

Структурные сдвиги и единичные корни: различение моделей нестационарности временных рядов *

Бродский Б.Е.

1. Введение

Одной из наиболее сложных проблем современной эконометрической теории нестационарных временных рядов является различение моделей структурного сдвига и единичного корня. Как замечает Перрон (2005) в обзоре современного состояния этой теории, присутствие структурного сдвига резко снижает мощность известных тестов на единичные корни (например, ADF (Augmented Dickey-Fuller test) Dickey, Fuller (1979, 1981)), KPSS (Kwiatkowski-Phillips-Schmidt-Shin (1992))), и наоборот, тестирование структурного сдвига в моделях с единичным корнем крайне затруднено.

Причина этой ситуации, по мнению Перрона, состоит в том, что большинство статистических тестов, предназначенных для решения задач обнаружения единичного корня и структурного сдвига в эконометрических данных, обладают выраженной статистической "сенситивностью" ко всем этим видам нестационарности наблюдений. Проблема в том, чтобы предложить тесты, избирательно реагирующие на тот или иной тип нестационарности данных, или тесты "проявляющие" различные "отпечатки" этих видов нестационарности. На сегодняшний день эти "сенситивные тесты" не разработаны (см. Перрон (2005)).

Вместе с тем проблема различения различных видов нестационарности наблюдений продолжает вызывать значительный исследовательский интерес. По сути дела, она стала актуальной уже после знаменитой статьи Nelson, Plosser (1982), в которой на 14 макроэкономических временных рядах в США было продемонстрировано, что ADF-тест не позволяет отвергнуть гипотезу о наличии стохастического тренда в данных. Эти выводы были подвергнуты критике в работе Rappoport, Reichlin (1989) и в особенности в статьях Perron (1989, 1994), в которых было показано,

*Работа поддержана грантом РФФИ 07-06-00181

что присутствие структурного сдвига в данных чаще всего не позволяет эффективно диагностировать наличие стохастического тренда в полученной выборке.

Работы Перрона (1989,1994) базировались на следующей спецификации регрессионной модели временного ряда:

$$y_t = \mu + \beta t + \alpha y_{t-1} + \sum_{i=1}^k c_i \Delta y_{t-i} + \epsilon_t.$$

Если в данных присутствует структурный сдвиг по уровню (μ), то оценка авторегрессионного коэффициента α в асимптотике сдвинута к единице. Это означает, что ADF-тест с модифицированным порогом по MacKinnon (1980) будет диагностировать наличие стохастического тренда в данных.

В последующих работах Montanes, Reyes (1999,2000) было показано, что аналогичные проблемы возникают с тестом Phillips-Perron (1988) на единичные корни.

Ситуация еще более усложняется, если моменты структурных сдвигов в выборке данных неизвестны. В работах Christiano (1992), Banerjee et al. (1992), Zivot, Andrews (1992), Perron, Vogelsang (1992), Perron (1997), было показано, что в этой ситуации мощность известных тестов на единичные корни резко снижается. Некоторые предварительные подходы к решению этой проблемы были намечены в современных работах Perron, Yabu (2005), Kim, Perron (2005).

В целом, следует признать, что проблема эффективного различения нестационарностей типа единичного корня и структурного сдвига не получила окончательного решения в общем контексте неизвестных моментов структурных сдвигов и статистической зависимости наблюдений.

В данной работе исследуется статистическая эффективность двухшаговой процедуры различения нестационарности стохастического тренда и единичного корня, основанной на модифицированной статистике Колмогорова-Смирнова. На первом шаге предлагаемой процедуры тестируется гипотеза стационарности выборки против совместной альтернативы нестационарности наблюдений (структурный сдвиг или единичный корень). На втором шаге процедуры в случае принятия гипотезы нестационарности выборки тестируется гипотеза наличия структурного сдвига (нескольких структурных сдвигов) против альтернативы единичного корня. Показано, что вероятности ошибочной классификации этих видов нестационарности данных в рамках предложенной процедуры экспоненциально стремятся к нулю с увеличением объема выборки.

2. Постановка задачи

Далее будут рассмотрены следующие модели наблюдений на вероятностном пространстве $(\Omega, \mathcal{F}, P)$:

H_0 : модель стационарного временного ряда:

$$e_i = \rho e_{i-1} + u_i, \quad e_0 \equiv 0, \quad |\rho| < 1, \quad i = 1, \dots, N, \quad (1)$$

где u_i - последовательность независимых одинаково распределенных наблюдений с $Eu_i = 0$, $Eu_i^2 = \sigma^2$. Через E обозначим математическое ожидание по мере P . Процесс наблюдений: $y_i = e_i$.

