

Referee report on the paper, “Democracy and taxation”, by Balamatsias Pavlos

Not sure why the title of section 2.1 is what appears in the paper when it discusses mainly waves of democratisation. Also, the title of section 2.2 is bit misleading to me. These two sub-sections, as I see them, are more about what determines democracy (rather than other way around, i.e., how democracy affects taxation). The latter section does suggest that endogeneity is possibly an important problem in the type of empirical analysis that is conducted in such studies, which is useful. In fact, section 2.3 could be the (entire) Literature Review, and sections 2.1 and 2.2 could even be deleted (other than the bits dealing with the endogeneity issue there), as I do not see much value added to the paper from the inclusion of those two sub-sections. Also, section 2.3 can, and should be, shortened considerably, which can be achieved by compressing the review of some of the relevant papers, and also by omitting some references, for instance, Moutos (2001), Lee (2004), Mooij et al. (1998), etc. The word “writer” used many times in section 2 should be changed to “author”.

It is fine to use an approach similar to that used by Acemoglu et al. (2014) in defining the political regime in place in a country. It would, however, be useful to mention other approaches used in the literature and indicate why this particular one is preferred to others.

As regards the choice of control variables, some of the causations are not too clear to me: for instance, how investment, openness, education affect tax revenues – these need to be explained better. Besides, I could think of the causation going in the opposite direction. This should be clarified.

I think, after Table 1, you could present a correlation matrix showing the pairwise correlation coefficients between your chosen variables.

Information on neighbouring country regimes in a geographical area to create waves of democratisation is a good way to capture the desired relationship studied in this paper. The 2SLS estimations make sense. But why not run a regression of tax revenues on democracy directly (without capturing neighbourhood effects) and without doing this via two stages, and see what happens?

Is it worthwhile to regress tax *rates* on democracy and tax *revenues* on democracy separately? This is because even if the first effect is positive (negative), the second effect could be negative (positive).

Quite a few robustness checks have been done. This is more than adequate, I think. In fact, I am not sure whether Tables 5a, 5b and Tables 6a, 6b are required, especially the latter. Unless better justification is provided, Tables 6a, 6b in particular should be dropped, in my opinion.

Some of the controls could have been interacted with democracy – for instance, it would be interesting to see how investment in a democratic environment affects tax revenues or how openness in a democracy does the same. Could this be done here?

Here $N > T$: so there could be a compelling case for using the GMM technique rather than the 2SLS method, as Roodman (2006) has suggested. So, shouldn't GMM be tried (at least for the basic model), and the results compared with what has been obtained via 2SLS? If the author feels this is not required, then that should be adequately justified.

Overall, I feel the author should be given the chance to revise and resubmit this paper taking into account my comments and suggestions, but I appreciate that the final decision on this rests with the editor.