WLS MEMO 175

Constructing a Single Measure of Cognitive Ability from the Two Adolescent Test Scores Available in the Wisconsin Longitudinal Study Graduate Sample

Jeremy Freese

February 22, 2006

The WLS contains Henmon-Nelson test scores for respondents from their freshman and junior years. For many purposes, it would be preferred to use a single measure that combined this information. The preferred measure that has been recommended by WLS (GWIIQ_BM) has been used the junior year score when available and a transformation of the freshman year score otherwise. An alternative possibility is to create a measure that incorporates information from both years when available.

The most obvious way of doing this would be (1) to average the two scores when two scores are available and (2) to use just one score from either year when only one score is available. A problem with this approach is that the metric for the two scores (contained in variables GWIIQ_J and GWIIQ_F) are not comparable because they are based on centile ranks from their respective years (GWICR_J and GWICR_F) and the cognitive gradient in drop-outs implies a higher mean centile rank for WLS respondents in their freshman year (54.4) than their junior year (50.0). As with the existing preferred measure, however, one can readily transform the freshman score so that the metrics are comparable.

A second problem, however, is that observation for which the measure was constructed for only one score would have less variance than observations with two scores. An important consequence of this would be that observations with only one score would be disproportionately represented in the tails of the cognitive ability distribution for purely artifactual reasons, which is especially problematic given potential interest in respondents with either very low or very high ability.

As an alternative, I began by taking cases that had available scores for their freshman and junior years and converting them into non-integer centile ranks.[1] The ranks are the proportion of those with both freshman and junior scores who scored lower than the respondent plus half the proportion of those with the same score as the respondent.

Then, these centile ranks were applied who had missing scores for either year.[2] If one had a freshman score but not a junior score, for example, one was assigned the freshman centile rank of others with the same freshman score, and the junior score was left unassigned.

---

[1] Because the Henmon-Nelson is less than 100 questions, not all centile ranks are used. A presumption of what follows is that equivalent centile ranks are to be treated as equivalent raw scores, which seems both reasonable and necessary.

[2] A empirical point of potential consequence here is that both non-parametric (rank-sum) and parametric (t) tests indicate that there is no difference in junior-year scores among those who do and do not have freshman-year scores, presumably because the availability of freshman year scores is mostly determined by whether the school

New normalized variables were created based on these centile ranks. The simple combined measure would simply average these measures when both are available and use the only available measure otherwise.

Instead, new measure #1 replaces missing values for the junior(freshman) score by taking the median junior(freshman) score for those who had the same freshman(junior) score as the respondent. The expected result is that there will now be greater variance among those with two scores rather than one score. While not ideal, of course, it is presumably better for any analysis of extreme cases to be biased in the direction of looking disproportionately toward cases in which the extremity results from two separate test administrations rather than being disproportionately biased toward including cases with only one score.

In order to equalize variances between those with one score and one with two scores, one needs to impose some solution that introduces variation in imputed values. Measure #2 uses a hotdeck-style solution to replace missing values for the junior(freshman) score with a randomly selected case that has the same freshman(junior) score as the missing case. Measure #2 results in no systematic bias in the variance due to whether one or both scores is observed. However, measure #2 has the disadvantage of not being deterministic and so exact values change depending on how the cases are randomly selected. One could use multiple imputations and methods for incorporating those, but little seems like it would be gained in this situation given that the correlation between measure #1 and measure #2 is > .98. The algorithm used, however, is easily modified so that multiple imputations could be generated.

Additional matters

1. The measures are created such that they are normally distributed with a mean of zero and a variance of one. WLS has distributed measures in which values are put on an IQ-like scale, but the author of this measure distrusts this metric because it implicitly assumes that Wisconsin high school juniors provide an adequate population for norming tests, despite known variation in drop-outs. The resulting measure explicit takes the Wisconsin juniors as the norming population and uses standard deviations as the metric to express this. Those so inclined can transform the measure to a variable with a mean of 100 and a standard deviation of 15 if they wish.

2. Because freshman IQ data are expressed in centile ranks, it would seem possible to rescale the data so that values correspond to a population normed on the WLS freshman. That task is outside the scope of this memo.

3. The existing preferred measure in the WLS (GWIIQ_BM) includes 215 observations for which scores exist that do not contain either junior or freshman scores (GWICR_J and GWICR_F). I believe these scores are valid but were derived by other means. The cases are assigned the mean value of other cases in the WLS with the same value of GWIIQ_BM.

---

administered the test that year and this school decision was apparently unrelated to the average cognitive ability of students.

4. The existing preferred measure in the WLS (GWIIQ_BM) was scaled in part to maximize its comparability with the sibling measures. The sibling measures were not taken into consideration when constructing measures here. Until a similar procedure is applied to the sibling measure (if possible), then analyses that depend on comparable measures for graduates and siblings seem like they should still use GWIIQ_BM.

Measures created

hn1 = Measure #1 described above
hn2 = Measure #2 described above
hn3 = average of two scores and simply using one score (from re-normalized scores) when available

Summary statistics

All variables have mean 0 and sd 1:

```
. su hn1 hn2 hn3

    Variable |       Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         hn1 |     10317    .0000214    .9989981  -3.224925   3.183736
         hn2 |     10317    .0000258    .9989384  -3.099489   3.183736
         hn3 |     10317    .0000108     .999466  -3.726893   3.489135
```

All variables are very highly correlated with one another and highly correlated with the WLS preferred measure (gwiiq_bm):

```
. cor hn1 hn2 hn3 gwiiq_bm
(obs=10317)

             |      hn1      hn2      hn3 gwiiq_bm
-------------+------------------------------------
         hn1 |   1.0000
         hn2 |   0.9852   1.0000
         hn3 |   0.9987   0.9839   1.0000
    gwiiq_bm |   0.9711   0.9569   0.9723   1.0000
```

For the reasons described above, the variance of hn1 is greater for those with freshman test scores than without, while the variance of hn3 is greter for those without freshman scores. Only hn2 results in approximately equal variances for both.

```
. table has_fscore, c(sd hn1 sd hn2 sd hn3)

----------------------------------------------
has       |
freshman  |
year      |
score     |    sd(hn1)     sd(hn2)     sd(hn3)
----------+-----------------------------------
        0 |   .9720979     .995098    1.022481
        1 |   1.011431     1.00072    .9883494
----------------------------------------------
```