Структурные сдвиги в моделях коинтеграции

Антон Скроботов[a]
[a] РАНХиГС, СПБГУ

13 марта 2021

Structural breaks in cointegration models

Аннотация

В данном обзоре рассматриваются методы тестирования коинтеграции во временных рядах при наличии структурных сдвигов. Обсуждаются современные подходы, основанные как на одном уравнении, так и на нескольких уравнениях. Кроме этого приводятся различные методы оценивания датировки сдвигов и построение доверительных интервалов для полученных дат. В дополнение рассматриваются нелинейные модели коинтеграции с переключением режимов.

Ключевые слова: тестирование на коинтеграцию, тестирование ранга коинтеграции, структурные сдвиги, модель коррекции ошибок.

This review discusses methods of testing for a cointegration in a time series in the presence of structural breaks. The review covers a large number of recently developed testing methods based on both one equation and multiple equation frameworks. In addition, various methods for estimating the dating of break dates and constructing of their confidence intervals are presented. In addition, nonlinear cointegration methods with regime swithings are considered.

Keywords: testing for cointegration, testing for cointegration rank, structural breaks, error correction model.

JEL Codes: C12, C22

1 Введение

Исследование коинтеграции уже давно является важным элементом анализа данных для изучения долгосрочных соотношений между макроэкономическими переменными. Если временные ряды являются интегрированными первого порядка и существует их стационарная линейная комбинация, то говорят, что эти временные ряды являются коинтегрированными. Стационарная линейная комбинация временных рядов интерпретируется как долгосрочное положение равновесия в системе.

Исследование коинтеграции началось с важного вопроса о наличии ложной (spurious) связи между переменными. Существует известный факт, что при анализе зависимости одной нестационарной переменной от другой (или других), регрессионные оценки и соответствующие им t-статистики не будут иметь смысла – t-статистика будет расходиться к бесконечности, тем самым приводя к ложным выводам относительно наличия взаимосвязи, и кажущаяся значимость коэффициента может быть просто результатом случайной реализации данных. Мы можем говорить о действительной (долгосрочной) зависимости между нестационарными переменными только если они являются коинтегрированными, то есть если существует их линейная комбинация, которая является стационарным процессом. Если такая линейная комбинация существует, то оценки наименьших квадратов будут суперсостоятельными, то есть будут сходиться с более высокой скоростью к истинным коэффициентам, в отличие от случая со стационарными переменными.

Соответственно, стандартной методологией исследования временных рядов является тестирование на наличие единичного корня каждого из исследуемых временных рядов, анализ зависимости этих временных рядов и тестирование остатков на стационарность/единичный корень. Если есть свидетельство коинтеграции, то можно делать содержательные выводы об оцененных коэффициентах, а также строить модель коррекции ошибок. Можно пойти другим путем и с самого начала тестировать ранг коинтеграции (максимальное количество линейно независимых комбинаций временных рядов – коинтеграционных соотношений) и строить модель коррекции ошибок на основе полученного ранга коинтеграции. Последний подход имеет смысл, когда исследователь априорно не знает, сколько коинтеграционных соотношений может быть в модели.

Однако такая методология не исключает и некоторые проблемы, которые могут возникать при анализе экономических временных рядов. Одной из важных проблем является наличие структурных сдвигов, особенно если мы рассматриваем относительно длинные временные периоды. Длинные периоды нужны нам, среди прочего, из-за того, что коинтеграционное соотношение интерпретируется как некоторое долгосрочное положение равновесия между переменными. В этом случае мы должны учитывать, что экономическая структура может измениться в течение рассматриваемого периода.[1] Кроме этого, мы можем обнаружить структурные сдвиги для индивидуальных рядов (см. недавний обзор Скроботов (2020)).

В данном обзоре мы обсуждаем проблемы исследования коинтеграционных свойств временных периодов, если есть свидетельство о наличии структурных сдвигов. В разделе 2 мы описываем методы тестирования нулевой гипотезы о наличии/отсутствии коинтеграции, если в коинтегрирующей регрессии имеются структурные сдвиги, причем мы рассматриваем как структурные сдвиги в константе и/или наклоне тренда, так и структурные сдвиги в коинтегрирующем векторе (в долгосрочных коэффициентах). Мы также обсуждаем тесты на структурные сдвиги, тесты на определение количества сдвигов и методы построения доверительных множеств для дат сдвигов. В разделе 3 мы рассматриваем многомерные методы исследования коинтеграционных свойств временных рядов. Здесь многие методы основаны на теории Йохансена и определении ранга коинтеграции последовательным тестированием, учитывая наличие структурных сдвигов. Рассматриваются как известные даты структурного сдвига, так и методы, основанные на неизвестной дате структурного сдвига. Также обсуждается проблема изменения ранга коинтеграции в вы-

[1] Одной из первых работ по анализу структурной стабильности в коинтеграционных моделях была Hansen and Johansen (1999). В ней авторы анализируют рекурсивные коэффициенты и рекурсивные тесты модели.

борке, в частности, в конце выборки. В разделе 4 рассматриваются модели коинтеграции, в которых либо коинтеграционное соотношение является нелинейным, либо нелинейным является механизм коррекции (возвращения к долгосрочному равновесию).

2 Тестирование одного коинтеграционного соотношения при наличии структурных сдвигов

Тестирование коинтеграции при наличии структурных сдвигов было рассмотрено в работах (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) и (Arai and Kurozumi, 2007). Рассмотрим коинтегрирующую регрессию с одним уравнением с единственным структурным сдвигом. Пусть y_t – стохастический процесс, который порождается следующим образом:

$$y_t = f(t) + \beta' x_t + e_t,\ t = 1, \ldots, T, \tag{1}$$

где $f(t)$ – некоторая детерминированная функция от времени, отвечающая за структурный сдвиг, x_t – вектор $I(1)$ регрессоров размерности $(p \times 1)$, и

$$\begin{aligned} e_t &= \gamma_t + \nu_{1t} && (2)\\ x_t &= x_{t-1} + \nu_{2t} && (3)\\ \gamma_t &= \gamma_{t-1} + u_t, \gamma_0 = 0, && (4) \end{aligned}$$

где $u_t \sim i.i.d.(0, \sigma_u^2)$. В (Arai and Kurozumi, 2007) предполагается, что процесс $\nu_t = (\nu_{1t}, \nu_{2t}')'$ имеет нулевое среднее и обладает условием сильного перемешивания с коэффициентами $-p\beta/(p-\beta)$ и $E|\nu_t|^p < \infty$ для некоторых $p > \beta > 5/2$, и y_0 – случайный вектор с $E|y_0| < \infty$. Наша цель состоит в том, чтобы тестировать гипотезу о наличии коинтеграции и эффективно оценить коинтегрирующий вектор β.

Определим дамми-переменные $DU_t = I(t > T_1)$ и $DT_t = (t - T_1)I(t > T_1)$, где $I(\cdot)$ – индикатор-функция, $T_1 = \lambda_1 T$ - дата структурного сдвига, λ_1 - доля даты структурного сдвига, $0 < \lambda_1 < 1$, определяющая его местоположение в выборке (эта величина непрерывна, в отличие от дискретной даты сдвига). Рассмотрим несколько типов моделей, которые специфицированы через конкретный вид детерминированной функции $f(t)$:

Модель 0а. Сдвиг в уровнях без наличия тренда

$$f(t) = \mu_1 + \mu_2 DU_t,\ t = 1, \ldots, T. \tag{5}$$

Модель 0. Сдвиг в уровнях с наличием тренда без изменения его наклона

$$f(t) = \mu_1 + \mu_2 DU_t + \theta_1 t,\ t = 1, \ldots, T. \tag{6}$$

Модель I. Изменение наклона тренда без сдвига в уровнях

$$f(t) = \mu_1 + \theta_1 t + \theta_2 DT_t,\ t = 1, \ldots, T. \tag{7}$$

Модель II. Сдвиг в уровнях с изменением наклона тренда

$$f(t) = \mu_1 + \mu_2 DU_t + \theta_1 t + \theta_2 DT_t,\ t = 1, \ldots, T. \tag{8}$$

Также в некоторых практических ситуациях имеет смысл моделировать коинтеграционное соотношение, которое может само “сдвинуться” от своей долгосрочной траектории в определённый момент времени. Исходя из этого, целесообразно было бы исследовать иной тип структурных сдвигов – сдвигов в коэффициентах β, то есть рассматривать модель вида:

$$y_t = f(t) + \beta_1' x_t + \beta_2' x_t DU_t + e_t,\ t = 1, \dots, T. \tag{9}$$

В этой модели можно также рассмотреть две спецификации детерминированной компоненты.

Модель IIIa. Сдвиг в β при наличии сдвига в уровнях без наличия тренда

$$f(t) = \mu_1 + \mu_2 DU_t,\ t = 1, \dots, T. \tag{10}$$

Модель IIIb. Сдвиг в β при наличии сдвига в уровнях и наклоне тренда

$$f(t) = \mu_1 + \mu_2 DU_t + \theta_1 t + \theta_2 DT_t,\ t = 1, \dots, T. \tag{11}$$

В работе (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) были рассмотрены все указанные модели, в то время как в (Arai and Kurozumi, 2007) были рассмотрены только модели 0a, 0 и IIIa. Нулевая гипотеза о наличии коинтеграции со структурными сдвигами соответствует тому, что e_t – стационарный ряд, т.е. $\sigma_u^2 = 0$ в уравнении (4).

Если дата сдвига известна, то регрессии (1) и (9) оцениваются OLS, а затем полученные остатки тестируются на стационарность аналогом теста множителей Лагранжа (LM-тест, KPSS-тест):

$$V(\lambda_1) = T^{-2} \sum_{t=1}^{T} S_t^2 / \hat{\omega}_1^2 \tag{12}$$

где $S_t = \sum_{j=1}^{t} \hat{e}_j$, $\hat{e}_t$ – OLS-остатки в регрессии (1) или (9), а $\hat{\omega}_e^2$ – любая состоятельная оценка долгосрочной дисперсии ошибок e_t.[2] Остатки $\hat{e}_t$ от OLS-оценивания модели получены исходя из того, что дата структурного сдвига известна. Предельное распределение тестовой статистики (12) при нулевой гипотезе зависит как от датировки структурного сдвига, так и от количества элементов ряда x_t (то есть числа $I(1)$ регрессоров). При альтернативной гипотезе статистика расходится к бесконечности со скоростью $O_p(T/l)$, где l - ширина окна, используемая для оценивания долгосрочной дисперсии. Соответствующие критические значения приведены в (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) и (Arai and Kurozumi, 2007).

При коррелированности регрессора и ошибки, x_t и e_t, рекомендуется, например, дополнить регрессию (1) или (9) опережающими и запаздывающими разностями (“leads and lags”, см. (Saikkonen, 1991) и (Stock and Watson, 1993)). Этот подход также часто называют

[2] Чтобы эта оценка была состоятельной, рекомендуется использовать подход к выбору ширины окна, предложенный в (Kurozumi, 2002). Этот подход заключается в том, чтобы использовать ядро Бартлетта в оценке долгосрочной дисперсии, оцененной на нулевой частоте с параметром ширины окна l , выбранном автоматически следующим образом: $l = \min\left(1.1447\left\{\frac{4\hat{\rho}^2 T}{(1+\hat{\rho})^2(1-\hat{\rho})^2}\right\}^{1/3}, 1.1447\left\{\frac{4k^2 T}{(1+k)^2(1-k)^2}\right\}^{1/3}\right)$, где $\hat{\rho}$ – оценка AR(1)-коэффициента для остатков e_t, а k выбирается равным 0.8, как предложено в (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006).

динамическим методом наименьших квадратов, DOLS (Dynamic OLS), и рассматриваемые выше модели (1) и (9) переписываются следующим образом:

$$y_t = f(t) + \beta' x_t + \sum_{i=-K_L}^{K_U} \pi_i' \Delta x_{t-i} + \varepsilon_t \qquad (13)$$

$$y_t = f(t) + \beta_1' x_t + \beta_2' x_t DU_t + \sum_{i=-K_L}^{K_U} \pi_i' \Delta x_{t-i} + \varepsilon_t \qquad (14)$$

где K_L – число запаздывающих разностей, K_U – число опережающих разностей[3]. Можно показать, что выполняется свойство строгой экзогенности вида

$$E(v_{2t}\varepsilon_{t+k}) = 0 \text{ для } k = 0, \pm 1, \pm 2, \dots \qquad (15)$$

Параметры усечения K_U и K_L должны быть достаточно большим, чтобы эффект усечения стал незначительным, так как формально условие (15) выполняется при $K_L = K_U = \infty$, но не должны быть слишком большими, так как это увеличит неэффективность оценивания β. Важно ометить, что мы не включаем запаздывающие и опережающие значения переменной $\Delta x_t DU_t$, так как требование для получения асимптотической эффективности заключается в экзогенности ε_t по отношению к регрессорам. Однако заметим, что в (15) это требование выполняется, то есть ε_t строго экзогенен не только по отношению к Δx_{t-i}, но и по отношению к $\Delta x_{t-i} DU_t$.

Приведённый выше анализ строится на предположении, что даты структурных сдвигов известны априорно. Если дата сдвига неизвестна, в (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) и (Arai and Kurozumi, 2007) предлагается ее оценить на основе минимизации сумм квадратов остатков в уравнениях (1) и (9) по всем возможным датам сдвигов. Другими словами, оценка $\hat{\lambda}_1$ доли даты сдвига λ_1 определяются следующим образом:

$$\hat{\lambda}_1 = \arg\min_{\lambda_1 \in \Lambda} SSR(\lambda_1) \qquad (16)$$

где $SSR(\lambda_1)$ – сумма квадратов остатков в уравнении (1) или (9), Λ – замкнутое подмножество интервала $(0, 1)$. В (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) предлагается определять как $\Lambda = [2/T, (T-1)/T]$ для минимизации потери информации. В (Arai and Kurozumi, 2007) авторами в симуляциях рассматриваются два возможных варианта: $\Lambda = [0.05, 0.95]$ и $\Lambda = [0.15, 0.85]$. Свойства на малых выборках практически не отличаются для этих двух вариантов. Полученная оценка доли даты сдвига является T-состоятельной оценкой доли даты сдвига ($\hat{\lambda}_1 = \lambda_1 + O_p(1/T)$), из чего следует, что статистика (12) имеет то же самое распределение с оцененной датой сдвига, как если бы дата сдвига была известна. Также в регрессиях (13) и (14) мы можем тестировать гипотезы о коэффициентах,

[3]В (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006) и (Arai and Kurozumi, 2007) был рассмотрен частный случай $K_L = K_U$. В (Choi and Kurozumi, 2012) рекомендуется использовать различное число опережающих и запаздывающих разностей, выбираемое на основе информационных критериев, таких как, например, BIC, $BIC = n \ln\left(\frac{\sum_{t=K_L+2}^{T-K_U} \hat{e}_t^2}{n}\right) + (k)\ln(n)$, где k – число параметров в модели ($p(K_u + K_l + 1)$ связанных с запаздывающими и опережающими разностями плюс p для модели (13) и $2p$ для модели (14), а также число параметров детерминированной компоненты), $n = T - K_U - K_U - 1$. Также в работе (Kejriwal and Perron, 2008a) показывается, что использование информационных критериев предпочтительнее последовательного исключения незначимых опережающих и запаздывающих разностей.

используя стандартную технику. Подходы, описанные выше, легко обобщаются на случай нескольких структуных сдвигов.

