ACT Research Report Series 99-2

Accuracy of Course Placement Validity Statistics Under Various Soft Truncation Conditions

Jeff L. Schiel

Jason E. King

ACT

January 1999

For additional copies write:
ACT Research Report Series
PO Box 168
Iowa City, Iowa 52243-0168

Accuracy of Course Placement Validity Statistics Under Various Soft Truncation Conditions

Jeff L. Schiel
Jason E. King

Abstract

Analyses of data from operational course placement systems are subject to the effects of truncation: Students with low placement test scores may enroll in a remedial course, rather than a standard-level course, and therefore will not have outcome data from the standard course. In "soft truncation," some (but not all) students who score below the cutoff for a standard course enroll in and complete the course. Previous research, using one particular definition of soft truncation, showed that reasonably accurate validity statistics can be estimated under this condition. Alternative definitions of soft truncation could conceivably result in different validity statistics. This simulation study therefore examined an alternative definition of soft truncation, in which students who score just below the cutoff have a higher probability of enrolling in the standard course than do relatively lower-scoring students.

The effects of different combinations of soft truncation condition, logistic regression curve, test score distribution shape, and sample size on estimated optimal cutoff scores, accuracy rates, and success rates were summarized. Postsecondary institutions that experience a moderate degree of soft truncation (e.g., 20% to 60% of their respective placement groups) can expect to obtain acceptably accurate estimates of optimal cutoff scores, irrespective of the steepness of the logistic curve and the skewness of the marginal distribution of the predictor variable.

Accuracy of Course Placement Validity Statistics Under Various Soft Truncation Conditions

It is common practice for postsecondary institutions to use standardized test scores for placing students into courses. Of particular interest to these institutions is establishing cutoff scores that will result in a high percentage of correct placement decisions: Students scoring at or above a particular cutoff score are placed into a standard course and are ultimately successful in it; those scoring below the cutoff are placed into a remedial course and would not have succeeded in the standard course had they been placed into it. Incorrect placement decisions are likely to have negative consequences for both students and institutions. The student who is incorrectly placed into standard freshman English, for example, but lacks the required skills and knowledge to complete the course with a passing grade may become discouraged about his or her academic progress.

It is important for postsecondary institutions to establish statistical validity evidence for using test scores in course placement. Such evidence provides a rationale for using particular tests, other variables, and cutoff scores. The institution can then use the evidence to respond to potential criticism of its placement practices.

Logistic regression and decision theory are well suited for describing relationships between outcomes in standard college courses and test scores, for establishing cutoff scores, and for providing course placement validity evidence. With logistic regression, a binary outcome variable (e.g., course success or failure) can be modeled as a function of test score, yielding an estimated conditional probability of success ($\hat{P}$) in the standard course. Estimated conditional probabilities obtained from a logistic regression model can then be used with the marginal distribution of test scores to estimate other course placement validity statistics, such as accuracy rates and success rates, which in turn can be used to identify an optimal cutoff score. The *optimal cutoff score* is the cutoff score at which the estimated accuracy rate ($\hat{A}$) is maximized.

In the context of evaluating course placement systems, $\hat{A}$ is defined as the proportion of correct placement decisions, and focuses on estimated probabilities of success for two groups of students: 1) those students scoring at or above the cutoff score for a standard course who are adequately prepared for and successful in the course, and 2) those scoring below the cutoff score who need remedial instruction and therefore would not have been successful in the standard course had they enrolled in it. The estimated success rate ($\hat{S}$) is defined as the proportion of students succeeding in the standard course, among all students who could have been placed in that course. Additional information about using logistic regression and decision theory to estimate $\hat{P}$, $\hat{A}$, $\hat{S}$, and to evaluate course placement systems can be found in ACT (1994) and Sawyer (1989, 1996).

One inherent problem in evaluating course placement systems is that students who score below the cutoff often do not enroll in the standard course and therefore do not have standard course outcome data (e.g., grades). This situation results in a marginal course outcome distribution that is truncated below the cutoff score. The logistic regression function, which is computed from the data of students who completed the standard course, must therefore be extrapolated to test scores below the cutoff score in order to estimate $\hat{P}$, $\hat{A}$, and $\hat{S}$. Thus, these statistics will be useful only to the extent that their accuracy is not adversely affected by truncation.

Truncation of the marginal course outcome distribution may occur in varying degrees; moreover, increases in the severity of truncation are associated with decreases in the accuracy of validity statistics. The most severe truncation is *hard truncation*, a condition in which no standard course outcome data are available for students below a cutoff score in the marginal distribution of test scores. Hard truncation has been shown to reduce the accuracy of estimated validity statistics relative to those based on non-truncated data (Houston, 1993; Schiel & Noble, 1992; Schiel, 1998), sometimes with grave consequences. Schiel (1998), for example, found that the estimated optimal ACT Assessment cutoff score for one simulated sample was 23 before truncation, but under hard truncation, the optimal cutoff score decreased to 6, a difference of about 3½ standard deviations.

Fortunately, hard truncation is uncommon in practice. A more prevalent type of truncation is *soft truncation*, in which standard course outcomes are available for some of the students who score below a particular cutoff score. Soft truncation occurs in those course placement systems in which cutoff scores are used as guides and are not strictly enforced by an institution, thereby permitting some students who score below the cutoff score, but who are confident of their ability to succeed in the standard course, to enroll in that course. These students will have course outcome data, even though their test scores are below the placement cutoff score.

