

PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	A discrete choice experiment to determine facility-based delivery services desired by women and men in rural Ethiopia
AUTHORS	Beam, Nancy; Bekele Dadi, Gezehegn; Rankin, Sally; Weiss, Sandra; Cooper, Bruce; Thompson, Lisa

VERSION 1 – REVIEW

REVIEWER	Ruth Jackson Alfred Deakin Institute Deakin University Australia
REVIEW RETURNED	04-May-2017

GENERAL COMMENTS	<p>This paper will be of interest to policy makers in Ethiopia who are involved in identifying priorities for resource allocation in the health system, especially at the district (health center) level.</p> <p>Comments:</p> <p>In the Abstract under objectives it would be useful to add two or three words as to why men play an important role in increasing FB deliveries (re their role in decision-making).</p> <p>The paper should start with: The Maternal Mortality Ratio (MMR) not the MMRate which is a different measurement.</p> <p>MMR: maternal mortality ratio (maternal deaths per 100 000 live births)</p> <p>MMRate: maternal mortality rate (the number of maternal deaths divided by person-years lived by women of reproductive age)</p> <p>e: WHO, UNICEF, UNFPA, World Bank Group, United Nations Population Division, 2015. <i>Trends in Maternal Mortality: 1990 to 2015</i>. http://apps.who.int/iris/bitstream/10665/194254/1/9789241565141_eng.pdf?ua=1.</p> <p>Page 8, Line 44 mentions government efforts to improve facility based deliveries and EmONC but provides no evidence or government policy documents—perhaps refer to Federal Democratic Republic of Ethiopia Ministry of Health, 2015. <i>HSTP Health Sector Transformation Plan 2015/16 - 2019/20 (2008-2012 EFY)</i> or even the Federal Ministry of Health, 2015/16. <i>Health Sector Transformation Plan 1 Version 1 Annual Performance Report EFY 2008 (2015/16)</i>.</p> <p>4. Page 8, Line 48 refers to Peru which was unexpected. If you are going to do cross-country comparison, perhaps best to find a couple of sub-Saharan African countries to mention here as well as it might be more relevant to researchers from Africa (e.g. Rwanda, Kenya, Nigeria).</p> <p>5. Page 9, Lines 4-8. Perhaps here you could be a bit more circumspect by suggesting</p>
-------------------------	--

that policy makers and health facility staff should consider factors that underly delivery place choices that respond to family preferences given the 'coercive nature' of many government policies.

Other comments:

1. Re generalizability: I think you should consider making a point about religion as all women and 96% of men in your research were Protestants (as are most people in Sidama), this may limit generalizability as people in other areas of Ethiopia are generally Orthodox or Moslem. Also, much of Sidama has a much higher population density than other areas including some of the emerging regions such as Afar Region so distance may be less of a factor for women giving birth in some parts of Sidama than other regions.

2. There is a good argument for targeting men in research about the uptake of maternal health research as men are disproportionately involved in making household decisions (as indicated in the EDHS; and other research (references only provided for interest) such as:

Holden, S., Tefera, T., 2008. From Being Property of Men to Becoming Equal Owners? Early Impacts of Land Registration and Certification on Women in Southern Ethiopia. http://arken.umb.no/~steiho/HoldenTefera2008From_Being_Property_of_Men_to_becoming_Equal_Owners.pdf

Kumar, N., Quisumbing, A.R., 2015. Policy Reform toward Gender Equality in Ethiopia: Little by Little the Egg Begins to Walk. *World Development*, 67, 406-423. <http://dx.doi.org/10.1016/j.worlddev.2014.10.029>.

Mabsout, R., van Staveren, I., 2010. Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia. *World Development*, 38, 783-796. <http://dx.doi.org/10.1016/j.worlddev.2009.11.011>.)

Other strong influences on maternal health (not mentioned in your research) are mothers-in-law and other older women including TBAs who argue that institutional delivery care is necessary only if there are birth complications—thus leading to delays until the last moments of childbirth/delays in transportation/inability to reach a facility in time. See for example:

Gebrehiwot, T., Goicolea, I., Edin, K., San Sebastian, M., 2012. Making pragmatic choices: women's experiences of delivery care in Northern Ethiopia. *BMC Pregnancy & Childbirth*, 12, 113-123. doi: 10.1186/1471-2393-12-113.

Jackson, R., Tesfay, F., Godefay, H., Gebrehiwot, T., 2016. Health Extension Workers' and Mothers' Attitudes to Maternal Health Service Utilization and Acceptance in Adwa Woreda, Tigray Region, Ethiopia. *PLoS ONE*, 11, e0150747. [10.1371/journal.pone.0150747](https://doi.org/10.1371/journal.pone.0150747).

It is *not* surprising that women and men place a higher value on health facilities that always have medications/supplies and allow support people into the delivery room, nor that husbands are more likely to choose nearer, less expensive delivery services with friendly providers.

Limited research about the introduction of the free ambulance service shows that men are influenced by distance because of difficulties finding and paying for the return journey home and funding the cost of the wife's stay while she waits to give birth at the health center:

Jackson, R., Tesfay, F.H., Gebrehiwot, T.G., Godefay, H., 2017. Factors that hinder or enable maternal health strategies to reduce delays in rural and pastoralist areas in Ethiopia. *Tropical Medicine & International Health*, 22, 148-160. 10.1111/tmi.12818.

Godefay, H., Kinsman, J., Admasu, K., Byass, P., 2016. Can innovative ambulance transport avert pregnancy-related deaths? One-year operational assessment in Ethiopia. *Journal of Global Health*, 6:10.7189/jogh.06.010410.

