ECONOMICS BULLETIN

# Panel Unit Root Tests and the Specification of Cross-sectional Dependence

Andrea Cerasa
*University Carlos III, Madrid*

## *Abstract*

This paper analyzes, through Monte Carlo experiments, the robustness of several panel unit root tests to different specifications of the cross-sectional dependence. Since results show that the miss-specification of cross-correlation crucially affects the properties of the tests, a graphical approach is suggested in order to determine the model of dependence which is likely to have generated the original data.

**Citation:** Cerasa, Andrea, (2008) "Panel Unit Root Tests and the Specification of Cross-sectional Dependence." *Economics Bulletin*, Vol. 3, No. 37 pp. 1-13
**Submitted:** March 11, 2008. **Accepted:** July 18, 2008.
**URL:** http://economicsbulletin.vanderbilt.edu/2008/volume3/EB-08C30035A.pdf

# 1 Introduction

In recent years, the issue of testing for unit root in panel data has been a much debated topic. The literature about the development of such tests was initially based upon the assumption of cross-sectional independence between the units and it produced the so called "first generation panel unit root tests". However, in several empirical applications, this assumption is likely to be violated and O'Connell (1998) showed that not considering the possible dependence between units could introduce severe bias in the first generation panel unit root tests. Hence researchers were interested in developing tests invariant with respect to the cross-sectional dependence, the so called "second generation unit root tests".

The main problem behind such testing procedures is that cross-sectional dependence can be specified in several ways. Therefore, assuming a particular model for cross-correlation, could lead to a miss-specification of the original DGP which may influence the properties of panel unit root tests, as argued by Breitung and Pesaran (2005).

The aim of this paper is just to analyze the robustness of several panel unit root tests to the different specifications of the cross-sectional dependence. This is not an unimportant issue, since, as argued by Breitung and Pesaran (2005), "the application of factor models in the case of weak correlation does not yield valid testing procedures" (p. 24). Furthermore, a graphical approach is suggested to determine the model of dependence which is likely to have generated the original data.

Hence, Section 2 is about the presentation of different cross-sectional models whereas Section 3 deals with the consequence on panel unit root tests of a dependence miss-specification, investigated through a Monte Carlo experiment. Finally, Section 4 presents a graphical analysis which could help to determine the kind of dependence which affected the data.

# 2 Specifications of the cross-section dependence

Cross-correlation in a panel data may be due to a variety of factors. Above all in presence of macro-economic data, the variables may be influenced by observed common factors, spatial spill over effects, unobserved common factors, or general residual interdependence that could remain even when all the observed and unobserved common effects are taken into account. Starting from the expression

$$z_{it} = \phi_i z_{i,t-1} + u_{it} \tag{1}$$

one of the most general specification for cross-correlation in the error term $u_{it}$ can be written as:

$$u_{it} = \gamma_i' f_t + v_{it} \qquad \text{or} \qquad \mathbf{u_t} = \Gamma' f_t + \mathbf{v_t} \tag{2}$$

where $f_t$ is a $K \times 1$ vector of serially uncorrelated random common factors and $\Gamma$ is a $N \times K$ matrix of random factor loadings defined by $\Gamma = (\gamma_1, \ldots, \gamma_N)'$[1]. Without loss of generality, it is assumed that the covariance matrix of $f_t$ is $I_K$ and that $f_t$ and $v_{it}$ are independently distributed.

Hence, the structure of cross-correlation is described by the covariance matrix of the composite error $u_t$ which, under the above assumptions, is given by:

$$\Omega_u = E(\mathbf{u_t}\mathbf{u_t'}) = \Gamma\Gamma' + \Omega_v \tag{3}$$

where $\Omega_v$ could be diagonal (*strict factor model*) or not (*approximate factor model*). These specifications are usually named "strong dependence" in order to distinguish them from the "weak dependence" case which rules out the presence of unobserved common factors (that is, for a weak dependence specification, $\Omega_u = \Omega_v$).

