
МАТЕМАТИЧЕСКИЕ И ИНСТРУМЕНТАЛЬНЫЕ МЕТОДЫ ЭКОНОМИКИ

УДК 330.43, 519.6

НОВЫЙ СПОСОБ ОБНАРУЖЕНИЯ СТРУКТУРНЫХ СДВИГОВ В GARCH-МОДЕЛЯХ

Д. А. Борзых, М. А. Хасыков, А. А. Языков

Национальный исследовательский университет «Высшая школа экономики»

Поступила в редакцию 1 февраля 2017 г.

Аннотация: предлагается ранее не рассматривавшийся метод обнаружения структурных сдвигов временных рядов в рамках кусочно-заданных GARCH-моделей. Метод основан на анализе скользящей статистики отношения правдоподобия. В случае отсутствия структурных сдвигов для статистики отношения правдоподобия найдены нижняя и верхняя 95%- и 99%-границы. На основе этих границ выработан критерий наличия структурных сдвигов. Хорошие свойства предлагаемого метода подтверждаются численными экспериментами по методу Монте-Карло. В рамках проводимых испытаний получено, что метод обнаруживает правильное число структурных сдвигов примерно в 88 % случаев. При этом в случае верного обнаружения числа структурных сдвигов сами моменты структурных сдвигов устанавливаются достаточно точно. В случае отсутствия структурных сдвигов предлагаемый метод ошибочно обнаруживает структурные сдвиги достаточно редко – примерно в 2,5 % случаев. Метод апробирован на реальных данных в рамках решения задачи обнаружения структурных сдвигов волатильности доходности обыкновенных акций компании «Газпром».

Ключевые слова: GARCH, структурные сдвиги, волатильность, статистика отношения правдоподобия.

Abstract: the article proposes a new method of structural breaks detection in time series in the piecewise-specified GARCH-models. The method is based on the moving likelihood ratio statistics. In case of absence of structural breaks lower and upper 95%- and 99%-bounds were found for the likelihood ratio statistics. The criterion of structural breaks based on these bounds has been worked out. Good properties of the proposed method are supported by Monte Carlo numerical experiments. In the framework of performed calculations it is obtained that the method detects the correct number of structural changes approximately in 88 % of cases. In case of correct detection of number of structural changes the moments of the structural breaks are estimated quite accurately. In the absence of structural breaks the proposed method falsely detects structural breaks quite rarely – around 2,5 % of cases. The method is tested on the real data when detecting the structural breaks in the volatility of returns for “Gazprom” ordinary shares.

Key words: GARCH, structural breaks, volatility, likelihood ratio statistics.

1. Введение

Для получения более точных оценок коэффициентов эконометрических моделей требуется большее количество наблюдений. Однако при

расширении выборки исследователи сталкиваются с проблемой, называемой *структурными сдвигами* или *разладками* случайного процесса.

В данной статье мы предлагаем новый подход к обнаружению структурных сдвигов в рамках семейства кусочно-заданных GARCH-моделей вида:

$$\left\{ \begin{array}{l} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_1 + \delta_1 \cdot \sigma_{t-1}^2 + \gamma_1 \cdot \varepsilon_{t-1}^2, \\ t \leq \tau_1 - 1; \\ \dots \\ Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_2 + \delta_2 \cdot \sigma_{t-1}^2 + \gamma_2 \cdot \varepsilon_{t-1}^2, \\ \tau_1 \leq t \leq \tau_2 - 1, \\ \dots \\ \dots \\ Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_k + \delta_k \cdot \sigma_{t-1}^2 + \gamma_k \cdot \varepsilon_{t-1}^2, \\ \tau_{k-1} \leq t \leq \tau_k - 1, \\ \dots \\ Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_{k+1} + \delta_{k+1} \cdot \sigma_{t-1}^2 + \gamma_{k+1} \cdot \varepsilon_{t-1}^2, \\ t \geq \tau_k, \end{array} \right. \quad (1)$$

где $\omega_j, \delta_j, \gamma_j$ – неизвестные параметры модели, $j = 1, \dots, k + 1$, $(\xi_t)_{t \in Z}$ – последовательность независимых нормальных стандартных случайных величин, $\tau_1, \dots, \tau_k \in Z$ – моменты возможных структурных сдвигов, k – число возможных структурных сдвигов. Отметим, что моменты структурных сдвигов, а также их число считаются неизвестными.

Теперь перейдем к краткому обзору существующих методов обнаружения структурных сдвигов в рамках кусочно-заданных GARCH-моделей. Среди указанных методов наиболее известными являются процедуры, основанные на использовании CUSUM-статистики – кумулятивной суммы квадратов значения зависимой переменной или регрессионных остатков (см. [1–3]). Все статистики из работ [1–3] являются модификациями IT-статистики, описанной в статье [4], где кумулятивная сумма квадратов впервые была использована для обнаружения момента структурного сдвига дисперсии в модели с кусочно-постоянной дисперсией.

В статье [5] предложен подход *автосегментации* (Auto-Seg). Здесь задача обнаружения структурных сдвигов решается с помощью MDL-метода (Minimum Description Length). Согласно методу MDL, введенному в работе [6], из семейства возможных моделей выбирается та модель, которая наиболее компактно «сжимает» данные, т. е. позволяет получить самую короткую длину кода, восстанавливающую анализируемые данные.