Далее будем предполагать, что для последовательности u_i выполнено равномерное условие Крамера: существует константа $H > 0$ такая, что для всех i

$$Ee^{tu_i} < \infty, \quad |t| < H.$$

В силу центрированности последовательности u_i , это условие эквивалентно следующему: существуют числа $T > 0$, $g > 0$ такие, что для всех i

$$Ee^{tu_i} \leq \exp\left(\frac{1}{2}\sigma^2 gt^2\right), \quad |t| < T. \quad (2)$$

H_1 : модель структурного сдвига:

Процесс "состояния" e_i , $i = 1, 2, \dots$ описывается уравнением (1), а процесс наблюдений равен:

$$y_i = hI\{i \geq [\theta N]\} + e_i.$$

где $I(A)$ - характеристическая функция множества A ; $0 < \beta \leq \theta \leq 1 - \beta < 1$, $h \neq 0$, а момент структурного сдвига θ неизвестен.

H_2 : модель единичного корня:

$$\begin{aligned} e_i &= e_{i-1} + u_i, \quad e_0 \equiv 0, \quad i = 1, \dots, N \\ y_i &= e_i. \end{aligned}$$

Основная статистика, которая будет использоваться для тестирования гипотез H_0, H_1, H_2 , имеет вид:

$$Z_N(n) = \frac{1}{N^2} \left(n \sum_{j=1}^N y_j - N \sum_{j=1}^n y_j \right), \quad n = 1, \dots, N.$$

Зафиксируем параметр $0 < \alpha < 1/2$. Решающее правило, основанное на статистике

$$T_N = \max_{[\alpha N] \leq n \leq [(1-\alpha)N]} |Z_N(n)|$$

и пороговой границе $0 < C$, имеет вид:

$$d_N = \begin{cases} 0, & \text{если } T_N \leq C \text{ принимается гипотеза } H_0 \\ 1, & \text{если } T_N > C \text{ принимается совместная альтернатива } H_1 \& H_2 \end{cases}$$

На этом заканчивается первый этап метода. На втором этапе происходит различение альтернатив H_1 и H_2 .

В ситуации единственного момента структурного сдвига $n_0 = [\theta N]$, где $\alpha \leq \theta \leq (1 - \alpha)$, тестирование гипотез H_1 и H_2 осуществляется следующим образом:

1. Пусть на первом шаге метода выполнено неравенство $T_N > C$ и максимум модуля статистики $Z_N(n)$ достигается в точке $\hat{n}_1$ (в случае нескольких точек максимума $|Z_N(n)|$ выбирается наименьшая из этих точек).

2. На втором шаге фиксируется параметр $0 < \delta < \min(\theta, 1 - \theta)$ и формируются две подвыборки: $X_1(\hat{n}_1) = \{1, \dots, \hat{n}_1 - [\delta N]\}$ и $X_2(\hat{n}_1) = \{\hat{n}_1 + [\delta N], \dots, N\}$.

3. Для каждой из этих подвыборок рассчитываются соответствующие величины $T_N(X_1)$ и $T_N(X_2)$ - максимумы модуля статистики $|Z_N(\cdot)|$, построенной по этим подвыборкам.

4. Рассчитываются пороговые границы $C_1 = C_1(\hat{n}_1 - [\delta N])$ и $C_2 = C_2(N - \hat{n}_1 - [\delta N])$ для подвыборок X_1 и X_2 .

5. Решающее правило на втором шаге метода таково:

$$d_2(N) = \begin{cases} 1, & T_N(X_1) < C_1, \text{ или } T_N(X_2) < C_2 \\ 2, & \text{в противном случае} \end{cases}$$

Значение $d_2(N) = 1$ соответствует принятию гипотезы H_1 , $d_2(N) = 2$ - принятию гипотезы H_2 .

Таким образом, первый этап предлагаемого метода состоит в тестировании гипотезы стационарности наблюдений против совместной альтернативы единичного корня и стохастического тренда. При отклонении гипотезы стационарности выполняется второй этап метода, на котором тестируется гипотеза структурного сдвига против альтернативы единичного корня.