# 3 Panel unit root tests and miss-specification of the cross-correlation

Second generation panel unit root tests assume one of the three specifications previously introduced. Actually, only Bai and Ng (2004) developed a test based on an approximate factor model. However their test requires large panels, with the power only favorably affected when $T$ is increased. Therefore, the robustness of testing procedures to different specifications of the dependence will be investigated only for the $t_{rob}$ test provided by Breitung and Das (2005) and the $t_\alpha$ and $t_\beta$ proposed by Moon and Perron (2004). The first is a robust version of Dickey-Fuller $t$-statistic under the assumption of "weak dependence". The second ones assume a strict factor model specification and are based on eliminating the effect of cross-sectional dependence by projecting the panel data onto the space orthogonal to the factor loadings. The estimation of the common factors, of the factor loadings and of the choice of their number is very similar to the known procedure proposed in Bai and Ng (2002)[2].

[1] In the very particular case $\gamma_1 = \cdots = \gamma_N = \gamma$, the common factors components are equal to $\gamma' f_t$ for all $i$. Hence thhey reduce to a conventional "time effect" that can be removed by subtracting the cross-sectional mean from the data.

[2] The maximum number of common factors allowed is fixed to 8.

The 1000 Monte Carlo populations of $N$ units and $T$ observations are generated according to expression (1) with $\phi_i = 1$ for the size case and $\phi_i = 0.98$ for the power case[3]. The error terms $u_{it}$ are generated according to the following specifications:
-DGP 1 (*no cross-sectional dependence*): $u_{it}$ ~$N$(0,3);
-DGP 2 (*weak dependence*): $\mathbf{u_t} = \mathbf{w_t} S'$ where $\mathbf{w_t}$ ~$N(0,I_N)$ and $s_{ij}$ ~Uni(0, $9/N$);
-DGP 3 (*strict 3-factors model*): $\mathbf{u_t} = \Gamma' f_t + \mathbf{w_t}$ where $f_{jt}$ ~$N$(0,1), $\gamma_{ij}$ ~Uni(0, 6/K) with $K$=number of common factors;
-DGP 4 (*approximate 3-factors model*): $\mathbf{u_t} = \Gamma' f_t + \mathbf{v_t}$ where $f_{jt}$ ~$N$(0,1), $\gamma_{ij}$ ~Uni(0, 6/K) and $\mathbf{v_t} = \mathbf{w_t} S'$ where $\mathbf{w_t}$ ~$N(0,I_N)$ and $s_{ij}$ ~Uni(0, $3/N$);

This particular choice of settings guarantees that in each case the expected values of the variance of $u_{it}$ is equal to 3 whereas the expected values of the covariance between each couple of $ij$ with $i \neq j$ are 3/2 for DGP 2 and 4, and 3/4 for DGP 3.

It is seen (table 1) that the $t_{rob}$ test provided by Breitung and Das (2005) appears to be the most robust to the dependence specification: its estimated sizes are indeed always close to the nominal level. Also its power is rather satisfactory and increasing with $N$ and $T$, even though in DGP 2 and 4 its value is remarkably smaller than in the other cases.

The same robustness is not present in the tests based on a strict factor structure. The empirical results show evident over-size problems for both $t_\alpha$ and $t_\beta$ in all cases when $N$=10. When $N$>10, the size of both tests approaches to the nominal level only in DGP 1 (no dependence) and 3 (strict factor model), with a power significantly higher than the $t_{rob}$ test in case of strict factor specification. But when the assumption of strict factor model does not hold, the estimated sizes remain clearly over the nominal level, even with large $N$ and $T$.

Looking at the last column of the table, which lists the mean number of common factors estimated over the 1000 replications, it is possible to appreciate the robustness of the number of factors selection procedure. In case of strict factor model, it works quite properly when $N$ is at least 20. Instead, in case of approximate factor model, it tends to overestimate the true number of common factors. As expected, in cases of DGP 3 and 4, the procedure improves remarkably when $N$ and $T$ grows.

Concluding, it seems that simulation results confirm the suggestion of Breitung and Pesaran (2005). When $N = 10$, the use of $t_{rob}$ is recommended due to its robustness to all the specifications. But when $N > 10$, the use of $t_\alpha$ and $t_\beta$ in case of strict factor models guarantees good size accuracy

[3]Both tests are indeed optimal against the homogeneous alternative.

and high power. Then, it is very important to detect which model generates the cross-correlation in the data, in order to use the proper unit root tests. Unfortunately, the procedure suggested by Bai and Ng (2002) does not work properly when N and T are not enough large and when the common factors assumption does not hold. In the next paragraph, a graphical approach is introduced in order to determine the model of dependence which is likely to have generated the original data.

