AJAE appendix for ‘Defining Access to Health Care: Evidence on the Importance of Quality and Distance in Rural Tanzania’

Heather Klemick
Kenneth L. Leonard
Melkiory C. Masatu, M.D.

October 3, 2008

Note: The material contained herein is supplementary to the article named in the title and published in the American Journal of Agricultural Economics (AJAE).
Derivation of Competence and Performance Scores using Item Response Theory Analysis

For vignettes, the probability that clinician \( i \) implements protocol item \( j \) is modeled as a function of that clinician’s competence \( (A_i) \), an item-specific discrimination factor \( (\alpha_j) \) and an item-specific difficulty parameter \( (\beta_j) \). Discrimination is the degree to which more competent clinicians are likely to implement an item properly, and difficulty is the degree to which all clinicians are likely to implement an item correctly. To solve for these parameters, we use the logit model as follows:

\[
\text{prob}(a_{ij} = 1) = \frac{\exp(\hat{\alpha}_j A_i - \hat{\beta}_j)}{1 + \exp(\hat{\alpha}_j A_i - \hat{\beta}_j)}.
\]

For performance, we implement the same procedure, but add information about the characteristics of patient \( k \) \( (Z_k) \) and the length of time the clinician had been under observation \( (t) \). Patient characteristics include age (infant or child) and whether he or she showed multiple symptoms. We control for the number of previous consultations observed by the research team because clinicians in this sample exhibited high but declining quality while under observation—the Hawthorne effect. We estimate parameters for patient characteristics \( (\omega) \) and the Hawthorne effect \( (\delta) \), as well as for performance \( (A_i) \), item-specific discrimination \( (\alpha_j) \) and item-specific difficulty \( (\beta_j) \).

\[
\text{prob}(a_{jk} = 1) = \frac{\exp(\hat{\alpha}_j (A_i + \hat{\delta} t + Z_k \hat{\omega}) - \hat{\beta}_j)}{1 + \exp(\hat{\alpha}_j (A_i + \hat{\delta} t + Z_k \hat{\omega}) - \hat{\beta}_j)}
\]

1 The list of items and estimated discrimination and difficulty coefficients for the vignette and patient observation checklists are available from the authors upon request.

2 Leonard and Masatu (2006) demonstrated that doctors increased their effort when the research team first arrived and decreased it as they became used to the research team’s presence. By examining the quality of care provided by doctors who were never directly observed by the research team, the authors showed that declining quality was caused by the presence of the research team.
This methodology generates one competence and one performance score for each clinician that controls for the observable variation in case mix across clinicians. The IRT process does not identify the level or scale for either score, so we normalize them as follows. Some of the actual patients are similar to three of the vignettes (malaria, pneumonia and diarrhea). On average, clinicians in our sample completed 52.0% of the 26 comparable protocol items on the vignette (with a standard deviation of 17.2%) but only 39.6% of these items in practice (with a standard deviation of 10.0%). We normalize the distribution of the competence and performance scores to these means and standard deviations, respectively. These normalized scores can be interpreted as the percentage of items completed.

Figure 1 plots the competence and performance score pair for each clinician. It also shows the frontier that represents performance equal to competence. If competence is the maximum level of performance, then performance should always be less than or equal to competence. With the exception of five clinicians, this interpretation is reasonable: most clinicians are either close to the frontier, or well below it.

Figure 1 also indicates each clinician’s facility type, either public or non-public (private, parastatal or NGO). The dashed and dotted lines show the relationship between competence and performance for public and non-public facilities, respectively. The difference between the intercepts of these two lines is significant ($p = 0.01$) and indicates that the average non-public sector clinician implements approximately seven percentage points more items in practice than a public sector clinician with the same level of competence. The slope of performance with respect to competence is significantly
positive ($p = 0.01$), indicating that competence and performance are positively correlated, but the difference in the slopes is not significant ($p = 0.36$).
Note: Figure 1 shows the levels of competence and performance (derived from IRT analysis) for 95 clinicians who were observed with both vignette case studies and actual patients. The solid line represents the frontier where performance is equal to competence. The dashed line represents the trend line for clinicians at non-public facilities, and the dotted line represents the trend line for clinicians at public facilities.