Как отмечается в Carrion-i Silvestre and Kim (2019), тесты Arai and Kurozumi (2007) и Carrion-i-Silvestre and Sansó-i-Rosselló (2006) требуют наличия структурных сдвигов при нулевой гипотезе, то есть контролируют размер только в том случае, если нет согласованных сдвигов (cobreaking). Можно решить эту проблему, используя большие процентили, полученные из двух предельных распределений, с согласованными сдвигами и без. Carrion-i Silvestre and Kim (2019) разрабатывают тест, который определяет, есть ли линейная комбинация переменных, при которых устраняется нестационарность в данных: или путем добавления сдвигов в коинтеграционное соотношение, или при наличии согласованных сдвигов в нескольких переменных. Также рассматривается вариант возможного согласованного трендового поведения в нестационарных временных рядах. Carrion-i Silvestre and Kim (2019) обобщают подход Jansson (2005) для построения точечно-оптимального инвариантного теста (point optimal invariant, POI). Данный подход позволяет при помощи соответствуюющих ограничений (ограничение на согласованные сдвиги и ограничение на согласованные тренды) построить тесты на коинтеграцию.

Mogliani (2010) сравнивает поведение на конечных выборках нескольких тестов на коинтеграцию[4], анализируя их размер и мощность в моделях 0a, 0 и II. Хотя рассматриваемые им тесты в оригинале имеют только единственный структурный сдвиг, автор допускает появление многократных сдвигов. Если исследователь априори знает, что регрессоры строго экзогенны, то тогда рекомендуется использовать WE-тесты, которые имеют размер, близкий к номинальному, и высокую мощность во всех моделях и с любым количеством структурных сдвигов. Однако при нестрого экзогенных регрессорах такие тесты имеют серьёзные искажения размера и потерю мощности. В таких случаях рекомендуется использовать другие тесты. Симуляции показывают, что, когда модель правильно специфицирвоана при наличии единственного сдвига, во всех рассматриваемых моделях тесты CS_{DOLS} и CS_{DGLS}[5] работают лучше, а при наличие трёх сдвигов - CS_{DGLS}. Когда остатки неправильно специфицированы (некорректное число сдвигов), рекомендуется применять CS_{FM} и BLS_{SSR} при наличии одного сдвига, при наличие трёх сдвигов - CS_{FM} и BLS_{SSR} для моделей 0a, 0 и CS_{DOLS} и CS_{DGLS} для модели II. При пяти структурных сдвигах все тесты имеют большие искажения размера, и не рекомендуется их использовать при более чем трёх сдвигах в уровнях и в наклоне тренда и менее чем 200 наблюдениях.

Анализ мощности тестов показывает, что тесты CS и BLS имеют достаточно высокую мощность, в частности, наилучшим тестом является тест CS_{DGLS} в моделях 0a, 0 и CS_{DOLS} в модели II. Значительная потеря мощности происходит для теста WE, когда регрессоры не строго экзогенны.

Остаются важные вопросы определения количества структурных сдвигов в одном коинтеграционном соотношении, построения доверительных интервалов для дат сдвигов, а

[4]Были рассмотрены тесты WE из (Westerlund and Edgerton, 2007), CS_{DOLS} и CS_{DGLS} из (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006), CS_{FM} – из (Phillips and Hansen, 1990) и (Carrion-i-Silvestre and Sansó-i-Rosselló, 2006), BLS_{SSR} из (Bartley, Lee and Strazicich, 2001), основанный на доступной оценке канонической коинтегрирующей регрессии (feasible canonical cointegration regression estimator).

[5]Оценка CS_{DGLS} основана на регрессиях (13) и (14) с заменой x_t на $\tilde{x}_t = x_t'\hat{\phi}(L)$, где $\hat{\phi}(L)$ – лаговый полином остатков. Обычно $\hat{\phi}(L)$ строится, используя итеративную процедуру Кохрейна-Оркатта как AR(1) процесс. После этого проводится преобразование переменных, использующее полученные оценки для $\hat{\phi}(L)$.

также тестирования коэффициентов в модели. Эти проблемы были рассмотрены в двух взаимосвязанных работах (Kejriwal and Perron, 2008*b*), где анализируется коинтегрирующая регрессия с произвольным количеством структурных сдвигов, а также выводятся предельные распределения коэффициентов и дат структурных сдвигов, и (Kejriwal and Perron, 2010), где рассматриваются проблемы тестирования наличия структурных сдвигов и предлагаются различные тесты (в том числе и последовательная процедура на определение количества сдвигов). Хотя в этих работах допускается присутствие $I(0)$-регрессоров в модели, для простоты мы ограничимся описанием модели только с $I(1)$-регрессорами.

Пусть имеется следующая модель линейной регрессии с m структурными сдвигами ($m+1$ режимами):

$$y_t = c_j + z'_{ft}\delta_f + z'_{bt}\delta_{bj} + u_t,\ t = T_{j-1}+1,\dots,T_j \tag{17}$$

для $j = 1,\dots,m+1$, где $T_0 = 0$, $T_{m+1} = T$, $T_1,\dots,T_m$ – даты структурных сдвигов, которые предполагаюстя неизвестными. В этой модели y_t - $I(1)$-переменная, z_{ft} - $(q_f \times 1)$ и z_{bt} - $(q_b \times 1)$ векторы $I(1)$-регрессоров, определенные как $z_{ft} = z_{f,t-1} + u^f_{zt}$ и $z_{bt} = z_{b,t-1} + u^b_{zt}$, где начальные значения z_{f0} и z_{b0} предполагаются для упрощения либо $O_p(1)$, либо фиксированными константами. Переменные, у которых изменяется коэффициент в новом режиме, обозначаются индексом b, а индексом f обозначаются переменные, коэффициенты при которых постоянны во всех режимах. Такая модель часто называется моделью с частичным структурным сдвигом, когда дополнительно присутствуют не изменяющиеся на всей выборке переменные. В модели также допускается наличие детерминированного тренда в каждом из режимов в $I(1)$-регрессорах в форме $\tilde{z}_{ft} = \theta_f t + z_{ft}$ и $\tilde{z}_{bt} = \theta_b t + z_{bt}$ с $q_b > 1$ и $\theta_b \neq 0$.

Оценки параметров можно получить на основе минимизации суммы квадратов остатков по всем возможным датам сдвигов. Для каждого m-разбиения $(T_1,...,T_m)$ такого, что $T_i - T_{i-1} \geq \varepsilon T$ для некоторого ε, связанные с ним OLS-оценки δ_f и $\gamma = (\delta'_{b1},...,\delta'_{b,m+1})$ можно получить, минимизируя[6]

$$SSR_T(T_1,\dots,T_m) = \sum_{i=1}^{m+1}\sum_{t=T_{i-1}+1}^{T_i}\left[y_t - c - z'_{ft}\delta_f - z'_{bt}\delta_{bj}\right]^2. \tag{18}$$

Пусть H_0^a обозначает ограничения вида $\{c_j = c, \delta_{bj} = \delta_b$ для всех $j = 1,\dots,m+1\}$. Kejriwal and Perron (2010) предлагают следующие комбинации нулевых и альтернативных гипотез для тестирования общей модели (17):

1. $H_0^a(1) = \{H_0^a, q_f = 0\}$ против $H_1^a(1) = \{q_f = 0\}(y_t = c_j + z'_{bt}\delta_{bj} + u_t)$ – чистый структурный сдвиг, константа изменяется между режимами;

2. $H_0^a(2) = \{H_0^a, q_b = 0\}$ против $H_1^a(2) = \{q_b = 0\}(y_t = c_j + z'_{ft}\delta_f + u_t)$ – частичный структурный сдвиг, только константа изменяется между режимами;

3. $H_0^a(3) = \{H_0^a, q_f = 0\}$ против $H_1^a(3) = \{c_j = c$ для всех $j = 1,\dots,m+1, q_f = 0\}$ $(y_t = c + z'_{bt}\delta_{bj} + u_t)$ – частичный структурный сдвиг, константа не изменяется между режимами;

[6]Если сдвигов больше двух, для ускорения вычисления можно испоьзовать алгоритм (Bai and Perron, 2003).

4. $H_0^a(4) = \{H_0^a\}$ против $H_1^a(4)$ ={нет ограничений}$(y_t = c_j + z'_{ft}\delta_f + z'_{bt}\delta_{bj} + u_t)$ – блочно-частичный структурный сдвиг, подмножество $I(1)$-коэффициентов и константа изменяются между режимами;

5. $H_0^a(5) = \{H_0^a\}$ против $H_1^a(5) = \{c_j = c$ для всех $j = 1, \ldots, m+1\}$ $(y_t = c + z'_{ft}\delta_f + z'_{bt}\delta_{bj} + u_t)$ – блочно-частичный структурный сдвиг, подмножество $I(1)$-коэффициентов изменяется между режимами, константа не изменяется между режимами;

Во всех случаях нулевой гипотезой является предположение о том, что модель не имеет структурных сдвигов, а отличия заключаются лишь в количестве фиксированных (не изменяющихся между режимами) регрессоров.

Для тестирования указанных выше гипотез авторы рассматривают три типа тестов[7], допускающих серийную коррелированность ошибок и эндогенные регрессоры. Первый тип тестов применяется, когда альтернативная гипотеза включает фиксированное количество сдвигов $m = k$. Для этого используется следующая версия статистики Вальда, робастная к серийно коррелированным ошибкам:

$$F_T(\lambda, k) = \sup_{\lambda \in \Lambda_{k,\varepsilon}} \frac{SSR_0 - SSR_k}{\hat{\sigma}^2}, \tag{19}$$

где SSR_0 – сумма квадратов остатков при нулевой гипотезе, SSR_k – сумма квадратов остатков при альтернативной гипотезе для k структурных сдвигов, $\lambda = (\lambda_1, \ldots, \lambda_k)$, $\lambda_i = T_i/T$, а $\hat{\sigma}^2$ – оценка долгосрочной дисперсии, робастная к серийной корреляции и гетероскедастичности (HAC). Так как доли дат сдвигов состоятельны даже в случае серийно коррелированных ошибок, оценки дат сдвигов можно как даты, при которых достигается супремум. Тогда можно получить робастную версию, вычисляя $F_T(\lambda, k)$ в оцененных датах сдвигов.

Проблема с тестом состоит в том, что с инерционными ошибками возникают очень сильные искажения размера. Причина заключается в том, что при оценивании долгосрочной дисперсии используются остатки при альтернативной гипотезе. Поэтому при вычислении оценки долгосрочной дисперсии следует использовать подход, предложенный в (Kejriwal, 2009), который заключается в следующем: пусть оценка параметра σ есть

$$\hat{\sigma}^2 = T^{-1} \sum_{t=1}^{T} \tilde{u}_t^2 + 2T^{-1} \sum_{j=1}^{T-1} w(j/\hat{h}) \sum_{t=j+1}^{T} \tilde{u}_t \tilde{u}_{t-j} \tag{20}$$

где $\tilde{u}_t$ – остатки, полученные при нулевой гипотезе. Ядро $w(\cdot)$ является квадратичным спектральным и оценка ширины окна находится, следуя подходу (Andrews, 1991), т.е. $\hat{h} = 1.3221\,(\hat{a}(2)T)^{1/5}$, где $\hat{a}(2) = 4\hat{\rho}^2/(1-\hat{\rho})^4$ и $\hat{\rho}$ – OLS-оценка регрессии $\hat{u}_t$ на $\hat{u}_{t-1}$, где $\hat{u}_t$ – остатки, полученные при альтернативной гипотезе. Этот метод, названный гибридным (hybrid method), позволяет как контролировать размер в малых выборках, так и устранять проблему немонотонной мощности.

Второй тест, двойной максимальный тест (double maximum test), применяется тогда, когда альтернативная гипотеза включает неизвестное число сдвигов между 1 и некоторой

[7]Отметим, что если регрессоры не строго экзогенны, регрессия (17) просто дополняется опережающими и запаздывающими разностями регрессоров, сохраняя все асимптотические выводы неизменными.

верхней границей M. Основная причина введения таких тестов состоит в том, что если ряд содержит более одного структурного сдвига, то тест на единственный структурный сдвиг может иметь низкую мощность. Кроме того, тест для определённого (заданного) числа структурных изменений может иметь немонотонную мощность, когда действительное их количество больше, чем специфицировано в тесте. Соответственно, двойной максимальный тест рекомендуется применять в начале процедуры тестирования, чтобы обнаружить, что ряд действительно имеет структурные сдвиги, диапазон для количества которых изначально задаётся. Двойной максимальный тест определяется как

$$UD\max F_T(M) = \max_{1\le m\le M} F_T(\hat{\lambda}, m), \tag{21}$$

то есть как максимум статистик Вальда (20) для заданого k по всем k.

Третья тестовая процедура предназначена для последовательного тестирования на наличие дополнительного структурного сдвига в модели. Рассмотрим модель с k сдвигами с оценками дат сдвигов $(\hat{T}_1, \dots, \hat{T}_k)$. Процедура состоит в тестировании нулевой гипотезы о том, что имеется k сдвигов, против альтернативной гипотезы о том, что имеется $k+1$ сдвигов, так что дополнительный сдвиг может находиться в одном из $k+1$ режимов. Тестовая статистика определяется как

$$SEQ_T(k+1|k) = \\ \max_{1\le m\le k+1} \sup_{\tau\in\Lambda_{j,\varepsilon}} \frac{SSR_T(\hat{T}_1, \dots, \hat{T}_k) - SSR_T(\hat{T}_1, \dots, \hat{T}_{j-1}, \tau, \hat{T}_j, \dots, \hat{T}_k)}{\hat{\sigma}^2_{k+1}}$$

где $\Lambda_{j,\varepsilon} = \{\tau : \hat{T}_{j-1} + (\hat{T}_j - \hat{T}_{j-1})\varepsilon \le \tau \le \hat{T}_j + (\hat{T}_j - \hat{T}_{j-1})\varepsilon\}$, а $\hat{\sigma}^2_{k+1}$ – состоятельная оценка долгосрочной дисперсии при нулевой гипотезе, в которой ширина окна оценивается, используя остатки при альтернативной гипотезе, как в (20).

Рассмотрим теперь вопрос построения доверительных интервалов (в общем случае доверительных множеств) для даты сдвига в коинтегрирующих регрессиях.

В случае регрессий, использующих стационарные переменные, дата структурного сдвига оценивается путем минимизации суммы квадратов остатков или путем использования метода максимального квазиправдоподобия, как было предложено в Bai (1997) и Bai and Perron (1998). В этом случае основным предположением, сделанным для построения доверительного интервала, было предположение о том, что величина структурного сдвига сходится к нулю со скоростью, меньшей, чем $1/\sqrt{T}$. На практике такая величина структурного сдвига будет слишком большой, а полученная асимптотическая аппроксимация не будет хорошей на малых выборках (если сдвиг будет достаточно малым, доверительный интервал будет недонакрывать истинную дату сдвига с заданной вероятностью). Подобная теория также предполагалась в работах (Bai *et al.*, 1998) и (Kejriwal and Perron, 2008*b*), где рассматривались регрессии с нестационарными пременными: при нестационарных переменных величина структурного сдвига должна сходиться к нулю со скоростью, меньшей, чем $1/T^{1/4}$. В работе (Kurozumi and Skrobotov, 2018) был предложен метод, основанный на обращении теста на местоположение структурного сдвига. Этот метод не требует нереалистичного предположения о величине сдвига и, как показывают симуляции авторов, превосходит подход (Bai *et al.*, 1998) и (Kejriwal and Perron, 2008*b*), давая корректный уровень накрытия как для малых сдвигов, так и для больших.

Kurozumi and Skrobotov (2018) рассматривают линейную модель с единственным сдвигом:

$$y_t = w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + e_t \tag{22}$$

для $t = 1, \cdots, T$, где $w_{b,t}$, $w_{b,t}(\lambda_0)$ и $w_{f,t}$ являются p_b-, p_b- и p_f-мерными регрессорами соответственно, $w_{b,t}(\lambda_0) = 1(t > [\lambda_0 T])w_{b,t}$ с индикатор-функцией $1(\cdot)$, λ_0 – действительная доля структурного сдвига, $[a]$ обозначает наибольшую целую часть числа, меньшую чем a, e_t – ошибка, β_b, δ_b и β_f – p_b-, p_b- и p_f-мерные векторы неизвестных коэффициентов соответственно. Поскольку целью является построение доверительного множества для даты структурного сдвига, предполагается, что происходит одномоментный структурный сдвиг в выборке, и истинная дата структурного сдвига обозначается как $T_0 = [\lambda_0 T]$. Заметим, что в коэффициентах, связанных с $w_{b,t}$, также происходит структурный сдвиг: они изменяются с β_b до $\beta_b + \delta_b$, в то время как β_f фиксированы во всей выборке.

