**Response to Reviewer 2 of JORS-14-221, "Do State Business Climate Indicators Explain Relative Economic Growth at State Borders?"**

**Thank you for your helpful comments. We have worked to address your comments as well as those of the other reviewers'. Below, please find specific responses to your comments in bold.**

Comments to the Author

This paper in some ways extends the JRS paper by Kolko et al. (KNM), in this case estimating the effects of business climate indexes at state borders, on the argument (discussed below) that this gives cleaner identification of the effects of the policies captured by the indexes. It ultimately reaches conclusions fairly similar to KNM (although you wouldn’t know that from the intro), but its contribution could be much sharper if it hewed closer to the original paper and isolated the differences attributable to the identification strategy. Without that, it is very hard to know what to make of the results from the combined set of papers (and there is recent one by Neumark and Muz from the NBER). Comments follow:

Major comments

1. The key contention in this paper is that we can better identify the effects of the policies captured by business climate indexes by looking at narrow areas along state borders (matched cross-border counties). The idea has some intuitive appeal, but there are some problems.

As the author notes, this strategy has been used in recent minimum wage work (Dube et al.). However, the author seems unaware of a budding follow up literature that raises questions about whether this actually is a better identification strategy (and follow up defense by the original authors). Whatever one makes of that debate, I think there is no question that it establishes that one wants to test whether the assumption that cross-border controls are really better is true. There are numerous comments in the paper asserting that they are, but that is not enough. Note, by the way, that the argument in Dube et al. is quite different. It is about contemporaneous or recent shocks that they argue are similar on two sides of the border.

There are many studies that estimate policy effects using cross-border designs. The author should discuss these, and what does or doesn’t make sense about it in the current context.

The discussion of why exactly cross-border counties provide a better experiment is vague. See pp. 4-5 and elsewhere. What exactly is controlled for by doing the cross-border comparisons, and under what assumptions? The author instead lists everything that could possibly contaminate the state-level regression (endogeneity, heterogeneity, etc.), and then asserts without much argument (and no evidence!) that the cross-border approach is more reliable.

The author also refers to the value of comparisons with the many states a state may border. But I don’t see any way that comes out in the analysis.

There are clearly reasons the cross-border controls may not be good. There can clearly be positive spillovers, which would attenuate any effects. Does northern NJ not benefit from proximity to NYC? It might be useful to think about two extremes. Do we want to estimate policy effects using cross-border regions, or separate islands? That depends. If there are important spillovers then islands would be ideal. If there are important unmeasured sources of heterogeneity, that are more similar for cross-border regions, then the author’s design makes more sense. But how do we establish this? This comes back to the issue of working much harder to validate the design.

The author seems to have in mind the idea that state border regions “compete” with cross-border regions, but state interiors do not. What regions actually compete with each other, and in what industries? What markets are local and what markets are national, and do the results differ for the former (if there is competition across state borders)? This is all too vague.

There is an important issue of what the border regions look like. Near NY, they are very dense with lots of economic activity. In CA, the border regions are essentially empty. One would think this could matter a lot for the interpretation of the results. This is an issue that has been addressed in other work using cross-border regions. A related point is that if a cross-border region is urban, then cities may adopt policies to compete with cross-border regions, in which case the state business climate measure may simply be less relevant at the border, attenuating effects.

2. One thing the paper emphasizes is whether indexes do a better job predicting future growth or past growth. The author takes it as a sign that an index is “bad” if it appears to predict past growth just as well. In this particular context, that may not be very sensible. First, as the author notes, the intertemporal correlation of these indexes is very high, so there may be little difference between the index at t predicting growth from t-5 to t and the index at t-5 predicting this growth. Second, my sense is that the indexes have a lot of inertia and don’t always use up to date policy parameters. Given that, it would not be at all surprising if the indexes seem to predict past growth. Also, didn’t KNM do a similar kind of test as a way of assessing policy endogeneity, although predicting the indexes with past growth? The author should clarify the differences. Does that fact that only the New Economy Index does more to look backward suggest that in general there is not a policy endogeneity problem?

3. The paper at some points seems like a direct extension of KNM, and at other seems quite distinct. I think especially because JRS published the original paper, it would be far more useful to readers for the author to hew very closely, for at least part of the paper, to what KNM did, but then substituting the cross-border design to see if it makes a difference. Otherwise it is very hard to tell what is going on. There are many other places, too, where there is not sufficient attention to the overlap. For example, the paper discusses the content of the indexes, but that was discussed in detail in that paper. At the same time, things are a bit muddled with respect to comparing the two papers. The last paragraph on p. 2 sounds like the paper finds no evidence that the indexes predict anything, whereas the conclusion sounds much more like the findings from KNM. But since sample periods, regressions, and some indexes (including names of the same indexes) are not consistent, it is very hard to compare.

4. When the author discusses the fairly low contribution of the indexes to the R square, I don’t know what to make of this. Compared to what? For example, the regressions appear to have no other controls (which is weird in light of KNM), and we wouldn’t expect them to be important relative to the business cycle (there are no time effects). There aren’t many papers that make much of explanatory power. What is the rationale for doing so here?

And simply throwing them all in to establish an upper bound for explanatory power makes no sense at all. I can’t think of a statistical or conceptual rationale for this. Moreover, pointing to low R-squares to say the indexes are “just not that good” doesn’t really make sense. Again, what is the context and comparison?

5. On p. 8 the author suggests that somehow the state border design helps with the problem that states may react to their neighbors’ policies. How does the state border design help on this score? Doesn’t it make the endogeneity worse?

6. Most empirical research on county-level data uses QCEW employment, and not much else. How good are the BEA regional data used in this paper? What are the data sources?

7. The author notes that for some outcomes (like wages) there is no prediction that policies will create some differences at the border. Then why use these in the analysis?

**As a counterfactual**

Minor comments

1. What does the author have in mind as state specific productivity shocks (p. 1)?

2. There needs to be more careful discussion of what we are estimating when we put the BC indexes on the right hand side of these kinds of regressions. KNM have some nice discussion of this, whereas it is left a little vague in this paper. This is another place where I can’t see any disadvantage to sticking much closer to the original paper and highlighting key differences in empirical strategy.

3. What is the value of the many quotes arguing that the indexes compete? I don’t think the interesting question is simply which index is better.

4. On p. 5, the author says that indexes that included outcomes were not used in the analysis. But KNM showed that you can get the raw data and just strip these out. The same issue applies to those reported as rankings. Why not use the raw data as they did?

5. P. 6 says that all the indexes consider various aspects of quality of life. I disagree. Some narrowly focus on taxes and costs.

6. Figure 1 was not included, and perhaps because of this I have no idea what the paragraph following the reference to the figure (p. 10) means. And why is the comparison to clustering of interest? Do these two approaches in any way address the same problem? (Of course clustering doesn’t affect the point estimates anyway.)

7. P. 18 states that the indexes lack “the scientific rigor typically required of social science research.” What does this mean, exactly? How can one possible assess rigor based on whether or not we get predictive power? Maybe the theory is wrong.

8. I have no idea what the author’s assertion about indexes performing best immediately after issuance (p. 15) means.

9. P. 16 describes the effect of a 10% increase in a BC index as increasing growth by less than 1% as small. Is a 1% effect small, if this is a growth rate? That seems large to me but I may be misinterpreting the magnitude.

10. I don’t see any basis in the paper for the conclusion in the last sentence. What does the author have in mind regarding forecasting tools? Using the indexes differently? If so, then why not do that?