Существуют также непараметрические методы поиска структурных сдвигов. Например, SLEX-метод [7] и NPCPM-метод [8]. SLEX-метод является модификацией представления Крамера стационарных случайных процессов в виде ряда Фурье, а NPCPM основан на классическом ранговом критерии Муда для проверки гипотез об однородности распределения наблюдений в различных выборках [9].

После краткого перечисления существующих подходов к обнаружению структурных сдвигов в кусочно-заданных GARCH-моделях перейдем к содержанию и основным результатам данной статьи.

В настоящей работе мы предлагаем ранее не рассматривавшийся метод обнаружения структурных сдвигов, основанный на вводимой в данной статье скользящей статистике отношения правдоподобия $LR_{\tau^*} \in Z$. Определение данной статистики приведено в разделе 2.

В разделе 3 для скользящей статистики отношения правдоподобия найдены нижняя \underline{q}_{LR} и верхняя \bar{q}_{LR} 95%- и 99%-границы. Данные границы используются в предлагаемом алгоритме обнаружения структурных сдвигов.

Раздел 4 посвящен описанию указанного алгоритма, идея которого состоит в следующем. Пусть в некоторой точке τ^* функция LR_{τ} имеет локальный максимум. Если $LR_{\tau^*} > \bar{q}_{LR}$, мы считаем, что в точке τ^* есть структурный сдвиг; если $LR_{\tau^*} < \underline{q}_{LR}$, то в точке τ^* структурного сдвига нет; если же $\underline{q}_{LR} \leq LR_{\tau^*} \leq \bar{q}_{LR}$, то значение LR_{τ^*} попало в зону неопределенности и в этом случае нельзя сделать окончательный вывод о наличии структурного сдвига.

В разделе 5 на симулированных данных исследуются свойства предлагаемого метода. В частности, в рамках проводимых испытаний получено, что данный метод обнаруживает правильное число структурных сдвигов примерно в 88 % случаев. При этом в случае верного обнаружения числа структурных сдвигов сами моменты структурных сдвигов устанавливаются достаточно точно. В случае отсутствия структурных сдвигов предлага-

емый метод ошибочно обнаруживает структурные сдвиги достаточно редко – примерно в 2,5 % случаев.

В разделе 6 метод апробируется на реальных данных в рамках решения задачи обнаружения структурных сдвигов волатильности доходности обыкновенных акций компании «Газпром».

2. Скользящая статистика отношения правдоподобия

В данном разделе мы опишем способ вычисления скользящей статистики отношения правдоподобия $LR_\tau, \tau \in Z$. Зафиксируем параметр h , отвечающий за ширину «скользящего окна», и рассмотрим фрагмент временного ряда $(Y_t)_{t=\tau-h}^{t=\tau+h}$, содержащего h наблюдений до и h наблюдений после текущего момента времени τ . Будем предполагать, что этот фрагмент $(Y_t)_{t=\tau-h}^{t=\tau+h}$ порожден моделью

$$(2) \quad \begin{cases} \left\{ \begin{array}{l} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_1 + \delta_1 \cdot \sigma_{t-1}^2 + \gamma_1 \cdot \varepsilon_{t-1}^2, \\ t \in [\tau - h; \tau - 1]; \end{array} \right. \\ \left\{ \begin{array}{l} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_2 + \delta_2 \cdot \sigma_{t-1}^2 + \gamma_2 \cdot \varepsilon_{t-1}^2, \\ t \in [\tau; \tau + h], \end{array} \right. \end{cases}$$

допускающей структурный сдвиг в момент времени τ . В этом случае логарифмическая функция правдоподобия имеет вид:

$$l(\theta_1, \theta_2) = -\frac{1}{2} \sum_{t=\tau-h}^{\tau-1} \left(\ln 2\pi + \ln \sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2} \right) - \frac{1}{2} \sum_{t=\tau}^{\tau+h} \left(\ln 2\pi + \ln \sigma_t^2 + \frac{\varepsilon_t^2}{\sigma_t^2} \right),$$

где

$$\theta_j := (\omega_j, \delta_j, \gamma_j) \in \Theta, j = 1, 2,$$

$$\Theta := \{(\omega, \delta, \gamma) : \omega > 0, \delta \geq 0, \gamma \geq 0, \delta + \gamma < 1\};$$

$$\varepsilon_t = \begin{cases} 0 & \text{при } t = \tau - h - 1, \\ Y_t & \text{при } t \in [\tau - h; \tau + h]; \end{cases}$$

$$\sigma_t^2 = \begin{cases} \frac{1}{h} \sum_{t=\tau-h}^{\tau-1} (Y_t - \bar{Y})^2 & \text{при } t = \tau - h - 1, \\ \omega_1 + \delta_1 \sigma_{t-1}^2 + \gamma_1 \varepsilon_{t-1}^2 & \text{при } t \in [\tau - h; \tau - 1], \\ \omega_2 + \delta_2 \sigma_{t-1}^2 + \gamma_2 \varepsilon_{t-1}^2 & \text{при } t \in [\tau; \tau + h]. \end{cases}$$

Определение. Скользящей статистикой отношения правдоподобия назовем величину:

$$LR_\tau := -2 \left(\max_{\theta_1, \theta_2 \in \Theta, \theta_1 = \theta_2} l(\theta_1, \theta_2) - \max_{\theta_1, \theta_2 \in \Theta} l(\theta_1, \theta_2) \right), \tau \in Z. (3)$$

Иными словами, скользящая статистика отношения правдоподобия в момент времени τ – это обычная статистика отношения правдоподобия соответствующая модели (2) при тестировании гипотезы

$$H_0 : \Theta_1 = \Theta_2$$

об отсутствии структурного сдвига в момент времени τ .