3. Основные результаты

Качество методов тестирования моделей целесообразно оценивать вероятностями ошибок $P_i(H_j)$, $i, j = 0, 1, 2$, $i \neq j$. На первом шаге процедуры тестируется гипотеза стационарности H_0 против совместной альтернативы $H_1 \& H_2$. Поэтому на первом шаге используются вероятности ошибок $P_0\{H_1 \& H_2\}$, $P_1\{H_0\}$, $P_2\{H_0\}$.

Справедлива следующая

Теорема 1.

Пусть $0 < C < \sigma|h|\theta(1 - \theta)$. Тогда вероятности ошибок $P_0\{H_1 \& H_2\}$, $P_1\{H_0\}$, $P_2\{H_0\}$ стремятся к нулю с увеличением объема выборки N .

Более того, при гипотезе H_1 отнормированная оценка момента структурного сдвига $\hat{n}/N$ сходится почти наверно к истинному параметру структурного сдвига θ при увеличении объема выборки N .

Доказательство.

Рассмотрим свойства статистики $Z_N(\cdot)$ при различных альтернативах. Вначале запишем для гипотезы H_0 :

$$e_i = \rho(\rho e_{i-2} + u_{i-1}) + u_i = \dots = \sum_{k=0}^{i-1} u_{i-k} \rho^k.$$

Поэтому

$$\sum_{j=1}^n e_j = \sum_{j=1}^n \sum_{k=0}^{j-1} u_{j-k} \rho^k = \sum_{k=0}^{n-1} \rho^k \sum_{m=1}^{n-k} u_m = \sum_{m=1}^n u_m \sum_{k=0}^{n-m} \rho^k.$$

При гипотезе H_0 статистика $Z_N(i)$ имеет вид:

$$\begin{aligned} Z_N(i) &= (i \sum_{m=1}^N u_m \frac{1 - \rho^{N-m+1}}{1 - \rho} - N \sum_{m=1}^i u_m \frac{1 - \rho^{i-m+1}}{1 - \rho}) / N^2 \\ &= \frac{1}{(1 - \rho)N^2} (\sum_{m=1}^i u_m [i(1 - \rho^{N-m+1}) - N(1 - \rho^{i-m+1})] + \sum_{m=i+1}^N u_m i(1 - \rho^{N-m+1})). \end{aligned}$$

В силу независимости u_i :

$$E_0 Z_N^2(i) \leq \frac{\sigma^2}{(1 - \rho)^2 N^4} (\sum_{m=1}^i (N - i)^2 + \sum_{m=i+1}^N i^2) = \sigma^2 \frac{i(N - i)}{N^3(1 - \rho)^2}$$

Поэтому

$$\max_i E_0 Z_N^2(i) \leq O(\frac{1}{N}).$$

С использованием условия Крамера для последовательности u_i отсюда получим:

$$P_0\{H_1 \& H_2\} = P_0\{\max_n |Z_N(n)| > C\} \leq \begin{cases} \exp(-\frac{(1 - \rho)^2 C^2 N}{2\sigma^2 g}), & 0 < C < NTg\sigma^2 \\ \exp(-\frac{(1 - \rho)CNT}{2}), & C \geq NTg\sigma^2. \end{cases}$$

При гипотезе H_2 статистика $Z_N(i)$ принимает вид:

$$\begin{aligned} Z_N(i) &= (i \sum_{m=1}^N u_m(N-m+1) - N \sum_{m=1}^i u_m(i-m+1))/N^2 \\ &= (\sum_{m=1}^i u_m(N-i)(m-1) + \sum_{m=i+1}^N u_m i(N-m+1))/N^2. \end{aligned}$$

Поэтому при гипотезе H_2 в силу независимости u_i имеем:

$$\begin{aligned} E_2 Z_N^2(i) &= \frac{1}{N^4} \left(\sum_{m=1}^i (N-i)^2(m-1)^2 + \sum_{m=i+1}^N i^2(N-m+1)^2 \right) \\ &= \frac{1}{N^4} \left((N-i)^2 \sum_{k=0}^{i-1} k^2 + i^2 \sum_{k=1}^{N-i} k^2 \right) = \frac{i(N-i)}{6N^3} (2i(N-i) + 1) \end{aligned}$$

Поэтому существует предел

$$\lim_{N \rightarrow \infty} \frac{1}{[Nt]} E_2 Z_N^2([Nt]) = \frac{t(1-t)^2}{3} = \sigma_t^2.$$

С использованием результата Peligrad (1992) имеем:

$$U_t = \frac{1}{\sigma_t \sqrt{N}} Z_N([Nt]) \rightarrow W(t),$$

где $W(t)$ - винеровский процесс на отрезке $[0, 1]$.

