Author's response to reviews

Title: Job Satisfaction and Motivation of Health Workers in Different Settings in India

Authors:

David H Peters (dpeters@jhsph.edu)
Subrata Chakraborty (sc@jiml.ac.in)
Prasanta Mahapatra (pmahapat@ihs.org.in)
Laura Steinhardt (lsteinha@jhsph.edu)

Version: 2 Date: 28 April 2010

Author's response to reviews: see over
Thank you for our chance to respond to the reviewers’ feedback, which we found very thoughtful and constructive. Our responses are shown to each reviewer’s comments below in italics, and incorporated into the manuscript showing “track changes” where appropriate.

We hope you will find the responses satisfactory, and look forward to your reply.

Reviewer 1 (Shiv Dutt Gupta):

**Reviewer’s report:**

Most Compulsory Revisions
None

Minor Essential Revisions Objectives

Objectives

The objectives of the study are though implicit and understood, but not explicitly described.

*We have revised the manuscript to be explicit about the objectives: “to identify important aspects of health worker satisfaction and motivation in two Indian states working in public and private sectors.”*

**Methodology**

1. What were the criteria for calculating sample size?

*The study was part of a larger research program conducted by Indian institutions to examine options for India’s future health systems, with AP and UP representing distinct regions of the country [ref 12], so the sample sizes were not determined by the objectives of this study. We note that the number of respondents is larger than most found in the literature, we did not suffer from under-powering to find statistically significant results. We’ve also revised the description of the sampling methods in the paper.*

2. Purpose of stratifying districts on the basis of socioeconomic status is not clear. Is it associated with job satisfaction and or motivation? Performance of the districts would have been more appropriate as it is likely to be associated with job satisfaction and motivation levels.

*The purpose of stratification was to ensure that each major SES region of the states was included as part of the larger assessment of health services across the state, but was not expected to be directly related to job satisfaction/motivation. At the time of sampling, we did not have good measures of district performance.*

Discretionary Revisions

Methodology

1. What approach was taken to ensure quality and reliability of data collection?
This is an important issue as more than one organization was involved in the study. The authors mention about use of a common protocol, but whether the data collection personnel were given training to use the protocols.

There is a difficulty in having the research conducted by three different institutions with slightly different study objectives, but we used standardized procedures to minimize error. We tried to minimize differences by using study protocols that jointly developed, pre-tested and revised before implementation. Survey teams that had experience in social sciences and interviewing skills were selected, trained, and tested in pilot studies outside the study sites, and provided with survey manuals and definitions. Multiple visits to health care facilities were frequently required to complete the larger questionnaires for the study, although the health provider interviews were generally conducted in one day. Standardized supervision protocols were followed, followed by double-entry data and reconciliation. We’ve added more details to the text.

2. It is not clear whether the doctors and health workers who abstained on the day of survey (abstention from duty station is quite frequent in UP particularly), were included/excluded. If they are not included, the possibility of selection bias cannot be ruled out. It is quite possible that those who were motivated or had satisfaction with their jobs were present and completed the survey.

Only providers known to be usually present on the day(s) of the first survey visit were included in the sample, unless they were on formal leave. If necessary, multiple visits were made to each facility to complete the survey, with follow up visits scheduled at the convenience of the health facility. We tried to minimize the risks by giving advance notice of the study team visit to the health facilities involved. We obtained a 100% response rate from staff who usually attend the facility, but were unable to include those staff who might have been assigned to the facility but never, or only rarely showed up. We agree that some degree of selection bias cannot be ruled out, particularly in the public sector, where we could not find reliable lists of people who were supposed to be on duty. However, we believe we have a good sample of those providers who usually work at the health facility.

3. It is important to address the question that who (doctors) decide to go to the private sector. May be those who went to private sector, had such perception of the public sector priory.

We agree. We tried to address this by identifying the differences in perceptions of ideal job characteristics, which were different between public and private providers.