3. Критические точки скользящей статистики отношения правдоподобия

На множестве допустимых значений параметров

$$\Theta = \{(\omega, \delta, \gamma) : \omega > 0, \delta \geq 0, \gamma \geq 0, \delta + \gamma < 1\}$$

зададим сетку Θ_{net} :

– параметр ω пробегает все значения из отрезка $[\underline{\omega}; \bar{\omega}]$, где $\underline{\omega} = 0,0001$ и $\bar{\omega} = 0,031$, с шагом $\Delta\omega = 0,001$;

– параметр δ пробегает все значения из отрезка $[0; 1 - \Delta\delta]$ с шагом $\Delta\delta = 0,03$;

– при каждом фиксированном значении параметра δ параметр γ пробегает все значения из отрезка $[0; (1 - \Delta\gamma) - \delta]$ с шагом $\Delta\gamma = 0,03$.

Проекция множества Θ и сетки Θ_{net} на плоскость $\delta\theta\gamma$ схематично изображены на рис 1.

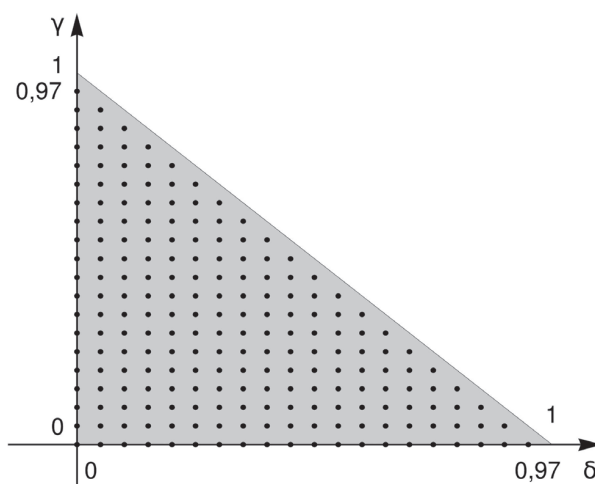


Рис. 1. Проекция множества Θ и сетки Θ_{net} на плоскость $\delta\theta\gamma$

Для каждого узла $(\omega', \delta', \gamma')$ сетки Θ_{net} проведем серию из 10 000 симуляций, на основе которых найдем 95%- и 99%-е квантили статистики отношения правдоподобия.

Опишем одну такую симуляцию. Зададим значение параметра h , отвечающего за ширину «окна». Согласно модели

$$\begin{cases} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega' + \delta' \cdot \sigma_{t-1}^2 + \gamma' \cdot \varepsilon_{t-1}^2, \\ t \in [-h; h] \end{cases}$$

сгенерируем ряд данных $(Y_t)_{t=-h}^{t=h}$. Используя полученный ряд $(Y_t)_{t=-h}^{t=h}$, оценим две модели:

модель без ограничения ((UR)-модель):

$$\begin{cases} \left\{ \begin{array}{l} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_1 + \delta_1 \cdot \sigma_{t-1}^2 + \gamma_1 \cdot \varepsilon_{t-1}^2, \\ t \in [-h; -1]; \end{array} \right. \\ \left\{ \begin{array}{l} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_2 + \delta_2 \cdot \sigma_{t-1}^2 + \gamma_2 \cdot \varepsilon_{t-1}^2, \\ t \in [0; h], \end{array} \right. \end{cases}$$

и модель с ограничением ((R)-модель), получаемую подстановкой ограничения $H_0 : \omega_1 = \omega_2, \delta_1 = \delta_2, \gamma_1 = \gamma_2$ в (UR)-модель:

$$\begin{cases} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega + \delta \cdot \sigma_{t-1}^2 + \gamma \cdot \varepsilon_{t-1}^2, \\ t \in [-h; h]. \end{cases}$$

Отметим, что ограничение, записанное в гипотезе H_0 , означает отсутствие структурного сдвига в точке $t = 0$.

С помощью (UR)- и (R)-моделей рассчитываем статистику отношения правдоподобия:

$$LR = -2(l_R(\hat{\omega}, \hat{\delta}, \hat{\gamma}) - l_{UR}(\hat{\omega}_1, \hat{\delta}_1, \hat{\gamma}_1, \hat{\omega}_2, \hat{\delta}_2, \hat{\gamma}_2)),$$

где $l_{UR}(\hat{\omega}_1, \hat{\delta}_1, \hat{\gamma}_1, \hat{\omega}_2, \hat{\delta}_2, \hat{\gamma}_2)$ – наибольшее значение логарифмической функции правдоподобия (UR)-модели, а $l_R(\hat{\omega}, \hat{\delta}, \hat{\gamma})$ – наибольшее значение логарифмической функции правдоподобия (R)-модели.

Для заданного узла сетки $(\omega', \delta', \gamma')$, выполнив серию из 10 000 симуляций, получаем 10 000 значений статистики отношения правдоподобия, с помощью которых находим 95 %- и 99 %-е квантили: $q_{LR}^{(\omega', \delta', \gamma')}(u, h)$, где $u = 0,95$ или $u = 0,99$.

