Fredj Jawadi

Guest Editor

Computational Economics

1st February, 2017

Dear Dr. Jawadi,

Please find attached a revised version of our manuscript CSEM-D-16-00137 entitled “Fast and adaptive cointegration based model

for forecasting high frequency financial time series”.

We are very grateful for the reviewer’s comments that have enabled us to improve the quality of our manuscript. We have addressed each comment point-by-point, highlighting the changes we made in each case.

We hope that we have responded appropriately to each comment in order to make our manuscript suitable for publication in Computational Economics.

**Reviewer #1 comment:** The paper is very well written.

**Response:** Thank you for your comment.

**Reviewer #1 comment:** The paper has a real contribution from a technical point of view. However, I’m not convinced by whether determining long run equilibrium in high frequency data make sense. In other words, when we have 5 minutes data, what is the long term and what is the short term? This is my principle subtlety for this analysis.

**Response:** In high frequency data, the long run equilibrium found by cointegration has its foundation in statistical arbitrage theory. In the case of using a window of 5 minutes, the VECM will determine the long-term relationship between variables which could change if the window size changes. We have added a paragraph in the Introduction section (page 2 paragraph 2) to address this comment and the one below.

**Reviewer #1 comment:** In line to my previous comments, authors did not motivate any where in the paper, the importance of long-run equilibrium and short term interaction for economic and/or financial point of view, under high frequency data. In low frequency data (monthly, quarterly) the long run has an economic interpretation (for example economic cycle, Phillips curve relationship, and others economic background). Authors should highlight whether this information coming from long-run equilibrium for the case of high frequency data are useful from a theoretical point of view. In other words, whether this information are useful for policymakers, investors, hedgers and economic agents?

**Response:** We are grateful for this comment and agree that cointegration in high frequency data was not highlighted. We have added that cointegration is mainly motivated by statistical arbitrage theory. VECM can help to detect long-term relationship between high frequency assets and this is relevant information for high frequency trading. We have also added three new references that show that cointegration could appear with increasing data frequency.

1. Miao GJ (2014) High frequency and dynamic pairs trading based on statistical arbitrage using a two-stage correlation and cointegration approach. International Journal of Economics and Finance 6(3)
2. Zhou S (2001) The power of cointegration tests versus data frequency and time spans. Southern Economic Journal 67(4):906-921.
3. Rittler D (2012) Price discovery and volatility spillovers in the European union emissions trading scheme: A high-frequency analysis. Journal of Banking & Finance 36(3):774-785.

**Reviewer #1 comment:** Authors say “In order to show that the number of cointegration vectors depends on the amount L of historical data and the number of lags p in the VECM, we used a grid search. We defined a grid of possible values for L and p. L goes throughout [2, 14] hours (1 hour = 360 data points) with a step size of 4 hours and p throughout [1, 5] with step size of 1.” My comment is as follows: what are the motivations behind the choice of these different grids. For example, for the choice of L goes through [2, 14] hours, what are the raisons, as in the financial markets trading is made through cession of 6 hours or 24hours.

**Response:** The grid was defined arbitrarily. Our objective was only to show the variability of the number of cointegration vectors when we changed parameters L and p. We have added this explanation in the Methodology section (page 6 paragraph 1).

**Reviewer #1 comment:** How you can interpret the figure 1. For example, the probability of having one cointegration relationship is equal to one for 10 hours.

**Response:** Figure 1 shows the distribution of the number of cointegration vectors for different values of L in 1000 iterations. When L=2 hours, there was no cointegration in more than 60% of the iterations. When L was 10 hours, the probability of having one cointegration vector was 100%. A more detailed explanation can be found in the Methodology section (page 6 paragraph 2).

**Reviewer #1 comment:** After presenting the PC ratio formula, authors run an experiment with L go through [700, 1500] data points. For 700 data points is approximately corresponds to 2 hours. However, why the choice of 1500 data points and not 5000 which corresponds to 14 hours?

**Response:** We have extended the axis to 5040, corresponding to 14 hours. Figure 2 was updated (page 8).

**Reviewer #1 comment**: For figure 2, authors interpret the first one “We found that better cointegration percentage leads to better performance accuracy in terms of reducing MSE ». What about the second figure with L. indeed, we observe that MSE is lowest with L equal to 1050. How you explain this result?

**Response:** The second figure shows how MSE varies with respect to L. We found that increasing the size of L doesn’t necessarily help to reduce MSE. We have added the explanation of this comment (page 7 paragraph 2).

**Reviewer #1 comment:** The empirical application is well done.

**Response:** We really appreciate this comment.

**Reviewer #1 comment:** The authors used four currencies markets in order to analysis their relationship based on the proposed approach AVECM. However, the economic analysis still very scarce or absent.

**Response:** We chose the most traded rates related to USD so they were likely to be cointegrated. This analysis could be done with two or more assets. We chose more than two in order to show that it is possible to find cointegration with more assets in short periods of time. The advantage of using AVECM is we don’t need to restrict the amount of data or the number of assets to obtain a response in a short time.

We are really grateful for all the highly insightful comments that enabled us to improve the quality of our manuscript. We hope we have responded appropriately to all of your comments.

**Reviewer #2 comment:** Congratulations. I think this article is relevant, clearly written and uses a correct methodology. I have recommended its acception to the editor.

**Response:** Thank you for your comment and your recommendation for acceptance.