Поскольку мы рассматриваем модель коинтегрирующей регрессии, регрессоры $w_{b,t}$ и/или $w_{f,t}$ включают в себя I(1) переменные, которые задаются как

$$\begin{aligned} z_{b,t} &= z_{b,t-1} + u^z_{b,t} \\ z_{f,t} &= z_{f,t-1} + u^z_{f,t}, \end{aligned}$$

где $z_{b,t}$ и $z_{f,t}$ являются p^z_b- и p^z_f-мерными векторами соответственно. Для простоты далее мы продолжаем изложение при предположении об отсутствии сноса (drift) I(1) переменных, даже когда $w_{b,t}$ включает в себя линейный тренд. Как обсуждается в конце данного раздела, основной результат в регрессиях с I(1) переменными со сносом остается неизменным, если слегка изменить статистику теста.

В общем случае ошибки $u^z_{b,t}$ и $u^z_{f,t}$ являются коррелированными с ошибками e_t в (22), и эта корреляция становится мешающим параметром. Для устранения этой корреляции мы включаем опережающие и запаздывающие разности I(1) переменных в качестве регрессоров в модель (22), как предлагается, среди прочих, в (Saikkonen, 1991). Такая модель принимает вид

$$y_t = w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + \sum_{j=-l}^{l}\pi'_{b,j}\Delta z_{b,t-j} + \sum_{j=-l}^{l}\pi'_{f,j}\Delta z_{f,t-j} + u_t, \tag{23}$$

где Δ обозначает оператор первых разностей, и глубина запаздывающих и опережающих разностей обозначается как l для удобства обозначения, но они могут быть отличными, как, например, в (Choi and Kurozumi, 2012). Далее мы предполагаем, что u_t некоррелированы с $u^z_{b,t-j}$ и $u^z_{f,t-j}$ для всех j.

Для удобства обозначения объединим дополнительные стационарные регрессоры $\Delta z_{b,t-j}$ и $\Delta z_{f,t-j}$ и обозначим их как p^x_f-мерный вектор $x_{f,t}$. Поскольку $x_{f,t}$ состоит из первых разностей I(1) регрессоров, можно всюду предположить, что $E[x_{f,t}] = 0$. Кроме того, поскольку коэффициент, связанный с $x_{f,t}$, фиксирован во всей выборке, мы включаем $x_{f,t}$ в $w_{f,t}$, так что модель можно просто выразить как

$$\begin{aligned} y_t &= w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta_b + w'_{f,t}\beta_f + u_t \\ &= w_t(\lambda_0)'\beta + u_t, \end{aligned} \tag{24}$$

где $w_{f,t}$ включает $x_{f,t}$, и $w_t(\lambda_0) = [w'_{b,t}, w_{b,t}(\lambda_0)', w'_{f,t}]'$.

Kurozumi and Skrobotov (2018) рассматривают следующий широко используемые в практическом анализе спецификации:

Модель II-a : Константа, линейный тренд и I(1) регрессоры включаются в коинтеграционное соотношение, и на все регрессоры, за исключением регрессоров первых разностей I(1) регрессоров, оказывает влияние структурный сдвиг. То есть $w_{b,t} = [1, t, z'_{b,t}]'$, $w_{f,t} = x_{f,t}$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}, \beta'_{b,z}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}, \delta'_{b,z}]'$ и $\beta_f = \beta_{f,x}$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + z'_{b,t}\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t + z'_{b,t}\delta_{b,z}\right) + x'_{f,t}\beta_{f,x} + u_t.$$

Модель II-b : Константа, линейный тренд и I(1) регрессоры включаются в коинтеграционное соотношение, и сдвиг происходит только в константе и линейном тренде. То есть $w_{b,t} = [1, t]$, $w_{f,t} = [z'_{f,t}, x'_{f,t}]'$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}]'$ и $\beta_f = [\beta'_{f,z}, \beta'_{f,x}]'$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t\right) + z'_{f,t}\beta_{f,z} + x'_{f,t}\beta_{f,x} + u_t.$$

Модель II-c : Константа, линейный тренд и I(1) регрессоры включаются в коинтеграционное соотношение, и некоторые коэффициенты, связанных с I(1) регрессорами, являются фиксированными во всей выборке. То есть $w_{b,t} = [1, t, z'_{b,t}]'$, $w_{f,t} = [z'_{f,t}, x'_{f,t}]'$, $\beta_b = [\beta_{b,c}, \beta_{b,\tau}, \beta'_{b,z}]'$, $\delta_b = [\delta_{b,c}, \delta_{b,\tau}, \delta'_{b,z}]'$, and $\beta_f = [\beta'_{f,z}, \beta'_{f,x}]'$:

$$y_t = \beta_{b,c} + \beta_{b,\tau}t + z'_{b,t}\beta_{b,z} + 1(t > [\lambda_0 T])\left(\delta_{b,c} + \delta_{b,\tau}t + z'_{b,t}\delta_{b,z}\right) + z'_{f,t}\beta_{f,z} + x'_{f,t}\beta_{f,x} + u_t.$$

Также авторыми были рассмотрены Модели I-a, I-b, I-c, которые отличаются от Моделей II-a, II-b, II-c отсутствием линейного тренда.

В этих моделях ошибка u_t предполагается некоррелированной со всеми опережающими и запаздывающими разностями $\Delta z_{b,t}$ и $\Delta z_{f,t}$.Также не допускается коинтеграция среди I(1) регрессоров: на практике у нас может быть несколько I(1) регрессоров, и в этом случае необходимо сначала тестировать на коинтеграцию и сделать вывод, что нет коинтеграционного соотношения среди I(1) регрессоров.

Для неизвестной даты структурного сдвига Kurozumi and Skrobotov (2018) проверяют гипотезу

$$H_N:\ T_0 = T_1 \quad \text{против} \quad H_A:\ T_0 = T_2 \tag{25}$$

или

$$H_N:\ \lambda_0 = \lambda_1 \quad \text{vs.} \quad H_A:\ \lambda_0 = \lambda_2 \tag{26}$$

где $\lambda_1 = T_1/T$ и $\lambda_2 = T_2/T$, на уровне значимости α, и если нулевая гипотеза не отвергается, то мы включаем T_1 в доверительное множество; в противном случае мы исключаем T_1 из доверительного множества. Проводя этот тест для всех допустимых дат структурных сдвигов, мы получаем доверительное множество для даты структурного сдвига с уровнем доверия $1 - \alpha$.

В этой процедуре доверительное множество становится меньше, если тест становится более мощным, и поэтому мы должны построить как можно более мощный тест. Однако нетрудно видеть, что не существует равномерно наиболее мощного теста для тестирования гипотезы (25). Вместо этого, следуя литературе, авторы рассматривают тест, который максимизирует взвешенное среднее мощности.

Сначала отметим, что мы не можем непосредственно оценить (24), используя $w_{b,t}(\lambda_0)$, поскольку $w_{b,t}(\lambda_0)$ зависит от неизвестной доли даты сдвига λ_0. Поскольку тестовая задача задается как (25), мы рассмотрим оценивание модели при нулевой гипотезе и построим тестовую статистику. Пусть $w_{b,t}(\lambda_1) = 1(t > [\lambda_1 T])w_{b,t}$ и $r_t(\lambda_0, \lambda_1) = w_{b,t}(\lambda_0) - w_{b,t}(\lambda_1)$. Тогда модель (24) можно записать как

$$\begin{aligned} y_t &= w'_{b,t}\beta_b + w_{b,t}(\lambda_0)'\delta + w'_{f,t}\beta_f + u_t \\ &= w'_{b,t}\beta_b + w_{b,t}(\lambda_1)'\delta + w'_{f,t}\beta_f + u_t + r_t(\lambda_0, \lambda_1)'\delta \\ &= w_t(\lambda_1)'\beta + u_t(\lambda_0, \lambda_1), \end{aligned} \tag{27}$$

где $w_t(\lambda_1) = [w'_{b,t}, w_{b,t}(\lambda_1)', w'_{f,t}]'$ и $u_t(\lambda_0, \lambda_1) = u_t + r_t(\lambda_0, \lambda_1)'\delta$. Можно показать, что не существует равномерно наиболее мощного теста, и мощность теста будет зависеть от величины структурного сдвига, δ, и местоположения структурного сдвига при альтернативной гипотезе, λ_2, и, таким образом, мощность можно выразить как вероятность $P\,(\varphi$ отвергает $H_N|\delta, \lambda_2)$, где φ – тест для проверки (26) на уровне значимости α. Следуя (Andrews and Ploberger, 1994), (Elliott and Müller, 2007) и (Kurozumi and Yamamoto, 2015), среди прочих, Kurozumi and Skrobotov (2018) рассматривали максимизацию взвешенного среднего $P\,(\varphi$ отвергает $H_N|\delta, \lambda_2)$ по δ и λ_2, используя некоторые взвешивающие функции, которые задаются как

$$\int\int P\,(\varphi \text{ отвергает } H_N|\delta, \lambda_2)\, dQ_{\lambda_2}(\delta) dJ(\lambda_2) \tag{28}$$

где $Q_{\lambda_2}(\delta)$ и $J(\lambda_2)$ являются неотрицательными мерами на $\mathbb{R}^{p_b}$ и $(0, 1)$ соответственно. Обычно эти взвешивающие функции выбираются таким образом, чтобы асимптотическое распределение тестовой статистики не зависело от мешающих параметров.

Авторы выводят тест отношения правдоподобий, который максимизирует среднюю мощность, заданную в (28) и получают его предельное распределение. Предельные распределения для моделей без I(1) регрессоров с фиксированными коэффициентами имеют более простой вид, поскольку в этих моделях регрессор $w_{b,t}$, в коэффициентах которого происходит структурный сдвиг, становится асимптотически ортогональным другим регрессорам $w_{f,t}$ с фиксированными коэффициентами. С другой стороны, когда I(1) регрессоры включены в $w_{f,t}$, $w_{b,t}$ коррелирован с $w_{f,t}$ даже в пределе, и, таким образом, предельное распределение тестовой статистики зависит от p_f^z, числа I(1) регрессоров с фиксированными коэффициентами.

Критические значения этих распределений зависят от λ_1, и неудобно их табулировать для всех допустимых долей дат структурных сдвигов λ_1. Вместо этого Kurozumi and Skrobotov (2018) приводят регрессии поверхности отклика для вычисления критических значений.

Предельное распределение теста зависит также от локализующего параметра c, который контролирует вес величины структурного сдвига, так что тест может лучше обнаружить небольшой структурный сдвиг, когда c близко к нулю, в то время как тест с большим значением c подходит для большого структурного сдвига. Следуя (Andrews and Ploberger, 1994), Kurozumi and Skrobotov (2018) рассматривают тест типа среднее, для которого вес накладывается на небольшое изменение ($c \to 0$) и тест типа экспоненциального, для которого принимается во внимание большая величина сдвига ($c \to \infty$), а также тест типа супремум, следуя (Andrews, 1993).

В итоге, тесты Kurozumi and Skrobotov (2018) имеют вид

$$
\begin{aligned}
\text{sup-}LR_T(T_1) &= \max_{T_2 \in \mathcal{T}_\epsilon} F_{T_2}(T_1), && (29)\\
\text{avg-}LR_T(T_1) &= \frac{1}{T^*} \sum_{T_2 \in \mathcal{T}_\epsilon} F_{T_2}(T_1), && (30)\\
\text{exp-}LR_T(T_1) &= \log\left\{ \frac{1}{T^*} \sum_{T_2 \in \mathcal{T}_\epsilon} \exp\left(\frac{1}{2} F_{T_2}(T_1)\right)\right\}, && (31)
\end{aligned}
$$

где $\mathcal{T}_\epsilon = \{T_2 : \ \epsilon T \le T_2 < T_1 - \epsilon T, \ T_1 + \epsilon T < T_2 \le (1-\epsilon)T\}$, T^* – количество значений T_2, включенных в $\mathcal{T}_\epsilon$ и $F_{T_2}(T_1)$ является тестовой статистикой для простой нулевой гипотезы $T_0 = T_1$ против простой альтернативы $T_0 = T_2$, которая задается как

$$
F_{T_2}(T_1) = \begin{cases} \left(\sum_{t=T_2+1}^{T_1} w_{b,t}\hat{u}_t\right)' \left(\tilde{\omega}_{uu}\hat{H}\right)^{-1} \left(\sum_{t=T_2+1}^{T_1} w_{b,t}\hat{u}_t\right) & : \ T_2 < T_1 \\ \left(\sum_{t=T_1+1}^{T_2} w_{b,t}\hat{u}_t\right)' \left(\tilde{\omega}_{uu}\hat{H}\right)^{-1} \left(\sum_{t=T_1+1}^{T_2} w_{b,t}\hat{u}_t\right) & : \ T_1 < T_T \end{cases},
$$

где $\hat{u}_t$ являются регрессионными остатками от регрессии y_t на $w_{b,t}$, $w_{b,t}(\lambda_1)$ и $w_{f,t}$,

$$
\hat{H} = \sum_{t=1}^{T} \hat{r}_t(\lambda_2, \lambda)\hat{r}_t(\lambda_2, \lambda_1)'
$$

с регрессионными остатками $\hat{r}_t$ от регрессии $r_t(\lambda_2, \lambda_1)$ на $w_t(\lambda_1)$ и $\tilde{\omega}_{uu}$ является состоятельной оценкой ω_{uu}. Эту оценку можно вычислить, применяя подход Yamamoto (2016), так чтобы она была состоятельной и при нулевой, и при альтернативной гипотезе.

Когда $I(1)$ регрессоры имеют компоненту сноса, мы обычно включаем линейный тренд в регрессии. Поскольку константа и линейный тренд (со структурным сдвигом) включены как регрессоры в модели от II-a до II-c, тестовая статистика $\widetilde{LR}_T(\lambda_1)$ является инвариантной к компоненте сноса I(1) регрессоров, если мы заменим $\Delta z_{b,t-j}$ и $\Delta z_{f,t-j}$ в $x_{f,t}$ на усредненные версии, $\Delta z_{b,t-j} - \overline{\Delta z_b}$ и $z_{f,t-j} - \overline{\Delta z_f}$ соответственно, чтобы гарантировать, что $E[x_{f,t}] = 0$. Заметим, что когда I(1) регрессоры имеют компоненту сноса в моделях с I-a до I-c, предельное распределение будет отличаться от того, которое приведено в теореме 1 Kurozumi and Skrobotov (2018), но этот случай не рассматривается, поскольку естественно и практически более релевантно включить линейный тренд в модели с II-a до II-c, когда I(1) переменные имеют снос. Мы также замечаем, что компонента сноса I(1) регрессоров сама может иметь структурный сдвиг, поскольку она может быть поглощена линейным трендом (со структурным сдвигом). В этом случае усредненные версии I(1) регрессоров можно построить, используя оцененную дату сдвига, как в оценивании ω_{uu}, объясняемое в конце этого подраздела.

3 Многомерные методы

3.1 Тестирование ранга коинтеграции при наличии структурных сдвигов в данных

3.1.1 Известная дата структурного сдвига

В работе (Saikkonen and Lütkepohl, 2000a) рассматривается тестирование гипотезы о том, что имеется r коинтегрирующих соотношений, против альтернативы, что их более чем r, при допущении сдвига в функции тренда отдельного ряда при нулевой и альтернативной гипотезе. Их процедура, однако, ограничивается только сдвигом в уровнях, происходящим в известное время.

