} e^{-\frac{2k(k+1)v^2}{(2k+1)^2}}, \quad (15)$$

dla $n \in \mathbb{Z}, n > 0, y > 0, v \in \mathbb{R}$,

$$\zeta(n) := Z(n, 0, 0) = \sum_{k=1}^{\infty} \frac{1}{k^n}, \quad (16)$$

² Formuła może zostać udostępniona na życzenie Czytelnika.

dla $n \in \mathbb{Z}$, $n > 1$ $\zeta(n)$ jest tzw. funkcją Zeta-Riemanna,

$$Z_1(v) := Z(1, 0, v) - 2Z\left(1, \frac{1}{2}, v\right) + Z(1, 1, v), \quad (17)$$

$$\begin{aligned} Z_2(v) &:= 1 - Z(2, 0, v) + Z(2, 1, v) = \\ &= 2(3 - 4 \ln 2)v^2 + 2v^4 + \\ &+ \sum_{m=2}^{\infty} v^{2m} \sum_{j=1}^{m-1} \frac{(-1)^{m+j} (2^{2j+1} - 1) \zeta(2j+1)}{2^{m+2j-3} m(m-1)(m-j-1)!(j-1)!} = \\ &= 2(3 - 4 \ln 2)v^2 + \left(2 - \frac{7\zeta(3)}{4}\right)v^4 + \left(\frac{7\zeta(3)}{24} - \frac{31\zeta(5)}{96}\right)v^6 + \\ &\quad + \left(-\frac{7\zeta(3)}{192} + \frac{31\zeta(5)}{384} - \frac{127\zeta(7)}{3072}\right)v^8 + \\ &\quad + \left(\frac{7\zeta(3)}{1920} - \frac{31\zeta(5)}{2560} + \frac{127\zeta(7)}{10240} - \frac{511\zeta(9)}{122880}\right)v^{10} + \dots, \end{aligned} \quad (18)$$

$$Z_3(v) := Z(3, 0, v) + Z(3, 1, v), \quad (19)$$

$$Z_4(v) := 1 - Z(4, 0, v) + Z(4, 1, v), \quad (20)$$

$$H_1(v) := \frac{1}{2} + \frac{1}{2v^2}(-3 + Z_2(v) + Z_3(v)) + \frac{1}{4v^4}(8Z_2(v) - Z_4(v)), \quad (21)$$

$$\begin{aligned} H_2(v) &:= -Z\left(3, \frac{1}{2}, v\right) + 4Z_1(v) + \frac{1}{2}Z_2(v) - \\ &\quad - \frac{3}{2v^2}(4Z_1(v) + Z_2(v)) + \frac{3}{v^4}Z_2(v), \end{aligned} \quad (22)$$

$$\begin{aligned} H_3(v) &:= -\frac{3}{2} - \frac{v^2}{2} + 6Z_1(v) + \frac{3}{4}Z_2(v) - \frac{3}{2}Z\left(3, \frac{1}{2}, v\right) + \\ &\quad + \frac{1}{v^2}\left(\frac{9}{4} - 9Z_1(v) - 3Z_2(v) - \frac{3}{4}Z_3(v)\right) + \\ &\quad + \frac{1}{v^4}\left(-\frac{3}{2} + 3\zeta(3)e^{-\frac{v^2}{2}} + 3Z_2(v) - \frac{3}{2}Z_3(v) + \frac{3}{4}Z_4(v)\right). \end{aligned} \quad (23)$$

Korzystając z formuł (11), (12), (13) można obliczyć wartości oczekiwane dowolnych potęg zmiennych A_t , C_t , X_t . Poniżej przedstawiono wyprowadzone wzory dla wybranych momentów do czwartego rzędu włącznie.

Momenty pierwszego rzędu dla $v \neq 0$ ($\mu \neq 0$) dane są w postaci:

$$E[X_t] = \mu t = v\sigma\sqrt{t}, \quad (24)$$

$$E[C_t] = \frac{\sigma\sqrt{t}}{2} \left(\left(\frac{1}{v} + v \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) + \sqrt{\frac{2}{\pi}} e^{-\frac{v^2}{2}} + v \right), \quad (25)$$

$$E[A_t] = \frac{\sigma\sqrt{t}}{2} \left(- \left(\frac{1}{v} + v \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) - \sqrt{\frac{2}{\pi}} e^{-\frac{v^2}{2}} + v \right). \quad (26)$$

Momenty drugiego rzędu dla $v \neq 0$ ($\mu \neq 0$) opisują formuły:

$$E[X_t^2] = \sigma^2 t + \mu^2 t^2 = \sigma^2 t(1 + v^2), \quad (27)$$

$$E[C_t^2] = \sigma^2 t \left[\left(-\frac{1}{2v^2} + 2v^2 + 1 \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) + \frac{(1 + v^2)e^{-\frac{v^2}{2}}}{\sqrt{2\pi}v} + 1 + \frac{v^2}{2} \right], \quad (28)$$

$$E[A_t^2] = \sigma^2 t \left[\left(\frac{1}{2v^2} - 2v^2 - 1 \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) - \frac{(1 + v^2)e^{-\frac{v^2}{2}}}{\sqrt{2\pi}v} + 1 + \frac{v^2}{2} \right], \quad (29)$$

$$E[A_t C_t] = \sigma^2 t \left[-\frac{1}{2} + \frac{1}{4v^2} Z_2(v) \right], \quad (30)$$

$$E[C_t X_t] = \sigma^2 t \left[\left(-\frac{1}{2v^2} + 2v^2 + 1 \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) + \frac{(1 + v^2)e^{-\frac{v^2}{2}}}{\sqrt{2\pi}v} + \frac{1}{2} + \frac{v^2}{2} \right], \quad (31)$$

$$E[A_t X_t] = \sigma^2 t \left[\left(\frac{1}{2v^2} - 2v^2 - 1 \right) \operatorname{erf} \left(\frac{v}{\sqrt{2}} \right) - \frac{(1 + v^2)e^{-\frac{v^2}{2}}}{\sqrt{2\pi}v} + \frac{1}{2} + \frac{v^2}{2} \right]. \quad (32)$$

Moment trzeciego rzędu dla $v \neq 0$ ($\mu \neq 0$) opisuje zależność:

$$E[X_t^3] = 3\sigma^2 \mu t^2 + \mu^3 t^3 = \sigma^3 t^{\frac{3}{2}} (3v + v^3). \quad (33)$$