Research on soft truncation suggests that this condition has much less severe effects on the accuracy of estimated validity statistics. In a recent simulation study, for example, soft truncation was defined as omitting fixed percentages of observations from each of several score intervals below the estimated optimal cutoff score (Schiel, 1998). Given this particular definition, reasonably accurate optimal cutoff scores (e.g., accurate to within one ACT scale score point) and validity statistics were able to be estimated even under 40%, and in some cases 60% to 80%, soft truncation.

In Schiel's (1998) definition of soft truncation, it is possible for even very low-scoring students to have standard course outcome data, a situation which may not often be encountered in practice. Alternative definitions of soft truncation are possible, of course, and may better represent the actual environment of some course placement systems. It is conceivable that these other types of soft truncation could result in estimated validity statistics that differ somewhat from those observed by Schiel (1998).

The results of Schiel's soft truncation research are based on relatively large (non-truncated) initial samples consisting of 500 observations each. In practice, however, samples of actual data with fewer than 500 observations are common. Some research has been conducted on the relationship between sample size and validity statistic accuracy, but only under hard truncation conditions. For example, in a study in which a bootstrap method was used to estimate confidence intervals for validity statistics, Crouse (1996) found that given a normally distributed predictor variable and steep logistic regression slope, larger samples with relatively less hard truncation yield relatively more accurate estimated validity statistics. It is likely that, under soft truncation, large samples would similarly give relatively more accurate results. The current study was therefore implemented to determine the extent to which alternately defined soft truncation conditions, applied to initial non-truncated samples of various sizes, affect the accuracy of estimated validity statistics.

Method

Placement Group Simulation

A *placement group* consists of all students for whom a placement decision needs to be made and is, by definition, non-truncated (ACT, 1994). Data for five placement groups, each containing 500 observations previously simulated for use in the Crouse (1996) and Schiel

(1998) studies, were also used in the present study. Data for six additional placement groups were simulated in samples of 100 and 150 observations. These sample sizes were chosen to represent placement group sample sizes of relatively small postsecondary institutions participating in ACT's Course Placement Service (CPS), which helps institutions identify optimal cutoff scores. In addition, these six additional placement groups were sufficiently large to ensure that an adequate number of observations were available for estimating the parameters of logistic regression functions, even after hard truncation was implemented. For example, the smallest truncated sample that occurred for the size n=100 placement groups consisted of 37 observations. It has been shown that the accuracy of estimated logistic regression parameters declines significantly when very small sample sizes (e.g., n=25) are involved (Houston, 1993).

Each placement group was intended to be representative of data that ACT receives from institutions participating in the CPS. It consisted of the joint distribution of two random variables: *X*, which reflects the ACT Assessment score scale (1-36) and *Y*, which reflects a standard course outcome (i.e., success or failure). Additional information about the simulation procedure is provided in Schiel (1998).

Two additional factors, other than sample size, were varied in the simulations: the slope of the logistic regression curve and the skewness of the marginal distribution of the ACT score variable *X*. The five simulated placement groups of size n=500 are described in Table 1, which shows, for each group, the resulting skewness and logistic regression parameters *a* and *b*. In addition, this table shows the estimated optimal cutoff score for each group. The simulated data for placement Groups 1-3 had steep logistic regression curves and either high (-.62) or medium (-.29) negative skewness, or virtually zero (.03) skewness, respectively.

Groups 4 and 5, on the other hand, consisted of simulated data with relatively flat logistic regression curves for two of the three categories of skewness. A sixth placement group of size n=500, which had a flat slope and medium skewness, was also examined. However, the optimal cutoff score that was estimated for this placement group was only 14, which prevented soft truncation from being successfully simulated (see the description of soft truncation below).

TABLE 1

Simulated Placement Groups

Placement group	Size	Estimated optimal cutoff	Slope of logistic curve		Skewness of marginal distribution	
1	500	20	Steep	(a = -2.18, b = .11)	High	(-.62)
2	500	20	Steep	(a = -2.46, b = .12)	Medium	(-.29)
3	500	21	Steep	(a = -2.22, b = .11)	Zero	(.03)
4	500	23	Flat	(a = -.79, b = .03)	High	(-.61)
5	500	18	Flat	(a = -.53, b = .03)	Zero	(-.01)
6	100	20	Steep	(a = -2.43, b = .12)	High	(-.44)
7	100	17	Steep	(a = -2.44, b = .14)	Medium	(-.16)
8	100	21	Steep	(a = -2.80, b = .13)	Zero	(.13)
9	150	20	Steep	(a = -3.92, b = .17)	High	(-.52)
10	150	20	Steep	(a = -1.79, b = .10)	Medium	(-.26)
11	150	20	Steep	(a = -1.60, b = .08)	Zero	(-.13)

Table 1 also shows the "intercept" (a) and "slope" (b) parameters of the logistic regression curves and the skewness for placement groups of size n=100 and n=150. Placement groups with flat logistic curves for these sample sizes were also simulated. In each case, however, the $\hat{A}$ achieved a maximum value for one or more very low ACT scores (e.g., < 5), indicating that these were optimal cutoff scores. Very low cutoffs such as these would not likely be employed in actual course placement systems. As a consequence, data from placement groups with flat logistic curves were not analyzed further for sample sizes of 100 and 150.