Проведенные расчеты показывают, что в различных узлах (ω, δ, γ) сетки Θ_{net} квантили $q_{LR}^{(\omega, \delta, \gamma)}(u, h)$ могут весьма сильно различаться:

$$u = 0,95: \quad \underline{q}_{LR}^{\text{net}}(u, h) = 7,03 < \bar{q}_{LR}^{\text{net}}(u, h) = 67,35, \quad (4)$$

$$u = 0,99: \quad \underline{q}_{LR}^{\text{net}}(u, h) = 10,00 < \bar{q}_{LR}^{\text{net}}(u, h) = 157,66, \quad (5)$$

где

$$\begin{aligned} \underline{q}_{LR}^{\text{net}}(u, h) &:= \min_{(\omega, \delta, \gamma) \in \Theta_{\text{net}}} q_{LR}^{(\omega, \delta, \gamma)}(u, h), \\ \bar{q}_{LR}^{\text{net}}(u, h) &:= \max_{(\omega, \delta, \gamma) \in \Theta_{\text{net}}} q_{LR}^{(\omega, \delta, \gamma)}(u, h). \end{aligned}$$

Столь большое различие между левой и правой частями в формулах (4) и (5) (даже потенциально) не позволяют нам построить доверительную область, общую для всех возможных значений параметров $(\omega, \delta, \gamma) \in \Theta$, в которой гипотеза $H_0 : \omega_1 = \omega_2, \delta_1 = \delta_2, \gamma_1 = \gamma_2$ не отвергается.

Для решения указанной проблемы нам пришлось сузить множество допустимых значений параметров Θ до множества

$$\Theta_{\text{actual}} := \{(\omega, \delta, \gamma) : \underline{\omega} \leq \omega \leq \bar{\omega}, \underline{\delta} \leq \delta, \underline{\gamma} \leq \gamma, \delta + \gamma < 1\},$$

где $\underline{\omega} = 0,0001, \bar{\omega} = 0,031, \underline{\delta} = 0,7, \underline{\gamma} = 0$.

Тем не менее множество Θ_{actual} оказалось достаточно широким, чтобы обеспечить потребности реальных экономических задач. Дело в том, что для тех реальных финансовых временных рядов, для которых имеет место GARCH-эффект, параметр δ не бывает меньше 0,7. Далее, принимая во внимание известную формулу для дисперсии ошибок

$$D(\varepsilon_t) = \frac{\omega}{1 - \delta - \gamma} \text{ в GARCH-моделях, мы получаем}$$

соотношение $D(\varepsilon_t) \geq \omega \geq \underline{\omega}$, из которого следует, что на множестве Θ_{actual} минимальное значение волатильности в годовом выражении в процентах равно

$$\sqrt{\underline{\omega}} \times \sqrt{250} \times 100\% = \sqrt{0,0001} \times \sqrt{250} \times 100\% \approx 15,81\%$$

(здесь в качестве приближенного числа торговых дней в году взято число 250). Максимальное же значение волатильности на множестве Θ_{actual} в го-

довом выражении в процентах может быть оценено снизу как

$$\sqrt{\bar{\omega}} \times \sqrt{250} \times 100\% = \sqrt{0,031} \times \sqrt{250} \times 100\% = 278,39\%.$$

Таким образом, диапазон значений волатильности на множестве Θ_{actual} от 15,81 до 278,39 % покрывает те возможные значения волатильности, которые встречаются для реальных финансовых временных рядов.

На введенном выше множестве Θ_{actual} зададим сетку $\Theta_{\text{actual net}}$:

- параметр ω пробегает все значения из отрезка $[\underline{\omega}; \bar{\omega}]$ с шагом $\Delta\omega = 0,001$;
- параметр δ пробегает все значения из отрезка $[\underline{\delta}; 1 - \Delta\delta]$ с шагом $\Delta\delta = 0,03$;
- при каждом фиксированном значении параметра δ параметр γ пробегает все значения из отрезка $[\underline{\gamma}; (1 - \Delta\gamma) - \delta]$ с шагом $\Delta\gamma = 0,03$.

Проекция множества Θ_{actual} и сетки $\Theta_{\text{actual net}}$ на плоскость $\delta\gamma$ схематично изображены на рис. 2.

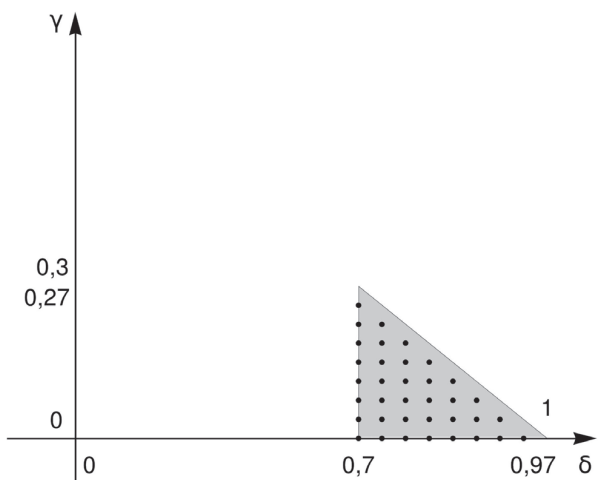


Рис. 2. Проекция множества Θ_{actual} и сетки $\Theta_{\text{actual net}}$ на плоскость $\delta\gamma$

Так же, как это сделано для сетки Θ_{net} , для сетки $\Theta_{\text{actual net}}$ рассчитаем нижнюю и верхнюю границы 95%- и 99%-х квантилей статистики отношения правдоподобия (табл. 1).