Отсюда получим для вероятности ошибки

$$\begin{aligned} P_2\{H_0\} &= P_2\left\{ \max_{[\alpha N] \leq n \leq [(1-\alpha)N]} |Z_N(n)| < C \right\} \\ &\leq P_2\left\{ \max_{\alpha \leq t \leq 1-\alpha} \frac{1}{\sigma_t \sqrt{N}} |Z_N(n)| < \frac{C}{\alpha \sqrt{N}} \left(\frac{3}{1-\alpha} \right)^{1/2} \right\} \rightarrow 0 \quad \text{при } N \rightarrow \infty. \end{aligned}$$

При гипотезе H_1 статистика $Z_N(i)$ равна:

$$Z_N(i) = a_\theta\left(\frac{i}{N}\right) + \eta_N(i),$$

где

$$a_\theta(t) = \begin{cases} t(1-\theta)h, & 0 \leq t \leq \theta \\ \theta(1-t)h, & \theta < t \leq 1, \end{cases}$$

$$\eta_N(i) = \frac{1}{(1-\rho)N^2} \left(\sum_{m=1}^i u_m [i(1-\rho^{N-m+1}) - N(1-\rho^{i-m+1})] + \sum_{m=i+1}^N u_m i(1-\rho^{N-m+1}) \right).$$

Поэтому при $C(N) < \sigma|h|\theta(1 - \theta)$ получим для вероятности ошибки

$$P_1\{H_0\} = P_1\{\max_i |Z_N(i)| \leq C\} \leq P_1\{\max_i |\eta_N(i)| \geq \sigma|h|\theta(1 - \theta) - C\}.$$

Обозначим $\kappa = \sigma|h|\theta(1 - \theta) - C$. Заметим, что для последней вероятности справедлива экспоненциальная оценка сверху, аналогичная доказанной выше для гипотезы H_0 :

$$P_1\{\max_i |\eta_N(i)| \geq \kappa\} \leq \begin{cases} \exp(-\frac{(1 - \rho)^2 \kappa^2 N}{2\sigma^2 g}), & 0 < \kappa < NTg\sigma^2 \\ \exp(-\frac{(1 - \rho)\kappa NT}{2}), & \kappa \geq NTg\sigma^2. \end{cases}$$

Таким образом, все вероятности ошибок для предложенного теста стремятся к нулю с увеличением объема выборки N .

Докажем теперь второе утверждение теоремы. Пусть $\hat{n}$ - оценка момента структурного сдвига, полученная на этапе 1 метода. Тогда для любого $0 < \epsilon < \alpha$ рассмотрим событие $\{|\frac{\hat{n}}{N} - \theta| > \epsilon\}$. Имеем

$$P_1\{|\frac{\hat{n}}{N} - \theta| > \epsilon\} \leq P_1\{\max_t |\eta_N(t)| > \frac{1}{2}\epsilon|h|\min(\theta, 1 - \theta)\}.$$

Пусть $\kappa_0 = \frac{1}{2}\epsilon|h|\min(\theta, 1 - \theta)$. Для вероятности в правой части справедлива экспоненциальная оценка сверху, доказанная выше:

$$P_1\{\max_t |\eta_N(t)| > \kappa_0\} \leq \begin{cases} \exp(-\frac{(1 - \rho)^2 \kappa_0^2 N}{2\sigma^2 g}), & 0 < \kappa_0 < NTg\sigma^2 \\ \exp(-\frac{(1 - \rho)\kappa_0 NT}{2}), & \kappa_0 \geq NTg\sigma^2. \end{cases}$$

Из этой экспоненциальной оценки следует, что отнормированная оценка момента структурного сдвига $\hat{n}/N$ сходится почти наверно к истинному параметру структурного сдвига θ при увеличении объема выборки N .

Теорема 1 доказана.

На втором этапе метода происходит различение гипотез H_1 и H_2 . На содержательном уровне возможность различения этих гипотез основана на различных свойствах решающей статистики $Z_N(\cdot)$ при гипотезе структурного сдвига и единичного корня.

Справедлива следующая

Теорема 2.