In practice, when very low optimal cutoff scores are identified, at least initially, one possible solution is to adjust the criterion variable, because a result such as this suggests that more than half of the low-scoring students are succeeding in the course. For example, if the criterion (successful course performance) is defined as earning a grade of C or higher, but does not result in an $\hat{A}$ curve that has a definite peak, then successful course performance may be redefined as a grade of B or higher. This redefinition typically results in an $\hat{A}$ curve with a definite peak, allowing an optimal cutoff score to be identified. Sawyer (1996) provides additional information about this situation. A redefinition of successful course performance was not simulated in this study, but could be a topic for future research.

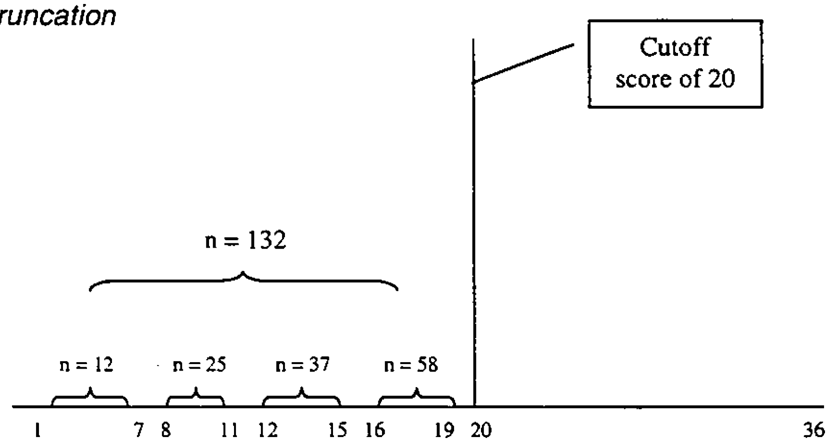
Note that the skewness for Placement Groups 8 and 11, although specified in each simulation to be very close to zero, turned out to be somewhat different than expected. Placement Group 8 has slight positive skewness, and Placement Group 11 has slight negative skewness. This results from the relatively small size of the samples being simulated; similar slight positive or negative skewness occurred when additional simulations of these placement groups were performed. The problems inherent in simulating small samples have important implications for interpreting the results of this study, and are described further in a following section (see "Discussion").

Former Definition of Soft Truncation

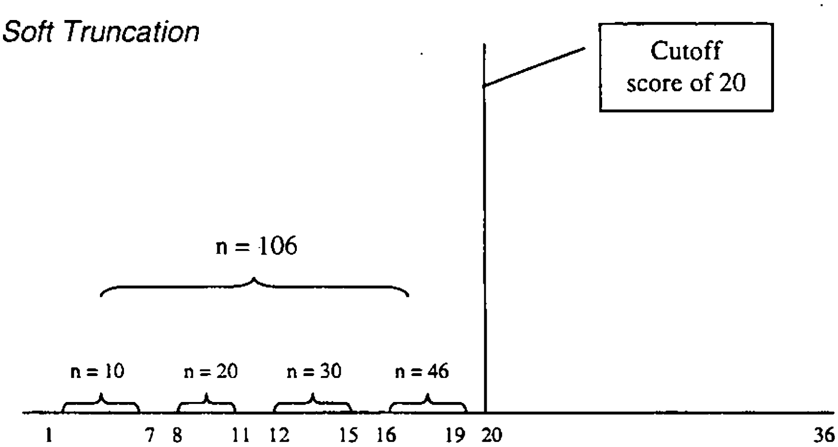
Panel A of Figure 1 depicts Schiel's (1998) definition of soft truncation for Placement Group 1 (n=500). This panel shows that there are 132 observations (consisting of (x,y) pairs) below an ACT cutoff score of 20. In Panel B, 20% of observations within each of four score intervals have been randomly selected and then deleted, reducing the total number of observations below the cutoff to 106. However, 10 of the 12 observations in the lowest score interval (i.e., "1 – 7") still remain in the sample. It is plausible that in many course placement systems few, if any, students obtaining ACT Assessment scores below 8 would enroll in and complete a standard course. Moreover, it is possible that students whose scores are below the cutoff but nearer to it would be more likely to enroll in the standard course than those whose scores are both below and farther from the cutoff. The alternative definition of soft truncation described below considers these factors.

FIGURE 1. Former Definition of Soft Truncation
(Placement Group 1: Steep slope, high skewness, n=500)

A. No Truncation



B. 20% Soft Truncation



Alternative Definition of Soft Truncation

The alternative definition of soft truncation has two key elements that distinguish it from the former definition: 1) It allows for observations below the cutoff score, but nearer to it, to have a higher probability of being retained in the truncation sample than those observations that are below the cutoff, but farther from it, and 2) it reflects the situation in which the lowest-scoring students do not enroll in the standard course; that is, all observations below a certain "lower bound" are deleted.

