Author's response to reviews

Title: Wrong Schools or Wrong Students? - The potential role of medical education in regional imbalances of the health workforce in Tanzania.

Authors:

Beatus K Leon (beatusleon@yahoo.co.uk)
Julie R Kolstad (julie.kolstad@econ.uib.no)

Version: 2 Date: 18 October 2009

Author's response to reviews: see over
We would like to thank the reviewer for valuable comments. Our responses are reported below under each comment respectively.

Responses to the referee report

Major Compulsory Revisions

1) The authors describe every variable in detail, except the dependent variable: the willingness to accept a rural post. Is this a yes/no question? Or are there degrees of willingness? I see from the probit analysis that 62% of respondents answered yes to this question, which implies that 62% of soon-to-graduate students would be willing to accept a rural post. This number may seem too high, but hiding it is not the right way to go.

The dependent variable has now been given proper attention in section 2.2 “Model specifications”, where we describe the choice of dependent variable. The dependent variable, as the referee suggests, is a binomial variable.

The predicted probability of answering yes to the question of whether or not the respondents would accept an offer of a job in a rural area is indeed very high. This fact is now presented in section 3.2.4 “Predicted probabilities of accepting a rural job”. In the same section we discuss implications of this finding, with a particular focus on the interpretation of the other results in our analysis.

2) The authors should state the core problem with interviews; that they do not necessarily represent preferences and that they are likely to be biased towards what are seen as socially acceptable views. This would go a long way towards explaining the 62% mentioned above. We know that many fewer than 62% will actually accept the posting. However, it is probably true that 38% would refuse a posting. Thus, the question does contain information about willingness even if it is biased towards socially acceptable views. Since the probit analysis includes a constant, it is not the raw percentage that matters, but whether or not we can describe the difference between those who say yes and those who say no. Here the authors need to be honest and to discuss what they think was happening, but the analysis is still valid.

We have now incorporated a paragraph where we discuss the strengths and weaknesses of survey data (the second paragraph of section 2.1 “The data”).

The implications of using survey data for the interpretation of our results are discussed, as described under point 1), in section 3.2.4 “Predicted probabilities of accepting a rural job”.

Minor Essential Revisions

1) There is no such thing as a simple multinomial probit. In fact, it is essentially impossible to do a multinomial probit without very high tech simulation tools. I suspect that the dependent variable is simply yes/no, which means it is binary.
Thus, the authors have almost certainly done a binomial probit, not a multinomial probit. This is straightforward. In addition, the sample mean for the dependent variable is different for each of the three regressions, which means the sample size is also changing. The authors need to be clear about how the sample size is changing across the three regressions.

The referee is right; our dependent variable is binomial, and we have therefore done a binomial probit analysis, not a multinomial one. The model is now presented more thoroughly in section 2.2 “Model specifications”, where we also describe the dependent variable in detail.

Information about the sample size, the goodness of fit and degrees of freedom are now included in Table 5 (the table showing our regression results, Table 3 in the initially submitted version) in the last four rows. Variations between the models are discussed in section 3.2.5 “General comments to the regression analysis”.

2) In the third paragraph of the background section, the authors imply that poor access leads to low health outcomes and high levels of poverty. It is safe to say that poor access, poor health outcomes and poverty are all high correlated, but it is completely unjustified to imply that poor access causes both poor outcomes and poverty. This needs to be reworded to remove any suggestion of causality between health care and poverty.

We certainly have no evidence that allow us to take any stand in this ongoing debate. As improved health is among the development goals in itself and poverty is not the main focus in this paper, we have simply excluded the sentences that refer to the relationship between health and poverty. The focus in the third paragraph of the background section is now solely on the relationship between the density of health personnel and improved health.

3) The authors make much of the fact that medical schools do not recruit rural students. But is this a problem with medical schools, or with higher education in general? Is there any reason to believe that engineering or business schools do any better? It seems likely that the underrepresentation of rural students is a problem everywhere in the world. Your opinion on this matter would be interesting to any reader of the paper. Is this problem something special to medicine?