# 4 IC function and the specification of cross-section dependence

The estimates of the common factors and the factor loadings are obtained through the principal components methods applied to the $T \times T$ matrix $\hat{\mathbf{u}}\hat{\mathbf{u}}'$, where $\hat{\mathbf{u}}$ is the $T \times N$ matrix of the OLS residuals from regression (1). The choice of the number of common factors is treated as a model selection problem and it is based on the optimization of an information criteria. Hence:

$$\bar{K} = arg\min_{r}\{\ln[tr(\hat{V}_r'\hat{V}_r)/NT] + r \cdot g(N,T)\} \tag{4}$$

where $\hat{V}_r$ is the estimation of the matrix of residuals $V$ with $r$ factors and $g(N,T)$ is the penalty term such that: $g(N,T) \to 0$ as $N,T \to \infty$.

Bai and Ng (2002) proved that, in case of strict or approximate common factor structure, the principal components consistently estimate the common factors and the factor loadings. Furthermore, they show that the minimization of $IC(\cdot)$ consistently estimates the true number of common factors $r$. Empirical simulations have shown that this method offers good performance also in small samples.

But what happens when the data are characterized by weak dependence or no dependence? Non trivial calculations show that:

$$tr(\hat{V}_r'\hat{V}_r) = \sum_{i=r+1}^{N} \hat{a}_i \tag{5}$$

where $\hat{a}_i$ is the $i$-th eigenvalue of the matrix $\hat{\mathbf{u}}'\hat{\mathbf{u}}$. Therefore:

$$IC(r) = \ln(\sum_{i=r+1}^{N} \hat{a}_i) + r \cdot g(N,T) \tag{6}$$

$$IC(r) - IC(j) = \ln\left(\frac{\sum_{i=r+1}^{N} \hat{a}_i}{\sum_{i=j+1}^{N} \hat{a}_i}\right) + (r-j) \cdot g(N,T) \tag{7}$$

Bai and Ng (2002) proved that, in case of factors models (approximate or strict), this quantity is asymptotically bigger than 0 whenever $j = K$ and $r \neq K$. When, instead, weak dependence (or no dependence) affects the data, we have:

$$\lim_{T\to\infty}[IC(r) - IC(j)] = \ln\left(\frac{\sum_{i=r+1}^{N} a_i}{\sum_{i=j+1}^{N} a_i}\right) + (r - j)\cdot g(N, T) \tag{8}$$

since $\mathbf{\hat{u}'\hat{u}}$ is a consistent estimate of the matrix $T\Omega_u$[4]. The first term is smaller than zero whereas the second is bigger than zero, hence this quantity can be either positive or negative. If now $T \to \infty$ is followed by $N \to \infty$, given that the eigenvalues of $\Omega_u$ are bounded as $N \to \infty$ as argued by Breitung and Pesaran (2005), we have:

$$\lim_{N\to\infty}\{\lim_{T\to\infty}[IC(r) - IC(j)]\} = 0 \tag{9}$$

Therefore, we can conclude that when a factor structure is wrongly assumed, the $IC(\cdot)$ function assumes asymptotically a constant value and then it does not have a minimum. To analyze the behavior of the function in small samples, we can perform another Monte Carlo experiment in which the path of $IC(r)$ for $1 < r < (N-1)$ is analyzed considering the four different specifications for the cross-sectional dependence of $\mathbf{u_t}$ previously introduced. On this subject, the figures 1, 2, 3, 4, 5, 6 and 7 report the paths of the averages of $IC(r)$ obtained over 1000 Monte Carlo replications. As suggested in several applications, the penalty term used for the simulations is:

$$g(N,T) = \frac{N+T}{NT}\ln\left(\frac{NT}{N+T}\right) \tag{10}$$

It is seen that the behavior of $IC(\cdot)$ functions differ depending on the dependence specification. Starting from the case of no dependence (DGP 1), the path of $IC(\cdot)$ is convex. In particular, it is strictly decreasing when $N$=10 and firstly increasing and then decreasing when $N$>10. In line with the results of table 1, this means that imposing $r_{max} = 8$, the procedure tends to select in general $\hat{K} = 1$.

In case of weak dependence (DGP 2), the path of $IC(\cdot)$ is decreasing, at least for $N$<40. Hence, again in accordance with table 1, the procedure tends to choose $\hat{K} = r_{max}$ when $N$<40, $\hat{K} = 1$ when $N$=80, whereas when $N$=40 the possible presence of local minimum in the first values of the function makes the expected behavior of the selection process highly unpredictable.

