Referee's report on "Degree selection methods for curve approximation via Bernstein polynomials"

This paper aims to obtain minimum possible degree of the approximating Bernstein Polynomials(BPs) using probabilistic methods that are robust to various shapes of the unknown continuous function. The authors address the issue of the choice of the degree of the BP and claim the necessity of data-dependent automatic choice of m. Then, they propose a probabilistic method based on a previous knowledge of a possibly existing turning point on the target function and derive two criteria to serve as stopping rule for the degree selection. Simulation studies and real data with Berkeley growth study are performed to illustrate the merit of the proposed approaches. The manuscript is well written but there remain a few things to be complemented and clarified in the proposed approaches. Here are some of my detailed comments on the manuscript.

- It is not clear how critical or important the issue of selecting the optimal degree for BP ch improvement is from the empirical, computational or practical point of view. It is generally guided to use large number of degree for BP, for example, half of the sample size n as given by default in the R package `bisoreg` of Curtis and Ghosh (2011), also cited as [9] in the paper.

- The starting point is given with the prior knowledge of turning point, but there are other functions to be tuned than the starting point, for which turning points do not exist, including monotone functions.

- They use the optimal value of m as $m_{opt} = \{m > 4 : \text{p-value } > 0.1\} + 1$. However, there is no rule for fixing the significance level as 0.1. With different significance level, one can obtain the different optimal values for the degree m. If the selection of m is robust in terms of the significance level, I think the authors may need to provide more evidence or empirical results to convince readers of their reasoning with the selection of m.

- Regarding the real data analysis with Berkeley growth study, I think the overall mean curve oversimplifies the pattern of growth rate considering the results in Table 4. It seems that the detection of turning points is too sensitive and final results in their data analysis also depends on different choices of beta distributions for U_1 and U_2.

- Even though they argue the proposed methods are robust to various shapes of the unknown continuous function, they do not consider shape-restricted curve estimation. It would be more interesting to apply their methods to the shape-restricted curve estimation, for example, for the monotone function estimation with BP as given in the work of Curtis and Ghosh (2011). Also, some further considerations or comparisons with the conventional model selection approaches such as marginal likelihoods, BIC, WAIC, or DIC are expected. Finally, it would be useful to provide code or R package for their methods, so that implementation and reproduction of the authors work can be useful to practitioners and other statisticians.