

## ТЕСТ ПРИЧИННОСТИ ПО ГРЕНДЖЕРУ В ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ МОДЕЛЯХ НА ОСНОВЕ ПАНЕЛЬНЫХ ДАННЫХ

© 2009 А.А. Сняков\*

**Ключевые слова:** тест причинности по Гренджеру в панельных данных, векторные авторегрессионные модели, динамические модели на основе панельных данных, метод инструментальных переменных, обобщенный метод моментов.

Рассматриваются особенности теста причинности по Гренджеру в линейных моделях на основе панельных данных, обусловленные динамической спецификацией зависимой переменной и малым количеством наблюдений во времени, характерном для большинства микроэкономических моделей. Представлен обзор существующих эконометрических методов, учитывающих данные особенности при оценивании векторных авторегрессионных моделей на основе панельных данных и альтернативные процедуры реализации теста.

Тест причинности по Гренджеру является одним из этапов<sup>1</sup> исследования оценок коэффициентов векторной авторегрессионной модели и формулирования выводов относительно динамических взаимосвязей в системе временных рядов. К настоящему времени в эконометрике сформулированы определения причинности, на этой основе предложен ряд тестов причинности, которые широко используются в прикладных работах<sup>2</sup>. Современные исследования ведутся в направлении формулировки тестов в нелинейных моделях<sup>3</sup>, приложениях с ограниченными зависимыми переменными<sup>4</sup>, моделях с долгосрочной причинностью<sup>5</sup>, моделях на основе панельных данных<sup>6</sup>.

Существует большое количество прикладных эконометрических гипотез, тестируемых на панельных данных. Некоторые из них допускают динамическую спецификацию. Это обстоятельство делает актуальной разработку процедуры тестирования причинности в динамических панелях.

Первые шаги в направлении соответствующей методологии были сделаны в 1981 году Andersen и Hsiao<sup>7</sup>, и в настоящее время выработано два подхода к тесту Гренджера в панельных данных.

Первый подход исходит из однородности индивидов (фирм) относительно их реакции на изменение объясняющих переменных, а индивидуальные особенности учитываются через введение фиксированных или случайных эффектов в константу. Так, описанный в

работе Holtz-Eakin (1988)<sup>8</sup> тест Гренджера основан на модели

$$y_{it} = a_0 + \sum_{l=1}^m \alpha_l * y_{it-l} + \sum_{l=1}^m \delta_l * x_{it-l} + \varphi * f_i + u_{it}, \quad (1)$$

где  $f_i$  - ненаблюдаемый индивидуальный эффект. Ошибка характеризуется стандартными свойствами.  $i = 1, 2, \dots, N$ ;  $t = 1, 2, 3, \dots, T$ . Тест состоит в применении теста Wald к проверке линейного ограничения:

$$H_0 : \delta' = 0.$$

Другой подход основан на индивидуальности реакции каждой фирмы на изменение объясняющих переменных, при этом ставится задача оценить модель Hurlin (2004)<sup>9</sup>:

$$y_{it} = a_0 + \sum_{l=1}^m \alpha_{li} * y_{it-l} + \sum_{l=1}^m \delta_{li} * x_{it-l} + \varphi_i * f_i + u_{it}. \quad (2)$$

Преимущества первой спецификации очевидны: увеличение числа степеней свободы при оценивании коэффициентов, простая интерпретация коэффициентов, простота интерпретации результатов теста причинности.

Включение эндогенной зависимой переменной создает ряд проблем, которые, должны быть учтены для получения несмещен-

\* Сняков Андрей Александрович, аспирант Саратовского государственного социально-экономического университета.

