

Прикладная эконометрика, 2021, т. 61, с. 110–139.

Applied Econometrics, 2021, v. 61, pp. 110–139.

DOI: 10.22394/1993-7601-2021-61-110-139

А. В. Костырка, Д. И. Малахов¹

А был ли сдвиг: эмпирический анализ тестов на структурные сдвиги в волатильности доходностей

В данной работе рассматриваются два популярных метода выявления структурных сдвигов в волатильности доходностей: ICSS-алгоритм и оценивание методом наименьших квадратов (МНК). Показано, что ICSS-алгоритм чувствителен ко многим характеристикам рядов, и использование асимптотических критических значений не всегда обоснованно. В сравнительных симуляциях МНК более точно оценивал количество и время сдвигов, особенно когда их несколько. При анализе реальных данных также получается, что МНК дает более согласованные с экономической интуицией результаты, хотя они чувствительны к длине выборки.

Ключевые слова: структурный сдвиг; волатильность; GARCH.

JEL classification: C12; C22; C58.

1. Введение

Модели условной волатильности из семейства GARCH крайне популярны в риск-менеджменте. Как показано в работах (Galeano, Tsay, 2010; Hillebrand, 2005; Lamoureux, Lastrapes, 1990; Mikosch, Stărică, 2004), при оценке GARCH-моделей на реальных данных часто получается, что доходности не являются слабо стационарными. Из нестационарности напрямую следует, что безусловная (долгосрочная) дисперсия доходностей не существует, а цены опционов должны быть очень чувствительными к начальным условиям. Однако оба этих следствия слабо согласуются с наблюдаемыми фактами. Некоторые исследователи предположили, что эта нестационарность является ложной, т. к. структурные сдвиги (т. е. изменения параметров моделей, вызванные сменой глубинных экономических механизмов) создают поведение рядов, похожее на нестационарное (Galeano, Tsay, 2010; Hillebrand, 2005). С общеэкономической точки зрения этот подход выглядит более реалистичным. При этом, как показано в (Galeano, Tsay, 2010), учет структурных сдвигов позволяет улучшить прогноз GARCH-моделей, а также может помочь сделать статистические выводы относительно параметров более надежными. Таким образом, после обнаружения сдвигов исследователь может оценивать модели на подвыборках или использовать более гибкие модели, например, с марковскими переключениями.

¹ Костырка Андрей Викторович — Университет Люксембурга, Люксембург; andrei.kostyrka@uni.lu.

Малахов Дмитрий Игоревич — Национальный исследовательский университет «Высшая школа экономики»; dmalakhov@hse.ru.

Основная цель статьи — продемонстрировать особенности поведения тестов на структурные сдвиги в безусловной дисперсии доходностей на конечных выборках путем сопоставления их поведения на симулированных данных, содержащих часто встречающиеся на практике типы сдвигов. Также даются некоторые прикладные рекомендации по поводу выбора вспомогательных параметров, на которые опираются тесты. Данные результаты позволят точнее определить наличие сдвигов и стационарность доходностей, а значит, помогут сделать статистические выводы в моделях ценообразования активов и риск-менеджмента более надежными. Существуют десятки тестов² на структурные сдвиги, некоторые из которых можно использовать для анализа дисперсии. Их область применения не ограничивается параметрическими моделями условной гетероскедастичности, но в статье сконцентрировано внимание именно на этом семействе структурных моделей, т. к. оно является одним из самых популярных на практике среди всех семейств моделей волатильности доходностей, и при этом наиболее простым для изучения. На практике достаточно широко используются два способа выявления сдвигов в среднем, адаптированных для поиска сдвигов в безусловной дисперсии, на сравнении которых сосредоточено внимание в данной статье. Первый тест, который рассматривается — *скорректированный тест Inclán–Tiao* (1994) (АИТ-тест), встроенный в *ICSS-алгоритм* (Rapach et al., 2008). Этот тест основан на простой идее, что если происходит сдвиг в волатильности доходности, то накопленная сумма квадратов доходностей должна отразить этот сдвиг через резкое изменение своего значения. Второй способ выявления сдвигов основан на методе наименьших квадратов (МНК) — минимизации суммы квадратов отклонений квадратов доходностей от локальных средних значений при учете штрафа за введение дополнительных степеней свободы (Andreou, Ghysels, 2002). В данной статье *МНК-процедурой* для краткости называется выбор числа сдвигов на основе информационных критериев. Строго говоря, МНК-процедура является не тестом, а методом оценивания числа сдвигов³, поэтому исследователь получает только оценки числа и точек сдвига, а самого статистического тестирования не проводится. Здесь будем рассматривать только два этих метода в том виде, как они описаны в вышеуказанных работах, а изучение их возможных расширений и комбинаций выходит за рамки настоящего исследования⁴.

При анализе АИТ-теста показано, что в конечных выборках его эмпирический размер может быть далек от номинального, а мощность может быть невысокой из-за нескольких факторов: почти нестационарного поведения процесса, метода робастного оценивания дисперсии в АИТ-статистике и типа модели, которая порождает данные (data-generating process). Поэтому лучше пользоваться не асимптотическими критическими значениями, а значениями, полученными при симуляциях процесса, близкого к исследуемому. МНК-оценивание более робастно из-за того, что оно, в отличие от ICSS-алгоритма с многократным

² Число тестов на структурные сдвиги во временных рядах крайне велико, см., например, недавний обзор (Скроботов, 2020), поэтому проблема выбора наиболее подходящего теста (по статистической мощности, по близости эмпирического размера к теоретическому, по типу данных, к которым он применяется) стоит достаточно остро.

³ В работе (Andreou, Ghysels, 2002) процедура оценивания числа и времени сдвигов из оригинальной статьи (Lavielle, Moulines, 2000), не содержащей слова «тест», для единообразия переименована в «Lavielle and Moulines (2000) test».

⁴ Среди прочих тестов на обнаружение сдвигов в волатильности стоит отметить работы (Xu, 2013a, b; Wu, Xiao, 2018) и, отчасти, (Perron et al., 2020).

проведением АПТ-теста, производится за один шаг, не требует расчета критических значений, зависящих от вспомогательных методов, и позволяет проверить гипотетически любое число сдвигов. При этом оценивание истинного числа сдвигов сводится к выбору всего лишь конкретной штрафной функции, а в случае наличия в ряде нескольких немонотонных сдвигов (например, в реальных финансовых данных при переключении между режимами высокой и низкой волатильности) МНК дает менее смещенную оценку числа сдвигов.

В сравнительных симуляциях с 8 различными типами процессов показывается, что АПТ-тест часто оказывается недостаточно статистически мощным (не отвергается нулевая гипотеза об отсутствии сдвига при его наличии). Поэтому даже на «либеральном» уровне значимости 10% (при котором должна чаще отвергаться нулевая гипотеза) ICSS-алгоритм с этим тестом выявляет лишь небольшую часть сдвигов, если их действительно много и они расположены не в середине рассматриваемого периода выборки. Если же сдвиг найден не слишком близко к краю выборки, то оценка даты сдвига будет центрирована вокруг истинного значения. МНК-процедура точнее оценивает число множественных сдвигов, если использовать ее вариант с умеренным регуляризирующим штрафом. После получения оценки числа сдвигов МНК-процедура определяет даты сдвигов также достаточно точно. При анализе реальных данных по доходностям американского, немецкого и японского индексов обнаружено больше доводов в пользу МНК-процедуры по сравнению с АПТ-тестом с общеэкономической точки зрения, однако ее результаты оказались более чувствительными к длине используемой выборки.

Структура работы следующая. В разделе 2 дается краткое описание семейства GARCH-моделей и обсуждается проблема ложной нестационарности в этих моделях. В разделе 3 приводится описание тестов на структурные сдвиги и анализируется чувствительность АПТ-теста к свойствам ряда. В разделе 4 приводятся результаты сравнительных симуляций двух методов. В разделе 5 эти методы применяются к реальным данным. Общие выводы приводятся в последнем разделе.

2. GARCH-модели: спецификации, распределения и (ложная) нестационарность

Пусть имеется временной ряд наблюдений цены некоторого актива $\{p_t\}_{t=0}^T$. Доходности этого актива могут быть определены как $r_t = \ln(p_t/p_{t-1})$. Их можно разложить на долгосрочное среднее и шумовую компоненту:

$$r_t = \mu + \varepsilon_t. \quad (1)$$

Шум ε_t можно разложить на произведение стандартизированной ошибки z_t и условного стандартного отклонения σ_t , т. е. $\varepsilon_t = \sigma_t z_t$. Предполагается, что z_t — независимые и одинаково распределенные случайные величины, такие что $E z_t = 0$, $\text{Var } z_t = 1$.

В данной работе рассматриваются три популярные функциональные формы для уравнения условной дисперсии доходностей.

1. $\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2$ — классическая модель GARCH(1,1) (Bollerslev, 1986).
2. $\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \gamma \varepsilon_{t-1}^2 I(\varepsilon_{t-1} < 0) + \beta \sigma_{t-1}^2$ — асимметричная модель GJR-GARCH(1,1) (Glosten et al., 1993), в которой коэффициент γ определяет разницу в воздействии

отрицательных и положительных рыночных шоков, так называемый эффект рычага; $I(\cdot)$ — индикаторная функция, равная 1, если условие в скобках выполнено, и 0 иначе.

3. $\ln \sigma_t^2 = \omega + \alpha z_{t-1} + \gamma (|z_{t-1}| - E|z_{t-1}|) + \beta \ln \sigma_{t-1}^2$ — экспоненциальная модель EGARCH(1,1) (Nelson, 1991), в которой коэффициент γ имеет тот же смысл, что и в предыдущей модели. Также EGARCH(1,1) позволяет лучше учитывать нелинейный отклик дисперсии на шоки предыдущих периодов.

Кроме того, условное распределение ε_t может зависеть от некоторых не меняющихся во времени дополнительных параметров (скошенность, форма и др.). В качестве распределения базовых шоков будем предполагать нормальное или скошенное обобщенное нормальное распределение⁵ (далее — СОН). Скошенность распределения задается по методологии (Fernández, Steel, 1998). Если $f(x; \dots)$ — симметричная вокруг значения m функция плотности распределения некоторой непрерывной случайной величины, которая может зависеть от других параметров (обозначены многоточием), то ее *скошенное обобщенное* распределение, зависящее от тех же параметров и параметра скошенности s , задается функцией плотности

$$g(x; m, s, \dots) = \frac{2}{s + 1/s} \left[f\left(\frac{x-m}{s}; \dots\right) \cdot I(x > m) + f\left(s(x-m); \dots\right) \cdot I(x \leq m) \right].$$

Теперь перейдем к обсуждению критериев слабой стационарности GARCH-процессов⁶. В работе (Ling, McAleer, 2002) приводится условие, необходимое для стационарности процессов GARCH и GJR-GARCH, а в (McAleer, Hafner, 2014) — условие стационарности процесса EGARCH. Эти условия опираются на некоторую функцию $c(z_{t-1})$, являющуюся частью структурного уравнения GARCH. Для существования у процесса семейства GARCH безусловного момента порядка $2m$ достаточно выполнения следующего условия (при некоторых неограничительных предположениях): $E c(z_{t-1})^m < 1$ (см. (Ling, McAleer, 2002)). Это означает, что для существования конечного второго момента доходностей достаточно, чтобы $E c(z_{t-1}) < 1$. Приведем выражения для функций $c(\cdot)$ рассматриваемых в данной работе моделей. В GARCH(1,1) $c(z_{t-1}) = \beta + \alpha z_{t-1}^2$, следовательно, достаточным условием стационарности будет $E c(z_{t-1}) = \alpha + \beta < 1$. В модели GJR-GARCH(1,1) $c(z_{t-1}) = \beta + \alpha z_{t-1}^2 + \gamma I(z_{t-1} < 0) z_{t-1}^2$, откуда требуется выполнение $\alpha + \beta + b\gamma < 1$, где $b = z_t^2 P(z_t < 0)$; $b = 1/2$ для симметричных распределений (в том числе для нормального), а значение b для СОН-распределения зависит от параметров скошенности и формы. В модели EGARCH(1,1) требуется только условие $E c(z_{t-1}) = \beta < 1$.

Из таблицы 1, содержащей оценки указанных выше моделей по выборке дневных данных S&P500 за период последних 50 лет, видно, что практически во всех моделях $E c(z_{t-1}) \approx 1$, поэтому доходности американского рынка, скорее всего, нестационарны. Однако, как показано в работах (Galeano, Tsay, 2010) и (Hillebrand, 2005), процесс, выглядящий нестационарным,

⁵ Функция плотности обобщенного нормального распределения со средним m определяется как $f(x; m, \eta, \sigma) = \eta \cdot \exp\left(-(|x-m|/\sigma)^\eta\right) / (2\sigma \Gamma(1/\eta))$, где $\Gamma(\cdot)$ — гамма-функция, η — параметр формы. При этом дисперсия равна $\sigma^2 \Gamma(3/\eta) / \Gamma(1/\eta)$, а коэффициент эксцесса — $\Gamma(5/\eta) \Gamma(1/\eta) / \Gamma(3/\eta)^2 - 3$. Стандартное распределение Гаусса соответствует случаю $\eta = 2$, $\sigma = \sqrt{2}$.

