

Report on MS 1146 *Testing for near $I(2)$ trends when the signal to noise ratio is small* by Katarina Juselius.

General comments

The paper analyses why the univariate Dickey-Fuller test may wrongly reject the null of a unit root (in the differenced trending series) more often than the multivariate (trace based) tests.

The analysis and the results are based both on analytical results and on Monte Carlo analysis. The paper is well organized and it is well written with clearly presented results. The topic of the paper is of interest since the degree of integration is an important modelling concept in dynamic econometric models that often include the real exchange rate as an endogenous variable.

The economic theory used to motivate that the real exact rate is likely to contain two (long-run?) units-root is rather special. Other theoretical approaches will lead to other null hypotheses to test; and structural breaks can intervene into the story as well. This is not a critique of the analysis, but it suggests that a clearer delineation of the scope of the paper can be made.

Some detailed comments

The following sentences on page 3 struck me as important, but ambiguous.

When testing the order of integration of the simulated series the results show that the univariate DF test tends to reject the second (near) unit root in almost all cases, whereas the multivariate test almost always finds it. The former result can be explained by the low power of the univariate DF tests to detect a second unit root

The received wisdom, I thought, was that the DF test has low power (high probability of type-II error), and the second sentence is therefore confusing. The problem disclosed by the results in the paper is then related to the *level* of the DF test, relative to the multivariate test. The problem is that the DF test wrongly rejects the null of a unit root (in the differences) more often than the formal significance level suggests. The multivariate test on the other hand has about the right level.

I must admit that my “prior” was that multivariate tests have better power in rejecting unit-roots that are not in the DGP (while maintaining the correct level) than the DF family of tests. Does this only apply to cointegration, and not to univariate tests? Maybe some clarification can be considered?

Last paragraph on page 2. The conclusion that both the real and the nominal exchange rates are $I(2)$ deserves more comment, since it is unexpected to readers with a macroeconomic background. If both s_t and $(p_t - p_t^*) \equiv pp_t$ are $I(2)$,

and also $rex_t = pp_t - s_t$ is $I(2)$, then pp_t and s_t cannot be cointegrated. This conclusion seems to contradict the statement earlier in the Introduction: “nominal exchange rate has shown a tendency to move in long persistent swings around its long-run purchasing power parity (PPP)”, since without cointegration there is no well defined equilibrium. Alternatively, $rex \sim I(2)$ would be logically consistent with $s_t \sim I(2)$ and $pp_t \sim I(1)$, and lack of cointegration would be “due to” different degree of integration (which does not say much without further economic interpretation).

Page 4, l 3 from bottom. Under the H_0 a “long-run value”, which here must mean unconditional expectation, does not exist.

Page 5: last paragraph of section 2.1. If the second unit root is hard to detect, how damaging is it if one gets the conclusion wrong? For example when one investigate empirically the relationship between the real exchange rate and other macro series, such as the profit-share or the rate of unemployment? A short section, either here or in the conclusion for example, would be of interest I think.

Figure 1, caption. I suggest using the same notation (symbols) as in the main text (i.e. s_{12})

page 7, line 2 from bottom: I suggest write “moving-average” instead of “MA” here (first instance)

Section 2.3. The discussion shows that the problem is that the DF test rejects too frequently (relative to the chosen significance level) when the null of a unit-root is true. as already noted, I found it confusing, initially, that this was associated with low power of the DF test. More constructively: It is interesting that the nature of the “problem of the DF” changes from low test power to wrong test level when the a null hypothesis of a unit root is formulated for the differenced series. Is there an intuitive explanation?

Section 2.4, first paragraph. My interpretation is that the level of the multivariate test is better (more correct than the DF test), so that the probability of wrongly rejecting the second unit root is e.g., 5 %, when a 5 % significance level is used.