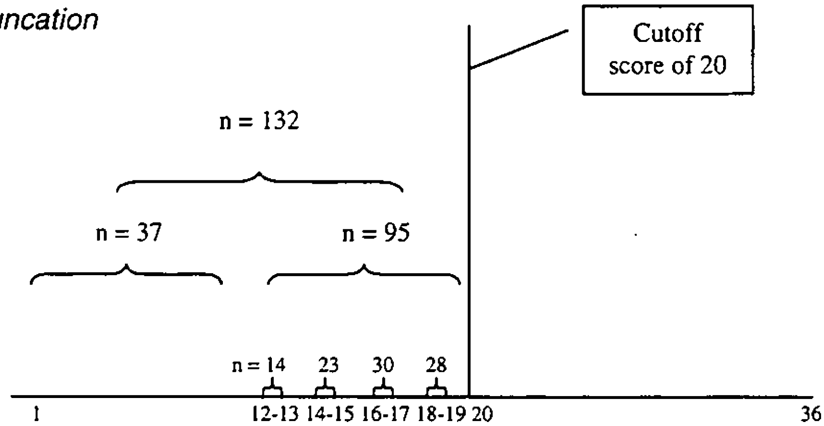
The lower bound was chosen based on typical chance-level scores on the ACT Assessment. A chance-level score is the score that an examinee is likely to obtain if he or she simply guesses when responding to each item. On the ACT Reading test, for example, each item has four possible responses. Therefore, an examinee has a one-in-four chance of correctly answering any one item by guessing. The chance-level raw (number-correct) score for this test is equal to the sum of the individual probabilities for the 40 Reading items: 40 x 1/4 = 10. Depending on the particular form of the ACT Assessment, this raw score converts to a scale score of approximately 11.

Chance-level ACT scores vary across the four subject-area tests and across forms. Across 21 recently administered forms of the ACT, for example, the average chance-level score for the English test was determined to be about 10; for the Mathematics, Reading, and Science Reasoning tests, average chance-level scores were 13, 11, and 13, respectively. Thus, the average chance-level Composite score was approximately 12. For consistency across subject-area tests, a score of 12 was chosen as the lower bound below which all scores in the simulated truncation samples were deleted.

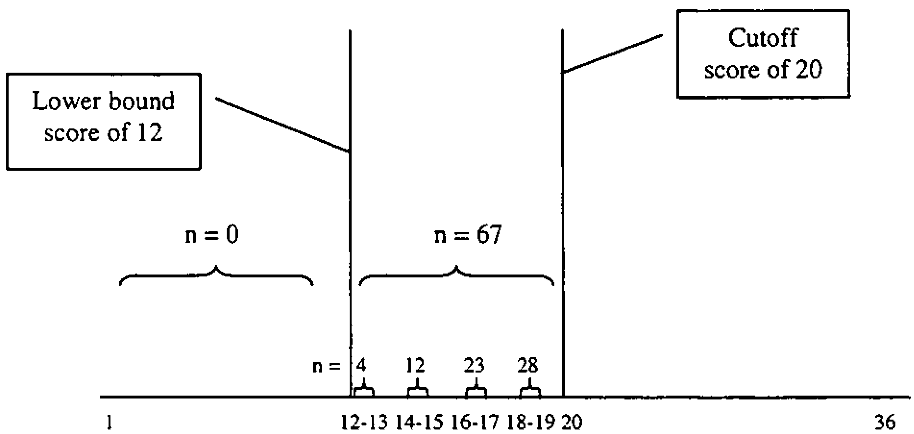
For each placement group, four score intervals were established between the lower bound and the estimated optimal cutoff score (see Panel A of Figure 2). Because optimal cutoff scores vary somewhat over placement groups, the intervals were not always of equal width as depicted in Figure 2. Panel B shows the graduated *baseline condition* in which (a) all observations below the lower bound of 12 have been deleted, and (b) more observations have been retained in intervals closer to and below the cutoff score. Specifically, no observations have been deleted from the first interval below the cutoff; however, 25% have been randomly selected and then deleted from the second interval, 50% have been deleted from the third, and 75% have been deleted from the fourth. These adjustments result in a distribution whose left side is short-tailed and relatively steep below the cutoff. Panel C of Figure 2 portrays a 20% additional soft truncation condition in which 20% of the observations in each interval in Panel B have been randomly selected and then deleted (e.g., 20% of the observations in interval "12-13" have been deleted, reducing the count from 4 to 3). Next, 40%, 60%, and 80% truncation conditions were implemented (but are not shown).

FIGURE 2. Alternative Definition of Soft Truncation
(Placement Group 1: Steep slope, high skewness, n=500)

