行政院國家科學委員會專題研究計畫 成果報告

自我迴歸下偏誤的降低：簡單組合估計式之建構(第2年)
研究成果報告(完整版)

計 畫 類 別 ： 個別型
計 畫 編 號 ： NSC 98-2410-H-004-039-MY2
執 行 期 間 ： 99年08月01日至100年12月31日
執 行 單 位 ： 國立政治大學國際貿易學系

計 畫 主 持 人 ： 郭炳伸

報 告 附 件 ： 出席國際會議研究心得報告及發表論文

公 開 資 訊 ： 本計畫涉及專利或其他智慧財產權，2年後可公開查詢

中 華 民 國 101年05月03日

中 文 摘 要： 本文發現二階差分自我迴歸模型，若加入簡單轉換後，自我迴歸係數（autoregressive coefficient）無論是 1 或絕對值小於 1，皆可證明其極限為常態分配。而且該常態極限不會因為模型內有無趨勢項（time trend）而有所改變。這不同於 Phillips and Han（2008）將一階差分自我迴歸模型加以轉換後所獲得的極限性質。

中文關鍵詞： 自我迴歸模型、一階差分、二階差分、單根

英 文 摘 要：

英文關鍵詞：

Gaussian Inference in General AR(1) Models Based on Second Difference*

Jhih-Gang Chen †
National Chengchi University

Biing-Shen Kuo‡
National Chengchi University

March. 2012

Abstract

This paper develops a simple second-difference transformation for estimation and inference in general AR(1) models. As in Phillips and Han (2008, *Econometric Theory* 24, 631−650), a Gaussian limit theory with a convergence rate of $\sqrt{T}$ is available, whether or not a unit root is present in the process. Yet, the novelties of our limit results are that the same weak convergence applies to the models with or without a trend, and that the asymptotic distribution is characterized by a constant variance of value 2. The merits promise usefulness of the second-difference transformation in applications to dynamic panels.

Keywords: AR model, first difference, second difference, unit root.

JEL codes: C13, C22.

*Both Chen and Kuo gratefully acknowledge financial support from the National Science Council of Taiwan.

†Dept. of International Business, National Chengchi University, Taipei 116, Taiwan; Tel: (02)29393091 ext. 81262, Fax: (02)29387699, E-mail: jgchen@nccu.edu.tw.

‡Corresponding author, Dept. of International Business, National Chengchi University, Taipei 116, Taiwan; Tel: (02)29393091 ext. 81029, Fax: (02)29387699, Email: bsku@nccu.edu.tw.

1 Motivation

Difference-type transformation has been one of the commonly-employed practices in time series-related research. In studies with dynamic panel data, differencing eliminates the need to estimate fixed-effect parameters. In cases where the data are generated by a unit root series, by doing so, the stationarity of the series under transformation can be achieved. The transformations, such as the Prais-Winsten estimator and Cochrane-Orcutt estimator, are also useful in reducing the degree of autocorrelation in the regression errors. The efficiency of the resulting estimates can be made close to the bound attained by the Gauss-Markov theorem. Furthermore, Phillips and Han (2008) as well as Paparoditis and Politis (2000) show that the difference-based method leads to the standard Gaussian asymptotic theory for the AR(1) series, irrespective of whether the series is a unit root process.

The clever transformation of Phillips and Han (2008) comes from the observation that the autoregressive coefficient of concern is in a one-to-one linear relationship with the first-order autocorrelation coefficient for the difference series. Instead of the level regressions, the autoregressive coefficient can now be recovered by working on the transformed regressions. Because of the stationarity nature of the difference series, the unusual limit normality is able to be obtained for all values of the autoregressive coefficient, including the case of a unit root. It thus avoids the discontinuity problem in the limit distributions associated with the level regressions.

This paper investigates whether the aforementioned transformation is unique to the considered AR(1) models. Of particular concern is whether any linear relations that could yield standard limit results remain to be found, when a "higher-order-difference" technique applies. Our inquiry into the question arises from the potential uses of the difference-based estimator. Generally speaking, regardless of which order difference is taken for the AR(1) process, to obtain consistency for the parameter entails the estimation of the method of moments. While the method of moments estimator is useful to achieve the consistency, it is found to suffer from small-sample bias. On the contrary, Han and Phillips (2010) show that their difference-based estimator, when applied to dynamic panel data models, basically incurs no bias, even in the cases where the time dimension is very short. More than this, their estimator is immune to the weak instrument problem that occurs in the cases of some of the widely-used method of moments estimators where the autoregressive coefficient is close to unity. The Phillips and Han estimator is actually based on the notion of processing the first-difference time series under study. Seeing these practical advantages, whether the notion with processing data is applicable to higher-order-difference time series or not equally deserves careful investigation.

We demonstrate that, for the second-difference AR(1) series, another transformation which gives rise to some form of linear relationship between the AR coefficient and the autocorrelation coefficient does in fact exist. Our second-difference estimator possesses a number of interesting properties. As with Phillips and Han (2008), the transformed second-difference estimator of the autoregressive coefficient has a Gaussian limit distribution that is continuous as the autoregressive coefficient passes through unity to local departures from unity. When a time trend in the model is entertained, the aforementioned Gaussian limit theory still holds. The limiting result that is invariant to the presence of a trend arises as a consequence of applying a simple de-meaning technique to the second-difference series, instead of the double-difference method used by Phillips and Han (2008). The normality limit of our estimator for all the considered models is characterized by a constant variance of value 2. Our simulations further reveal that the estimator displays negligible bias for very small samples, as opposed to the conventional least squares estimator. The small bias will prove an advantage when applying the estimator to dynamic panels with or without incidental trends.

