

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form ([see an example](#)) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Analysis of Variation in Charges and Prices Paid for Vaginal and Cesarean Section Births: A Cross-Sectional Study
<b>AUTHORS</b>	Hsia, Renee; Akosa Antwi, Yaa; Weber, Ellerie

### VERSION 1 - REVIEW

<b>REVIEWER</b>	Prof Graham Cookson Department of Health Care Management & Policy Surrey Business School University of Surrey, Guildford, UK
<b>REVIEW RETURNED</b>	22-Oct-2013

<b>GENERAL COMMENTS</b>	<p><b>Overview</b></p> <p>I enjoyed reading this paper and I found the enormous, and apparently inexplicable, variation in pricing to be both worrying and noteworthy. Therefore I would strongly recommend this paper for publication, once my comments are satisfactorily addressed by the authors.</p> <p><b>Main comments</b></p> <p>There are two major issues with this paper that need to be addressed prior to publication. First, the paper is very light on any theoretical justification for the models fitted. The discussion includes an attempt to justify and explain the findings from the analysis but it would be better to start off with a strong theoretical model from which to base the empirical work. At the moment the paper is too descriptive and exploratory.</p> <p>The second major issue is regarding the statistical analysis. I have some comments regarding the variables. If available ethnicity should be included as an explanatory variable, as well as some measure of the capital costs or overheads of the hospitals. Ethnicity along with age tend to predict complications and additional resource use in maternities, while the overheads or capital costs should play some role in explaining hospital level charges. I also wonder whether the</p>
-------------------------	---

	<p>urban/rural distinction is too coarse and may not adequately reflect the cost structures of the hospitals, which I presume is why it's included as you already have a market/competition based measure in the model. In general, it would be better to have an explanation of the theory or hypothesis that drives the inclusion of your variables which relates to my first major comment (see last paragraph). Finally, in relation to the variables perhaps the models should use log dependent variables as prices seem heavily skewed.</p> <p>In terms of the analysis, it is not clear to me why the data are not modelling using a hierarchical or multilevel model as you have patients nested within hospitals/ This would allow you to investigate the variation within and between hospitals more explicitly, and to control for the variation in the average number of patients within trusts. At the very least I think the errors need to be clustered and interaction effects included.</p> <p><b>Detailed Comments</b></p> <p>While the English is readily understandable, the article could do with a very careful editing to find the persistent grammatical errors. For instance, there appears to be some confusion about the use of nouns and adjectives, inter alia, California and Californian, and health care and healthcare. Please find some minor but detailed comments below:</p> <p>P5, Line 1 – clarify what expenditures (I presume US).</p> <p>P5. Line 18 – clumsy English please rewrite</p> <p>P5, line 22 “seemingly inexplicable” – delete seemingly or clarify extent.</p> <p>P5, lines 32-38 sentence doesn't make sense.</p> <p>P5, line 46 “opaque idiosyncratic proprietary formulas” too verbose</p> <p>P6, line 7 “indexes” should be indices</p> <p>P6, line 27 “of” should be in</p> <p>P7, line 10 – is this full reference necessary? Consider footnote?</p> <p>P7, line 25, insert ‘a’ before hospital wage index</p> <p>P7, line 34 delete ‘of’ before 775</p>
--	--

	<p>P7, line 37 provide % of total episodes that involve a privately insured patient</p> <p>P7, line 44, how many women are excluded by excluding KP patients?</p> <p>P7, line 53 why is the patient's gender relative for births?</p> <p>P8, line 1 charitable not charity</p> <p>P8, line 1, clarify "cell size limit" comment</p> <p>P8, lines 24-7. Please explain the secondary outcome measure derivation more clearly.</p> <p>P8, line 51. Please justify why age was split into these two groups</p> <p>P9, line 51 insert 'a' before wage index</p> <p>P10, line 1, insert 'the' before degree of..</p> <p>P10, lines 24-7. Mention over how many hospitals in sample</p> <p>P13, line 32 replace "past literature" with "the extant literature" or similar.</p> <p>P13, lines 32-34. The justification for the assertion regarding 'noise' needs much more attention and detail. It seems fairly spurious.</p> <p>P13, lines 34-35. Can't you test the hypothesis about historical prices by including some measures of historical prices or by exploring time trends in the pricing data?</p> <p>P13, line 53. "Miscorrelation" What does this mean? This seems to be made up.</p>
--	--