⁶ Заметим, что поскольку GARCH-модели имеют более толстые хвосты, чем модели с постоянной условной дисперсией, часто условия слабой стационарности бывают более жесткие, чем условия строгой стационарности.

на самом деле может быть последовательностью нескольких стационарных процессов с разными значениями параметров. Таким образом, в контексте анализа слабой стационарности структурные сдвиги могут исказить качественные выводы.

Таблица 1. Оценки моделей по всей выборке дневных доходностей S&P500 1970–2020 гг.

Параметры	Модель и распределение					
	GARCH		GJR-GARCH		EGARCH	
	Нормальное	СОН	Нормальное	СОН	Нормальное	СОН
$\hat{\mu}$	0.0005 (0.0001)	0.0004 (0.0001)	0.0003 (0.0001)	0.0003 (0.0001)	0.0003 (0.0001)	0.0002 (0.0001)
$\hat{\omega}$	$1.59 \cdot 10^{-6}$ ($4.97 \cdot 10^{-6}$)	$1.11 \cdot 10^{-6}$ ($2.26 \cdot 10^{-6}$)	$1.83 \cdot 10^{-6}$ ($4.06 \cdot 10^{-6}$)	$1.47 \cdot 10^{-6}$ ($1.57 \cdot 10^{-6}$)	-0.2091 (0.0030)	-0.1759 (0.0147)
$\hat{\alpha}$	0.0910 (0.0581)	0.0805 (0.0413)	0.0218 (0.0067)	0.0218 (0.0866)	-0.0968 (0.0090)	-0.0913 (0.0064)
$\hat{\beta}$	0.8949 (0.0656)	0.9097 (0.0419)	0.8994 (0.0340)	0.9071 (0.2223)	0.9773 (0.0003)	0.9814 (0.0016)
$\hat{\gamma}$	—	—	0.1199 (0.0503)	0.1131 (0.2332)	0.1504 (0.0136)	0.1416 (0.0281)
$\hat{\sigma}$	—	0.9427 (0.0172)	—	0.9352 (0.0520)	—	0.9308 (0.0115)
$\hat{\eta}$	—	1.3630 (0.0557)	—	1.3992 (0.3490)	—	1.3998 (0.0477)
Среднее лог-правдоподобие	3.3059	3.3306	3.3172	3.3383	3.3181	3.3400
Средний AIC	-6.6111	-6.6603	-6.6336	-6.6755	-6.6354	-6.6789
b	—	—	0.5000	0.5232	—	—
$E c(z_{t-1})$	0.9859	0.9902	0.9812	0.9881	0.9773	0.9814

Примечание. Обозначения соответствуют параметрам процессов из описанных выше спецификаций 1–3. В скобках указаны робастные (QML) стандартные ошибки. s — параметр скошенности, η — параметр формы СОН-распределения (определены выше).

Проиллюстрируем это на простом примере. Пусть доходности порождаются стационарным GARCH(1,1)-процессом с нулевым математическим ожиданием, математическим ожиданием квадрата доходности 0.0001 до момента времени t^* и 0.0004 после t^* . Заметим, что скачок в значении безусловной дисперсии может произойти из-за изменения значения любого параметра условной дисперсии, например константы ω (безусловная дисперсия для такой модели будет равна $\omega/(1-\alpha-\beta)$). Если исследователь не знает о наличии сдвига и предполагает, что параметры процесса постоянны на всем промежутке, а потом оценивает GARCH-модель, то из-за неправильной спецификации могут получиться оценки, которые будут свидетельствовать о бесконечной безусловной дисперсии (IGARCH-модель). Тем не менее, настоящая безусловная дисперсия процесса будет находиться в интервале (0.0001; 0.0004) в зависимости от положения момента t^* , т. е. она не стремится к бесконечности при увеличении временного горизонта.

Теперь рассмотрим пример образования псевдонестационарного (или почти нестационарного) поведения GARCH(1,1)-моделей. Пусть данные для первой подвыборки длиной

500 наблюдений сгенерированы согласно GARCH(1,1)-модели (с нормальным распределением) с параметрами $(\mu, \omega, \alpha, \beta)$, равными $(0, 10^{-6}, 0.1, 0.8)$. Данные для второй подвыборки длиной 500 наблюдений сгенерированы аналогичной моделью с параметрами $(0, 5 \cdot 10^{-6}, 0.1, 0.8)$, т.е. изменение происходит только в константе ω . В обоих случаях модели стационарны, и если на двух данных подвыборках оценить модели по отдельности, то оценки получаются примерно равными истинным значениям. Однако если оценить GARCH(1,1)-модель (с нормальным распределением) на всей выборке ($t \in [1; 1000]$), то оценки получаются равными $(-0.00002, 1.9 \cdot 10^{-8}, 0.0577, 0.9462)$. При этом $\hat{\alpha} + \hat{\beta} = 1.004$. Это означает, что согласно данным результатам доходности можно считать нестационарными. Таким образом, из-за изменения параметров модели, порождающей данные, при оценивании на всей выборке модели, не учитывающей этого изменения, возникает псевдостационарность поведения доходностей.

Продемонстрируем эмпирическую релевантность проблемы псевдо- и почти нестационарного поведения GARCH-моделей, оцененных на всей выборке. Для этого в предположении СОИ-распределения оценим более общую модель GJR-GARCH(1,1) скользящим окном на тех же 50-ти годах дневных данных S&P500 (оцениваем достаточно общую модель, чтобы исключить риск получения нестабильных оценок из-за упрощенной спецификации⁷). В качестве ширины окна выберем 1, 3, 5 и 8 лет. Динамика оценок продемонстрирована на рис. 1. Графики отображают значение соответствующего параметра (или логарифма правдоподобия), оцененного по выборке с этой точкой в середине. Для экономии места оценки μ и параметров распределения s и η опущены.

Оценки моделей меняются во времени. Для полноты картины также проанализированы доверительные интервалы: почти во всех моделях существовали продолжительные периоды времени, где 95 и даже 99%-ные доверительные интервалы для параметров, основанные на оценках по непересекающимся (во времени) выборкам, также не пересекались. Так, в модели GJR-GARCH с СОИ-распределением точечная оценка параметра γ по трехлетней выборке

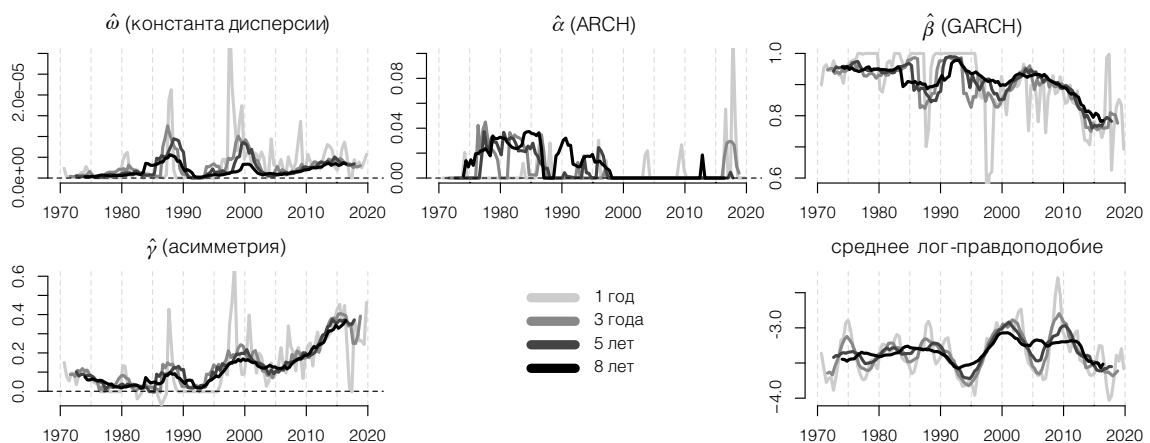


Рис. 1. Динамика оценок коэффициентов и значений функции правдоподобия на данных S&P500

⁷ Были оценены скользящим окном все шесть приведенных в табл. 1 моделей, однако качественно (с точки зрения нестабильности и флуктуации оценок) результаты были практически одинаковыми.

за 2003–2005 гг. равна 0.089 с 99%-ным доверительным интервалом [0.026; 0.152], а по выборке за 2015–2017 гг. — 0.399 с доверительным интервалом [0.198; 0.602]. Это означает, что 99%-ный доверительный интервал для коэффициента асимметрии для второго периода не включает ни одного значения из такого же доверительного интервала для первого периода. Следовательно, оценивать единую GJR-GARCH-модель с 2003 по 2017 г. с фиксированными коэффициентами было бы не вполне корректно.

Оценки, полученные на выборках длиной в 1 год (примерно 250 точек), более вариабельные, чем на выборках длиной 8 лет (2000 точек), что ожидаемо с точки зрения статистики. В симуляциях обнаружено, что гибридный алгоритм максимизации⁸ работает точно и стабильно, и проблем с отсутствием сходимости (из-за гипотетических участков целевой функции, похожих на плато) не возникает. Исходя из этого, можно заключить, что оценки, полученные на выборках длиной 8 лет, слишком сильно меняются во времени, и их поведение сложно объяснить выборочным шумом или проблемами с оптимизатором. Следовательно, естественно предположить нестабильность значений коэффициентов во времени. Динамика функции правдоподобия в данных моделях (правый нижний график на рис. 1), содержащая существенные колебания, также является свидетельством того, что в некоторые периоды поведение рынка хуже описывается стандартными моделями условной гетероскедастичности. Дополнительные расчеты 95%-ных доверительных интервалов для правдоподобия на непересекающихся выборках (на графиках не указаны) также свидетельствуют о том, что колебания функции правдоподобия являются не случайными, а систематическими, вызванными изменениями структуры процесса и отражающими изменяющееся качество подгонки модели, аппроксимирующей данные. Это говорит о том, что разные экономические механизмы сменяют друг друга, поэтому даже общая спецификация динамики шоков может существенно меняться.

3. Тесты на наличие структурных сдвигов

3.1. Общий контекст тестов на наличие структурных сдвигов в дисперсии доходностей

Сначала рассмотрим общий контекст анализируемых моделей. Пусть доходности задаются процессом условной гетероскедастичности со средним μ согласно формуле (1). Если $\mu = 0$ (что является частым и не сильно ограничивающим предположением), то $r_t^2 = \sigma_t^2 z_t^2$, $E_{t-1} r_t^2 = \sigma_t^2$, поэтому можно рассматривать r_t^2 как прокси-переменную для σ_t^2 . Также заметим, что если σ_t^2 не меняется во времени, то она эквивалентна безусловной дисперсии доходностей.

Структурными сдвигами называют изменения популяционных значений параметров модели из-за смены глубинных экономических механизмов. Существуют два основных типа

⁸ Для получения оценок GARCH-моделей были использованы статистический пакет *R* и библиотека *rugarch* (Ghalanos, 2020). Гибридный оптимизатор включает в себя последовательность процедур максимизации на основе алгоритмов *solnp* (метод расширенных множителей Лагранжа), *nlnmb* (PORT-процедура адаптивной нелинейной минимизации), затем опять *solnp* с множественным стартом из случайных значений и, наконец, *nloptr* (общий оптимизатор, по умолчанию использующий симплексный алгоритм SBPLX).

изменений: медленное и резкое. Остановимся на резком изменении (скачке) параметров условной дисперсии, которые в свою очередь вызывают резкие сдвиги в безусловной дисперсии — такие сдвиги легче найти, и для них разработана большая часть существующих процедур. Также отметим, что рассматриваемые процедуры являются *ex-post* тестами, т. е. они ищут сдвиги в уже имеющихся данных. В современной литературе преобладают два популярных подхода к нахождению сдвигов в динамике волатильности: с помощью накопленных сумм (Rapach et al., 2008; Inclán, Tiao, 1994) и наименьших квадратов (Andreou, Ghysels, 2002). Ни в одном из них не делается предположений о функциональной форме динамики условной дисперсии, поэтому не существует ограничения на их применение (например, не только к стандартным моделям GARCH(1,1)⁹). В этих тестах проверяется стабильность значений безусловного второго момента рассматриваемого ряда, т. е. определяются лишь факты изменения значения безусловной дисперсии на основе стабильности поведения условной дисперсии. В тестах на сдвиг неявно предполагается, что сдвиг в условной дисперсии может быть вызван только изменением значений параметров, а значит, они повлияют и на безусловную дисперсию.

Отметим, что в статье не рассматриваются нестационарные процессы, формирующие доходность, поскольку, во-первых, значение долгосрочной дисперсии доходностей и чувствительность цен опционов к начальным условиям говорят о конечности дисперсии доходностей; во-вторых, доходность есть разность логарифмов цен, которые, в свою очередь, имеют поведение, похожее на случайное блуждание, что свидетельствует о стационарности доходностей.