This is a valid observation. Since the recruitment base is essentially the same, it may be assumed that the problem applies to other higher educations in Tanzania. However, we have no evidence to support this notion, and think this might be an avenue for further research rather than a topic in this paper. Moreover, there is one important difference between health personnel and other professions requiring high education levels; the impact of maldistribution of health personnel on societies in the rural areas is probably higher and with more immediate consequences. It may be all right to have one engineer visiting several districts offering advice, but medical staff must reside in the district to handle emergencies. Our main concern in this paper is that the under-representation of rural students in medical schools is related to later rural-urban imbalances in the distribution of medical staff. We have therefore decided not to discuss the recruitment of rural students into other professional trainings.
Discretionary revisions (but highly recommended)

1) The fact that you have students in three medical schools is very important and you must at least take this difference into account. The fact that students who grew up in urban areas outside of Dar are different than those who grew up in Dar, is probably completely driven by the fact that these students attend KCMC. You need to include a variable indicating the school they attended, at the very least. However: …

This is a valid point; due to differences in intake policies, financing possibilities and curricula, one could imagine that the school the respondents are attending could be related to their willingness to work in rural areas. To investigate this, we have run regressions where we have included dummies for the three schools as independent variables. It turns out that we are not able to identify a relationship between any of the schools and the probability of accepting a rural remote job (see Model 4 in the attached document “reg_wrong_schools_17092009”). None of the school dummies, neither by themselves nor in pairs with the other school dummies, were significant. Furthermore, the size and significance of the coefficients for the other variables did hardly change when the school dummies were included. Neither did the predicted probability of accepting the rural job.

Moreover, there was practically no correlation between the school dummies and the variables initially included in the regression analysis (see the attached file “corr_wrong_schools_17092009”), so there should not be a problem of these variables picking up possible school effects when school dummies are not included.

Since the results reported in the initially submitted paper turn out to be solid and unaffected by the inclusion of med schools, and there seem to be no relationship between which school a respondent is attending and the probability of accepting a rural job, we decided not to include school dummies in the regression analysis. However, we have now included a footnote that explains that there is no such relationship in our data. Footnote 7:

"It could be that due to differences in intake policies or financing possibilities, and thus a direct or indirect selection of students, which school a student is attending was related to the probability of accepting a rural job. However, we were not able to establish any such relationship in our data."

2) Study the differences between the three schools. They must be different from each other and they must have different outcomes. Does one attract more rural students than the others. Is one more likely to graduate students who are interested in working in the rural areas? Does one manage to not discourage students? If they are different in outcomes, why? What are they doing differently. You should examine the differences between the schools to learn more about policy.

Since we were not able to find any relationship between the schools and the probability of accepting a rural job, we feel that such a discussion will be a bit misplaced here.
General revisions

We have restructured some parts of the paper so that is more consistent and focused. The background section now consists of three subsections: 1.1 The number and distribution of Tanzanian Doctors, 1.2 Medical students and the HRH imbalance, and 1.3 Previously identified predictors of willingness for rural medical practice. All the information that was there before is still available; we have only restructured the section in order to make the background even clearer.

As a direct result of the first three comments from the reviewer, we found that a general clarification of the methods used was necessary. The data and the methods of the analysis are now removed from the results section and are more thoroughly explained and discussed in section 2. This section now consists of two subsections; 2.1 The data, and 2.2 Model specifications. We have also incorporated more discussion related to our model choices and in particular their implications, in the result and discussion section (section 3). In section 3 we have furthermore added a simple correlation matrix, Table 6, in order to discuss the validity of our results.

Instead of the six figures that were serving as descriptive statistics in the originally submitted version of the paper, we have now chosen to incorporate all the descriptives in one table, Table 1. This choice was made because we think that the descriptive statistics should give as whole a picture as possible and that the information in the six figures was too scattered to give this whole picture.

We have furthermore updated the list of references and added the following references to the already existing text: