Reviewer's report

Title: Are waiting times for hospital admissions affected by patients' choices and mobility?

Version: 3 Date: 28 December 2010

Reviewer: Ansgar Wuebker

Reviewer's report:

The authors tried hard to improve the paper in line with some of my suggestions. They are more cautious in their argumentation, implemented some methodical suggestions and discussed more the limitations of their results. However, there is still room to improve the paper in order to justify publication in BMC Health Services research. I have two major suggestions:

Statistical Methods:

I would suggest that the authors extend the section “Statistics and estimation method”. I appreciate that the authors followed my suggestion to implement an IV-approach. However, it is not clear for me how the authors conduct the instrumental variables approach. Good instruments should show a considerable explanatory power for the “choice dummy” but must not affect the dependent variable (waiting times) once the remaining explanatory variables are controlled for. Did you run a two-stage procedure? (compare for example Cameron/Trivedi 2005 or Cameron/Trivedi 2009 “Microeconometrics using Stata regarding the methods used within the IV-approach).

What is your identification strategy? I.e. what instruments are used to predict the “choice variable” but not included in the original regression? Please provide tests for the instrument validity and relevance. Is it an over identified model? Are the regressors endogenous? You can test these issues (compare the literature hints above). If you do not provide this information reproducibility of results is difficult and trust in results suffers. I am not convinced of the IV-results presented in Table 3, since usually standard errors largely increase.

Did you calculate robust standard errors: I strongly recommend that you should calculate cluster robust standard errors. This is especially important at the hospital level, i.e. you should cluster at hospital level, since waiting times should be correlated within a hospital. Neglecting clustering could be a great problem, since it can lead to great downward bias in default standard errors. (Compare for example Cameron/Trivedi 2009)
how to implement these methods). Finally, did you test for functional specification as suggested in the previous review?

Results:
As asked in the previous review: How many observations are in each regression model?
I suppose that the number of observation differs between the model specifications, because – from table 2 - you can derive that the number of observations changes for the different variables. That is a problem for comparison of the models, if there are some systematic dropouts. In that case you compare apples with oranges. Please indicate in table 3 the numbers of observations and the R-squared statistic. Please provide in addition to the point estimates standard errors in Table 3. Please provide some quantitative interpretation of the log-linear-models. E.g. one year more education leads to two percent decrease in waiting times (compare model 4).

A little bit irritating is the result difference of the ICD Chapter 8 between models 1 and 2 and models 3 and 4. The results changed from highly significant to non significant.

Please check it or discuss it? If it is true, it hints that there could be specification problems. Therefore I requested/recommended some specification tests (see above or last review).

**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'