<b>REVIEWER</b>	Christopher Tompkins Brandeis University, USA
<b>REVIEW RETURNED</b>	22-Oct-2013

<b>GENERAL COMMENTS</b>	<p>In many ways, this study is well done. The data and methods are reasonable and appropriate to address and explicate the authors' goals, including the formulation of adjusted amounts, and then regressing those amounts on variables of interest.</p> <p>The discussion section shows impressive familiarity with the context and implications of the findings. Focusing on episodes or products is a useful conceptual approach, rather than the item-by-item critiques that have been made many times. The episodes are the meaningful units for consumers or purchasers in terms of what they are getting</p>
-------------------------	--

	<p>for the stated price.</p> <p>Interesting choice of episodes, however, given that the market for birthing services is one of the few recognized successes in term of consumer-directed shopping and transparent pricing. Moreover, in the real market, the price for the delivery is not the sum of nominal prices in the chargemaster, but the result of strategic choices driven by market forces that lead to tangible results (i.e., purchases by consumers). A huge gap here is the extent to which the chargemaster dollar amounts, together with discount estimates, correlate with the real prices offered to consumers; suspecting very little.</p> <p>The storyline in the article is familiar. Many tout the importance of price transparency, and the presumed consumer-directed improvements to efficiency from making price comparisons and rewarding lower priced hospitals in the marketplace. Birthing is not very typical, being highly “elective” and very standardized basic product, with plenty of time to compare options with respect to pricing, services, perceived quality, and convenience. Most of hospital care is much less straightforward, with heterogeneity and uncertainty of need, wide-ranging indicators of quality, and finally opaque pricing for consumers who care.</p> <p>As is frequently the case, the storyline is disjointed and comes at the issues also from a different direction, namely, the practical suitability of the chargemaster to determine realistic pricing of definable products. Although once a vehicle for “realistic pricing,” it has long been the case that the chargemaster has become an internal document housing interim or working accounting figures that are driven significantly by payer-mix, service utilization by payer, and aggregate revenue requirements for the hospital. That is combined with regulatory obligations to charge all patients the same amounts, nominally, for the same services.</p> <p>So, a significant strength is the focus on products that consumers can understand and would be willing and able to shop-around for. A weakness is that the study essentially sums the admittedly arbitrary charges in order to pin a “price” on each hospital, even though those predicted prices probably bear no relation to what uninsured patients actually pay.</p>
--	---

**VERSION 1 – AUTHOR RESPONSE**

Reviewer Name Prof Graham Cookson  
 Institution and Country Department of Health Care Management & Policy  
 Surrey Business School  
 University of Surrey, Guildford, UK  
 Please state any competing interests or state ‘None declared’: No competing interests

Overview

I enjoyed reading this paper and I found the enormous, and apparently inexplicable, variation in pricing to be both worrying and noteworthy. Therefore I would strongly recommend this paper for publication, once my comments are satisfactorily addressed by the authors.

#### Main comments

1. There are two major issues with this paper that need to be addressed prior to publication. First, the paper is very light on any theoretical justification for the models fitted. The discussion includes an attempt to justify and explain the findings from the analysis but it would be better to start off with a strong theoretical model from which to base the empirical work. At the moment the paper is too descriptive and exploratory.

Thank you for this important comment. The model that underpins our paper was grounded in the economics literature and therefore we omitted it out of concern that it would not be of interest to the average, medical-oriented reader of the BMJ Open journal. To address these concerns, however, we have added a brief description of the theoretical model in the introduction, which reads:

“Our analysis is based on the theory that the variation in charges for these homogeneous episodes of care, once adjusted for patient characteristics, is potentially explained by measurable hospital and market characteristics that govern hospital charge-setting behavior.”