Our estimator is connected to the aggregation estimator developed by Han, Phillips and Sul (2010a). Their estimator aggregates moment conditions formulated in differences with different orders, and attains efficiency when all possible moment conditions are included. One specific moment condition exploited in their estimation can be viewed as a temporally balanced version of the moment condition associated with our estimator. The resulting estimator based on the specific moment condition may carry the same limiting properties as ours.

The remainder of this paper is organized as follows. Section 2 presents the linear transformation for the second-difference series. In Section 3, the limiting results for the estimator are derived and spelled out. Section 4 extends the second-difference estimator to models with a time trend. Section 5 concludes.

2 Model and the Transformation

Consider a simple autoregressive model

$$y_t = \alpha + u_t,\ u_t = \rho u_{t-1} + \varepsilon_t,\ t = -1, 0, \cdots, T, \tag{1}$$

where $\varepsilon_t \overset{iid}{\sim} (0, \sigma^2)$, and $\rho \in (-1, 1]$. The initial condition is set to be $u_{-2} = \sum_{j=0}^{\infty} \rho^j \varepsilon_{-2-j}$ when $|\rho| < 1$, or $u_{-2} = \sum_{j=0}^{\kappa} \varepsilon_{-2-j}$ for some fixed integer κ when $\rho = 1$ (e.g., Phillips and Magdalinos, 2009). The model corresponds to a reduced form written as

$$y_t = (1-\rho)\alpha + \rho y_{t-1} + \varepsilon_t \tag{2}$$

from which the data generated constitute a simple unit root process for the boundary case, or a stationarity process when $|\rho| < 1$.

We begin by considering a k^{th}-difference transformation of (2):

$$\triangle_k y_t = \rho \triangle_k y_{t-1} + \triangle_k \varepsilon_t, \tag{3}$$

where k is some positive integer with $k/T \to 0$, and $\triangle_k = (1 - \mathscr{L}^k)$ with $\mathscr{L}$ being the lag operator. Note that $\triangle_k y_{t-1}$ and $\triangle_k \varepsilon_t$ are generally correlated except for the cases where $\rho = 0$ and $k \geq 2$, and inconsistent estimates of the coefficient are to be expected when the least squares principle applies to (4). As in Phillips and Han (2008), we intend to re-establish the orthogonality condition by considering the following transformation:

$$\triangle_k y_t = (\rho - \phi)\triangle_k y_{t-1} + (\triangle_k \varepsilon_t + \phi \triangle_k y_{t-1}) = \theta \triangle_k y_{t-1} + \eta_t, \tag{4}$$

where $\theta = \rho - \phi$, $\eta_t = \triangle_k \varepsilon_t + \phi \triangle_k y_{t-1}$, and the coefficient ϕ should satisfy $\mathrm{E}(\triangle_k y_{t-1} \eta_t) = 0$. Given the fact that $\mathrm{E}(\triangle_k y_{t-1})^2 = 2(1-\rho^k)\sigma^2/(1-\rho^2)$ and $\mathrm{E}[\triangle_k y_{t-1} \triangle_k \varepsilon_t] = -\rho^{k-1}\sigma^2$, we can solve for $\phi = \rho^{k-1}(1-\rho^2)/[2(1-\rho^k)]$, which is only a function of the autoregressive coefficient, given any k. Because $\mathrm{E}(\triangle_k y_{t-1} \eta_t) = 0$, the regression coefficient θ in (4) can be consistently estimated, and the estimated autoregressive coefficient can be recovered from the relationship whereby

$$\theta = \rho - \frac{\rho^{k-1}(1-\rho^2)}{2(1-\rho^k)}.$$

The transformation is made possible by the information regarding the moments from the autoregressive models. When deriving the aforementioned relationship, the two important pieces of information needed are the expectation regarding the square of the explanatory variable $\mathrm{E}(\triangle_k y_{t-1})^2$, and the expectation regarding the cross product of the explanatory variable and the regression error $\mathrm{E}(\triangle_k y_{t-1} \triangle_k \varepsilon_t)$. Without knowing the structure of the autoregression, the expected values of these moment conditions can not be uncovered *a priori*. Thus, the transformation method is not applicable to more general regression models.

Inversion of the autoregressive coefficient from the derived relationship requires solving nonlinear equations. Exceptions are the cases for $k = 1$ and 2. When $k = 1$, $\theta = (\rho - 1)/2$ which leads to the transformation of Phillips and Han (2008), $2\triangle y_t + \triangle y_{t-1} = \rho \triangle y_{t-1} + 2\eta_t$. Furthermore, when $k = 2$, $\theta = \rho/2$. This simple linear relationship, however, yields a new second-difference regression:

$$2\triangle_2 y_t = \rho \triangle_2 y_{t-1} + \zeta_t, \tag{5}$$

where $\zeta_t = 2\triangle_2 y_t - \rho \triangle_2 y_{t-1}$. The regression can be exploited to produce consistent estimates of the autoregressive coefficient with the orthogonality condition that

$$\mathrm{E}(\triangle_2 y_{t-1} \zeta_t) = \mathrm{E}[\triangle_2 y_{t-1}(2\triangle_2 y_t - \rho \triangle_2 y_{t-1})] = 0. \tag{6}$$

3 Large Sample Properties

The least-squares estimator of ρ based on the second-difference regression (5) is found to be

$$\hat{\rho}_{\mathrm{SD}} = \frac{2\sum_{t=1}^{T} \triangle_2 y_{t-1} \triangle_2 y_t}{\sum_{t=1}^{T} \triangle_2 y_{t-1}^2}. \tag{7}$$

We are now in a position to demonstrate the first set of the limit theory concerning the transformed second-difference estimator:

Theorem 1. *If* $\rho \in (-1, 1\,]$, $\sqrt{T}(\hat{\rho}_{\mathrm{SD}} - \rho) \Rightarrow N(0, 2)$, *as* $T \to \infty$.

