The estimation of the autocorrelation coefficient in panel data models with AR(1) disturbances.*

Alexandre Cazenave-Lacroutz[a,b,c] and Vieu Lin[c]

23 septembre 2019

Version 1.0.4 - comments are welcome !

This paper proposes new estimators of the autocorrelation coefficient in fixed- and random-effects model with serial correlation of order one in the perturbations. In balanced panels, it indeed shows that the current estimators are strongly biased, even when the number of individuals tends to infinity, and quantifies their bias. Then it proposes an estimator that is consistent and asymptotically unbiased in both balanced and unbalanced panels. It also proposes additional estimators which are asymptotically equivalent in long panels and very simple to estimate. Monte-Carlo simulations eventually illustrate that the new estimators are much more reliable than current estimates.

*We would like to thank Tomáš Jagelka, Fanny Godet, Sébastien Roux, Fabian Stürmer-Heiber and Joachim Winter for comments and suggestions provided while writing this article. This research was conducted while implementing new wage equations in the microsimulation model Destinie 2, see Cazenave-Lacroutz *et al.* (2019). Alexandre thanks Prof.Dr. Winter for having hosted him at LMU University of Munich when this research was started.

[a] Institut National de la Statistique et des Études Économiques, 88 Avenue Verdier, Montrouge. This document does not reflect the position of Insee, Université Paris Dauphine or Crest, but only its authors' views.

[b] Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75016 Paris

[c] Crest, 5 Avenue Henry Le Chatelier, 91120 Palaiseau

1 Introduction

This articles improves the estimation of the autocorrelation coefficient in linear unobserved effects panel data models[1] with AR(1) disturbances, that is processes of the form :

$$y_{it} = x'_{it}\beta + \nu_i + \varepsilon_{it} \tag{1}$$

$$\varepsilon_{it} = \rho\varepsilon_{it-1} + \eta_{it} \tag{2}$$

where $|\rho| < 1$ and the η_{it}'s are i.i.d. disturbances with mean 0 and variance σ^2_η.

Such estimations are common in the wage equation literature (with even higher order of correlation), and are not infrequent in the general economics literature. They are performed in Stata with the command *xtregar*, which has been used in influential and recent economics articles such as Dafny (2010) or Hau *et al.* (2013).

The estimation of the autocorrelation ρ is key in such models : First, these models are often specifically chosen for their explicit modeling of serial correlations in error terms, something which is accounted for by ρ. Second, the other parameters of the model are estimated based on a Cochrane–Orcutt transformation using the given estimate of ρ.

A standard estimator of ρ is obtained with the Durbin-Watson statistic d :

$$\hat{\rho} = \rho_d = 1 - \frac{d}{2} \tag{3}$$

Section 2 shows that, in balanced panels, this standard estimator is biased towards zero in $O(\frac{1}{T})$, where T is the number of observations per individual.
In section 3, an alternative estimate of ρ, that we call ρ_{BFN} since it was suggested with much intuition by Bhargava *et al.* (1982), is shown to be consistent and asymptotically unbiased as N becomes large. We also generalize it to the unbalanced setting, where it keeps these desirable properties under reasonable hypotheses. This new estimator eventually performs much better than the existing estimates in Monte Carlo simulations both in balanced and unbalanced panels.

[1] That is in both fixed and random effects models.

We further define two additional estimators that can yield advantages over this estimator in long panels. In balanced panels, just dividing ρ_d estimator by $1 - \frac{2}{T}$, with T the number of periods, is enough to get an estimator whose performances are almost indiscernible to the estimator initially suggested by Bhargava *et al.* (1982) (the bias is in $\frac{1}{T^2}$). In unbalanced panels, a similar approximation yields a less precise estimator whose bias tends however to zero when the minimal number of period per individual tends to infinity, something which does not seem guaranteed by current estimators. Section 4 concludes.

2 Overlooked consequences of Bhargava *et al.* (1982).

2.1 The balanced case

2.1.1 Definitions

In a perfectly balanced panel (Bhargava *et al.*, 1982), the Durbin-Watson statistics (used by the Stata command *xtregar* to estimate ρ) writes :

$$d_p = \frac{\sum_{i=1}^{N}\sum_{t=2}^{T}(\tilde{u}_{it} - \tilde{u}_{it-1})^2}{\sum_{i=1}^{N}\sum_{t=1}^{T}\tilde{u}_{it}^2} \tag{4}$$

where $\tilde{u}_{it}$ are the residuals of the within-estimation of model (1) - the residuals of the OLS regression of $y_{it} - \overline{y}_i$ on $x_{it} - \overline{x}_i$ with $\overline{y}_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$ and $\overline{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{it}$ -, N is the number of individuals and T the number of (equally spaced) periods.

Bhargava *et al.* (1982) have generalized the Durbin-Watson statistics to get a Uniformly Most Powerful test that ρ is significantly different from zero. However, they first note that one can use equation (3) with this formula to generate a consistent estimate of ρ as $T \to \infty$. This is what *xtregar* currently performs.

2.1.2 An estimate of the bias

In their MonteCarlo simulations (with finite $T = 10$), Bhargava *et al.* (1982) however point out that the corresponding estimates ρ_d can be significantly different from the true value of ρ (see their Table IV, page 541 ; or our own Table 1). This bias can be high in magnitude : in their simulations, it can amount to -0.26 for a value of

$\rho = 0.9$.

Although they do not highlight it, the existence of such a bias theoretically derives from the below relation (see Bhargava *et al.* (1982))[2] :

$$\mathbb{E}(\rho_d) = 1 - \frac{(1-\rho)(T-1)}{T - \frac{1}{T}\sum_{i,j=1}^{T} \rho^{|i-j|}} \tag{5}$$

The bias does not depend on the number of individuals N, as equation (5) does not depend on N[3]. It does however depend on the number of period T. We therefore provide an approximation of the bias when $T \to \infty$. One first notices (as demonstrated by Appendix B of Bhargava *et al.* (1982)) :

$$\sum_{j,k=1}^{T} \rho^{|j-k|} = \frac{1+\rho}{1-\rho}T - \frac{2\rho}{1-\rho}\frac{1-\rho^T}{1-\rho} \tag{6}$$

which leads to :

$$\mathbb{E}(\rho_d) = 1 - \frac{(1-\rho)\frac{(T-1)}{T}}{1 - \frac{1+\rho}{1-\rho}\frac{1}{T} + \frac{2\rho(1-\rho^T)}{(1-\rho)^2}\frac{1}{T^2}} \tag{7}$$

We develop this equation at the second order in $\frac{1}{T}$[4] :

$$\mathbb{E}(\rho_d) = \rho - \frac{2\rho}{T} + O(\frac{1}{T^2}) \tag{9}$$

Hence, in the balanced case, the bias is of the order of $\frac{1}{T}$[5]. For instance, for $T = 10$, it is of the order of 0.1. Which is high in magnitude if, for instance, ρ equals 0.5.

[2] Indeed, if ρ_d were unbiased, then $\mathbb{E}(\rho_d) = \rho$ for any ρ. From equation (5), this would imply, for any ρ :

$$\rho = 1 - \frac{(1-\rho)(T-1)}{T - \frac{1}{T}\sum_{i,j=1}^{T} \rho^{|i-j|}}$$

which contradicts the fact this equation has a finite number of roots.

[3] Note that relation (5) is however valid for N large enough.

[4] More precisely :

$$\mathbb{E}(\rho_d) = \rho - \frac{2\rho}{T} + \frac{1}{T^2}\frac{\frac{2\rho^2}{(1-\rho)}(2\frac{1-\rho^T}{T(1-\rho)} - 1 - \rho^{T-1})}{1 - \frac{1+\rho}{1-\rho}\frac{1}{T} + \frac{2\rho(1-\rho^T)}{(1-\rho)^2}\frac{1}{T^2}} \tag{8}$$

[5] And, as the coefficient in front of $\frac{1}{T}$ is negative, it is probably negative (if ρ positive) even for moderate values of T, as could be already observed in the Monte Carlo simulations of Baltagi et Wu (1999).

This bias tends to 0 as the time dimension gets larger. For T = 50, it is of the order of 0.02, which might be perceived (or not) as negligible.

2.2 A generalization to the unbalanced case

In the unbalanced case, an estimator of $d_p = 2(1-\rho)$ is given by :

$$d_p = \frac{\sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} (\tilde{u}_{it_{ij}} - \tilde{u}_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1}}{\sum_{i=1}^{N} \frac{1}{n_i} \sum_{j=1}^{n_i} \tilde{u}^2_{it_{ij}}} \tag{10}$$

where $t_{i1} < ... < t_{in_i}$ denote the dates at which individual i is observed and $K_i = \sum_{j=2}^{n_i} \mathbb{1}_{t_{ij}-t_{ij-1}=1}$ the number of observations separated by one period for the individual i. It is easy to see that, in the balanced case, one gets equation (4) back [6].