Husbands do not want other men examining their wives which is a strong barrier to facility birth in some areas such as Afar Region. The government is attempting to educate men through public meetings and the introduction of the Health Development Army—which aims to empower women: 'in the government's view many husbands were holding back their wives from becoming "model women," because doing so meant going against their conceptions of a homebound, quiet woman' (Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015. Using community health workers: Discipline and Hierarchy in Ethiopia's Women's Development Army. *Annals of Anthropological Practice*, 39, 42-57. 10.1111/napa.12064.)

See also: Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015. A Women's Development Army: Narratives of Community Health Worker Investment and Empowerment in Rural Ethiopia. *Studies in Comparative International Development*, 50, 455-478. 10.1007/s12116-015-9197-z.

A further study you might consider underscores the importance of targeting men when designing interventions that are intended to raise the awareness and use of prenatal health-care services:

Biratu, B., Lindstrom, D., 2006. The influence of husbands' approval on women's use of prenatal care: Results from Yirgalem and Jimma towns, south west Ethiopia. *Ethiopian Journal of Health Development*, 20, 84-92.

3. Preference for HEWs (page 20). I would suggest that the preference for HEWs to perform deliveries is not related to their skill but directly related to the fact that they are highly appreciated by pregnant women in the community especially for their home based visits for antenatal and postnatal care, and assistance with referral. Women still largely prefer to deliver at home and research shows HEWs were largely viewed as kind, sometimes likened to family, and sometimes regarded as competent enough to manage uncomplicated deliveries. Unpublished research evaluating safe and clean delivery training for HEWs (Federal Ministry of Health) and the role of gender equality and women's empowerment from HEWs shows that women appreciated HEWs because of their cultural sensitivity, support for traditional coffee ceremonies and provision of porridge following delivery—some of these practices have been introduced at some health centers to encourage women to give birth there. That HEWs are female was reflected by comments that they care for women like mothers, sister, or other close family members or friends.

Recently, the role of HEWs in referral serving as the linkage between rural women and health facilities has been supported by the Health Development Army (or Women's Development Army).

Complaints about health centers are well known (as your research also shows), including lack of privacy, male providers, barriers to performing traditional celebratory ceremonies, lack of family support during delivery, and social pressure (not only from husbands but also mothers-in-law) to deliver at home.

For interest, some of the research about HEWS and the HDA are:

Kok, M., Kea, A.Z., Datiko, D., Broerse, J., Dieleman, M., Taegtmeier, M., et al., 2015. A qualitative assessment of health extension workers' relationships with the community and health sector in Ethiopia: opportunities for enhancing maternal health performance. *Human Resources for Health*, 13, 80.

Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015a. A Women's Development Army: Narratives of Community Health Worker Investment and Empowerment in Rural Ethiopia.

	<p>Studies in Comparative International Development, 50, 455-478. 10.1007/s12116-015-9197-z.</p> <p>Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015b. Using community health workers: Discipline and Hierarchy in Ethiopia's Women's Development Army. <i>Annals of Anthropological Practice</i>, 39, 42-57. 10.1111/napa.12064.</p> <p>Jackson, R., Kilsby, D., (under review). Unexpected consequences in a community health intervention: Insights on gender equality and women's empowerment from Health Extension Workers in rural Ethiopia.</p> <p>Jackson, R., Tesfay, F., Godefay, H., Gebrehiwot, T., 2016. Health Extension Workers' and Mothers' Attitudes to Maternal Health Service Utilization and Acceptance in Adwa Woreda, Tigray Region, Ethiopia. <i>PLoS ONE</i>, 11, e0150747. 10.1371/journal.pone.0150747.</p> <p>Jackson, R., Hailemariam, A., 2016. The Role of Health Extension Workers in Linking Pregnant Women With Health Facilities for Delivery in Rural and Pastoralist Areas of Ethiopia. <i>Ethiopian Journal of Health Science</i>, 26, 471-478. http://dx.doi.org/10.4314/ejhs.v26i5.9.</p> <p>4. I wondered why you did not differentiate between doctors, nurses and midwives as the accelerated Midwifery Training Program aims to place two midwives into each health center (perhaps this was too confusing for the participants but you should explain why you just use the term 'nurse'). In Ethiopia do nurses or midwives offer the lowest cost solution to providing skilled care? (note: 49% of nurses, 71% of midwives are female).</p> <p>5. Could the authors consider expanding the Conclusion to include more implications that would be useful for policymakers in Ethiopia, especially around the role of gender and trust in health providers (and the need to provide respectful care). Other implications could consider culturally sensitivity at birth (e.g. coffee ceremonies) and being surrounded by family and friends. Some of men's concerns around distance and cost stem from the cost of their wives staying in a health facility (with questions about who looks after the other children, cattle etc.) and costs related to returning home—especially in areas where there is little or no transportation.</p>
--	--

REVIEWER	Heather Sipsma Benedictine University, USA
REVIEW RETURNED	11-May-2017

GENERAL COMMENTS	<p>The authors present an important piece of work to understand preferences among men and women in rural Ethiopia needed to increase facility-based births. I applaud the contribution and include a few comments and recommendations to strengthen the quality of the manuscript.</p> <p>Abstract: The objective in the abstract was a bit unclear and seemingly inconsistent with the results the present in the abstract. It says “to determine differences between the facility-based delivery care attributes...” whereas the results present preferences among women and men, not necessarily the differences. I understand that in the paper, these results are unpacked further, but in the abstract, it is a bit confusing.</p> <p>Abstract: Remove the commas in the phrase, “women, who were pregnant or who had a child < 2 years old and their male partners” as this phrase is necessary to understand who the women are.</p> <p>I recommend using consistent language throughout the manuscript. For example “FB,” “facility based,” “facility-based” are used</p>
-------------------------	--

interchangeably as are “birth,” “delivery,” and “childbirth services.” Consistency of terms would improve its readability.

Instead of “a 54-study literature review” on Page 9 (line 10), I recommend rewriting this as “a literature review including 54 studies...”

In lines 13-15 of Page 9, two terms in your list are nouns and the other two are adjectives. I recommend ensuring these are parallel with one another. Maybe it should be something like “Changeable facility-level factors included cultural barriers, perceived benefits and barriers, economic barriers, and physical barriers.”

Please explain how “support of family and friends” is a barrier on Page 9, line 20. Additionally, I would combine the paragraphs describing the qualitative and quantitative findings. Furthermore, the ideas of “cultural norms,” “cultural barriers,” and “cultural factors” get a little confusing. I recommend revising this section to streamline these ideas better.

Additionally, better topic sentences for the paragraphs on Pages 9 and 10 would help link each type of barrier back to the initial list presented at the top of the Page 9.

On Page 10, line 41 says “women who had delivered in the last 5 years,” although the abstract and methods say women who had delivered in the past 2 years.

In the aims presented in the manuscript, it says “combinations of preferences” – could this be explained further? From the remainder of the paper, it appears as though these preferences are independent of one another or did I misinterpret? This idea of “combinations of preferences” comes up again on Page 15, lines 30-32. Please clarify either in the aims or in the methods/results.

Page 11, line 46: I am not convinced that provider gender and support persons represent cultural barriers. Can the authors provide additional support for this categorization?

Page 15, lines 46-48. The manuscript reads “Predictor interactions with involvement in household decision-making (Level 3) were also tested.” It sounds as though involvement in household decision-making was a Level 3 variable; if this is the case, which gender’s response for involvement in household decision-making represented the household?

Page 16, line 3, The authors state “percent deprivation greater than or equal to 33.3%, the definition for multidimensional poverty...” Is this cutpoint still appropriate given the missing indicators in the MPI?

The findings with respect to provider type is a bit confusing. The preference of provider gender and type is difficult to sort through and make meaning from. Is it possible to examine provider gender separately from provider type? The comparison between HEWs, nurses, and doctors is easier to interpret.

My last comment should be considered potential limitation of the study design. The extent to which women’s preferences matter is dependent on her decision-making power. If she does not have decision making power, and her partner makes all decisions –

	including where she should give birth – her preferences have little weight. Do these findings differ among
--	--

REVIEWER	Isabelle Bray University of the West of England UK
-----------------	--

REVIEW RETURNED	07-Dec-2017
------------------------	-------------

GENERAL COMMENTS	<p>This paper addresses an important research question and the design is appropriate to answer that question. The following minor comments are for consideration by the authors:</p> <ol style="list-style-type: none"> 1. In some places it seems results are missing from the text e.g. OR for 'doctors performed the delivery' on line 37, p5; percentages for the statements 'women were significantly younger and less educated than their husbands' on line 8, p17; and 'men had greater exposure to mass media and participated more in household decisions' on line 11, p17. 2. The Conclusion of the Abstract seems a bit strong - this research may be one step towards improving facility births, but as stated in the Background, there are other factors as well as decision-making about seeking care. The conclusion of the Discussion section is more balanced. 3. Although the paper is generally well-referenced, the first four paragraphs of the Background section are quite light on references, meaning that many statements are not supported by literature. 4. Many abbreviations are used, so an abbreviation list may be useful (if it fits within the journal style guide) but be careful to always use the term in full the first time. For example, I cannot see where 'ANC' is explained, and 'QR decomposition' is mentioned on line 13, p15, with no explanation. Please check all other abbreviations. 5. Inclusion criteria – in most places this is women who had given birth within the last 2 years, whereas line 41, p10, refers to "the last five years". 6. In the Research Design section, you explain that the survey questions were drawn from the EDHS, but lines 32-34 on p10 say that 'descriptive studies that base data collection on the EDHS limit new knowledge by asking the same questions in the same way'. 7. The sample size section seems unsatisfactory to me, partly because the required sample size was not achieved, and partly because the methods used to do the calculations are not easy to understand. What were the prior probabilities, and how were they chosen? Given the second paragraph of the 'Sample Size' section, did the sample size calculations not take into account the fact that the design was based on prior parameters? Why are the mean and median sample size so different, and which should you be aiming for? If this can be disregarded so easily, what was the point of doing the sample size calculation. 8. Although this is mentioned as a limitation, no attempt is made to quantify or even describe the likely completeness of clinic and home visit records as a sampling frame (line 53, p12). 9. Validity - if the EDHS questionnaires have been adapted for this study, are they valid? (line 20, p13). 10. Reliability – why was the back-translation done for Amharic but not Sadaminya? 11. How was pre-testing of questionnaires carried out e.g. how many people, and how selected? (line 46, p13). 12. Multilevel modelling – I am not sure why gender was included as a level (line 22, p15) given that men and women's responses were
-------------------------	---