Like the first-difference estimator (FDE) of Phillips and Han (2008), a Gaussian limit theory applies to the second-difference estimator (SDE), regardless of whether or not a unit root is present in the process. [1] The limit result is also featured by a constant asymptotic variance, as opposed to $2(1+\rho)$ for the FDE. Thus, for the cases with positive autoregressive coefficients, typically seen in applications, the SDE exhibits an increasingly smaller asymptotic variance than that of the FDE, as ρ increases to 1.

Nevertheless, in the case of a unit root, neither of the estimators are as efficient as the ordinary least squares estimator, due to information loss from differencing. This is unlike the 'full aggregation estimator' (FAE) proposed by Han, Phillips and Sul (2010a). The FAE aggregates a set of moment conditions produced by the cross-differencing of different orders:

$$\mathrm{E}[(y_{t-1} - y_{t-1-l})\{(y_t - y_{t-2-l}) - \rho(y_{t-1} - y_{t-1-l})\}] = 0, \;\; l = 1, \cdots, T.$$

In the unit root case, the FAE has been shown to converge at a rate of T, and thus is as efficient as the OLS, as it preserves the signal from various conditions. In the stationary case, the estimator not only retains a usual convergence rate at $\sqrt{T}$ but it is also large-T efficient.

In fact, each of the moment condition used by FAE may yield another new consistent estimator, and the one using that with $l = 2$ is the same as the temporally balanced version of $\hat{\rho}_{\mathrm{SD}}$ obtained by replacing $2\triangle_2 y_t$ by $\triangle_2 y_t + \triangle_2 y_{t-2}$. More specifically, we can define the so-called single-lag estimator for $l = 2$ as $\left(\sum_{t=2}^{T} \triangle_2 y_{t-1}^2\right)^{-1} \left[\sum_{t=2}^{T} (\triangle_2 y_t + \triangle_2 y_{t-2})\triangle_2 y_t\right]$, which has a very similar expression to that in (7). Given that $\mathrm{E}(2\triangle_2 y_t \triangle_2 y_{t-1}) = \mathrm{E}[(\triangle_2 y_t + \triangle_2 y_{t-2})\triangle_2 y_{t-1}]$ by covariance stationarity, the single-lag estimator indeed shares the same asymptotics as our estimator.

When $\rho > 1$, the SDE appears to be inconsistent, because $\triangle_2 y_t$ is not stationary. However, if the system is mildly explosive as long as the condition that $\rho^{2T}/\sqrt{T} \to 0$ is satisfied, the same limit result as in the theorem continues to hold. [2] The condition should include the local-to-unity case as a special one. The theorem thus establishes the continuity of the limit distribution as the autoregressive coefficient passes through unity to the local area of unity.

In Table 1, we provide some simulation results for a finite-sample performance comparison between FDE, SDE, and FAE. The artificial data are generated from (1) with $\alpha = 1$ and $\varepsilon_t \overset{iid}{\sim} N(0,1)$. 100,000 replications are performed. As the table shows, all the 3 estimators exhibit very little bias, and no one seems to outperform the others for all values of ρ. In addition, both the re-scaled finite sample variances of FDE and SDE are very close to their asymptotic values, namely, $2(1+\rho)$ and 2, respectively. The t-ratio based on FDE or SDE is then expected to be well characterized as a standard normal. However, the same is not true of FAE for larger ρ cases in which its re-scaled variance seems to converge to the asymptotic value slowly. [3] Finally, the FAE displays a dominance over the other estimators in terms of the RMSE, this being largely attributed to the reduction in estimation variations. When the observations increase, the FAE exploits more information by including more moment conditions. The marginal effects of additional observations could be quite substantial in small samples for the cases where the values of the coefficient are close to unity.

The very little bias lends to the SDE the applicability in dynamic panel data models with fixed effects. In some detail, suppose cross-sectional data are also available such that $y_{it} = (1-\rho)\alpha_i + \rho y_{it} + \varepsilon_{it}$, with α_i being the individual effects, where $i = 1, \cdots, N$ denotes the cross-section units, and $\varepsilon_{it} \overset{i.i.d.}{\sim} (0,\sigma^2)$. Based on the second-difference approach, we can make use of the moment conditions, $\mathrm{E}[\triangle_2 y_{it-1}(2\triangle_2 y_{it} - \rho\triangle_2 y_{it-1})] = 0$ for all $\rho \in (-1,\ 1]$, to obtain a consistent estimate of ρ. The associated panel second-difference estimator (PSDE) takes the form $\hat{\rho}_{\mathrm{PSD}} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}\triangle_2 y_{t-1}^2\right]^{-1}\left(2\sum_{i=1}^{N}\sum_{t=1}^{T}\triangle_2 y_{t-1}\triangle_2 y_t\right)$. Of significance is that the PSDE may escape from the weak instrument problem from which the conventional generalized method of moments approach suffers, because the moment conditions that the estimator uses are strong for all ρ in the interval. Moreover, a Gaussian limit theory with estimable variances may be available for cases with fixed T but large N, given the panel version law of large numbers and central limit theory established in Han and Phillips (2010). Panel inferences of the autoregressive coefficient thus can be conducted through standard procedures.