Momenty czwartego rzędu dla $v \neq 0$ ($\mu \neq 0$) przedstawione są równaniami:

$$E[X_t^4] = 3\sigma^4 t^2 + 6\sigma^2 \mu^2 t^3 + \mu^4 t^4, \quad (34)$$

$$E[A_t^2 X_t^2 + C_t^2 X_t^2] = 4\sigma^4 t^2 + 7\sigma^2 \mu^2 t^3 + \mu^4 t^4, \quad (35)$$

$$E[A_t^4 + C_t^4] = 6\sigma^4 t^2 + 8\sigma^2 \mu^2 t^3 + \mu^4 t^4, \quad (36)$$

$$E[A_t^3 X_t + C_t^3 X_t] = \frac{9}{2}\sigma^4 t^2 + \frac{15}{2}\sigma^2 \mu^2 t^3 + \mu^4 t^4, \quad (37)$$

$$E[A_t X_t^3 + C_t X_t^3] = 3\sigma^4 t^2 + 6\sigma^2 \mu^2 t^3 + \mu^4 t^4, \quad (38)$$

$$E[A_t^2 C_t^2] = \sigma^4 t^2 H_1(v), \quad (39)$$

$$E[A_t^2 C_t X_t + A_t C_t^2 X_t] = \sigma^4 t^2 \left(H_2(v) - \frac{v^2}{2} - \frac{1}{2} \right), \quad (40)$$

$$E[A_t C_t X_t^2] = \sigma^4 t^2 \left(\frac{1}{2} H_2(v) - \frac{v^2}{2} - \frac{1}{2} \right), \quad (41)$$

$$E[A_t^3 C_t + A_t C_t^3] = \sigma^4 t^2 H_3(v). \quad (42)$$

Dla $v = 0$ ($\mu = 0$) przedstawione w równaniach (24-42) momenty zmiennych losowych można wyznaczyć analogicznie korzystając z tych samych formuł (11), (12), (13). Jednak w tym wypadku konieczne jest wyznaczenie granic funkcji po prawej stronie dla $v \rightarrow 0$ ($\mu \rightarrow 0$). W większości wypadków, prostszym sposobem jest jednak wykorzystanie wzoru Vasicka (patrz Garman, Klass, 1980, str. 78). Gotowe wyniki można znaleźć w pracy Garmana, Klasa (1980, str. 74).

3. ESTYMATORY ZMIENNOŚCI

3.1. PRZEGLĄD ESTYMATORÓW WARIANCJI KONSTRUOWANYCH NA PODSTAWIE ZNAJOMOŚCI CEN MINIMALNYCH, MAKSYMALNYCH I ZAMKNIĘCIA

Tożsamości (24-42) zostały wyprowadzone przy wykorzystaniu znajomości rozkładu łącznego zmiennych losowych A_t , C_t , X_t . Na ich podstawie można uzyskać wiele informacji na temat znanych estymatorów wariancji ruchu Browna. Autorom publikacji nie jest znane żadne opracowanie, w którym przedstawiona byłaby kom-

pleksowa analiza estymatorów wariancji przy wykorzystaniu tego samego aparatu matematycznego. W punkcie 3.1 wyznaczone zostaną m.in. obciążenia wybranych, popularnych estymatorów wariancji dla niezerowego dryfu, a także sformułowane zostaną dokładne wartości wariancji estymatorów. Te charakterystyki nie były dotąd nigdzie publikowane.

Najbardziej popularny, nieobciążony estymator wariancji procesu konstruowany jest na podstawie jedynie cen zamknięcia, ale wymaga również znajomości wartości dryfu. Można go przedstawić w postaci formuły:

$$\widehat{s}_0^2(t) = (X_t - \mu t)^2. \quad (43)$$

Jego wariancja jest równa:

$$\text{Var}[\widehat{s}_0^2(t)] = E[(X_t - \mu t)^2 - \sigma^2 t]^2 = E[(X_t - \mu t)^4 - \sigma^4 t^2] = 2\sigma^4 t^2. \quad (44)$$

Wysoka wartość wariancji tego estymatora wynika oczywiście z częściowego tylko wykorzystania zbioru danych A_t , C_t , X_t .

Podobną cechę ma **estymator Parkinsona** (1980) dany wzorem:

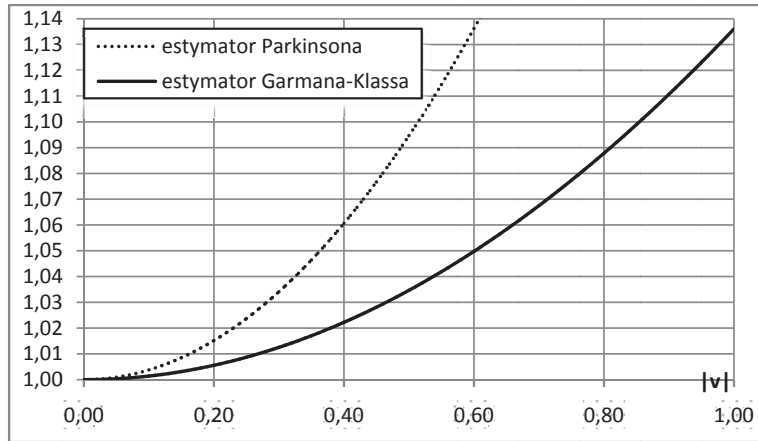
$$\widehat{s}_1^2(t) = \frac{(C_t - A_t)^2}{4 \ln 2}, \quad (45)$$

przy konstrukcji którego nie wykorzystuje się znajomości X_t . Dla $\mu = 0$ jest on nieobciążonym estymatorem wariancji $\sigma^2 t$ arytmetycznego ruchu Browna. W praktyce, estymator ten wykorzystywany jest często w sytuacji, gdy wartość dryfu μ jest nieznaną, bądź też, gdy $\mu \neq 0$. Korzystając z równań (28-30) wyprowadzono formułę na wartość oczekiwaną estymatora Parkinsona dla niezerowego dryfu:

$$\begin{aligned} E[\widehat{s}_1^2(t)] &= \frac{1}{4 \ln 2} \left(3 - \frac{1}{2v^2} Z_2(v) + v^2 \right) \sigma^2 t = \\ &= \left(1 + \frac{1}{4 \ln 2} \sum_{m=1}^{\infty} v^{2m} \sum_{j=1}^m \frac{(-1)^{m+j} (2^{2j+1} - 1) \zeta(2j+1)}{2^{m+2j-1} m(m+1)(m-j)! (j-1)!} \right) \sigma^2 t = \\ &= \left[1 + \frac{7\zeta(3)}{32 \ln 2} v^2 + \left(-\frac{7\zeta(3)}{192 \ln 2} + \frac{31\zeta(5)}{768 \ln 2} \right) v^4 + \right. \\ &\quad \left. + \left(\frac{7\zeta(3)}{1536 \ln 2} - \frac{31\zeta(5)}{3072 \ln 2} + \frac{127\zeta(7)}{24576 \ln 2} \right) v^6 + \right. \\ &\quad \left. + \left(-\frac{7\zeta(3)}{15360 \ln 2} + \frac{31\zeta(5)}{20480 \ln 2} - \frac{127\zeta(7)}{81920 \ln 2} + \frac{511\zeta(9)}{983040 \ln 2} \right) v^8 + \dots \right] \sigma^2 t, \quad (46) \end{aligned}$$

gdzie v oraz $\zeta(\cdot)$ zostały zdefiniowane w równaniach (14) i (16).