Пусть $y_t = (y_{1t}, ..., y_{nt})'$ – n-мерный временной ряд ($t = 1, ..., T$), порождённый следующим образом:

$$y_t = \mu_0 + \mu_1 t + \delta_0 D_t + \delta_1 DU_t + x_t, \tag{32}$$

где μ_i и δ_i ($i = 0, 1$) – неизвестные параметры размерности ($n \times 1$), $D_t = \mathbb{I}(t = T_0)$ – импульсная дамми-переменная, $DU_t = \mathbb{I}(t \geq T_1)$ – дамми-переменная, отвечающая за сдвиг в уровнях. Предполагается, что компонента x_t (ненаблюдаемый процесс ошибок) имеет VAR(p) представление

$$x_t = A_1 x_{t-1} + \cdots + A_p x_{t-p} + \varepsilon_t. \tag{33}$$

Здесь A_j – ($n \times n$) матрица коэффициентов, $\varepsilon_t \sim NID(0, \Omega)$. Условие нормальности ошибок наложено для удобства и может быть ослаблено до *i.i.d.* и надлежащих условий на моменты. Представление (33) в форме модели коррекции ошибок выглядит следующим образом:

$$\Delta x_t = \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \; t = p+1, p+2, \ldots, \tag{34}$$

где $\Pi = -(I_n - A_1 - \cdots - A_p)$ и $\Gamma_j = -(A_{j+1} - ... - A_p)$ ($j = 1, \ldots, p-1$) – ($n \times n$) матрицы.

В предположении, что x_t – интегрированный первого порядка временной ряд, $I(1)$, и его компоненты коинтегрированы с рангом коинтеграции r, матрицу Π можно записать как

$$\Pi = \alpha\beta', \tag{35}$$

где α и β матрицы полного ранга размерности ($n \times r$).

Кроме того, используя теорему представления Грейнджера, процесс x_t может быть представлен в виде

$$x_t = C \sum_{j=1}^{t} \varepsilon_j + \xi_t, \; t = 1, 2, \ldots, \tag{36}$$

где $C = \beta_\perp(\alpha_\perp' \Gamma \beta_\perp)^{-1} \alpha_\perp'$ (здесь $\alpha_\perp$ и $\beta_\perp$ – матрицы (ортогональные дополнения) размерности $n \times (n-r)$ такие, что $\alpha' \alpha_\perp$ и $\beta' \beta_\perp = 0$)[8] и $\Gamma = I_n - \Gamma_1 - ... - \Gamma_{p-1}$, ξ_t – (асимптотически) стационарный процесс с нулевым средним.[9]

[8]Ортогональное дополнение невырожденной матрицы есть ноль, а ортогональное дополнение нулевой матрицы есть единичная соответствующей размерности.

[9]Обсуждение теоремы представления Грейнджера см., например, в (Johansen, 1996, Chapter 4) или (Juselius, 2006, Chapter 5).

Определим лаговый полином

$$A(L) = I_n - A_1 L - ... - A_p L^p = I_n \Delta - \Pi L - \Gamma_1 \Delta L - ... - \Gamma_{p-1} \Delta L^{p-1} \quad (37)$$

где, исходя из представления (34),

$$\begin{array}{rcl} A_1 & = & I_n + \alpha\beta' + \Gamma_1 \\ A_j & = & \Gamma_j - \Gamma_{j-1},\ j = 2, \dots, p-1 \\ A_p & = & -\Gamma_{p-1} \end{array} \quad (38)$$

Умножая (32) на $A(L)$, получим

$$\begin{aligned} \Delta y_t = \upsilon_0 + \upsilon_1 t + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \delta_0 D_t - \sum_{j=1}^{p} A_j \delta_0 D_{t-j} + \delta_1 \Delta DU_t \\ - \sum_{j=1}^{p-1} \Gamma_j \delta_1 \Delta DU_{t-j} - \Pi \delta_1 DU d_{t-1} + \varepsilon_t,\ t = p+1, p+2, \dots, \end{aligned} \quad (39)$$

где $\upsilon_0 = -\Pi\mu_0 + (\Gamma + \Pi)\mu_1$ и $\upsilon_1 = -\Pi\mu_1$.

Уравнение (39) можно переписать в более простом виде:

$$\begin{aligned} \Delta y_t = \upsilon + \alpha(\beta' y_{t-1} - \tau(t-1) - \theta DU_{t-1}) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p} \gamma_{0j} D_{t-j} \\ + \sum_{j=0}^{p-1} \gamma_{1j} \Delta DU_{t-j} + \varepsilon_t,\ t = p+1, p+2, \dots, \end{aligned} \quad (40)$$

где $\upsilon = -\Pi\mu_0 + \Gamma\mu_1$, $\tau = \beta'\mu_1$, $\theta = \beta'\delta_1$, а

$$\gamma_{0j} = \begin{cases} \delta_0, & j = 0 \\ -A_j \delta_0, & j = 1, \dots, p \end{cases}, \text{ и } \gamma_{1j} = \begin{cases} \delta_1, & j = 0 \\ -\Gamma_j \delta_1, & j = 1, \dots, p-1 \end{cases}.$$

Заметим, что здесь ΔDU_{t-j} – импульсная дамми $(= \mathbb{I}(t = T_1 + j))$, то есть, возможно, что некоторые импульсные дамми появляются дважды и могут быть объединены.

Уравнение (40) специфицирует модель коррекции ошибок (ECM) для наблюдаемого процесса y_t. Оценивая эту модель, мы можем получить предварительные оценки параметров процесса ошибок α, β, Γ_j $(j = 1, \dots, p-1)$ и Ω.

В рамках рассматриваемой модели следует проверить правильность предположения о ранге матрицы Π. С этой целью для некоторого специфицированного ранга r_0 мы тестируем нулевую гипотезу:

$$H_0(r_0) : rank(\Pi) = r_0 \text{ против } H_1(r_0) : rank(\Pi) > r_0. \quad (41)$$

Идея состоит в том, чтобы провести проверку этой гипотезы в рамках модели коррекции ошибок для выборочного аналога ненаблюдаемого процесса x_t, получаемого доступным детрендированием наблюдаемого процесса y_t с использованием оценок параметров μ_i

и δ_i $(i = 0, 1)$ модели (32). Для оценивания параметров μ_i и δ_i можно использовать доступную GLS (FGLS) процедуру, аналогичную (Saikkonen and Lütkepohl, 2000*b*), которая заключается в следующем.

Введем дополнительные переменные:

$$a_{0t} = \begin{cases} 1, & t \geq 1 \\ 0, & t \leq 0 \end{cases}, \text{ и } a_{1t} = \begin{cases} t, & t \geq 1 \\ 0, & t \leq 0 \end{cases}.$$

Умножая (32) слева на $A(L)$, получим

$$A(L)y_t = H_{0t}\mu_0 + H_{1t}\mu_1 + K_{0t}\delta_0 + K_{1t}\delta_1 + \varepsilon_t, \ t = 1, 2, \ldots, \tag{42}$$

где y_t полагаются равными 0 для $t \leq 0$, $H_{it} = A(L)a_{it}$, $K_{0t} = A(L)D_t$ и $K_{1t} = A(L)DU_t$.

Определим матрицу

$$Q = \left[\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1/2} : \alpha_\perp(\alpha'_\perp\Omega^{-1}\alpha_\perp)^{-1/2}\right] \tag{43}$$

имеющую свойство $Q'Q = \Omega^{-1}$. Умножая (42) слева на матрицу Q' (квадратный корень из матрицы Ω^{-1}), получим справа преобразованный вектор ошибок, ковариационная матрица которого является единичной матрицей, как это и требуется в GLS-оценивании. Таким образом, мы находим преобразование, которое приводит к регрессионной модели с обычными свойствами ошибок. Заметим, что представление $Q'Q = \Omega^{-1}$ неоднозначно; использование в качестве Q матрицы, указанной в (43) удобно для получения асимптотических свойств оценок.

Подходящие оценки параметров α, β, Γ_j $(j = 1, \ldots, p-1)$ и Ω можно получить, используя регрессию с пониженным рангом (40)[10]. Тогда оценки $\tilde{A}_j$ для A_j можно получить, используя (38). При этом, получим оцененный лаговый полином $\tilde{A}(L) = I_n - \tilde{A}_1L - \ldots - \tilde{A}_pL^p$, оцененные параметры $\tilde{H}_{it} = \tilde{A}(L)a_{it}$, $\tilde{K}_{0t} = \tilde{A}(L)D_t$, $\tilde{K}_{1t} = \tilde{A}(L)DU_t$ и оцененную матрицу $\tilde{Q}$. Оценивая уравнение регрессии:

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{Q}'\tilde{H}_{0t}\mu_0 + \tilde{Q}'\tilde{H}_{1t}\mu_1 + \tilde{Q}'\tilde{K}_{0t}\delta_0 + \tilde{Q}'\tilde{K}_{1t}\delta_1 + \eta_t, t = 1, \ldots, T, \tag{44}$$

получим оценки $\hat{\mu}_i$ и $\hat{\delta}_i$ $(i = 0, 1)$.

Saikkonen and Lütkepohl (2000*b*) установили, что $\beta'(\hat{\mu}_0 - \mu_0) = O_p(T^{-1/2})$, $\beta'_\perp(\hat{\mu}_0 - \mu_0) = O_p(1)$, $\hat{\delta}_0 - \delta_0 = O_p(1)$, $\beta'(\hat{\delta}_1 - \delta_1) = O_p((T - T_1)^{-1/2})$, $\beta'_\perp(\hat{\delta}_1 - \delta_1) = O_p(1)$, $\beta'(\hat{\mu}_1 - \mu_1) = O_p(T^{-3/2})$, $T^{1/2}\beta'_\perp(\hat{\mu}_1 - \mu_1) \Rightarrow N(0, \beta'_\perp C\Omega C'\beta_\perp)$, где $C\beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp$. То есть, $\hat{\mu}_1$ является состоятельной оценкой, в то время как $\hat{\mu}_0$ состоятельна только в направлении β. Ограниченность по вероятности этой оценки в направлении $\beta_\perp$, однако, достаточна для дальнейших целей. Оценка параметра $\hat{\delta}_0$ несостоятельна, но ограничена по вероятности, $\hat{\delta}_1$ состоятельна только в направлении β, когда $T - T_1 \to \infty$, другими словами, когда информация о параметре δ_1 в направлении β возрастает при увеличении выборки).

Построим выборочный аналог процесса x_t:

$$\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - \hat{\delta}_0 D_t - \hat{\delta}_1 DU_t \tag{45}$$

[10]Смотрите, например, (Johansen, 1996, Chapter 6) или (Juselius, 2006, Chapter 7).

и определим выборочный аналог уравнения (34):

$$\Delta\hat{x}_t = \Pi\hat{x}_{t-1} + \sum_{j=1}^{p-1}\Gamma_j\Delta\hat{x}_{t-j} + e_t,\ t = p+1,\dots,T. \tag{46}$$

Saikkonen and Lütkepohl (2000*a*) предлагают использовать для тестирования гипотезы (41) как обычную LR-статистику (см. Johansen (1996)), так и введенную в (Saikkonen and Lütkepohl, 2000*b*) LM-статистику. Однако Saikkonen and Lütkepohl (2000*b*) показывают, что, в целом, LR-статистика имеет лучшие свойства на конечных выборках, кроме некоторых частных случаев. Обе эти статистики асимптотически эквивалентны.

Полученное авторами предельное распределение LR-статистики не зависит от параметров структурных сдвигов, поэтому можно использовать критические значения, приведенные в (Lütkepohl and Saikkonen, 2000, Table 1) для случая отсутствия структурных сдвигов.

Кроме того, если наложено априорное ограничение $\mu_1 = 0$, все результаты продолжают выполняться[11]. Также, как уже было замечено выше, результаты легко обобщаются на случай произвольного количества структурных сдвигов, с возможным добавлением сезонных дамми [12].

Обобщение теста (Saikkonen and Lütkepohl, 2000*a*), допускающее изменение наклона тренда в модели, было рассмотрено в (Trenkler, Saikkonen and Lütkepohl, 2008). В этом случае в процесс (32) добавляется компонента, отвечающая за сдвиг в тренде, то есть

$$y_t = \mu_0 + \mu_1 t + \delta_1 DU_t + \delta_2 DT_t + x_t, \tag{47}$$

где $DT_t = \mathbb{I}(t - T_1 + 1)I(t \geq T_1)$ (заметим, что в уравнение (47) не включаются импульсные дамми).

Модель коррекции ошибок (40) переписывается тогда следующим образом:

$$\begin{aligned}\Delta y_t = \upsilon + \alpha(\beta' y_{t-1} - \tau(t-1) - \theta DT_{t-1}) + \sum_{j=1}^{p-1}\Gamma_j\Delta y_{t-j} + \sum_{j=0}^{p-1}\gamma_{1j}\Delta DU_{t-j}\\ + \eta_0\Delta DT_t + \varepsilon_t,\ t = p+1, p+2, \dots,\end{aligned} \tag{48}$$

где $\upsilon = -\Pi\mu_0 + \Gamma\mu_1$, $\tau = \beta'\mu_1$, $\theta = \beta'\delta_1$, $\eta_0 = \Gamma\delta_2 - \Pi\delta_1$ и

$$\gamma_{1j} = \begin{cases}\delta_1 + \Pi\delta_1 + \Gamma_1\delta_2 + ... + \Gamma_{p-1}\delta_2, & j = 0\\ -\Gamma_j\delta_1 + \Gamma_{j+1}\delta_2 + ... + \Gamma_{p-1}\delta_2, & j = 1,\dots,p-2\\ -\Gamma_{p-1}\delta_1, & j = p-1\end{cases}.$$

Здесь ΔDU_{t-j} – импульсная дамми, ΔDT_t – дамми сдвига в уровнях.

Далее применяется уже описанный выше подход Saikkonen and Lütkepohl (2000*a*). То есть уравнение (42) переписывается в виде

$$A(L)y_t = H_{0t}\mu_0 + H_{1t}\mu_1 + K_{1t}\delta_1 + K_{2t}\delta_2 + \varepsilon_t,\ t = 1,2,\dots, \tag{49}$$

[11] Только в соответствующем распределении LR-статистики процесс Броуновского моста заменяется на обычное Броуновское движение.

[12] Как отмечается в (Johansen, 1991), добавление сезонных дамми не меняет предельного распределения тестовой статистики.

где y_t полагаются равными 0 при $t \leq 0$, $H_{it} = A(L)a_{it}$ $(i = 0, 1)$, $K_{1t} = A(L)DU_t$ и $K_{2t} = A(L)DT_t$.

Далее находим оценки для α, β, Γ_j $(j = 1, \ldots, p-1)$ и Ω, используя метод Йохансена для регрессии (48) и получаем оценки $\tilde{A}_j$, лагового полинома $\tilde{A}(L)$, параметров $\tilde{H}_{it} = \tilde{A}(L)a_{it}$ $(i = 0, 1)$, $\tilde{K}_{1t} = \tilde{A}(L)DU_t$, $\tilde{K}_{2t} = \tilde{A}(L)DT_t$ и матрицы $\tilde{Q}$, где Q определяется, как в (43). Тогда регрессия (44) переписывается в виде:

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{Q}'\tilde{H}_{0t}\mu_0 + \tilde{Q}'\tilde{H}_{1t}\mu_1 + \tilde{Q}'\tilde{K}_{1t}\delta_1 + \tilde{Q}'\tilde{K}_{2t}\delta_2 + \eta_t,\ t = 1, \ldots, T. \tag{50}$$

Оценивая её OLS, получаем оценки $\hat{\mu}_i$ и $\hat{\delta}_i$. Авторы установили, что также как и в Saikkonen and Lütkepohl (2000a), $\hat{\mu}_0$ и $\hat{\delta}_1$ не состоятельны в направлении $\beta_\perp$, хотя ограниченность по вероятности достаточна для дальнейших целей. Совместное асимптотическое распределение $\hat{\mu}_1$ и $\hat{\delta}_2$ в направлении $\beta_\perp$ зависит от доли даты структурного сдвига ($\beta'(\hat{\mu}_1 - \mu_1) = O_p(T^{-3/2})$, $\beta'(\hat{\delta}_2 - \delta_2) = O_p(T^{-3/2})$, $[T^{1/2}\beta'_\perp(\hat{\mu}_1 - \mu_1) : T^{1/2}\beta'_\perp(\hat{\delta}_2 - \delta_2)] \Rightarrow \beta'_\perp C[\zeta_1 : \zeta_2]$, где $[\zeta_1 : \zeta_2] = [W_n(1) : W_n(1) - W_n(\lambda_1)] \begin{bmatrix} \lambda_1 & 1 - \lambda_1 \\ 1 - \lambda_1 & 1 - \lambda_1 \end{bmatrix}^{-1}$, $W_n(s)$ – n-мерное Броуновское движение с ковариационной матрицей Ω и $C = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp$).