3.2. Метод на основе накопленных сумм

В данном разделе под нулевой гипотезой будет подразумеваться гипотеза об отсутствии структурного сдвига в безусловной дисперсии ряда (ряд стационарен). Метод накопленных сумм идейно похож на широко применяемый тест на единичный корень (Kwiatkowski et al., 1992): если динамика ряда инвариантна во времени, то центрированные и масштабированные частичные суммы квадратов значений ряда должны сходиться к броуновскому мосту. Рассмотрим модифицированный итеративный метод накопленных сумм, описанный в работе (Rapach et al., 2008). Он опирается на идею из работы (Kokoszka, Leipus, 1999) и является обобщением подхода, впервые описанного в (Inclán, Tiao, 1994). Если безусловная дисперсия процесса стабильна и не имеет сдвигов, то при наличии T наблюдений доля суммы квадратов доходностей до момента k (в полной сумме за весь период) будет приближенно равна k/T . Следовательно, резкое изменение параметров, влияющее на условную

⁹ Тем не менее, как будет показано далее на симулированных данных, спецификация процесса важна для конечных выборок. В тесте на основе накопленных сумм предполагается лишь конечность первых восьми моментов базовых *i.i.d.* шоков $z_t = \varepsilon_t/\sigma$, и существование стационарного решения GARCH-уравнений, из которых следует выполнение необходимых условий теоремы Донскера. А в МНК-процедуре предполагается выполнение условий теоремы Гаека–Реньи для процесса (Lavielle, Moulines, 2000).

Популяционный восьмой момент базовых шоков, используемых в наших симуляциях, Ez^8 , равен 105 для нормального распределения и (685.0, 917.6, 517.9) для COH-распределений с параметрами из табл. 2 (столбцы GARCH-COH, GJR-GARCH-COH, EGARCH-COH соответственно). Оценка абсолютной ошибки численного интегрирования менее 10^{-10} . Симулируемые GARCH-процессы стационарны.

и безусловную дисперсии, приведет к существенному отклонению доли этих частичных сумм от k/T . Таким образом, анализ поведения квадратов значений ряда позволит судить о стабильности безусловной дисперсии во времени. Исходя из этого, можно рассматривать так называемую АИТ-статистику (Adjusted Inclán–Tiao):

$$\text{AIT} = \sup_k \sqrt{C_T \hat{\sigma}^2 / \hat{S}} \cdot |D_k|, \quad (2)$$

где $D_k = C_k / C_T - k/T$, $C_k = \sum_{t=1}^k \tilde{r}_t^2$, $C_T = \sum_{t=1}^T \tilde{r}_t^2$, \tilde{r}_t — центрированные доходности, $\hat{\sigma}^2 = C_T / T$.

\hat{S} — НАС-оценка (Heteroskedasticity- and Autocorrelation-Consistent, состоятельная при гетероскедастичности и автокорреляции) безусловной дисперсии *квадратов* доходностей,

т. е. $\hat{S} = \sum_{j=-m}^m w(j, m) \hat{\Omega}_j$, где $\hat{\Omega}_j = \widehat{\text{Cov}}(\tilde{r}_t^2, \tilde{r}_{t-j}^2) = T^{-1} \sum_{t=j+1}^T (\tilde{r}_t^2 - \hat{\sigma}^2)(\tilde{r}_{t-j}^2 - \hat{\sigma}^2)$, $w(j, m)$ — ядер-

ные веса выборочных автоковариаций (популярные ядра w описаны в работе (Andrews, 1991)), m — ширина окна (параметр сглаживания; правила его выбора есть в работах (Andrews, 1991; Newey, West, 1994))¹⁰. Упрощенная версия данной статистики, $\text{IT} = \sup_k \sqrt{T/2} \cdot |D_k|$, приводится в более ранней работе (Inclán, Tiao, 1994), но на практике ее не следует использовать по той причине, что приращения процессов $\{\tilde{r}_t\}$ и $\{\tilde{r}_t^2\}$ являются зависимыми. При верной нулевой гипотезе о том, что безусловная дисперсия доходностей постоянна, статистики АИТ и ИТ ведут себя как $\sup_r |W^*(r)|$, где $W^*(r)$ — броуновский мост, а критические значения для тестирования гипотез можно получить с помощью численных симуляций.

Особенностью данной статистики является то, что она дает информацию о времени только одного, самого большого потенциального структурного сдвига в дисперсии доходностей. Это позволяет быстро рассчитать выражение под знаком модуля во все моменты времени, а затем найти момент времени, когда это значение достигает максимума на выборке. При наличии нескольких сдвигов требуется применять особый итеративный алгоритм (Iterated Cumulative Sum of Squares, ICSS) для выявления всех мест их расположения.

Детальное описание ICSS-алгоритма приводится в (Inclán, Tiao, 1994) и в онлайн-приложении¹¹.

3.3. Калибровка АИТ-теста на конечных выборках

Следует отметить, что АИТ-статистика, на которую опирается ICSS-алгоритм, обладает рядом недостатков, один из которых — зависимость распределения ее критических значений на конечных выборках от типа процесса и от используемой НАС-оценки дисперсии

¹⁰ Статистика АИТ из (Rapach et al., 2008) идентична статистике $\sup_k |U_N(k) / \hat{s}_{N,q}|$ из (Kokoszka, Leipus, 1999, формула 2.3); обозначение $\hat{s}_{N,q}$ соответствует $\sqrt{\hat{S}}$, а при расчете $U_N(k)$ в качестве случайной величины X_t (обозначение из оригинальной статьи (Kokoszka, Leipus, 1999)) берется \tilde{r}_t^2 .

¹¹ См. https://github.com/Fifis/WasThereEverAShift_OnlineAppendix/blob/main/online-appendix.pdf или https://github.com/Fifis/WasThereEverAShift_OnlineAppendix.

квадратов доходностей. Следовательно, использование асимптотических критических значений может привести к недо- или переоценке истинного числа сдвигов, а получение скорректированных значений для конечных выборок требует наличия некоторых знаний о характере истинного процесса. Ниже продемонстрируем чувствительность данного теста и, следовательно, всей процедуры к особенностям процесса даже на длинных выборках.

Предположим, что исследователь изучает временной ряд доходностей на предмет сдвигов и использует 50 лет дневных данных S&P500 для получения оценок различных моделей из семейства GARCH (табл. 1). Поскольку ему требуется аппроксимировать распределение при нулевой гипотезе, а оценки моделей на всей выборке могут свидетельствовать о ложной нестационарности (при которой АИТ-тест вообще неприменим ввиду бесконечности дисперсии), то у него есть возможность брать для симуляций значения параметров, похожие на оценки по коротким подвыборкам, порождающим стационарные или почти нестационарные процессы. Для выделения таких подвыборок требуются дополнительные знания о характере процесса, что порождает порочный круг в анализе. Рекомендуется выделять некоторый фрагмент данных, для которого есть априорная информация об отсутствии сдвигов (к примеру, 2–3 года дневных данных без обширных экономических кризисов). Во всех проведенных симуляциях почти нестационарный случай примерно соответствует значениям параметров из табл. 1, стационарный — значениям из табл. 2 (стационарность процесса гарантируется за счет того, что $E c(z_{t-1}) = 0.9$). Симуляции были проведены для всех возможных комбинаций следующих параметров (10000 Монте-Карло симуляций для каждой комбинации):

- процессы — GARCH(1,1), GJR-GARCH(1,1), EGARCH(1,1);
- распределения ошибок — нормальное и COH-распределение;
- значения параметров в уравнении условной дисперсии — наборы из таблиц 1 и 2;
- длина выборки — 500, 1000, 2000, 4000, 8000 наблюдений;
- методы получения HAC-оценки \hat{S} — все оптимальные правила из статей (Andrews, 1991; Newey, West, 1994) и (Lumley, Heagerty, 1999);
- доходности — центрированные и нецентрированные.

Таблица 2. Значения параметров в просимулированных стационарных моделях

Параметры	Модель и распределение					
	GARCH		GJR-GARCH		EGARCH	
	Нормальное	COH	Нормальное	COH	Нормальное	COH
α	0.2000	0.2000	0.0500	0.0500	-0.2518	0.9000
β	0.7000	0.7000	0.7000	0.6900	0.9000	0.9000
γ	—	—	0.3000	0.2700	0.1477	0.0836
s	—	0.8203	—	0.7500	—	0.7478
η	—	1.3564	—	1.3500	—	1.5152
b	—	—	0.5000	0.5948	—	—

Примечание. Обозначения соответствуют параметрам процессов из спецификаций 1–3. Параметры μ и ω в данном случае не имеют значения, поскольку производится центрирование; используются соответствующие $\hat{\mu}$ и $\hat{\omega}$ из табл. 1.

Критические значения АИТ-статистики рассчитывались как эмпирические квантили уровней 90, 95 и 99%. Таблица 3 содержит средние абсолютные отклонения (MAE) этих

Таблица 3. Средние абсолютные отклонения симуляционных критических значений АИТ-статистики от асимптотических

Ядро + правило	Уровень значимости			Ядро + правило	Уровень значимости		
	10%	5%	1%		10%	5%	1%
<i>Симуляции с коэффициентами из табл. 2 (стационарный случай)</i>							
Quadratic spectral (NW) \tilde{r}	0.258	0.276	0.309	Tukey–Hanning (A) \tilde{r}	0.846	0.928	1.082
Quadratic spectral (NW)	0.263	0.281	0.312	Tukey–Hanning (A)	0.851	0.932	1.086
Quadratic spectral (A) \tilde{r}	0.866	0.951	1.108	Truncated (A) \tilde{r}	0.994	1.100	1.305
Quadratic spectral (A)	0.871	0.956	1.114	Truncated (A)	0.998	1.104	1.308
Bartlett (NW) \tilde{r}	0.138	0.141	0.147	LH truncate \tilde{r}	0.058	0.067	0.093
Bartlett (NW)	0.142	0.145	0.149	LH truncate	0.054	0.061	0.085
Bartlett (A) \tilde{r}	0.938	1.034	1.228	LH smooth \tilde{r}	0.150	0.194	0.304
Bartlett (A)	0.944	1.042	1.234	LH smooth	0.146	0.188	0.299
Parzen (NW) \tilde{r}	0.153	0.159	0.163	Неробастная \tilde{r}	1.295	1.434	1.699
Parzen (NW)	0.158	0.164	0.167	Неробастная	1.294	1.434	1.698
Parzen (A) \tilde{r}	0.789	0.864	1.008				
Parzen (A)	0.794	0.870	1.014				
<i>Симуляции с коэффициентами из табл. 1 (почти нестационарный случай)</i>							
Quadratic spectral (NW) \tilde{r}	1.276	1.380	1.538	Tukey–Hanning (A) \tilde{r}	2.508	2.751	3.160
Quadratic spectral (NW)	1.276	1.382	1.539	Tukey–Hanning (A)	2.508	2.750	3.158
Quadratic spectral (A) \tilde{r}	2.545	2.795	3.213	Truncated (A) \tilde{r}	2.734	3.014	3.494
Quadratic spectral (A)	2.546	2.794	3.212	Truncated (A)	2.734	3.014	3.490
Bartlett (NW) \tilde{r}	0.823	0.875	0.943	LH truncate \tilde{r}	0.158	0.195	0.280
Bartlett (NW)	0.825	0.876	0.944	LH truncate	0.158	0.196	0.282
Bartlett (A) \tilde{r}	2.664	2.932	3.394	LH smooth \tilde{r}	0.110	0.147	0.294
Bartlett (A)	2.664	2.932	3.390	LH smooth	0.109	0.145	0.290
Parzen (NW) \tilde{r}	0.966	1.034	1.131	Неробастная \tilde{r}	3.160	3.503	4.127
Parzen (NW)	0.967	1.036	1.131	Неробастная	3.157	3.501	4.126
Parzen (A) \tilde{r}	2.409	2.641	3.027				
Parzen (A)	2.409	2.640	3.025				

Примечание. (NW) — правило выбора ширины окна из (Newey, West, 1994).

(A) — правило выбора ширины окна из (Andrews, 1991).

LH — ядерная оценки и правило выбора ширины окна из (Lumley, Heagerty, 1999).

Неробастная — оценка дисперсии без HAC-коррекции.

\tilde{r} — использование центрированных доходностей вместо оригинальных.

Ядра: Quadratic spectral — квадратическое спектральное, Bartlett — Бартлетта, Parzen — Парзена,

Tukey–Hanning — Тьюки–Хэннинга, Truncated — усеченное, LH truncate — усеченное по Lumley–Heagerty,

LH smooth — сглаженное по Lumley–Heagerty.

Асимптотические значения для уровней значимости 10, 5 и 1% составляют 1.224, 1.358 и 1.628 соответственно.