ных и состоятельных оценок. Во-первых, независимо от характера индивидуальных эффектов (фиксированные или случайные) OLS оценки уравнения (1) смещены. Если индивидуальные эффекты случайны - регрессор коррелирован с ошибкой. Действительно:

$$y_{it} = a_0 + \alpha y_{it-1} + \beta' * X_{it} + \eta_i + \varepsilon_{it}, \quad (3)$$

где  $X_{it}$  - вектор экзогенных или predetermined переменных,  $y_{it}$  - зависимая переменная,  $\eta_i$  - случайная величина, отражающая индивидуальные эффекты (фиксированные во времени),  $\varepsilon_{it}$  - шок,  $E\eta_i = E\varepsilon_{it} = 0$ ,

$$E\eta_i\eta_j = 0 \forall i \neq j, \quad E\eta_i\varepsilon_{jt} = 0 \forall i, j, t, \quad \text{Var}(\eta_i) = \sigma_{\eta_i}^2,$$

$$E\varepsilon_{it}\varepsilon_{js} = 0 \forall i \neq j, t \neq s, \quad \text{Var}(\varepsilon_i) = \sigma_{\varepsilon_i}^2.$$

Так как

$$y_{it-1} = a_0 + \alpha y_{it-2} + \beta' * X_{it-1} + \eta_i + \varepsilon_{it-1},$$

то корреляция регрессора  $y_{it-1}$  с элементом ошибки  $\eta_i$  в уравнении (3) очевидна. Эта корреляция не уменьшается ни с ростом  $T$ , ни с ростом  $N$ .

Альтернативой является рассмотрение  $\eta_i$  как фиксированных эффектов. В этом случае естественный метод оценивания LSDV<sup>10</sup>. Но, как установил Nickell (1981)<sup>11</sup>, применение этого метода к динамическим панелям приводит к смещенным оценкам из-за прежней проблемы: корреляции регрессора и ошибки. Причина смещения в используемом методе устранения фиксированных эффектов - взятии временной средней для каждого индивидуального наблюдения. Действительно, суть метода состоит в применении OLS к модифицированному уравнению:

$$y_{it} - \bar{y}_i = \alpha(y_{it-1} - \bar{y}_{i,-1}) + \beta' * (X_{it} - \bar{X}_i) + (\varepsilon_{it} - \bar{\varepsilon}_i) \quad t = 1, T, \quad (4)$$

$$\text{где } \bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}, \quad \bar{y}_{i,-1} = \frac{1}{T-1} \sum_{t=2}^{T-1} y_{it-1},$$

$$\bar{X}_i = \frac{1}{T} \sum_{t=1}^T X_{it}.$$

И хотя в этом случае фиксированные эффекты элиминируются, но  $(y_{it-1} - \bar{y}_{i,-1})$  и

$(\varepsilon_{it} - \bar{\varepsilon}_i)$  оказываются коррелированными. При малом  $T$ , характерном для микроэкономических приложений, величина смещения будет значительной. Для  $T \rightarrow \infty$  LSDV является состоятельным методом, но для  $N \rightarrow \infty$  при фиксированном  $T$ , он не состоятелен.

В качестве способа преодоления этих проблем и состоятельного оценивания при  $N \rightarrow \infty$  и  $T$  фиксированном было предложено несколько методов.

В 1981 г. Andersen и Hsiao<sup>12</sup> (далее АН) предложили дифференцировать данные и представить модель в первых разностях:

$$y_{it} - y_{it-1} = \alpha(y_{it-1} - y_{i,t-2}) + \beta' * (X_{it} - X_{it-1}) + (\varepsilon_{it} - \varepsilon_{it-1}). \quad (5)$$

Так как  $(y_{it-1} - y_{i,t-2})$  коррелированы с  $(\varepsilon_{it} - \varepsilon_{it-1})$  из-за корреляции  $y_{it-1}$  с  $\varepsilon_{it-1}$ , АН использовали метод инструментальных переменных. В качестве инструментов предложили брать или лагированные уровни зависимой переменной  $y_{it-2}$  или лагированные первые разности  $y_{it-2} - y_{i,t-3}$ . Оценка АН состоятельна при фиксированном  $T$ , но не является эффективной, так как не использует все возможные инструменты и не учитывает, что ошибка является процессом скользящей средней первого порядка, а не белым шумом.

В 1991 г. Arellano и Bond<sup>13</sup> (далее: АВ) предложили использовать обобщенный метод моментов (GMM) и увеличили число инструментов, для совершенствования процедуры АН. Увеличение количества инструментов основано на предположении, что вся предшествующая история может быть инструментом для  $Y$ . Так, при  $T = 3$  существует один инструмент:  $y_{i1}$ . При  $T = 4$  существует два инструмента:  $y_{i1}; y_{i2}$  и т.д.