Далее, после оценивания регрессии (50), можно тестировать гипотезу о ранге коинтеграции (41) на основе выборочного аналога ряда x_t:

$$\hat{x}_t = y_t - \hat{\mu}_0 - \hat{\mu}_1 t - \hat{\delta}_1 DU_t - \hat{\delta}_2 DT_t. \tag{51}$$

LR-статистика определяется тогда на основе уравнения (46), используя скорректированный на детерминированную компоненту ряд $\hat{x}_t$ в (51).

Предельное распределение полученной тестовой статистики зависит от доли даты структурного сдвига, в отличие от распределения, полученного в Saikkonen and Lütkepohl (2000a), и зависимость от этого параметра возникает из-за введения сдвига в тренде. Авторы получают также предельное распределение для случая произвольного количества структурных сдвигов через независимые Броуновские мосты. Предельное распределение будет зависеть только от относительных длин подвыборок между сдвигами, но не от их порядка.

Другой подход к анализу коинтеграции со структурными сдвигами был рассмотрен в (Johansen, Mosconi and Nielsen, 2000), где сразу анализировалась более общая задача с несколькими известными структурными сдвигами, без применения процедуры GLS-детрендирования. Этот подход является обобщением стандартного теста Йохансена (Johansen, 1996), используемого в ситуации без структурных сдвигов.

Рассмотрим модель с $(m - 1)$ структурными сдвигами (m режимами):

$$\Delta y_t = (\Pi, \Pi_i) \begin{pmatrix} y_{t-1} \\ t \end{pmatrix} + \mu_i + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t \tag{52}$$

где $i = 1, ..., m$ и $T_{i-1} + p < t \leq T_i$, где T_i – даты структурных сдвигов, $\varepsilon_t \sim NID(0, \Omega)$. Параметры Π, Γ_j и Ω предполагаются неизменными во всех режимах, тогда как Π_i и μ_i относятся к детерминированной компоненте и различны в каждом из режимов.

Для анализа гипотез о ранге коинтеграции запишем упрощенную модель (без запаздывающих разностей и без сдвигов):

$$\Delta y_t = \Pi y_{t-1} + \Pi_1 t + \mu + \varepsilon_t \tag{53}$$

Коинтеграция подразумевает пониженный ранг матрицы $\Pi = \alpha\beta'$. Модель (53) в общем случае имеет квадратичный тренд, который маловероятен при исследовании реальных данных, поэтому наложим ограничение $\Pi_1 = \alpha\gamma'$. Модель (53) при этом принимает вид:

$$\Delta y_t = \alpha(\beta' y_{t-1} + \gamma' t) + \mu + \varepsilon_t \tag{54}$$

Модель (54) является самой общей в нашем рассмотрении[13] и обозначается как $H_l(r)$: процесс имеет линейный тренд, коинтеграционное соотношение стационарно относительно линейного тренда. Если $\gamma = 0$, то процесс имеет линейный тренд, но коинтеграционное соотношение – нет. Такая модель обозначается как $H_{lc}(r)$. Если $\gamma = 0$ и $\mu = \alpha\rho'$, то процесс не имеет линейного тренда ни в одном направлении. Эта модель обозначается как $H_c(r)$.

Таким образом, гипотезы о ранге коинтеграции для (52) формулируются следующим образом:

$H_l(r)$: $rk(\Pi, \Pi_1, ..., \Pi_m) \leq r$;

$H_c(r)$: $rk(\Pi, \mu_1, ..., \mu_m) \leq r$ и $\Pi_1, ..., \Pi_m = 0$;

$H_{lc}(r)$: $rk(\Pi) \leq r$ и $\Pi_1, ..., \Pi_m = 0$.

Модель (52) описывается m уравнениями, но её можно записать и в виде одного уравнения.

Определим дамми-переменные: $D_{i,t} = \mathbb{I}(t = T_{i-1})$ (то есть $D_{i,t-p} = \mathbb{I}(t = T_{i-1} + p)$) и $DU_{i,t} = \mathbb{I}(T_{i-1} + p + 1 \leq t \leq T_i)$.

Таким образом, модель (52) принимает вид

$$\Delta y_t = \alpha \begin{pmatrix} \beta \\ \gamma \end{pmatrix}' \begin{pmatrix} y_{t-1} \\ tDU_t \end{pmatrix} + \mu DU_t + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=1}^{p} \sum_{i=2}^{m} \kappa_{ij} D_{i,t-j} + \varepsilon_t, \tag{55}$$

где $DU_t = (DU_{1,t}, ..., DU_{m,t})'$, $\mu = (\mu_1, ..., \mu_m)'$ и $\gamma = (\gamma_1', ..., \gamma_m')'$. Видно, что это представление совпадает с представлением (48) и, в частном случае, с (40) при некоторых переобозначениях. Ранг коинтеграции можно найти, применяя обычную процедуру Йохансена к этой модели.

В отличие от подхода Saikkonen and Lütkepohl (2000*a*), предельное распределение LR-статистики зависит от точной спецификации детерминированной компоненты и действительных дат сдвигов. Кроме того, оно отличается от полученных в Saikkonen and Lütkepohl (2000*a*) и Trenkler *et al.* (2008) распределениях дополнительной χ^2 компонентой. Это происходит из-за сохранения размерности в корректирующей компоненте ECM в Johansen *et al.* (2000). Тесты Saikkonen and Lütkepohl (2000*a*) и Trenkler *et al.* (2008) основаны на скорректированном ряде $\hat{x}_t$, в котором детерминированная функция, включающая сдвиг, содержится лишь косвенно. Следовательно, размерность корректирующей компоненты в (48) не зависит от числа сдвигов.

В работе Lütkepohl *et al.* (2003) сравнивались два рассмотренных выше подхода для случая единственного сдвига в уровнях (для модели $y_t = \mu_0 + \mu_1 t + \delta_1 d_{1t} + x_t$). Авторы дают предельные распределения тестовых статистик этих подходов при локальных альтернативах и сравнивают локальную мощность критериев[14], показывая, что их локальная мощность зависит от разности $(n - r_0)$, а не от n и r_0 в отдельности. Кроме того, она не

[13] Общее рассуждение о детерминированных компонентах см. в (Juselius, 2006, Chapter 6).

[14] Локальная альтернатива записывается в форме $\Pi = \alpha\beta' + T^{-1}\alpha_1\beta_1'$, где α и β – матрицы размера $(n \times r_0)$ и ранга r_0 , а α_1 и β_1 – матрицы размера $(n \times (r - r_0))$ и ранга $(r - r_0)$.

зависит от параметров среднего, тренда и сдвига (то есть от μ_0, μ_1 и δ_1). Для случая $\mu_1 = 0$ подход Saikkonen and Lütkepohl (2000*a*) имеет равномерно более высокую локальную мощность, чем подход Johansen *et al.* (2000). Если сравнивать тесты в ситуации с линейным трендом (то есть $\mu_1 \neq 0$), то подход Saikkonen and Lütkepohl (2000*a*) в большинстве случаев также имеет более высокую локальную мощность. Если сравнивать локальную мощность критериев для случаев $\mu_1 = 0$ и $\mu_1 \neq 0$, то для последнего наблюдается более низкая мощность, и различие здесь может быть весьма существенным. К потере мощности для всех тестов приводит возрастание величины $(n - r_0)$ – числа стохастических трендов при $H_0(r_0)$. То есть, единственное коинтеграционное соотношение сложнее определить в трёхмерном, нежели в двумерном процессе (результат, аналогичный известному для процессов без структурных сдвигов). Исследуя поведение на малых выборках, авторы установили, что мощность рассмотренных тестов примерно одинакова, хотя каждый из тестов имеет преимущество в отдельных специфических случаях. Однако, в общем, тест Saikkonen and Lütkepohl (2000*a*) имеет более благоприятные свойства в отношении размера. Кроме того, предельное распределение тестовой статистики не зависит от датировки структурного сдвига, и поэтому нет необходимости использовать специальные критические значения, что упрощает использование этого теста[15].

Симуляции Trenkler *et al.* (2008) показали, что тест с учётом сдвига в тренде обладает лучшими свойствами на конечных выборках, чем тест Johansen *et al.* (2000), то есть имеет меньшие искажения размера и сопоставимую или более высокую мощность. В некоторых ситуациях, однако, более высокую мощность имеет последний тест, поэтому рекомендуется основывать своё решение относительно числа коинтеграционных соотношений, используя оба теста одновременно.

Наконец, Andrade, Bruneau and Gregoir (2005) рассмотрели общий случай модели, в которой структурные сдвиги могут происходить не только в параметрах детерминированной функции времени, но и в коинтеграционных соотношениях.

3.1.2 Неизвестная дата сдвига

Ранее предполагалось, что датировка структурных сдвигов в многомерных моделях известна априорно. Если даты сдвигов неизвестны, Lütkepohl, Saikkonen and Trenkler (2004) (далее LST) предложили подход к их оцениванию на основе рассмотренного выше подхода SL для известных дат сдвигов. LST рассмотрели случай только одного сдвига в уровнях, то есть модель

$$y_t = \mu_0 + \mu_1 t + \delta_1 DU_t + x_t, \tag{56}$$

аналогичную модели (32) с такими же свойствами.

Оценку даты сдвига можно получить на основе VAR-модели в уровнях, полученной из (39), не накладывая никаких ранговых ограничений на Π и приводя модель к виду

$$y_t = \upsilon_0 + \upsilon_1 t + \delta_1^* DU_t + \sum_{j=0}^{p-1} \gamma_{1j}^* \Delta DU_{t-j} + A_1 y_{t-1} + \ldots. + A_p y_{t-p} + \varepsilon_t, \tag{57}$$

где снова $\upsilon_0 = -\Pi\mu_0 + (\Gamma + \Pi)\mu_1$ и $\upsilon_1 = -\Pi\mu_1$, а $\delta_1^* = -\Pi\delta_1$, $\gamma_{10}^* = \delta_1 - \delta_1^*$ и $\gamma_{1j}^* = \gamma_{1j}$.[16]

[15] Однако это не так для более общего случая сдвига в тренде.

[16] Можно было оценивать модель и в разностях, как в (Saikkonen, Lütkepohl and Trenkler, 2006).

Тогда дата сдвига T_1 определяется как

$$\hat{T}_1 = \arg\min_{T_1 \in \mathrm{T}} \det\left(\sum_{t=p+1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}'_t\right), \tag{58}$$

где T – область возможных дат сдвигов, а $\hat{\varepsilon}_t$ – остатки при OLS-оценивании (57).

Наличие p импульсных дамми в (57) может весьма затруднить датировку сдвигов. Дело в том, что постановка дамми на отдельные наблюдения (а не на группу наблюдений) приводит к тому, что эти наблюдения не учитываются в процессе минимизации, так что при этом может потеряться информация о действительных датах сдвигов. Поэтому можно провести оценивание дат сдвигов в модели без импульсных дамми, то есть в модели

$$y_t = \upsilon_0 + \upsilon_1 t + \delta_1^* DU_t + A_1 y_{t-1} + \cdots + A_p y_{t-p} + \varepsilon_t^*,\ t = p+1, \ldots, T \tag{59}$$

где $\varepsilon_t^* = \sum_{j=0}^{p-1} \gamma_{1j}^* \Delta DU_{t-j} + \varepsilon_t$. Тогда оценка даты сдвига $\tilde{T}_1$ будет равна

$$\tilde{T}_1 = \arg\min_{T_1 \in \mathrm{T}} \det\left(\sum_{t=p+1}^{T} \hat{\varepsilon}_t^* \hat{\varepsilon}_t^{\prime *}\right), \tag{60}$$

где $\hat{\varepsilon}_t^*$ – OLS-остатки при оценивании (59). Авторы показывают, что обе оценки являются состоятельными оценками действительных дат сдвигов и сходятся с одинаковой скоростью, даже при том, что $\tilde{T}_1$ основана на неправильно специфицированной модели.

Следует отметить одну особенность: возможен случай $\delta_1^* = -\Pi\delta_1 = 0$, который может выполняться даже если $\delta_1 \neq 0$. Тогда процесс $\beta' y_t$ не имеет сдвига, и, таким образом, мы имеем случай согласованных сдвигов (co-breaking). В этом случае ни одна из оценок доли даты сдвига не является состоятельной. Однако если сдвиг достаточно большой, то очень вероятна реакция на него импульсных дамми, и поэтому есть возможность найти дату сдвига, опираясь на оценку $\hat{T}_1$. С другой стороны, оценка $\tilde{T}_1$ может лишь случайно найти дату сдвига в этой ситуации. Таким образом, если нельзя исключить случай $\delta_1^* = 0$ (всегда выполняется, если ранг коинтеграции равен 0), использование только оценки $\tilde{T}_1$ может быть проблематичным.

После нахождения даты сдвига можно тестировать ранг коинтеграции, используя подходы, описанные в предыдущем разделе, предполагая оцененную дату сдвига как известную. При этом сохраняются все полученные результаты и получается то же самое предельное распределение тестовой статистики. При этом также возможно добавление в DGP сезонных и импульсных дамми. В то же время LST указывают, что свойства рассмотренных оценок и теста на коинтеграцию при малых выборках могут существенно зависеть от величины сдвига.

Saikkonen, Lütkepohl and Trenkler (2006) (далее SLT) было рассмотрено обобщение описанного выше подхода в нескольких направлениях. Были рассмотрены дополнительные оценки даты сдвига, а также асимптотические свойства всех оценок с введением зависимости величины сдвига от размера выборки. При этом анализировались асимптотические результаты, получаемые в ситуациях, когда величина сдвига возрастала (или убывала) при неограниченном увеличении объема выборки.

Более конкретно, параметр δ_1 в рассматриваемой работе определяется как

$$\delta_1 = T^a \delta_*,\ a \leq 1/2. \tag{61}$$

Таким образом, величина сдвига может быть убывающей, постоянной или возрастающей при росте выборки, эта величина зависит от a, которая меньше нуля, равна нулю или больше нуля, соответственно. Заметим, что когда структурный сдвиг достаточно большой, состоятельное оценивание даты сдвига, используя (58), требует накладывания дополнительного условия $\gamma^*_{1p-1} \neq 0$, что в противном случае приводит к тому, что асимптотически оценка даты сдвига даёт либо истинную дату сдвига, либо дату в предыдущий момент времени. Таким образом, "переспецификация" порядка VAR может привести к несостоятельной оценке даты сдвига . С другой стороны, когда размер сдвига достаточно мал, необходимо предположить, что $\delta_1 \neq 0$, то есть происходит фактическое изменение уровня в модели, а не только импульсные дамми. Это не является необходимым, если структурный сдвиг большой, так как тогда даже импульсные дамми могут быть использованы для точного оценивания датировки структурного сдвига. Однако даже при том, что состоятельная оценка даты сдвига не возможна в случае малого структурного сдвига, состоятельная оценка доли этого сдвига все еще возможна, если размер сдвига не является слишком маленьким.