	<p>analysed separately (line 27, p15) and are presented separately in Tables 5 and 6. In any case, I would have thought that gender would be a fixed effect in a model that used multi-levels to allow for clustering of couples within households.</p> <p>13. Please explain what RedCap is (preferably with a reference) and add a reference for Stata 14.</p> <p>14. Were any exploratory analyses carried out to check the relationship between continuous variables, such as cost of care and the outcome? Did you consider modelling this variable as categorical?</p> <p>15. 'Poverty and facility choice' on p19 – a p-value of 0.055 should not be interpreted as 'not significant', firstly because the cut-off of 0.05 is entirely arbitrary so the classification of significant/not significant should be avoided, but also because there is some evidence here that deprivation affected facility choice.</p> <p>16. Figure 2 – is it possible to start this diagram with the (approximate) size of the population from which the sample is drawn e.g. how many households are there in the study area?</p> <p>17. Lack of generalisability – differences between the final study sample and the general population of Ethiopia are well described, but there is little discussion about why the sample are different from the general population, or what efforts were made to draw a representative sample.</p>
--	---

VERSION 1 – AUTHOR RESPONSE

The authors present an important piece of work to understand preferences among men and women in rural Ethiopia needed to increase facility-based births. I applaud the contribution and include a few comments and recommendations to strengthen the quality of the manuscript.

Thank you

Abstract: The objective in the abstract was a bit unclear and seemingly inconsistent with the results presented in the abstract. It says “to determine differences between the facility-based delivery care attributes...” whereas the results present preferences among women and men, not necessarily the differences. I understand that in the paper, these results are unpacked further, but in the abstract, it is a bit confusing.

Corrected to read “This study aimed to determine the FBD care attributes preferred by women and men...”

Abstract: Remove the commas in the phrase, “women, who were pregnant or who had a child < 2 years old and their male partners” as this phrase is necessary to understand who the women are.

Done

On Page 10, line 41 says “women who had delivered in the last 5 years,” although the abstract and methods say women who had delivered in the past 2 years.

This is a description of a different study, not this study. Because this was confusing, the authors' names have been added to read “A discrete choice experiment (DCE) conducted by Kruk, Paczkowski, et al. in rural Ethiopia overcame this weakness. Women who had delivered in the last five years were...”

1. In the Abstract under objectives it would be useful to add two or three words as to why men play an important role in increasing FB deliveries (re their role in decision-making).

This now reads “Men have an important role in increasing FBD due to their decision-making power, but this is largely unexplored.”

I recommend using consistent language throughout the manuscript. For example “FB,” “facility based,” “facility-based” are used interchangeably as are “birth,” “delivery,” and “childbirth services.” Consistency of terms would improve its readability.

“facility based,” “facility-based” delivery changed to “FBD”
“birth” and “childbirth” changed to “delivery” when appropriate

Instead of “a 54-study literature review” on Page 9 (line 10), I recommend rewriting this as “a literature review including 54 studies...”

Done

In lines 13-15 of Page 9, two terms in your list are nouns and the other two are adjectives. I recommend ensuring these are parallel with one another. Maybe it should be something like “Changeable facility-level factors included cultural barriers, perceived benefits and barriers, economic barriers, and physical barriers.”

Changed to “Changeable FBD factors included cultural barriers, perceived benefits and barriers, economic accessibility, and physical accessibility.”

Please explain how “support of family and friends” is a barrier on Page 9, line 20. Additionally, I would combine the paragraphs describing the qualitative and quantitative findings. Furthermore, the ideas of “cultural norms,” “cultural barriers,” and “cultural factors” get a little confusing. I recommend revising this section to streamline these ideas better.

Changed to “facility rules limiting support from family and friends during delivery”

Additionally, better topic sentences for the paragraphs on Pages 9 and 10 would help link each type of barrier back to the initial list presented at the top of the Page 9.

Done

In the aims presented in the manuscript, it says “combinations of preferences” – could this be explained further? From the remainder of the paper, it appears as though these preferences are independent of one another or did I misinterpret? This idea of “combinations of preferences” comes up again on Page 15, lines 30-32. Please clarify either in the aims or in the methods/results. Corrected to read “Our study aims were to determine: (a) the FBD care attributes preferred by women and men, (b) whether gender differences exist in attribute preferences; and (c) whether poverty levels or household decision-making involvement are associated with facility choice.” And “Women’s and men’s responses were analyzed separately to determine the utility of specific Level 1 attributes for each group that significantly contributed to facility choice”

Page 11, line 46: I am not convinced that provider gender and support persons represent cultural barriers. Can the authors provide additional support for this categorization?

These are discussed in the 4th paragraph of the Background section and references are given.

Page 15, lines 46-48. The manuscript reads “Predictor interactions with involvement in household decision-making (Level 3) were also tested.” It sounds as though involvement in household decision-making was a Level 3 variable; if this is the case, which gender’s response for involvement in household decision-making represented the household?

This should have said Level 2. It now reads “Predictor interactions with involvement in household decision-making (Level 2) were also tested.”

Page 16, line 3, The authors state “percent deprivation greater than or equal to 33.3%, the definition for multidimensional poverty...” Is this cutpoint still appropriate given the missing indicators in the MPI?

In the analysis, the variables were re-weighted to reflect use of fewer variables. This information has been added to the 5th paragraph of the Data Management and Analysis section

The findings with respect to provider type is a bit confusing. The preference of provider gender and type is difficult to sort through and make meaning from. Is it possible to examine provider gender separately from provider type? The comparison between HEWs, nurses, and doctors is easier to interpret.

Yes, the results were not straight forward. An explanation has been added to the discussion: "In order to have a reasonable number of scenarios, provider type and gender were linked in the study design, making provider preferences difficult to interpret."