Т а б л и ц а 1

Значения $q_{LR}^{\text{actual net}}(u, h)$ и $\bar{q}_{LR}^{\text{actual net}}(u, h)$ при $h = 200$

	$q_{LR}^{\text{actual net}}(u, h)$	$\bar{q}_{LR}^{\text{actual net}}(u, h)$
$u = 0,95$	7,03	11,09
$u = 0,99$	10,00	17,78

здесь

$$q_{LR}^{\text{actual net}}(u, h) := \min_{(\omega, \delta, \gamma) \in \Theta_{\text{actual net}}} q_{LR}^{(\omega, \delta, \gamma)}(u, h),$$

$$\bar{q}_{LR}^{\text{actual net}}(u, h) := \max_{(\omega, \delta, \gamma) \in \Theta_{\text{actual net}}} q_{LR}^{(\omega, \delta, \gamma)}(u, h).$$

Как видно из табл. 1, при переходе от множества Θ к более узкому множеству Θ_{actual} «зазор» между значениями $q_{LR}^{\text{actual net}}(u, h)$ и $\bar{q}_{LR}^{\text{actual net}}(u, h)$ существенно уменьшился по сравнению с прежними значениями $q_{LR}^{\text{net}}(u, h)$ и $\bar{q}_{LR}^{\text{net}}(u, h)$ (см. формулы (4) и (5)). Это обстоятельство позволило нам построить доверительную область, общую для всех возможных значений параметров $(\omega, \delta, \gamma) \in \Theta_{\text{actual}}$, в которой гипотеза $H_0: \omega_1 = \omega_2, \delta_1 = \delta_2, \gamma_1 = \gamma_2$ не отвергается.

4. Алгоритм обнаружения структурных сдвигов, основанный на скользящей статистике отношения правдоподобия

Предположим, что в точке $\tau' \in \mathbf{Z}$ есть структурный сдвиг. Ясно, что в случае, когда диапазон наблюдений $[\tau - h; \tau + h]$ не содержит точку τ' (и не содержит других структурных сдвигов), величины $\max_{\theta_1, \theta_2 \in \Theta} l(\theta_1, \theta_2)$ и $\max_{\theta_1, \theta_2 \in \Theta, \theta_1 = \theta_2} l(\theta_1, \theta_2)$ из формулы (3) принимают достаточно близкие значения, и, стало быть, статистика LR_τ близка к нулю. Если же этот диапазон содержит структурный сдвиг τ' (и не содержит других структурных сдвигов), статистика LR_τ принимает достаточно большое значение; при этом самое большое значение скользящая статистика отношения правдоподобия LR_τ принимает приблизительно в момент времени $\tau = \tau'$.

На рис. 3 приведен типичный график скользящей статистики отношения правдоподобия в случае двух структурных сдвигов в моменты времени τ'_1 и τ'_2 .

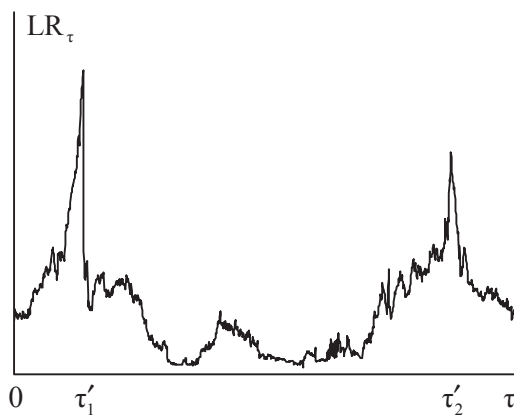


Рис. 3. Типичный вид графика скользящей статистики отношения правдоподобия в случае двух структурных сдвигов в моменты времени τ'_1 и τ'_2

Алгоритм обнаружения структурных сдвигов. Пусть в некоторой точке $\tau^* \in \mathbf{Z}$ функция LR_τ имеет h -локальный¹ максимум ($\forall t \in [\tau^* - h; \tau^* + h] \setminus \{\tau^*\}: LR_t < LR_{\tau^*}$), не являющийся выбросом (определение выброса дано ниже). Если $LR_{\tau^*} > \bar{q}_{LR}$, мы считаем, что в точке τ^* есть структурный сдвиг; если $LR_{\tau^*} < \underline{q}_{LR}$, то в точке τ^* структурного сдвига нет; если же $\underline{q}_{LR} \leq LR_{\tau^*} \leq \bar{q}_{LR}$, то значение LR_{τ^*} попало в зону неопределенности и в этом случае окончательно вывода о наличии структурного сдвига сделать нельзя.

Под выбросами скользящей статистики отношения правдоподобия мы подразумеваем те наблюдения, в которых значения статистики LR_τ скачкообразно принимают аномально высокие значения. Визуально выбросы определяются по графику скользящей статистики отношения правдоподобия как резко выдающиеся вверх тонкие пики (4–5 наблюдений в ширину).

Пример одного такого выброса приведен на рис. 4.