[4] In case of no dependence, the $N$ eigenvalues $a_i$ are the variances of $u_i$

It is not worthless to notice how the $IC(\cdot)$ paths for DGP 1 and DGP 2 differ in their final part: in both cases they are indeed decreasing but with a deeper intensity in the case of weak dependence.

When a factors structure affects the data, the shape of the $IC$ function in its first values is remarkably different. As expected, in case of DGP 3 it has a local minimum in $r = K$ when $N$ is at least 20, whereas the remaining path is similar to that of DGP 1. Instead, in case of DGP 4 it has a change of slope in $r = K + 1$ which becomes a local minimum when $N$=80 whereas the remaining path is similar to that of DGP 2.

Therefore, above all when $N$ is not big enough, the analysis of the path of $IC$ function could help to determine the model of dependence which affected the actual data. A presence of a local minimum or a change of slope in its first values points out the presence of a common factors structures in the data. The possible cross-correlation in the idiosyncratic component could instead be investigated by the analysis of the last part of the $IC(\cdot)$ path. If it is firstly increasing and then decreasing, it is likely to assume independence in the idiosyncratic term. Whereas, if it is decreasing with an high negative slope in its final part, it could be better to assume cross-sectional dependence in the idiosyncratic component.

# 5 Conclusion

In this paper, a Monte Carlo experiment shows the consequences of miss-specification of the cross-sectional specification on several panel unit root tests. Results confirm the hint of Breitung and Pesaran (2005), highlighting size bias when miss-specification of the dependence arises, above all for the tests which assumes strict factors models.

Therefore, the determination of cross-sectional model becomes crucial in the choice of panel unit tests. Another Monte Carlo experiment suggests that the graphical analysis of IC path could be useful in order to determine the kind of dependence. This will help to choose the proper and most powerful panel unit root tests.

# References


Bai, J. and S. Ng (2002), "Determining the number of common factors in approximate factor model", *Econometrica*, 70, 191-221.

Bai, J. and S. Ng (2004), "A PANIC attack on Unit Roots and Cointegration", *Econometrica*, 72, 1127-1177.

Breitung, J. and S. Das (2005), "Panel Unit Root tests under Cross-Sectional Dependence", *Statistica Neerlandica*, 59, 414-433.

Breitung, J. and M.H. Pesaran (2005), "Unit Roots and Cointegration in Panels",IEPR Working Papers 05.32, Institute of Economic Policy Research (IEPR), revised.

Moon, R. and B. Perron (2004), "Testing for Unit Root in Panels with Dynamic Factors", *Journal of Econometrics*, 122, 81-126.

O'Connell, P. (1998) "Overvaluation of Purchasing Power Parity", *Journal of International Economics*, **44**, 1-19.

Table 1: Size and power of 5% panel unit root tests for different specifications of the cross-sectional dependence