Также авторы предлагают другую оценку даты сдвига, которая принимает во внимание связь между коэффициентами авторегрессии и коэффициентами, относящимися к дамми-переменным. Соответственно, вместо (57), рассматривается другая спецификация модели:

$$\Delta y_t = \upsilon_0 + \upsilon_1 t + \delta_1^* DU_{t-1} + \sum_{j=0}^{p-1} \gamma_{1j} \Delta DU_{t-j} + \Pi y_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t, \tag{62}$$

где γ_{1j} определяются как в (39). Уравнение (62) можно переписать в виде

$$\Delta y_t = \upsilon_0 + \upsilon_1 t + (I_n \Delta DU_t - \sum_{j=1}^{p-1} \Gamma_j \Delta DU_{t-j} - \Pi DU_{t-1})\delta_1 + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t. \tag{63}$$

В отличие от модели (57), импульсные дамми больше не появляются отдельно в представлении (63), а включены в компоненту, которая включает дамми, отвечающую за сдвиг. Таким образом, единственный вектор параметров δ_1 связан со всеми дамми, что позволяет оценить дату сдвига более точно. Оценка даты сдвига (обозначим как $\hat{T}_1^R$) строится как (58) для остатков в модели (63), оцененных нелинейным OLS. В отличие от предыдущей оценки даты сдвига, для больших сдвигов состоятельность выполняется без каких-либо дополнительных ограничений на коэффициенты. Для малых сдвигов, однако, результаты остаются аналогичными. Моделирование Монте-Карло показало, что оценка даты сдвига $\hat{T}_1^R$, основанная на регрессии (63), лучше всего находит истинную дату сдвига. Для решения проблемы вычислительной сложности предлагается на первом шаге найти оценку $\hat{T}_1$, а затем вычислить оценку $\hat{T}_1^R$ на основе $\hat{T}_1$, то есть таким образом, что $\left|\hat{T}_1 - \hat{T}_1^R\right| \leq p$ или $\left|\hat{T}_1 - \hat{T}_1^R\right| \leq 2p$. Другой вариант оценивания заключается в том, чтобы сначала оценить параметры Π, Γ_j $(j = 1, ..., p-1)$ в модели (62) без ограничений, а затем подставить их в уравнение (63) и применить обычный OLS. Оба варианта на малых выборках работают примерно так же, как если бы нахождение $\hat{T}_1^R$ производилось посредством полной оптимизации нелинейной функции.

Асимптотические результаты для оценки (60) также были обобщены на случай зависимости параметра сдвига от величины выборки.

Тестирование на ранг коинтеграции несколько отличается от подхода SL из-за неидентифицируемости параметра μ_0 в направлении $\beta_\perp$, поэтому SLT оценивают только параметры μ_1 и δ_1 в детерминированной компоненте. Опишем далее процедуру оценивания этих параметров.

Аналогично SL на первом шаге мы оцениваем параметры α, β, Γ_j $(j = 1, ..., p-1)$ и Ω, а также υ и τ в модели

$$\Delta y_t = \upsilon + \alpha(\beta' y_{t-1} - \tau(t-1) - \theta DU_{t-1}) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \sum_{j=0}^{p-1} \gamma_{1j} \Delta DU_{t-j} + \varepsilon_t, \tag{64}$$

где $\upsilon = -\Pi\mu_0 + \Gamma\mu_1$, $\tau = \beta'\mu_1$, $\theta = \beta_1'\delta_1$ и$\gamma_{1j} = \begin{cases} \delta_1, & j = 0 \\ -\Gamma_j\delta_1, & j = 1, ..., p-1 \end{cases}$.

Равенство $\upsilon = -\Pi\mu_0 + \Gamma\mu_1$ можно записать как $\upsilon = -\Pi\mu_0 + \Gamma\beta(\beta'\beta)^{-1}\beta'\mu_1 + \Gamma\beta_\perp(\beta_\perp'\beta_\perp)^{-1}\beta_\perp'\mu_1$. Так как $\alpha_\perp'\Pi = \alpha_\perp'\alpha\beta' = 0$, умножение этого равенства слева на $\alpha_\perp'$ приводит к $\alpha_\perp'(\upsilon - \Gamma_\beta\varphi) = \alpha_\perp'\Gamma_{\beta_\perp}\varphi_*$, где $\varphi = \beta'\mu_1$, $\varphi_* = \beta_\perp'\mu_1$, $\Gamma_\beta = \Gamma\beta(\beta'\beta)^{-1}$, $\Gamma_{\beta_\perp} = \Gamma\beta_\perp(\beta_\perp'\beta_\perp)^{-1}$. Матрица $\alpha_\perp'\Gamma_{\beta_\perp}$ невырождена, обратная ей матрица есть $(\alpha_\perp'\Gamma_{\beta_\perp})^{-1} = \beta_\perp'\beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}$. Таким образом, $\varphi_* = \beta_\perp'C(\upsilon - \Gamma_\beta\varphi)$, где снова $C = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}\alpha_\perp'$. Значит можно построить оценки $\tilde{C}$ и $\tilde{\Gamma}_\beta$, используя оценки для α, β и Γ_j $(j = 1, ..., p-1)$, полученные из (64), после чего получить оценку $\tilde{\varphi}_* = \tilde{\beta}'_\perp\tilde{C}(\tilde{\upsilon} - \tilde{\Gamma}_\beta\tilde{\varphi})$ для φ_*.

Используя оценки $\tilde{\varphi}$ и $\tilde{\varphi}_*$, оценку для μ_1 получим как $\tilde{\mu}_1 = \tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\varphi} + \tilde{\beta}_\perp(\tilde{\beta}'_\perp\tilde{\beta}_\perp)^{-1}\tilde{\varphi}_*$.

Оценку параметра δ_1 можно получить аналогичным способом. Из (64) получаем, что $[\gamma_0, \gamma_1, ..., \gamma_{p-1}]' = [I_n, -\Gamma_1, ..., -\Gamma_{p-1}]'\delta_1$. Умножая обе части этого уравнения слева на матрицу $[\alpha_\perp' : \cdots : \alpha_\perp']$, получим $\alpha_\perp'\sum_{j=1}^{p-1}\gamma_j = \alpha_\perp'\Gamma\delta_1 = \alpha_\perp'\Gamma_\beta\theta + \alpha_\perp'\Gamma_{\beta_\perp}\theta_*$, где $\theta = \beta'\delta_1$ и $\theta_* = \beta_\perp'\delta_1$. Решая уравнение для θ_*, находим: $\theta_* = \beta_\perp'C\left(\sum_{j=1}^{p-1}\gamma_j - \Gamma_\beta\theta\right)$, соответствующая оценка равна $\tilde{\theta}_* = \tilde{\beta}'_\perp\tilde{C}\left(\sum_{j=1}^{p-1}\tilde{\gamma}_j - \tilde{\Gamma}_\beta\tilde{\theta}\right)$. Таким образом, оценка для δ_1 есть $\tilde{\delta}_1 = \tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\theta} + \tilde{\beta}_\perp(\tilde{\beta}'_\perp\tilde{\beta}_\perp)^{-1}\tilde{\theta}_*$. Авторы выводят асимптотические свойства полученных оценок, которые можно использовать для тестирования на ранг коинтеграции.

Рассматриваемые гипотезы специфицированы так же, как и в (41). Рассмотрим скорректированный на детерминированные компоненты ряд

$$\tilde{y}_t^{(0)} = y_t - \tilde{\mu}_1 t - \tilde{\delta}_1 DU_t = \mu_0 + x_t - (\tilde{\mu}_1 - \mu_1)t - \tilde{\delta}_1 DU_t + \delta_1 DU_t \tag{65}$$

Таким образом, отвлекаясь от отличия оценок от истинных значений параметров, мы имеем $\tilde{y}_t^{(0)} \sim \mu_0 + x_t$, поэтому естественно использовать тест, основанный на следующей искусственной регрессии:

$$\Delta\tilde{y}_t^{(0)} = \Pi^+\tilde{y}_{t-1}^{(+)} + \sum_{j=1}^{p-1}\Gamma_j\Delta y_{t-j}^{(0)} + e_t \tag{66}$$

где $\tilde{y}_{t-1}^{(+)} = [\tilde{y}_{t-1}^{(0)'}, 1]'$, и Π^+ получается добавлением дополнительного столбца к матрице Π. Трактуя эту модель как истинную, можно получить LR-статистику для специфицированного ранга коинтеграции, используя обычный метод Йохансена.

Из-за того, что в модель (66) включена константа, предельное распределение тестовой статистики отличается от полученного ранее в LST. Новые критические значения приведены авторами для случая $\mu_1 \neq 0$. Если $\mu_1 = 0$, то предельное распределение тестовой статистики то же самое, что и в Johansen (1996, Theorem 6.3), поэтому можно использовать критические значения из Johansen (1996, Table 15.2). Заметим также, что и здесь модель может быть обобщена добавлением произвольного количества импульсных дамми, сезонных дамми и т.п. По сравнению с тестом LST, тест SLT имеет тенденцию отвергать истинную нулевую гипотезу чаще.

Harris, Leybourne and Taylor (2016) обобщают подход Trenkler, Saikkonen and Lütkepohl (2008), описанный выше, на случай неопределённости относительно того, присутствует ли сдвиг в тренде в данных или нет. Если структурного сдвига в действительности нет в данных, нет смысла его включать в модель и тем самым терять мощность. С другой стороны, если сдвиг есть, но не учитывается в модели, размер тестов на коинтеграцию не будет корректным при нулевой гипотезе, и тест не будет состоятельным при альтернативной гипотезе.

Harris, Leybourne and Taylor (2016) предлагают три подхода для выбора модели: SC-VECM, SC-DIFF и SC-VAR. Процедура SC-VECM состоит в тестировании гипотезы, что ранг коинтеграции равен r, $H(r)$, против альтернативы $H(n)$. Пусть информационный критерий Шварца для заданной модели обозначается как $BIC(p, r; \lambda)$, где p – количество запаздывающих разностей в VECM с рангом коинтеграции r и долей даты сдвига $\lambda = T_1/T$. Пусть также $BIC(p, r)$ – информационный критерий Шварца для модели без сдвига. На первом шаге оценивается дата сдвига для всех возможных значений количества лагов в VECM, $p = 1, \ldots, \bar{p}$, для заданного ранга коинтеграции r путем максимизации функции правдоподобия по всем возможным датам сдвига. Затем для оцененной доли даты сдвига $\hat{\lambda}(p)$ число запаздывающих разностей выбирается на основе информационного критерия Шварца $BIC(p, n; \hat{\lambda}(p))$ для неограниченной модели VAR с $r = n$, получая $\hat{p}_r$. Количество лагов для модели без сдвига выбирается на основе минимизации $BIC(p, n)$ по всем p, получая $\hat{p}_0$. Наконец, итоговая модель выбирается на основе сравнения информационных критериев с и без сдвига, $BIC(\hat{p}_r, r; \hat{\lambda}(p))$ и $BIC(\hat{p}_0, r)$. После этого строится LR-тест для проверки гипотезы $H(r)$. Данная процедура последовательно повторяется для всех $r = 0, 1, \ldots,$ пока гипотеза $H(r)$ будет отвергаться.

Процедура SC-DIFF состоит в оценивании даты сдвига в модели в разностях, полагая $p = 1$ и $r = 0$. Итоговая модель выбирается на основе сравнения информационных критериев с и без сдвига, $BIC(1, 0; \hat{\lambda}(1))$ и $BIC(1, 0)$. Затем количество запаздывающих разностей p выбирается на основе выбранной модели со сдвигом или без при $r = n$. Далее тестируется гипотеза $H(r)$. Отметим, что хотя процедура основана на неправильно специфицированной модели с $r = 0$ и $p = 1$, она позволяет состоятельно выбрать корректную модель.

В процедуре SC-VAR дата сдвига и количество лагов оцениваются при ограничении $r = n$, в отличие от процедуры SC-DIFF. Модель выбирается на основе сравнения $BIC(\hat{p}_n, n; \hat{\lambda}(1))$ и $BIC(\hat{p}_n, n)$. Кроме этого, на конечных выборках Harris, Leybourne and Taylor (2016) рекомендуют заменять оценку даты сдвига в SC-VAR процедуре (имеющую плохие свойства) на оценку из SC-DIFF полагая $p = 1$ и $r = 0$.

Harris, Leybourne and Taylor (2016) показывают, что оценка доли даты сдвига является состоятельной вне зависимости от выбранного ранга коинтеграции и выбранного количества лагов. Все процедуры состоятельно выбирают верную модель (с или без сдвига) и имеют асимптотически корректный размер теста на ранг коинтеграции. Среди всех трех процедур авторы рекомендую использовать SC-VECM.

3.2 Тестирование коинтеграции при возможном изменении ранга в подвыборках

На практике тесты на коинтеграцию могут не выявлять наличие коинтеграции, предполагаемое экономической теорией. Одной из причин, из-за которых это может происходить, является наличие некоторой подвыборки в данных, в которой ранг коинтеграции может быть другим, нежели в оставшейся части выборки. Qu (2007) предлоил тест (однако для данных, не содержащих тренда), в котором нулевая и альтернативная гипотеза специфицируются следующим образом:

1. H_0: существует r_0 стабильных коинтегрирующих соотношений в системе, то есть ранг коинтеграции равен r_0;

2. H_1: существует $r > r_0$ коинтегрирующих соотношений в некоторой подвыборке $[T_1, T_2]$, то есть существует по крайней мере один режим с рангом коинтеграции, выше чем r_0.

Qu (2007) предлагает тест, допускающий m изменений режимов против альтернативы, что происходит $m+1$ изменений. Для состоятельности необходимо, чтобы хотя бы один режим имел ранг коинтеграции, выше чем r_0. Также накладывается обычное условие на минимальную длину режима, 20% от выборки ($\varepsilon = 0.2$).

Для заданного разбиения выборки $(T_0, T_1, \ldots, T_{m+1})$, где для удобства $T_0 = 0$ и $T_{m+1} = T$, построим усредненный (demeaned) процесс $\hat{U}_t$, полученный от регрессии y_t на константу сегмент за сегментом (так как необходимо построить состоятельный тест против сдвигов в константе). Другими словами, для $t \in [T_k + 1, T_{k+1}]$ процесс $\hat{U}_t$ определяется как

$$\hat{U}_t = y_t - \sum_{T_k+1}^{T_{k+1}} y_t = y_t - \bar{y}_k.$$

Для всей выборки можно записать этот процесс как

$$\hat{U}_t = \sum_{k=0}^{m} \mathbb{I}(T_k + 1 \leq t \leq T_{k+1})(y_t - \bar{y}_k). \tag{67}$$

Рассмотрим следующие матрицы, соответствующие второму моменту $\hat{U}_t$ и второму моменту частичных сумм:

$$A_k = (T_{k+1} - T_k)^{-2} \sum_{t=T_k+1}^{T_{k+1}} \hat{U}_t \hat{U}_t', \tag{68}$$

$$B_k = (T_{k+1} - T_k)^{-4} \sum_{t=T_k+1}^{T_{k+1}} \left(\sum_{j=T_k+1}^{t} \hat{U}_j \right) \left(\sum_{j=T_k+1}^{t} \hat{U}_j \right)', \tag{69}$$

для $k = 0, \ldots, m$ и определим следующее отношение моментных матриц:

$$Q_k = (B_k)^{-1/2} A_k (B_k)^{-1/2}. \tag{70}$$

Пусть $\rho_i(Q_k)$, $i = 1, \ldots, n$ обозначают i-е наименьшие собственные значения Q_k. Следовательно, $(\rho_1(Q_k), \ldots, \rho_n(Q_k))$ – упорядоченные решения следующей задачи нахождения собственных значений:

$$|\rho B_k - A_k| = 0. \tag{71}$$

Когда существует r_0 коинтегрирующих векторов, $n - r_0$ наименьших собственных значений сходится к невырожденному распределению, а r_0 наибольших собственных значений расходится к бесконечности. Следовательно, когда ранг коинтеграции больше r_0, сумма наименьших $n - r_0$ собственных значений расходится к бесконечности.

