Author's response to reviews

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

Authors:

Mark W Smith (mark.w.smith@truvianhealth.com)
Bernard Friedman (econbarry@gmail.com)
Zeynal Karaca (zeynal.karaca@ahrq.hhs.gov)
Herbert S Wong (herbert.wong@ahrq.hhs.gov)

Version: 4 Date: 29 May 2015

Author's response to reviews:

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.

> The number of observations appears at the top of Table 2.

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.

> We have added a paragraph at the start of the Results section:
“The contributing states had 1,144 hospitals in 2006 [2], about 20 percent of all US community hospitals. After eliminating hospitals with missing or outlier PCRs, the count of hospitals by category was 1,110 for Medicare, 1,061 for Medicaid, 1,105 for private insurance, 899 for Self-Pay, and 868 for other insurance.
The correlation table has been updated as well, as noted in the Methods and Results and described below.

2. 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.

> Correct, the number of hospitals varies by payer. This is indicated in the row labeled “No. Obs.” We have moved it from the bottom of the table to the top.

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.

> Here are two calculations to suggest that the 9.5% figure is reasonable.

In November 2014 there were 1,325 CAHs (http://www.raconline.org/topics/critical-access-hospitals#how-many), roughly 25% of all hospitals in the U.S. Although CAH’s are small on average, given the large number of CAHs it seems reasonable that they could account for 9.5% of Medicare-funded hospital stays.

We can also make a back-of-the-envelope calculation using 2010 numbers. According to MedPAC, there were 48,420,576 Medicare beneficiaries in 2010, and 23% (11,136,732) of them lived in rural areas. Based on 2010 HCUP data, there were 13,862,942 hospitalizations registered to Medicare enrollees, and the residency for 2,972,596 (21.4%) of patients were rural areas. As few as 44.3% of Medicare-funded stays for rural residents could be in CAHs and still we would get 9.5% of Medicare stays being in CAHs. (That is, 9.5% ~ .443 * 2,972,596 / 13,862,942.) Prima facie it seems reasonable that 44% of Medicare-funded stays for rural residents are in critical-access hospitals.

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.

> We have removed “(%)” from the subheading and have inserted it after the first four entries.

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 Paragraph 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 do 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?
We have removed the in-sample predictions (old Table 4). The new Table 4 represents the out-of-sample predictions (old Appendix A).

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.

> That sentence was misleading and has been dropped.

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.

> We regret the omission and have added Average DRG Weight to the methods and to Table 1. It is the average DRG weight of discharges at the hospital within each payer category. Thus the weighting is by discharge rather than by person, because a single individual could have more than one discharge in our dataset.

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

> Table 5 indicates that the Medicare PCR is highly correlated with the PCR for private insurance (#=0.75). Whether this is acceptably high is a matter of personal discretion for the researcher. The tradeoff is accuracy versus ease.

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.

> The PCRs for Medicare were not based simply on the pricer algorithm. We have added this paragraph on page 19:

The Medicare PCR represents an alternative to the Medicare CCR and to the payment estimated by the pricer software. Levit et al. [2] explains the development of the PCR in detail, noting that the average price of hospitals stays in the 10-state sample is considerably different under the Medicare pricer ($9,850) and the Medicare PCR ($8,418). This difference is enough to suggest that simply using the pricer output could lead to a substantially different answer to most research questions.
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.

> We have dropped the sentence.

Discretionary revisions

9. p. 5 last paragraph of Background. It is stated that the goal is to estimate payment to charge ratios (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.

> We have added the following sentence in that paragraph: “Data sources for the PCR estimates include patient demographic and clinical information from administrative claims data, as well as publicly available information on the hospital and its market, and selected state policies.”

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?

> We appreciate the suggestion and have rephrased the sentence that way.

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.

> We respectfully decline this addition, as it would likely go beyond the statistical knowledge and interest of many readers.

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.

> The sentence now indicates that it refers to Medicare payments.

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.

> The sentence now reads “All of the variables were associated with…”

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.

> We appreciate the suggestion and have added a similar sentence: “Including the square enables us to estimate whether payments relate to scale and efficiency 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.

The section on page 8 describes the variables, while the section on page 10 (referring to Table 1) describes the sources of those variables. We have written out the sentence on page 10 a bit more to clarify the distinction.

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.

We see the reviewer’s point but feel that there is too much variation in the significance and magnitude of the APR-DRG regression coefficients to make a blanket statement like this.

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?

We could not determine the context of this question. Does it refer to a particular sentence in the Results section?

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.

The former Appendix A is now Table 4. The first part of Appendix B is now Table 5, and the rest of Appendix B was dropped.

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.

We have replaced “encounter” with “stay” throughout and now focus on the out-of-sample predictions.

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.

We have rephrased the sentence as suggested.

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

We have done so.