Analysis
1. A separate analysis of general medical doctors and specialist doctors, and doctors and nurses may be attempted, particularly in the public sector. The attributes of satisfaction and motivation in these categories may differ in their rankings of importance. Further, the nurses in district hospitals and primary health centres (Auxiliary Nurse Midwife) may have different perceptions (our own unpublished have shown).
We agree that these analysis would be good to conduct, but did not have the data available at all sites, and did not have space in the article for more detailed breakdowns. We agree that different professions and types of facility are likely to have different results. For nurses, for example, the ideal ratings of time for family life; good income; a superior who recognized their work; and not having to pay bribes were more important compared to the ideal for doctors. Regarding the actual presence of conditions, nurses rated as significantly more present than doctors all conditions except: time challenging work (marginally more present, p=0.06); independence from interference (no difference); keeping job as long as they want (significantly less present than for doctors); not having work influenced by political decisions (no difference); trusted by clients (marginally higher among nurses, p=0.07); knowing what you're expected to do (no difference) ; and not having to pay bribes to get what you want (only marginally more present for nurses, p=0.06).

Similarly, satisfaction with actual/current conditions tended to be lower among those working at small and big hospitals, compared to PHCs (significantly lower than PHCs for time for family life (big hospitals); training opportunities (small hospitals); good income (big hospitals); keeping job as long as you want (big hospitals). Slightly higher than PHCs were independence from interference (big hospitals at p=0.056).

2. In the methodology section, the authors have mentioned about logistic regression but the analysis of the same is not presented.

We expand the description of the methods, and make clear that the analysis of statistical significance is based on those models.

3. Duration in job also influences the satisfaction. The authors may like to check it.

We agree, but we did not collect the variable, but are limited to analyzing the data by age.

Reviewer 2 (Andrew Mitchell)

Major Compulsory revisions
1. Please provide detail on response rates in terms of the sampling frame. That is, were any providers who were randomly selected from either the public or private lists not available for/refused to participate in the survey? The sentence in the Discussion: “In AP and UP, it may also be that those workers for whom remuneration is most important may be more likely to be absent from work in order to pursue other sources of income, particularly in the public sector” suggests that there may have been some number of respondents randomly sampled but unavailable for interview (and, methodologically, what happened if someone was unavailable?). It’s unclear whether the reader should be concerned that the final sample suffers from selection bias or whether that is not likely to be affecting results. If, in fact, response/respondent availability rates were particularly
low by any of the provider characteristics highlighted in the study (e.g., AP vs UP, public vs. private, by type of facility), please also comment on how that may affect/bias results.

Only 4 of the private facilities refused to participate (1.2%); all other sampled facilities participated. We had 100% participation of health care providers sampled for the survey. As described above, the sampling frame was based on health workers usually present at the time of the survey (minus those on leave), and where advance notice was given about the study. When staff who are usually present were known to be absent, repeat visits were made to complete the surveys. In public facilities, over 90% of the data from the health facilities surveys were obtained in two visits, with the remainder requiring three or more visits. In private facilities, about 20% of health facilities required three or more visits to complete the surveys (Estimates based on AP data – UP data on number of visits were not retrieved due to changes in data personnel). We are unable to assess the number of staff who are not usually present, but who should have been assigned to the facility but had never or rarely showed up. We are also unable to assess how they would influence the results. Based on informal discussions and past experience, we suspect that in some cases in the public sector, particularly in UP, staff have been assigned to work at a particular health facility but have never or only rarely shown up. However, we were unable to quantify this. Given that advance notice was given and multiple visits that were made, we believe that we had a good sample of those who regularly show up for work. As noted above, we agree that some degree of selection bias cannot be ruled out, particularly in the public sector, where we could not find reliable lists of people who were supposed to be on duty.

2. Please provide detail JDI instrument completion rate (i.e., were all questions answered by all respondents?) and how authors treated missing data, if any.