A. No Truncation



B. Baseline Graduated Soft Truncation



C. 20% Additional Soft Truncation

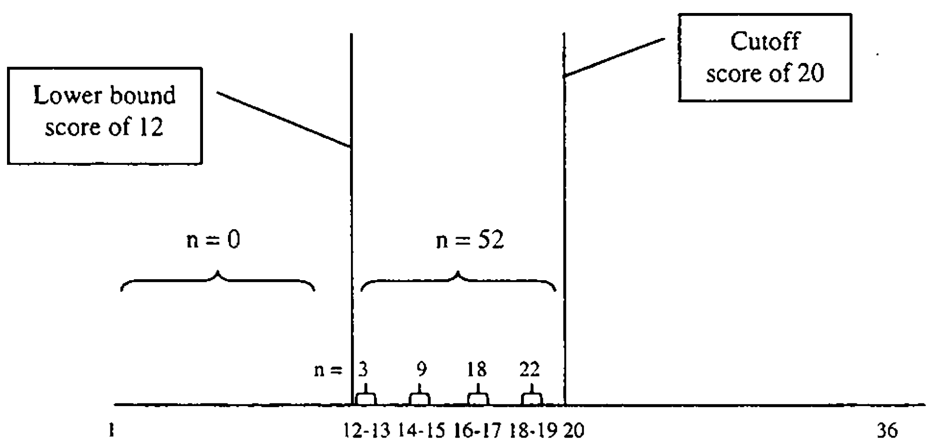


Table 2 contains truncation sample sizes, by placement group and truncation condition. Depending on the shape of the distribution and the location of the optimal cutoff score, truncation samples varied considerably in size. For example, for placement groups of size n=500 (Groups 1-5), truncation sample sizes under the 80% condition varied from 212 (Group 3; steep slope, zero skewness) to 382 (Group 1; steep slope, high skewness). The smallest truncation sample occurred for Placement Group 8 (steep slope, zero skewness, n=100) under the hard truncation condition; only 37 observations remained in this sample after hard truncation was implemented.

TABLE 2

Truncation Sample Sizes, by Placement Group and Truncation Condition

	Truncation condition					
Placement group	**Baseline**	**20%**	**40%**	**60%**	**80%**	**Hard**
1	435	420	408	394	382	368
2	401	381	361	341	321	301
3	333	303	273	244	212	182
4	438	415	395	372	352	330
5	386	366	346	325	305	285
6	85	81	79	78	75	72
7	86	81	79	76	72	70
8	70	63	56	51	44	37
9	126	119	113	109	102	96
10	126	121	116	112	105	101
11	99	90	80	72	61	53

Soft Truncation Simulation and Estimation of Validity Statistics

The procedure used to simulate soft truncation and compare the resulting estimated validity statistics to those of the non-truncated placement group was similar to that described in Schiel (1998) and consisted of the following steps:

1. *Calculate estimated validity statistics for the placement group.*

2. *Simulate soft truncation.* This step is repeated 500 times to obtain 500 truncation samples for a particular soft truncation condition (e.g., 20% soft truncation).

3. *Repeat Step 2 for different soft truncation conditions.*

4. *Calculate estimated $\hat{P}$, $\hat{A}$, and $\hat{S}$ at each ACT scale score point (representing different hypothetical cutoff scores) for each truncation sample generated in Steps 2 and 3.*

5. *For each truncation condition, calculate median estimated $\hat{P}$, $\hat{A}$, and $\hat{S}$ across 500 truncation samples, by ACT score.*

6. *Compare the median estimates from Step 5 to those of the placement group (Step 1).*

7. *Compute estimated $\hat{P}$, $\hat{A}$, and $\hat{S}$ at each ACT scale score point for a hard truncation condition (i.e., no observations below the optimal cutoff score) and compare these statistics to those obtained in Steps 1 and 5.*

8. *Repeat the entire procedure (Steps 1-7) 11 times, once for each of the 11 simulated placement groups.*

This procedure yielded, for each combination of simulated placement group and truncation condition, estimated $\hat{P}$s, $\hat{A}$s, and $\hat{S}$s for the 36 ACT scale score points. These validity statistics were plotted for comparison purposes. In addition, differences between the validity statistics estimated from the simulated placement groups and the truncation samples were calculated. For example, the $\hat{A}$ for a (non-truncated) placement group ($\hat{A}_N$) was subtracted from the $\hat{A}$ for a baseline soft truncation condition ($\hat{A}_B$) for each possible ACT scale score point[1]:

---

[1]A subscript of N will henceforth denote a non-truncated, simulated placement group (e.g., $\hat{S}_N$ is the success rate for this group), and a subscript of B will denote the baseline soft truncation condition.

$$\Delta \hat{A}_B^{(i)} = \hat{A}_B^{(i)} - \hat{A}_N^{(i)} .$$

A total of 36 accuracy rate differences were calculated. This also pertained to the calculation of $\Delta\hat{P}_B$ and $\Delta\hat{S}_B$. These statistics were used to evaluate the accuracy of estimated $\hat{P}$, $\hat{A}$, and $\hat{S}$.

Similar calculations were performed for the 20%, 40%, 60%, 80%, and hard truncation conditions. Mean differences were then calculated, and means of the absolute value of the differences were also calculated. The mean of the absolute values of the $\Delta\hat{A}_B^{(i)}$, for example, may be expressed as

$$| \overline{\Delta\hat{A}_B} | = \frac{1}{36} \sum_{i=1}^{36} | \Delta \hat{A}_B^{(i)} | .$$