эмпирических значений от асимптотических критических значений, усредненные по всем типам процессов, распределениям ошибок и размерам выборки. Это дает возможность сравнить, как в целом распределение на конечных выборках отличается от асимптотического, и имеет ли смысл проводить симуляции при использовании этого теста на практике. В таблице 4 приводятся эмпирические квантили АИТ-статистики для стационарных и близких к нестационарным процессов для наиболее популярных или дающих стабильные результаты HAC-оценок \hat{S} . Таблицу 4 следует интерпретировать так же, как табл. 1 из (Inclán, Tiao, 1994, p. 914), т. е. сравнивать полученные на практике расчетные статистики с приведенными в таблице критическими значениями, соответствующими длине исследуемого ряда, типу оцениваемого процесса и используемой HAC-оценке, для тестирования гипотезы о наличии в ряде сдвига на уровне значимости 10, 5 или 1%.

Стационарность доходностей. Расчет АИТ-статистики предполагает конечность безусловной дисперсии рассматриваемого ряда. Ниже демонстрируется, что для рядов, у которых безусловная дисперсия очень велика ввиду высокой инерционной составляющей в условной дисперсии, требуется использовать иные критические значения, чем для рядов, в которых влияние предыдущих значений затухает быстрее.

Таблица 3 подтверждает, что хотя во всех симуляциях чисто формально ряды были стационарными, на самом деле поведение АИТ-статистики отличалось для двух наборов параметров очень сильно. Получилось, что большие расхождения с асимптотическими критическими значениями у почти нестационарных процессов связаны с «растяжением» распределения вправо. Из таблицы 4, которая содержит именно критические значения для теста, видно, что для почти нестационарных рядов при использовании ядра Бартлетта и оптимальной по Ньюи–Уэсту ширины окна требуется брать более консервативные критические значения (примерно в 1.5 раза большие) по сравнению со стационарными рядами. Напротив, при использовании HAC-оценки по Lumley–Heagerty критические значения для стационарных и почти нестационарных рядов весьма похожи, причем для последних они на 5–20% меньше, чем для первых. Наконец, убывание критических значений при увеличении размера выборки является результатом, полностью противоположным приведенному в (Inclán, Tiao, 1994, p. 914), где с ростом числа наблюдений критические значения возрастают.

Тип HAC-оценки. Следующая трудность состоит в том, что робастная оценка матрицы \hat{S} из формулы (2) зависит от ширины окна и от ядерной функции, причем различные авторы (Andrews, 1991; Lumley, Heagerty, 1999; Newey, West, 1994) дают оптимальные в разных смыслах правила выбора ширины окна для каждого вида ядра. Поэтому перед началом анализа следует проверить эмпирически, какое из описанных выше правил расчета ширины окна дает наиболее близкие к асимптотическим распределения АИТ-статистики на выборках разной длины.

Сравниваются HAC-оценки со следующими ядрами: квадратическое спектральное (quadratic spectral); Бартлетта (Bartlett); Парзена (Parzen); усеченное индикаторное (truncated); Тьюки–Хэннинга (Tukey–Hanning); а также WEAVE-оценки Lumley–Heagerty из работы (Lumley, Heagerty, 1999) с плавно затухающим (smooth) и усеченным (truncated) ядром. Для первых трех ядер сравниваются правила выбора оптимальной ширины из статей (Newey, West, 1994; Andrews, 1991), а для усеченного и Тьюки–Хэннинга — только из (Andrews, 1991), т. к. Newey, West (1994) не приводят правил для этих ядер. Кроме того, изучим поведение нескорректированной (неробастной) статистики \hat{S} , рассчитанной как обыкновенная выборочная дисперсия без непараметрической коррекции. Для всех методов оценки

Таблица 4. Эмпирические квантили АПТ-статистики

Тип процесса, распределение шоков	Длина ряда и уровень значимости								
	T = 500			T = 2000			T = 8000		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
<i>Стационарный случай:</i>									
<i>НАС-оценка Lumley–Heagerty с усеченным ядром</i>									
GARCH, нормальное	1.260	1.396	1.666	1.182	1.293	1.533	1.174	1.294	1.539
GJR-GARCH, нормальное	1.245	1.377	1.646	1.162	1.279	1.509	1.160	1.272	1.509
EGARCH, нормальное	1.362	1.505	1.823	1.253	1.391	1.625	1.235	1.368	1.614
GARCH, COH	1.273	1.406	1.692	1.165	1.291	1.507	1.157	1.280	1.525
GJR-GARCH, COH	1.213	1.346	1.582	1.138	1.249	1.470	1.130	1.248	1.484
EGARCH, COH	1.386	1.532	1.847	1.259	1.377	1.619	1.219	1.354	1.630
<i>Почти нестационарный случай:</i>									
<i>НАС-оценка Lumley–Heagerty с плавно затухающим (smooth) ядром</i>									
GARCH, нормальное	1.219	1.367	2.820	1.156	1.312	2.164	1.078	1.170	1.421
GJR-GARCH, нормальное	1.173	1.272	1.703	1.094	1.194	1.479	1.062	1.149	1.297
EGARCH, нормальное	1.157	1.248	1.520	1.094	1.184	1.382	1.072	1.166	1.307
GARCH, COH	1.192	1.289	1.643	1.135	1.252	1.784	1.066	1.154	1.359
GJR-GARCH, COH	1.160	1.252	1.576	1.098	1.192	1.414	1.034	1.116	1.280
EGARCH, COH	1.162	1.244	1.444	1.089	1.176	1.346	1.064	1.153	1.303
<i>Стационарный случай:</i>									
<i>НАС-оценка с ядром Бартлетта, оптимальная ширина окна по Ньюи–Уэсту</i>									
GARCH, нормальное	1.527	1.676	1.966	1.393	1.534	1.793	1.288	1.416	1.676
GJR-GARCH, нормальное	1.470	1.610	1.907	1.341	1.479	1.761	1.247	1.369	1.624
EGARCH, нормальное	1.505	1.652	1.955	1.406	1.556	1.826	1.309	1.454	1.721
GARCH, COH	1.471	1.613	1.908	1.346	1.465	1.736	1.245	1.380	1.624
GJR-GARCH, COH	1.427	1.573	1.855	1.306	1.425	1.697	1.212	1.339	1.586
EGARCH, COH	1.509	1.653	1.944	1.398	1.536	1.798	1.297	1.431	1.724
<i>Почти нестационарный случай:</i>									
<i>НАС-оценка с ядром Бартлетта, оптимальная ширина окна по Ньюи–Уэсту</i>									
GARCH, нормальное	2.241	2.417	2.679	2.241	2.433	2.781	1.900	2.104	2.471
GJR-GARCH, нормальное	2.111	2.282	2.549	2.019	2.212	2.576	1.735	1.898	2.220
EGARCH, нормальное	2.101	2.267	2.547	2.010	2.213	2.607	1.714	1.903	2.238
GARCH, COH	2.284	2.454	2.742	2.344	2.539	2.905	2.004	2.201	2.577
GJR-GARCH, COH	2.161	2.334	2.631	2.121	2.335	2.664	1.802	1.978	2.350
EGARCH, COH	2.142	2.319	2.602	2.109	2.303	2.669	1.795	1.980	2.362

ядерные веса меньше 10^{-3} обнулялись с целью ускорения расчетов (это играло самую большую роль для квадратического спектрального ядра). Если число членов суммы превышало $\max(500, \lceil T/5 \rceil)$, то расчет НАС-оценки ускорялся с помощью быстрого преобразования Фурье по алгоритму из работы (Heberle, Sattarhoff, 2017), что давало существенный прирост производительности для оценщиков Lumley–Heagerty ввиду медленного затухания автокорреляционной функции, рассчитанной по их методологии.

Из таблицы 3 видно, что для почти всех ядер и почти всех правил выбора ширины окна (кроме Lumley–Heagerty с плавно затухающим ядром) эмпирическое распределение критических значений для почти нестационарных рядов отклоняется от асимптотических значений намного больше, чем для стационарных. При отсутствии возможности провести большое число симуляций и в условиях неопределенности относительно порождающего процесса, надежнее всего опираться на асимптотические критические значения, если HAC-оценка \hat{S} получена с ядром Бартлетта или Парзена при оптимальной ширине окна по Ньюи–Уэсту или по методологии Lumley–Heagerty (кроме случаев стационарного ряда и тестирования на уровне 1%). Правило выбора ширины окна по Andrews (1991), являющееся, например, умолчанием в популярном пакете *sandwich* для языка *R*, применять не следует.

Тип модели. Нельзя обойти стороной вопрос спецификации. Валидность асимптотических критических значений при достаточно общих предположениях АИТ-теста не отменяет того факта, что на конечных выборках при наличии в уравнении дисперсии асимметричной составляющей или при асимметричном распределении ошибок критические значения выборочного распределения могут существенно отличаться от асимптотических.

Несмотря на то что функциональная форма дисперсии доходностей не играет роли в распределении асимптотических критических значений, на конечных выборках критические значения для рядов, порожденных процессами GARCH(1,1), GJR-GARCH(1,1) и EGARCH(1,1) с нормальным и COH-распределением ошибок, ощутимо различаются (см. табл. 4). При различных фиксированных размерах выборки и правилах выбора ширины окна соотношение критических значений в разных моделях меняется, причем выделить единую закономерность невозможно. Например, для рядов, порожденных стационарным GJR-GARCH(1,1)-процессом, при использовании популярной HAC-оценки с ядром Бартлетта по правилу Ньюи–Уэста критические значения будут систематически ниже порожденных стационарным GARCH(1,1)-процессом, а те, в свою очередь, ниже порожденных стационарным EGARCH(1,1)-процессом. У нестационарных же процессов самые высокие критические значения на конечных выборках имеют GARCH(1,1)-модели. Такое же соотношение критических значений между типами модели прослеживается и для HAC-оценок по методу Lumley–Heagerty.

Центрирование. В статьях (Rapach et al., 2008) и (Inclán, Tiao, 1994) в явном виде предполагается, что доходности имеют нулевое среднее, но в формулах среднее не вычитается. В реальности же средняя доходность слегка отличается от нуля, поэтому требуется проверить, велика ли разница между критическими значениями, полученных на симулированных данных без центрирования и с ним.

По результатам симуляций из табл. 3 видно, что использование центрированных доходностей для расчета частичных сумм квадратов не оказывает такого сильного влияния на разницу между симулированными и асимптотическими критическими значениями, как выбор ядра и расчет оптимальной ширины окна. Для стационарных рядов с центрированием средние абсолютные отклонения критических значений оказались на 1–5% меньше, чем без центрирования, кроме HAC-оценок Lumley–Heagerty, где центрирование слегка ухудшило результаты в терминах MAE. Для почти нестационарных рядов центрирование практического значения не имеет, т. к. различия возникают только на третьей значащей цифре.

На основании сказанного выше можно заключить, что исследователь с большей вероятностью получит неточные результаты, если будет применять ядро Бартлетта (используемое многими авторами), а истинной будет стационарная модель GARCH(1,1) или EGARCH(1,1)

с нормальным распределением ошибок или почти нестационарная модель GARCH(1,1) со скошенным распределением ошибок.

В Приложении дана таблица с эмпирическими размерами тестов для асимптотических критических значений по аналогии с табл. 4. Получилось, что для стационарных процессов HAC-оценки по Lumley–Neagerty и с ядром Бартлетта дают доли отклонений верной нулевой гипотезы, близкие к номинальному размеру теста, только для больших выборок (2000 и 8000 наблюдений). Для почти нестационарного случая только HAC-оценка Lumley–Neagerty для 8000 наблюдений дает относительно корректные эмпирические размеры тестов. Таким образом, это является дополнительным свидетельством в пользу того, что асимптотические критические значения не являются надежными для использования в прикладных исследованиях.

