Reply to the referee's report

Summary

We do not agree that *time series* are called linear or nonlinear. *Models*, not time series, can be linear or nonlinear.

Main Comments/Suggestions

Although we also possess *feelings* and *impressions* we refrain from disclosing them in this reply.

Since no reasons and specific comments and criticisms are provided, we cannot understand the reason why *"the contribution of the paper (...) is very limited."*. Furthermore, while it may be true that *"the existing literature on nonlinearity in business cycles is extensive"*, there is no relation between this and the second part of the sentence, stating that *"the authors should spend more effort to convince the reader..."*. Since the validity of the first part of the sentence has nothing to do and therefore cannot support the statement of its second part and since no justification and/or detail is provided about this one, it is devoid of any basis.

It is true that we describe some tests. But these descriptions are very far from corresponding to *"most parts of the paper"*. Indeed, first we are very selective both about the tests that we choose to use and about the ones that we describe. Second, the descriptions that we provide are rather brief, not *"lengthy"*, and third, in most cases they refer to tests that are neither *"well-known"* nor *"widely used"*. This is the case, for instance, of the CDR test, that was proposed many years ago but appears to remain rather unknown and has been rarely used; this is also the case of Tsay's (1989) test. The only test that is well known but, as far as we know, has been rarely applied to business cycles data is the LM-STAR test. Furthermore, we have relegated the description of some (selected) unit root tests against nonlinear alternatives to a separate appendix.

The presentation of the empirical results, in sections 3 and 4, is really somewhat lengthy and detailed. As a matter of fact, the empirical analysis must be thorough and carefully done to support our claim: a battery of (selected) tests is applied to a dataset covering many countries, precisely the opposite of a test result for a single country. In other words, we are really making a large *"effort to convince the reader"*.

Unfortunately, the current version of the abstract was not completely written by us. Surprisingly, our version was changed by the editor who, for instance, replaced the usual *"we"* with *"the authors"*. We cannot understand these changes because

we could not find any instruction forbidding the use of "*we*" and we tend to agree with the referee: our original version was better than this one.

Minor Comments/Suggestions

- Indeed the sample could *now* be somewhat extended. However, according to our reading of the literature and our own simulation experience, extending the sample with 4 or even 8 more observations produces almost no effects on the power of the tests. To obtain a significant — but still small — increase in power one would need at least 20 more observations, i.e., about 5 more years of data.

- The suggestion to use (linear and non-linear) panel unit root tests is rather strange. First, in every case the sample size (T) is large enough to permit reliable country-specific inference, i.e., there is no justification to pool countries into a panel. Second and more importantly, the panel approach lacks the flexibility and insightfulness on individual countries of the (individual) time series approach. Performing panel unit root tests is precisely the opposite that we need to do. To assess whether linear models are dismissible we need to analyse the larger and the most diverse set of countries. Only this way we can build a really solid argument and try to "*convince the reader*" of its goodness. For instance, the panel approach would preclude obtaining the evidence that is summarized in the last paragraph of the paper concerning Canada, Germany and the U.S. .

- A serious discussion on business cycle indicators would be a very long and probably "*exhausting*" task, much beyond the purposes of this paper, and this suggestion clearly contradicts the previous one because it would increase "*the number of results*". We simply follow the bulk of the literature, while simultaneously adopting the classical approach to business cycles. Moreover, the availability of data on "*hours*" or "*employment*" is much more limited than for GDP. Further still, since "*production*" is mentioned without the restriction to the industrial sector, let us recall that GDP means gross domestic *product* (and that it is not available at the monthly frequency).

  Anyway, in section 2, our paper already contains a discussion justifying the use of the data that is the focus of our attention. A further discussion on its limitations is presented in a paragraph of the final section.

- We believe that figure 1 is useful, at least because it illustrates the variety of shapes for output growth around the world. It is also useful to support

the references to heteroskedasticity that we make later (for instance, in the discussion of the final section).

The scaling is not different for the euro area because in every case $\Delta y_t = \Delta \log(GDP_t)$. The problem here is one of a graphical typo and we are grateful to the referee for pointing it to us. Indeed, the problem refers to the values of the vertical axis where some digits are missing: "20" should be 0.020, "10" should appear as 0.010, etc. This problem sometimes occurs when transferring graphics to the word processor and we will correct it in the revised version.

Finally, the answer is yes, it is, according to the information provided by Datastream: the GDP for India is seasonally adjusted.