Linking AzMERIT Scores to ACT Scores

This study aims to examine the relationship between student performance on the previous AzMERIT ELA and mathematics assessments and their subsequent performance on the ACT reading and mathematics sections.

Sample

The sample used to investigate relationships between AzMERIT and ACT was based on records for students who took grade 11 ELA and Algebra II tests in spring 2015 and the ACT at an appropriate time for graduation in 2016. From among the full set of spring 2015 grade 11 ELA and Algebra II test takers, there were 58,888 (93%) and 32,945 (56%) grade 11 students, respectively. These records represent the target sample for the analyses reported in this study.

The ACT data that we received from ADE was verified by comparing the percentage of students that were considered "college ready" from the data set with the ACT Profile Report for Arizona for the Graduating Class 2016. We note first that there were 33,871 records in the data file; the Profile Report showed 36,285 students in the 2016 graduating class took the ACT. Second, the ACT Profile Report showed that 39% students met ACT college readiness benchmarks for Reading and 38% for mathematics based on Benchmark Scores of 22. The data received from ADE shows 38% for reading and 37% for mathematics met the same Benchmark Scores. Last, approximately 8,500 records in the ACT data file had missing SAISID's. The merging between the ACT records and the AzMERIT records was done by student name and birthday for the cases where the SAISID didn't match between the two data files.

Because a large number of students did not take the ACT and the two subgroups differed systematically across demographic and achievement variables, the imputing approach is often employed to handle missing data in the analysis of the relationship between the AZMERIT scores and subsequent performance on the ACT. However, previous studies conducted in Minnesota and Ohio showed that imputing or deleting the missing records did not impact the linkage identified between their graduation tests and the ACT test. In this study, the regression model that links the AzMERIT scale score to the ACT scale score was built based on the merged ACT and AzMERIT records. The records were than randomly split between a model building and validation sample to estimate the variability of the model fit and to estimate the error rate of the model when applied in new previously unseen data.

Multiple Regression Model

ELA. Examinees with missing ACT or AzMERIT scale scores were removed from the merged dataset. Then the ACT reading scale score for the remaining 25,977 students were regressed onto the applicable grade 11 ELA scale score and demographic variables. Stepwise selection was used to identify the prediction model. The following regression equation, which has the smallest AIC, smallest RMSE and largest adjusted R^2, was identified as the best model to predict ACT reading from prior performance on the AzMERIT ELA test:

$\hat{Y} = -290.65+0.12*X1+0.26*X2-2.35*X3-0.79*X4+0.57*X5-2.32*X6-1.79*X7-2.40*X8-1.82*X9-2.07*X10$, where

Ŷ =ACT Reading Scale Score
X1 = AZMERIT ELA Scale Score
X2 = Female-Male Contrast
X3=American Indian-White Contrast
X4 = Multi-ethnic Contrast
X5 = Asian Contrast
X6 = Hispanic-White Contrast
X7 = African American-White Contrast
X8= Native Hawaiian-White Contrast
X9 = Free and Reduced Lunch Contrast
X10= ELL Contrast

The overall model was statistically significant ($F_{(10,\ 20388)}$ = 1,704.70, p < .0001; adjusted R^2 = 0.46). Table 1 shows the estimated model parameters. Application of this regression model indicates that an AzMERIT ELA scale score 2585 is associated with the ACT reading college ready cut score of 22.