3.4. Метод на основе наименьших квадратов

Недостатки ICSS-алгоритма на конечных выборках, возникающие из-за последовательного выявления потенциальных моментов сдвига, можно попытаться исправить, выявив эти моменты с помощью одношагового метода (который не является формальным тестом). Описанная ниже процедура приводится по статье (Andreou, Ghysels, 2002). В ней отсутствует проблема подбора критических значений, сильно зависящих от описанных в предыдущем разделе особенностей процесса, поскольку минимизируемый критерий (сумма) в статистическом смысле робастнее экстремальных статистик в АПТ-тесте. Еще одной сильной стороной данного метода является отсутствие сглаживания и гиперпараметров, сильно меняющих результат и требующих оптимального выбора. Такой метод оценивания числа и времени сдвигов основывается на предположении, что рассматриваемый процесс $\{Y_t\}_{t=1}^T$ (в данном случае в качестве процесса берутся квадраты доходностей) является кусочно-стационарным вокруг среднего, однако во временном ряде B раз в моменты $\{k_i\}_{i=1}^B$ происходит дискретное изменение среднего:

$$Y_t = \mu_j + \varepsilon_t, \quad j = 0, \dots, B, \quad k_j + 1 \leq t \leq k_{j+1}, \quad k_0 = 0, \quad k_{B+1} = T.$$

В этом методе рассматривается статистика, представляющая собой сумму квадратов случайных величин, центрированных вокруг средних значений на каждом интервале:

$$LS(k_1, \dots, k_B, \mu_0, \dots, \mu_B) = \sum_{j=0}^B \sum_{t=k_j+1}^{k_{j+1}} (Y_t - \mu_j)^2. \quad (3)$$

Решением $(\{\hat{k}_i\}_{i=1}^B, \{\hat{\mu}_j\}_{j=0}^B)$ задачи наименьших квадратов является точка минимума (по $\{k_i\}_{i=1}^B$ и $\{\mu_j\}_{j=0}^B$) следующего модифицированного критерия:

$$Q(\{k_i\}_{i=1}^B, \{\mu_j\}_{j=0}^B) = \ln LS(\{k_i\}_{i=1}^B, \{\mu_j\}_{j=0}^B) + \text{штраф}(B, T),$$

где $\text{штраф}(\cdot, \cdot)$ обозначает штрафную функцию, зависящую от размера выборки, числа оцениваемых параметров и некоторых числовых констант. В работе (Liu et al., 1997) приводятся две штрафные функции:

$$\begin{aligned} \text{штраф}_1(B, T) &= -\ln(T - B) + (2B + 1) \frac{\ln T}{T}, \\ \text{штраф}_2(B, T) &= -\ln(T - 2B - 1) + a_0(2B + 1) \frac{(\ln T)^{2+a_1}}{T}, \end{aligned}$$

где a_0 и a_1 — калибровочные константы¹². Liu et al. (1997) рекомендуют значения $a_0 = 0.299$, $a_1 = 0.1$. Выражение штраф_1 является штрафным компонентом в классическом критерии Шварца (BIC), а второй — более консервативной модификацией для выбора более «коротких» моделей. Данная модификация необходима во избежание выбора излишне длинной модели. Критерий Шварца, основанный на предположении нормальности и примененный к не нормальным данным, приводит на коротких выборках к завышению оценки числа сдвигов (Liu et al., 1997), поэтому эти авторы рекомендуют использовать штраф_2 для получения оценки числа сдвигов, более близкой к реальной. Наконец, если исследователь подозревает, что число сдвигов может быть недооценено из-за их малой магнитуды или близости, то вместо критерия Шварца можно взять критерий Акаике, в котором $\text{штраф}_0 = -\ln T + 2(2B + 1)/T$. Таким образом, указанная выше процедура является не формальным статистическим тестом, а именно некоторой процедурой, которая позволяет оценить число сдвигов, а при заданном числе — даты сдвигов. При этом распределение оценок моментов сдвигов является «нестандартным» и опирается на броуновское движение, а аналитические выражения для статистических выводов даны в работе (Lavielle, Moulines, 2000).

Симуляции, проведенные в работе (Andreou, Ghysels, 2002), демонстрируют, что на финансовых данных одношаговая МНК-процедура для рядов с сильной инерцией в динамике условной дисперсии — а именно такая чаще всего наблюдается в реальных данных — дает меньше ложноположительных результатов, чем ICSS-алгоритм с АИТ-тестом на уровне значимости 5%. МНК-процедура (особенно с классическим критерием Шварца) немного лучше обнаруживала небольшие и монотонные сдвиги в процессе волатильности на коротких рядах, в то время как АИТ-тест обладал хорошей мощностью только для немонотонных (разнонаправленных) скачков большой амплитуды. Данная работа дополняет эти результаты, поскольку проводится две серии симуляций. Сначала рассматривается широкий набор спецификаций и находятся откалиброванные критические значения для АИТ-теста, а затем рассматривается набор из восьми разных порождающих данные процессов, отражающих особенности обоих тестов, связанные со спецификой множественных сдвигов.

На практике в МНК-критерий зачастую вводят дополнительные ограничения на минимальную длину периода между сдвигами (см., например, (Bai, Perron, 2003)). Чтобы избежать получения скученных оценок моментов сдвига (а структурные сдвиги не могут происходить слишком часто) разумно добавить в штрафную функцию слагаемое, равное нулю, если $\min_i (k_{i+1} - k_i - 1) < p$, $i = 0, \dots, B$ (p — некоторая заданная константа), и бесконечности в противном случае. Константа p выбирается на основе экономических предпосылок. Так, логично предположить, что структурные сдвиги происходят не чаще, чем раз в месяц ($p = 10$ торговых дней между моментами сдвига — это значение используется в настоящей работе).

¹² В работе (Liu et al., 1997) в формуле для исходного критерия Шварца присутствует опечатка (лишний множитель 2, который отсутствует, например, в (Yao, 1988)), здесь приведена правильная формула.

Процедуру минимизации суммы квадратов можно разложить на несколько этапов, т. к. задача осложняется тем, что число неизвестных параметров заранее не определено. Если априорно зафиксировать B (число сдвигов), то штрафной компонент в определении критерия Q не будет зависеть от k_i и μ_j , а потому его можно не учитывать при фиксированном B . Далее, если зафиксировать и моменты сдвигов $\{k_i\}_{i=1}^B$, то оптимальная оценка $\{\hat{\mu}_j\}_{j=0}^B$ тривиально получается как выборочное среднее Y_t на подвыборках (\bar{Y}_j) . Следовательно, можно снизить размерность вектора оптимизируемых параметров, фиксируя B и используя указанные

выше оценки $\hat{\mu}_j$; поэтому можно рассматривать статистику $LS^*(k_1, \dots, k_B) = \sum_{j=0}^B \sum_{t=k_j+1}^{k_{j+1}} (Y_t - \bar{Y}_j)^2$.

Наконец, поскольку при таком подходе неизвестными остаются только B и $\{k_i\}_{i=1}^B$, можно перебирать различные B , оценивать соответствующие им моменты сдвигов, а затем выбирать такое B , которое давало бы наименьшее значение критерия Q . На практике можно перебирать возможные значения B из некоторого разумного диапазона: например с помощью ICSS-алгоритма оценить вероятное число сдвигов B' и найти оптимальные значения моментов сдвигов для каждого $B \in [B'/3; 3B']$, а затем исследовать поведение «концентрированного» критерия Q как функции от B .

Следует заметить, что существуют расширения данного подхода (например, взвешенный МНК с заданными весами), которые делают LS-статистику, определенную в (3), эквивалентной монотонно преобразованной АПТ-статистике. Поэтому АПТ-тест и МНК имеют некоторые сходства, и при поиске одного, наиболее вероятного сдвига, скорее всего, дадут схожие результаты. Решение же задачи «найти ровно N сдвигов», часто встречающейся на практике, намного легче получить с помощью МНК-процедуры, чем ICSS-алгоритма.

4. Сравнение точности ICSS-алгоритма и МНК-процедуры на симулированных данных

Для сравнения точности методов обнаружения сдвигов было проведено 10000 численных Монте-Карло симуляций, в которых были получены оценки числа и времени сдвигов для рядов при различных особенностях процесса. Эти процессы выбирались так, чтобы, во-первых, помогать выявлять сильные и слабые стороны методов, а во-вторых, дополнять уже существующие результаты. Рассмотрим следующие варианты дизайна процессов, порождающих данные (во всех случаях предполагается процесс белого шума с нулевым средним μ).

1. $r_t \sim \mathcal{N}(\sigma = 0.01)$ — без сдвигов.
2. $\{r_t\}_{t=1}^{500} \sim \mathcal{N}(\sigma = 0.010)$, $\{r_t\}_{t=501}^{1000} \sim \mathcal{N}(\sigma = 0.011)$ — один слабый сдвиг (рост волатильности на 10%) посередине выборки.
3. $\{r_t\}_{t=1}^{500} \sim \mathcal{N}(\sigma = 0.01)$, $\{r_t\}_{t=501}^{1000} \sim \mathcal{N}(\sigma = 0.02)$ — один сильный сдвиг (рост волатильности на 100%) посередине выборки.
4. $\{r_t\}_{t=1}^{60} \sim \mathcal{N}(\sigma = 0.01)$, $\{r_t\}_{t=61}^{1000} \sim \mathcal{N}(\sigma = 0.02)$ — один сильный сдвиг в самом начале выборки.
5. $\{r_t\}_{t=200(i-1)+1}^{200i} \sim \mathcal{N}(\sigma = 0.01 \cdot i)$, $i = 1, \dots, 11$ — сдвиг через каждые 200 точек и стандартное отклонение увеличивается на 0.01 (всего $B = 10$ монотонных сдвигов).

6. $\{r_t\}_{t=200(i-1)+1}^{200i} \sim \mathcal{N}(\sigma = 0.01 + 0.01 \cdot I(i - \text{четное}))$, $i = 1, \dots, 11$ — через каждые 200 точек стандартное отклонение меняет значение попеременно с 0.01 до 0.02 и наоборот (всего $B = 10$).
7. $\{r_t\}_{t=60(i-1)+1}^{60i} \sim \mathcal{N}(\sigma = 0.01 + 0.01 \cdot I(i - \text{четное}))$, $i = 1, \dots, 10$, $\{r_t\}_{t=601}^{1600} \sim \mathcal{N}(\sigma = 0.01)$ — через каждые 60 точек в начале выборки стандартное отклонение меняет значение попеременно с 0.01 до 0.02 и наоборот, а затем становится равным 0.01 (всего $B = 10$ сдвигов, размещенных в первых 30% ряда).
8. $\{r_t\}_{t=1}^{2300} \sim \mathcal{N}(\sigma \in \{0.01, 0.02\})$ — в моменты сдвига σ переключается с 0.01 на 0.02 или обратно. Моменты сдвигов: 500, 600, 700, 800, 900, 1400, 1500, 1600, 1700, 1800 (всего $B = 10$ сдвигов, сконцентрированных в двух кластерах).

Во всех симуляциях из содержательных соображений задавалось ограничение вида «два сдвига не могут быть расположены на расстоянии менее $p = 10$ точек друг от друга». Данные симуляции, в отличие от приведенных в разделе 3.3, проводились для более простых порождающих процессов, чтобы выявить иную проблематику — чувствительность (или ее отсутствие) тестовых процедур к слабым или частым сдвигам, а не к гиперпараметрам метода оценивания. Чтобы исключить влияние выбора НАС-оценки на результат, симулировался белый шум, для которого не требуется НАС-коррекция и для которого справедлива упрощенная версия АИТ-статистики (простая ИТ-статистика из раздела 3.2). Также использовались критические значения из работы (Inclán, Tiao, 1994) для некоррелированных приращений, чтобы выделить именно эффект частых или слабых сдвигов. Поскольку в МНК-процедуре оценивание моментов сдвига производится при условии их заданного числа, то при каждом фиксированном числе сдвигов поиск их дат производили оптимизатором Нелдера–Мида¹³. Если предполагаемых сдвигов меньше, чем на самом деле в дизайне, начальные значения оптимизатора по параметру дат брались случайным образом из набора истинных дат; если предполагается истинное количество сдвигов в дизайне, старт производился из истинных дат, а если предполагается больше сдвигов, чем на самом деле в дизайне, в качестве начальных значений брались все истинные даты, дополненные случайными датами, равномерно взятыми на всем временном промежутке. Подобным образом решалась проблема локальных оптимумов, благодаря чему приведенные в табл. 5 распределения оценок соответствуют настоящему выборочному распределению метода. В таблице приведено распределение долей симуляций, в которых ICSS-алгоритм с АИТ-тестом обнаружил на соответствующем уровне значимости указанное (в заголовке) число сдвигов.

¹³ В данных симуляциях не используется рекурсивный алгоритм глобальной оптимизации (Bai, Perron, 2003), имеющий сложность $O(T^2)$, где T — длина выборки, т. к. истинные даты сдвигов известны. Исходя из этого, для получения выборочного распределения оценщика в Монте-Карло симуляциях использовался более быстрый алгоритм оптимизации, стартующий из истинных значений. Он обладает двумя преимуществами: хорошо работает с негладкими функциями благодаря изменению конфигурации симплекса в многомерном пространстве; хорошо адаптирован к оптимизации с ограничениями в виде неравенств (сдвиги упорядочены по времени и удалены друг от друга как минимум на 10 точек). В работе (Lavielle, Moulines, 2000) показано, что даже если число сдвигов переоценено, то оценки дат будут сходиться к истинным значениям. На практике при одновременном оценивании моментов сдвига в некотором временном ряде использование алгоритма глобальной оптимизации (Bai, Perron, 2003) дает такое преимущество, как перебор всех возможных разбиений временного ряда на отрезки и выбор оптимального.

