August 8, 2018

To the Editor:

Please find enclosed the revised version of our jointly authored paper “Demand and Supply Inflation in Floating Catchment Areas (FCA) Methods” that you kindly invited us to revise and resubmit for publication consideration to PLoS ONE.

We would like to begin by thanking two reviewers for their constructive comments and suggestions.

We thank you for the attention paid to this submission and look forward to hearing back from you in due course.

Sincerely,

Antonio Páez

(on behalf of the coauthors)

I have another dinner to go to… this time at the boss’ house. But some quick thoughts. Beyond just the inflation factors… When you do a correlation on the un- vs adjusted accessibility numbers, they are very highly correlated ~.98-.99. In absolute terms, it doesn’t appear like we add that much as it all sorts itself out in the 2SFCA. That said, if you normalize the values and compare them with the standard (xi – xmin)/(xmax – xmin), you do see some differences across DAs. For example, the unadjusted has estimated access that is about 50% greater for some DAs than in our adjusted method. Many are around 5-10%. On the other side of the coin, the unadjusted numbers can be around half of what we estimate in the adjusted method – so here the 2SFCA underestimates access. This is a more nuanced angle than the 8x inflation factor story. I don’t have time to plot at the moment but tomorrow I can see the spatial distribution of this relationship.

BTW I have done these numbers just for Hamilton DAs at the moment (just city) as a test run. Similar things should be seen across the CMA.

**Reviewer #1**

The authors propose a very interesting adjustment to the 2SFCA method. I’m still unsure that these adjustments are the panacea as suggested by the authors (i.e. inflation issues are ‘eliminated’) – I’d like to do my own testing (in the future); however, they are a worthy addition to the literature for further academic debate.

First of all, many thanks for your thoughtful response to our submission. If you are interested in conducting testing of your own, we would like to remark that the data and code are placed in a public GitHub repository as follows:

<https://github.com/paezha/Demand-and-Supply-Inflation-in-Floating-Catchment-Area-FCA-Methods->

In what follows we respond to your comments and indicate the actions that we have taken in response.

One fundamental problem throughout the paper is their referral to step 1 as if is measuring the actual or real demand on that supply point. Clearly that is not the case because, in most scenarios, populations will have multiple supply points that are accessible and thus their demand will be ‘averaged’ across these.

Thank you for this comment. This is something that we have clarified. “Demand” in a model of accessibility represents potential demand – since we do not really know how many members of the population actually have a family doctor. Even in a universal care system such as Canada, there are members of the public who do not have a regular health care provider – approximately 15.8% of the population in 2016, for example (Statistics Canada, 2017). Assuming, however, that all members of the population *do* demand health care (universal coverage as the design parameter, so to speak), one thing that becomes clear from our approach is that the “averaging” that intuitively takes place across the multiple supply points tends to be biased in earlier 2SFCA approaches. The reason for this is that the distribution of population and points of service is usually not uniform across space.

Typically the 2SFCA is measuring ‘potential demand’, but the authors effectively are presenting it as ‘actual demand’ (when in reality this is far from true). This very important distinction needs clarification upfront. The authors present an adjustment method that I distinguish as recalculating the ‘likely demand on supply points’. Similarly, accessibility (Step 2) is ‘averaged’ across multiple supply points. The effect of demand inflation and supply inflation to the 2sfca method really should be looking at the combined effect – i.e. at the Ai (accessibility) scores.

Their simple examples in Figure 1 and 2 are highly extreme (wholly un-real) scenarios, which are very effective at demonstrating how the 2sfca method can fail. However, their extremeness also makes it hard for the reader to evaluate whether, in a more ‘normal’ scenario, the same deficiencies still exist or are a significant concern. I believe these examples need to introduced as such or similar: “These examples have been purposefully designed to demonstrate key weaknesses of the 2sfca method; we acknowledge these scenarios are not representative of typical demand and supply landscapes”

Overall, I like the concept of the demand vector summing to equal the original population (demand) size – even after the application of impedance weights. In theory, this enables the generation of accessibility scores that more truly are interpretable ratios (e.g. health providers per population). Having said that, their adjusted weights – which divide the original weight by the sum of all weights for (potential) service points – still assumes that all service points are equally ‘attractive’. I strongly believe this is not the case at urban fringe locations where larger urban populations are unlikely to be attracted to nearby rural services, but the reverse scenario is very different.