В 1998 г. Blundell и Bond<sup>14</sup> (далее: ВВ) указали на недостатки метода АВ: в случае малого числа наблюдений (относительно  $T$ ) имитации по методу Монте-Карло показывают большую величину смещения из-за того, что лагированные уровни являются слабыми инструментами для первых разностей. ВВ (1998) предложили метод, получивший название system GMM, который состоит в использовании лагированных разностей объясняющих переменных как инструментов для

уравнений в уровнях в дополнение к лагированным уровням как инструментам для уравнения в разностях.

Kiviet<sup>15</sup> (1995, 2001) исследовал асимптотические свойства IV оценок и установил их состоятельность при фиксированном  $T$  и  $N$ . На основе серии имитационных расчетов Kiviet показал, что в конечных выборках (При малом  $T = 4, 5, 7, 10$ ) оценки на основе IV методов оказываются смещенными. Поэтому Kiviet предложили альтернативный метод оценивания коротких динамических панелей, получивший название corrected LSDV. Суть метода состоит в коррекции LSDV оценок уравнения (1) в уровнях на асимптотическую величину смещения, установленную при различных  $T$ . Преимущество метода состоит в том, что наряду с несмещенностью оценок, он обеспечивает их большую эффективность, относительно IV методов, так как использование инструментов всегда менее эффективно использования исходных переменных.

Pesaran и Smith<sup>16</sup> (1995) показали, что если панель не однородна, то применение методологии оценивания при предположении однородности коэффициентов приведет к смещенным оценкам коэффициентов. Причина смещения в случае неоднородности коэффициентов состоит в том, что для истинной модели

$$y_{it} = \alpha_i + \gamma_i y_{it-1} + \beta' * X_{it} + \varepsilon_{it},$$

при оценивании модели

$$y_{it} = \alpha_i + \gamma y_{it-1} + \beta' * X_{it} + \varepsilon_{it}.$$

Разность  $(\gamma_i - \hat{\gamma})y_{it-1}$  будет частью ошибки и эта корреляция приведет к смещению и несостоятельности оценок. Что более важно: смещение не устраняется использованием IV методов и требует либо свести процедуру к применению техники (seemingly unrelated regressions), либо разработать некоторые ограничения, которые бы сохраняли степени свободы, но и допускали разнородность реакции.

Erdil<sup>17</sup> (2000) предложил процедуру последовательного тестирования причинности по Гренджеру в панельных данных (на основе модели

$$y_{it} = \alpha_0 + \sum_{k=1}^p \alpha_k * y_{it-k} + \sum_{k=1}^p \beta_{ik} * x_{it-k} + \varphi * f_i + u_{it} ) \quad c$$

учетом неоднородности. В этом случае, причинность может быть присуща не всем единицам совокупности, а только их отдельной группе.

Альтернативную процедуру, состоящую в наложении частичных ограничений на коэффициенты для сохранения степеней свободы, разработала Weinhold<sup>18</sup> (2000). Метод получил название "MFR-модель". В модели

$$y_{it} = \alpha_i + \gamma_i * y_{it-1} + \beta_i * x_{it} + \varepsilon_{it}.$$

Weinhold предложила считать коэффициенты  $\gamma_i$  как фиксированные, а коэффициенты при  $x_{it}$  как случайные величины, т.е.

$$\beta_i = \beta + \eta_i, \quad E[\eta_i] = 0, \quad Var[\eta_i] = \sigma_\eta^2.$$

Эксперименты по методу Монте-Карло показали, что MFR метод характеризуется существенно меньшим смещением в малых выборках при разнородности панели (даже при  $T = 5$ ) по сравнению с моделями однородных коэффициентов, и обеспечивает более эффективные, чем SUR оценки. Weinhold разработала процедуру тестирования причинности по Гренджеру в динамической разнородной панели, состоящую в определении степени причинности в выборке, посредством расчета доли наблюдений превышающих  $2\sqrt{N * \sigma_\varepsilon^2}$  интервал для коэффициентов при переменной причинности которой тестируется.