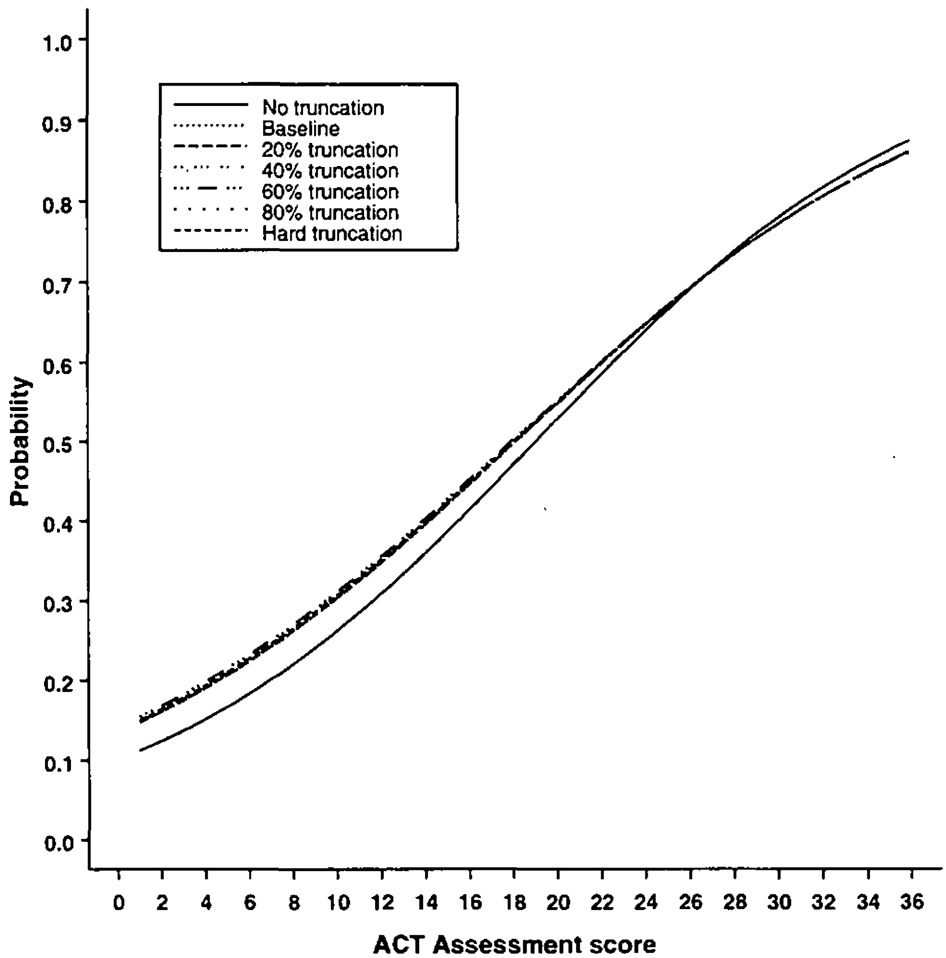
Results

Estimated Conditional Probabilities of Success

The effects of soft truncation on the estimated probabilities of success for Placement Group 1 (steep slope, high skewness, n=500) are displayed in Figure 3.1. The solid curve in this figure represents probabilities for the non-truncated placement group. Probabilities for the five soft truncation conditions and the hard truncation condition are shown by either dashed or dotted curves which, for this particular placement group, have considerable overlap and are therefore difficult to distinguish. Figure 3.1 illustrates that the effects of soft truncation on the estimated conditional probabilities of success were small and fairly consistent over the different truncation conditions. Under soft truncation, conditional probabilities for Placement Group 1 were overestimated for ACT Assessment scores of 25 or lower, and were underestimated slightly for higher scores (e.g., > 28).

FIGURE 3.1. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 1: Steep slope, high skewness, n=500)