Missingness on the survey ranged from 0.3% to 3.5% per item. We believe that this level of missing values is small and unlikely to substantially change results. Missing values were dropped during the analysis. We’ve added this information to the text.

3. Please provide slightly more detail on the logistic regression methodology described in the Methods section as well as how Tables 3 – 5 relate to those models. In terms of detail on the logistic regression methodology, a simple presentation of the estimating equation should be sufficient (I take it that a dummy variable was included to represent public or private providers with the other group omitted?).

We’ve changed the text as follows:

Logistic regression models were used for each binary outcome variable to test the significance of differences in the public vs. private sectors for each state, adjusting for sex, job, and age category using dummy indicator variables in the following equation:

\[
\log\left( \frac{P_j(Job \ Characteristic \ Present)}{1 - P_j(Job \ Characteristic \ Present)} \right) = B_0 + B_1 \text{Public/Private} + B_2 \text{Sex} + B_3 \text{Job} + B_4 \text{Age(CATEGORY)}
\]
In terms of how Tables 3 – 5 relate to the logistic models, do “significant differences” refer to the coefficients from the estimating equation?

Yes – we’ve revised the text to make this clear

Minor essential revisions
1. Please clarify construction of the dichotomous discordant variable: does “Important” refer to categories 3 and 4 while “not present” refer to categories 0 – 2? The preceding sentence beginning “We conducted exploratory analysis…” appears to imply this but it’s not entirely clear.

This is correct – and we’ve revised the text to be more explicit: “We conducted exploratory analysis analyzing the Likert scales as continuous, ordinal, and binary variables (created from the highest two outcomes – (3) & (4) – as “Important”/”Present” compared with the low and neutral responses – (0) to (2) – as “Not Important”/”Not Present”).”

Discretionary Revisions
1. Please briefly state the empirical reliability of the Job Descriptive Index (JDI) (e.g., provide a Cronbach’s alpha coefficient) both in its previous/original applications and for the current study. Does this reliability differ by State?

The JDI had average Cronbach alpha of 0.88 when initially reported, though we were able to find reliability coefficients for the JDI ranging from 0.75 to 0.92 in the US, and 0.64 to .84 in other countries. In our study, Cronbach’s alpha, by state, for the ideal and actual scales, separately, by state, are reported as:

17-ITEM IDEAL SCALE: 0.754 FOR AP, 0.712 FOR UP, 0.760 COMBINED
17-ITEM ACTUAL SCALE: 0.802 FOR AP, 0.791 FOR UP, 0.792 COMBINED

We’ve added some of this information to the text.

2. It would be interesting to know if PCAs stratified by public/private sector providers yielded results similar to those of the pooled analysis presented in the report or exhibit any particularly interesting differences in terms of factors/factor loadings.

The results were very similar. A fifth component with an eigenvalue > 1 emerged when the sample was stratified by state and sector for all groups except UP Public Sector. When the components were constrained to four for each group, they appeared similar across the groups. There were some slight variations among these four components by state and public/private, but no discernable patterns emerged. Results are available from the authors upon request.

3. Was there a theoretical motivation for choosing orthogonal rotation for the
factors? If not, are there any qualitative changes to results if an oblique rotation is
used instead?

There was not a theoretical motivation for the orthogonal rotation of factors, merely that
it provided the most clear distinction of factors. With other rotations, we did not identify
any qualitative changes.

4. Table 1 suggests that inherent differences are likely to exist between public
and private providers in each State, particularly in terms of age (i.e., private
providers tend to be younger) and position (i.e., private providers more likely to
be nurses). While the logistic regression models account for these kind of
observable differences, there may be remaining unobservable differences that
are not entirely correlated with the observable differences affecting results. Were
any matching techniques considered to explore that possibility (e.g., propensity
score matching)? If not, this reviewer suggests exploring that possibility so the
authors can be reassured about the robustness of their findings.

