Comments on "A Critical Investigation on Detrending Procedures for Non-Linear Processes"

Richard Ashley*
Virginia Tech

Randal J. Verbrugge
U.S. Bureau of Labor Statistics

October 18, 2005

abstract

We agree that either mistaking a stochastic trend for a deterministic trend (or vice-versa) is consequential for unit root tests and for tests of nonlinear serial dependence. In addition, we comment that similar results obtain for ordinary parameter inference in simple linear models. In particular, we note that detrending stochastically trended data with a deterministic polynomial or by applying the Hodrick-Prescott filter yields notably mis-sized hypothesis tests, even with a sample length of 200 observations. Interestingly, we further find that these size distortions persist even with stationary – albeit highly persistent – data.

JEL Classification: C12, C22, C32
Key words: nonlinearity, detrending, unit roots, spurious regression, Hodrick-Prescott filtering

* Economics Department (0316), Virginia Tech, Blacksburg, VA 24061. Telephone: (540) 231 6220. Fax: (540)231 5097. E-mail: ashleyr@vt.edu.

This paper is a valuable contribution to the literature on nonlinearity. The authors convincingly make an important point: the detrending choice is consequential for subsequent inference. They prove that inappropriate detrending distorts the population autocorrelation function. And they use simulations to show that inference results with regard to unit roots and non-linear serial dependence will be distorted if one either erroneously applies polynomial detrending to a series with a stochastic trend ('under-differencing') **or** if one erroneously differences a series which is nonstationary due to a deterministic polynomial trend ('over-differencing'). Their results focus mainly on the power of such tests but (absent bootstrapping) one will see, as they point out, similarly substantial distortions in test sizes.[1] This paper is particularly useful because, aside from Kapetanios and Shin (2003), the literature has paid scant attention to these issues in the context of nonlinear modeling.

It is worth noting that this kind of result on detrending is not limited to unit root and nonlinearity tests: we have obtained similar results (Ashley and Verbrugge (2004)) in the context of ordinary parameter inference in simple linear models. For example, we simulated 10,000 samples of length 200 from the process:

$$\begin{aligned} x_t &= 1.0 + \phi x_{t-1} + .70\{x_{t-1} - \phi x_{t-2}\} + \epsilon_t \\ y_t &= 1.0 + \phi y_{t-1} + .70\{y_{t-1} - \phi y_{t-2}\} + \eta_t \end{aligned} \tag{1}$$

in which ϵ_t and η_t are both NIID[0, 1] and ϕ equals {1.00, .99, .95, .50, .10.}. Clearly, x_t and y_t are independently distributed, but they are dominated by stochastic trends if ϕ is one. We then estimated several linear models for the dependence of y_t on past values of x_t in these data and computed the size of simple hypothesis tests of the null hypothesis that no such dependence exists. The models considered were straightforward regression equations: in the levels of x_t and y_t, in the levels of x_t and y_t but including a linear time trend term, and in the first differences (Δx_t and Δy_t) of the two series. Since the HP filter is often used to detrend nonstationary data in the macroeconomics literature, a model in the levels of HP filtered data was also considered.

In addition we estimated a "lag-augmented VAR" regression equation relating y_t to past values of x_t. The lag-augmented regression model – suggested in a more general (vector autoregression) context by Toda and Yamamoto (1995) – perhaps requires further comment. It is simply:

$$y_t = \alpha_o + \alpha_1 \Delta y_{t-1} + \alpha_2 \Delta x_{t-1} + \alpha_3 y_{t-2} + \alpha_4 x_{t-2} + \eta_t \qquad (2)$$

This formulation requires estimation of one more parameter (α_4) on the additional ("augmented") lag in x_t. Note, however, that it nests the model in levels, in which y_t depends on lagged values of x_t and y_t. Its advantage for inference is that $\alpha_2 = 0$ is a necessary condition in order for y_t to be independent of past values of x_t. But α_2 by construction multiplies the differenced series $\left(\Delta x_{t-1}\right)$, which will be I(0) even if x_t is itself nonstationary. Consequently, the least squares estimator of α_2 will have a standard asymptotic distribution.

Table 1 below gives the empirical sizes of hypothesis tests that y_t does not depend on past values of x_t based on these simulations. It is commonplace to observe that such tests are mis-sized in levels models for independent I(1) processes – i.e., for $\phi = 1$: this is the familiar "spurious regressions" phenomenon. And Hodrick-Prescott (HP) filtering is clearly not a satisfactory approach to detrending. Note, however, that including the linear trend in the model instead of first differencing ("under-differencing") yields a test which is still seriously over-sized.

Moreover, note that (for these simulations at N = 200) these tests are still seriously over-sized even for data which are fairly persistent (ϕ = .99 or .95) but actually stationary!

Evidently, de-trending stochastically trended data with a deterministic polynomial (under-differencing) is consequential for inference even in simple linear contexts. What about over-differencing? Other simulations in Ashley and Verbrugge (2004) show that over-differencing – i.e., modeling in differences when the data are already stationary except for a linear trend – yields very distorted estimated impulse response functions. Disturbingly, the distortions become increasingly severe as the sample length increases.[2] We conjecture that the same holds true for impulse response functions from nonlinear models.

Table 1. Empirical Size of 5% Test That y_t Is Independent of Past x_t					
	Generating Processes				
Estimated Model	Stochastic Trend	Independent Stationary Processes			
	$\phi = 1.00$	$\phi = .99$	$\phi = .95$	$\phi = .50$	$\phi = .10$
Levels model	.246	.163	.084	.058	.056
HP-filtered levels	.092	.098	.098	.085	.077
Levels with linear trend	.153	.171	.106	.060	.058
Differences	.047	.060	.058	.060	.057
Lag-augmented levels	.053	.058	.056	.048	.049

Footnotes

1. Other papers which study the distorting effects of commonly-used detrending procedures are Phillips (2002), Harvey and Jaeger (1993), and King and Rebelo (1993). Verbrugge (1997) discusses the impact of detrending on the type or form of asymmetry which is detectable; see also Psaradakis and Sola (2002).

2. We note in passing that the lag-augmented estimated model does well in both settings.

References

Ashley, R. and R. J. Verbrugge (2004) " To Difference or Not to Difference: A Monte Carlo Investigation of Inference in Vector Autoregressive Models," VPI Economics Department Working Paper E99-15, http://ashleymac.econ.vt.edu/working_papers/9915.pdf.

Harvey, A.C., and A. Jaeger (1993) "Detrending, stylized facts and the business cycle," *Journal of Applied Econometrics 8*, 231-247.

Kapetanious, G. and Y. Shin (2003) "GLS Detrending-based Unit Root Tests in Nonlinear STAR and SETAR Frameworks," mimeo, University of Edinburgh.

King, R.G., and S. Rebelo (1993) "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control 17*, 207-231.

Phillips, P. C.B. (2002) "New unit root asymptotics in the presence of deterministic trends," *Journal of Econometrics 111,* 323-53.

Psaradakis, Z. and M. Sola (2002) "On Detrending and Cyclical Asymmetry," *Journal of Applied Econometrics 18*, 271-289..

Toda, H. Y. and T. Yamamoto (1995) "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes," *Journal of Econometrics 66*, 225-250.

Verbrugge, R. (1997) "Investigating Cyclical Asymmetries," *Studies in Nonlinear Dynamics and Econometrics 2*, 15-22