Таблица 5. Распределение (частотное) оценки числа сдвигов, полученной с помощью ICSS-алгоритма

Дизайн	Значимость, %	Число сдвигов				Среднее количество
		0	1	2	> 2	
1	10	0.912	0.073	0.013	0.002	0.104
	5	0.959	0.038	0.003	0.000	0.045
2	10	0.429	0.487	0.074	0.009	0.665
	5	0.557	0.411	0.029	0.002	0.476
3	10	0.000	0.830	0.135	0.035	1.211
	5	0.000	0.915	0.078	0.007	1.093
4	10	0.581	0.396	0.017	0.007	0.450
	5	0.757	0.238	0.005	0.001	0.250

		Число сдвигов										Среднее количество
		0	1–3	4–5	6–7	8	9	10	11	12	13–16	
5	10	0.000	0.003	0.114	0.656	0.158	0.052	0.013	0.003	0.001	0.000	6.769
	5	0.000	0.010	0.227	0.701	0.056	0.006	0.001	0.000	0.000	0.000	6.086
6	10	0.992	0.006	0.001	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.020
	5	0.999	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002
7	10	0.000	0.977	0.017	0.006	0.000	0.000	0.000	0.000	0.000	0.000	1.254
	5	0.000	0.995	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.095
8	10	0.000	0.585	0.066	0.111	0.068	0.033	0.075	0.041	0.014	0.007	4.562
	5	0.000	0.804	0.068	0.069	0.036	0.008	0.012	0.002	0.000	0.000	2.917

Примечание. Жирным шрифтом выделены значения, стоящие в столбце, соответствующем истинному числу сдвигов для данного дизайна. В последнем столбце указано среднее (по Монте-Карло симуляциям) количество сдвигов, обнаруженное данным методом.

Распределение оценок времени сдвига (по ICSS-алгоритму) представлено на рис. 2. Около каждого графика сверху в круглых скобках указан соответствующий номер дизайна. Для дизайнов 1–4 цвет и тип линии соответствуют типу оценки моментов сдвига в симуляциях: «=1» — был найден ровно один сдвиг, «>1» — было найдено более одного сдвига (для дизайна 3 детализированы три распределения по аналогичному принципу). Для дизайнов 5–8 приводятся плотности распределения оценок времени сдвига вне зависимости от их числа. В квадратных скобках указан уровень значимости, для которого проводился АПТ-тест. Вертикальными линиями обозначены истинные моменты сдвига.

Из таблицы 5 и рисунка 2 можно сделать ряд заключений. Во-первых, если сдвигов в ряду нет, то при тестировании на «популярных» уровнях значимости ($\alpha = 5\%$ или 10%) нулевая гипотеза об отсутствии сдвигов не будет отвергаться в близком к $(100 - \alpha)\%$ случаев. Интерпретация процента отвержений по итогам работы ICSS-алгоритма как размера теста несколько затруднена, т. к. АПТ-тест сам по себе является частью этого алгоритма, и нулевая гипотеза в итерационной процедуре может отвергаться один, два или более раз. Во-вторых, если сдвиг слабый (рост волатильности на 10%), то приблизительно в половине случаев он не будет обнаружен, однако момент сдвига, если он найден, определяется верно. В-третьих, если сдвиг сильный, то он будет почти наверняка обнаружен, однако в 8–17% случаев алгоритм обнаружит более одного сдвига; тем не менее, основная масса оценок множественных моментов

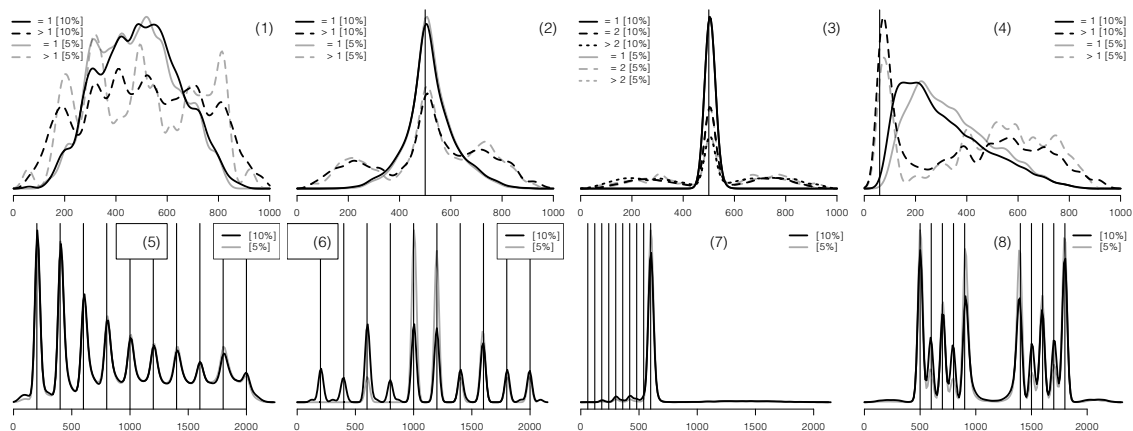


Рис. 2. Плотность распределения оценок времени сдвига, полученных с помощью ICSS-алгоритма в симуляциях

сдвигов будет приходиться на истинный момент. В-четвертых, если сдвиг находится с краю выборки (на расстоянии 6% от начала), то он, скорее всего, не будет обнаружен, а если и будет, то оценка момента сдвига может существенно запаздывать. В-пятых, при десяти равномерных монотонных сдвигах процедура на 30–40% недооценивает число сдвигов, однако моменты найденных сдвигов определяются достаточно точно. В-шестых, если сдвиги чередуются по направлению и сильные, то алгоритм работает очень неточно: в подавляющем большинстве симуляций не обнаруживается ни одного сдвига. В-седьмых, при концентрации немонотонных сдвигов в начале процесса алгоритм, опять же, может выявить менее 13% сдвигов; при этом моменты сдвигов определяются только к концу нестабильного периода. В-восьмых, при наличии двух кластеров алгоритм обнаруживает в среднем только 29–46% сдвигов. Таким образом, АПТ-тест склонен в целом недообнаруживать сдвиги, особенно в случае множественных сдвигов, однако даты сдвигов определяются более или менее точно.

В таблице 6 приведено распределение долей симуляций, в которых МНК-процедура обнаружила указанное (в заголовке) число сдвигов. Штрафные функции, определенные в разделе 3.4, обозначены как: МВІС — модифицированный критерий Шварца, ВІС — критерий Шварца, АІС — критерий Акаике.

Распределение МНК-оценок времени сдвига представлено на рис. 3. Около каждого графика сверху в круглых скобках указан соответствующий номер дизайна. Для дизайнов 1–4 цвет и тип линии соответствуют типу оценки моментов сдвига в симуляциях: «= 1» — найден ровно один сдвиг, «> 1» — найдено более одного сдвига. Черная линия отсутствует в случаях, где не было обнаружено ни одного сдвига при штрафе 2. Для дизайнов 5–8 приводятся плотности распределения оценок времени сдвига вне зависимости от их количества. В квадратных скобках указан тип штрафа (2 — модифицированный критерий Шварца, 1 — критерий Шварца). Вертикальной линией обозначены истинные моменты сдвига.

Теперь попробуем выявить закономерности нахождения сдвигов в тех же симуляциях с помощью МНК-процедуры на основе результатов из табл. 6 и рис. 3. Во-первых, при отсутствии сдвигов в данных при всех трех штрафных функциях МНК-процедура также свидетельствует об их отсутствии. Во-вторых, в дизайне с одним слабым сдвигом МНК склонен не находить его, однако при использовании критерия Акаике метод ошибался гораздо меньше, хотя сама точка сдвига определялась не совсем точно. В третьем дизайне, с одним

Таблица 6. Распределение (частотное) оценки числа сдвигов с помощью МНК-процедуры

Дизайн	Штраф	Число сдвигов				Среднее количество						
		0	1	2	> 2							
1	MBIC	1.000	0.000	0.000	0.000	0.000						
	BIC	0.997	0.002	0.001	0.000	0.004						
	AIC	0.765	0.072	0.059	0.103	0.583						
2	MBIC	1.000	0.000	0.000	0.000	0.000						
	BIC	0.917	0.077	0.004	0.002	0.090						
	AIC	0.323	0.427	0.108	0.142	1.177						
3	MBIC	0.000	0.999	0.001	0.000	1.002						
	BIC	0.000	0.959	0.026	0.015	1.065						
	AIC	0.000	0.579	0.141	0.280	1.946						
4	MBIC	1.000	0.000	0.000	0.000	0.000						
	BIC	0.136	0.859	0.003	0.002	0.871						
	AIC	0.000	0.760	0.091	0.149	1.509						
		Число сдвигов										Среднее количество
		0	1–3	4–5	6–7	8	9	10	11	12	13–16	
5	MBIC	0.000	0.998	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	1.735
	BIC	0.000	0.766	0.225	0.008	0.000	0.000	0.000	0.000	0.000	0.000	2.903
	AIC	0.000	0.115	0.377	0.288	0.052	0.017	0.005	0.028	0.038	0.081	6.345
6	MBIC	0.997	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.008
	BIC	0.000	0.000	0.000	0.000	0.001	0.001	0.982	0.014	0.001	0.001	10.017
	AIC	0.000	0.000	0.000	0.000	0.000	0.000	0.722	0.127	0.067	0.083	10.547
7	MBIC	0.009	0.786	0.158	0.039	0.006	0.002	0.001	0.000	0.000	0.000	2.412
	BIC	0.000	0.003	0.014	0.045	0.058	0.091	0.628	0.101	0.035	0.024	9.792
	AIC	0.000	0.000	0.001	0.004	0.006	0.016	0.329	0.231	0.168	0.246	11.373
8	MBIC	0.040	0.840	0.116	0.004	0.000	0.000	0.000	0.000	0.000	0.000	2.422
	BIC	0.000	0.001	0.018	0.034	0.041	0.060	0.818	0.022	0.003	0.003	9.677
	AIC	0.000	0.000	0.000	0.000	0.001	0.002	0.574	0.170	0.105	0.147	10.883

Примечание. Жирным шрифтом выделены значения, стоящие в столбце, соответствующем истинному числу сдвигов для данного дизайна. В последнем столбце указано среднее (по Монте-Карло симуляциям) количество сдвигов, обнаруженное данным методом.

сильным сдвигом, МНК-процедура находила его намного лучше; при этом критерий Акаике был наименее точным из всех (а сама дата сдвига находилась достаточно точно). При сильном сдвиге в начале выборки (четвертый дизайн) МНК с модифицированным критерием Шварца практически никогда не находил сдвига, однако версии с двумя другими критериями были гораздо ближе к истине (опять же, в случае определения сдвига его дата была оценена достаточно точно). Из малых монотонных сдвигов в пятом дизайне метод в среднем находит лишь 17–63%, и все три критерия, как следствие, давали заниженные оценки числа сдвигов. В шестом дизайне, когда дисперсия принимает два значения попеременно, метод с обычными критериями Шварца и Акаике находил в большинстве случаев корректное число сдвигов, однако модифицированный критерий Шварца не позволял обнаружить практически ни одного сдвига (в этом и последующих дизайнах даты сдвигов определялись точно).

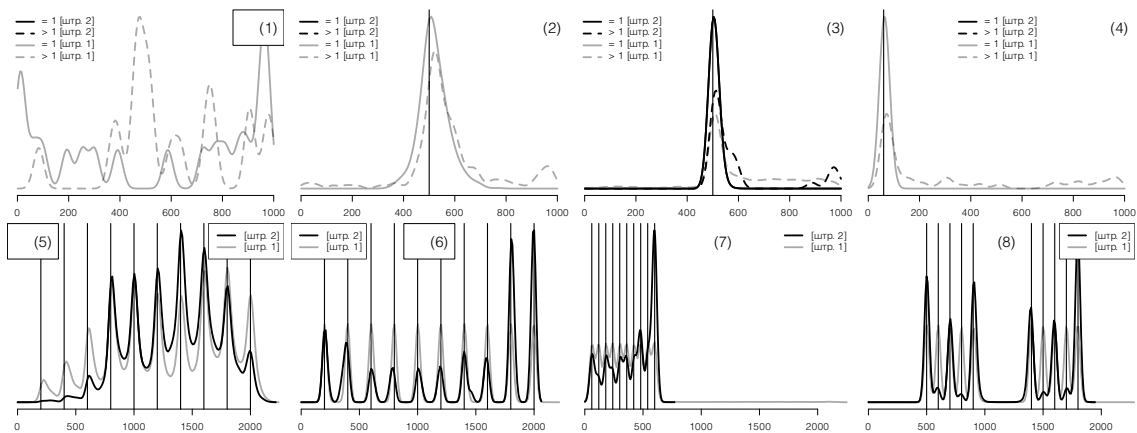


Рис. 3. Плотность распределения МНК-оценок времени сдвига в симуляциях

Со сдвигами, сосредоточенными в начале выборки, ситуация в целом аналогична: модифицированный критерий Шварца их почти не находит; обычный критерий Шварца дает очень точное, а критерий Акаике — завышенное в среднем на 5% число сдвигов. С немонотонными сдвигами в двух кластерах также получается, что модифицированный критерий занижает число сдвигов, а два остальных более точны (особенно обычный критерий Шварца). Исходя из результатов, можно утверждать, что МНК со штрафом в виде обычного критерия Шварца в большинстве дизайнов достаточно точно оценивает число сдвигов и их даты.

Причина такого результата состоит в том, что использование модифицированного критерия Шварца является оптимальным с точки зрения сходимости оценок к истинным значениям для негауссовских и зависимых процессов. В данном случае истинный процесс был гауссовским некоррелированным, поэтому даже в условиях гетероскедастичности обычный критерий Шварца чаще выбирал истинное число сдвигов. В среднем в симуляциях с дизайнами 5–8 критерий Шварца как функция от B убывал до $B = 10$ и затем выходил на плато, в то время как минимум модифицированного критерия Шварца (по B) находился в точках 0, 1 или 2, что говорит о том, что недооценка числа сдвигов при использовании модифицированного критерия делает оригинальный критерий Шварца более предпочтительным.