We explain in Annex A why d_p can be perceived as a natural estimator of ρ also in the unbalanced case. However, we show in Annex C that $\rho_d = 1 - \frac{d_p}{2}$ is asymptotically close to its expectancy as $N \to \infty$ [7], and in Annex B that the generalization of relation (5) writes :

$$\mathbb{E}(\rho_d) = 1 - \frac{(1-\rho)\sum_{i=1}^{N} \frac{K_i}{1+K_i}}{N - \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|}} \tag{11}$$

Unlike what was observed in the balanced case, such bias does not necessarily tend towards zero if the number of observations per individual tends to infinity. This depends on the structure of the missing observations.

[6] Note that such generalization (in the unbalanced case) of the Durbin-Watson statistics of Bhargava *et al.* (1982) differs from the d1-statistics of Baltagi et Wu (1999) (which is another generalization to the unbalanced case of the the Durbin-Watson statistics of Bhargava *et al.* (1982)). In this latter, the dummy variable in the numerator is straightly in the parentheses along with $\tilde{u}_{it_{ij-1}}$, and there is no mention of K_i.

[7] Note that this extends Bhargava *et al.* (1982) work in the balanced case as they only noted that ρ_d was consistent when $T \to \infty$. True, when only $N \to \infty$, ρ_d is consistent towards something else as ρ, which may be why Bhargava *et al.* (1982) have not noticed it.

Indeed, since $|t_{ij} - t_{ik}| \geq |j-k|$ and $|\rho| < 1$:

$$|\sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|}| \leq \sum_{j,k=1}^{n_i} |\rho|^{|j-k|} = \frac{1+|\rho|}{1-|\rho|} n_i - \frac{2|\rho|}{1-|\rho|} \frac{1-|\rho|^{n_i}}{1-|\rho|}$$

and using a development of order 1 in $\frac{1}{m}$ with $m = \min(n_i)$, we notice :

$$\begin{aligned} \mathbb{E}(\rho_d) &= 1 - \frac{\frac{1-\rho}{N} \sum_{i=1}^{N} \frac{K_i}{1+K_i}}{1+O(\frac{1}{m})} \\ &= 1 - \frac{1-\rho}{N} \sum_{i=1}^{N} \frac{K_i}{1+K_i} (1 + O(\frac{1}{m})) \qquad (12) \\ &= (1 - \frac{1}{N} \sum_{i=1}^{N} \frac{K_i}{1+K_i}) + \frac{\rho}{N} \sum_{i=1}^{N} \frac{K_i}{1+K_i} + O(\frac{1}{m}) \end{aligned}$$

For instance, consider a dataset where we observe the two first consecutive observations, and only every other observation thereafter. Hence $K_i = 1$ and $\mathbb{E}(\rho_d) \xrightarrow[m \to \infty]{} \frac{1+\rho}{2} \neq \rho$.

3 The ρ_{BFN} estimate of ρ and its approximations

Bhargava *et al.* (1982) also suggest to estimate ρ by solving for equation (7) after substituting $\mathbb{E}(\rho_d)$ with ρ_d. Using rather equation (11), ρ_{BFN} is thus implicitly defined in the general case by :

$$\rho_d = g_N(\rho_{BFN}) \qquad (13)$$

with :

$$g_N : r \mapsto 1 - \frac{(1-r) \sum_{i=1}^{N} \frac{K_i}{1+K_i}}{N - \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}} \qquad (14)$$

Note that in the balanced case, g_N does not depend on N, in which case we write it f with :

$$f : r \mapsto 1 - \frac{(1-r)(T-1)}{T - \frac{1}{T} \sum_{i,j=1}^{T} r^{|i-j|}}$$

In the balanced case, Bhargava *et al.* (1982) perform Monte Carlo simulations to provide an assessment of this method : it seems to deliver unbiased estimates of ρ. We build on this remark to improve the estimation of ρ.

3.1 Formal definition and theoretical properties

As noted by Bhargava *et al.* (1982), this correction procedure for estimating ρ is "*somewhat unconventional*". Even though their Monte Carlo study is somewhat conclusive, they do not establish the theoretical properties of their estimator, such as whether it is unbiased or consistent.

We first check that the definition of the ρ_{BFN} is unambiguous, in that there is at most one solution to the defining equation (13). We establish it formally in the case $0 < \rho < 1$. In both the balanced and unbalanced case, this new estimator is consistent under reasonable hypotheses.

We further show that ρ_{BFN} is not biased for large N in both balanced an unbalanced panels - unlike the bias in ρ_d which is, for instance, of order $\frac{1}{T}$ whatever the value of N in balanced panel. This confirms that this estimator of ρ should be considered instead of ρ_d.
As their is some ambiguity regarding the definition of ρ_{BFN} when $\rho < 0$, and as the computation of ρ_{BFN} may become numerically intractable as the time dimension of the panel increases, we also provide two approximations of ρ_{BFN} : one in the balanced case, and one in the unbalanced case, which are well-defined and have good theoretical properties in long panels.

3.1.1 Unambiguity of the definition of ρ_{BFN}

The following lemma shows there exists at most one estimator taking values between 0 and 1 that solves equation (13) (see Annex D for a demonstration).

Lemma 1 : $g_N : r \to 1 - \frac{(1-r)\sum_{i=1}^{N} \frac{K_i}{1+K_i}}{N-\sum_{i=1}^{N} \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}}$ establishes a bijection from $[0,1]$

to $[1 - \frac{\sum_i^N \frac{K_i}{1+K_i}}{N - \sum_i^N \frac{1}{n_i}}, g_N(1)]$.

Section (3.1.4) deals with the case $\rho < 0$.

3.1.2 Consistency of ρ_{BFN}

We want to prove :

$$\rho_{BFN} \xrightarrow{\mathbb{P}} \rho \tag{15}$$

Under mild hypotheses, we have established in Annex C the following convergence as N grows to $+\infty$:

$$\rho_d - g_N(\rho) \xrightarrow{\mathbb{P}} 0$$

which implies, from the defining equation (13) :

$$g_N(\rho_{BFN}) - g_N(\rho) \xrightarrow{\mathbb{P}} 0$$

Balanced case :

In the balanced case, the above convergence writes :

$$f(\rho_{BFN}) \xrightarrow{\mathbb{P}} f(\rho)$$

In Annex B, we show that f is continuous and bijective over $[0,1]$. Hence, there exists a continuous inverse function f^{-1}. From the continuous mapping theorem :

$$\rho_{BFN} = f^{-1}(f(\rho_{BFN})) \xrightarrow{\mathbb{P}} f^{-1}(f(\rho)) = \rho$$

Unbalanced case :

We shall prove the convergence (15) under the following assumption :

Assumption 3 : For all $r \in [0,1]$, both sequences $\frac{1}{N}\sum_{i=1}^N \frac{K_i}{1+K_i}$ and $\frac{1}{N}\sum_{i=1}^N \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}$ converge as $N \to \infty$.

One notices that, in the general case, this assumption holds for $\frac{1}{N}\sum_{i=1}^{N}\frac{K_i}{1+K_i}$, provided the K_i's are realizations of any distribution. Furthermore, this assumption holds in the balanced case for both sequences.

Such an assumption ensures the pointwise convergence of g_N to some function g_∞ on the segment $[0,1]$. Note that, in Annex B, we show that g_N is bijective over $[0,1]$, and its inverse g_N^{-1} is C^1 over its image [8]. Hence, g_N admits a C^1 inverse function g_N^{-1}. g_∞ is assumed to verify the following assumption :

Assumption 4 : g_∞ is bijective from $[0,1]$ to its image and its inverse g_∞^{-1} is C^1 over its image.

Lemma 1 above shows that g_N is increasing on $[0,1]$, hence according to the second Dini theorem, g_N converges uniformly to g_∞. Moreover, g_∞ is continuous as it is the case for g_∞^{-1}. The rest of the proof is now straightforward.