Hurlin<sup>19</sup> (2004) предложил в случае неоднородной пространственной структуры модели использовать среднюю статистику Wald из статистик, определяющих тест Гренджера в каждом из временных рядов. В результате того, что средняя статистика при фиксированном  $T$  не имеет стандартного распределения, Hurlin, на основе имитационного моделирования, затабулировал критические значения статистики проверки нулевой гипотезы отсутствия причинности по Гренджеру в  $N$  единицах совокупности для разных  $N$  и  $T$ .

<sup>1</sup> Наряду с получением оценок функций отклика на единичный импульс (Impulse response functions) и разложением дисперсии ошибки прогноза (variance decomposition).

<sup>2</sup> См., например: Sims C.A. Money, Income and causality // The American economic review. Vol. 62. № 4. Sep. 1972. P. 540-552; Toda H.Y., Phillips P.C.B. Vector autoregressions and causality // Econometrica. Vol. 61. № 6. Nov. 1993. P. 1367-1393; Introduction to modern time series analysis by Kirchgassner G., Wolters J. Springer, 2007.

<sup>3</sup> Hiemstra C., Jonathan D.J. Testing for linear and nonlinear Granger causality in the stock price-

volume relation // *The Journal of finance*. Vol. 49. № 5. Dec. 1994. P. 1639-1664.

<sup>4</sup> *Arent J.* Endogeneity and heterogeneity in LDV panel data models. 2001.

<sup>5</sup> *Dufour J.-M., Pelletier D., Renault E.* Short run and long run causality in time series: inference // *CIREQ*. 2003. № 14. P. 1-27.

<sup>6</sup> См.: *Bond S.* Dynamic panel data models: a guide to micro data methods and practice // *CEM-MAP Working papers*. 2002. № 9. P. 1-36; *Hurlin C.* Testing Granger causality in heterogeneous panel data models with fixed coefficients. 2004.

<sup>7</sup> *Anderson T.W., Hsiao Ch.* Estimation of dynamic models with error components // *Journal of American statistical association*. Vol. 76. № 375. Sep. 1981. P. 598-606.

<sup>8</sup> *Holtz-Eakin D., Newey W., Rosen H.* Estimating vector autoregressions with panel data // *Econometrica*. Vol. 56. № 6. Nov. 1988. P. 1371-1395.

<sup>9</sup> *Hurlin C.* Testing Granger causality in heterogeneous panel data models with fixed coefficients. 2004.

<sup>10</sup> См.: Least squares dummy variable (или Within group estimator). Analysis of panel data by *Hsiao Ch.*, Cambridge University press, 1986; *Bun M., Kiviet J.* The accuracy of inference in small samples of dynamic panel data models. Tinbergen institute discussion paper, 2001.

<sup>11</sup> *Nickell S.* Biases in dynamic models with fixed effects // *Econometrica*. Vol. 49. № 6. Nov. 1981. P.1417-1426.

<sup>12</sup> *Anderson T.W., Hsiao Ch.* Estimation of dynamic models with error components // *Journal of American statistical association*. Vol. 76. № 375. Sep. 1981. P. 598-606.

<sup>13</sup> *Arellano M., Bond S.* Some tests of specification for panel data: Monte Carlo evidence and an Application Employment equations // *The review of economic studies*. Vol. 58. № 2. Apr. 1991. P. 277-297.

<sup>14</sup> *Blundell R.W., Bond S.R.* Initial conditions and moment restrictions in dynamic panel data models // *Journal of econometrics*. 1998. Vol. 87. P. 115-143.

<sup>15</sup> См.: *Kiviet J.* On bias, inconsistency, and efficiency of various estimators in dynamic panel data models // *Journal of econometrics*. 1995. Vol. 68. P. 53-78; *Kiviet J.* The accuracy of inference in small samples of dynamic panel data models. Tinbergen institute discussion paper, 2001.

<sup>16</sup> См.: *Pesaran H.M., Smith R.* Estimating long-run relationships from dynamic heterogeneous panels // *Journal of econometrics*. 1995. Vol. 68. P. 78-93.

<sup>17</sup> *Erdil E., Yetkiner H.* A panel data approach for income-health causality. 2000.

<sup>18</sup> *Nair-Reichert N., Weinhold D.* Causality tests for cross-country panels: new look at FDI and Economic growth in developing countries. 2000.

<sup>19</sup> *Hurlin C.* Testing Granger causality in heterogeneous panel data models with fixed coefficients. 2004.