Следует отметить, что в реальности может быть намного больше сценариев, чем рассмотрено в данном симуляционном эксперименте. Так, оба метода определения сдвига, расположенного близко к краю выборки, могут работать с большей или меньшей надежностью в зависимости от направления скачка волатильности и соотношения доли периодов низкой и высокой волатильности в имеющемся ряде. Если изменить симуляционный дизайн 4 таким образом, что скачок волатильности вверх будет располагаться не в начале, а в конце на том же расстоянии от края выборки, то он будет лучше определяться обоими методами. А именно, в дизайне 4 первые 6% точек до сдвига дают лишь 1.6% общей дисперсии ряда, а если сдвиг произойдет в конце, то последние 6% точек дадут 20.3% общей дисперсии ряда, поэтому отклонения накопленных сумм от ожидаемого значения будут более значительными. Таким образом, можно заключить, что описываемые в данной работе тесты чувствительны ко многим характеристикам ряда, включая форму распределения, направление, частоту и интенсивность сдвигов, поэтому ограничиваемся лишь наиболее типичными случаями, представляющими интерес для исследователей.

5. Тестирование реальных данных на структурные сдвиги

Для проверки тестов на сдвиги на реальных данных возьмем три набора дневных доходностей: индекс S&P500 с 3 января 1970 г., индекс Nikkei 225 (далее — Nikkei) с 5 января 1970 г. и индекс DAX с 4 января 1988 г. (все три по 27 июля 2020 г.). В данном разделе приведем только краткие результаты для экономии места.

Применим ICSS-алгоритм с АИТ-тестом к указанным выше рядам. Выберем метод НАС-оценивания \hat{S} в АИТ-статистике Ньюи–Уэста с ядром Бартлетта, а также более консервативного и стабильного метода Lumley–Heagerty. Мощность данного теста на практике недостаточно высока, что подтверждается симуляциями из работы (Andreou, Ghysels, 2002) и из предыдущего раздела, поэтому проведем анализ двумя способами. Сначала — с симуляционно рассчитанными критическими значениями для уровня значимости 10% из табл. 4 для используемого ядра и метода выбора ширины окна в предположении стационарности, а затем — с максимально либеральным¹⁴ критическим значением (чтобы чаще отвергалась гипотеза об отсутствии сдвигов) 1.034 (минимальное обнаруженное во всех Монте-Карло симуляциях для уровня значимости 10%). Делается это по той причине, что для исследователя чаще менее критично обнаружить лишние сдвиги, чем упустить существенные (ошибка второго рода важнее). Результаты данной процедуры приводятся в табл. 7. Здесь и далее даты обнаруженных сдвигов будут записываться в формате ISO 8601 (год-месяц-день).

Таблица 7. Оценки дат предполагаемых сдвигов, обнаруженных с помощью ICSS-алгоритма

НАС: Бартлетт + Ньюи–Уэст	S&P500	1997-10-15
	Nikkei	1975-01-30, 1990-02-20
	DAX	1988-07-26, 1989-10-13, 1991-08-21, 1997-07-10, 2002-06-13, 2003-06-16
НАС: Lumley–Heagerty + усечен.	S&P500	Не обнаружено
	Nikkei	1975-01-30, 1990-02-20
	DAX	Не обнаружено

Примечание. АИТ-тест в ICSS-алгоритме проводился с использованием симуляционных и либеральных критических значений. Симуляционные критические значения взяты из табл. 4 для GJR-GARCH(1,1) — модели с СОН-распределением ошибок. Либеральные критические значения зафиксированы на уровне 1.034 (минимальное для уровня значимости 10% из наблюдавшихся в симуляциях). Жирным начертанием выделены сдвиги, которые были обнаружены при обоих критических значениях, нежирным — обнаруженные только при либеральных.

Несмотря на большую длину периода (50 лет), ICSS-алгоритм обнаружил в ряде S&P500 только один, значимый на уровне 10% потенциальный момент сдвига — 1997-10-15 (только для НАС-ошибок с ядром в форме Бартлетта), причем снижение критического значения

¹⁴ По аналогии с консервативной статистикой, на основании которой может слишком редко отвергаться нулевая гипотеза, из-за чего эмпирический размер теста будет не более номинального, либеральной называется статистика, на основании которой может слишком часто отвергаться нулевая гипотеза, из-за чего эмпирический размер теста будет не менее номинального. Например, для классического t -теста в простой линейной регрессии более либеральной является t -статистика на основе робастных стандартных ошибок в форме Уайта HC_0 (из-за недооценивания дисперсии в знаменателе), а консервативной — t -статистика на основе завышенных стандартных ошибок в форме HC_3 Дэвидсона–МакКиннона.

до либерального не позволило выявить дополнительных сдвигов. У ряда Nikkei ICSS-алгоритм на шаге выявления потенциальных моментов сдвига находил 3–4 потенциальных сдвига, но затем оставалось только два момента сдвига, и их оценки были стабильными. У индекса DAX алгоритм нашел 4 или 6 сдвигов (в последнем случае четыре из них совпали с найденными при консервативных критических значениях) для HAC-оценок с ядром в форме Бартлетта. АИТ-тест с HAC-оценками по методу Lumley–Heagerty, дающему самые стабильные результаты в симуляциях при различных видах процессов, сдвигов для двух индексов не обнаружил, а оценки дат сдвигов для Nikkei совпали с указанными выше результатами. Таким образом, данная тестовая процедура на длинных рядах выявляла количество сдвигов, не соответствующее базовым экономическим представлениям о структуре рынка. Например, мировой экономический кризис 2008–2009 гг. и кризис, связанный с COVID-19, не были обнаружены ни в одном из рядов, поэтому можно заключить, что, скорее всего, имеет место ошибка второго рода и структурные сдвиги недообнаруживаются.

Далее попробуем применить метод наименьших квадратов, в котором все потенциальные сдвиги ищутся одновременно. Предположим, что истинное число сдвигов не превышает 25, и вычислим значение критерия Q из раздела 3.4 для каждого B от 0 до 25, а затем выберем B , соответствующее наименьшему значению Q . На практике довольно сложно минимизировать функцию, содержащую множество локальных минимумов (даже при $B = 11$ критерий Q как функция от момента сдвига k_1 содержит более 10 локальных минимумов). Поэтому следует провести предварительную диагностику, чтобы минимизировать вероятность нахождения локального оптимума, не являющегося глобальным, и при этом избежать необходимости использовать $O(T^2)$ -алгоритм, где число возможных операций может измеряться миллионами или даже миллиардами для длинных выборок или высокочастотных данных. Воспользуемся методом роя частиц (Zambrano-Bigiarini, Rojas, 2013) из пакета *hydroPSO* с количеством частиц порядка $\lceil 500\sqrt{B} \rceil$ на $\lceil 50\sqrt[3]{B} \rceil$ итерациях для выбора начального значения и оптимизатором Нелдера–Мида для детерминированной оптимизации из данного старта¹⁵. Путем отслеживания динамики поведения роя в процессе решения задачи было обнаружено, что данный метод дает неплохие результаты и демонстрирует устойчивую сходимость при $1 \leq B \leq 25$, даже если принудительно ранжировать значения $k_{(1)} < \dots < k_{(B)}$ после каждой итерации. А при наличии многих процессоров данный процесс может быть ускорен или робастифицирован посредством увеличения числа итераций и, во вторую очередь, увеличением численности роя или ограничением максимальной скорости¹⁶. В таблице 8 приведены результаты поиска сдвигов с помощью МНК-процедуры. Авторы использовали два

¹⁵ Можно использовать любой другой стохастический оптимизатор, который бы достигал эффекта «прочесывания» многомерного пространства. В этом плане метод дифференциальной эволюции несколько хуже, т. к. он комбинирует лучшие точки популяции, в то время как в данной задаче потенциальные моменты сдвига ранжированы, поэтому мутация набора точек почти не имеет смысла. Напротив, перемещение в некотором направлении с большей вероятностью сохранит порядок потенциальных моментов сдвига, особенно при ограниченной максимальной скорости частицы.

¹⁶ Авторы пробовали провести стохастическую процедуру с одной итерацией и одной очень большой популяцией частиц размером $200B^2$ на тех же данных. Число найденных сдвигов на рядах было схожее, функция $Q(B)$ выглядела аналогично, однако ее абсолютные значения были выше, что говорит о попадании оптимизатора в локальный оптимум. Следовательно, количество итераций важнее, чем численность роя. Например, процедура с начальной популяцией порядка 1000 частиц и 100 итерациями дает лучший результат, нежели с начальной популяцией 100000 частиц и одной итерацией.

критерия: обычный критерий Шварца, как показавший наилучшую точность в симуляциях, и модифицированный критерий Шварца, как рекомендованный в (Andreou, Ghysels, 2002).

Таблица 8. Оценки дат предполагаемых сдвигов, обнаруженных с помощью МНК-процедуры

Индекс	Штраф	
S&P500	Модифиц. Шварц	1983-11-08, 1987-10-13 , 1987-11-10 , 1991-07-19, 1992-01-15, 1996-03-07, 2000-03-02, 2001-11-19, 2008-09-26 , 2008-12-02 , 2019-03-19, 2020-02-21
	Шварц	1973-04-03, 1975-03-19, 1975-10-03, 1979-12-28, 1981-04-24, 1986-08-28, 1987-10-13 , 1987-11-10 , 1992-11-09, 1993-07-14, 2001-09-05, 2003-06-17, 2005-03-10, 2008-09-26 , 2008-12-02 , 2009-08-24, 2012-01-10, 2012-06-13, 2017-08-09, 2020-02-28 , 2020-03-30
Nikkei	Модифиц. Шварц	1990-02-20, 2008-10-07 , 2008-11-10
	Шварц	1976-11-24, 1979-10-05, 1983-02-17, 1984-09-14, 1985-06-06, 1987-10-16, 1987-11-17, 1989-09-05, 1990-09-25, 1991-01-17, 1993-01-05, 1993-03-31, 1998-07-30, 2000-11-30, 2001-07-31, 2003-08-19, 2004-12-24, 2005-11-25, 2006-08-22, 2008-10-07 , 2008-11-10 , 2009-06-16, 2012-10-01, 2013-06-05
DAX	Модифиц. Шварц	1998-04-23, 2002-07-01 , 2003-04-08, 2008-10-03 , 2008-12-08 , 2020-03-05
	Шварц	1993-02-08, 1995-10-18, 1997-01-20, 1998-08-07, 1999-01-22, 2002-06-25 , 2003-05-09, 2004-10-25, 2008-10-03 , 2008-12-08 , 2012-08-03, 2015-04-14, 2017-03-28, 2020-03-04 , 2020-04-01

Примечание. Жирным начертанием выделены даты, которые обнаруживались при обоих видах штрафа с разницей менее одной недели.

Из таблицы 8 можно увидеть, что число найденных точек сдвига больше по сравнению с найденным ICSS-алгоритмом для всех индексов. Более того, им легче дать содержательную интерпретацию. Стоит, однако, отметить, что нахождение точной даты события для этого метода затруднительно. Например, если проанализировать даты для индекса S&P500, то можно предположить, что дата 2008-09-26 соответствует банкротству Lehman Brothers, а дата 2020-02-28 — нижней точке кризиса, связанного с COVID-19 (2 марта S&P500 достиг локального минимума), поэтому результаты не являются абсолютно точными. Использование оригинального критерия Шварца вместо модифицированного, как и в симуляциях с множественными сдвигами, позволило выявить намного больше экономически интерпретируемых сдвигов, в том числе средней амплитуды. В целом все эти факты означают, что МНК-процедура дает более надежные результаты на реальных данных по сравнению с ICSS-алгоритмом.

Также отметим, что если не очень сильно менять выборку, например, удалить последние три недели наблюдений, то АПТ-тесты дадут те же результаты, однако стабильность результатов у МНК-процедуры с критерием Шварца будет ниже: сильно меняется оценка числа сдвигов. Так, согласно модифицированному и оригинальному критериям Шварца, на выборке без последних 15 точек было найдено сдвигов (в квадратных скобках — число сдвигов для урезанной выборки): у S&P500—12 [14] и 21 [14], у Nikkei — 3 [4] и 24 [23], у DAX — 6 [6] и 17 [9]. Изменения оценок временных точек, впрочем, были небольшими. Таким образом, хотя МНК в симуляциях более надежен, на практике он более чувствителен к изменениям выборки, а в целом его результаты выглядят более согласованными с общеэкономической логикой.