Вероятности ошибочных решений на втором шаге метода стремятся к нулю при увеличении объема выборки N .

Идея доказательства теоремы 2 состоит в следующем. Из теоремы 1 следует, что нормированная оценка момента структурного сдвига $\hat{n}_1/N$ стремится почти наверно к истинному параметру структурного сдвига θ с увеличением объема выборки N . Поэтому для любого $0 < \delta < \min(\theta, 1 - \theta)$ эта отнормированная оценка попадает в δ -окрестность параметра θ при достаточно большом N .

Поэтому две подвыборки: $X_1(\hat{n}_1) = \{1, \dots, \hat{n}_1 - [\delta N]\}$ и $X_2(\hat{n}_1) = \{\hat{n}_1 + [\delta N], \dots, N\}$, формируемые на втором этапе метода, в ситуации единственного структурного сдвига с вероятностью, стремящейся к единице, будут классифицироваться как статистически однородные (гипотеза H_0).

Вместе с тем при гипотезе единичного корня подвыборки $X_1(\hat{n}_1) = \{1, \dots, \hat{n}_1 - [\delta N]\}$ и $X_2(\hat{n}_1) = \{\hat{n}_1 + [\delta N], \dots, N\}$ сохраняют свойство статистической неоднородности (гипотеза H_2). В этом и состоит содержательная идея этапа 2 метода: для каждой из этих подвыборок рассчитываются соответствующие величины $T_N(X_1)$ и $T_N(X_2)$ - максимумы модуля статистики $|Z_N(\cdot)|$, построенной по этим подвыборкам. Далее рассчитываются пороговые границы $C_1 = C_1(\hat{n}_1 - [\delta N])$ и $C_2 = C_2(N - \hat{n}_1 - [\delta N])$ для подвыборок X_1 и X_2 .

Решающее правило на втором шаге метода таково:

$$d_2(N) = \begin{cases} 1, & T_N(X_1) < C_1, \text{ или } T_N(X_2) < C_2 \\ 2, & \text{в противном случае} \end{cases}$$

Из теоремы 1 следует, что для этого решающего правила вероятности ошибочных решений будут стремиться к нулю при увеличении объема выборки N .

Предложенный метод легко обобщается на случай нескольких моментов структурных сдвигов.

4. Эксперименты

Имитационное моделирование

Целью имитационных экспериментов была проверка свойств предложенного теста и качества различения основных рассматриваемых гипотез - стационарности, структурного сдвига и единичного корня - для конечных объемов выборки. При имитационном моделировании рассматривались следующие гипотезы и модели:

- H_0 - гипотеза стационарности наблюдений:

$$\begin{aligned} x_i &= \rho x_{i-1} + \sigma v_i, \quad v_i \sim \mathcal{N}(0, 1) \\ y_i &= x_i, \quad i = 1, \dots, N, \end{aligned}$$

где N - объем выборки; $-1 < \rho < 1$, $\sigma > 0$; $\mathcal{N}(0, 1)$ - стандартное нормальное распределение;

- H_1 - гипотеза структурного сдвига по среднему:

$$y_i = x_i + hI(i \leq [\theta N]), \quad i = 1, \dots, N;$$

где $0 < \theta < 1$, $h > 0$;

- H_2 - гипотеза единичного корня:

$$y_i = y_{i-1} + \sigma v_i, \quad i = 1, \dots, N.$$

При тестировании этих гипотез все параметры этих моделей, а именно, σ, ρ, h, θ - предполагались неизвестными. Предполагалось лишь, что $\theta \geq \theta_0 = 0.1$, $-0.99 \leq \rho \leq 0.9$.

Оценивание коэффициента авторегрессии ρ осуществлялось следующим образом. Вначале рассчитывались эмпирические оценки дисперсии и автоковариации 1-го порядка: для $n = [\theta_0 N]$

$$\begin{aligned} r(0) &= \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1} \\ r(1) &= \frac{\sum_{i=2}^n (y_i - \bar{y})(y_{i-1} - \bar{y})}{n-2}, \end{aligned}$$

где $\bar{y} = \sum_{i=1}^n y_i / n$, $0 < \theta_0 < \theta < 1$.

Тогда оценка коэффициента авторегрессии равна $\hat{\rho} = r(1)/r(0)$. Оцен-

ка дисперсии $\hat{\sigma}^2 = \frac{1}{n-2} \sum_{i=2}^n (y_i - \hat{\rho} y_{i-1})^2$.