| | | | size | | | power | | | |
|---|---|---|---|---|---|---|---|---|---|
| **N** | **T** | **DGP** | $t_\alpha$ | $t_\beta$ | $t_{rob}$ | $t_\alpha$ | $t_\beta$ | $t_{rob}$ | $\bar{K}$ |
| **10** | **50** | **1** | 15.6 | 26.3 | 4.0 | 60.5 | 66.6 | 66.9 | 8.0 |
| | | **2** | 30.1 | 34.1 | 5.5 | 68.5 | 68.3 | 20.9 | 8.0 |
| | | **3** | 23.0 | 33.0 | 5.8 | 68.2 | 73.7 | 31.8 | 8.0 |
| | | **4** | 31.0 | 38.6 | 6.4 | 72.0 | 72.7 | 20.5 | 8.0 |
| | **100** | **1** | 22.1 | 25.0 | 4.5 | 88.0 | 85.4 | 97.6 | 8.0 |
| | | **2** | 36.2 | 35.8 | 5.7 | 89.2 | 83.2 | 33.7 | 8.0 |
| | | **3** | 26.8 | 30.3 | 4.3 | 90.5 | 87.0 | 53.4 | 8.0 |
| | | **4** | 34.0 | 34.1 | 5.2 | 90.3 | 85.8 | 30.3 | 8.0 |
| **20** | **50** | **1** | 2.1 | 6.7 | 3.3 | 92.9 | 95.7 | 93.2 | 1.0 |
| | | **2** | 25.7 | 40.0 | 6.1 | 87.9 | 91.4 | 26.9 | 8.0 |
| | | **3** | 3.5 | 10.7 | 5.8 | 94.3 | 96.4 | 37.2 | 2.6 |
| | | **4** | 26.7 | 40.5 | 6.9 | 88.1 | 91.0 | 23.5 | 8.0 |
| | **100** | **1** | 4.6 | 7.9 | 5.1 | 100.0 | 100.0 | 100.0 | 1.0 |
| | | **2** | 26.8 | 33.7 | 5.0 | 99.2 | 98.7 | 35.5 | 8.0 |
| | | **3** | 5.9 | 8.2 | 6.3 | 100.0 | 99.9 | 58.7 | 2.7 |
| | | **4** | 26.9 | 33.1 | 6.6 | 99.4 | 99.1 | 35.1 | 8.0 |
| **40** | **50** | **1** | 0.9 | 5.1 | 2.6 | 100.0 | 100.0 | 99.9 | 1.0 |
| | | **2** | 24.0 | 35.0 | 6.6 | 97.4 | 98.5 | 29.7 | 5.1 |
| | | **3** | 2.1 | 9.8 | 5.5 | 99.7 | 99.9 | 45.2 | 2.9 |
| | | **4** | 32.1 | 46.0 | 7.2 | 97.2 | 98.5 | 21.4 | 7.2 |
| | **100** | **1** | 3.1 | 5.7 | 3.6 | 100.0 | 100.0 | 100.0 | 1.0 |
| | | **2** | 28.4 | 34.9 | 9.0 | 100.0 | 100.0 | 42.8 | 5.4 |
| | | **3** | 3.2 | 8.0 | 6.4 | 100.0 | 100.0 | 60.8 | 3.0 |
| | | **4** | 30.5 | 37.2 | 7.4 | 100.0 | 100.0 | 37.9 | 7.3 |
| **80** | **100** | **1** | 1.1 | 4.2 | 2.4 | 100.0 | 100.0 | 100.0 | 1.0 |
| | | **2** | 12.1 | 17.4 | 7.7 | 100.0 | 100.0 | 47.5 | 1.1 |
| | | **3** | 3.5 | 9.1 | 8.7 | 100.0 | 100.0 | 66.6 | 3.0 |
| | | **4** | 25.0 | 31.9 | 8.7 | 100.0 | 100.0 | 42.8 | 4.2 |

*Notes*: the DGPs are generated according to the expressions (1) with $\phi_i$=1 $\forall i$ in the size case and $\phi_i = 0.98$ for the power case. The residuals $u_{it}$ are defined as:
DGP 1 (*no cross-sectional dependence*): $u_{it} \sim N(0,3)$;
DGP 2 (*weak dependence*): $\mathbf{u_t} = \mathbf{w_t} S'$ where $\mathbf{w_t} \sim N(0,I_N)$ and $s_{ij} \sim$Uni(0, $9/N$);
DGP 3 (*strict 3-factors model*): $\mathbf{u_t} = \Gamma' f_t + \mathbf{w_t}$ where $f_{jt} \sim N(0,1)$, $\gamma_{ij} \sim$Uni(0, 6/K) with $K$=number of common factors;

DGP 4 (*approximate 3-factors model*): $\mathbf{u_t} = \Gamma' f_t + \mathbf{v_t}$ where $f_{jt} \sim N(0,1)$, $\gamma_{ij} \sim$Uni(0, 6/K) and $\mathbf{v_t} = \mathbf{w_t} S'$ where $\mathbf{w_t} \sim N(0,I_N)$ and $s_{ij} \sim$Uni(0, $3/N$)

Figure 1: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=10 and T=50 (DGP as in table 1)

