Posted by PER PETTERSSON-LIDBOM

Posted on 2/24/2014 4:28 AM

First, l would like to acknowledge that I was in contact with the editor, Alan Auerbach, (May 20, 2012) regarding the submission of my comment to this journal. He responded: "I believe that this comment would be appropriate to post on the journal's web site instead. Let me stress that I am not suggesting that your comment is not useful or important. Rather, I believe that, given its nature, it would be best for it to appear on the web site with the paper, which will make it immediately available to any reader of the paper's electronic version, which as you know is a growing share of the overall readership." Hence, this is the reason for posting my comment on the Journals Web site.

Regarding the authors' response to my comment, it is important to stress that footnote 25 was added only after they received my comment in May 2012. At this time, the paper was already accepted for publication and posted on the Web site as forthcoming. Such type behavior raises concerns about misreporting. Similarly, it also seems quite suspicious that 4 of the total of 12 entries of Table 2 in their paper are left out intentionally by the authors although these results could easily been reported. Most importantly, the omitted results do not lend support the authors' conclusion about the stability of the estimated first-stage effect made on page 181 (On this point, see my original comment). Furthermore, the authors have also failed to follow by now firmly established RD (kink or not) procedure of reporting results from an optimal bandwidth choice as well as from half the optimal bandwidth and twice the optimal bandwidth (e.g., see Imbens and Lemieux 2008). The important point here is that the choice of bandwidth should not be selected by a researcher so as to avoid concerns about data mining. Had they reported these standard sets of estimates, it had been very clear that their first-stage results are highly non-robustness since these estimates are both positive and negative. For example, the optimal bandwidth is 4.4 according to the Imbens and Kalyanaraman (2012) procedure (it is even smaller, 2.9, according to the Calonico et al. (2013) bandwidth selection procedure). In this case the estimated first-stage effect is 1.72 with a standard error of 1.23. The estimated effect for half the optimal bandwidth is -2.83 with a standard error of 2.08. It is also noteworthy that the statistically significant results for the smallest bandwidth of 5 in their Table 2 is no longer significant at conventional levels (5% or 10%) if the bandwidth is only marginally decreased to 4.8. Again, reinforcing the extreme fragility of their results to the choice of bandwidth.

Turning to their argument that the regression kink (RKD) estimator is biased for small bandwidth. This statement is simply wrong since an OLS estimator (i.e., local linear regression with a rectangular kernel) is always unbiased independent of sample size. However, they are right in pointing out that often the sample size is often not very big in an RD design. In such a

case, if the analysis is based on a large bandwidth (or if the degree of the polynomial is so small) then it isn't really regression discontinuity that is identifying things. To conclude the results of their study do not survive the conventional set of robustness checks of an RD design (kink or not) and their analysis is not based on a local discontinuity since their results does not hold for the optimal or smaller bandwidths.