При этих оценках рассчитывалось значение пороговой границы:

$$C(N) = 1.3581 \frac{\hat{\sigma}}{\sqrt{N}} \frac{1 + \frac{|\hat{\rho}|}{4}}{1 - \hat{\rho}},$$

где значение 1.3581 соответствует доверительной вероятности 0.95. Во избежание недоразумений, подчеркнем, что этот выбор пороговой границы, зависящей от объема выборки N , обеспечивает постоянство вероятности ошибки 1-го рода на уровне 0.05.

В экспериментах рассчитывались оценки следующих характеристик качества тестирования:

на первом шаге метода:

- вероятность ошибки 1-го рода: $\alpha = P_0\{T_N > C(N)\}$,

- вероятность ошибки 2-го рода: $\beta = P_{12}\{T_N \leq C(N)\}$,

на втором шаге метода:

- вероятности перепутывания гипотез H_1 и H_2 : $\gamma = P_1\{H_2\}$, $\delta = P_2\{H_1\}$.

Оценки вероятностей $\alpha, \beta, \gamma, \delta$ в Таблице 1 рассчитывались как средние в $k = 5000$ повторениях имитационных экспериментов.

В экспериментах были получены следующие результаты:

Таблица 1.

N		100	200	300	500	700	1000
$h = 1$	α	0.114	0.056	0.04	0.072	0.056	0.038
$\rho = 0.3$	β	0.412	0.282	0.106	0.03	0	0
$\sigma = 0.5$	γ	0.05	0.054	0.06	0.04	0.04	0.028
$\theta = 0.5$	δ	0.496	0.56	0.52	0.34	0.172	0.090
$h = 0.5$	α	0.082	0.062	0.052	0.040	0.038	0.048
$\rho = 0.3$	β	0.64	0.23	0.082	0.026	0.006	0
$\sigma = 0.5$	γ	0.05	0.052	0.058	0.044	0.03	0.024
$\theta = 0.5$	δ	0.50	0.566	0.568	0.354	0.168	0.072
$h = 0.5$	α	0.26	0.124	0.064	0.072	0.038	0.044
$\rho = 0.7$	β	0.95	0.86	0.52	0.35	0.136	0.04
$\sigma = 0.5$	γ	0.044	0.026	0.028	0.012	0.018	0.020
$\theta = 0.5$	δ	0.468	0.524	0.518	0.308	0.19	0.090
$h = 0.5$	α	0.046	0.038	0.058	0.032	0.026	0.030
$\rho = -0.7$	β	0.37	0.29	0.088	0.026	0.004	0
$\sigma = 0.5$	γ	0.024	0.032	0.058	0.024	0.024	0.032
$\theta = 0.5$	δ	0.494	0.528	0.502	0.346	0.176	0.084

Полученные результаты требуют нескольких комментариев. Во-первых, вероятность ошибки первого рода α была выбрана на уровне 5 процента, чем и объясняются стабильно низкие значения этого показателя в таблице. Вероятность ошибки 2-го рода β , как видно из таблицы, быстро убывает при возрастании объема выборки. В таблице 1 приведены выборки данных с отношением сигнал/шум 1-2, что является довольно низким показателем. Но даже в этих сложных условиях тестирования удовлетворительное качество наблюдается при объемах выборки порядка 500 наблюдений. В практических задачах отношение сигнал/шум, как правило, гораздо выше (10-100), что позволяет достигать удовлетвори-

тельного качества тестирования при объемах выборки 100-150 наблюдений.

Во-вторых, качество тестирования сильно зависит от степени коррелированности наблюдений. Удовлетворительное качество тестирования при высоких значениях коэффициента автокорреляции (0.7-0.8) достигается при больших объемах выборки данных (800-1000).

На втором этапе метода происходит различение гипотез структурного сдвига и единичного корня. Из данных, приведенных в таблице, видно, что вероятность ошибки при тестировании гипотезы структурного сдвига (γ), как правило, довольно мала (3-5 процентов), тогда как вероятность ошибки при тестировании гипотезы единичного корня несколько выше, но также быстро убывает при возрастании объема выборки.

Практические применения

Предложенный метод был использован в задачах проверки стационарности временных рядов эконометрических наблюдений. В качестве этих временных рядов были выбраны важнейшие макроэкономические индикаторы по России за период 1994-2007 гг., измеренные с месячным интервалом.