My last comment should be considered potential limitation of the study design. The extent to which women's preferences matter is dependent on her decision-making power. If she does not have decision making power, and her partner makes all decisions – including where she should give birth – her preferences have little weight. Do these findings differ among women with decision-making power compared to women without decision-making power?

No, no difference was found between women with decision-making power compared to women without decision-making power in this study. As stated in the DCE Results/Decision-making section: "While Table 4 illustrated significant differences between women and men's involvement in decision-making, decision-making involvement did not significantly influence facility choice, whether measured as none vs. any ($p=0.496$); involved in healthcare decisions for self vs. not involved ($p=0.653$); involved in healthcare decisions for self vs. not involved, women vs. men ($p=0.189$); number of decisions involved in (continuous) ($p=0.930$); or number of decisions involved in (categorical) ($p=0.133$)."

This paper addresses an important research question and the design is appropriate to answer that question. The following minor comments are for consideration by the authors:

Thank you

1. In some places it seems results are missing from the text e.g. OR for 'doctors performed the delivery' on line 37, p5; percentages for the statements 'women were significantly younger and less educated than their husbands' on line 8, p17; and 'men had greater exposure to mass media and participated more in household decisions' on line 11, p17.

We have added the results in the abstract. We summarized the results from Table 4 in the text and therefore didn't add the percentages in the text but did qualitatively discuss the results/

2. The Conclusion of the Abstract seems a bit strong - this research may be one step towards improving facility births, but as stated in the Background, there are other factors as well as decision-making about seeking care. The conclusion of the Discussion section is more balanced.

We changed the abstract conclusion to state: "Differences in women's and men's preferences may influence delivery service choices. Considering these choices is one way the Ethiopian government and health facilities may encourage FBD in rural areas."

3. Although the paper is generally well-referenced, the first four paragraphs of the Background section are quite light on references, meaning that many statements are not supported by literature.

We added several more references.

4. Many abbreviations are used, so an abbreviation list may be useful (if it fits within the journal style guide) but be careful to always use the term in full the first time. For example, I cannot see where 'ANC' is explained, and 'QR decomposition' is mentioned on line 13, p15, with no explanation. Please check all other abbreviations.

ANC corrected and an abbreviation list is provided if the journal chooses to include it.

QR decomposition is not an abbreviation. It is the name of a statistical method, which would require exceeding the word limit to explain. Reference #74 is provided that explains this method in more detail.

5. Inclusion criteria – in most places this is women who had given birth within the last 2 years, whereas line 41, p10, refers to “the last five years”.

This refers to a different study. Since this was confusing, the authors’ names have been included

6. In the Research Design section, you explain that the survey questions were drawn from the EDHS, but lines 32-34 on p10 say that ‘descriptive studies that base data collection on the EDHS limit new knowledge by asking the same questions in the same way’.

Demographic data was collected using questions from the EDHS, but DCE methodology was used to collect data on preferences

7. The sample size section seems unsatisfactory to me, partly because the required sample size was not achieved, and partly because the methods used to do the calculations are not easy to understand. What were the prior probabilities, and how were they chosen? Given the second paragraph of the ‘Sample Size’ section, did the sample size calculations not take into account the fact that the design was based on prior parameters? Why are the mean and median sample size so different, and which should you be aiming for? If this can be disregarded so easily, what was the point of doing the sample size calculation.

The following explanation has been added:

“Examining the equation for sample size provides an explanation for the wide range:

$$N_k = (T_{k2} * \text{sek}2) / \text{betak}$$

where N_k is the sample size, T_{k2} is the t-ratio required for significance, $\text{sek}2$ is the standard error for the prior parameter, and betak is the prior parameter. Therefore, as β approaches zero, the sample size needed to detect statistical difference increases.

Several of the priors range from -1 to 1, reflecting the degree of uncertainty in the priors, which in turn results in a large sample size requirement (J. Rose, personal communication, August 18, 2015).”

Prior parameters were chosen based on a review of the literature. A reference to the dissertation that this paper is based on has been added (#56), which provides further details.

8. Although this is mentioned as a limitation, no attempt is made to quantify or even describe the likely completeness of clinic and home visit records as a sampling frame (line 53, p12).

Expanded to read: “The household list used to select participants came from paper-based registers and patient charts, which made identifying eligible participants difficult. Families who lived near health posts or attended clinic may have been over-represented. Although the health workers were expected to visit every home, staffing limitations make this difficult to accomplish. This may limit generalizability.”

9. Validity - if the EDHS questionnaires have been adapted for this study, are they valid? (line 20, p13).

Changed to read: “Questions from the EDHS were used for this study, thus building upon EDHS’ strong validity.”

10. Reliability – why was the back-translation done for Amharic but not Sadaminya?

Back-translation was done for Sidaminya. A deletion was made during editing. A correction has been made.

11. How was pre-testing of questionnaires carried out e.g. how many people, and how selected? (line 46, p13).

Added “In addition to pre-testing with male and female community members that took place during the translation and testing of the DCE pictures, the entire instrument was pre-tested during a day of field-testing. Pre-testing was conducted at households that had not been selected as part of the sample. Approximately twelve men and twelve women participated in pre-testing. “

12. Multilevel modelling – I am not sure why gender was included as a level (line 22, p15) given that men and women's responses were analysed separately (line 27, p15) and are presented separately in Tables 5 and 6. In any case, I would have thought that gender would be a fixed effect in a model that used multi-levels to allow for clustering of couples within households.

