

Referee report

Manuscript title: Twitter's daily happiness sentiment and the predictability of stock returns

Comments:

- An appropriate base of theories is missing. There are no concepts presented and the motivation by other ideas is limited to the introduction. The paper is imbalanced between presenting interesting statistical results and the theoretical/practical side. The authors should clearly explain why Twitter's daily "happiness index" can be regarded as "investor sentiment". It is not clear to me what drives reported impact of Twitter's daily happiness index on stock return. It is important to discuss it. I understand that the authors follow previous articles in using Twitter's daily happiness index as a proxy for investor sentiment, but the rationale should be much stronger than that.
- One of the weaknesses of the paper is its methodology. The selection of asset pricing model matters. In the main results, the authors use the 3-factor Fama-French and Carhart momentum factor. I am surprised that authors do not report the main results based on the 5-factor Fama and French model, but shortly mention it in the robustness test. Maybe, there is a value in the selected model. However, a reader can make his mind after comparison with state of art asset pricing model -5 factor model. I suggest the authors put the results of the 5-factor Fama and French model, not the 3-factor Fama-French and Carhart momentum factor, in the main empirical result section. The authors should report appropriate diagnostic tests (i.e., autocorrelation, heterogeneity). Although Durbin-Watson (DW) is mentioned in the article, the authors should realize the limitation of it. DW can be test autocorrelation at 1 lag only. As the daily data is used in the study, the authors should consider other methods of autocorrelation tests based on more than 1 lag. Do not the authors think that the reported adjusted R-squared (0.99, 0.95) is too high for the selected model? I suggest the authors compare the adjusted R-squared reported in the manuscript with that of comparable published articles. The error term should be added in equations (1) – (3). Skewness and kurtosis should be reported in the descriptive statistics. The sample period (2008-2021) covers different market conditions (i.e., financial crisis, covid-19 pandemic). The analysis in subperiods might be interesting to see the different impacts of investor sentiment on stock return.
- I acknowledge the work-effort the authors put into the empirical result section, but the authors use this in the findings and discussion only in a ticking-off way. Please use the identified theories for a more reflective discussion. Concerning the relationship to literature the authors need to improve their paper. In the present form, the paper focuses too much on statistical findings and offers no theory/practical-led explanations.

- In the introduction, the authors need to focus more (much more) on the why aspect of the paper. There is too much emphasis on the literature aspect and not enough on what motivated the authors to write this paper. For example, it reads like the authors wrote it just to see what happens adding the Twitter's daily happiness index into the 3-factor Fama-French and Carhart momentum factor, but does anyone actually care about that? Do not make the readers do the work, tell them exactly in the early part of the introduction section. As there are many similar published papers regarding the relationship between the Twitter's daily happiness index and the stock return, the authors need to work more on the contribution in the introduction.
- The writing, especially the results section, needs professional editing. To make the paper more readable in the results section, add sentences in addition to the authors' straight description of the tables. For example, in addition to saying that coefficient X is negative and highly significant, explain what that means. Some discussion of the results, theoretical and practical implication should be addressed. Moreover, we conventionally put asterisk at the estimated coefficient, not at the t-stat (see In Tables 4-5).
- Overall, I do not suggest publishing this paper chiefly because a whole paper is just a statistical exercise that is not well motivated, executed, or explained.