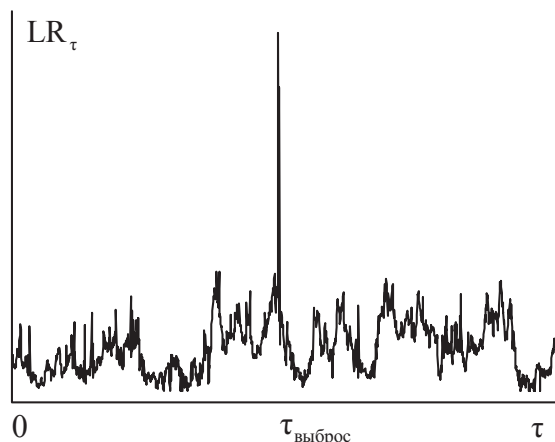


Рис. 4. Типичный вид выброса в графике скользящей статистики отношения правдоподобия

Более формально момент τ^* назовем *выбросом*, если

$$LR_{\tau^*} > 1,5 \times \frac{LR_{\tau^*-3} + LR_{\tau^*-2} + LR_{\tau^*+2} + LR_{\tau^*+3}}{4},$$

т. е. если значение статистики LR_{τ^*} более чем в полтора раза превышает среднее арифметическое значений LR_τ , где $t \in \{\tau^* - 3, \tau^* - 2, \tau^* + 2, \tau^* + 3\}$.

¹ В работе используется достаточно большое значение h : $h = 200$. Это связано с тем, что для адекватного оценивания параметров модели на сегментах, не содержащих структурных сдвигов, необходимо достаточное количество наблюдений.

5. Апробирование метода на симулированных данных

В данном разделе выполнено два численных эксперимента по методу Монте-Карло, которые позволяют судить о достаточно хороших свойствах предлагаемого способа обнаружения структурных сдвигов.

В первом эксперименте истинная модель содержит два структурных сдвига. В рамках проведенного эксперимента показано, что предлагаемый метод обнаруживает правильное число структурных сдвигов примерно в 88 % случаев. При этом в случае верного обнаружения числа структурных сдвигов сами моменты структурных сдвигов устанавливаются достаточно точно.

Во втором численном эксперименте истинная модель, напротив, не содержит структурных сдвигов. Проведенные расчеты показывают, что метод ошибочно обнаруживает структурные сдвиги только примерно в 2,5 % случаев. В остальных же 97,5 % случаев метод верно указывает на отсутствие структурных сдвигов.

Для удобства мы ввели следующую терминологию.

Определение. Будем считать, что очередная симуляция завершилась *слабым успехом*, если в данной симуляции метод обнаружил правильное количество структурных сдвигов, независимо от точности обнаружения этих сдвигов. Если же в очередной симуляции произошел слабый успех и при этом каждый из структурных сдвигов был обнаружен с погрешностью не более Δ , в этом случае мы будем говорить о (Δ)-*сильном успехе*.

Численный эксперимент № 1 (два структурных сдвига). В данном численном эксперименте рассматривается модель

$$\begin{cases} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_1 + \delta_1 \cdot \sigma_{t-1}^2 + \gamma_1 \cdot \varepsilon_{t-1}^2, \\ t \in [0; \tau_1 - 1]; \\ Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_2 + \delta_2 \cdot \sigma_{t-1}^2 + \gamma_2 \cdot \varepsilon_{t-1}^2, \\ t \in [\tau_1; \tau_2 - 1]; \\ Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega_3 + \delta_3 \cdot \sigma_{t-1}^2 + \gamma_3 \cdot \varepsilon_{t-1}^2, \\ t \in [\tau_2; T], \end{cases} \quad (6)$$

содержащая два структурных сдвига – в точках $\tau_1 = 501$ и $\tau_2 = 1501$. На промежутках времени $[0; \tau_1 - 1]$, $[\tau_1; \tau_2 - 1]$ и $[\tau_2; T]$ параметры в модели (6) выбраны следующим образом:

$$\begin{aligned} \omega_1 &= 0,0001, & \delta_1 &= 0,98, & \gamma_1 &= 0, \\ \omega_2 &= 0,0006, & \delta_2 &= 0,98, & \gamma_2 &= 0, \\ \omega_3 &= 0,0001, & \delta_3 &= 0,98, & \gamma_3 &= 0. \end{aligned}$$

Таким образом, параметры δ и γ на всем промежутке $[0; T]$ остаются неизменными, а параметр ω в момент времени τ_1 совершает скачок вверх на величину 0,0005, а в момент времени τ_2 он совершает обратный скачок вниз на ту же величину 0,0005.

Нами проведен численный эксперимент, состоящий из 10 000 симуляций, в каждой из которых согласно модели (6) генерировался ряд длины $T = 2000$. В расчетах использовались границы $q_{LR}^{actual\ net}(u, h)$, $\bar{q}_{LR}^{actual\ net}(u, h)$ при $u = 0,99$ и $h = 200$. Среди 10 000 симуляций произошло:

- слабых успехов: 8755 (87,55 %);
- (100)-сильных успехов: 8109 (81,09 %);
- (50)-сильных успехов: 7880 (78,80 %);
- (25)-сильных успехов: 7512 (75,12 %);
- (10)-сильных успехов: 6349 (63,49 %);
- (5)-сильных успехов: 4454 (44,54 %).

Для изучения качества оценивания моментов структурных сдвигов τ_1 и τ_2 рассмотрим две характеристики:

$$\text{Mean}(\hat{\tau}_i) := \frac{1}{N} \sum_{j=1}^N \hat{\tau}_i^{(j)} \text{ – среднее,}$$

$$\text{RMSE}(\hat{\tau}_i) := \sqrt{\frac{1}{N} \sum_{j=1}^N (\hat{\tau}_i^{(j)} - \text{Mean}(\hat{\tau}_i))^2} \text{ – стандартная ошибка,}$$

где $\hat{\tau}_i^{(j)}$ – оценка момента структурного сдвига τ_i , $i = 1, 2$, в j -й симуляции; N – общее число симуляций, по которым вычисляются характеристики Mean и RMSE.