Table 1. Estimated Parameters of the ELA Regression Model

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-290.65	2.72	-87.06	<.0001
X1-AzMERIT	1	0.12	0.00	95.44	<.0001
X2-Gender	1	0.26	0.07	3.90	<.0001
X3-American Indian	1	-2.35	0.17	-13.80	<.0001
X4-Multi	1	-0.79	0.26	-3.11	0.0019
X5-Asian	1	0.57	0.17	3.26	0.0011
X6-Hispanic	1	-2.32	0.08	-28.03	<.0001
X7-African American	1	-1.79	0.16	-11.33	<.0001
X8-Native Hawaiian	1	-2.40	0.82	-2.94	0.0033
X9-FRL	1	-1.82	0.08	-23.50	<.0001
X10-ELL	1	-2.07	0.40	-5.23	<.0001

Mathematics. Moving on to mathematics, all records with missing ACT or AzMERIT scale scores were excluded from the analysis. Then the ACT mathematics scale scores for the remaining 13,777 students were regressed onto the applicable AzMERIT Algebra II test and demographic variables. Stepwise selection was used to identify the prediction model. The following regression equation, which has the smallest AIC, smallest RMSE and largest adjusted R^2, was identified as the best model to predict ACT mathematics scores from prior performance on the AzMERIT Algebra II test:

$\hat{Y}$ = – 305.7 + 0.08*X1 -0.55*X2 – 1.55*X3 – 0.48*X4 -0.44*X5 – 1.44*X6 – 1.41*X7 – 0.83*X8– 1.22*X9– 1.57*X10, where

$\hat{Y}$ =ACT Mathematics Scale Score
X1 = AZMERIT Reading Scale Score
X2 = Female-Male Contrast
X3=American Indian-White Contrast
X4 = Multi-ethnic Contrast
X5 = Asian Contrast
X6 = Hispanic-White Contrast
X7 = African American-White Contrast
X8= Native Hawaiian-White Contrast
X9 = Free and Reduced Lunch Contrast
X10= ELL Contrast

The overall model was statistically significant ($F_{(10,\ 13768)}$ =1,764.13, p < .0001; adjusted R^2 = 0.5). Table 2 shows the estimated model parameters. Application of this regression model indicates that an AzMERIT mathematics score of 3727 is associated with the ACT mathematics college ready cut score of 22.

Table 2. Estimated Parameters of the Mathematics Regression Model

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	-305.70	2.92	-98.08	<.0001
X1-AzMERIT	1	0.09	0.00	104.98	<.0001
X2-Gender	1	-0.55	0.05	-12.05	<.0001
X3-American Indian	1	-1.55	0.10	-15.72	<.0001
X4-Multi	1	-0.49	0.17	-2.79	0.01
X5-Asian	1	-0.44	0.16	-2.79	0.01
X6-Hispanic	1	-1.44	0.05	-28.21	<.0001
X7-African American	1	-1.41	0.10	-13.56	<.0001
X8-Native Hawaiian	1	-0.83	0.49	-1.70	0.09
X9-FRL	1	-1.22	0.03	-18.08	<.0001
X10-ELL	1	-1.57	0.38	-7.60	<.0001

Validation Set

Reserving a random sample of records as a validation set allows for better estimation of model fit. Estimates of model fit derived from the sample on which the model is constructed is always inflated, because any idiosyncrasies in the model building sample are incorporated into the model parameters. By holding out a subset of the data from the model building process, the performance of the model is evaluated on the independent validation sample, which leads to more accurate calculation of model error. This approach allows for more accurate estimates of model fit.

In this study, the training dataset contained 50% randomly selected merged records and the testing dataset included the other 50% of students. Following identification of the regression model, the predictive model was applied onto the validation set. The Root Mean Squared Error (RMSE) was calculated as the square root of the average squared errors found between the actual ACT score point and the model fitted values. We repeated this sampling and model fitting process 100 times to see how the RMSE varies. For ELA, the average RMSE was 5.03 with a standard deviation of 0.02 across 100 replications. For mathematics, the average RMSE was 2.79 with a standard deviation of 0.02. The standard deviation of the RMSE is very small indicates that the sample selected for the modeling had no significant impact on the model fitting.

Equipercentile Equating

The equipercentile equating method was used to verify the linkage between the ACT and AzMERIT test scores. Table 3 shows the ACT reading scale score to AzMERIT grade 11 ELA scale score conversion. Table 4 provides the ACT mathematics to AzMERIT Algebra II conversion table. Using the equipercentile approach, the AzMERIT scale score associated with the ACT college ready cut score was 2585.72 for ELA and 3727.46 for mathematics. This is consistent with the cut scores identified using regression models.