Extending our approach to the AR(p) models would be desirable but technically costly. In general, it requires solving a system of p nonlinear equations. Taking AR(2) as an instance, our approach would transform the series to $\triangle_2 y_t = (\rho_1 - \phi_1)\triangle_2 y_{t-1} + (\rho_2 - \phi_2)\triangle_2 y_{t-2} + (\triangle_2\varepsilon_t + \phi_1\triangle_2 y_{t-1} + \phi_2\triangle_2 y_{t-2}) = \theta_1\triangle_2 y_{t-1} + \theta_2\triangle_2 y_{t-2} + \eta_t$, where $\theta_1 = \rho_1 - \phi_1$, $\theta_2 = \rho_2 - \phi_2$, $\eta_t = (\triangle_2\varepsilon_t + \phi_1\triangle_2 y_{t-1} + \phi_2\triangle_k y_{t-2})$. Estimating the coefficients ϕ_i now implies that the sample moment conditions of $\mathrm{E}(\triangle_2 y_{t-1}\eta_t) = \mathrm{E}(\triangle_2 y_{t-2}\eta_t) = 0$, which are nonlinear functions of the coefficients, have to be jointly met. The FAE , on the other hand, can be easily extended to the AR(p) models (Han, Phillips and Sul, 2010b). The extension is possible because of the linear structure of the moment conditions regarding the parameters that the FAE is associated with. Our approach does not usually lead to such moment conditions.

4 Models with Trend

This section applies the SDE to models with a linear trend. The model under consideration is given as follows:

$$y_t = \alpha + \gamma t + u_t,$$

where $u_t = \rho u_{t-1} + \varepsilon_t$, $\rho \in (-1,1]$, and $\varepsilon_t \overset{iid}{\sim} (0,\sigma^2)$ with initial conditions at $t = -2$. The model has the following alternative expression:

$$y_t = (1-\rho)\alpha + \rho\gamma + (1-\rho)\gamma t + \rho y_{t-1} + \varepsilon_t. \tag{8}$$

Following the same approach as in the simple model, we first remove the intercept by second-differencing, and re-arranging the terms gives

$$\triangle_2 y_t - 2\gamma = \rho\,(\triangle_2 y_{t-1} - 2\gamma) + \triangle_2 \varepsilon_t, \tag{9}$$

where γ is understood to be $\mathrm{E}(\triangle y_t)$. With the expression, when $\gamma = 0$ and is known, (9) is identical to (4) for $k = 2$. So the strategy to deliver consistent estimates of the autoregressive coefficient in (4) applies equally well to that in (9). The same argument applies to the cases with a known $\gamma \neq 0$. As in (7), the proposed transformed estimator in this case is defined as

$$\hat{\rho}_{\mathrm{SD}} = \frac{2\sum_{t=1}^{T}(\triangle_2 y_{t-1} - 2\gamma)(\triangle_2 y_t - 2\gamma)}{\sum_{t=1}^{T}(\triangle_2 y_{t-1} - 2\gamma)^2}.$$

Note that the orthogonality condition that $\mathrm{E}[(\triangle_2 y_{t-1} - 2\gamma)\{2(\triangle_2 y_t - 2\gamma) - \rho(\triangle_2 y_{t-1} - 2\gamma)\}] = 0$ continues to hold and the consistency is expected as a result.

When γ is unknown, replacing it with any consistent estimate in $\hat{\rho}_{\mathrm{SD}}$ will work in the same fashion asymptotically. One natural choice for the consistent estimators of γ is the sample mean of the first-difference series, by its very definition. This leads to the "feasible second-difference estimator" (FSDE):

$$\hat{\rho}_{\mathrm{FSD}} = \frac{2\sum_{t=1}^{T}(\triangle_2 y_{t-1} - 2\hat{\gamma})(\triangle_2 y_t - 2\hat{\gamma})}{\sum_{t=1}^{T}(\triangle_2 y_{t-1} - 2\hat{\gamma})^2}, \text{ where } \hat{\gamma} = \frac{\sum_{t=-1}^{T}\triangle y_t}{T+2}.$$

The following limit result closely resembles that associated with the simple models.

Theorem 2. *For all* $\rho \in (-1, 1]$, $\sqrt{T}(\hat{\rho}_{\mathrm{FSD}} - \rho) \Rightarrow \mathrm{N}(0, 2)$ *as* $T \to \infty$.

The limit result is thus invariant to whether a unit root presents itself in the process, and to whether a time trend appears in the model. The de-meaning technique adopted here only plays the role of eliminating the trend effect in the second-difference series. The asymptotic auto-covariance functions for the de-meaned series are left unchanged, and thus the same weak convergence is attained for either the drift model or the trend model.

While adopting the de-meaning practice leads to the invariance property of the FSDE in the limit, the method might introduce a small-sample bias when applying the estimator. The potential bias comes into existence because the moment condition does not exactly hold in finite samples, as a result of using the consistent estimate of the trend parameter ($\mathrm{E}[(\triangle_2 y_{t-1} - 2\hat{\gamma})\{2(\triangle_2 y_t - 2\hat{\gamma}) - \rho(\triangle_2 y_{t-1} - 2\hat{\gamma})\}] \neq 0$).

The size that the bias of the FSDE could be is worth examining by means of simulations. Table 2 reports the finite-sample simulation results for the FSDE, including the magnitudes of the bias for each sample size of concern. The simulation setup follows that for the drift model, and thus the results can be compared with those of the SDE. The results are similar to what we found for the model without a trend, where, for example, normality yields a good approximation for the estimator. The cost of having an additional trend parameter to estimate, however, is the slightly higher RMSE, compared with that for the simple model. The source of the higher RMSE is mainly due to the larger bias produced. While the observation is more significant when the autoregressive coefficients of interest fall into the local area of the unity, the bias is still not noticeable at all.