Added for clarification: “The analysis was conducted in four parts, which are described in more detail below. First, separate multi-variate analyses of women's and men's data was conducted to determine their preferences (Table 5 and 6). Second, the data was combined, and gender was introduced as a Level 2 variable to determine whether a statistical difference existed between women and men's preferences. Third, a Level 2 analysis of various decision-making measures were tested to determine their effect on facility choice. Finally, the effect of household poverty on preferences was tested in a Level 3 analysis.”

13. Please explain what RedCap is (preferably with a reference) and add a reference for Stata 14.
References added

14. Were any exploratory analyses carried out to check the relationship between continuous variables, such as cost of care and the outcome? Did you consider modelling this variable as categorical?

Cost can be modelled as either continuous/linear or categorical. We treated these as continuous/linear for three reasons: First, we were interested in marginal willingness to pay estimates, which involves taking the ratio of the derivatives towards the x's. We wanted to know how much change in attribute k will be compensated by a change in the price attribute that results in a zero change in utility. Derivatives assume continuous variables – mathematically, it involves an infinitesimal change in the function. If you dummy code, you must take a unit change, not an infinitesimal change. Second, you can test for non-linearity by squaring, exponentiating, logging, etc. the variables so even though you may treat it a continuous, you don't need to assume that it is linear. Along these lines, it is more parsimonious in terms of number of parameters estimated. Thirdly, if you have interaction effects, it is easier to handle, as you don't have to go through every level by level combination.

15. 'Poverty and facility choice' on p19 – a p-value of 0.055 should not be interpreted as 'not significant', firstly because the cut-off of 0.05 is entirely arbitrary so the classification of significant/not significant should be avoided, but also because there is some evidence here that deprivation affected facility choice.

We changed the wording as follows: “Facility choice did not differ between multidimensionally poor and not multidimensionally poor households ($p=0.170$), but facility choice was associated weakly with percent household deprivation ($p = 0.055$).”

16. Figure 2 – is it possible to start this diagram with the (approximate) size of the population from which the sample is drawn e.g. how many households are there in the study area?

Added info on # of women in study area

17. Lack of generalisability – differences between the final study sample and the general population of Ethiopia are well described, but there is little discussion about why the sample are different from the general population, or what efforts were made to draw a representative sample.

In discussion, we added discussion of religion and population density and differences from general population. We also added a description about limitations of clinic records used to identify our sample.

2. The paper should start with: The Maternal Mortality Ratio (MMR) not the MMRate which is a different measurement.

MMR: maternal mortality ratio (maternal deaths per 100 000 live births) MMRate: maternal mortality rate (the number of maternal deaths divided by person-years lived by women of reproductive age).

See: WHO, UNICEF, UNFPA, World Bank Group, United Nations Population Division, 2015. Trends in Maternal Mortality: 1990 to 2015.

http://apps.who.int/iris/bitstream/10665/194254/1/9789241565141_eng.pdf?ua=1.

Corrected.

3. Page 8, Line 44 mentions government efforts to improve facility based deliveries and EmONC but provides no evidence or government policy documents—perhaps refer to Federal Democratic Republic of Ethiopia Ministry of Health, 2015. HSTP Health Sector Transformation Plan 2015/16 - 2019/20 (2008-2012 EFY) or even the Federal Ministry of Health, 2015/16. Health Sector Transformation Plan 1 Version 1 Annual Performance Report EFY 2008 (2015/16).

Done. Thank you.

4. Page 8, Line 48 refers to Peru which was unexpected. If you are going to do cross-country comparison, perhaps best to find a couple of sub-Saharan African countries to mention here as well as it might be more relevant to researchers from Africa (e.g. Rwanda, Kenya, Nigeria).

Added the following: “Kenya’s program to increase dialogue between communities and health services increased FBD in the rural community by 6.1%.¹⁰ Community mobilization increased FBD by 30% in Burkina Faso.¹¹ No studies were found that tested community directed facility-based interventions to improve FBD in Ethiopia.

5. Page 9, Lines 4-8. Perhaps here you could be a bit more circumspect by suggesting that policy makers and health facility staff should consider factors that underly delivery place choices that respond to family preferences given the 'coercive nature' of many government policies.

Changed to “An increased understanding of factors underlying delivery place choice may help Ethiopian health facilities to better respond to families’ preferences.”

Other comments:

1. Re generalizability: I think you should consider making a point about religion as all women and 96% of men in your research were Protestants (as are most people in Sidama), this may limit generalizability as people in other areas of Ethiopia are generally Orthodox or Moslem. Also, much of Sidama has a much higher population density than other areas including some of the emerging regions such as Afar Region so distance may be less of a factor for women giving birth in some parts of Sidama than other regions.

Added.

2. There is a good argument for targeting men in research about the uptake of maternal health research as men are disproportionately involved in making household decisions (as indicated in the EDHS; and other research (references only provided for interest) such as:

Holden, S., Tefera, T., 2008. From Being Property of Men to Becoming Equal Owners? Early Impacts of Land Registration and Certification on Women in Southern Ethiopia.

http://arken.umb.no/~steiho/HoldenTefera2008From_Being_Property_of_Men_to_becoming_Equal_Owners.pdf

Kumar, N., Quisumbing, A.R., 2015. Policy Reform toward Gender Equality in Ethiopia: Little by Little the Egg Begins to Walk. *World Development*, 67, 406-423.

<http://dx.doi.org/10.1016/j.worlddev.2014.10.029>

Mabsout, R., van Staveren, I., 2010. Disentangling Bargaining Power from Individual and Household Level to Institutions: Evidence on Women's Position in Ethiopia. *World Development*, 38, 783-796. <http://dx.doi.org/10.1016/j.worlddev.2009.11.011> .)