Rysunek 2 daje pogląd na temat wielkości obciążenia tego estymatora w zależności od wartości parametru v .



Iloraz wartości oczekiwanej estymatora i wariancji ruchu Browna w zależności od parametru v .

Rysunek 2. Obciążenie estymatorów Parkinsona i Garmana-Klassa dla niezerowego dryfu

Źródło: opracowanie własne.

Korzystając z tożsamości podanych we wzorach (28–30), (36), (39) i (42) można wyznaczyć wariancję estymatora Parkinsona:

$$\begin{aligned} \text{Var}[\widehat{s}_1^2(t)] &= E \left[\left(\frac{(C_t - A_t)^2}{4 \ln 2} - E \left[\frac{(C_t - A_t)^2}{4 \ln 2} \right] \right)^2 \right] = \\ &= \left[\frac{-3 + 2v^2 + Z_2(v) + 6H_1(v) - 4H_3(v)}{16 \ln^2 2} + \frac{3Z_2(v)}{16v^2 \ln^2 2} - \frac{Z_2^2(v)}{64v^4 \ln^2 2} \right] \sigma^4 t^2. \quad (47) \end{aligned}$$

Na podstawie tożsamości zawartych w Garman, Klass (1980, str. 74), dla $\mu = 0$ otrzymujemy:

$$\text{Var}[\widehat{s}_1^2(t)] = \left(-1 + \frac{9\zeta(3)}{16 \ln^2 2} \right) \sigma^4 t^2 \approx 0,407332 \sigma^4 t^2. \quad (48)$$

Wynik ten zgodny z rezultatem podanym przez Parkinsona (1980, str. 63), lecz odbiega od szacunku podanego przez Garmana, Klassa (1980, str. 71), według których estymator $\widehat{s}_1^2(t)$ jest 5.2 razy efektywniejszy od $\widehat{s}_0^2(t)$. W żadnej z omawianych prac nie podano dokładnych formuł, na podstawie których zostały wyznaczone te wartości.

Garman, Klass (1980) rozważali również tylko przypadek, gdy $\mu = 0$. Skonstruowali oni estymator wariancji arytmetycznego ruchu Browna dany formułą (określany dalej jako estymator G-K):

$$\widehat{s}_2^2(t) = \frac{(C_t - A_t)^2}{2} - (2 \ln 2 - 1)X_t^2. \quad (49)$$

Na podstawie równań (28-30) skonstruowano formułę na wartość oczekiwaną estymatora G-K dla $v \neq 0$ ($\mu \neq 0$):

$$\begin{aligned} E[\widehat{s}_2^2(t)] &= \left(\left(\frac{5}{2} - 2 \ln 2 \right) + \left(\frac{3}{2} - 2 \ln 2 \right) v^2 - \frac{1}{4v^2} Z_2(v) \right) \sigma^2 t = \\ &= [1 + (1 - 2 \ln 2)v^2 + \\ &\quad + \sum_{n=1}^{\infty} v^{2n} \sum_{j=1}^n (-1)^{n+j} \frac{(2^{2j+1} - 1)\zeta(2j+1)}{2^{n+2j} n(n+1)(n-j)!(j-1)!}] \sigma^2 t = \\ &= \left[1 + \left(1 - 2 \ln 2 + \frac{7}{16} \zeta(3) \right) v^2 + \left(-\frac{7}{96} \zeta(3) + \frac{31}{384} \zeta(5) \right) v^4 + \right. \\ &\quad + \left(\frac{7\zeta(3)}{768} - \frac{31\zeta(5)}{1536} + \frac{127\zeta(7)}{12288} \right) v^6 + \\ &\quad \left. + \left(-\frac{7\zeta(3)}{7680} + \frac{31\zeta(5)}{10240} - \frac{127\zeta(7)}{40960} + \frac{511\zeta(9)}{491520} \right) v^8 + \dots \right] \sigma^2 t. \quad (50) \end{aligned}$$

Na podstawie rysunku 2 można ocenić wielkość obciążenia tego estymatora w zależności od wartości parametru v .

Wariancja estymatora G-K jest równa:

$$\begin{aligned} Var[\widehat{s}_2^2(t)] &= \\ &= E \left[\left(\frac{(C_t - A_t)^2}{2} - (2 \ln 2 - 1)X_t^2 - E \left[\frac{(C_t - A_t)^2}{2} - (2 \ln 2 - 1)X_t^2 \right] \right)^2 \right] = \\ &= \left[\frac{13}{4} - 12 \ln 2 + 8 \ln^2 2 + \left(\frac{17}{2} - 24 \ln 2 + 16 \ln^2 2 \right) v^2 + \frac{3H_1(v)}{2} - \right. \\ &\quad \left. - (2 \ln 2 - 1)H_2(v) - H_3(v) + \left(\frac{3}{4} - \ln 2 \right) Z_2(v) + \right. \\ &\quad \left. + \frac{(5 - 4 \ln 2)Z_2(v)}{4v^2} - \frac{Z_2^2(v)}{16v^4} \right] \sigma^4 t^2. \quad (51) \end{aligned}$$

Przy wyprowadzeniu równania (51) korzystano z formuł (28-30), (34-36), (39), (41) i (42). Na podstawie tożsamości zawartych w Garman, Klass (1980, str. 74), dla $\mu = 0$ otrzymujemy:

$$\text{Var}[\widehat{s}_2^2(t)] = \left(2 - 8 \ln 2 + 4 \ln^2 2 + \left(4 - \frac{7}{2} \ln 2 \right) \zeta(3) \right) \sigma^4 t^2 \approx 0,2687 \sigma^4 t^2. \quad (52)$$

Uzyskany w równaniu (52) rezultat jest zgodny z szacunkami podanymi przez Garmana, Klassa (1980, str. 74), według których estymator $\widehat{s}_2^2(t)$ jest 7,4 razy bardziej efektywny od $\widehat{s}_0^2(t)$ oraz z wynikami Rogersa, Satchella (1991, str. 505), według których wariancja estymatora $\widehat{s}_2^2(t)$ jest równa $0,27 \sigma^4 t^2$.

