Reviewer's report

Title: The supply of General Practitioners across local areas: accounting for spatial heterogeneity

Version: 2 Date: 8 June 2015

Reviewer: Maria Pia Fantini

Reviewer's report:

The study by McIsaac et al. aims to increase the understanding of the key drivers of general practitioner (GP) distribution in Australia, by accounting for the geographical referencing of GP supply data. The study is well designed, the methodology is interesting and the obtained results are discussed in depth. However, the article is excessively long, the structure should be massively rearranged (it resembles more a PhD thesis), and a discussion on the limitations of the adopted methodology should be added.

Major Compulsory Revisions

1. The first section should be rearranged. In particular, I suggest moving the second paragraph (Ln. 90-100) to the bottom of the Introduction (Ln. 129-130). Also, I would remove some sentences which are not suitable for this section of the paper (Ln. 133-139; Ln 141-144). To sum up, the last paragraph of the Introduction should be as follows:

“A major shortcoming of this earlier literature is that it largely ignores that the data are geographically referenced. The geographical referencing, inherent in GP supply data, results in several methodological issues that might influence the validity of results. These include the potential for omitting a variable that is spatially correlated, treating variables as independent when they are in fact related to the values of their neighbours, or using inappropriately aggregated data. The first two problems lead to econometric biases, whereas the third can, as pointed out by Birch et al., lead to incorrect inferences regarding individual behaviour as associations may vary according to the level of aggregation of the data. Ignoring the geographical referencing of GP supply data can therefore lead to erroneous conclusions regarding the effectiveness of policies aimed at redistributing GPs.

This study aims to increase the understanding of the key drivers of GP distribution by accounting for the geographical referencing of GP supply data, and to explore how the factors associated with GP supply are aligned with the arbitrary geographic boundaries of GP supply data. In particular, the independent variables are analysed to determine their relationship to the same variables for their proximate neighbours. This paper represents the first attempt to map the factors influencing GP supply to the appropriate geographic level at which GPs may be considering that factor”.

By doing so, the reading will be easier and the aim of the study will be clearer.
2. I think that the journal does not allow to write two introductory sections (i.e., Introduction and Background). Please move the first paragraph of the Background (Ln. 148-159) after the first paragraph of the Introduction. As for the second and third paragraph of the Background, you may move them to the Conclusions and (in a very shortened version) to the beginning of the Methods, respectively. Alternatively, you may delete them.

3. Please entitle the first section of the Methods (Ln. 176-255) as “Statistical analyses”.

4. For the sake of consistency, please replace “small areas” with “postal areas” (Ln. 226).

5. Please specify that the number of GPs per 1,000 persons is 2008 data (Ln. 258).

6. Please state at the beginning of the Results that the number of postal areas you have analysed is 2032 (Ln. 322-328).

7. The Results section actually ends on P. 14, Ln. 354. Everything after that should be considered as part of the Discussion. As a whole, the authors must try their best to condense their thoughts from P. 14 to 21, because the paper is very long, and many concepts seem to be reiterated.

8. The authors state that “the variable capturing the proportion of indigenous persons is significant in the non-spatial Tobit model and the 10-neighbour band, but not significant in the 5-neighbour band” (Ln. 434-436). This appears inconsistent with the figures presented in Table 3 – the authors seem to have inverted 10 and 5.

9. Another inconsistency between text and tables. The authors state that the coefficient of mortality rate “is negative in the spatial models” (Ln. 444); however, it seems to be positive (Table 3).

10. The authors should address some limitations of the adopted methodology, which, in my opinion, has room for improvement. For example, the choice of a fixed number of neighbours for each postal area is debatable: coastal areas may have less, because in some cases the farthest areas will result to be “too upcountry and far”; rural, remote and very large areas also may have less, because villages that are more than 1,000 kilometres apart cannot be considered “neighbours”. I think that a possible solution is to consider neighbours of coastal and remote areas the 3 closest, and neighbours of all the other postal areas the 5 closest. Or, as an alternative, 5 and 10 respectively.

Minor Essential Revisions

11. I have found some punctuation typos (e.g., Ln. 150, 179, 260). Please check all the paper.
Discretionary revisions

12. Figure 1 clearly shows that there is considerable variation in GP supply across postal areas, but no significant spatial clustering. However, it may be of some interest to present, as additional files, also the detailed maps of the five Australia’s largest cities, where postal areas are very small.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.