First, we observe that, for all $0<\rho<1$, $g_N(\rho_{BFN})$ converges to $g_\infty(\rho)$ in probability.
Second, the following inequality shows $g_\infty(\rho_{BFN}) \overset{\mathbb{P}}{\to} g_\infty(\rho)$:

$$\begin{aligned}|g_\infty(\rho_{BFN})-g_\infty(\rho)| &\leq |g_\infty(\rho_{BFN})-g_N(\rho_{BFN})|+|g_N(\rho_{BFN})-g_\infty(\rho)|\\ &\leq ||g_\infty-g_N||_\infty+|g_N(\rho_{BFN})-g_\infty(\rho)|\end{aligned}$$

Besides, applying the mean value theorem to $h=g_\infty^{-1}$, we may write

$$h(g_\infty(\rho_{BFN}))=h(g_\infty(\rho))+(g_\infty(\rho_{BFN})-g_\infty(\rho))h'(c_N)$$

for some scalar c_N lying between $g_\infty(\rho)$ and $g_\infty(\rho_{BFN})$. This relation writes down to

$$\rho_{BFN}=\rho+(g_\infty(\rho_{BFN})-g_\infty(\rho))h'(c_N)$$

Since $g_\infty(\rho_{BFN}) \overset{\mathbb{P}}{\to} g_\infty(\rho)$ and h' is bounded, we have $\rho_{BFN} \overset{\mathbb{P}}{\to} \rho$.

[8] As soon there is at least one individual that is observed at least three times.

3.1.3 The asymptotic bias of ρ_{BFN}

Here, we establish the following convergence :

$$\mathbb{E}(\rho_{BFN}) \xrightarrow[N\to\infty]{} \rho \tag{16}$$

That is, ρ_{BFN} is asymptotically unbiased as the number of individuals is becoming large.

As $\rho_d = g_N(\rho_{BFN})$ and $\mathbb{E}(\rho_d) = g_N(\rho)$, we have :

$$\mathbb{E}(g_N(\rho_{BFN})) = g_N(\rho) \tag{17}$$

Denoting $A_N = \frac{1}{N}\sum_{i=1}^{N} \frac{K_i}{1+K_i}$ which satisfies $\frac{1}{2} \leq |A_N| \leq 1$, we notice :

$$g_N(r) = 1 - A_N + rA_N - A_N p_N(r) \tag{18}$$

where :

$$p_N(r) = \frac{1-r}{\frac{N}{\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}} - 1}$$

We treat the balanced and the unbalanced case separately.

The balanced case :

In the balanced case, relation (17) translates into :

$$\mathbb{E}(f(\rho_{BFN})) = f(\rho) \tag{19}$$

Relation (18) becomes :

$$f(r) = 1 - \frac{T-1}{T} + r\frac{T-1}{T} - \frac{T-1}{T}q(r) \tag{20}$$

where

$$q(r) = \frac{1-r}{\frac{1}{\frac{1}{T^2}\sum_{j,k=1}^{T} r^{|j-k|}} - 1}$$

Combining relations (19) and (20), we get :

$$\mathbb{E}(\rho_{BFN} - \rho) = \mathbb{E}(q(\rho_{BFN}) - q(\rho))$$

It is easy to show that the right hand side tends to zero as N grows to infinity, which establishes (16). Indeed, as shown in Annex E, q may be continously prolounged over the segment $[0,1]$, hence q is bounded by some constant $B > 0$ and, according to the continuous mapping theorem, $q(\rho_{BFN}) \xrightarrow{\mathbb{P}} q(\rho)$. The rest of the proof is straightforward. Let $\varepsilon > 0$. There exists $N \geq 0$ such that for all $n \geq N$, $\mathbb{P}(|q(\rho_{BFN}) - q(\rho)| > \varepsilon) \leq \varepsilon$. Then,

$$\begin{aligned}
|\mathbb{E}(q(\rho_{BFN}) - q(\rho))| &\leq \mathbb{E}(|q(\rho_{BFN}) - q(\rho)|) \\
&= \mathbb{E}(|q(\rho_{BFN}) - q(\rho)|\mathbb{1}_{|q(\rho_{BFN})-q(\rho)|>\varepsilon})+ \\
&\mathbb{E}(|q(\rho_{BFN}) - q(\rho)|\mathbb{1}_{|q(\rho_{BFN})-q(\rho)|\leq\varepsilon}) \\
&\leq 2B\varepsilon + \varepsilon
\end{aligned}$$

The convergence (16) is then established in the balanced case.

The unbalanced case

By applying equations (17) and (18), and considering that A_N is deterministic[9], it comes :

$$\mathbb{E}(\rho_{BFN} - \rho) = \mathbb{E}(p_N(\rho_{BFN}) - p_N(\rho))$$

Contrary to the balanced case, the continuous mapping theorem does not apply here, as p_N depends on N. However, the above demonstration in the balanced case would apply if we were able to show that : $p_N(\rho_{BFN}) - p_N(\rho) \xrightarrow{\mathbb{P}} 0$.

We notice :

$$g_N(\rho) - g_N(\rho_{BFN}) = A_N(\rho - \rho_{BFN}) - A_N(p_N(\rho) - p_N(\rho_{BFN}))$$

which may be rewritten

$$p_N(\rho) - p_N(\rho_{BFN}) = \rho - \rho_{BFN} - \frac{1}{A_N}(g_N(\rho) - g_N(\rho_{BFN}))$$

As we have :

$$\rho_{BFN} - \rho \xrightarrow{\mathbb{P}} 0$$

[9] That is : the pattern of missing data is supposed to be given here, just like N, m or ρ.

$$g_N(\rho_{BFN}) - g_N(\rho) \xrightarrow{\mathbb{P}} 0$$
$$\frac{1}{2} \leq A_N$$

We indeed get :

$$p_N(\rho_{BFN}) - p_N(\rho) \xrightarrow{\mathbb{P}} 0$$

Besides, Annex (E) shows that p_N is bounded uniformly in N :

$$0 \leq p_N(r) \leq \frac{1}{1 - \frac{1}{m}}$$

Hence, we get the announced result (16).

This result is of first-order importance. Section 2.2 showed that the bias of ρ_d could heavily depend on the patterns of the missing values and does not necessarily converge to zero, even when the minimal number of periods m is becoming large. All the opposite, ρ_{BFN} is asymptotically unbiased as $N \to \infty$ even for small values of m.

3.1.4 Two approximations of ρ_{BFN} : ρ_{BFN2B} and ρ_{BFN2U}

What is the rationale behind additional estimates of ρ ?

When $-1 \leq \rho \leq 0$, we do not provide a demonstration that ρ_{BFN} is well defined. In the remaining cases, one might have a use of an estimator of ρ_{BFN} to determine which root is the most appropriate estimator of ρ. Such estimate ρ_{BFN2U} can be obtained thanks to Formula (12). If its last term can be neglected, Formula (12) [10] becomes indeed linear in ρ and provides a unique root when one replaces $E(\rho_d)$ with ρ_d (see formulas (21) and (22) below, respectively for the balanced and unbalanced case).

One may even consider ρ_{BFN2B} and ρ_{BFN2U} rather than ρ_{BFN} as estimators of ρ. Note these approximations could also be interesting when $\rho > 0$ as their computation is numerically almost as simple as the estimation of ρ_d.

Definition :

$$\rho_{BFN2B} = \frac{\rho_d}{(1 - \frac{2}{T})} \tag{21}$$

[10] Or Formula (9) in the balanced case. In which case we call the new estimator ρ_{BFN2B}.

$$\rho_{BFN2U} = \frac{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i} - 1 + \rho_d}{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i}} \tag{22}$$

Asymptotic behavior :
In the balanced case, from the asymptotic behavior of ρ_d as $N \to \infty$, it comes :

$$\rho_{BFN2B} \xrightarrow{\mathbb{P}} \frac{f(\rho)}{1-\frac{2}{T}}$$

where :

$$\frac{f(\rho)}{1-\frac{2}{T}} = \rho + O(\frac{1}{T^2})$$

In the unbalanced case, from the asymptotic behavior of ρ_d as $N \to \infty$, it comes :

$$\rho_{BFN2U} - \frac{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i} - 1 + g_N(\rho)}{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i}} \xrightarrow{\mathbb{P}} 0$$

where :

$$\frac{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i} - 1 + g_N(\rho)}{\frac{1}{N}\sum_i^N \frac{K_i}{1+K_i}} = \rho + O(\frac{1}{m})$$

Expectations :
By dividing equation (9) by $1-\frac{2}{T}$, one gets in the balanced case :

$$E(\rho_{BFN2B}) = \rho + O(\frac{1}{T^2})$$

Similarly, it comes in the unbalanced case :

$$E(\rho_{BFN2U}) = \rho + O(\frac{1}{m})$$

The above properties show that ρ_{BFN2B} and ρ_{BFN2U} have good properties in long panels. Both their limit and their asymptotic expectancy (for large N) are a term whose difference to ρ tends to 0 when the time dimension of the panel (T in the balanced case, or for instance m in the unbalanced case) tends to $+\infty$.