Следует отметить, что преимущества предложенного метода проявляются начиная с объемов выборки 100-150 наблюдений. Поэтому попытки использования данного метода для тестирования стационарности наблюдений при объемах выборки 20-50 точек будут в большинстве случаев давать некорректные результаты.

В экспериментах были использованы следующие динамические ряды в диапазоне 1994.1-2007.9:

CPI - месячный индекс инфляции на потребительском рынке (в процентах);

M2 - денежный агрегат М2 (трлн.руб);

E - официальный обменный курс доллара (руб./долл.);

PEL - месячный индекс тарифов на электроэнергию для конечных потребителей (в процентах);

pi - темп инфляции $pi = CPI/100 - 1$;

eps - темп изменения курса доллара $eps = E/E(-1) - 1$;

mi - темп изменения денежного агрегата М2 $mi = M2/M2(-1) - 1$;

$piel$ - темп изменения тарифов на электроэнергию для конечных потребителей $piel = PEL/100 - 1$;

$pcit$ - уровень инфляции (базовый индекс) $pcit = pcit(-1) * CPI/100$;

$pcel$ - уровень тарифов на электроэнергию для конечных потребителей $pcel = pcel(-1) * PEL/100$;

$lpcum$ - логарифм уровня инфляции $lpcum = \log(pcum)$;
 $lpcel$ - логарифм уровня тарифов на электроэнергию $lpcel = \log(pcel)$;
 $lm2$ - логарифм денежного агрегата M2 $lm2 = \log(M2)$;
 lE - логарифм обменного курса доллара $lE = \log(E)$;
 $Dlpcum$ - приращение логарифма уровня цен $Dlpcum = lpcum - lpcum(-1)$;
 $Dlpcel$ - приращение логарифма уровня тарифов на электроэнергию $Dlpcel = lpcel - lpcel(-1)$;
 DlE - приращение логарифма обменного курса доллара $DlE = lE - lE(-1)$.

Проверка этих динамических рядов на стационарность с использованием расширенного теста Дики-Фуллера (ADF) с тремя лагами и без включения константы и тренда дала следующие результаты:

- ряды $CPI, M2, E, PEL, pcum, pcel, lpcum, lpcel, lm2, lE$ имеют порядок нестационарности $I(1)$ с доверительным уровнем 95 процентов;
- ряды $pi, eps, mi, piel, Dlpcum, Dlpcel, DlE$ стационарны с доверительным уровнем 95 процентов.

Следует отметить, что проверка на наличие структурного сдвига в стандартных эконометрических пакетах для отдельных рядов, как правило, не проводится.

Исследование данных динамических рядов на стационарность с использованием предложенного в статье метода дало следующие результаты:

- ряды $pi, eps, mi, piel, Dlpcum, Dlpcel, DlE$ являются стационарными при уровне ошибки 5 процентов. Таким образом, получено полное совпадение результатов с тестом ADF;
- ряды $CPI, M2, E, PEL, pcum, pcel, lpcum, lpcel, lm2, lE$ являются нестационарными при уровне ошибки 5 процентов. Из этих рядов только $pcum, lpcum$ классифицируются по типу структурного сдвига (в точке 1998.9); для остальных рядов принимается гипотеза наличия единичного корня.

5. Выводы

В работе предложен метод обнаружения и тестирования статистических нестационарностей типа структурного сдвига и единичного корня для одномерных временных рядов. Метод основан на модифицированной статистике Колмогорова-Смирнова и представляет собой двухшаговую процедуру: на первом шаге тестируется гипотеза стационарности временного ряда против совместной альтернативы неизвестного структурного сдвига и единичного корня; на втором этапе при отклонении гипотезы

стационарности тестируется гипотеза структурного сдвига против альтернативы единичного корня.

Для предложенного метода в работе доказаны результаты об экспоненциальной сходимости к нулю вероятностей ошибочных решений в зависимости от возрастающего объема выборки. Для конечных выборок данных свойства предложенной двухшаговой процедуры исследованы методом имитационного моделирования.

В заключение попытаемся дать ответ на некоторые "наивные" вопросы, которые неизбежно возникают при чтении этой статьи:

- почему предложенный метод эффективнее широко известных в "узких" эконометрических кругах процедур DF, ADF (Dickey-Fuller), KPSS тестирования единичных корней и стохастических трендов в динамических рядах эконометрических наблюдений? А также метода Chow (1960), метода CUSUM (Ploberger, Kramer (1992)) и др. для тестирования структурных сдвигов?