Rogers, Satchell (1991) prowadzili badania dla niezerowego dryfu. Zaproponowali oni następujący estymator wariancji procesu X_t (określany dalej jako R-S):

$$\widehat{s}_3^2(t) = C_t(C_t - X_t) + A_t(A_t - X_t), \quad (53)$$

który nie tylko jest nieobciążony dla wszystkich wartości $\mu \neq 0$, ale znajomość tego parametru nie jest konieczna do estymacji wariancji procesu. Wariancja tego estymatora jest równa:

$$\begin{aligned} \text{Var}[\widehat{s}_3^2(t)] &= E[(C_t(C_t - X_t) + A_t(A_t - X_t) - \sigma^2 t)^2] = \\ &= (2H_1(v) - H_2(v)) \sigma^4 t^2. \end{aligned} \quad (54)$$

Równanie (54) zostało wyprowadzone z wykorzystaniem formuł (35-37), (39-42). Dla $\mu = 0$ uzyskujemy:

$$\text{Var}[\widehat{s}_3^2(t)] = \left(1 - 4 \ln 2 + \frac{7}{4} \zeta(3) \right) \sigma^4 t^2 \approx 0,331011 \sigma^4 t^2. \quad (55)$$

Otrzymana wartość $\text{Var}[\widehat{s}_3^2(t)]$ jest zgodna z wynikami Rogersa, Satchella (1991, str. 505). Także w tym wypadku nie podano formuł, na podstawie których wartość ta została wyznaczona.

W niniejszej pracy rozważane są estymatory, które konstruowane są na podstawie jedynie pojedynczego wektora zmiennych losowych (A_t, C_t, X_t) . W badaniach finansowych wykorzystuje się szereg innych estymatorów zmienności, które konstruowane są na podstawie szerszego zbioru informacji. Jednym z nich jest estymator o bardzo wysokiej efektywności, zaproponowany przez Kunitomo (1992). Wykorzystuje on tzw. skorygowaną maksymalną i minimalną wartość, lecz do ich wyznaczenia konieczna jest znajomość wszystkich śróddziennych stóp zwrotu. Z tego względu estymator ten nie będzie bliżej omawiany.

Jak zostało to już zaznaczone, przedstawione dotychczas estymatory stosowane są w praktyce do wyznaczania wariancji dla jednego dnia notowań. Przy określeniu

ich własności przyjęto założenie, że dzienne logarytmiczne stopy zwrotu są realizacją ruchu Browna o ustalonych na dany dzień wartościach dryfu μ oraz wariancji σ^2 . Czasami jednak estymatory te wykorzystywane są do obliczenia wariancji dla okresu, którego długość przekracza jeden dzień. Wówczas zakłada się, że wariancja stóp zwrotu jest stała w danym okresie, ponadto przyjmuje się, że jej szacunek w skali jednego dnia jest równy średniej arytmetycznej dziennych wariancji wyznaczonych dla kolejnych dni należących do przyjętego okresu. W takim wypadku mamy zatem do czynienia z bardzo istotnym wzmocnieniem założeń przyjętego modelu. Należy jednak zauważyć, że dla większości szeregów finansowych tak wzmocnione założenie jest nieprawdziwe w świetle wyników badań empirycznych (patrz np. prace przeglądowe dotyczące zmienności warunkowej wariancji – Bollerslev, Chou, Kroner, 1992, Bollerslev, Engle, Nelson, 1994). Z tego powodu, taki sposób wykorzystania przedstawionych estymatorów nie był przedmiotem zainteresowania niniejszego opracowania.

Przy założeniu stałości wariancji dla okresu dłuższego niż jeden dzień skonstruowany został również estymator zaproponowany przez Yanga, Zhanga (2000). Wykorzystuje on znajomość wartości ciągu wektorów o długości większej od 1, których współrzędnymi są ceny dziennego otwarcia, maksimum, minimum i zamknięcia. Zatem, dodatkowo uwzględniane są tzw. zwroty nocne, czyli stopy zwrotu pomiędzy ceną zamknięcia poprzedniego dnia i ceną otwarcia dnia następnego (patrz rysunek 1). Ze względu na wysoką efektywność i powszechną dostępność danych, na podstawie których jest konstruowany, jest on jednym z najpowszechniej stosowanych estymatorów zmienności wykorzystujących cenę otwarcia, minimalną, maksymalną i zamknięcia. Nie jest natomiast możliwe wykorzystanie tego estymatora do szacowania wariancji tylko dla jednego dnia. Z tego względu, a także z powodu przyjętych bardzo silnych założeń, estymator ten nie będzie również bliżej omawiany.

3.2. PROPOZYCJA NOWEGO ESTYMATORA ZMIENNOŚCI

Według wiedzy autorów, w przypadku gdy nie jest znana wartość dryfu μ , estymator R-S jest najbardziej efektywnym ze znanych estymatorów wariancji ruchu Browna dla niezerowej wartości dryfu konstruowanym na podstawie realizacji wektora zmiennych (A_t, C_t, X_t) (w przypadku $\mu = 0$ najbardziej efektywnym znanym estymatorem jest estymator G-K). Można postawić pytanie, czy w sytuacji, gdy znana jest wartość parametru μ , istnieje możliwość wykorzystania tej dodatkowej informacji tak, aby poprawić efektywność estymatora R-S.

Poniżej zostanie zaproponowany estymator wariancji ruchu Browna przy znanej wartości parametru μ , konstruowany na podstawie realizacji (A_t, C_t, X_t) . Rozważmy następujące estymatory wariancji:

$$\widehat{s}_{11}^2(t) = A_t X_t + C_t X_t - \mu^2 t^2, \quad (56)$$

$$\widehat{s}_{12}^2(t) = \frac{1}{2}(A_t^2 + C_t^2 - \mu^2 t^2), \quad (57)$$

$$\widehat{s}_{13}^2(t) = X_t^2 - \mu^2 t^2. \quad (58)$$

Wykorzystując formuły (31), (32), (27), (28), (29), można łatwo pokazać, że $\widehat{s}_{11}^2(t)$, $\widehat{s}_{12}^2(t)$, $\widehat{s}_{13}^2(t)$ są nieobciążonymi estymatorami wariancji $\sigma^2 t$ arytmetycznego ruchu Browna. W związku z tym estymator:

$$\widehat{s}^2(a, b, t) = a\widehat{s}_{11}^2(t) + b\widehat{s}_{12}^2(t) + (1 - a - b)\widehat{s}_{13}^2(t) \quad (59)$$