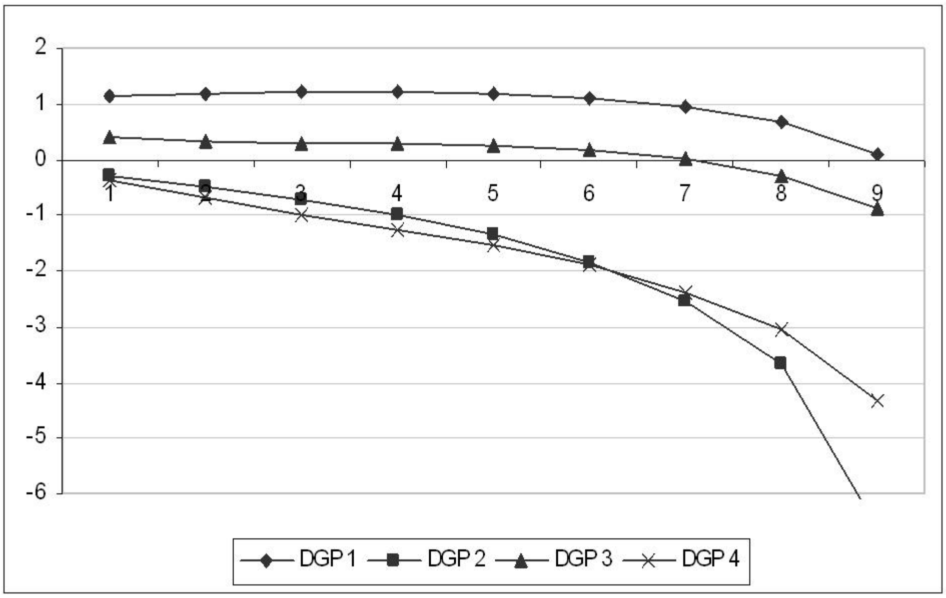


Figure 2: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=10 and T=100 (DGP as in table 1)

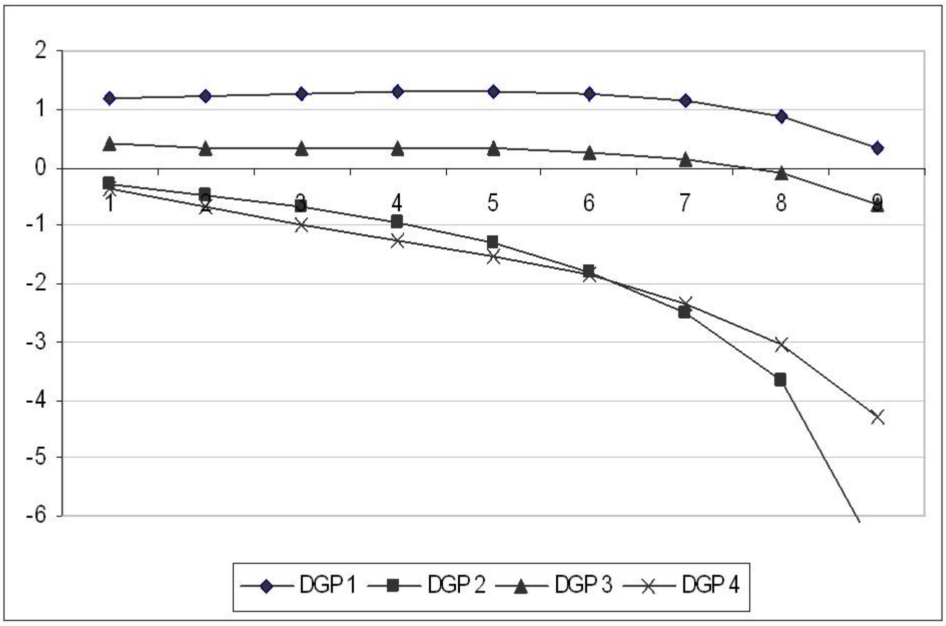

Figure 3: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=20 and T=50 (DGP as in table 1)