Так как альтернативная гипотеза предполагает ранг, больший чем r_0 в некоторых подвыборках, необходимо рассмотреть поведение $n - r_0$ собственных значений по всем допустимым режимам в выборке. То есть для допустимого разбиения $(T_1, \ldots, T_m)$ строится следующая статистика:

$$Q^m(T_1, \ldots, T_m) = \sum_{k=0}^{m} \sum_{i=1}^{n-r_0} \rho_i(Q_k). \tag{72}$$

При нулевой гипотезе эта статистика имеет невырожденное предельное распределение, при альтернативной она расходится из-за того, что расходится хотя бы в одном из режимов. Так как разбиение на режимы неизвестно, Qu (2007) предлагает использовать супремум от статистики (72) по всем возможным датам сдвигов:

$$Sup\text{-}Q^m = \sup_{(T_1,\ldots,T_m)\in \mathrm{T}_\varepsilon} \sum_{k=0}^{m} \sum_{i=1}^{n-r_0} \frac{\rho_i(Q_k)}{100(n - r_0)}, \tag{73}$$

где $\mathrm{T}_\varepsilon = \{(T_1, \ldots, T_m) : |T_{k+1} - T_k| \geq \varepsilon T, T_1 \geq \varepsilon T, T_m \leq (1 - \varepsilon)T\}$. Шкалирующий фактор $100(n - r_0)$ предназначен для контроля критических значений. Предельное распределение статистики (при нулевой гипотезе) не зависит от параметров, определяющих краткосрочную динамику. То есть, так как тест фактически непараметрический, нет необходимости оценивать количество лагов[17]. Однако на конечных выборках непараметрический тест автоматически корректирует серийную корреляцию только в некоторой степени, и эта коррекция может быть недостаточна (возникают либеральные искажения размера, когда силен эффект возвращения к среднему, и консервативные искажения размера, когда ряды в первых разностях положительно коррелированы). Также тест робастен к нелинейной переходной динамике.

Хотя число режимов неизвестно априорно, для состоятельности достаточно установить $m = 2$, так как в этом случае находится сегмент с более высоким рангом коинтеграции, и не нужно знать точное количество режимов при альтернативе. Поэтому автор предлагает построить следующие два теста:

$$WQ = \max\left(Sup\text{-}Q^1, Sup\text{-}Q^2 \frac{c(\alpha, 1, n - r_0, \varepsilon)}{c(\alpha, 2, n - r_0, \varepsilon)}\right), \tag{74}$$

$$SQ = Sup\text{-}Q^1 + Sup\text{-}Q^2 \frac{c(\alpha, 1, n - r_0, \varepsilon)}{c(\alpha, 2, n - r_0, \varepsilon)}, \tag{75}$$

[17] Оценивание количества лагов также становится проблематичным, производя поиск по всем возможным датам сдвигов.

где $c(\alpha, i, n-r_0, \varepsilon)$ обозначает критическое значение теста для i сдвигов на уровне значимости α.

Возможны дополнительные модификации тестовой статистики (73) для увеличения мощности, когда имеется некоторая априорная информация о местоположении сдвига. Если известно, что сдвиг происходит в конце выборки, используется $Sup\text{-}Q^F$ (forward recursive test), если в начале – $Sup\text{-}Q^R$ (reverse recursive test), если в середине – $Sup\text{-}Q^W$ (rolling window test), которые определяются следующим образом:

$$Sup\text{-}Q^F = \frac{1}{100(n-r_0)} \sup_{t\in[T_1+1,T]} \sum_{i=1}^{n-r_0} \rho_i((B_{1,t})^{-1/2} A_{1,t} (B_{1,t})^{-1/2}), \tag{76}$$

$$Sup\text{-}Q^R = \frac{1}{100(n-r_0)} \sup_{t\in[1,T_1]} \sum_{i=1}^{n-r_0} \rho_i((B_{t,T})^{-1/2} A_{t,T} (B_{t,T})^{-1/2}), \tag{77}$$

$$Sup\text{-}Q^W = \frac{1}{100(n-r_0)} \sup_{(t_1,t_2)\in \mathrm{T}_\varepsilon^W} \sum_{i=1}^{n-r_0} \rho_i((B_{t_1,t_2})^{-1/2} A_{t_1,t_2} (B_{t_1,t_2})^{-1/2}), \tag{78}$$

где $\mathrm{T}_\varepsilon^W = \{(T_1, T_2) : T_2 - T_1 \geq \varepsilon T, T_1 \geq 1, T_2 \leq T\}$, и $A_{i,j}$ и $B_{i,j}$ построены, используя наблюдения с i до j.

Qu (2007) получает при $r_0 = 0$ для независимых случайных блужданий в DGP, что размер теста $Sup\text{-}Q^2$ больше всех реагирует на размерность системы, хотя размер всех тестов близок к номинальному. Вводя переходную динамику в систему (AR или MA компоненты), размер тестов становится сильно нестабильным, что говорит о том, что на конечных выборках переходная динамика становится важной, хотя и не входит в предельное распределение статистик. Если ряды в первых разностях показывают сильную положительную корреляцию, тесты становятся очень консервативными. С другой стороны, если ряды в первых разностях показывают сильное возвращение к среднему, тест становится слишком либеральным[18].

Qu (2007) предлагает процедуру для исключения серийной зависимости из переходной компоненты системы для улучшения свойств на конечных выборках, используя разложение Бевериджа-Нельсона. Эта процедура заключается в следующем:

1. Оценить регрессию $\Delta y_t = \mu + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + e_t$, используя всю выборку;
2. Построить полиномы $\hat{\Gamma}(1) = I_n - \hat{\Gamma}_1 - \hat{\Gamma}_2$ и $\hat{\Gamma}^*(L) = \hat{\Gamma}_1^* + \hat{\Gamma}_2^* L$, где $\hat{\Gamma}_1^* = \hat{\Gamma}_1 + \hat{\Gamma}_2$ и $\hat{\Gamma}_2^* = \hat{\Gamma}_2$;
3. Построить скорректированный процесс $\tilde{y}_t = y_t + \hat{\Gamma}(1)^{-1} \hat{\Gamma}^*(L) \Delta y_t$;
4. Использовать $\tilde{y}_t$, чтобы построить тестовую статистику.

Заметим, что данное преобразование просто добавляет стационарный процесс к интегрированному, поэтому оно имеет тот же самый стохастический тренд, что и y_t, и коррекция не влияет на ранг коинтеграции. Также нет необходимости определить точное количество лагов (хотя это и может несколько улучшить свойства на конечных выборках) и

[18] Последнее происходит из-за того, что переходная динамика доминирует компоненту случайного блуждания.

использовать $p = 2$, так как это позволяет очистить ряд от сильной серийной корреляции. Использовании полной выборки для оценивания корректирующих параметров улучшает размер. Добавление константы и лагированной переменной (то есть, используя оценивание регрессии с пониженным рангом) применяется для того, чтобы учитывать возможное влияние коинтегрированности рядов при нулевой гипотезе.

Если нулевая гипотеза оказывается отвергнута, необходимо оценить местоположение сдвигов. Однако разбиение, которое максимизирует тестовую статистику, не обязательно состоятельно оценивает даты сдвигов, так как при альтернативе соответствующие нормализованный тест сходится к случайной величине вместо константы, и максимум не соответствует точке сдвига. Поэтому применяется подход Bai, Lumsdaine and Stock (1998), в котором состоятельная оценка доли даты сдвига вычисляется путем максимизации функции правдоподобия при некоторых слабых условиях на эту долю даты сдвига[19]. Тогда скорость сходимости доли даты сдвига достаточна для того, чтобы гарантировать стандартную $\sqrt{T}$-асимптотику для оцененных коэффициентов. Для обобщения результата Bai, Lumsdaine and Stock (1998) на случай нескольких сдвигов предлагается использовать процедуры Bai and Perron (1998) и Qu and Perron (2007).

Существует другой частный случай, в котором коинтеграция может изменяться в подвыборке. Речь идет о случае, в котором коинтеграционное соотношение может изменяться в конце выборки, причем коинтеграция может пропадать. Конец выборки, как предполагается, имеет конечную длину при стремящемуся к бесконечности объему выборки.

Одно из решений задачи тестирования нарушения коинтеграции в конце выборки было предложено в Andrews and Kim (2006). Авторы предложили тестировать нулевую гипотезу о стабильности во всей выборке против альтернативы о том, что происходит сдвиг в коэффициентах/распределении ошибок. Точнее, рассмотрим регрессионную модель вида

$$y_t = \begin{cases} x_t'\beta_0 + u_t, & \text{for } t = 1, \dots, T \\ x_t'\beta_t + u_t, & \text{for } t = T+1, \dots, T+m. \end{cases} \tag{79}$$

при нулевой гипотезе $\beta_t = \beta_0$ для $t = T+1, \dots, T+m$, и $\{u_t\}_{t=1}^{T+m}$ стационарный и эргодический процесс. При альтернативной гипотезе $\beta_t \neq \beta_0$ для некоторых $t = T+1, \dots, T+m$ и/или распределение $\{u\}_{T+1}^{T+m}$ и $\{u\}_1^T$ отличачаеются. Это означает, что при альтернативной гипотезе имеется структурный сдвиг в коэффициентах и/или распределении ошибок.

Andrews and Kim (2006) предложили P-тест и R-тест. P-тест можно построить как

$$P = P_{T+1}(\hat{\beta}_{1:(T+m)}) = \sum_{t=T+1}^{T+m} (y_t - x_t\hat{\beta}'_{1:(T+m)})^2, \tag{80}$$

где $\hat{\beta}_{1:(T+m)}$ есть OLS-оценка по всей выборке $t = 1, \dots, T+m$. Для проверки нлевой гипотезы о стабильности Andrews and Kim (2006) предлагают использовать методы сабсемплинга и эмпирическую функцию распределения $\{P_j(\beta),\ j = 1, ..., T-m+1\}$ с состоятельной оценкой для β. Вместо использования простой оценки для β, $\hat{\beta}_{1:(T+m)}$, чтобы исключить очень частое отвержение нулевой гипотезы Andrews and Kim (2006) предлагают использовать

[19] Условия заключаются в том, что сдвиг в среднем имеет порядок, выше чем $T^{-1/2}\log T$ и/или сдвиг в коинтегрирующем векторе имеет порядок, выше чем $T^{-1}\log T$.

оценку “leave-$m/2$-out” как

$$\hat{\beta}_{2(j)} = \text{оценка для } \beta\text{, использую наблюдения,} \\ \text{индексированные как } t = 1, \ldots, T \text{ с } t \neq j = j, \ldots, j + [m/2] - 1 \quad (81)$$

Тогда критическое знаечние на уровне значимости α можно получить как выборочную квантиль уровня $1 - \alpha$ величины $\{P_j(\hat{\beta}_{2(j)}) : j = 1, \ldots, T - m + 1\}$.

Также Andrews and Kim (2006) предложили другой тест вида

$$R = \sum_{t=T+1}^{T+m} \left(\sum_{s=t}^{T+m} (y_s - x_s \hat{\beta}'_{1:(T+m)}) \right)^2 \quad (82)$$

с тем же самым методом получения критических значений, как для P-теста. Этот тест асимптотически обоснован для более широко класса моделей. Как показывают результаты симуляций Монте-Карло, оба метода ведут себя несколько отличным образом от модели к модели.

Kim (2010) обобщил подход Andrews and Kim (2006), используя квази-разности в данных (квази-GLS), получая более робастный к серийной корреляции ошибок тест, чем тесты Andrews and Kim (2006).[20]

4 Нелинейные модели

Модели со структурными сдвигами, рассмотренные ранее, предполагают фиксированное число сдвигов на всём временном периоде, и эти сдвиги являются детерминированными по своей природе. Однако структурные сдвиги можно моделировать, исходя из динамики самого процесса, например, когда модель состоит из нескольких режимов (в простейшем случае из двух), и переход из одного в другой регулируется на основании прошлых значений самих временных рядов.[21]

Пороговая коинтеграция (Threshold cointegration) была введена в работе Balke and Fomby (1997), в которой авторы обобщили стандартную линейную коинтеграцию, допускающую нелинейную корректировку долгосрочного равновесия. Balke and Fomby (1997) предложили сначала тестировать гипотезу об отсутствии коинтеграции против альтернативы о наличии коинтеграции, а затем тестировать нулевую гипотезу о коинтеграции против альтернативы о пороговой коинтеграции. Hansen and Seo (2002) обобщили подход Balke and Fomby (1997) на многомерный случай. Hansen and Seo (2002) рассмотрели двухрежимную модель коррекции ошибок следующего вида[22]:

$$\Delta y_t = \begin{cases} A'_1 X_{t-1}(\beta) + u_t, \text{если } w_{t-1}(\beta) \leq \gamma \\ A'_2 X_{t-1}(\beta) + u_t, \text{если } w_{t-1}(\beta) > \gamma \end{cases}, \quad (83)$$

[20] См. также мониторинговые процедуры Wagner and Wied (2017) и Trapani and Whitehouse (2020), а также Zeileis *et al.* (2005).

[21] Обзор экономических приложений пороговых моделей см. в Hansen (2011).

[22] В работе Saikkonen (2005) выводятся условия, гарантирующие стационарность линейного долгосрочного соотношения, когда краткосрочная динамика в модели коррекции ошибок является нелинейной достаточно произвольного вида. В Saikkonen (2008) обобщаются результаты Saikkonen (2005), а также выводится аналог теоремы представления Грейнджера в контексте нелинейных векторных авторегрессионных моделей.

где $X_{t-1}(\beta) = (1, w_{t-1}(\beta), \Delta y_{t-1}, \ldots, \Delta y_{t-k})'$, $w_t(\beta) = \beta' y_t$, γ – параметр порога. Соответственно, параметры в модели зависят от того, превышает ли стационарная линейная комбинация $w_t(\beta)$ в предыдущий момент времени некоторый порог γ. В данной модели предполагается, что все переменные в модели изменяются между режимами, хотя можно и наложить разумные ограничения, например, допустить, что константа и краткосрочные параметры при запаздывающих разностях являются постоянными. Эту модель можно переписать в линейном виде как

$$\Delta y_t = A_1' X_{t-1}(\beta)\mathbb{I}(w_{t-1}(\beta) \leq \gamma) + A_2' X_{t-1}(\beta) + \mathbb{I}(w_{t-1}(\beta) > \gamma) + u_t. \tag{84}$$

Данную модель можно оценить как на основе метода максимального правдоподобия, так и на основе более вычислительно быстрого двухшагового метода, когда на первом шаге вычисляются оценки $\hat{A}_1(\beta, \gamma)$, $\hat{A}_2(\beta, \gamma)$ и ковариационная матрица ошибок $u_t(\beta, \gamma)$ для всех мозможных комбинаций (β, γ) (простым поиском по сетке, используя допустимое множество для параметров β на основе состоятельной оценки по линейной модели), а затем минимизируется определитель логарифма этой ковариационной матрицы по всем возможным (β, γ). Асимптотические свойства параметров оцененной модели, в том числе параметра порога, были исследованы в Seo (2011). Seo (2011) также предлагает сглаженное оценивание (smoothed least squares, SLS), позволяющее получить асимптотически нормальные оценки порога, но сходящиеся к нормальному распределению с более низкой скоростью. Kristensen and Rahbek (2010) и Kristensen and Rahbek (2013) предлагают использовать для оценивания QMLE, тем самым обобщая подход Johansen (1996) на случай, в котором коррекция в сторону долгосрочного соотношения происходит нелинейно посредством гладкой переходной функции, которая может быть и асимметричной. Kristensen and Rahbek (2010) и Kristensen and Rahbek (2013) также допускают наличие более одного коинтегарционного соотношения.