Yet, when the panel is short (in the balanced case) or only moderately long (in the unblanced case), ρ_{BFN} should be strictly preferred when it is possible. For instance, MonteCarlo simulations show that ρ_{BFN} is a much better estimator than ρ_{BFN2U}, even for relatively long panels (e.g. $T = 50$). Hence, the former should be preferred to the later.

3.2 Monte Carlo simulations

3.2.1 In the balanced case

Monte Carlo simulations confirm that, in a given balanced setting ($N = 500$; $T = 10$; $\rho = 0.6$; $\sigma_\eta = 0.3$; $\sigma_\nu = 0.35$), ρ_{BFN} and ρ_{BFN2U} are unbiased[11], unlike all other estimators currently provided by the *xtregar* command (see Table 1). Estimates of σ_η derived from these two estimators seem also unbiased. Yet, they do not seem to improve the estimation of σ_ν.

Comparing the two alternative estimators ρ_{BFN2B} and ρ_{BFN2U} that we suggested in Section 3.1.4, simulation results follow the theoretical properties : the alternative estimator adapted to the balanced case ρ_{BFN2B} performs as well as ρ ; the alternative estimator built for the unbalanced case has a bad performance, due to the small time dimension of the panel (T=10).

3.2.2 In the unbalanced case

Similarly to the balanced case, ρ_{BFN} provides a much better estimate of ρ than all other alternatives currently provided by *xtregar*. However, it does not improve the estimates of σ_η and of σ_ν that are highly biased in all cases.

As in the balanced case, the approximate estimator ρ_{BFN2U} has a very poor performance due to the low time dimension of the panel (T=10). We study in section 3.2.3 what happens when the time dimension of the panel increases.

[11] For ρ_{BFN}, this was already observed by Bhargava *et al.* (1982).

Table 1. Monte Carlo simulations on a balanced panel

	ρ	σ_η	σ_ν
true values	0.6	0.3	0. 35
dw (default)	.464***	.293**	.415***
	(.012)	(3.0e-03)	(.012)
regress	.389***	.292**	.415***
	(.015)	(3.0e-03)	(.012)
freg	.389***	.292**	.415***
	(.014)	(3.0e-03)	(.012)
tscorr	.341***	.293**	.415***
	(.013)	(3.0e-03)	(.012)
theil	.379***	.292**	.415***
	(.014)	(3.0e-03)	(.012)
nagar	.464***	.293**	.415***
	(.012)	(3.0e-03)	(.012)
onestep	.379***	.292**	.415***
	(.014)	(3.0e-03)	(.012)
ρ_{BFN}	**.598**	**.299**	.415***
as in (Bhargava *et al.*, 1982))	**(.017)**	**(3.0e-03)**	(.012)
ρ_{BFN2B}	**.597**	**.299**	.415***
	(.016)	**(3.0e-03)**	(.012)
ρ_{BFN2U}	.405***	.292**	.415***
	(.014)	(3.0e-03)	(.012)

Legend : The average estimators should not be significantly different from the true values. It is the case only for those in bold. Significance levels for the differences with the true values are otherwise pinpointed by stars : * ($p < 0.10$), ** ($p < 0.05$), *** ($p < 0.01$)

Note : This Monte Carlo simulation was performed on a panel of 500 individuals observed each over 10 periods, with 50 replications.

The estimates of σ_η and of σ_ν in the last lines are obtained by estimating first ρ_{BFN}, and then by imposing it as the estimate of ρ in *xtregar*.

Table 2. Monte Carlo simulations on an unbalanced panel

	ρ	σ_η	σ_ν
true valuess	0.6	0.3	0. 35
dw (default)	$.691^{***}$	$.555^{***}$	$.442^{***}$
	(.013)	(.014)	(.013)
regress	$.271^{***}$	$.561^{***}$	$.442^{***}$
	(.027)	(.025)	(.013)
freg	$.272^{***}$	$.562^{***}$	$.442^{***}$
	(.029)	(.025)	(.013)
tscorr	$.115^{***}$	$.396^{***}$	$.442^{***}$
	(.013)	(.017)	(.013)
theil	$.248^{***}$	$.541^{***}$	$.442^{***}$
	(.028)	(.027)	(.013)
nagar	$.691^{***}$	$.555^{***}$	$.442^{***}$
	(.013)	(.014)	(.013)
onestep	$.248^{***}$	$.541^{***}$	$.442^{***}$
	(.028)	(.027)	(.013)
ρ_{BFN}	**.601**	$.616^{***}$	$.442^{***}$
(our generalization)	**(.035)**	(.021)	(.013)
ρ_{BFN2U}	$.326^{***}$	$.603^{***}$	$.442^{***}$
	(.032)	(.022)	(.013)

Legend : The average estimators should not be significantly different from the true values. It is the case only for those in bold. Significance levels for the differences with the true values are otherwise pinpointed by stars : * $(p < 0.10)$, ** $(p < 0.05)$, *** $(p < 0.01)$

Note : Approximately half of a panel of 500 individuals observed each over 10 periods has been randomly deleted, before the Monte Carlo process has been implemented with 50 replications.

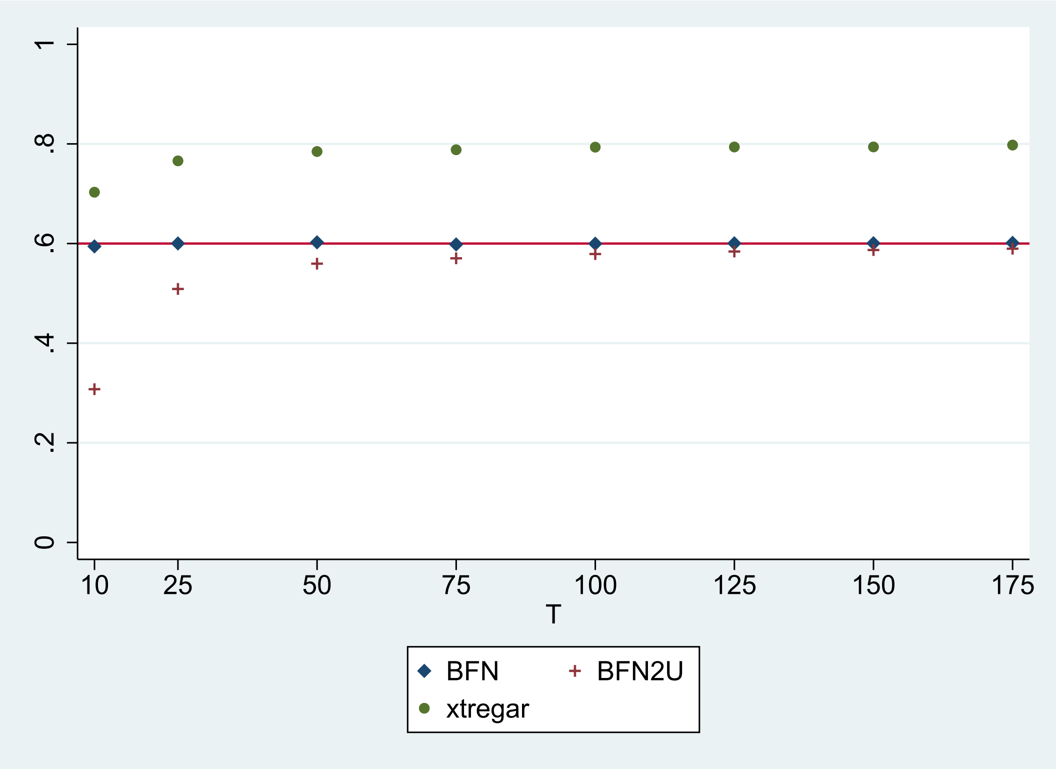
The estimates of σ_η and of σ_ν in the two last lines are obtained by estimating first ρ_{BFN} (or ρ_{BFN2U}), and then by imposing it as the estimate of ρ in *xtregar*.

3.2.3 Monte Carlo simulations when $T \to \infty$

In our previous Monte Carlo simulations with $T = 10$, we observed that ρ_{BFN} (and ρ_{BFN2B} in the balanced case) provided much better estimates than all other existing alternatives.

However, the performances of ρ_{BFN2U} was poor, even relative to currently existing alternatives. As we establish that : $\mathbb{E}(\rho_{BFN2U}) = \rho + O(\frac{1}{m})$, we empirically study how this relative performance evolves when $T \to \infty$, where T is the time span of the panel before we drop observations [12].