Thank you. Several of these have been added as references

Other strong influences on maternal health (not mentioned in your research) are mothers-in-law and other older women including TBAs who argue that institutional delivery care is necessary only if there are birth complications—thus leading to delays until the last moments of childbirth/delays in transportation/inability to reach a facility in time. See for example:

Gebrehiwot, T., Goicolea, I., Edin, K., San Sebastian, M., 2012. Making pragmatic choices: women's experiences of delivery care in Northern Ethiopia. *BMC Pregnancy & Childbirth*, 12, 113-123. doi: 10.1186/1471-2393-12-113.

Jackson, R., Tesfay, F., Godefay, H., Gebrehiwot, T., 2016. Health Extension Workers' and Mothers' Attitudes to Maternal Health Service Utilization and Acceptance in Adwa Woreda, Tigray Region, Ethiopia. *PLoS ONE*, 11, e0150747. 10.1371/journal.pone.0150747.

We did not include mothers-in-law in the study due to the added complexity of doing so, and the limitations of time and funding, although it would have been quite interesting. This has been added to the limitations section.

It is not surprising that women and men place a higher value on health facilities that always have medications/supplies and allow support people into the delivery room, nor that husbands are more likely to choose nearer, less expensive delivery services with friendly providers. Limited research about the introduction of the free ambulance service shows that men are influenced by distance because of difficulties finding and paying for the return journey home and funding the cost of the wife's stay while she waits to give birth at the health center:

Jackson, R., Tesfay, F.H., Gebrehiwot, T.G., Godefay, H., 2017. Factors that hinder or enable maternal health strategies to reduce delays in rural and pastoralist areas in Ethiopia. *Tropical Medicine & International Health*, 22, 148-160. 10.1111/tmi.12818.

Godefay, H., Kinsman, J., Admasu, K., Byass, P., 2016. Can innovative ambulance transport avert pregnancy-related deaths? One-year operational assessment in Ethiopia. *Journal of Global Health*, 610.7189/jogh.06.010410.

At the community meeting held after data collection was complete, the men primarily wanted to discuss the difficulties in accessing ambulance services in a timely manner. They used the forum to ask government representatives to provide more ambulances, especially for outlying areas.

This has been added to the Discussion section

Husbands do not want other men examining their wives which is a strong barrier to facility birth in some areas such as Afar Region. The government is attempting to educate men through public meetings and the introduction of the Health Development Army—which aims to empower women: "in the government's view many husbands were holding back their wives from becoming "model women," because doing so meant going against their conceptions of a homebound, quiet woman" (Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015. Using community health workers: Discipline and Hierarchy in Ethiopia's Women's Development Army. *Annals of Anthropological Practice*, 39, 42-57. 10.1111/napa.12064.)

See also: Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015. A Women's Development Army: Narratives of Community Health Worker Investment and Empowerment in Rural Ethiopia. *Studies in Comparative International Development*, 50, 455-478. 10.1007/s12116-015-9197-z.

Thank you.

A further study you might consider underscores the importance of targeting men when designing interventions that are intended to raise the awareness and use of prenatal health-care services:

Biratu, B., Lindstrom, D., 2006. The influence of husbands' approval on women's use of prenatal care:

Results from Yirgalem and Jimma towns, south west Ethiopia. *Ethiopian Journal of Health Development*, 20, 84-92.

Thank you. Added as reference.

3. Preference for HEWs (page 20). I would suggest that the preference for HEWs to perform deliveries is not related to their skill but directly related to the fact that they are highly appreciated by pregnant women in the community especially for their home-based visits for antenatal and postnatal care, and assistance with referral. Women still largely prefer to deliver at home and research shows HEWs were largely viewed as kind, sometimes likened to family, and sometimes regarded as competent enough to manage uncomplicated deliveries. Unpublished research evaluating safe and clean delivery training for HEWs (Federal Ministry of Health) and the role of gender equality and women's empowerment from HEWs shows that women appreciated HEWs because of their cultural sensitivity, support for traditional coffee ceremonies and provision of porridge following delivery—some of these practices have been introduced at some health centers to encourage women to give birth there. That HEWs are female was reflected by comments that they care for women like mothers, sister, or other close family members or friends. Recently, the role of HEWs in referral serving as the linkage between rural women and health facilities has been supported by the Health Development Army (or Women's Development Army). Complaints about health centers are well known (as your research also shows), including lack of privacy, male providers, barriers to performing traditional celebratory ceremonies, lack of family support during delivery, and social pressure (not only from husbands but also mothers-in-law) to deliver at home.

Yes, the cultural sensitivity and kindness of HEWs is key. I'm excited to hear that cultural birth practices are beginning to be allowed at health centers.

References have been added to strengthen the section on HEWs

For interest, some of the research about HEWS and the HDA are:

Kok, M., Kea, A.Z., Datiko, D., Broerse, J., Dieleman, M., Taegtmeier, M., et al., 2015. A qualitative assessment of health extension workers' relationships with the community and health sector in Ethiopia: opportunities for enhancing maternal health performance. *Human Resources for Health*, 13, 80.

Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015a. A Women's Development Army: Narratives of Community Health Worker Investment and Empowerment in Rural Ethiopia. *Studies in Comparative International Development*, 50, 455-478. 10.1007/s12116-015-9197-z.