By contrast, Phillips and Han (2008) use a double-difference method to remove both the intercept and the trend from the models. Their estimator is found to exhibit different Gaussian limiting distributions for models with a trend from those without a trend. Thus, their estimator is not associated with the invariance property. This is certainly not necessarily a disadvantage in applications. The important justification for applying the double differencing to the trend model is to have the resulting moment condition hold exactly, so that the problem of bias can be minimized. [4]

We can also apply the idea of the double-difference method to (9), and derive the corresponding limiting distribution accordingly. The transformed estimator can be shown to have a normality limit as well. [5] Inevitably, in this case, a complicated nonlinear inversion is generally required to obtain the estimated autoregressive coefficient.

5 Concluding Remarks

For typical estimators to have a normality limit is generally not an exception but a rule for the stationary autoregression cases, while the reverse is true for the unit root cases. The transformed second-difference estimator proposed in this paper nevertheless displays quite a different characteristic. It is found to have identical Gaussian asymptotics across the stationary, unity and local-to-unity cases. The normality theory is further characterized by a constant asymptotic variance. The normality limit can even hold when a trend is included in the autoregression. The limit result comes as the result of an introduction to a de-meaning procedure in the process of eliminating the nuisance parameter. The simple technique works by preserving the asymptotic auto-covariance functions of the transformed series.

The properties of the proposed estimators are expected to be more useful in panel contexts. The second-difference approach accompanied by the de-meaning technique that is designed to remove the intercept and trend in time series can deal with the fixed and time effects in dynamic panel models. Furthermore, the estimator is characterized as having little bias in very small samples. This will prove to be an important advantage for panels with short time spans. This is a subject that merits future research.

Notes

[1]Although we just focus on the case with $k = 2$ where the relationship between θ and ρ is linear, the asymptotic normality can be obtained as well for other nonlinear cases by applying the delta method.

[2]The proof of the result simply follows the arguments given for that of the FDE as in Phillips and Han (2008), and is available in an earlier version of the paper.

[3]Han, Phillips and Sul (2010a) show that the full aggregation estimator has the limiting representations characterized by $\sqrt{T}(\hat{\rho}_{\text{FA}} - \rho) \Rightarrow N\left(0, 1-\rho^2\right)$ if $|\rho| < 1$, and $\sqrt{T}(\hat{\rho}_{\text{FA}} - 1) \Rightarrow \frac{\int_0^1 \tilde{B}_r dBr + \int_0^1 B_r^2 dr}{\int_0^1 \tilde{B}_r^2 dr}$ if $\rho = 1$, where B_r is standard Brownian motion and $\tilde{B}_r$ is the corresponding de-meaned process.

[4]To be precise, the double difference procedure involves having the regressor and the error exactly uncorrelated in the transformed regression associated with the derived moment condition. The procedure proves to be very useful in reducing the bias in the context of short panels with incidental trends (Han and Phillips, 2010). We would like to thank one of the referees for pointing this out.

[5]The proof of the claimed result is straightforward by applying Theorem 3.7 of Phillips and Solo (1992) and Theorem 6 of Phillips and Han (2008). The details of the proof and the specific limiting representations for the case have been given in an earlier version of the paper.

References

Anderson, T.W. (1971) *The Statistical Analysis of Time Series*. NewYork: Wiley.

Han, C. and Phillips, P. C. B. (2010) GMM estimation for dynamic panels with fixed effects and strong instruments at unity. *Econometric Theory* **26**, 119−151.

Han, C., Phillips, P. C. B. and Sul, D. (2010a) Uniform asymptotic normality in stationary and unit root autoregression. *Econometric Theory*, forthcoming.

Han, C., Phillips, P. C. B. and Sul, D. (2010b) X-Differencing and dynamic panel model estimation. Cowles Foundation Discussion Papers No. 1747.

Paparoditis, E. and Politis, D. N. (2000) Large sample inference in the general AR(1) model. *Test* **9**, 487-509.

Phillips, P.C.B. and Han, C. (2008) Gaussian inference in AR(1) time series with or without a unit root. *Econometric Theory* **24**, 631−650.

Phillips, P.C.B. and Magdalinos, T. (2009) Unit root and cointegrating limit theory when initialization is in the infinite past. *Econometric Theory* **25**, 1682−1715.

Phillips, P.C.B. and Solo, V. (1992) Asymptotics for linear processes. *The Annals of Statistics* **20**, 971−1001.

White, H. (1984). *Asymptotic Theory for Econometricians*. Orlando, Fla.: Academic Press.

Appendix

A Proof of Theorem 1

First,

$$\sqrt{T}(\hat{\rho}_{\mathrm{SD}} - \rho) = \frac{T^{-1/2}\sum_{t=1}^{T} \triangle_2 y_{t-1}\zeta_t}{T^{-1}\sum_{t=1}^{T} \triangle_2 y_{t-1}^2}. \qquad (10)$$

When $\rho = 1$, $\triangle_2 y_t = \varepsilon_t + \varepsilon_{t-1}$, $\zeta_t = 2\triangle_2 y_t - \triangle_2 y_{t-1} = 2\varepsilon_t + \varepsilon_{t-1} - \varepsilon_{t-2}$,

$$\sqrt{T}(\hat{\rho}_{\mathrm{SD}} - \rho) = \frac{T^{-1/2}\sum_{t=1}^{T}(\varepsilon_{t-1}+\varepsilon_{t-2})(2\varepsilon_t+\varepsilon_{t-1}-\varepsilon_{t-2})}{T^{-1}\sum_{t=1}^{T}(\varepsilon_{t-1}+\varepsilon_{t-2})^2}.$$