jest również nieobciążonym estymatorem wariancji $\sigma^2 t$ dla wszystkich $a, b \in \mathbb{R}$. Łatwo zauważyć, że dla $a = -1$ i $b = 2$ otrzymujemy estymator R-S. Wariancja estymatora $\widehat{s}^2(a, b, t)$ dana jest formułą:

$$\begin{aligned} \text{Var}[\widehat{s}^2(a, b, t)] &= \left(H_2(v) \left(a + \frac{b}{2} \right)^2 + \right. \\ &+ \left. \frac{1}{4} H_{12}(v) \left(b - \frac{4(1+2v^2)}{H_{12}(v)} \right)^2 - \frac{4(1+2v^2)^2}{H_{12}(v)} + 2 + 4v^2 \right) \sigma^4 t^2, \end{aligned} \quad (60)$$

gdzie:

$$H_{12}(v) = 2H_1(v) - H_2(v) + 2 + 4v^2, \quad (61)$$

natomiast funkcje H_1 i H_2 zdefiniowano odpowiednio w równaniach (21) i (22). Funkcje $H_2(v)$ i $H_{12}(v)$ przyjmują wartości dodatnie dla $v \in \mathbb{R}$. Z uwagi na ograniczoną objętość artykułu dowód tego faktu zostanie pominięty³. W związku z tym funkcja $V[\widehat{s}^2(a, b)]$ osiąga minimum dla:

$$b = b_{\min}(v) := \frac{4(1+2v^2)}{H_{12}(v)}, \quad (62)$$

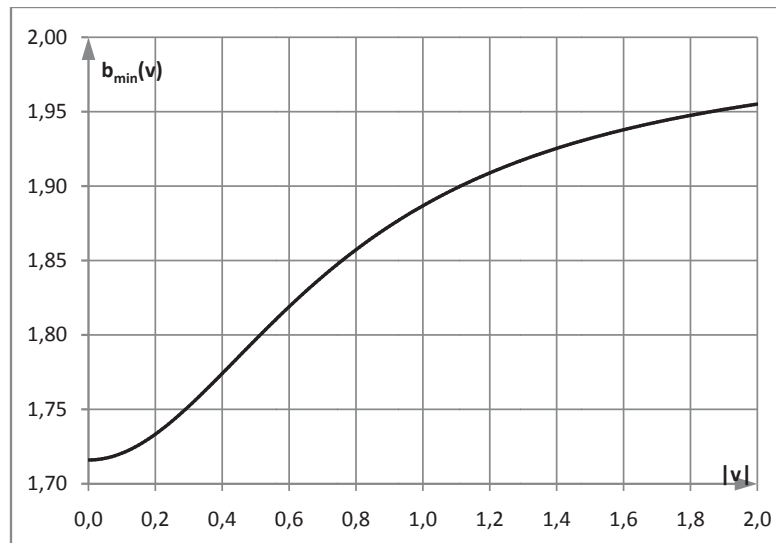
$$a = a_{\min}(v) := -\frac{b_{\min}(v)}{2} = -\frac{2(1+2v^2)}{H_{12}(v)}. \quad (63)$$

Jeżeli $v = 0$, wówczas wartość parametru b , dla której wariancja estymatora osiąga minimum wynosi:

³ Wyniki mogą zostać udostępnione na życzenie Czytelnika.

$$b_{min}(0) = \frac{16}{12 - 16 \ln 2 + 7\zeta(3)} \approx 1,715994. \quad (64)$$

Funkcja b_{min} jest parzysta. Rosnące wartości bezwzględne argumentu v powodują monotoniczny wzrost b_{min} ; dla $v \rightarrow \pm\infty$ funkcja ta asymptotycznie dąży do 2 (patrz rysunek 3).



Rysunek 3. Wartości parametru b w zależności od wartości v , dla których wariancja zaproponowanego estymatora osiąga minimum

Źródło: opracowanie własne.

Dla $v = 0$ wariancja estymatora jest równa:

$$V[\widehat{s^2}(a_{min}(0), b_{min}(0), t)] = \left(2 - \frac{16}{12 - 16 \ln 2 + 7\zeta(3)}\right) \sigma^4 t^2 \approx 0,284 \sigma^4 t^2. \quad (65)$$

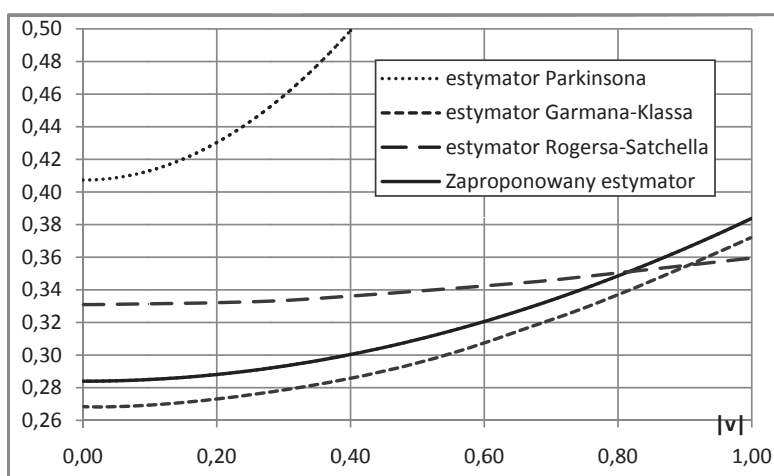
Jest ona zatem mniejsza od wariancji estymatora R-S $\widehat{s^2}_3$ i tylko nieznacznie większa od estymatora G-K $\widehat{s^2}_2$.

Sytuacja, w której postać estymatora o minimalnej wariancji zależy od wartości estymowanego parametru powoduje, że niemożliwe jest skonstruowanie najefektywniejszego estymatora. Innymi słowy, aby ustalić postać estymatora o minimalnej wariancji, trzeba znać z góry wartość estymowanego parametru. Fakt ten nie oznacza jednak, że nie ma możliwości poprawienia efektywności znanych estymatorów. W praktyce, w badaniach finansowych, w ogromnej większości przypadków występuje zależność: $|\mu t| \ll \sigma\sqrt{t}$, czyli wartość v jest bliska 0. Do praktycznych zastosowań

wań można przyjąć wartości $b = 1,72$ i $a = -0,86$. Proponowany estymator ma zatem następującą postać:

$$\widehat{s}^2(-0,86, 1,72, t) = 0,86(C_t(C_t - X_t) + A_t(A_t - X_t)) + 0,14(X_t^2 - \mu^2 t^2). \quad (66)$$

Rysunek 4 przedstawia wykres zależności wartości wariancji estymatorów $\widehat{s}^2_1(t)$, $\widehat{s}^2_2(t)$, $\widehat{s}^2_3(t)$ i $\widehat{s}^2(-0,86, 1,72, t)$ od parametru v .



