Reviewer’s report

Title: Predicting Inpatient Hospital Payments in the United States: A Retrospective Analysis

Version: 3
Date: 27 February 2015

Reviewer: Paul Barnett

Reviewer’s report:

Major Compulsory Revisions.

1. The number of observations and number of excluded observations are not provided, except in the supplementary material. This needs to be added. Page 11 Paragraph 3 What is the number of hospitals in the dataset? How many observations were excluded? It appears from appendix B tables 1 and 2 that 49 of 1039 hospitals were dropped. It is worth noting that the sample includes about 20% of U.S. hospitals.

2. The information in Table 2 is confusing, and seems to conflict with the description in the text. Page 12 Paragraph 1. Table 2 and the description of Table 2 are very confusing. Table 2 is labeled “hospital level mean”. If these are hospital level mean values, and if every hospital has a PCR for every payer, why do hospital and market characteristics differ by payer? Perhaps this is because not every hospital has a PCR for every payer, that is, there is a different number of hospitals in each column? It would be helpful to indicate the number of hospitals in the data, and if they are not the same for each payer, then the number by payer. The following illustrates the reviewer’s confusion. Table 2 is labeled “hospital level means.” In the column labeled Medicare the value of critical access hospital is 0.95. This suggests that 9.5% of the hospitals were critical access hospitals. Every critical access hospital has either 100% critical access stays, or 0% critical access stays. The description of data in table 2 suggests that it is a mean of the value for the number of stays, that is 9.5% of Medicare stays were in critical access hospitals. This is entirely different from saying that 9.5% of the hospitals were critical access hospitals. Its seems unlikely at nearly 10% of Medicare stays take place in these small facilities. In Table 2 below the entry “hospital and market (%)” the first 4 entries appear to be the percent of hospitals that have this characteristic, but the values below this are the mean values of continuous variables. Consider changing the labels.

3. Table 4 appears to provide prediction estimates “within sample,” that is, parameters are derived from the study hospitals, and predictions are how well the payment to charge ratio is predicting is obtained using the same hospitals. If this is correct, the analysis is overly optimistic. Appendix A has the out of sample predictions: these need to be summarized so that we have a single estimate of the out of sample accuracy for each payer. This information is the essence of the paper and should not be relegated to supplementary material. Page 15
4. It is not clear what the “predicted PCRs” are being show in Table 4. Is this the within sample prediction? Are the predictions for the same observations as were used to estimate regression model parameters? This is not as interesting or as relevant as the out of sample predictions in Appendix A. The root mean square errors in appendix A are always larger, so the values in Table 4 does not appear to be the performance of out of sample prediction. Don’t the values given in table 4 for the mean percent absolute value of error actually overstate that accuracy of the method, as the same observations are being used to develop the model and then to test its fit? Isn’t the information in Table A 1 actually address the major point of the paper?

Minor essential revisions.

4. Page 8 Paragraph 3. “We eliminated the small number of hospitals with PCR values below zero or above 1.” Presumably in some cases, these outliers were present in less than all 5 PCR values. There should be a justification for dropping all other PCR values from these hospitals from the dataset.

5. Page 14 Paragraph 3. Average DRG weight is a regressor in table 3, but is not described in the methods. Average DRG weight appears in Table 2 as a hospital & market characteristic. Is this the average DRG weight of all patients in the market, for all patients in the hospital, or for all patients of this payer at this hospital? If the latter, why isn’t it considered an “average patient characteristic?” DRG weight is used to determine payments, and is highly correlated with charges, and thus it is not surprising that it is not correlated.

6. How does the method described in this paper compare to a simpler method, simply using the Medicare PCR to estimate private insurance payments?

7. Page 18 Paragraph 3. Estimating Medicare payments is not so difficult using the CMS supplied pricer software, in fact, it would be quite a bit easier that building the dataset of hospital level values needed to use the parameters from this regression. The authors may have more justification in using this method to estimate private payments, and should probably lead the discussion session with this.

8. Page 19. Paragraph continued from page 18. This reviewer does not see how this method will be useful in understanding if hospitals decline to admit patients with public insurance or provide them with lower quality care. This sentence should be supported, or else dropped.

Discretionary revisions

9. p. 5 last paragraph of Background. It is stated that the goal is to estimate payment to charge rations (PCR) for hospitals in states that do not release complete information based on information on hospitals in the states that do release information. It would be helpful to the reader to briefly indicate in the introduction the source of data that would be used to make the prediction.
10. p. 5 paragraph 1. “Our goal was to model the ratio of payments from all payers to the hospital charge.” Isn’t the goal to model a separate ratio of payments to charges for each payer?

11. Page 6. Paragraph 3. It would be helpful to briefly understand how the inclusion of residuals from other equations as independent variables corrects for endogeneity.

12. Page 7. Paragraph 3. The description of the critical access designation might be clearer if it indicated that this is a part of the Medicare reimbursement method.

13. Page 8. Paragraph 2. Consider rephrasing so that it is clear that it is variables, not data sources, that were likely to have an association with PCRs.

14. Page 9. Paragraph 1. Including the volume of stays and its square are attempt to see if payment is related to scale without the assumption of linearity.

15. Page 10. Paragraph 1. Table 1 lists sources of state policies. This statement belongs with the description of the state policy variables that begins on page 8.

16. Does the APR DRG weight show that, within DRG, there is a loss in caring for sicker patients sicker and a gain (profit) in caring for patients who are less sick? This seems like a worthwhile point to be made.

17. In OLS regression the mean of the dependent variable and the mean of the fitted value of the dependent variable are equal. Is this a consequence of the log link function?

18. Page 16. The extensive discussion of the appendices is problematic if the reader can’t see them. Appendix A seems essential to the purpose of the paper, and not supplementary material. The % absolute value error from out of sample prediction should be reported in Table 4, not relegated to the appendix.

19. Page 19. Paragraph 3. The estimated prediction errors are “within sample” predictions, when the goal of the paper is to see if the method can predict PCR out of sample. The example uses $1,000 payment for an “encounter.” Aren’t the authors trying to predict the payment for a “stay”? Encounter seems like the wrong word as it implies interaction with a specific provider. The example would be more realistic if the payment were $10,000. The reported errors in prediction should focus on private insurance, which is what we are interested in predicting.

20. Page 20. Paragraph 1. It is stated that using CCR instead of PCR will introduce significant errors. It should clarify “in estimating the payment.” CCR would be superior in terms of estimating the cost of a stay from the hospital’s perspective.

21. The range of PCR values in Table B1 could be easily added to Table 2.

Level of interest: An article of limited interest
Quality of written English: Needs some language corrections before being published

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

Declaration of competing interests:

I am the former colleague of the first author, and I am the first author of a paper now under review at a different journal that includes the first author of this paper as a co-author.