Note that

$$T^{-1}\sum_{t=1}^{T}(\varepsilon_{t-1}+\varepsilon_{t-2})^2 \to 2\sigma^2,$$
$$T^{-1/2}\sum_{t=1}^{T}(\varepsilon_{t-1}+\varepsilon_{t-2})(2\varepsilon_t+\varepsilon_{t-1}-\varepsilon_{t-2}) = T^{-1/2}\sum_{t=1}^{T} 2(\varepsilon_{t-1}+\varepsilon_{t-2})\varepsilon_t + \mathrm{O}(T^{-1/2}).$$

Let $v_t = (\varepsilon_{t-1}+\varepsilon_{t-2})\varepsilon_t$ and $\Omega_t = \{\varepsilon_t, \varepsilon_{t-1}, \cdots\}$. Since $\mathrm{E}(v_t|\Omega_{t-1}) = 0$, v_t is a martingale difference sequence with respect to Ω_t. Using Corollary 5.25 of White (1984),

$$T^{-1/2}\sum_{t=1}^{T} v_t \Rightarrow N(0, 2\sigma^4).$$

Collecting all the aforementioned results, we obtain $\sqrt{T}(\hat{\rho}_{\text{SD}} - \rho) \Rightarrow N(0,2)$.

For the cases with $|\rho| < 1$,

$$\triangle_2 y_{t-1} = \rho \triangle_2 y_{t-2} + \triangle_2 \varepsilon_{t-1} = \sum_{j=0}^{\infty} \rho^j \triangle_2 \varepsilon_{t-1-j} = \varepsilon_{t-1} + \rho \varepsilon_{t-2} - (1-\rho^2) \sum_{j=3}^{\infty} \rho^{j-3} \varepsilon_{t-j}.$$

Thus, $\mathrm{E}(\triangle_2 y_{t-1})^2 = 2\sigma^2$. By Theorem 3.7 of Phillips and Solo (1992),

$$T^{-1} \sum_{t=1}^{T} (\triangle_2 y_{t-1})^2 \xrightarrow{a.s.} 2\sigma^2. \tag{11}$$

In addition,

$$\zeta_t = \triangle_2 \varepsilon_t + \triangle_2 y_t = 2\varepsilon_t + \rho \varepsilon_{t-1} - (2-\rho^2)\varepsilon_{t-2} - (1-\rho^2) \sum_{j=3}^{\infty} \rho^{j-2} \varepsilon_{t-j}.$$

Now let $\triangle_2 y_{t-1} = \sum_0^\infty c_j \varepsilon_{t-j}$, $\triangle_2 \varepsilon_t = \sum_0^\infty d_j \varepsilon_{t-j}$, in which

$$\begin{aligned}
&c_0 = 0, \;\; d_0 = 2, \\
&c_1 = 1, \;\; d_1 = \rho, \\
&c_2 = \rho, \;\; d_2 = -(2-\rho^2), \\
&c_k = -(1-\rho^2)\rho^{k-3}, \;\; d_k = -(1-\rho^2)\rho^{k-2}, \;\; k \geq 3.
\end{aligned}$$

Since $\mathrm{E}(\triangle_2 y_{t-1} \zeta_t) = 0$, applying Theorem 6 of Phillips and Han (2008) leads to

$$T^{-1/2} \sum_{t=1}^{T} \triangle_2 y_{t-1} \zeta_t \to N\left(0, \sigma^4 \sum_{k=1}^{\infty} \psi_k^2\right), \;\; \psi_k = \sum_{j=0}^{\infty} (c_j d_{k+j} + c_{k+j} d_j). \tag{12}$$

Further calculations reveal that

$$\psi_1^2 = 4, \psi_2^2 = 4\rho^2, \text{ and } \psi_k^2 = 4(1-\rho^2)^2 \rho^{2(k-3)} \text{ for } k \geq 3,$$

which in turn imply that the value of the asymptotic variance in (12) is $8\sigma^4$. Together with (10), (11) and (12), we have the desired result.

B Proof of Theorem 2

Rewite $\hat{\rho}_{\text{FSD}}$ as

$$\hat{\rho}_{\text{FSD}} = 2 \times \frac{\sum_{t=1}^{T} (\triangle_2 y_{t-1} - 2\gamma + 2\gamma - 2\hat{\gamma})(\triangle_2 y_t - 2\gamma + 2\gamma - 2\hat{\gamma})}{\sum_{t=1}^{T} (\triangle_2 y_{t-1} - 2\gamma + 2\gamma - 2\hat{\gamma})^2}.$$

Note that $\triangle y_t$ is covariance stationary and $\mathrm{E}(\triangle y_t) = \gamma$. The central limit theorem for stationary process (Anderson, 1971) implies that $(\hat{\gamma} - \gamma) = \mathrm{O}_p(T^{-1/2})$. Then it is trival to obtain $\sqrt{T}(\hat{\rho}_{\text{FSD}} - \rho) \overset{d}{\equiv} \sqrt{T}(\hat{\rho}_{\text{SD}} - \rho)$, where $\overset{d}{\equiv}$ means equal in the asymptotic distribution.

Table 1: Simulation Comparisons for Models with Drift

		100×Bias			T×Variance			RMSE		
ρ	T	FDE	SDE	FAE	FDE	SDE	FAE	FDE	SDE	FAE
0.00	40	2.341	-0.085	-0.031	1.993	1.952	0.957	0.222	0.221	0.155
0.00	80	1.227	-0.044	-0.025	1.980	1.978	0.979	0.157	0.157	0.111
0.00	160	0.640	-0.023	0.006	1.997	1.994	0.988	0.112	0.112	0.079
0.00	320	0.345	-0.018	0.011	2.001	1.989	0.992	0.079	0.079	0.056
0.3	40	1.641	-0.774	-1.411	2.527	1.961	0.918	0.249	0.221	0.152
0.3	80	0.859	-0.415	-0.757	2.542	1.981	0.918	0.177	0.157	0.107
0.3	160	0.451	-0.216	-0.365	2.580	1.993	0.913	0.127	0.112	0.076
0.3	320	0.246	-0.116	-0.179	2.590	1.990	0.909	0.090	0.079	0.053
0.50	40	1.117	-1.229	-2.360	2.888	1.967	0.819	0.266	0.222	0.145
0.50	80	0.612	-0.654	-1.251	2.923	1.983	0.792	0.190	0.158	0.100
0.50	160	0.326	-0.335	-0.613	2.969	1.990	0.770	0.136	0.112	0.070
0.50	320	0.180	-0.176	-0.304	2.981	1.992	0.760	0.096	0.079	0.049
0.90	40	0.224	-2.154	-4.425	3.618	1.978	0.467	0.297	0.223	0.118
0.90	80	0.108	-1.144	-2.316	3.699	1.991	0.339	0.214	0.158	0.069
0.90	160	0.086	-0.564	-1.150	3.747	1.990	0.264	0.153	0.112	0.042
0.90	320	0.054	-0.294	-0.568	3.766	1.994	0.227	0.108	0.079	0.027
0.95	40	0.097	-2.282	-4.744	3.711	1.983	0.424	0.301	0.224	0.113
0.95	80	0.040	-1.211	-2.485	3.800	1.996	0.271	0.217	0.158	0.063
0.95	160	0.056	-0.595	-1.239	3.847	1.994	0.182	0.155	0.112	0.036
0.95	320	0.039	-0.308	-0.611	3.866	1.994	0.139	0.110	0.079	0.022
1.00	40	0.095	-2.337	-5.200	3.848	2.010	0.401	0.306	0.225	0.113
1.00	80	0.085	-1.112	-2.767	3.931	2.000	0.220	0.220	0.158	0.059
1.00	160	0.067	-0.541	-1.439	3.934	1.993	0.116	0.156	0.112	0.031
1.00	320	0.049	-0.294	-0.731	3.953	1.995	0.059	0.111	0.079	0.016

Table 2: Simulations for FSDE in Models with Trend

	$\rho = 0$			$\rho = 0.9$		
T	100×Bias	T×Variance	RMSE	100×Bias	T×Variance	RMSE
40	0.149	1.939	0.220	-3.415	2.021	0.227
80	0.017	1.975	0.157	-1.448	2.003	0.159
160	-0.001	1.993	0.112	-0.638	1.993	0.112
320	-0.014	1.989	0.079	-0.313	1.995	0.079

	$\rho = 0.3$			$\rho = 0.95$		
T	100×Bias	T×Variance	RMSE	100×Bias	T×Variance	RMSE
40	-0.719	1.952	0.221	-4.555	2.055	0.231
80	-0.398	1.979	0.157	-1.837	2.023	0.160
160	-0.212	1.992	0.116	-0.749	2.000	0.112
320	-0.114	1.990	0.079	-0.346	1.996	0.079

	$\rho = 0.5$			$\rho = 1$		
T	100×Bias	T×Variance	RMSE	100×Bias	T×Variance	RMSE
40	-1.313	1.962	0.222	-7.489	2.091	0.241
80	-0.672	1.981	0.158	-3.653	2.046	0.164
160	-0.340	1.990	0.112	-1.780	2.017	0.114
320	-0.178	1.992	0.079	-0.920	2.008	0.080

行政院國家科學委員會補助國內專家學者出席國際學術會議報告

100 年 7 月 13 日

報告人姓名	郭炳伸	服務機構及職稱	國立政治大學國貿系教授
會議 時間 地點	July 4 - July 7, 2011 University of Adelaide, Adelaide, SA, Australia	本會核定補助文號	NSC 98-2410-H-004-039-MY2
會議名稱	Econometrics Society Australasian Meeting 2011		
發表論文題目	Gaussian Inference in General AR(1) Models Based on Second Difference		

一、參加會議經過

ESAM 是 Econometric Society 每年夏天在澳洲各城市舉辦的國際會議。今年的主辦大學是 Adelaide University 的經濟與管理學院。Adelaide 是南澳的中型城市，學術環境優良，也是澳洲著名的大學之一。今年的主辦單位大力邀集中生代計量學家，成為這次大會的一大特色。這次中生代計量學家大約畢業於 90 年代前後，大部份皆由 Yale 與 UCSD 獲取學位。經過 10 幾年的粹煉，這些訓練有素，歷經美國嚴謹升等制度並取得正教授的學者，事實上正主宰並影響未來經濟學術計量領域的走勢與思潮。在這 4 天的會議中，由於邀請學者眾多，其場次皆與投稿場次並列，常造成與會者不知參與哪一場次的抉擇問題。除此之外，整個大會，無論邀請場次或投稿場次的發表論文，均有很高的水準，是非常值得參加且收穫豐盛的國際會議。

二、與會心得

這次在大會所發表的文章，之前也曾在北京清華大學報告，也得到參會者一定的迴響。這篇文章已獲得 Econometric Theory 邀請修改，希望不久的未來會有好的結果，也可為參加這些國際會議劃下完美的句點。

在這次會議中，在聆聽不同場次的論文發表後，似乎可以歸納出一些未來的研究走向。除此之外，我也在這次會議發表論文中，找到新的研究議題，以及過去既有研究的重新定位。以下將逐點說明這些觀察與想法。

(一) 緩長記憶 (long memory) 的估計具有嚴重的偏誤現象。在文獻中，已有多篇重量級文章分析不同估計式下的小樣本偏誤解析式。這些解述式固然幫助學者瞭解偏誤的來源，但無助於偏誤的修正。這次會議有學者發表以 bootstrap 方式修正不同估計式偏誤之想法。但恐因近似模型不優良，修正結果並不理想，而必須輔以其他估計式，方得以進一步改善。不過這可能是 indirect inference 可以有所發揮的時候。初步的構想是給定任一緩長記憶參數，indirect inference 可以充分模擬該既定參數值下的數列，並將之與既有觀察數列比較。