Iloraz wariancji estymatora i kwadratu wariancji ruchu Browna w zależności od parametru v .

Rysunek 4. Efektywność estymatorów wariancji w zależności od wartości parametru v

Źródło: opracowanie własne.

Jak zostało to wcześniej zaznaczone, stosowanie estymatora $\widehat{s}^2(-0,86, 1,72, t)$ wymaga znajomości parametru dryfu μ . W analizach szeregów finansowych, jego dokładna wartość niemal nigdy nie jest znana i konieczna jest jego estymacja. Powstaje zatem problem określenia wpływu estymacji dryfu na własności estymatora $\widehat{s}^2(-0,86, 1,72, t)$.

Niech $\hat{\mu}$ będzie estymatorem dryfu, określonym na podstawie dużej próby o liczebności n . Ponieważ estymacja zmienności przeprowadzana jest na podstawie informacji z pojedynczego dnia tej próby, zatem estymator wariancji $\widehat{s}^2(-0,86, 1,72, t)$ jest zależny od estymatora dryfu $\hat{\mu}$. Zależność ta maleje jednak wraz ze wzrostem liczebności próby, na podstawie której konstruowany jest estymator dryfu. W dalszej części wywodu zostanie przyjęte, że zależność ta jest na tyle słaba, że można ją pominąć. W tej sytuacji zachodzi zależność $E[\hat{\mu}^2] = \mu^2$. Można łatwo pokazać, że

obciążenie estymatora $\widehat{s}^2(-0,86, 1,72, t)$ będzie wówczas równe $0,14(\mu^2 t^2 - \hat{\mu}^2 t^2)$. Z tego wynika, że maksymalna bezwzględna wartość obciążenia estymatora $\widehat{s}^2(-0,86, 1,72, t)$ spowodowana estymacją dryfu, wynosi $0,14 \max(\mu^2 t^2, \hat{\mu}^2 t^2)$. Oznacza to, że w sytuacji gdy np. $\hat{\mu}^2 = 0$, wielkość obciążenia tego estymatora wyniesie $0,14v^2\sigma^2 t$ i tylko nieznacznie przekroczy obciążenie estymatora G-K $\widehat{s}^2_3(t)$ (patrz równanie (50) i rysunek 2). Ponadto konieczność estymacji dryfu nie spowoduje wzrostu wariancji estymatora $\widehat{s}^2(-0,86, 1,72, t)$. Ocenę obciążenia i jego wpływu na wariancję estymatora $\widehat{s}^2(-0,86, 1,72, t)$, w przypadku gdy przyjęte założenie nie będzie spełnione, można przeprowadzić dla określonych estymatorów dryfu na podstawie metody Monte Carlo.

3.3. ESTYMACJA ZMIENNOŚCI STÓP ZWROTU INDEKSU WIG20

W celu ilustracji przydatności rozważanych estymatorów zastosowano je dla indeksu rynku akcji WIG20 do szacowania wariancji logarytmicznych stóp zwrotu pomnożonych przez 100. Indeks ten został wybrany celowo, z uwagi na fakt, że obejmuje największe i najbardziej płynne spółki notowane na GPW w Warszawie. Jest to istotne, ponieważ w badaniu korzystano z danych śróddziennych, których jakość zależy w dużym stopniu od płynności rynku (Doman, 2011). Zastosowano następujące estymatory: konstruowany na podstawie cen zamknięcia $\widehat{s}^2_0(t)$, G-K $\widehat{s}^2_2(t)$, R-S $\widehat{s}^2_3(t)$ oraz zaproponowany $\widehat{s}^2(-0,86, 1,72, t)$ (opisane odpowiednio równaniami (45), (49), (53), (66)). Analizę przeprowadzono dla 10-letniego okresu od 30 września 2002 r. do 28 września 2012 r., czyli okresu obejmującego zarówno okresy hossy, jak i bessy a także, co jest istotne, kryzysu finansowego. Skonstruowano 2513 wektorów dziennych stóp zwrotu: $(a_1, c_1, x_1), (a_2, c_2, x_2), \dots, (a_{2513}, c_{2513}, x_{2513})$. Następnie przeprowadzono estymację dryfu wg następującej formuły⁴: $\hat{\mu} = \frac{1}{2513} \sum_{i=1}^{2513} x_i \approx 0,00032613502$. Kolejnym krokiem było oszacowanie wariancji dla każdego dnia w próbie, czyli dla $1 \leq i \leq 2513$:

$$\widehat{s}^2_{0,i}(1) = (x_i - \hat{\mu})^2, \quad (67)$$

$$\widehat{s}^2_{2,i}(1) = \frac{(c_i - a_i)^2}{2} - (2 \ln 2 - 1)x_i^2. \quad (68)$$

⁴ Założono stałość dryfu w czasie. Przyjęto najprostszą postać estymatora dryfu. Gdyby zastosować efektywniejszy estymator dryfu, to należałoby się spodziewać, że uzyskane wyniki szacowania zmienności na podstawie zaproponowanego estymatora nie będą mniej dokładne od tych jakie przedstawiono w pracy.

$$\widehat{s}_{3,i}^2(1) = c_i(c_i - x_i) + a_i(a_i - x_i), \quad (69)$$

$$\widehat{s}_i^2(-0,86, 1,72, 1) = 0,86(c_i(c_i - x_i) + a_i(a_i - x_i)) + 0,14(x_i^2 - \hat{\mu}^2). \quad (70)$$

Dla porównania skonstruowano również szacunki wariancji na podstawie modelu GARCH⁵. Na podstawie kryterium informacyjnego Schwartza wybrano model GARCH(1,3) z warunkowym rozkładem t-Studenta (oznaczony jako GARCH-t(1,3)). Dodatkowo zastosowano również najczęściej stosowaną w badaniach empirycznych postać modelu GARCH(1,1) z warunkowym rozkładem normalnym. Do estymacji parametrów wspomnianych modeli wykorzystano odpowiednio metodę największej wiarygodności oraz metodę quasi-największej wiarygodności.

Do oceny dokładności szacunków zmienności jako realizacje wariancji przyjęto sumy kwadratów śróddziennych stóp zwrotu. Istotnym problemem przy zastosowaniu takich danych jest wybór odpowiedniej częstotliwości obserwacji (patrz np. Andersen, Bollerslev, Diebold, Labys, 2000 oraz Oomen, 2001). Z tego względu zmienność zrealizowaną szacowano w różnych wariantach poprzez wyznaczenie sum kwadratów 5-, 10-, 15-, 20-, 30- i 60-minutowych stóp zwrotu. Oceny dokładności szacunków zmienności konstruowanych na podstawie rozważanych metod dokonano za pomocą następujących miar⁶: odsetek przeszacowań, błąd średni ME, pierwiastek błędu średniokwadratowego RMSE, pierwiastek błędu średniokwadratowego skorygowany o heteroskedastyczność HRMSE, funkcja straty LINEX ($a = -0,01$ i $0,01$) oraz współczynnik determinacji R^2 dla równania regresji realizacji wariancji względem szacunków wariancji.