Table 3 ACT Reading Scale Score to AzMERIT Grade 11 ELA Scale Score conversion.

ACT	AzMERIT	SE	SE.bootstrap
1	2464.3	0.01	0
2	2464.3	0.01	0
3	2464.32	0.03	0
4	2464.36	0.05	0.01
5	2464.45	0.08	0.01
6	2464.66	0.14	0.03
7	2467.42	0.21	0.52
8	2475.11	0.31	1.13
9	2487.7	0.38	1.24
10	2500.62	0.44	0.99
11	2512.43	0.42	0.77
12	2522.88	0.41	0.57
13	2532.15	0.4	0.49
14	2540.41	0.37	0.43
15	2547.89	0.35	0.4
16	2554.69	0.34	0.4

17	2560.92	0.32	0.38
18	2566.66	0.31	0.38
19	2571.96	0.3	0.37
20	2576.85	0.29	0.35
21	2581.37	0.28	0.33
22	2585.72	0.27	0.26
23	2589.45	0.27	0.3
24	2593.07	0.26	0.28
25	2596.42	0.26	0.26
26	2599.53	0.26	0.25
27	2602.45	0.26	0.25
28	2605.74	0.24	0.37
29	2609.82	0.26	0.21
30	2612.41	0.27	0.32
31	2615.19	0.25	0.32
32	2618.22	0.26	0.37
33	2621.53	0.28	0.42
34	2625.47	0.28	0.46
35	2630.84	0.3	0.49
36	2640.88	0.4	0.51

Table 4 ACT Mathematics Scale Score to AzMERIT Algebra II Scale Score conversion.

ACT	AzMERIT	SE	SE.bootstrap
5	3628.62	0	0
7	3628.62	0	0
8	3628.63	0.01	0
9	3628.68	0.01	0
10	3628.83	0.02	0.01
11	3629.24	0.04	0.02
12	3630.43	0.06	0.15
13	3637.46	0.06	0.3
14	3648.71	0.06	0.3
15	3659.58	0.07	0.29
16	3669.8	0.08	0.25
17	3679.88	0.08	0.24
18	3690.91	0.09	0.24
19	3699.56	0.11	0.27
20	3709.07	0.12	0.29
21	3718.41	0.14	0.35

22	3727.46	0.15	0.4
23	3736.3	0.19	0.47
24	3744.74	0.2	0.52
25	3752.7	0.23	0.6
26	3760.15	0.26	0.68
27	3767	0.3	0.8
28	3773.15	0.31	0.89
29	3778.86	0.37	1.03
30	3784.1	0.39	1.11
31	3789.07	0.42	1.2
32	3794.02	0.45	1.3
33	3799.38	0.48	1.4
34	3805.91	0.5	1.47
35	3815.31	0.49	1.44
36	3835	0.56	0.82

Impact

To evaluate the impact applying the ACT college ready cut to the AzMERIT test scores, we observe that 34% of all the AzMERIT grade 11 ELA test takers scored at or above 2585. Moreover, the similar percentage (35%) of students who are at or above ACT score of 22 in the analysis sample. Next, we observe that 16% of students scored at or above 3727 on the AzMERIT Algebra II test. However, there are 21% of students score at or above ACT score of 22 in the analysis sample. The discrepancy in mathematics may be explained by the fact that only 56% Algebra II test takers were grade 11 students while 93% grade 11 ELA test takers are on grade.

Differential Prediction across Subgroups

We also explored whether the relationship between AzMERIT and ACT tests is consistent across different subgroups. The subgroups of interest included gender, ethnicity, socioeconomic status (Free or Reduced Price Lunch, FRL) and English language learner (ELL). The question of interest is whether the predictive relationship between AzMERIT and ACT differs across student demographic and ability subgroups. To evaluate predictive effects of student ability subgroups, students were categorized by their performance levels reported by AzMERIT tests, which are minimally proficient, partially proficient, proficient and highly proficient. The same multiple regression model ran for the whole sample was repeated in each of the subgroup samples for both ELA and mathematics tests.