Figure 1.
Estimations of ρ with Monte Carlo simulations on an unbalanced panel with $T \to \infty$



Note : This Monte Carlo simulation was performed on a panel of 500 individuals observed each over T periods, with 5 replications. Around half of all observations are each time randomly dropped to get an unbalanced panel. True value of the parameter ρ is 0.6.

[12] Another way could have been to both increase T and decrease the probability that an observation goes missing (currently 0.5).

Unsurprisingly, simulations show that ρ_{BFN2U} converges as T tends to infinity. For T=50, it starts delivering a prediction relatively close to the true value of ρ. More surprisingly, the current estimation of $\hat{\rho}$ does not seem to converge to the true value of ρ. If anything, it would rather converge towards 0.8. Hence, even for $\rho > 0$, if the computations for estimating ρ_{BFN} were computationally intractable due to a too large value of T, ρ_{BFN2U} should be preferred to the current implemented estimator of xtregar.

4 Conclusion

We built upon the work and an intuition of Bhargava *et al.* (1982) to provide a new estimator of the autocorrelation parameter in fixed- or random-effects models with AR(1) disturbances. We show that the suggested estimation method defines at most one estimator, denoted ρ_{BFN}, which is consistent and less biased than current estimates of ρ. For $0 < \rho < 1$, it is asymptotically unbiased as $N \to \infty$, with N the number of individuals.

To take into account specific situations (e.g. when the computation of ρ_{BFN} is numerically too demanding), we defined two additional estimators of ρ that approximate ρ_{BFN} in long panels : ρ_{BFN2B} in balanced panels, and ρ_{BFN2U} in unbalanced panels. Their bias tends however to zero when the time dimension of the panel (i.e. the minimal number of observations per individual) tends to infinity. They are easier to compute than ρ_{BFN}, and perform as well as ρ_{BFN} in our Monte Carlo simulations in long and very long panels. In case computations of ρ_{BFN} are numerically demanding, ρ_{BFN2B} and ρ_{BFN2U} would be our preferred choice in long balanced and very long unbalanced panels respectively.

Monte-Carlo simulations highlight these estimators are usually much better than the current methods used to estimate ρ.[13]

These results highlight a different focus from the canonical works of Bhargava *et al.* (1982) and Baltagi et Wu (1999). While they were mostly interested in testing

[13] A new Stata command **rho_xtregar** has been implemented to enable other researchers to benefit from these improvements in the estimation of ρ.

the nullity (or equality to 1) of the autocorrelation coefficient, this article tries to get the most precise estimate of ρ possible. These approaches are similar but distinct as clearly shown here.

As a caveat in unbalanced panels, the performed estimation methods heavily depend on consecutive observations. In sparse datasets with no consecutive observations, it cannot be implemented. In such cases, other methods like those of Magnac *et al.* (2018) are to be implemented. Moreover, our MonteCarlo simulations show that there is room for improvement in the estimations of other possible parameters of interest such as the variance of the perturbations in unbalanced panels. This is let for further research.

Références

Baltagi, B. H. et Wu, P. X. (1999). Unequally spaced panel data regressions with AR (1) disturbances. *Econometric Theory*, 15(6):814–823.

Bhargava, A., Franzini, L. et Narendranathan, W. (1982). Serial correlation and the fixed effects model. *The Review of Economic Studies*, 49(4):533–549.

Cazenave-Lacroutz, A., Godet, F. et Lin, V. (2019). Modélisation des trajectoires de revenus d'activité pour le modèle destinie 2.

Dafny, L. S. (2010). Are health insurance markets competitive? *American Economic Review*, 100(4):1399–1431.

Hau, H., Langfield, S. et Marques-Ibanez, D. (2013). Bank ratings: what determines their quality? *Economic Policy*, 28(74):289–333.

Magnac, T., Pistolesi, N. et Roux, S. (2018). Post-Schooling Human Capital Investments and the Life Cycle of Earnings. *Journal of Political Economy*, 126(3): 1219–1249.

Nagar, A. L. (1959). The bias and moment matrix of the general k-class estimators of the parameters in simultaneous equations. *Econometrica*, pages 575–595.

Annexes

A Intuitions behind the generalisation of d_p to the unbalanced case

We give here the intuition lying behind formula (10) that defines d_p.

Firstly :

$$\mathbb{E}(\sum_{j=2}^{n_i}(u_{it_{ij}}-u_{it_{ij-1}})^2\mathbb{1}_{t_{ij}-t_{ij-1}=1}) = \sum_{j=2}^{n_i}\mathbb{E}((u_{it_{ij}}-u_{it_{ij-1}})^2)\mathbb{1}_{t_{ij}-t_{ij-1}=1} = K_i((1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2) \tag{23}$$

hence $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{K_i+1}\sum_{j=2}^{n_i}(\tilde{u}_{it_{ij}}-\tilde{u}_{it_{ij-1}})^2\mathbb{1}_{t_{ij}-t_{ij-1}=1}$ is a natural estimate of $(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2$. Let us note A this estimate.

Secondly :

$$\mathbb{E}(\frac{1}{n_i}\sum_{j=1}^{n_i}(u_{it_{ij}}-\overline{u}_i)^2) = \mathbb{E}(\frac{1}{n_i}\sum_{j=1}^{n_i}u_{it_{ij}}^2-(\overline{u}_i)^2) = \sigma_u^2-\mathbb{E}(\overline{u}_i^2) \tag{24}$$

One develops $\mathbb{E}(\overline{u}_i^2)$:

$$\mathbb{E}(\overline{u}_i^2) = \frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\mathbb{E}(u_{it_{ij}}u_{it_{ik}}) = \frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\frac{\rho^{|t_{ij}-t_{ik}|}}{1-\rho^2}\sigma_\varepsilon^2 = \frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\sigma_u^2 \tag{25}$$

hence :

$$\mathbb{E}(\frac{1}{n_i}\sum_{j=1}^{n_i}(u_{it_{ij}}-\overline{u}_i)^2) = \sigma_u^2(1-\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}) \tag{26}$$

By using relation (6), we now show that, as $n_i\to\infty$, $\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\to 0$.
We indeed notice that : $|t_{ij}-t_{ik}|\geq|j-k|$; hence $|\rho|^{|t_{ij}-t_{ik}|}\leq|\rho|^{|j-k|}$ as $|\rho|\leq 1$. Therefore :

$$|\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}|\leq\sum_{j,k=1}^{n_i}|\rho|^{|t_{ij}-t_{ik}|}\leq\sum_{j,k=1}^{n_i}|\rho|^{|j-k|}=\mathcal{O}(n_i) \tag{27}$$

hence $\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}=\mathcal{O}(\frac{1}{n_i})$. The later terms tends to zero when $n_i\to\infty$. Therefore $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{u}_{it_{ij}}^2$ is a natural estimate of σ_u^2. We note B this estimateur.

By making the ratio $\frac{A}{B}$, one gets formula (10). Therefore, it provides a natural estimate of $\frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{\sigma_u^2} = 2 - 2\rho$.

B Derivation of the expectancy of ρ_d in the unbalanced case

We follow the demonstration of Bhargava *et al.* (1982). For N large enough :

$$\mathbb{E}(d_p) \simeq \mathbb{E}\left(\frac{\sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} (u_{it_{ij}} - u_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1}}{\sum_{i=1}^{N} \frac{1}{n_i} \sum_{j=1}^{n_i} (u_{it_{ij}} - \overline{u}_i)^2}\right) \tag{28}$$

Adapting to our data pattern the Nagar (1959) approximation used by Bhargava *et al.* (1982), one gets :

$$\mathbb{E}(d_p) \simeq \frac{\mathbb{E}\left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} (u_{it_{ij}} - u_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1}\right)}{\mathbb{E}\left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i} \sum_{j=1}^{n_i} (u_{it_{ij}} - \overline{u}_i)^2\right)} \tag{29}$$

On the one hand, the numerator is :

$$\begin{aligned} &\mathbb{E}\left(\frac{1}{N} \sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} (u_{it_{ij}} - u_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1}\right) = \\ &\frac{1}{N} \sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} \mathbb{E}((u_{it_{ij}} - u_{it_{ij-1}})^2) \mathbb{1}_{t_{ij}-t_{ij-1}=1} = \\ &\frac{1}{N} \sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} ((1-\rho)^2 \sigma_u^2 + \sigma_\varepsilon^2) \mathbb{1}_{t_{ij}-t_{ij-1}=1} = \\ &\frac{(1-\rho)^2 \sigma_u^2 + \sigma_\varepsilon^2}{N} \sum_{i=1}^{N} \frac{K_i}{1+K_i} \end{aligned} \tag{30}$$

