THE GRANGER CAUSALITY TEST IN LINEAR DYNAMIC PANEL-DATA MODELS

© 2009 A.A. Sinyakov*

Keywords: Granger causality test in panel data models, vector autoregressions, dynamic panel data models, instrumental variables' model, generalized method of moments.

This article presents the features of the Granger causality test in panel data models. The features are caused by a dynamic specification of the dependent variable in the left-hand side of the model and short *T* and large *N*. The paper includes some econometrics estimators for vector autoregressions based on panel-data and additional tests.

The Granger causality test is the following stage of estimation of vector autoregression models. Causality definitions have been formulated, and on this basis a number of tests of causality which are widely used in applied works are offered. Modern research is conducted in a direction of the formulation of tests in nonlinear models, appendices with the limited dependent variables, models with long-term causality, models on the basis of panel data.

There is a considerable quantity of applied econometric hypotheses, which are tested in panel data. Some of them suppose the dynamic specification. This circumstance calculates the procedure for causality testing in dynamic panels.

The first step in the direction of corresponding methodology was made in 1981 by Andersen and Hsiao and now two approaches for the testing in panel data have been developed.

The first approach starts with uniformity of individuals (firms) concerning their reaction to change of explaining variables, and specific features are considered through introduction fixed or random effects in the constant term of the model. The test, described in the article of Holtz-Eakin (1988), is based on the model:

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

where f_i - the unobservable individual effect. The error term is characterized by standard properties; $i = 1, 2, ..., N$; $t = 1, 2, 3, ..., T$. The test consists in application of Wald test to check the linear restriction:

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

Other approach starts with individuality of reaction of each firm on change of explaining variables, the problem is thus put to estimate model of Hurlin (2004):

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

Advantages of the first specification are obvious: increase of degrees of freedom, simple interpretation of factors, and simple interpretation of test's results.

Inclusion of endogenous dependent variable creates a number of problems, which should be considered for unbiased estimation. First, irrespective of character of individual effects (fixed or random) OLS estimations of the equation (1) are unbiased. If individual effects are random, then the repressor is correlated with an error. For example:

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

where X_{it} - predetermined variables, y_{it} - a dependent variable, η_i - a random variable reflecting individual effects (fixed in time), ε_{it} - a random variable, $E\eta_i\eta_j = 0 \forall i \neq j$, $E\eta_i\varepsilon_{jt} = 0 \forall i, j, t$, $Var(\eta_i) = \sigma^2_{\eta_i}$, $E\varepsilon_{it}\varepsilon_{js} = 0 \forall i \neq j, t \neq s$, $Var(\varepsilon_i) = \sigma^2_{\varepsilon_i}$.

Because of:

$$y_{it-1} = a_0 + \alpha y_{it-2} + \beta' * X_{it-1} + \eta_i + \varepsilon_{it-1} \,. \quad (3)$$

So, the correlation of y_{it-1} with an element of the error term η_i in the equation (3) is obvious. This correlation decreases neither with growth *T*, nor with growth *N*.

An alternative is consideration of η_i as fixed effects. In this case a natural method of estimation is least squares dummy variable. But, as it was established by Nickell (1981), an application of this method to dynamic panels leads to the biased estimations because of the problem: correlation of a repressor with an error. The reason is, used method of elimination of the fixed effects, a capture of time average for each individual observation. The method consists in

* Andrei A. Sinyakov, post-graduate student of Saratov State Socio-Economical University.

application ordinary least squares to the modified equation:

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

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

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

Although in this case the fixed effects are being eliminated, but $(y_{it-1} - \bar{y}_{i,-1})$ and $(\varepsilon_{it} - \bar{\varepsilon}_i)$ appear to be correlated. With short T the size of bias is considerable. For $T \to \infty$ least squares dummy variable is accurate, but for $N \to \infty$ with fixed T it is biased.

Some methods were offered for overcoming these problems and unbiased estimation for $N \to \infty$ and fixed T.

In 1981 Andersen and Hsiao (further AH) suggested to differentiate data and to present model in the first differences:

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

$(y_{it-1} + y_{i,t-2})$ $\bar{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$

Because of correlation with $(\varepsilon_{it} - \varepsilon_{it-1})$ AH used a instrumental variable method. Instruments could be presented by lagged levels of dependent variable y_{it-2} or lagged first differences ($y_{it-2} - y_{i,t-3}$). Estimations of AH are unbiased with fixed T, but are not effective because method does not use all tools and does not consider that the error is moving average process.