Для подвыборки слабых успехов и подвыборки (100)-сильных успехов мы рассчитали характеристики Mean и RMSE (табл. 2 и 3).

Т а б л и ц а 2

Среднее значение и стандартное отклонение $\hat{\tau}_1$ и $\hat{\tau}_2$ по подвыборке слабых успехов, содержащей 8755 наблюдений

	$\hat{\tau}_1$	$\hat{\tau}_2$
Mean	503,79	1492,50
RMSE	28,51	33,15

Как видно из табл. 2, по подвыборке слабых успехов момент первого структурного сдвига $\tau_1 = 501$ в среднем оценивается как $\text{Mean}(\hat{\tau}_1) = 503,79$, а момент второго структурного сдвига $\tau_2 = 1501$ оценивается как $\text{Mean}(\hat{\tau}_2) = 1492,50$. При этом стандартные ошибки оценивания составляют $\text{RMSE}(\hat{\tau}_1) = 28,51$ и $\text{RMSE}(\hat{\tau}_2) = 33,15$.

Если же из подвыборки слабых успехов исключить 646 наиболее «неудачных» наблюдений, в которых абсолютная погрешность оценивания хотя бы одного из параметров τ_1 или τ_2 превышает 100, и по оставшейся части выборки рассчитать характеристики Mean и RMSE, то получим следующие результаты (см. табл. 3).

Т а б л и ц а 3

Среднее значение и стандартное отклонение $\hat{\tau}_1$ и $\hat{\tau}_2$ по подвыборке (100)-сильных успехов, содержащей 8109 наблюдений

	$\hat{\tau}_1$	$\hat{\tau}_2$
Mean	501,04	1497,88
RMSE	11,36	11,50

Результаты, приведенные в табл. 3, трактуются аналогично результатам табл. 2.

Если сравнить результаты из табл. 2 и 3, видим, что при отбрасывании небольшой части выборки (приблизительно 7 % наблюдений), стандартные ошибки RMSE снижаются практически в три раза. Это означает, что сравнительно небольшая часть выборки слабых успехов (7 %) существенно «портит» общую картину точности оценивания. По большей же части подвыборки слабых успехов (93 % наблюдений) оценивание параметров τ_1 и τ_2 проводится достаточно точно (см. табл. 3).

Численный эксперимент № 2 (структурные сдвиги отсутствуют). В отличие от предыдущего численного эксперимента в данном численном эксперименте рассматривается модель, не содержащая структурных сдвигов:

$$\begin{cases} Y_t = \varepsilon_t, \\ \varepsilon_t = \sigma_t \cdot \xi_t, \\ \sigma_t^2 = \omega + \delta \cdot \sigma_{t-1}^2 + \gamma \cdot \varepsilon_{t-1}^2, \\ t \in [0; T]. \end{cases} \quad (7)$$

Параметры в модели выбраны следующим образом:

$$\omega = 0,0001, \quad \delta = 0,98, \quad \gamma = 0.$$

Нами проведен эксперимент, состоящий из 10 000 симуляций, в каждой из которых согласно модели (7) генерировался ряд длины $T = 2000$ наблюдений. Так же как и в численном эксперименте № 1, мы использовали границы $q_{LR}^{\text{actual net}}(u, h)$, $\bar{q}_{LR}^{\text{actual net}}(u, h)$ при $u = 0,99$ и $h = 200$. В данном эксперименте:

- в 253 случаях из 10 000 метод ошибочно обнаружил структурные сдвиги (2,53 %);
- в 9747 случаях из 10 000 метод правильно указал на отсутствие структурных сдвигов (97,47 %).

Полученные результаты говорят о том, что предлагаемый метод ошибочно обнаруживает структурные сдвиги достаточно редко.

Таким образом, результаты численных экспериментов № 1 и 2 позволяют судить о достаточно хороших свойствах рассматриваемого в статье метода.

6. Апробирование метода на реальных данных

Данный раздел посвящен апробированию предлагаемого метода на реальных данных. В качестве исследуемого ряда был взят ряд логарифмических доходностей обыкновенных акций компании «Газпром» за промежуток времени с 23 января 2006 г. по 3 ноября 2016 г. Структурные сдвиги волатильности оценивались в рамках кусочно-заданной GARCH(1,1)-модели с применением описанного выше алгоритма. В расчетах использовались границы $q_{LR}^{\text{actual net}}(u, h)$, $\bar{q}_{LR}^{\text{actual net}}(u, h)$ при $u = 0,99$ и $h = 200$.

Алгоритмом выявлено четыре структурных сдвига (9 марта 2007 г., 19 сентября 2008 г., 8 сентября 2009 г. и 10 июля 2013 г.), а также два момен-

та, подозрительных на структурные сдвиги (10 февраля 2011 г. и 8 декабря 2014 г.), для которых значение скользящей статистики отношения правдоподобия попало в зону неопределенности.

Ниже приведен график скользящей статистики отношения правдоподобия с отмеченными моментами структурных сдвигов (рис. 5).