Further, we have included, as requested, references to the theoretical justification for the models fitted in our methods section at the top of page 9.

2. The second major issue is regarding the statistical analysis. I have some comments regarding the variables. If available ethnicity should be included as an explanatory variable, as well as some measure of the capital costs or overheads of the hospitals. Ethnicity along with age tend to predict complications and additional resource use in maternities, while the overheads or capital costs should play some role in explaining hospital level charges.

Unfortunately, hospital overhead or capital costs are not available in our data. Further, as we are using a public dataset masked for identifiers in order to prevent identification of individual patient location. Race and ethnicity is a masked variable, as it can be so predictive of a patient's identity, especially in small areas. Therefore we did not have sufficient reliable data on race or ethnicity to include it in our model.

3. I also wonder whether the urban/rural distinction is too coarse and may not adequately reflect the cost structures of the hospitals, which I presume is why it's included as you already have a market/competition based measure in the model. In general, it would be better to have an explanation of the theory or hypothesis that drives the inclusion of your variables which relates to my first major comment (see last paragraph).

See response to comment 1. We specifically include citations regarding the sources from which we drew our variables on pages 9-10.

4. Finally, in relation to the variables perhaps the models should use log dependent variables as prices seem heavily skewed.

We apologize for the misunderstanding. Indeed, we are using logged charges as the main dependent variable for precisely the reason mentioned, and have edited the following sentence to our methods in the statistical analysis section on page 9:

“For all analyses, charges are logged to account for the right skew.”

5. In terms of the analysis, it is not clear to me why the data are not modelling using a hierarchical or multilevel model as you have patients nested within hospitals/ This would allow you to investigate the variation within and between hospitals more explicitly, and to control for the variation in the average number of patients within trusts. At the very least I think the errors need to be clustered and interaction effects included.

Thank you for this comment. We have run the hierarchical model and include the results below for comparison's sake. We also discuss below why there is negligible difference between hierarchical linear model (HLM) and our two-stage model.

For clarity, we want to reiterate our methodology: In the first stage, we regress (log) charges on patient characteristics and hospital fixed effects (with errors clustered at the hospital level), and then predict the (log) charge for the average patient at each hospital. Then in the second stage, we regress each hospital's predicted (log) charge on hospital characteristics. Statistically, this two-stage model is nearly equivalent to a hierarchical model.

As seen, the two-stage model we use in our paper maps to a hierarchical model, as can be verified by comparing the HLM results with the results from the first and second stage regressions. We show the HLM results in the uploaded PDF version of this document in Response Tables 1a & 1b, which can be compared to the results of our two-stage models, which, for convenience we include in Response Tables 2a & 2b and 3a & 3b. All Response Tables can be found in the uploaded PDF version of this document, under the filename "Response to Reviewers." As you can see, the results of the two models are quantitatively very similar.

Theoretically, the reason we prefer our two-stage model over HLM is because we believe it more clearly explains the outcome we are interested in: namely, the variation in the charge for the average patient across hospitals.

We think that some confusion on this point may have stemmed from the fact that our submitted paper did not include a table showing results from our first-stage regression. To clarify this issue for the reader, we have added the results of the two first-stage regressions as supplementary tables, and have attempted to clarify our language in describing our statistical analyses.

#### Detailed Comments

While the English is readily understandable, the article could do with a very careful editing to find the persistent grammatical errors. For instance, there appears to be some confusion about the use of nouns and adjectives, inter alia, California and Californian, and health care and healthcare. Please find some minor but detailed comments below:

6. P5, Line 1 – clarify what expenditures (I presume US).

Done.

7. P5. Line 18 – clumsy English please rewrite

The sentence now reads:

“At a time when out of pocket payments for healthcare are increasing,<sup>3</sup> and growing numbers of “consumer directed” high deductible health plans put more pressure on patients to make cost-efficient healthcare decisions,<sup>4 5</sup> the opacity of healthcare pricing is increasingly concerning.”

8. P5, line 22 “seemingly inexplicable” – delete seemingly or clarify extent.

We have deleted seemingly.