On the other hand, the denominator is :

$$
\begin{aligned}
\mathbb{E}\left(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}(u_{it_{ij}}-\overline{u}_i)^2\right) &= \\
\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}(u_{it_{ij}}-\overline{u}_i)^2\right) &= \\
\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}u_{it_{ij}}^2-(\overline{u}_i)^2\right) &= \\
\frac{1}{N}\sum_{i=1}^{N}(\sigma_u^2-\mathbb{E}(\overline{u}_i^2)) &= \\
\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_u^2-\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\sigma_u^2\right) &= \\
\sigma_u^2\left(1-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\right)
\end{aligned}
\tag{31}
$$

Hence :

$$
\begin{aligned}
\mathbb{E}(d_p) &\simeq \frac{\frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{N}\sum_{i=1}^{N}\frac{K_i}{1+K_i}}{\sigma_u^2\left(1-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\right)} \\
&= \frac{2(1-\rho)}{N}\frac{\sum_{i=1}^{N}\frac{K_i}{1+K_i}}{1-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}}
\end{aligned}
\tag{32}
$$

Therefore

$$
\begin{aligned}
\mathbb{E}(\rho_d) &= 1-\frac{1}{2}\mathbb{E}(d_p) \\
&\simeq 1-\frac{(1-\rho)\sum_{i=1}^{N}\frac{K_i}{1+K_i}}{N-\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}}
\end{aligned}
\tag{33}
$$

Note this relation is valid provided N is large enough.

C Asymptotic behavior of ρ_d

We study the asymptotic behavior of the estimator ρ_d defined via formula (10). More precisely, we show that ρ_d is close to the right member of formula (11) as $N \to \infty$. In what follows, we denote $\ddot{y}_{it} = y_{it} - \overline{y}_i$, $\ddot{x}_{it} = x_{it} - \overline{x}_i$ and $\ddot{u}_{it} = u_{it} - \overline{u}_i$. We denote by $\tilde{u}_{it}$ the OLS residuals from the regression of $\ddot{y}_{it}$ on $\ddot{x}_{it}$.

C.1 The case of the numerator of d_p

Writing

$$\begin{aligned}(\tilde{u}_{it_{ij}} - \tilde{u}_{it_{ij-1}})^2 &= (u_{it_{ij}} - u_{it_{ij-1}} + (\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'(\beta - \hat{\beta}))^2 \\ &= (u_{it_{ij}} - u_{it_{ij-1}})^2 \\ &+ (\beta - \hat{\beta})'(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'(\beta - \hat{\beta}) \\ &+ 2(u_{it_{ij}} - u_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'(\beta - \hat{\beta})\end{aligned} \tag{34}$$

one has

$$\begin{aligned}&\frac{1}{K_i+1}\sum_{j=2}^{n_i}(\tilde{u}_{it_{ij}} - \tilde{u}_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1} \\ &= \frac{1}{K_i+1}\sum_{j=2}^{n_i}(u_{it_{ij}} - u_{it_{ij-1}})^2 \mathbb{1}_{t_{ij}-t_{ij-1}=1} \\ &+ \frac{1}{K_i+1}\sum_{j=2}^{n_i}(\beta - \hat{\beta})'(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'(\beta - \hat{\beta})\mathbb{1}_{t_{ij}-t_{ij-1}=1} \\ &+ \frac{2}{K_i+1}\sum_{j=2}^{n_i}(u_{it_{ij}} - u_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'(\beta - \hat{\beta})\mathbb{1}_{t_{ij}-t_{ij-1}=1}\end{aligned} \tag{35}$$

Denoting by A_i, B_i, C_i the first, the second and the third term of the right hand side of this equality respectively, we shall prove the following probability convergences, as $N \to \infty$:

$$\frac{1}{N}\sum_{i=1}^{N} A_i - \frac{(1-\rho)^2\sigma_u^2 + \sigma_\varepsilon^2}{N}\sum_{i=1}^{N}\frac{K_i}{K_i+1} \to 0 \tag{36}$$

$$\frac{1}{N}\sum_{i=1}^{N} B_i \to 0 \tag{37}$$

$$\frac{1}{N}\sum_{i=1}^{N} C_i \to 0 \tag{38}$$

which will prove

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{K_i+1}\sum_{j=2}^{n_i}(\tilde{u}_{it_{ij}}-\tilde{u}_{it_{ij-1}})^2\mathbb{1}_{t_{ij}-t_{ij-1}=1}-\frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{N}\sum_{i=1}^{N}\frac{K_i}{K_i+1}\to 0 \quad (39)$$

We make the following assumption

Assumption 1 : There exists $\delta > 0$ such that for all i and t, $\mathbb{E}(|u_{it}|^{2+\delta}) < \infty$.

C.1.1 Proof of Assertion (36)

Since $\mathbb{E}((u_{it_{ij}} - u_{it_{ij-1}})^2)\mathbb{1}_{t_{ij}-t_{ij-1}=1} = ((1-\rho)^2\sigma_u^2 + \sigma_\varepsilon^2)\mathbb{1}_{t_{ij}-t_{ij-1}=1}$, one has

$$\mathbb{E}\left(\frac{1}{K_i+1}\sum_{j=2}^{n_i}(u_{it_{ij}}-u_{it_{ij-1}})^2\mathbb{1}_{t_{ij}-t_{ij-1}=1}\right)=\frac{K_i}{K_i+1}((1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2)$$

.

Using the independence of the A_i's and assumption 1, we may apply some version of the law of large numbers (see below) to obtain assertion (36).

Law of large number for independent non-identically distributed random variables : Let (X_i) be a sequence of independent random variables such that $\mathbb{E}[|X_i|^{1+\delta}] < \infty$ for some $\delta > 0$ and all i. Then, almost surely,

$$\frac{1}{N}\sum_{i=1}^{N}X_i-\frac{1}{N}\sum_{i=1}^{N}\mathbb{E}(X_i)\xrightarrow[N\to\infty]{}0$$

C.1.2 Proof of Assertion (37)

We make the following assumption

Assumption 2 : Let K denote the number of regressors. Then the $K \times K$ semi-definite positive matrix

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{K_i+1}\sum_{j=2}^{n_i}(\ddot{x}_{it_{ij}}-\ddot{x}_{it_{ij-1}})(\ddot{x}_{it_{ij}}-\ddot{x}_{it_{ij-1}})'\mathbb{1}_{t_{ij}-t_{ij-1}=1}$$

is uniformly bounded in N. That is, if M_N denotes this matrix, then there exists

$C > 0$ such that for all $N \geq 0$,

$$||M_N|| \leq C,$$

where $||.||$ denotes some norm over the $K \times K$ matrix space.

Under this assumption and recalling that $\hat{\beta}$ is consistent, we get assertion (37).

C.1.3 Proof of Assertion (38)

Using strict exogeneity, we have

$$\begin{aligned}
&\mathbb{E}\left((u_{it_{ij}} - u_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})|x_{it_{i1}}, ..., x_{it_{in_i}}\right) \\
&= \mathbb{E}\left((u_{it_{ij}} - u_{it_{ij-1}})|x_{it_{i1}}, ..., x_{it_{in_i}}\right)(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}}) \\
&= 0.
\end{aligned}$$

hence $\mathbb{E}\left((u_{it_{ij}} - u_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})\right) = 0$.
Then, using the law of large numbers as previously,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{K_i+1}\sum_{j=2}^{n_i}(u_{it_{ij}} - u_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})'\mathbb{1}_{t_{ij}-t_{ij-1}=1} \to 0.$$

We conclude by recalling that $\hat{\beta}$ is consistent.