6. Заключение

В данной работе были проанализированы два популярных метода выявления структурных сдвигов в волатильности доходностей: ICSS-алгоритм с АИТ-тестом и МНК-процедура. Показано, что АИТ-тест чувствителен ко многим характеристикам процесса на конечных выборках. В сравнительных симуляциях МНК-процедура дала более надежные результаты при наличии множественных немонотонных сдвигов по сравнению с ICSS-алгоритмом, который недооценивал число разнонаправленных сдвигов. Количество монотонных сдвигов было недооценено обоими методами. При оценивании на реальных данных МНК также дал более согласованные с общеэкономической точки зрения число и даты сдвигов, однако результаты оказались чувствительны к длине выборки. Таким образом, исследователям рекомендуется пользоваться МНК-процедурой с повторением оценивания на выборках, отличающихся несколькими крайними точками.

Благодарности. Авторы выражают благодарность Дмитрию Борзых за полезные комментарии. Исследование осуществлено в рамках Программы фундаментальных исследований НИУ ВШЭ в 2020 году (ТЗ 13). Проект реализован в научно-учебной лаборатории макроструктурного моделирования экономики России НИУ ВШЭ. А. В. Костырка благодарит за финансовую поддержку Национальный фонд исследований Люксембурга (FNR) и выделенный им PRIDE-грант для докторской подготовки по направлению MINLAB.

Список литературы

- Скроботов А. А. (2020). Структурные сдвиги и тестирование на единичный корень. *Прикладная эконометрика*, 58, 96–141.
- Andreou E., Ghysels E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17 (5), 579–600.
- Andrews D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59 (3), 817–858.
- Bai J., Perron P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18 (1), 1–22.
- Bollerslev T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31 (3), 307–327.
- Fernández C., Steel M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93 (441), 359–371.
- Galeano P., Tsay R. S. (2010). Shifts in individual parameters of a GARCH model. *Journal of Financial Econometrics*, 8 (1), 122–153.
- Ghalanos A. (2020). Rugarch: Univariate GARCH models. [R-пакет версии 1.4-2]. <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>.
- Glosten L. R., Jagannathan R., Runkle D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48 (5), 1779–1801.
- Heberle J., Sattarhoff C. (2017). A fast algorithm for the computation of HAC covariance matrix estimators. *Econometrics*, 5 (1), 9.

Hillebrand E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, 129 (1–2), 121–138.

Inclán C., Tiao G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89 (427), 913–923.

Kokoszka P., Leipus R. (1999). Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal*, 39 (2), 182–195.

Kwiatkowski D., Phillips P. C. B., Schmidt P., Shin Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54 (1–3), 159–178.

Lamoureux C. G., Lastrapes W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, 45 (1), 221–229.

Lavielle M., Moulines E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis*, 21 (1), 33–59.

Ling S., McAleer M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106 (1), 109–117.

Liu J., Wu S., Zidek J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, 497–525.

Lumley T., Heagerty P. (1999). Weighted empirical adaptive variance estimators for correlated data regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61 (2), 459–477.

McAleer M., Hafner C. M. (2014). A one-line derivation of EGARCH. *Econometrics*, 2 (2), 92–97.

Mikosch T., Stărică C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, 86 (1), 378–390.

Nelson D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 347–370.

Newey W. K., West K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61 (4), 631–653.

Perron P., Yamamoto Y., Zhou J. (2020). Testing jointly for structural changes in the error variance and coefficients of a linear regression model. *Quantitative Economics*, 11 (3), 1019–1057.

Rapach D. E., Strauss J. K., Wohar M. E. (2008). Chapter 10: Forecasting stock return volatility in the presence of structural breaks. In: Rapach D. E., Wohar M. E. (Eds.) *Forecasting in the Presence of Structural Breaks and Model Uncertainty (Frontiers of Economics and Globalization, Vol. 3)*, 381–416.

Wu J., Xiao Z. (2018). Testing for changing volatility. *The Econometrics Journal*, 21 (2), 192–217.

Xu K. L. (2013a). Powerful tests for structural changes in volatility. *Journal of Econometrics*, 173 (1), 126–142.

Xu K. L. (2013b). Power monotonicity in detecting volatility levels change. *Economics Letters*, 121 (1), 64–69.

Yao Y.-C. (1988). Estimating the number of change-points via Schwarz' criterion. *Statistics and Probability Letters*, 6 (3), 181–189.

Zambrano-Bigiarini M., Rojas R. (2013). A model-independent particle swarm optimization software for model calibration. *Environmental Modelling and Software*, 43, 5–25

Поступила в редакцию 31.07.2020;
принята в печать 14.12.2020.

Приложение

Таблица III. Эмпирический размер AIT-теста при использовании асимптотических критических значений

Тип процесса, распределение шоков	Длина ряда и уровень значимости								
	T = 500			T = 2000			T = 8000		
	10%	5%	1%	10%	5%	1%	10%	5%	1%
<i>Стационарный случай:</i>									
<i>НАС-оценка Lumley–Heagerty с усеченным (truncated) ядром</i>									
GARCH, нормальное	0.120	0.061	0.013	0.077	0.034	0.004	0.076	0.034	0.005
GJR-GARCH, нормальное	0.109	0.055	0.011	0.069	0.031	0.004	0.068	0.029	0.003
EGARCH, нормальное	0.177	0.101	0.029	0.115	0.062	0.001	0.104	0.053	0.009
GARCH, COH	0.123	0.066	0.015	0.075	0.03	0.004	0.068	0.032	0.005
GJR-GARCH, COH	0.094	0.047	0.007	0.059	0.024	0.003	0.059	0.025	0.004
EGARCH, COH	0.196	0.112	0.032	0.118	0.057	0.010	0.098	0.049	0.010
<i>Почти нестационарный случай:</i>									
<i>НАС-оценка Lumley–Heagerty с плавно затухающим (smooth) ядром</i>									
GARCH, нормальное	0.307	0.197	0.062	0.206	0.118	0.031	0.134	0.070	0.015
GJR-GARCH, нормальное	0.264	0.162	0.046	0.171	0.094	0.020	0.111	0.053	0.010
EGARCH, нормальное	0.288	0.176	0.058	0.207	0.121	0.035	0.143	0.079	0.019
GARCH, COH	0.268	0.161	0.046	0.171	0.095	0.020	0.112	0.057	0.010
GJR-GARCH, COH	0.231	0.138	0.037	0.152	0.074	0.015	0.094	0.045	0.008
EGARCH, COH	0.288	0.180	0.057	0.201	0.120	0.029	0.137	0.074	0.018
<i>Стационарный случай:</i>									
<i>НАС-оценка с ядром Бартлетта, оптимальная ширина окна по Ньюи–Уэсту</i>									
GARCH, нормальное	0.096	0.053	0.027	0.071	0.041	0.020	0.034	0.013	0.004
GJR-GARCH, нормальное	0.070	0.030	0.012	0.041	0.018	0.006	0.024	0.006	0.001
EGARCH, нормальное	0.060	0.022	0.007	0.035	0.012	0.004	0.027	0.005	0.000
GARCH, COH	0.082	0.032	0.010	0.058	0.028	0.013	0.028	0.010	0.004
GJR-GARCH, COH	0.062	0.025	0.009	0.038	0.014	0.005	0.018	0.004	0.001
EGARCH, COH	0.060	0.018	0.004	0.032	0.009	0.002	0.025	0.005	0.000
<i>Почти нестационарный случай:</i>									
<i>НАС-оценка с ядром Бартлетта, оптимальная ширина окна по Ньюи–Уэсту</i>									
GARCH, нормальное	0.759	0.658	0.452	0.748	0.630	0.414	0.568	0.436	0.224
GJR-GARCH, нормальное	0.689	0.578	0.365	0.646	0.515	0.301	0.454	0.329	0.146
EGARCH, нормальное	0.694	0.584	0.369	0.628	0.504	0.288	0.426	0.305	0.135
GARCH, COH	0.773	0.677	0.479	0.794	0.686	0.474	0.635	0.511	0.296
GJR-GARCH, COH	0.720	0.609	0.399	0.700	0.580	0.360	0.518	0.383	0.180
EGARCH, COH	0.721	0.612	0.392	0.676	0.558	0.339	0.468	0.342	0.164

Примечание. Доля отвержений рассчитана при использовании асимптотических критических значений из (Inclán, Tiao, 1994) для независимых приращений на 10000 симуляций.

Kostyrka A., Malakhov D. Was there ever a shift: Empirical analysis of structural-shift tests for return volatility. *Applied Econometrics*, 2021, v. 61, pp. 110–139.

DOI: 10.22394/1993-7601-2021-61-110-139

Andreï Kostyrka

University of Luxembourg, Luxembourg;
andrei.kostyrka@uni.lu

Dmitry Malakhov

National Research University Higher School of Economics, Moscow, Russian Federation;
dmalakhov@hse.ru

Was there ever a shift: Empirical analysis of structural-shift tests for return volatility

In this article, two popular tests for structural breaks are considered for return volatilities: the ICSS algorithm employing the AIT test, and the least-squares (LS) estimator. We show that the AIT test is sensitive to many features of the time series, and the use of asymptotic critical values is not always justified. The LS method was found to detect breaks more accurately, especially if there are many, in comparative simulations. Real data analysis revealed that LS estimation yields results in better accordance with general economic intuition, although its results are somewhat sensitive to the sample length. In general, we recommend the LS estimator for practical purposes.

Keywords: structural breaks; volatility; GARCH.

JEL classification: C12; C22; C58.

References

- Skrobotov A. A. (2020). Survey on structural breaks and unit root tests. *Applied Econometrics*, 58, 96–141 (in Russian).
- Andreou E., Ghysels E. (2002). Detecting multiple breaks in financial market volatility dynamics. *Journal of Applied Econometrics*, 17 (5), 579–600.
- Andrews D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica*, 59 (3), 817–858.
- Bai J., Perron P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18 (1), 1–22.
- Bollerslev T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31 (3), 307–327.
- Fernández C., Steel M. F. J. (1998). On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, 93 (441), 359–371.
- Galeano P., Tsay R. S. (2010). Shifts in individual parameters of a GARCH model. *Journal of Financial Econometrics*, 8 (1), 122–153.
- Ghalanos A. (2020). Rugarch: Univariate GARCH models. [R-пакет версии 1.4-2]. <https://cran.r-project.org/web/packages/rugarch/rugarch.pdf>.
- Glosten L. R., Jagannathan R., Runkle D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, 48 (5), 1779–1801.

- Heberle J., Sattarhoff C. (2017). A fast algorithm for the computation of HAC covariance matrix estimators. *Econometrics*, 5 (1), 9.
- Hillebrand E. (2005). Neglecting parameter changes in GARCH models. *Journal of Econometrics*, 129 (1–2), 121–138.
- Inclán C., Tiao G. C. (1994). Use of cumulative sums of squares for retrospective detection of changes of variance. *Journal of the American Statistical Association*, 89 (427), 913–923.
- Kokoszka P., Leipus R. (1999). Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal*, 39 (2), 182–195.
- Kwiatkowski D., Phillips P. C. B., Schmidt P., Shin Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54 (1–3), 159–178.
- Lamoureux C. G., Lastrapes W. D. (1990). Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance*, 45 (1), 221–229.
- Lavielle M., Moulines E. (2000). Least-squares estimation of an unknown number of shifts in a time series. *Journal of Time Series Analysis*, 21 (1), 33–59.
- Ling S., McAleer M. (2002). Stationarity and the existence of moments of a family of GARCH processes. *Journal of Econometrics*, 106 (1), 109–117.
- Liu J., Wu S., Zidek J. V. (1997). On segmented multivariate regression. *Statistica Sinica*, 497–525.
- Lumley T., Heagerty P. (1999). Weighted empirical adaptive variance estimators for correlated data regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 61 (2), 459–477.
- McAleer M., Hafner C. M. (2014). A one-line derivation of EGARCH. *Econometrics*, 2 (2), 92–97.
- Mikosch T., Stărică C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, 86 (1), 378–390.
- Nelson D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 347–370.
- Newey W. K., West K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies*, 61 (4), 631–653.
- Perron P., Yamamoto Y., Zhou J. (2020). Testing jointly for structural changes in the error variance and coefficients of a linear regression model. *Quantitative Economics*, 11 (3), 1019–1057.
- Rapach D. E., Strauss J. K., Wohar M. E. (2008). Chapter 10: Forecasting stock return volatility in the presence of structural breaks. In: Rapach D. E., Wohar M. E. (Eds.) *Forecasting in the Presence of Structural Breaks and Model Uncertainty (Frontiers of Economics and Globalization, Vol. 3)*, 381–416.
- Wu J., Xiao Z. (2018). Testing for changing volatility. *The Econometrics Journal*, 21 (2), 192–217.
- Xu K. L. (2013a). Powerful tests for structural changes in volatility. *Journal of Econometrics*, 173 (1), 126–142.
- Xu K. L. (2013b). Power monotonicity in detecting volatility levels change. *Economics Letters*, 121 (1), 64–69.
- Yao Y.-C. (1988). Estimating the number of change-points via Schwarz' criterion. *Statistics and Probability Letters*, 6 (3), 181–189.
- Zambrano-Bigiarini M., Rojas R. (2013). A model-independent particle swarm optimization software for model calibration. *Environmental Modelling and Software*, 43, 5–25.

Received 31.07.2020; accepted 14.12.2020.