We agree that unobserved factors may affect results, but data limitations prevent us
from further analysis to assess further differences, which we agree would be helpful.
We also used robust estimators to reduce bias due to clustering. We believe that the
findings are robust, but have added text on this in the limitations of the study.
Reviewer 3 (Freddie Ssengooba)

1. Minor Essential Revisions

I). Methods used:
   In general the research methods used are relevant to the study question. However there are issues that need to be addressed or clarified:
   a) There is discordance between the sampling approach for the study units and the results reported in this work. In particular the analysis only presents findings on two dimensions ie by states (AP and UP) and public/private dimension. The implication of aggregating the findings need to be clarified especially since the sampling logic indicate that the research process anticipated differences related among these categorization.

   We agree, and this occurred because the study was part of a larger set of studies done at different times by different institutions, with somewhat different primary objectives. We've revised the text to explain this in a way we hope does not complicate the paper. This is also why we analyze the separately for examining differences. The Principal Components analysis was pooled because we did not find substantial differences from doing it separately.

   b) The facility sample for UP was very large (>x3) relative to AP but the actual response show that UP is only about twice as much. Authors should clarify this disparity and consider the option of weighted analysis or similar measures to correct this problem in the comparison.

   We've changed the text to clarify the differences in sampling procedures. Weighting is a good idea, but we found no differences when conducting weighted analysis.

   c) The use of a job description index (JDI) in this study may be justified due to its psychometric properties that are well documents. It is not clear how the adaptations made on the tool may have affected its documented properties. The research team may need to address the limitations related to tool modification.

   We agree, and also believe that applying a tool in a different context also affect its properties. As described above, we now report the reliability characteristics of the instrument found in our study.

   d) The ordinal scales used in the tools indicate a major overlap of response scales. For instance response categories such as “to a little extent” and “to some extent” may not be different to a native English speaker. Likewise, response category “to a large extent” and “fully present” may not be distinguishable. Although a “neutral” response category is mentioned on top of page 10, there was no evidence of such a category in the description on page 9. This may not affect the results since these similar categories are aggregated but it illustrates the weakness in the measurement instrument that need to be acknowledged.
We’ve added this consideration to the text on limitations of the study.

e) Theoretic consistence will be required. For instance two items were added to
the questionnaire to measure “corruption”. In table 2 these two items load on a
factor that is referred to as “transparency”, the conclusion section the same factor
is called corruption (page 21). Likewise the authors may need to make a
distinction between job satisfaction and motivation since these are distinct in the
theory upon which this work is grounded.

We agree, and note that the corruption/transparency dimension affects was added
because the formative work for the study identified this dimension as affecting
satisfaction and motivation in AP and UP.

2. Conclusion section:
The conclusion especially about the need for the public sector to copy strategies
of the private sector will need revisiting given the findings and background
information about the roles played by the different sectors. The authors need to
reconcile the information provided about the inadequacy of regulation and week
quality in the private sector with this recommendation. It would be useful to state
what each state can learn from each other. If the findings were disaggregated by
small and big providers or doctor/nurses additional lessons would probably be
unmasked. I wish to advise the authors to reflect on the foundamental roles for
the public sector as they craft recommendations from this study. Would the public
sector wish to copy wholesale the autonomy of private health workers to decide
where they whant to work? such recommendations require the authors to fully
appreciate the feasible implications of their findings.

One of the main points of the conclusion is “that understanding motivation and
satisfaction of health workers is highly dependent on the local context. Perhaps one of
the most important implications of this is that health managers ought to be asking their
own workers about their particular motivational factors, and developing plans locally to
address them.” We agree that the public sector should not and cannot copy wholesale
a strategy of autonomy from the private sector, but we do argue that “one strategy may
be to find ways to learn from the private sector”. In the example of autonomy of
choosing where to work, at the time of the survey, there was very little scope for public
sector employees to choose where they could work. We believe that strategies could
be developed to provide more choice over work location, and/or to provide better
incentives to work in less desirable locations. We’ve added text to this effect in the
paper.