9. P5, lines 32-38 sentence doesn't make sense.

We have re-written the sentence into three sentences, which now read:

“Further, charges indirectly affect healthcare pricing for all patients. Many private insurers use charges as the basis of price negotiations;<sup>7 12 13</sup> Medicare uses charges in determining inpatient diagnosis group weighting and outlier payments,<sup>14-17</sup> and also must adapt their prices to private-sector prices.<sup>18</sup> In addition, hospitals use charges in calculations of uncompensated care, which affect their non-profit status and thus many aspects of their business model, participating insurance plans, and prices.<sup>19 20</sup>”

10. P5, line 46 “opaque idiosyncratic proprietary formulas” too verbose

We have deleted opaque and added a comma between idiosyncratic and proprietary.

11. P6, line 7 “indexes” should be indices

Done.

12. P6, line 27 “of” should be in

Done.

13. P7, line 10 – is this full reference necessary? Consider footnote?

We appreciate the suggestion and have moved the sentence to a footnote.

14. P7, line 25, insert ‘a’ before hospital wage index

Done.

15. P7, line 34 delete ‘of’ before 775

Done.

16. P7, line 37 provide % of total episodes that involve a privately insured patient

Done. Privately insured patients represent 47.5% of all vaginal births and 47.1% of all cesarean sections. We have inserted these percentages in the associated paragraph at the bottom of page 7.

17. P7, line 44, how many women are excluded by excluding KP patients?

After limiting the sample to privately insured patients, excluding Kaiser patients eliminates 29.4% of vaginal births and 19.7% of cesarean sections. We have added these percentages to the paragraph in question on the top of page 8.

18. P7, line 53 why is the patient’s gender relative (sic) for births?

We excluded patients for whom gender was not recorded to ensure that our data did not include odd-appearing data (e.g., males) for whom vaginal or cesarean section would be improbable.

19. P8, line 1 charitable not charity

Done.

20. P8, line 1, clarify “cell size limit” comment

We have changed the phrase to now read:

“...charges too large to fit within the charge variable’s character limit.”

21. P8, lines 24-7. Please explain the secondary outcome measure derivation more clearly.

We have attempted to clarify our explanation. The passage now reads:

“We obtained this by multiplying the total charge billed to the patient by the hospital’s discount rate for privately insured patients. As done in previous literature, the discount factor for privately insured patients at each hospital was calculated using the formula: ...”

22. P8, line 51. Please justify why age was split into these two groups

We have added the following sentence after the delineation of our age groups:

"We chose the 18-34 and  $\geq 35$  maternal age groups because 35 years old is the cutoff at which the American Congress of Obstetricians and Gynecologists generally defines "advanced maternal age," which has been associated with increased risk of complications for both the mother and the infant."

23. P9, line 51 insert 'a' before wage index

We inserted a 'the' before wage index, as we intend it to indicate that it is a particular wage index (the one determined by the Center for Medicare and Medicaid Services (CMS), and used in their reimbursement calculations) of the corresponding hospital, and feel 'the wage index' better indicates that aim than 'a wage index.'

24. P10, line 1, insert 'the' before degree of.

Done.

25. P10, lines 24-7. Mention over how many hospitals in sample.

Done.

26. P13, line 32 replace "past literature" with "the extant literature" or similar.

We have changed "past literature" to "the existing literature."

27. P13, lines 32-34. The justification for the assertion regarding 'noise' needs much more attention and detail. It seems fairly spurious.

Thank you for the suggestion. We feel that the rest of the paragraph dives into more detail regarding why the 'noise' hypothesis for our variation is likely valid. However, we realize this initial sentence was very poorly worded and does not convey the true intention of the following paragraph. We have therefore revised it to read:

"Based on findings from the existing literature, we hypothesize that the variation we find is more likely random than due to unobservable hospital characteristics."

28. P13, lines 34-35. Can't you test the hypothesis about historical prices by including some measures of historical prices or by exploring time trends in the pricing data?

When we refer to historical prices in this paragraph, we are referring to prices preserved over the course of several decades. OSHPD data only goes back to the early 1990's, making it difficult to truly track trends in charge data over the necessary scale to explore that hypothesis.