Similarly the concept of saturation of intervening opportunities and variable catchment sizes seem pertinent here (i.e. why would a population travel further if sufficient opportunities are available at nearer locations?). Some recent 2sfca methods papers have introduced variable catchment size modelling, which I believe are critical for larger geographic scale applications (e.g. for national modelling).

As seen in Figures 5 and 6, the demand inflation factor is mostly a factor of urbanized populations only. I didn’t understand the Accessibility scores in Figures 7 and 8 or 11 and 12 – how do you get an Accessibility ratio of 1 (i.e. equal supply/demand ratio)?

Again, the values in Figure 9 and 13 make no sense to me. Are the authors truly claiming that ‘current’ accessibility scores (e.g. 2sfca, e2sfca) are inflated 4 to 9 times higher than they should be? In all of the published work on floating catchments, I’ve never seen any evidence that even vaguely suggests they are out by this factor level.

Figures 14 and 15 are very interesting in that they suggest a very systematic bias between metropolitan and fringe/rural locations (strongest reds in the most metropolitan, strongest blues in the outer edges) – I suggest this is a key finding that needs further discussion.

Statistics Canada. (2017). Primary health care providers, 2016 (Catalogue no. 82-625-X ). Retrieved from https://www150.statcan.gc.ca/n1/pub/82-625-x/2017001/article/54863-eng.htm

**Reviewer #2**

This is a very interesting paper that presents an issue that I've grappled with in the past. Per PLOS ONE's reviewer policy, I'm going to unblind myself in an effort to promote a constructive dialog (and because the content of my review would likely do this anyways).

Major concerns/issues/comments

L89: This equation is somewhat confusing as presented. Are these all from different metrics?

L104-105: Could this also be conceptualized as the cost required to overcome distance?... e.g., rather than making it about preferences, it seems like this passage should acknowledge that overcoming distance/time is a “cost” of using the service, rather than just framing it as preference-based.

L347-349: I feel like I'm missing something here. The overall ratio of the system does not appear to be preserved, given my understanding of what the numbers represent. I think that they are (for 2SFCA): Prov/Pop ratio for A = 0.066 and for B = 0.006. In that scenario, the overall “system” ratio would be 0.46 (not 0.02) in this calculation. (If I'm interpreting the output incorrectly, I apologize. However, if that is the case, I think it would be extremely helpful to extend the example calculation to clearly demonstrate the output values and how the overall system ratio is preserved in the resulting values)

Using the row standardized weight matrix to “allocate” the population to the various facilities in the first step appears to be very similar to the example I provided in Delamater, 2013 (citation 20 in the ms) on pages 33 and 34 and illustrated in System's #4 and #5. In this case, I was concerned about the overall effect of splitting populations into discrete “pieces” and it appears that this approach might have similar issues. If I'm understanding the authors' approach correctly, I think that the 2SFCA calculation for the Figure 1(I) system would result in an accessibility value of (0.03 + 0.03 + 0.03) = 0.09 facilities per person. This output doesn't really make sense to me, purely from a logical perspective – and it would concern me that the approach does not perform well in the most simple of systems.

My overall thought is that the problem of supply and demand overestimation might be “problem” when each is considered in a vacuum, but not in the output/results when the two steps are integrated. I would suggest that the Demand and Supply Inflation section be expanded to show more simple examples – and most importantly how inflation actually effects the final sp accessibility values (A) in these systems, as a “proof of concept”. As of right now, I can clearly see the issue in the two separate steps, but I'm having a hard time understanding what the demonstrated effects of this are on the final outcome [e.g., is A (sp accessibility) under/over estimated for Pop A?... Pop B?]. This was one of my concerns when I critiqued the Wan et al 3SFCA paper in Delamater, 2013... I was not sure that competition really “needed” to be accounted for to adequately capture potential sp accessibility. In this case, I'm not sure if the inflation happening in the two steps is having a detrimental effect on the results. I think that clearly demonstrating issues with the output calculations of the E2SFCA and 2SFCA (in more than a single system) would provide a much better justification for this approach (than is presented currently).

I did not do a deep dive into the Results yet, given my concerns about the approach. However, I do wonder if using a ratio to describe inflation is a good approach, especially given that the output has true units (people and supply/people), e.g., from 0.002 to 0.004 and 2 to 4 would have a similar inflation factor... but would have very different absolute differences.

Minor concerns/issues/comments

L90: Should this be Sj, not Si?