Table 5 summarizes the standardized coefficient estimates and the partial R-square of AzMERIT scale score variable, the model R-square and the ratio of the partial R-square to the model R-square produced by regression models built using each of the subgroup samples. For the ELA test, the models built using male or female sample were very similar in terms of the overall model fitting, the variation explained by AzMERIT scale score and the AzMERIT coefficient estimates.

Across all ethnicity groups, the R-square suggests that the regression models were fitted better using Asian and Multiracial datasets compared to using Hispanic or African American datasets. The partial R-square of AzMERIT scale score was higher for Asian, Multiracial and White groups than for the other groups, which means the predictive relationship between the grade 11 achievement score and ACT score was stronger for these groups. The sample size for Native Hawaiian was too small to build the regression model in this case. The lower R-square and partial R-square of AzMERIT scale score observed in the FRL model shows that the AzMERIT scale score did not predict the ACT score for the students receiving free or reduced lunch as well as it did for the other students. The model within ELL group was very poorly fitted due to the very small sample size available in grade 11 test.

With respect to ability subgroups, although the sample size for each of the performance groups was large, the models fitted within each of the ability subgroups were poor. This is an artifact of truncation of range of ability within each of the ability subgroups. That said, the higher the performance level, the better the model was fitted in that group. This effect was expected since the ACT, as a college entrance exam, targets test information only to the highest ability levels, while test information for the AzMERIT is designed to measure achieve across a broader range of student achievement. The models built using mathematics samples were generally fitted better than their ELA counterparts. However, the model fitting and the predictivity of AzMERIT showed very similar pattern.

Table 5. Regression Estimates across Subgroups for ELA and mathematics

	Standardized Coefficient Estimate of AzMERIT Scale Score	**Partial R-Square of AzMERIT Scale Score**	**Model R-Square (Adjusted)**	**Partial R-Square/Model R-Square**
	ELA			
Male	0.52	0.37	0.45	82%
Female	0.53	0.39	0.46	85%
Multiracial	0.61	0.40	0.42	95%
American Indian	0.52	0.29	0.31	94%
Asian	0.67	0.50	0.51	98%
Hispanic	0.50	0.29	0.32	91%
African American	0.52	0.30	0.33	91%
White	0.58	0.36	0.37	97%
Native Hawaiian	0.32	0.10	0.10	100%
Non-FRL	0.58	0.38	0.40	95%
FRL	0.50	0.29	0.33	88%
Non-ELL	0.53	0.38	0.45	84%
ELL	0.17	0.03	0.05	60%
Minimally Proficient	0.04	0.08	0.12	67%
Partially Proficient	0.15	0.06	0.12	50%

Proficient	0.25	0.07	0.15	47%
Highly Proficient	0.34	0.14	0.22	64%
	Mathematics			
Male	0.64	0.46	0.50	92%
Female	0.63	0.46	0.50	92%
Multiracial	0.65	0.42	0.42	100%
American Indian	0.60	0.37	0.37	100%
Asian	0.68	0.46	0.49	94%
Hispanic	0.64	0.42	0.44	95%
African American	0.56	0.34	0.36	94%
White	0.67	0.47	0.49	96%
Native Hawaiian	0.71	0.51	0.51	100%
Non-FRL	0.66	0.46	0.48	96%
FRL	0.64	0.42	0.44	95%
Non-ELL	0.63	0.45	0.50	90%
ELL	0.59	0.37	0.39	95%
Minimally Proficient	0.24	0.06	0.14	43%
Partially Proficient	0.23	0.06	0.14	43%
Proficient	0.38	0.17	0.24	71%
Highly Proficient	0.41	0.23	0.31	74%