29. P13, line 53. "Miscorrelation" What does this mean? This seems to be made up.

While miscorrelation is an English word indicating improper correlation, we agree that it was a poor word choice. We have replaced "miscorrelation" with "lack of correlation between cost and charge."

Reviewer Name Christopher Tompkins

Institution and Country Brandeis University, USA

Please state any competing interests or state 'None declared': None declared

30. In many ways, this study is well done. The data and methods are reasonable and appropriate to address and explicate the authors' goals, including the formulation of adjusted amounts, and then regressing those amounts on variables of interest. The discussion section shows impressive familiarity with the context and implications of the findings. Focusing on episodes or products is a useful conceptual approach, rather than the item-by-item critiques that have been made many times. The episodes are the meaningful units for consumers or purchasers in terms of what they are getting for the stated price.

We appreciate this reviewer's comments on focusing on episodes or products rather than service-line items, since these are the more meaningful units for patients who are potential consumers of the health care "product."

31. Interesting choice of episodes, however, given that the market for birthing services is one of the few recognized successes in term of consumer-directed shopping and transparent pricing. Moreover, in the real market, the price for the delivery is not the sum of nominal prices in the chargemaster, but the result of strategic choices driven by market forces that lead to tangible results (i.e., purchases by consumers). A huge gap here is the extent to which the chargemaster dollar amounts, together with discount estimates, correlate with the real prices offered to consumers; suspecting very little.

We agree that there is likely limited correlation between charges and the total amount a woman will pay directly for childbirth. However, we attempted to estimate the actual price private insurers would pay the hospital through the estimated discounted price. These prices also varied widely across hospitals. In addition, charges in the end do have some relation to the prices paid, as they sometimes represent starting points for discussions with insurers (or if not, at least with outlier payments), and also therefore indirectly through the premiums of privately insured individuals.

32. The storyline in the article is familiar. Many tout the importance of price transparency, and the presumed consumer-directed improvements to efficiency from making price comparisons and rewarding lower priced hospitals in the marketplace. Birthing is not very typical, being highly "elective" and very standardized basic product, with plenty of time to compare options with respect to pricing, services, perceived quality, and convenience. Most of hospital care is much less straightforward, with heterogeneity and uncertainty of need, wide-ranging indicators of quality, and finally opaque pricing for consumers who care. As is frequently the case, the storyline is disjointed and comes at the issues also from a different direction, namely, the practical suitability of the chargemaster to determine realistic pricing of definable products. Although once a vehicle for "realistic pricing," it has long been the case that the chargemaster has become an internal document housing interim or working accounting figures that are driven significantly by payer-mix, service utilization by payer, and aggregate revenue requirements for the hospital. That is combined with regulatory obligations to charge all patients the same amounts, nominally, for the same services. So, a significant strength is the focus on products that consumers can understand and would be willing and able to shop-around for. A weakness is that the study essentially sums the admittedly arbitrary charges in order to pin a "price" on each hospital, even though those predicted prices probably bear no relation to what uninsured patients actually pay.

Again we agree that charges and the associated estimated discounted prices are blunt instruments to approximate prices paid. However we also believe that charges are still a relevant concept. Though uninsured patients often receive eventual discounts or charitable care to help cover the costs of their care, they do by law have to be billed the full charges, and if delinquent, the bill collector will initially seek that charge. Further, as we mentioned in the text, understanding the wide and unexplained variability in charges across hospitals does have a bearing on the negotiated prices, Medicare payments, and non-profit status the hospital receives. Thus, while not a study of the variation in prices patients will pay, we believe our analysis still captures the variation in a relevant measure involved in healthcare pricing.

## VERSION 2 – REVIEW

<b>REVIEWER</b>	Prof Graham Cookson Department of Health Care Management & Policy, University of Surrey, UK
<b>REVIEW RETURNED</b>	20-Nov-2013

<b>GENERAL COMMENTS</b>	The authors have addressed my comments and concerns from the first review. I believe the paper is much improved and provides an interesting and worthy perspective.
-------------------------	---