Testing for Breaks in Cointegrated Panels – with an Application to the Feldstein-Horioka Puzzle

By Francesca Di Iorio and Stefano Fachin

Referee report

This paper offers panel version of the tests for structural break originally developed by Hansen (1992). To cope with the problem of cross-section dependence, the new tests are based on bootstrapping.

Overall impression

The paper deals with an interesting issue – at least in the view of this referee – and the problem of testing for structural change in cointegrated panels has not been addressed before. I also found the paper well-written.

The main critique I have is that the paper is not self-contained in the sense that based on the results provided it is actually not possible to compute the proposed tests. In fact, one still needs to look up the Hansen (1992) and Politis and Romano (1994) papers for all the calculations involved as well as for most of the theoretical results.

My second main critique concerns the way in which the estimated breaks are treated. In particular, what does it mean to take the median break as the breakpoint estimator? If it means imposing a common breakpoint in the middle of the sample, then there is really not much value in the proposed tests, at least not from a practical point of view.

In what follows I will first comment on those aspects I believe to be the most important weaknesses of the paper, and then I also have some minor comments. The comments are ordered in the way things appear in the paper.

Major comments

1) Section 2.1: This relates to my first overall comment, namely the level of the details provided, which is too low. For example, why is it that, even though the test takes no breaks as the null hypothesis, one still has to estimate breaks in step 2 of the bootstrap algorithm?

Similarly, in step 5 you propose using the stationary bootstrap as a means to accommodate cross-section dependence. But this bootstrap is never explained in detail and so the interested reader is forced to look up that paper as well.

2) Page 4, step 3 of the bootstrap algorithm: Here you explain how the individual test statistics can be combined into a summary statistic for the whole panel. In particular, the sum and the median of the individual statistics are proposed. Are there any particular reasons for using these? Many times one uses the sum because (appropriately normalized) it leads to an asymptotic distribution that is normal. However, you use the bootstrap so you are basically free to pick whatever summary you want, and then the minimum and maximum stand out as the most natural candidates. In particular, since the individual tests are left-tailed, the maximum would be suitable when testing the no break null against the alternative that there is at least one break, while the minimum would be suitable when the alternative is formulated as that all the units of the panel are breaking.

3) Page 4, remark (i): Again, is it really necessary to estimate the breaks? I can see that this might matter when doing asymptotic analysis but this is not how things are done here. In particular, it seems to me that the use of the bootstrap should enable you to dispense with the break estimation all together? Otherwise, if you still need to estimate the breaks, then your suggestion of taking the median seems to take away much of the appeal of the new tests. In fact, although there is no break under the null, by picking the median you are basically assuming that there is a single homogenous break located in the middle of the sample (at least to the extent that the individual breaks can be assumed to be evenly distributed across the length of the sample). But then you can just as well put it there from the outset and forget about the estimation. I really thing that you need to give this issue of the break estimation some more thought.

4) Section 3.1: The design of the simulations is overly complicated, and should be redone so that readers can more easily see what's going on. For example, since a is set to zero anyways, equation (1) can be deleted. Also, are you really generating x_{it} as in (1)? I'm wondering because if this is indeed the case, then x_{it} isn't breaking at all as claimed in the discussion surrounding (3) as a is set to zero. Moreover, the first lines of (5) and (6) contain the parameter θ, which I presume is there in order to allow for a nonzero mean and trend in x_{it}. Why don't you just include this as a constant and trend in the equation for x_{it} instead?

It will make the whole thing much easier to understand. Finally, the second lines of (5) and (6) are the same.

5) Page 9, case 5, line 11: Here you state that the null should be formulated as that there is a large number of units for which there are no breaks. How is the test of this null implemented? In particular, if the null is indeed a fixed point in the parameter space – as it usually is – then you need to prespecify the breaks of the units that are supposedly breaking, which seems like a very odd thing to do. Furthermore, this formulation does not in any way follow from how it is done in Pedroni (2004) where the null is always taken to be that the panel isn't cointegrated. The formulation of the alternative varies a little bit depending on the particular test used, but the null is always formulated in this way.

6) Page 9, paragraph 2 from the bottom, the last 2 lines: What is actually meant by this? In particular, do you really mean that you have made different cross-sectional drawing for each test? If this is the case then you can't really say whether the observed differences in performance are due to true performance differentials or just random variation (especially since you make only 500 replications), which of course makes the whole comparison rather irrelevant. The proper way to do this if you still want to compare the different tests is to either make a joint program or to use the same seed for generating the random numbers.

7) Table A1: There seem to be some anomalies in the results provided. For example, in Table A1 the size distortions are bigger when N is relatively small, which goes against the usual wisdom that the distortions should tend to accumulate with N. Even in absolute terms the distortions are quite substantial, especially when N is small. Why is that? The whole purpose with the bootstrap is to get the null distribution right but this does not appear to be the case here. It would be good with at least some explanation of these results.

8) Section 4, paragraph 2: In contrast to what is claimed in the text, the figures don't show much evidence in favor of breaks. In fact, my guess is that if one were to subtract one of the series (savings or investments) from the other, then that difference would be well-described by a stationary series with no breaks. Also, you say that the fact that while the individual tests generally don't reject the null of no cointegration the panel tests do can be attributed to low power. I disagree. For the panel tests to reject, it is sufficient that there is evidence of cointegration for a single unit, which just what you find. Thus, in this case, it actually

appears as that the panel and individual tests lead to exactly the same conclusion, and hence there is no evidence of low power in either of the tests.

Minor comments

1) Page 1, line 4: What "tests" are you referring to here?

2) Page 1, lines 18 to 21: Please reformulate this sentence.

3) Page 1, last 2 lines of footnote 1: Although simulation results are only reported for N as large as 40, this does not mean that it has to be that large. In fact, $N = 10$ is usually enough for this type of factor based tests to work properly.

4) Page 3, line 6: There is only one "1980's" – I think you mean the latter part of the 1980's?

5) Page 3, section 2.1, paragraph 1: Why not write $\mathbf{Z}_{it} = (Y_{it},\, X_{it}')'$ where $\mathbf{X}_{it}$ is $k \times 1$? Also, make sure you use a comma to separate the vector elements, which will otherwise look like product elements.

6) Page 4, remark (i), line 2: Here you define the breakpoint estimate, $\widehat{t}_i^b$. How is this estimate actually constructed? Are you taking the supremum over i or is it "$(Sup\widehat{F})_i$" and then the maximum over i, or?

7) Page 6, line 2 beneath (2): Don't you mean strictly exogenous? u_{it}^x and u_{it}^y are uncorrelated, right?

8) Page 8, case 2: If you want to see the effect of increasing T you should increase it while keeping N fixed – not while reducing N.

9) Page 10, section 3.2, paragraph 2, line 1: What is meant by the "speed of adjustment of the DGP's"? Also, in the last sentence of that paragraph I think you mean that a panel approach seems more preferable even when N is small.

10) Page 10, last paragraph: Please reformulate the last complete sentence.

11) Page 11, paragraph 2, last line: The loss in power is not likely to be due to the imprecise coefficient estimates but rather to the fact that you are estimating more parameters.

12) Page 21, paragraph 2, lines 15 to 18: The p-values you get are much more likely to be due to size distortions than to low power.