如果兩數列差異太大，則進一步改變原給定參數值，直到模擬數列與觀察數列相似為止。這是非常值得嘗試的研究課題。

(二) 如何近似 Least squares 估計式的動差也是值得關注的課題。在我個人的研究中，2SLS 或 IV 估計在小樣本具有嚴重偏誤，亟待改善。若以 indirect inference 修正，中間涉有若干動差必須加以估計或模擬，因此瞭解 LS 的動差近似或可幫助釐清前述問題。在這篇論文中，發表人建議以 Taylor expansion 方式近似 LS 估計式的諸多動差。這樣的近似得以瞭解動差結構，但如何應用於 indirect inference 的操作中，值得再討論。

(三) Averaging 應是將來計量的發展方向。這次大會主講人 Prof. Bruce Hansen 從降低估計風險的角度形成這樣估計步驟。降低風險似乎應是每個實證研究者在估計時致力的工作，這也是為什麼在各期刊所發表的實證文章中，不同的迴歸結果必須加以呈現，以利比較外，並確保實證結果的穩定。Averaging 其實是充分利用數據所可能透露的訊息，以確保風險的最低。Prof. Hansen 在此次會議報告其研究最近的理論成果，令人興奮。其證明 Averaging 有可能達到最有效率的估計。這些新的理論結果，也可能將對實證研究形成一定影響。目前我的一篇研究結合個別數列與橫斷面數列之訊息，對總體基本面變數是否具有預測力進行探討。該文的實證結果很明顯為 Prof. Hansen 的理論在實證上呼應。在這個時點，這篇實證研究似乎更值得在國際頂級期刊一試它的價值。

三、考察參觀活動(無是項活動者省略)

無。

四、建議

無。

五、攜回資料名稱及內容

本會議議程。

六、其他

無。

無研發成果推廣資料

98 年度專題研究計畫研究成果彙整表

計畫主持人：郭炳伸	計畫編號：98-2410-H-004-039-MY2
計畫名稱：自我迴歸下偏誤的降低：簡單組合估計式之建構	

成果項目			量化 實際已達成數（被接受或已發表）	量化 預期總達成數(含實際已達成數)	量化 本計畫實際貢獻百分比	單位	備註（質化說明：如數個計畫共同成果、成果列為該期刊之封面故事...等）
國內	論文著作	期刊論文	1	1	100%	篇	無。
		研究報告/技術報告	1	0	100%		無。
		研討會論文	0	0	100%		無。
		專書	0	0	100%		無。
	專利	申請中件數	0	0	100%	件	無。
		已獲得件數	0	0	100%		無。
	技術移轉	件數	0	0	100%	件	無。
		權利金	0	0	100%	千元	無。
	參與計畫人力（本國籍）	碩士生	0	0	100%	人次	無。
		博士生	0	0	100%		無。
		博士後研究員	0	0	100%		無。
		專任助理	0	0	100%		無。
國外	論文著作	期刊論文	0	0	100%	篇	無。
		研究報告/技術報告	2	1	100%		無。
		研討會論文	0	0	100%		無。
		專書	0	0	100%	章/本	無。
	專利	申請中件數	0	0	100%	件	無。
		已獲得件數	0	0	100%		無。
	技術移轉	件數	0	0	100%	件	無。
		權利金	0	0	100%	千元	無。
	參與計畫人力（外國籍）	碩士生	0	0	100%	人次	無。
		博士生	1	1	100%		莊珮玲。
		博士後研究員	1	1	100%		陳致綱。
		專任助理	0	0	100%		無。

其他成果 (無法以量化表達之成果如辦理學術活動、獲得獎項、重要國際合作、研究成果國際影響力及其他協助產業技術發展之具體效益事項等，請以文字敘述填列。)	無。

	成果項目	量化	名稱或內容性質簡述
科教處計畫加填項目	測驗工具(含質性與量性)	0	
	課程/模組	0	
	電腦及網路系統或工具	0	
	教材	0	
	舉辦之活動/競賽	0	
	研討會/工作坊	0	
	電子報、網站	0	
	計畫成果推廣之參與（閱聽）人數	0	

國科會補助專題研究計畫成果報告自評表

請就研究內容與原計畫相符程度、達成預期目標情況、研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）、是否適合在學術期刊發表或申請專利、主要發現或其他有關價值等，作一綜合評估。

1. 請就研究內容與原計畫相符程度、達成預期目標情況作一綜合評估

□達成目標

■未達成目標（請說明，以 100 字為限）

□實驗失敗

□因故實驗中斷

■其他原因

說明：

在研究過程中，我們發現自我迴歸模型的性質，遠比我們想像還要複雜。因此為了更徹底瞭解該模型性質，我們大部分研究皆以二階差分自我迴歸模型之性質為研究主軸。

2. 研究成果在學術期刊發表或申請專利等情形：

論文：■已發表 □未發表之文稿 □撰寫中 □無

專利：□已獲得 □申請中 ■無

技轉：□已技轉 □洽談中 ■無

其他：（以 100 字為限）

無。

3. 請依學術成就、技術創新、社會影響等方面，評估研究成果之學術或應用價值（簡要敘述成果所代表之意義、價值、影響或進一步發展之可能性）（以 500 字為限）

即使我們原來的研究設定目標未能如預期達成，但是我們對二階差分模型自我迴歸的統計性質卻有了非常好的理解，部分研究成果已獲期刊接受刊登，部分則仍等待回音。