Maes, K., Closser, S., Vorel, E., Tesfaye, Y., 2015b. Using community health workers: Discipline and Hierarchy in Ethiopia's Women's Development Army. *Annals of Anthropological Practice*, 39, 42-57. 10.1111/napa.12064.

Jackson, R., Kilsby, D., (under review). Unexpected consequences in a community health intervention: Insights on gender equality and women's empowerment from Health Extension Workers in rural Ethiopia.

Jackson, R., Tesfay, F., Godefay, H., Gebrehiwot, T., 2016. Health Extension Workers' and Mothers' Attitudes to Maternal Health Service Utilization and Acceptance in Adwa Woreda, Tigray Region, Ethiopia. *PLoS ONE*, 11, e0150747. 10.1371/journal.pone.0150747.

Jackson, R., Hailemariam, A., 2016. The Role of Health Extension Workers in Linking Pregnant Women With Health Facilities for Delivery in Rural and Pastoralist Areas of Ethiopia. *Ethiopian Journal of Health Science*, 26, 471-478. <http://dx.doi.org/10.4314/ejhs.v26i5.9>.

4. I wondered why you did not differentiate between doctors, nurses and midwives as the accelerated Midwifery Training Program aims to place two midwives into each health center (perhaps this was too confusing for the participants but you should explain why you just use the term 'nurse'). In Ethiopia do nurses or midwives offer the lowest cost solution to providing skilled care? (note: 49% of nurses, 71% of midwives are female).

Ethiopian colleagues advised that patients mostly did not understand the difference between nurses and midwives in this rural area, so for the sake of clarity, they were all called nurses. Clarification added in Table 1.

5. Could the authors consider expanding the Conclusion to include more implications that would be useful for policymakers in Ethiopia, especially around the role of gender and trust in health providers (and the need to provide respectful care). Other implications could consider culturally sensitivity at birth (e.g. coffee ceremonies) and being surrounded by family and friends. Some of men's concerns around distance and cost stem from the cost of their wives staying in a health facility (with questions about who looks after the other children, cattle etc.) and costs related to returning home—especially in areas where there is little or no transportation.

Implications and references have been added to the Discussion section.

VERSION 2 – REVIEW

REVIEWER	Ruth Jackson Alfred Deakin Institute for Citizenship and Globalisation, Deakin University, Australia
REVIEW RETURNED	19-Jan-2018

GENERAL COMMENTS	<p>I only have a few minor comments. On p 7 line 35, 'In Peru'. I suggest the sentence should start with something like 'Other research from...' or 'Research in...' or 'Studies of facility-based services in low- and middle-income countries' or similar as there is no direct link between Ethiopia (in sub-Saharan Africa) and Peru.</p> <p>P 20 Line 47. preference for HEWs is 'surprising'. While your next sentence qualifies the word 'surprising', is it really a 'surprise' that women may prefer to have someone they know in attendance as well as HEWs being more accommodating of cultural practices? Surprise to whom? Perhaps just think again about using the word 'surprise' as you then go on to suggest the importance of trust in providers who understand the cultural context...</p> <p>p 41 Table 1 in addition to the clarification of nurses/midwives could you add that was because patients mostly did not understand the difference between nurses and midwives in this rural area, so all were called nurses.</p> <p>p 23 line 45 Conclusion Can you consider adding two or three more words to the final sentence about the cultural preference for women to be surrounded by family and friends and how excluding support persons from the delivery room is incompatible with cultural norms and is likely to decrease FBD uptake.</p>
-------------------------	--

REVIEWER	Isabelle Bray UWE, UK
REVIEW RETURNED	29-Jan-2018

GENERAL COMMENTS	Thank you for addressing reviewers comments.
-------------------------	--

REVIEWER	Heather Sipsma Benedictine University, USA
REVIEW RETURNED	31-Jan-2018

GENERAL COMMENTS

The authors were very responsive to my comments and suggestions. Thank you.

VERSION 2 – AUTHOR RESPONSE

On p 7 line 35, 'In Peru'. I suggest the sentence should start with something like 'Other research from...' or 'Research in...' or 'Studies of facility-based services in low- and middle-income countries' or similar as there is no direct link between Ethiopia (in sub-Saharan Africa) and Peru.
Changed to “In a setting where home deliveries were also common, ...”

P 20 Line 47. preference for HEWs is 'surprising'. While your next sentence qualifies the word 'surprising', is it really a 'surprise' that women may prefer to have someone they know in attendance as well as HEWs being more accommodating of cultural practices? Surprise to whom? Perhaps just think again about using the word 'surprise' as you then go on to suggest the importance of trust in providers who understand the cultural context...

Changed to: “While appreciation of skilled providers is not uncommon, 15,16,18,21,22,36 HEWs’ ability to perform safe deliveries has been questioned.8,14,20,21 Preference for HEWs may reflect the desire to be delivered by someone they know, or greater flexibility by HEWs in accommodating cultural birth practices⁸⁵ such as allowing support persons to be present in the delivery room.”

p 41 Table 1 in addition to the clarification of nurses/midwives could you add that was because patients mostly did not understand the difference between nurses and midwives in this rural area, so all were called nurses.

Changed to: “a Nurse was used to indicate both nurses and midwives on the advice of Ethiopian staff as patients generally did not understand the difference between nurses and midwives.”

p 23 line 45 Conclusion Can you consider adding two or three more words to the final sentence about the cultural preference for women to be surrounded by family and friends and how excluding support persons from the delivery room is incompatible with cultural norms and is likely to decrease FBD uptake.

Added: “Facilities that respond to these preferences for higher quality and culturally appropriate care may increase FBD uptake.”