Tabela 1 przedstawia wyniki przeprowadzonej analizy w przypadku, gdy zmienność zrealizowaną obliczano na podstawie 10-minutowych stóp zwrotu. Bardzo zbliżone relacje pomiędzy rozważanymi metodami szacowania zmienności występowały również wtedy, gdy zmienność zrealizowana była wyznaczana dla 5-, 15-, 20-, 30- i 60-minutowych stóp zwrotu.

Otrzymane wyniki dla odsetka przeszacowań oraz błędu średniego wskazują, że szacunki zmienności konstruowane na podstawie estymatorów wykorzystujących dane o cenach minimalnych i maksymalnych są znacząco niedoszacowane. Wynika to jednak nie ze słabej jakości tych estymatorów, a z przyjętej miary zmienności zrealizowanej, w tym przypadku sumy kwadratów stóp zwrotu z danych śróddziennych. Podobnie jak dla danych dziennych kwadrat śróddziennych stopy zwrotu jest nieefektywnym estymatorem wariancji śróddziennych stopy zwrotu. Z tego względu w tabeli 2 przedstawiono wyniki oceny dokładności szacunków zmienności jednakże tym razem zmienność zrealizowaną szacowano na podstawie 10-minutowych

⁵ Nie zaobserwowano istotnej statystycznie autokorelacji, dlatego przyjęto następującą specyfikację równania dla średniej: $x_i = \gamma_0 + \varepsilon_i$.

⁶ Ich opis można znaleźć na przykład w pracy Fiszedera (2009).

stóp zwrotu i estymatora R-S. Przy takim podejściu stopień niedoszacowania szacunków zmienności, konstruowanych na podstawie tych estymatorów znacząco się zmniejsza.

Tabela 1.

Ocena dokładności szacunków zmienności dla indeksu WIG 20 – zmienność zrealizowana obliczana na podstawie kwadratów 10-minutowych stóp zwrotu

Estymator wariancji	Odsetek przeszacowań	ME	RMSE	HRMSE	LINEX (a = 0,01)	LINEX (a = -0,01)	R ²
s_0^2	34,06	0,0811	4,8953	1,2882	1,19E-03	1,23E-03	0,2472
s_2^2 (G-K)	22,24	0,7144	2,5562	0,4704	3,36E-04	3,21E-04	0,6532
s_3^2 (R-S)	28,21	0,6946	2,7346	0,5502	3,85E-04	3,66E-04	0,6230
s^2	26,46	0,6087	2,4212	0,4593	3,02E-04	2,86E-04	0,6848
GARCH-t(1,3)	66,77	0,0441	3,3739	1,0747	6,22E-04	5,26E-04	0,3392
GARCH(1,1)	67,29	0,0647	3,2964	1,0402	5,95E-04	5,01E-04	0,3821

Źródło: obliczenia własne.

Tabela 2.

Ocena dokładności szacunków zmienności dla indeksu WIG 20 – zmienność zrealizowana obliczana na podstawie estymatora R-S i 10-minutowych stóp zwrotu

Estymator wariancji	Odsetek przeszacowań	ME	RMSE	HRMSE	LINEX (a = 0,01)	LINEX (a = -0,01)	R ²
s_0^2	39,79	-0,5497	4,7246	2,0265	1,06E-03	1,19E-03	0,2480
s_2^2 (G-K)	40,79	0,0836	2,5591	0,6936	3,18E-04	3,41E-04	0,5255
s_3^2 (R-S)	43,97	0,0637	2,8469	0,7834	3,95E-04	4,20E-04	0,5004
s^2	45,13	-0,0221	2,4830	0,7065	3,00E-04	3,20E-04	0,5659
GARCH-t(1,3)	80,30	-0,5867	2,7366	1,5616	3,91E-04	3,61E-04	0,2965
GARCH(1,1)	81,66	-0,5661	2,6555	1,5214	3,69E-04	3,39E-04	0,3248

Źródło: obliczenia własne.

Wyniki pozostałych mierników przedstawione w tabeli 1 wskazują na wyraźną przewagę estymatorów wariancji wykorzystujących dane o cenach minimalnych i maksymalnych nad estymatorem konstruowanym na podstawie modelu GARCH. Rezultaty zaprezentowane w tabeli 2 sugerują, że przewaga badanych estymatorów nie jest już tak istotna, a estymator R-S wypada nawet nieco gorzej od estymatora wariancji konstruowanego na podstawie modelu GARCH. Wszystkie rozważane

miary (poza HRMSE, tabela 2) wskazują, że najbardziej dokładne szacunki zmienności konstruowane były na podstawie zaproponowanego w pracy estymatora wariancji. Oczywiście przewaga tego podejścia nad estymatorami G-K i R-S nie zawsze jest znacząca i w wielu przypadkach jest prawdopodobnie nieistotna statystycznie.

4. ZAKOŃCZENIE

W opracowaniu przedstawiony został rozkład łączny stóp zwrotu minimum, maksimum i wartości końcowej ruchu Browna. Dzięki wykorzystaniu tego rozkładu:

- a) Udowodniono nieobciążoność najbardziej popularnych estymatorów zmienności: Parkinsona, G-K dla zerowej wartości dryfu oraz R-S dla dowolnego dryfu; innymi słowy potwierdzono zasadnicze wyniki prac jakie uzyskali autorzy tych estymatorów;
- b) Wyznaczono wartości oczekiwane estymatorów Parkinsona i G-K, w sytuacji gdy wartość dryfu jest różna od zera;
- c) Skonstruowano formuły wariancji estymatorów Parkinsona, G-K i R-S w sytuacji gdy wartość dryfu jest różna od zera.

Autorom nie są znane publikacje, które zawierałyby wyniki, o których mowa jest w punktach b) i c).

Ponadto zaproponowano nowy estymator wariancji procesu dla znanej niezerowej wartości dryfu. Nie udało się skonstruować estymatora, dla którego wariancja osiągałaby wartość minimalną niezależnie od wartości dryfu. Podjęte badania umożliwiły natomiast sformułowanie estymatora, który dla obserwowanych na rynkach finansowych rzędów wielkości dryfu będzie miał mniejszą wariancję od estymatora R-S. Przyjęte w pracy założenie, że stopy zwrotu w ciągu dnia są realizacjami arytmetycznego ruchu Browna nie jest spełnione w przypadku większości szeregów finansowych, dlatego w przyszłości będą konieczne dalsze badania dotyczące innych założeń modelowych.