Для тестирования гипотезы о линейности (линейной коинтеграции), то есть $A_1 = A_2$, Hansen and Seo (2002) предлагают использовать Sup-LM статистику вида

$$Sup - LM = \sup_{\gamma_L \leq \gamma \leq \gamma_U} LM(\tilde{\beta}, \gamma), \tag{85}$$

где

$$LM(\beta, \gamma) = vec(\hat{A}_1(\beta, \gamma) - \hat{A}_2(\beta, \gamma))'(\hat{V}_1(\beta, \gamma) + \hat{V}_2(\beta, \gamma))^{-1} vec(\hat{A}_1(\beta, \gamma) - \hat{A}_2(\beta, \gamma)), \tag{86}$$

где $\hat{V}_1(\beta, \gamma)$ и $\hat{V}_2(\beta, \gamma)$ – робастные к гетероскедастичности оценки ковариационной матрицы оценок $vec\hat{A}_1(\beta, \gamma)$ и $vec\hat{A}_2(\beta, \gamma)$. В (85) используется $\tilde{\beta}$, оценка для β, полученная от оценивания модели при нулевой гипотезе, если параметр β неизвестен. Поскольку порог γ также может быть неизвестен, в (85) берется супремум по всем возможным порогам, поскольку не существует значения γ в условиях нулевой гипотезы (проблема, обозначенная в Davies (1987)). Важно отметить, что параметр γ, обеспечивающий максимально значение LM-статистики, может отличаться от оценки параметра γ, полученной при максимизации функции правдоподобия (которая максимизируется при альтернативной гипотезе). Для тестирования гипотезы о линейности при помощи Sup-LM теста Hansen and Seo (2002) предлагают использовать алгоритм бутстрапа с фиксированными регрессорами (fixed regressor bootstrap) или бутстрапа, основанного на остатках (residual bootstrap).

Hansen and Seo (2002), однако, не доказывают асимптотическую обоснованность бутстраповских тестов, ограничиваясь выведением асимптотического распределения тестовых статистик и демонстрации работы бутстрапа на симуляциях.

В работе Seo (2006) разрабатывается тест для проверки нулевой гипотезы о линейности и отсутствии коинтеграции в пороговой модели коррекции ошибок. В то время как Hansen and Seo (2002) предлагают тестировать линейность против нелинейности после того как протестирована гипотеза об отсутствии коинтеграции, тестирование последней гипотезы может давать некорректные выводы, если в действительности коинтеграция нелинейная. Таким образом, тест Seo (2006) учитывает, что при альтернативной гипотезе коинтеграция может быть как линейная, так и пороговая. Таким образом, нулевая гипотеза в модели (83) – $A_1 = A_2 = 0$. Seo (2006) предложил использовать максимум статистика Вальда с бутстраповскими критическими значениями (используя бутстрап, основанный на остатках), устанавливая состоятельность бутстрапа первого порядка. В Gonzalo and Pitarakis (2006*b*) рассматривалась похожая модель, как в Seo (2006), однако модель не ограничивалась тем, что пороговая переменная – это обязательно коинтеграционное соотношение. Также не делалось предположений о том, является ли модель полностью коинтеграционной (то есть могло быть так, что ранг коинтеграции равен 0 или равен числу переменных). Gonzalo and Pitarakis (2006*b*) предлагают использовать статистику Вальда для проверки гипоетзы о линейности против альтернативы, что модель нелинейная (пороговая). Предельное распределение, полученное авторами, не зависит от ранга коинтеграции.

Нелинейность может быть и в самом коинтеграционном соотношении. В работе Saikkonen and Choi (2004) была рассмотрена модель вида

$$y_t = g(x_t, \theta) + u_t, \tag{87}$$

где $g(x_t, \theta)$ – известная гладкая функция, зависящая от нестационарного процесса x_t. В работе Gonzalo and Pitarakis (2006*a*) была рассмотрена похожая пороговая модель вида

$$y_t = \beta' x_t + \lambda' x_t I(q_{t-d} > \gamma) + u_t, \tag{88}$$

где q_{t-d} – стационарная пороговая переменная.[23] В Saikkonen and Choi (2004) была разработана асимптотическая теория для коинтеграционной модели общего вида, охватывающей такую нелинейную переходную динамику, и существование нелинейности предполагалось изначально. В работе Choi and Saikkonen (2004) была предложена тестовая процедура для проверки гипотезы о линейности, основанная на разложении Тейлора функции $g(\cdot)$ в (87). В Gonzalo and Pitarakis (2006*a*) был предложен LM-тест для проверки гипотезы $\lambda = 0$ в (88). В работе Choi and Saikkonen (2010) предлагается тест на существование нелинейной коинтеграции, основанный на тесте Kwiatkowski *et al.* (1992), применяемом к подвыборкам остатков в регрессии (87). См. также общую теорию нелинейной коинтеграции в Wang (2015) и ссылки в ней.

Наконец, отметим работы Psaradakis *et al.* (2004) и Hu and Shin (2014), в которых нелинейность в компоненте коррекции ошибок следует процессу с Марковскими переключениями (Markov switching ECM – MS-ECM).

[23] В Medeiros *et al.* (2014) было исследовано поведение оценок при неучтенной нелинейности в коинтеграционном соотношении.

Список литературы

Скроботов, А.А. (2020). Структурные сдвиги и тестирование на единичный корень. *Прикладная эконометрика*, **58**, 96–141.

Andrade, P., Bruneau, C., and Gregoir, S. (2005). Testing for the cointegration rank when some cointegrating directions are changing. *Journal of Econometrics*, **124**(2), 269–310.

Andrews, D.W.K. (1991). Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation. *Econometrica*, **59**, 817–858.

Andrews, D.W.K. (1993). Tests for Parameter Instability and Structural Change with Unknown Change Point. *Econometrica*, **61**, 821–856.

Andrews, D.W.K. and Kim, J.-Y. (2006). Tests for cointegration breakdown over a short time period. *Journal of Business & Economic Statistics*, **24**(4), 379–394.

Andrews, D.W.K. and Ploberger, W. (1994). Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative. *Econometrica*, **62**, 1383–1414.

Arai, Y. and Kurozumi, E. (2007). Testing for the Null Hypothesis of Cointegration with a Structural Break. *Econometric Reviews*, **26**, 705–739.

Bai, J. (1997). Estimation of a Change Point in Multiple Regression Models. *The Review of Economics and Statistics*, **79**, 551–563.

Bai, J., Lumsdaine, R.L., and Stock, J.H. (1998). Testing for and Dating Breaks in Integrated and Cointegrated Time Series. *Review of Economic Studies*, **65**, 395–432.

Bai, J. and Perron, P. (1998). Estimating and Testing Linear Models with Multiple Structural Changes. *Econometrica*, **66**, 47–78.

Bai, J. and Perron, P. (2003). Computation and Analysis of Multiple Structural Change Models. *Journal of Applied Econometrics*, **18**, 1–22.

Balke, Nathan S and Fomby, Thomas B (1997). Threshold cointegration. *International economic review*, 627–645.

Bartley, W. A., Lee, J., and Strazicich, M. C. (2001). Testing the null of cointegration in the presence of a structural break. *Economics Letters*, **73**, 315–323.

Carrion-i-Silvestre, J.L. and Sansó-i-Rosselló, A. J. (2006). Testing the Null of Cointegration with Structural Breaks. *Oxford Bulletin of Economics and Statistics*, **68**, 623–646.

Carrion-i Silvestre, J. L. and Kim, D. (2019). Quasi-likelihood ratio tests for cointegration, cobreaking, and cotrending. *Econometric Reviews*, **38**(8), 881–898.

Choi, I. and Kurozumi, E. (2012). Model selection criteria for the leads-and-lags cointegrating regression. *Journal of Econometrics*, **169**, 224–238.

Choi, In and Saikkonen, Pentti (2004). Testing linearity in cointegrating smooth transition regressions. *The Econometrics Journal*, **7**(2), 341–365.

Choi, In and Saikkonen, Pentti (2010). Tests for nonlinear cointegration. *Econometric Theory*, 682–709.

Davies, R.B. (1987). Hypothesis testing when a nuisance parameter is present under the alternative. *Biometrika*, **74**, 33–43.

Elliott, G. and Müller, U.K. (2007). Confidence sets for the date of a single break in linear time series regressions. *Journal of Econometrics*, **141**, 1196–1218.

Gonzalo, Jesus and Pitarakis, Jean-Yves (2006*a*). Threshold effects in cointegrating relationships. *Oxford Bulletin of Economics and Statistics*, **68**, 813–833.

Gonzalo, Jesus and Pitarakis, Jean-Yves (2006*b*). *Threshold effects in multivariate error correction models*. Palgrave Macmillan.

Hansen, Bruce E (2011). Threshold autoregression in economics. *Statistics and its Interface*, **4**(2), 123–127.

Hansen, Bruce E and Seo, Byeongseon (2002). Testing for two-regime threshold cointegration in vector error-correction models. *Journal of econometrics*, **110**(2), 293–318.

Hansen, H. and Johansen, S. (1999). Some tests for parameter constancy in cointegrated var-models. *The Econometrics Journal*, **2**(2), 306–333.

Harris, D., Leybourne, S. J., and Taylor, A.M. R. (2016). Tests of the co-integration rank in var models in the presence of a possible break in trend at an unknown point. *Journal of Econometrics*, **192**(2), 451–467.

Hu, Liang and Shin, Yongcheol (2014). Testing for cointegration in markov switching error correction models. In *Essays in Honor of Peter CB Phillips*. Emerald Group Publishing Limited.

Jansson, M. (2005). Point optimal tests of the null hypothesis of cointegration. *Journal of Econometrics*, **124**(1), 187–201.

Johansen, S. (1991). Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models. *Econometrica*, **59**, 1551–1580.

Johansen, S. (1996). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford University Press, Oxford.

Johansen, S., Mosconi, R., and Nielsen, B. (2000). Cointegration analysis in the presence of structural breaks in the deterministic trend. *Econometrics Journal*, **3**, 216–249.

Juselius, K. (2006). *The Cointegrated VAR model: Methodology and Applications*. Oxford University Press, Oxford.

Kejriwal, M. (2009). Tests for a Mean Shift with Good Size and Monotonic Power. *Economics Letters*, **102**, 78–82.

Kejriwal, M. and Perron, P. (2008*a*). Data Dependent Rules for the Selection of the Number of Leads and Lags in the Dynamic OLS Cointegrating Regression. *Econometric Theory*, **24**, 1425–1441.

Kejriwal, M. and Perron, P. (2008*b*). The limit distribution of the estimates in cointegrated regression models with multiple structural changes. *Journal of Econometrics*, **146**, 59–73.

Kejriwal, M. and Perron, P. (2010). Testing for Multiple Structural Changes in Cointegrated Regression Models. *Journal of Business and Economic Statistics*, **28**, 503–522.

Kim, D. (2010). Improved and extended end-of-sample instability tests using a feasible quasi-generalized least squares procedure. *Econometric Theory*, 994–1031.

Kristensen, Dennis and Rahbek, Anders (2010). Likelihood-based inference for cointegration with nonlinear error-correction. *Journal of Econometrics*, **158**(1), 78–94.

Kristensen, Dennis and Rahbek, Anders (2013). Testing and inference in nonlinear cointegrating vector error correction models. *Econometric Theory*, 1238–1288.

Kurozumi, E. (2002). Testing for Stationarity with a Break. *Journal of Econometrics*, **108**, 63–99.

Kurozumi, E. and Skrobotov, A. (2018). Confidence sets for the break date in cointegrating regressions. *Oxford Bulletin of Economics and Statistics*, **80**(3), 514–535.

Kurozumi, E. and Yamamoto, Y. (2015). Confidence sets for the break date based on optimal tests. *Econometrics Journal*, **18**, 412–435.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., and Shin, Y. (1992). Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root: How Sure Are We That Economic Time Series Have a Unit Root? *Journal of Econometrics*, **54**, 159–178.

Lütkepohl, H. and Saikkonen, P. (2000). Testing for the cointegrating rank of a VAR process with a time trend. *Journal of Econometrics*, **95**, 177–198.

Lütkepohl, H., Saikkonen, P., and Trenkler, C. (2003). Comparison of tests for cointegrating rank of a VAR process with a structural shift. *Journal of Econometrics*, **113**, 201–229.

Lütkepohl, H., Saikkonen, P., and Trenkler, C. (2004). Testing for the cointegrating rank of a VAR process with level shift at unknown time. *Econometrica*, **72**, 647–662.

Medeiros, Marcelo C, Mendes, Eduardo, and Oxley, Les (2014). A note on nonlinear cointegration, misspecification, and bimodality. *Econometric Reviews*, **33**(7), 713–731.

Mogliani, M. (2010). Residual-Based Tests for Cointegration and Multiple Deterministic Structural Breaks: A Monte Carlo Study. Unpublished Manuscript, Paris School of Economics.

Phillips, P. C. B. and Hansen, B. E. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies*, **57**, 99–125.

Psaradakis, Zacharias, Sola, Martin, and Spagnolo, Fabio (2004). On markov error-correction models, with an application to stock prices and dividends. *Journal of Applied Econometrics*, **19**(1), 69–88.

Qu, Z. (2007). Searching for cointegration in a dynamic system. *Econometrics Journal*, **10**, 580–604.

Qu, Z. and Perron, P. (2007). *Estimating and Testing Structural Changes in Multivariate Regressions*. Volume 75.

Saikkonen, P. (1991). Asymptotically Efficient Estimation of Cointegration Regression. *Econometric Theory*, **7**, 1–21.

Saikkonen, Pentti (2005). Stability results for nonlinear error correction models. *Journal of Econometrics*, **127**(1), 69–81.

Saikkonen, Pentti (2008). Stability of regime switching error correction models under linear cointegration. *Econometric Theory*, 294–318.

Saikkonen, Pentti and Choi, In (2004). Cointegrating smooth transition regressions. *Econometric theory*, 301–340.

Saikkonen, P. and Lütkepohl, H. (2000*a*). Testing for the cointegrating rank of a VAR process with structural shifts. *Journal of Business and Economic Statistics*, **18**, 451–464.

Saikkonen, P. and Lütkepohl, H. (2000*b*). Trend adjustment prior to testing for the cointegrating rank of a vector autoregressive process. *Journal of Time Series Analysis*, **21**, 435–456.

Saikkonen, P., Lütkepohl, H., and Trenkler, C. (2006). Break Date Estimation For Var Processes With Level Shift With An Application To Cointegration Testing. *Econometric Theory*, **22**, 15–68.

Seo, Myunghwan (2006). Bootstrap testing for the null of no cointegration in a threshold vector error correction model. *Journal of Econometrics*, **134**(1), 129–150.

Seo, Myung Hwan (2011). Estimation of nonlinear error correction models. *Econometric Theory*, 201–234.

Stock, J. H. and Watson, M. W. (1993). A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems. *Econometrica*, **61**, 783–820.

Trapani, Lorenzo and Whitehouse, Emily (2020). Sequential monitoring for cointegrating regressions. *arXiv preprint arXiv:2003.12182*.

Trenkler, C., Saikkonen, P., and Lütkepohl, H. (2008). Testing for the Cointegrating Rank of a VAR Process with Level Shift and Trend Break. *Journal of Time Series Analysis*, **29**, 331–358.

Wagner, M. and Wied, D. (2017). Consistent monitoring of cointegrating relationships: The us housing market and the subprime crisis. *Journal of Time Series Analysis*, **38**(6), 960–980.

Wang, Qiying (2015). *Limit theorems for nonlinear cointegrating regression.* World Scientific.

Westerlund, J. and Edgerton, D. L. (2007). New improved tests for cointegration with structural breaks. *Journal of Time Series Analysis*, **28**, 188–224.

Yamamoto, Y. (2016). A modified confidence set for the structural break date in linear regression models. *Econometric Reviews*, forthcoming.

Zeileis, A., Leisch, F., Kleiber, C., and Hornik, K. (2005). Monitoring structural change in dynamic econometric models. *Journal of Applied Econometrics*, **20**(1), 99–121.