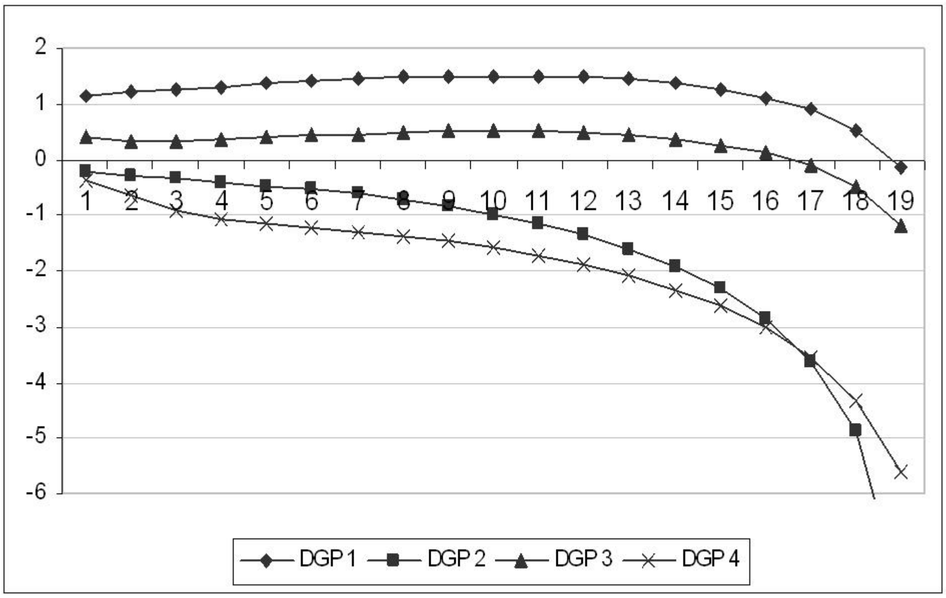


Figure 4: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=20 and T=100 (DGP as in table 1)

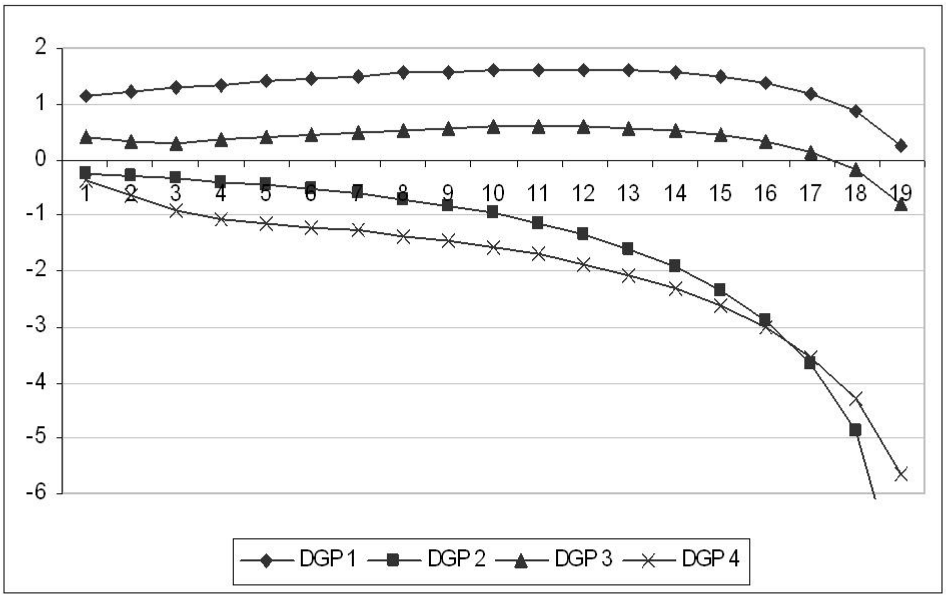

Figure 5: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=40 and T=50 (DGP as in table 1)

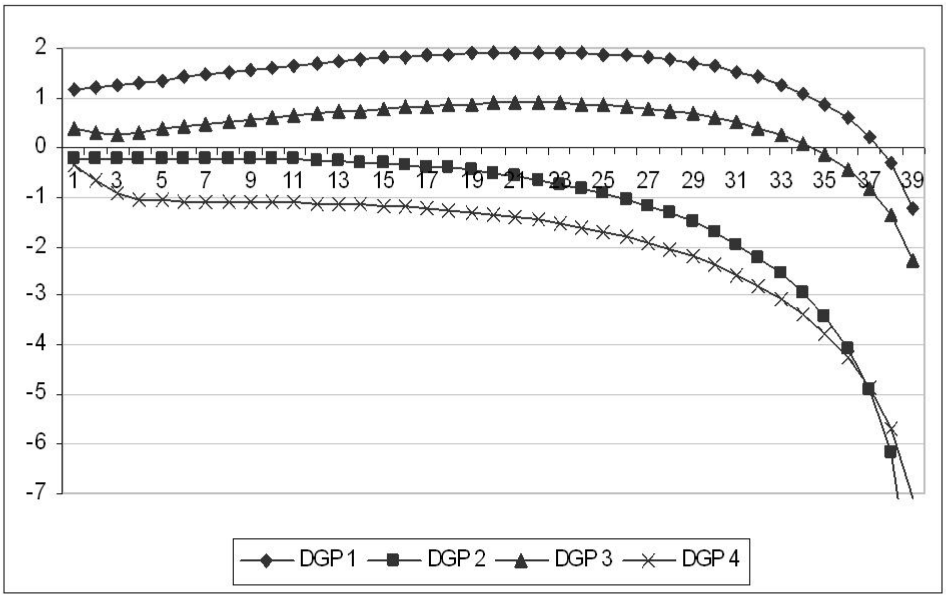


Figure 6: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=40 and T=100 (DGP as in table 1)