Wyniki przedstawionego w pracy badania dla indeksu WIG20 wskazują, że szacunki wariancji konstruowane na podstawie tego estymatora są znacząco dokładniejsze od tych, jakie można uzyskać na podstawie wariancji warunkowej modelu GARCH. Estymator ten może być w przyszłości wykorzystany do budowy modeli GARCH i SV, które są obecnie najczęściej stosowane do estymacji wariancji procesów obserwowanych na rynkach finansowych (patrz np. prace przeglądowe dotyczące polskiego rynku finansowego Pipień, Osiewalski, 2004; Doman, Doman, 2009; Fiszeder, 2009, Pajor, 2010), dzięki czemu będzie możliwe uzyskanie jeszcze dokładniejszych szacunków zmienności, a co ważniejsze konstruowanie trafniejszych prognoz zmienności.

LITERATURA

- [1] Andersen T. G., Bollerslev T., Diebold F. X., Labys P., (2000), Great Realizations, *Risk*, March, 105–108.
- [2] Bollerslev T., Chou R. Y., Kroner K. F., (1992), ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics*, 52, 5–59.
- [3] Bollerslev T., Engle R. F., Nelson D. B., (1994), ARCH Models, w: Engle R. F., McFadden D., (red.), *Handbook of Econometrics*, 4, Elsevier Science B. V., Amsterdam.
- [4] Cox D. R., Miller M. D., (1965), *The Theory of Stochastic Processes.*, Methuen and Co., London.
- [5] Doman M., (2011), *Mikrostruktura giełd papierów wartościowych*, Wydawnictwo Uniwersytetu Ekonomicznego w Poznaniu, Poznań.
- [6] Doman M., Doman R., (2009), *Modelowanie zmienności i ryzyka. Metody ekonometrii finansowej*, Wolters Kluwer Polska, Kraków.
- [7] Fiszedera P., (2009), *Modele klasy GARCH w empirycznych badaniach finansowych*, Wydawnictwo UMK, Toruń.
- [8] Garman M. B., Klass M. J., (1980), On the Estimation of Security Price Volatilities from Historical Data, *The Journal of Business*, 53 (1), 67–78.
- [9] Harrisson J. M., (1985), *Brownian Motion and Stochastic Flow Systems*, John Wiley & Sons, New York.
- [10] Jakubowski J., Palczewski A., Rutkowski M., Stettner Ł. (2006), *Matematyka finansowa, instrumenty pochodne*, WNT, Warszawa.
- [11] Kunitomo N., Ikeda M., (1992), Pricing Options with Curved Boundaries, *Mathematical Finance*, 2 (4), 275–298.
- [12] Kunitomo N. (1992), Improving the Parkinson Method of Estimating Security Price Volatilities, *Journal of Business*, 65, 295–302.
- [13] Li A., (1999), The Pricing of Double Barrier Options and Their Variations, *Advances in Futures and Options Research*, 10, 17–41.
- [14] Osiewalski J., Pipień M., (2004), Bayesian Comparison of Bivariate ARCH-Type Models for the Main Exchange Rates in Poland, *Journal of Econometrics*, 123, 371–391.
- [15] Oomen R., (2001), Using High Frequency Stock Market Index Data to calculate, Model, European University Institute, Discussion Paper No. 2001/6.
- [16] Pajor A., (2010), *Wielowymiarowe procesy wariancji stochastycznej w ekonometrii finansowej. Ujęcie bayesowskie*, Zeszyty Naukowe, Seria Specjalna: Monografie 195, Wydawnictwo Uniwersytetu Ekonomicznego w Krakowie.
- [17] Parkinson M., (1980), The Extreme Value Method for Estimating the Variance of the Rate of Return, *The Journal of Business*, 53 (1), 61–65.
- [18] Perczak G., (2013), Dodatkowe informacje o cenach minimalnych i maksymalnych w modelach klasy GARCH, rozprawa doktorska przygotowywana pod kierunkiem P. Fiszedera.
- [19] Rogers L. C. G., Satchell S. E., (1991) Estimating Variance From High, Low and Closing Prices, *The Annals of Applied Probability*, 1 (4), 504–512.
- [20] Yang D., Zhang Q., (2000), Drift-Independent Volatility Estimation Based on High, Low, Open, and Closing Prices, *Journal of Business*, 73, 477–491.

ESTYMACJA WARIANCJI ARYTMETYCZNEGO RUCHU BROWNA NA PODSTAWIE
ZNANYCH WARTOŚCI MINIMUM, MAKSIMUM, KOŃCOWEJ ORAZ DRYFU

Streszczenie

W opracowaniu poruszony jest problem wyznaczenia wariancji stopy zwrotu instrumentu finansowego na podstawie rynkowych notowań dziennych cen otwarcia, minimalnej, maksymalnej i zamknięcia. Wykorzystując znajomość łącznego rozkładu minimum, maksimum i wartości końcowej arytmetycznego ruchu Browna dokonano analizy porównawczej znanych estymatorów wariancji. Wyznaczono formuły wartości oczekiwanych bardzo wielu funkcji zmiennych losowych, które posłużyły do konstrukcji tych estymatorów. Ponadto, na ich podstawie zaproponowano nowy estymator wariancji. Dokonano analizy założeń, które przyjęto przy konstrukcji tego estymatora. Metodami analitycznymi porównano jego efektywność z efektywnością podstawowych znanych estymatorów zmienności dziennej.

Słowa kluczowe: estymator zmienności, cena otwarcia, minimum, maksimum, zamknięcia, ruch Browna

ESTIMATION OF ARITHMETIC BROWNIAN MOTION VARIANCE WHEN VALUES OF
MINIMUM, MAXIMUM, FINISH AND DRIFT ARE KNOWN

Abstract

This paper examines the problem of calculating the variance of returns of a financial instrument which is based upon the historical opening, closing, high, and low prices. For this purpose the joint distribution of minimum, maximum and final values of arithmetic Brownian motion was used. It gave a possibility to make a comparative analysis of the variance estimators. The formulae of expected values of many random variables, which were used for the construction of these estimators were calculated. Moreover, on the basis of those formulae, the new estimator of variance was proposed. The assumptions that were adopted for the construction of the estimator were examined. The efficiency of the proposed estimator was compared with the efficiency of the well-known estimators of daily volatility.

Key words: volatility estimator, open price, low price, high price, close price, Brownian motion