Дальнейшее тестирование осуществлялось путем соотнесения обнаруженных моментов структурных сдвигов (точек $A-F$ на рис. 5) с соответствующими значимыми событиями в экономике:

- (A: 9 марта 2007 г.) 9 февраля 2007 г. было объявлено о том, что «Газпром» и СУЭК (самый крупный производитель угля в России) намерены объединить свои энергоактивы в совместную компанию, контрольный пакет акций которой будет принадлежать газовому концерну;
- (B: 19 сентября 2008 г.) 15 сентября 2008 г. инвестиционный банк «Lehman Brothers» обратился в суд с заявлением о своем банкротстве;
- (C: 8 сентября 2009 г.) – начало роста цен на природный газ после годового падения;
- (D: 10 февраля 2011 г.) 16 февраля 2011 г. «Газпром» приобрел у крупнейшей итальянской нефтегазовой компании ENI 50%-ю долю участия в ливийском нефтяном проекте Elephant (стоимость сделки составила около 163 млн долл. США);
- (E: 10 июля 2013 г.) 11 июля 2013 г. Правительство России предоставило «Газпрому» лицензии на разведку и добычу газа на трех крупных месторождениях: Лудловском, Русановском и Белоостровском участках в Карском море;
- (F: 8 декабря 2014 г.) 1 декабря 2014 г. Президент России В. В. Путин заявил о том, что Россия не может продолжать реализацию проекта транснационального газопровода «Южный поток».

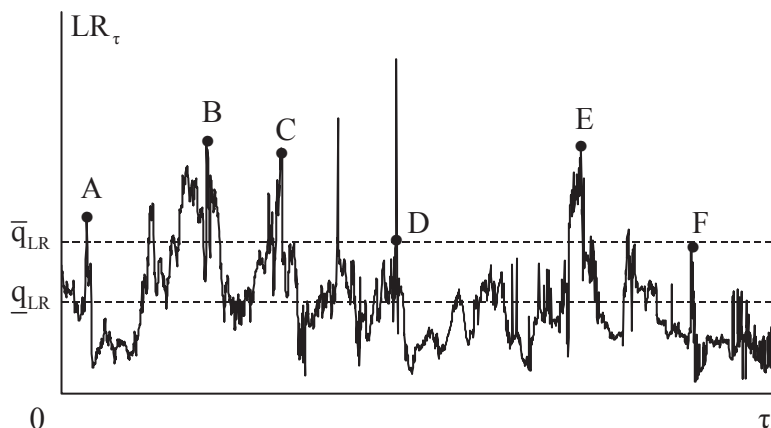


Рис. 5. График скользящей статистики отношения правдоподобия

ЛИТЕРАТУРА

1. Lee S. On the CUSUM test for parameter changes in GARCH(1,1) Models / S. Lee, S. Kim, S. Cho // Communications in Statistics – Theory and Methods. – 2000. – № 29(2). – P. 445–462.
2. Lee S. The CUSUM test for parameter change in regression models with ARCH errors / S. Lee, Y. Tokutsu, K. Maekawa // Journal of the Japanese Statistical Society. – 2004. – № 34. – P. 173–188.
3. Kokoszka P. Change-point estimation in ARCH models / P. Kokoszka, R. Leipus // Bernoulli. – 2000. – № 6(3). – P. 513–539.
4. Inclán C. Use of cumulative sums of squares for retrospective detection of changes of variance / C. Inclán, G. Tiao // Journal of the American Statistical Association. – 1994. – № 89 (427). – P. 913–923.
5. Davis R. Break detection for a class of nonlinear time series models / R. Davis, T. Lee, G. Rodriguez-Yam // Journal of Time Series Analysis. – 2008. – № 29(5). – P. 834–867.
6. Rissanen J. Stochastic Complexity in Statistical Inquiry / J. Rissanen // Singapore : World Scientific. – 1989.
7. Ombao H. The SLEX Model of a NonStationary Random Process / H. Ombao, J. Raz, R. von Sachs, W. Guo // Annals of the Institute of Statistical Mathematics. – 2002. – № 54(1). – P. 171–200.
8. Ross G. J. Modeling Financial Volatility in the Presence of Abrupt Changes / G. J. Ross // Physica A. Statistical Mechanics and its Applications. – 2013. – № 192(2). – P. 350–360.
9. Mood A. M. On the Asymptotic Efficiency of Certain Nonparametric Two-sample Tests / A. M. Mood // Annals of Mathematical Statistics. – 1954. – № 25 (3). – P. 514–522.

Национальный исследовательский университет «Высшая школа экономики»

Борzych Д. А., старший преподаватель департамента прикладной экономики факультета экономических наук

E-mail: dborzykh@hse.ru

Тел.: (495)772-95-90 (доб. 26050)

Хасыков М. А., студент 2-го курса магистратуры магистерской программы «Прикладная экономика» факультета экономических наук

E-mail: mkhasykov@gmail.com

Языков А. А., стажер-исследователь научной лаборатории макроструктурного моделирования экономики России

E-mail: ayazikov@hse.ru

Тел.: (495)772-95-90 (доб. 26050)

National Research University Higher School of Economics

Borzykh D. A., Senior Lecturer of Department of Applied Economics of Faculty of Economic Sciences

E-mail: dborzykh@hse.ru

Tel.: (495)772-95-90 (add. 26050)

Khasykov M. A., 2nd year Student of the Master's Program «Applied Economics» of Faculty of Economic Sciences

E-mail: mkhasykov@gmail.com

Yazikov A. A., Research Intern of Laboratory for Macro-Structural Modeling of the Russian Economy

E-mail: ayazikov@hse.ru

Tel.: (495)772-95-90 (add. 26050)