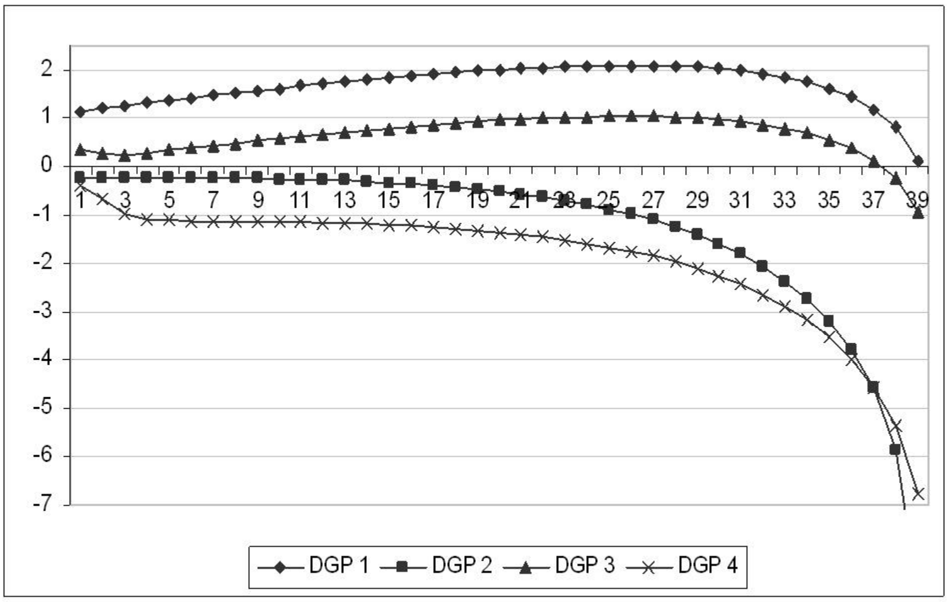

Figure 7: Path of $IC(r)$ function for $1 < r < (N-1)$ when N=80 and T=100 (DGP as in table 1)

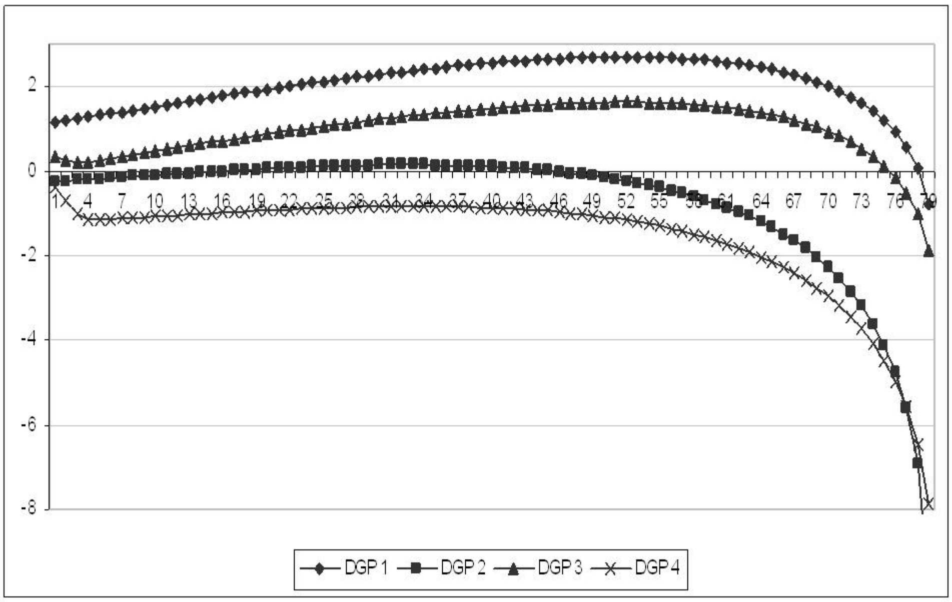


Figure 8: Zoom in of figure 7