C.2 The case of the denominator of d_p

Writing

$$\begin{aligned}
\tilde{u}^2_{it_{ij}} &= ((\ddot{x}_{it_{ij}})'(\beta - \hat{\beta}) + \ddot{u}_{it_{ij}})^2 \\
&= \ddot{u}^2_{it_{ij}} \\
&+ (\beta - \hat{\beta})'(\ddot{x}_{it_{ij}})(\ddot{x}_{it_{ij}})'(\beta - \hat{\beta}) \\
&+ 2\ddot{u}_{it_{ij}}(\ddot{x}_{it_{ij}})'(\beta - \hat{\beta})
\end{aligned}$$

one has

$$\begin{aligned}
&\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{u}_{it_{ij}}^2\\
&=\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}\ddot{u}_{it_{ij}}^2\\
&+\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}(\beta-\hat{\beta})'(\ddot{x}_{it_{ij}})(\ddot{x}_{it_{ij}})'(\beta-\hat{\beta})\\
&+\frac{1}{N}\sum_{i=1}^{N}\frac{2}{n_i}\sum_{j=1}^{n_i}\ddot{u}_{it_{ij}}(\ddot{x}_{it_{ij}})'(\beta-\hat{\beta})
\end{aligned}$$

Following the same arguments as for the proofs of (37) and (38), we show that the second and the third term of the right hand side of the equality converge to 0 as $N \to \infty$. Using the same argument as in the proof of (36) and the equality (31), we get, as $N \to \infty$,

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{u}_{it_{ij}}^2-\sigma_u^2\left(1-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\right)\to 0, \tag{40}$$

the convergence holding in probability.

C.3 Reconciliation

Let us now write the result of section C.1 with obvious notations :

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{K_i+1}\sum_{j=2}^{n_i}(\tilde{u}_{it_{ij}}-\tilde{u}_{it_{ij-1}})^2\mathbb{1}_{t_{ij}-t_{ij-1}=1}-\frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{N}\sum_{i=1}^{N}\frac{K_i}{K_i+1}=u_N-v_N\to 0$$

And the result of section C.2 :

$$\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i}\sum_{j=1}^{n_i}\tilde{u}_{it_{ij}}^2-\sigma_u^2\left(1-\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\rho^{|t_{ij}-t_{ik}|}\right)=w_N-x_N\xrightarrow{\mathbb{P}}0$$

Writing between braces differences that converge towards zero, we notice :

$$\frac{u_N}{w_N}=\frac{v_N}{x_N}(1+\frac{\{u_N-v_N\}}{v_N}-\frac{\{w_N-x_N\}}{x_N+\{w_N-x_N\}}-\frac{\{u_N-v_N\}\{w_N-x_N\}}{v_N(x_N+\{w_N-x_N\})})$$

We use the following lemma :

Lemma 2 : *There exists three numbers strictly positive K_1, K_2, and N_0 such as for all $N \geq N_0$:*

$$\textit{such as : } K_1 \leq x_N$$

$$\textit{and such as : } K_1 \leq v_N \leq K_2$$

From this lemma, we get immediately :

$$\frac{u_N}{w_N} - \frac{v_N}{x_N} \xrightarrow{\mathbb{P}} 0$$

That is :

$$\rho_d - \left(1 - \frac{(1-\rho)\sum_{i=1}^{N} \frac{K_i}{1+K_i}}{N - \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|}} \right) = \rho_d - g_N(\rho) \xrightarrow{\mathbb{P}} 0 \tag{41}$$

Note also that from Annex (B), one has :

$$g_N(\rho) - \mathbb{E}(\rho_d) \xrightarrow{\mathbb{P}} 0$$

Hence, we can even conclude :

$$\rho_d - \mathbb{E}(\rho_d) \xrightarrow{\mathbb{P}} 0 \tag{42}$$

C.4 Justification of Assumption 2 and demonstration of Lemma 2

We recall Assumption 2 and Lemma 2.

Assumption 2 : Let K denote the number of regressors. Then the $K \times K$ semi-definite positive matrix

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{K_i+1} \sum_{j=2}^{n_i} (\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})(\ddot{x}_{it_{ij}} - \ddot{x}_{it_{ij-1}})' \mathbb{1}_{t_{ij}-t_{ij-1}=1}$$

is uniformly bounded in N. That is, if M_N denotes this matrix, then there exists $C > 0$ such that for all $N \geq 0$,

$$||M_N|| \leq C,$$

where $||.||$ denotes some norm over the $K \times K$ matrix space.

This assumption is reasonnable because it is fair to assume that the vector of covariables x_{it} is bounded. For instance, in a wage equation model, x_{it} is typically made of variables such as experience, age, level of education, spell of unemployment, which obviously take a finite number of values.

Lemma 2 : *There exists three numbers strictly positive K_1, K_2 and N_0 such as for all $N \geq N_0$:*

$$\textit{such as : } K_1 \leq x_N = \sigma_u^2 \left(1 - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|} \right)$$

$$\textit{and such as : } K_1 \leq v_N = \frac{(1-\rho)^2\sigma_u^2 + \sigma_\varepsilon^2}{N} \sum_{i=1}^{N} \frac{K_i}{K_i+1} \leq K_2$$

The existence of K_1 and K_2 in the second equation is obvious. It is sufficient to choose $K_1 = \frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{2}$ as we restricted ourselves to individuals with $K_i \geq 1$, and $K_2 = (1-\rho)^2\sigma_u^2 + \sigma_\varepsilon^2$.
For the first equation, we notice

$$\frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|} \leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} |\rho|^{|t_{ij}-t_{ik}|} \leq \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} |\rho|^{|j-k|}$$

Hence :

$$x_N = \sigma_u^2 \left(1 - \frac{1}{N} \sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} \rho^{|t_{ij}-t_{ik}|} \right) \geq \sigma_u^2 \frac{1}{N} \sum_{i=1}^{N} h(|\rho|, n_i)$$

with :

$$h(r, n_i) = 1 - \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} r^{|j-k|}$$

and with :

$$\frac{\partial h(r, n_i)}{\partial r} = -\frac{1}{n_i^2} \sum_{|j-k|\geq 1}^{n_i} |j-k| r^{|j-k|-1} \leq 0$$

and $h(1, n_i) = 0$. Hence, for all n_i, for every $-1 < \rho < 1 : h(|\rho|, n_i) > h(1, n_i) = 0$.

The integer n_i takes only a finite number of values : those between 2 and T (defined here as the maximal length of the time dimension). For a given $|\rho| < 1$, let us write $M(\rho) = min_i(h(|\rho|, n_i)) = h(|\rho|, n_I) > h(1, n_I) = 0$.
Then, for every $-1 < \rho < 1 : x_N \geq \sigma_u^2 M(\rho) > 0$.

We may choose $K_1 = min(\frac{(1-\rho)^2\sigma_u^2+\sigma_\varepsilon^2}{2}, \sigma_u^2 M(\rho)) > 0$.

D Lemma 1 : unicity of the ρ_{BFN} estimator

We want to prove :

Lemma 1 : $g_N : r \rightarrow 1 - \frac{(1-r)\sum_{i=1}^{N}\frac{K_i}{1+K_i}}{N-\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}}$ establishes a bijection from $[0,1]$ to $[1 - \frac{\sum_i^N \frac{K_i}{1+K_i}}{N-\sum_i^N \frac{1}{n_i}}, g(1)]$.

It is clear that the denominator of the above function g_N is not nul over [0-1[; hence g_N is derivable. One derives it. The derivative has the same sign as :

$$h(r) = N - \sum_i^N \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{i,j}-t_{i,k}|} - (1-r)\sum_i^N \frac{1}{n_i^2}\sum_{|t_{i,j}-t_{i,k}|\geq 1} |t_{i,j}-t_{i,k}| r^{|t_{i,j}-t_{i,k}|-1}$$

with :

$$h(1) = 0$$

Its derivative is negative over $[0,1]$, and even strictly negative over $[0,1[$ (as soon as there is at least one individual with at least three observations, which we assume from now on) :

$$h'(r) = -(1-r)\sum_i^N \frac{1}{n_i^2}\sum_{|t_{i,j}-t_{i,k}|\geq 2}^{n_i} |t_{i,j}-t_{i,k}|(|t_{i,j}-t_{i,k}|-1) r^{|t_{i,j}-t_{i,k}|-2} < 0$$

Hence, h is strictly decreasing over [0; 1] ; h is always positive ; g' is always positive ; g is strictly increasing and goes from g(0) to g(1).

Eventually :

$$g_N(0) = 1 - \frac{\sum_i^N \frac{K_i}{1+K_i}}{N - \sum_i^N \frac{1}{n_i}}$$

Note also that **in the balanced case** :

$$g_N(0) = f(0) = 0$$

$$g_N(1) = f(1) = 1 - \frac{1}{1+\frac{T-2}{3}}$$

Moreover, Annex E shows that p_N (see Formula (45)) is bounded on $[0,1]$. It follows, from Formula (18), that g_N is also capped. As g_N is monotonous and continuous on $[0,1[$, it may thus be continuously extended on $[0,1]$.

Besides, note eventually that g_N is C^1 over $[0,1[$ and that $\lim_{r\to 1} g_N^{'}(r)$ exists (see Annex F). Hence, g_N is C^1 over $[0,1]$.