In 1991 Arellano and Bond (further: AB) suggested using the generalized method of moments (GMM) and increased number of instruments, for perfection of procedure AH. The increase in quantity of instruments is based on assumption, that all previous history can be the tool for Y. So for T=3 there is one instrument: , for T = 4 there are two tools: $y_{i1}; y_{i2}$ etc.

In 1998 Blundell and Bond (further: BB) pointed out the defects of AB's method: in case of small number of observations (for T) imitations on a method of Monte-Carlo show the big size of estimator's bias. As a result, legged levels are weak instruments for the first differences of dependent variable. BB (1998) offered the method which was called as system GMM. The method consists of using lagged differences of explaining variables as instruments for the equations in levels in addition to lagged differences as instruments for the equation in differences.

Kiviet (1995, 2001) investigated asymptotic properties of IV estimations and established their consistence for fixed T and N. On the basis of a series of imitating calculations Kiviet concluded, that in limited samples (for small T = 4, 5, 7, 10) estimations on the basis of IV methods appear to be unbiased. Therefore Kiviet offered the alternative estimator for short dynamic panels, which is called corrected LSDV. The method is based on correction LSDV in the equation (1) in levels on the size of bias, established at various T. The advantages of the method are unbiased and more efficient estimations, concerning IV methods.

Pesaran and Smith (1995) found out, that if the panel is not homogeneous, methodology which is based on the assumption of homogeneous of factors would lead to biased estimations. The reason of the bias in case of heterogeneity of factors' influences is explained by the fact, that when true model:

$$y_{it} = a_i + \gamma_i y_{it-1} + \beta^{'} * X_{it} + \varepsilon_{it} \ .$$

And econometrician estimates model:

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

The difference $(\gamma_i - \hat{\gamma}) y_{it-1}$ is a part of an error and this correlation leads to the bias and an inconsistency of estimations. What is more important: bias is not eliminated by IV methods. The procedure either should be reduced to techniques seemingly unrelated regressions or should be developed some restrictions which save degrees of freedom, but also allow heterogeneity.

Erdil (2000) offered procedure of consecutive Granger causality testing in panel data in case of heterogeneity.

The alternative procedure which consists of imposing partial restrictions on factors for preservation of degrees of freedom, was developed by Weinhold (2000). The method has received the name Mixed Fixed-Random Effects Model. In the model:

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

Weinhold suggested considering factors γ_i as fixed, and factors x_{it} as random variables, i.e. $\beta_i = \beta + \eta_i$, $E[\eta_i] = 0,$, $Var[\eta_i] = \sigma^2_\eta$.

Experiments by the Monte-Carlo method shown, that MFR method is characterised by a more smaller bias in small samples with panel heterogeneity (even at $T = 5$) in comparison with models of homogeneous factors, and provides more effective, than SUR estimations. Weinhold developed the procedure for causality testing in the dynamic heterogeneous panel-data model, consisting in definition of degree of causality in sample, by means of calculation of a share of observations exceeding $2\sqrt{N * \sigma^2_\varepsilon}$ interval for factors at a variable which causality is tested for.

Hurlin (2004) suggested using average Wald statistics in case of heterogeneity but the average statistics for fixed T has no standard distribution. Hurlin, on the basis of imitating modelling, tabulated critical values for test statistics.

---

1. *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.

2. *Arent J.* Endogeneity and heterogeneity in LDV panel data models. 2001.

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

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

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

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

7. *Hurlin C.* Testing Granger causality in heterogeneous panel data models with fixed coefficients. 2004.

8. 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.

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

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

11. *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.

12. *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.

13. *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.

14. *Pesaran H.M., Smith R.* Estimating long-run relashionships from dynamic heterogeneous panels // Journal of econometrics. 1995. Vol. 68. P. 78-93.

15. *Erdil E., Yetkiner H.* A panel data approach for income-health causality. 2000.

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

17. *Hurlin C.* Testing Granger causality in heterogeneous panel data models with fixed coefficients. 2004.