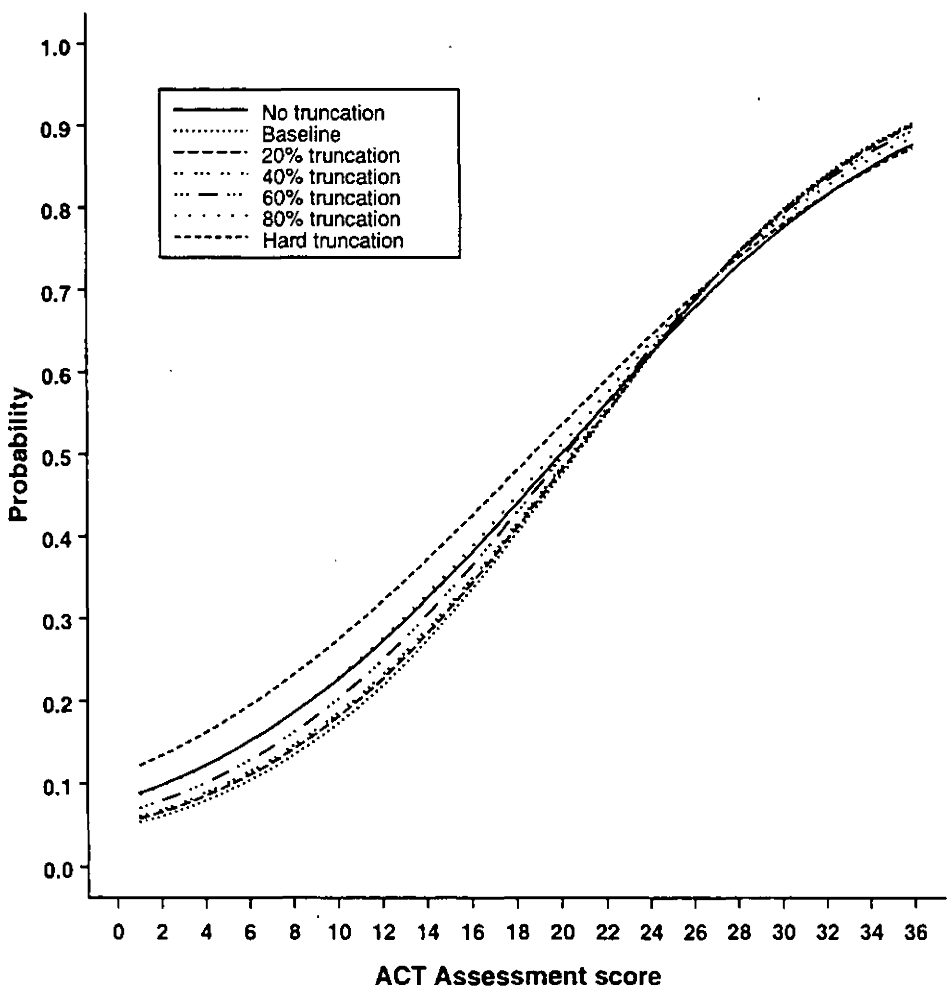


The results for Placement Group 2 (steep slope, medium skewness, n=500), shown in Figure 3.2a, are quite different from those of the previous figure in that there is less overlap of the logistic curves for the soft truncation conditions, and $\hat{P}_N$ is typically underestimated. An unusual finding shown in this figure that differs from the findings of previous research on soft truncation is that less severe soft truncation is associated with less accurate estimates of $\hat{P}$.

For example, the baseline soft truncation probability curve is relatively far from the curve representing the no truncation condition; in comparison, the 80% soft truncation probability curve lies almost directly on top of the no truncation curve, suggesting that the most accurate estimates of $\hat{P}_N$ were obtained under this soft truncation condition.

FIGURE 3.2a. Effects of Soft Truncation on Estimated Conditional Probability of Success

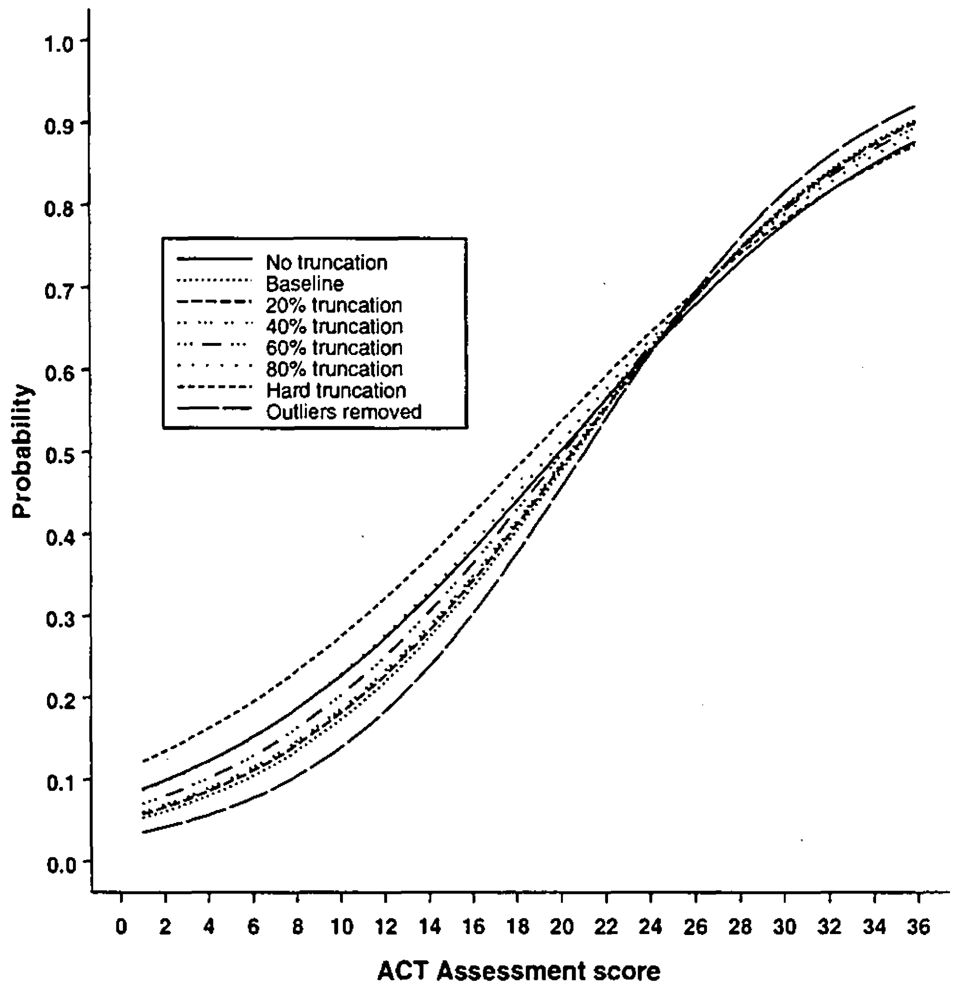
(Placement Group 2: Steep slope, medium skewness, n=500)



This apparent anomaly may be explained by the presence of influential outlying observations. A measure of the influence of the *i*th observation on the estimated regression coefficient (DFBETA) was used to detect these observations. DFBETA represents the standardized difference in the regression coefficients when the *i*th outlying observation is omitted. Figure 3.2b illustrates what happens when only those influential outliers below a score of 12 are omitted from the placement group prior to estimating the logistic function. This figure is identical to the previous figure, except that an extra logistic curve was plotted for the placement group after influential outliers (n=15) were removed from this group. Note that this logistic curve is steeper than the other curves, and that it lies relatively close to the curve representing the baseline soft truncation condition. If the "outliers removed" curve is substituted for the original $\hat{P}_N$ curve, then the relationship between soft truncation and estimate accuracy is similar to that of previous research (e.g., Schiel, 1998), in that more severe soft truncation is associated with less accurate estimates of $\hat{P}$.