Hence, there exists a continuous $(g_N)^{-1}$ over the image of g_N. It also quickly follows that (g_N^{-1}) is C^1 over $[(g_N)(0),(g_N(1)]$.

Indeed, with the notations introduced above :

$$g_N'(r) = (\sum_i^N \frac{K_i}{1+K_i})\frac{(h(r)}{(N - \sum_{i=1}^N \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|})^2}$$

As both the numerator and the denominator are strictly positive on $[0,1[$ (see above ; a sufficient and necessary condition for h to be strictly positive is that there is at least one individual with at least three observations, which we assumed to be the case) : for all $0 \leq r < 1$: $g_N'(r) > 0$.

For all $0 \leq r < 1$, one gets :

$$(g_N^{-1})'(g_N(r)) = \frac{1}{g_N'(r)}$$

Hence : g_N^{-1} is C^1 over the interior of the image of g_N.

Note that the above demonstration also shows that : $(g_N^{-1})'(g_N(0)) > 0$.

Note also that Annex F shows that :

$$g_N'(1) = A_N \frac{\frac{1}{N}\sum_{i=1}^N \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} \frac{|t_{ij}-t_{ik}|(|t_{ij}-t_{ik}|-1)}{2}}{(\frac{1}{N}\sum_{i=1}^N \frac{1}{n_i^2}\sum_{j,k=1}^{n_i} |t_{ij}-t_{ik}|)^2}$$

This latter term is strictly positive as soon as there is at least one observation with at least three observations. In those cases, we have also that $(g_N^{-1})'(g_N(1)) > 0$.

E Lemma 3 : Capping $p_N(r)$

Lemma 3 : for $m \geq 3$, for all $0 < r < 1$: $|p(r)| \leq \frac{1}{1-\frac{1}{m}}$ where :

$$p_N(r) = \frac{1-r}{\frac{N}{\sum_{i=1}^N \frac{1}{n_i^2} \sum_{j,k}^{n_i} r^{|t_{i,j}-t_{i,k}|}} - 1}$$

For $0 < r < 1$:

$$\frac{1}{n_i^2} \sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|} \leq \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} r^{|j-k|} = \frac{1+r}{(1-r)n_i} - \frac{2r}{(1-r)n_i} \frac{1-r^{n_i}}{(1-r)n_i} = l(r, n_i)$$

Hence :

$$0 \leq p_N(r) \leq \frac{1-r}{\frac{1}{\frac{1}{N}\sum_{i=1}^N l(r,n_i)} - 1} \tag{43}$$

$$= \frac{(1-r)\frac{1}{N}\sum_{i=1}^N l(r,n_i)}{\frac{1}{N}\sum_{i=1}^N (1 - l(r,n_i))} \tag{44}$$

We consider a given $M = n_i$. We apply in Annex G the Taylor's theorem, about the mean-value form of the remainder, at order 2, to $x \mapsto (1-x)^M$, where $1 - r = x$. One gets :

$$1 - l(r, M) \geq (1 - \frac{1}{M})(1-r)$$

We notice :

$$l(r, M) \leq 1$$

$$1 - l(r, M) \geq (1 - \frac{1}{M})(1-r)$$

Hence :

$$\frac{1}{N}\sum_{i=1}^N l(r, n_i) \leq 1$$

$$\frac{1}{N}\sum_{i=1}^N (1 - l(r, n_i)) \geq \frac{1}{N}\sum_{i=1}^N (1 - \frac{1}{n_i})(1-r)$$

$$\frac{(1-r)\frac{1}{N}\sum_{i=1}^N l(r,n_i)}{\frac{1}{N}\sum_{i=1}^N 1 - l(r,n_i)} \leq \frac{1-r}{\frac{1}{N}\sum_{i=1}^N (1 - \frac{1}{n_i})(1-r)}$$

We can conclude that over [0,1] :

$$0 \leq p_N(r) = \frac{1-r}{\frac{N}{\sum_{i=1}^{N} \frac{1}{n_i^2} \sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}} - 1} \leq \frac{1}{\frac{1}{N}\sum_{i=1}^{N}(1-\frac{1}{n_i})} \leq \frac{1}{1-\frac{1}{m}} \tag{45}$$

F Existence of a limit for g'_N in 1

We recall that :

$$g_N(r) = 1 - \frac{(1-r)A_N}{1 - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|}}$$

where : $\frac{1}{2} \leq A_N = \frac{1}{N}\sum_{i=1}^{N}\frac{K_i}{1+K_i} \leq 1$

$$g'_N(r) = A_N \frac{1 - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|} - (1-r)\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{|t_{ij}-t_{ik}|\geq 1}^{n_i} |t_{ij}-t_{ik}| r^{|t_{ij}-t_{ik}|-1}}{(1 - \frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i} r^{|t_{ij}-t_{ik}|})^2}$$

$$g'_N(r) = A_N \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}(1 - r^{|t_{ij}-t_{ik}|} - (1-r)|t_{ij}-t_{ik}| r^{|t_{ij}-t_{ik}|-1})}{(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}(1 - r^{|t_{ij}-t_{ik}|}))^2}$$

$$g'_N(r) = A_N \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\lambda(|t_{ij}-t_{ik}|, r)}{(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}(1 - r^{|t_{ij}-t_{ik}|}))^2} \tag{46}$$

where we note $\lambda(m,r) = 1 - r^m - (1-r)mr^{m-1}$, which satisfies $\lambda(0,r) = \lambda(1,r) = 0$. Let $m = |t_{ij} - t_{ik}|$. By using Taylor developments around 1, we get

$$\lambda(m,r) = \frac{m(m-1)}{2}(1-r)^2 + \ o((1-r)^2)$$

and

$$1 - r^m = m(1-r) + \ o(1-r)$$

which yields :

$$g'_N(r) = A_N \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\frac{|t_{ij}-t_{ik}|(|t_{ij}-t_{ik}|-1)}{2} + \ o(1)}{(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}|t_{ij}-t_{ik}| + \ o(1))^2}$$

$$g'_N(r) \xrightarrow[r\to 1]{} A_N \frac{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}\frac{|t_{ij}-t_{ik}|(|t_{ij}-t_{ik}|-1)}{2}}{(\frac{1}{N}\sum_{i=1}^{N}\frac{1}{n_i^2}\sum_{j,k=1}^{n_i}|t_{ij}-t_{ik}|)^2} \tag{47}$$

G A Taylor development

We consider : $l(r,n) = \frac{1}{n^2}\sum_{j,k=1}^{n} r^{|j-k|} = \frac{1+r}{(1-r)n} - \frac{2r}{(1-r)n}\frac{1-r^n}{(1-r)n}$

We consider a given $M = n_i$. We apply the Taylor's theorem, about the mean-value form of the remainder to $x- > (1-x)^M$, where $1-r = x; r = 1-x$. It shows : There exists $0 < \zeta < 1$ such that :

$$r^M = (1-x)^M = 1 - Mx + \frac{M(M-1)}{2}x^2 - \frac{M(M-1)(M-2)}{6}(1-\zeta)^{M-3}x^3$$

$$1 - r^M = Mx - \frac{M(M-1)}{2}x^2 + \frac{M(M-1)(M-2)}{6}(1-\zeta)^{M-3}x^3$$

$$\frac{1-r^M}{(1-r)M} = \frac{1-(1-x)^M}{Mx} = 1 - \frac{(M-1)}{2}x + \frac{(M-1)(M-2)}{6}(1-\zeta)^{M-3}x^2 \tag{48}$$

$$\begin{aligned}
&1 + r - 2r\frac{1-r^M}{(1-r)M} \\
&= 1 + r - 2r + 2r\frac{(M-1)}{2}x - 2r\frac{(M-1)(M-2)}{6}(1-\zeta)^{M-3}x^2 \\
&= x + (1-x)(M-1)x - (1-x)\frac{(M-1)(M-2)}{3}(1-\zeta)^{M-3}x^2 \\
&= Mx - (M-1)x^2 - (1-x)\frac{(M-1)(M-2)}{3}(1-\zeta)^{M-3}x^2
\end{aligned}$$

Hence :

$$l(r,M) = \frac{1}{(1-r)M}(1 + r - 2r\frac{1-r^M}{(1-r)M}) = 1 - (M-1)x - (1-x)\frac{(M-1)(M-2)}{3M}(1-\zeta)^{M-3}x$$

It comes :

$$1 - l(r,M) \geq (1 - \frac{1}{M})(1-r) \tag{49}$$