- в работе рассматривается довольно частная модель стационарного временного ряда:

$$e_i = \rho e_{i-1} + u_i, \quad e_0 \equiv 0, \quad |\rho| < 1, \quad i = 1, \dots, N,$$

где u_i - последовательность независимых одинаково распределенных наблюдений с $Eu_i = 0$, $Eu_i^2 = \sigma^2$.

Зависят ли результаты статьи от выбора именно этой модели?

Отвечая на первый вопрос, вновь подчеркнем, что предложенный метод предназначен не только для тестирования гипотезы стационарности наблюдений против альтернатив единичного корня и структурного сдвига, но также для различения гипотез единичного корня и структурного сдвига в случае отклонения нулевой гипотезы о стационарности наблюдений.

Ответ на второй вопрос короче: выбранная модель может обобщаться до произвольной модели ARMA (авторегрессии-скользящего среднего), в том числе с добавлением константы. Общие результаты статьи инвариантны к выбору этой модели.

Литература

Banerjee, A., Lumsdaine, R.L., Stock, J.H. (1992) Recursive and sequential tests of the unit root and trend-break hypothesis: theory and international evidence. *Journal of Business and Economic Statistics*, 10, 271-287.

Brodsky, B., Darkhovsky, B. (2000) *Non-parametric Statistical Diagnosis: Problems and Methods*, KLUWER, Dordrecht.

Chow G.C. (1960) Tests of equality between sets of coefficients in two linear regressions. *Econometrica*, 28, 591-605.

Christiano, L.J. (1992) Searching for breaks in GDP. *Journal of Business and Economic Statistics*, 10, 237-250

Dickey, D.A., Fuller, W.A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74, 427-431.

Dickey, D.A., Fuller, W.A. (1981) Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 49, 1057-1072.

Kim, D., Perron, P. (2005) Unit root tests with a consistent break fraction estimator. Department of Economics, Boston University.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., 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.

MacKinnon, J.B. (1996) Numerical distribution functions for unit roots and cointegration tests. *J. of Applied Econometrics*, 11, 601-618.

Montanes, A., Reyes, M. (1999) The asymptotic behavior of the Dickey-Fuller tests under the crash hypothesis. *Statistics and Probability Letters*, 48, 401-409.

Montanes, A., Reyes, M. (2000) Structural breaks, unit roots and methods for removing the autocorrelation pattern. *Statistics and Probability Letters*, 48, 401-409.

Nelson, C.R., Plosser, C.I. (1982) Trends and random walks in macroeconomic time series: some evidence and implications. *Journal of Monetary Economics*, 10, 139-162.

Peligrad M. (1992) On the central limit theorem for weakly dependent sequences with a decomposed strong mixing coefficient. *Stochastic Processes and their Applications*, v.42, pp.181-193.

Perron, P. (1989) The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57, 1361-1401.

Perron, P. (1994) Trend, unit root, and structural change in macroeconomic time series. In: *Cointegration for the Applied Economist*, Rao, B.B. (ed.), Basingstoke: Macmillan Press, 113-146.

Perron (2005) Dealing with structural breaks. *Palgrave Handbook of Econometrics*, vol.1, *Econometric Theory*.

Perron, P., Vogelsang T. (1992) Nonstationarity and level shifts with an application to purchasing power parity. *Journal of Business and Economic Statistics*, 10, 301-320.

Perron, P. (1997) Further evidence from breaking trend functions in macroeconomic variables. *Journal of Econometrics*, 80, 55-385.

Perron, P., Yabu, T. (2005) Testing for shifts in trend with an integrated or stationary noise component. Department of Economics, Boston University.

Phillips, P.C.B., Perron P. (1988) Testing for a unit root in a time series regression. *Biometrika*, 75, 335-346.

Ploberger W., Kramer W. (1992) The CUSUM test with OLS residuals. *Econometrica*, 60, 271-285.

Rappoport, P., Reichlin, L. (1989) Segmented trends and non-stationary time series. *Economic Journal*, 99, 168-177.

Zivot, E., & Andrews, D. (1992). Further evidence on the Great crash, the oil price shock and the Unit root hypothesis. *Journal of Business and Economic Statistics*, 10, 251–287.