FIGURE 3.2b. How Removing Outliers Increases the Slope of the Placement Group Logistic Curve

(Placement Group 2: Steep slope, medium skewness, n=500)



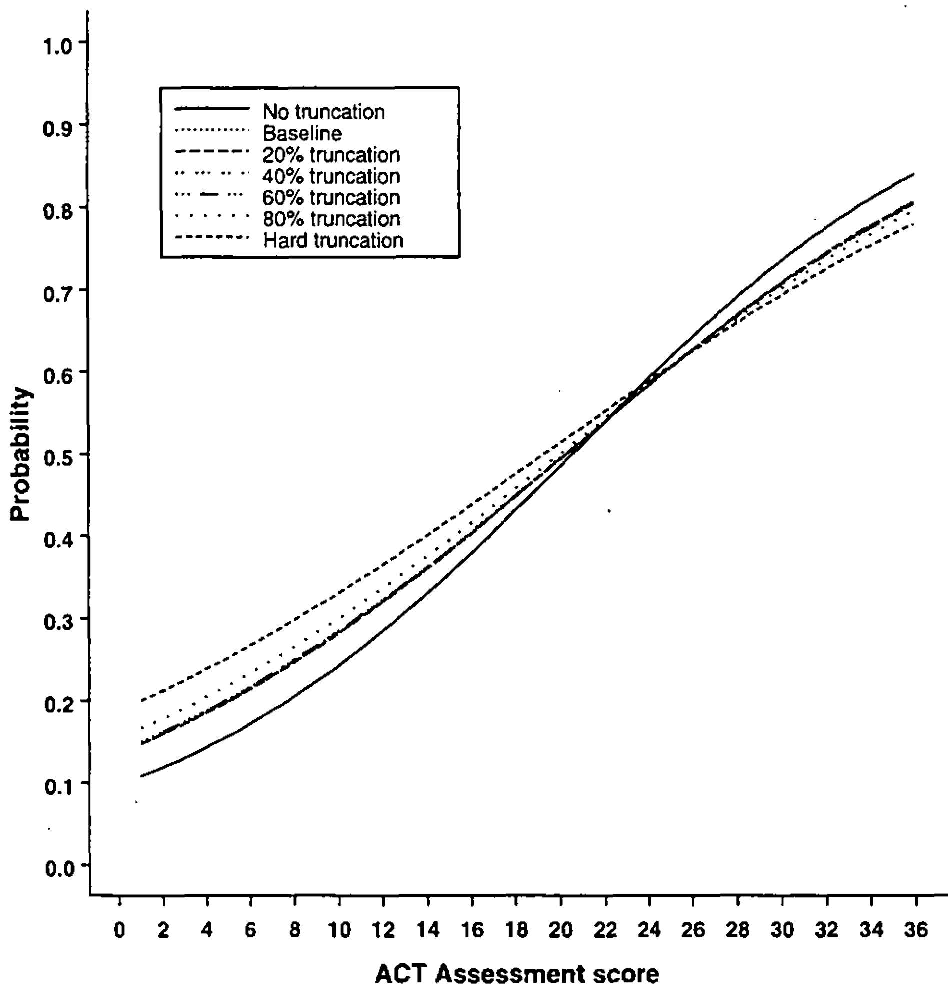
It is worth noting that, in principle, it was not necessary to remove influential outlying observations from the five truncation samples, because each of these was simulated 500 times and the results were then summarized over simulations, thereby averaging out any error due to such observations. Moreover, beginning with the baseline condition, all observations below an ACT Assessment score of 12 were consistently deleted; any influential outliers within the score interval of 1-12 would therefore always be deleted in each simulation.

Outlying observations were also removed from some of the other placement groups, for comparative purposes. As expected, this procedure typically produced a somewhat steeper $\hat{P}_N$ curve for these groups. The principal results described in this study, however, are based on $\hat{P}$ curves that were fitted to the placement group data without regard for outlying observations. This was done for the following reasons: A placement group includes *all* students for whom placement decisions must be made and thus, by definition, may contain outlying observations. Moreover, logistic regression curves are not, in practice, fit to the placement group data, but are fit only to the truncated data, thereby making outlier analyses of placement group data unnecessary.

The results for Placement Group 3 (steep slope, zero skewness, n=500; see Figure 3.3) suggest that although soft truncation has a slight effect on the accuracy of $\hat{P}$, the degree of accuracy does not decrease substantially as soft truncation increases (i.e., the logistic curves for the soft truncation conditions lie fairly close together).

FIGURE 3.3. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 3: Steep slope, zero skewness, n=500)



When the logistic regression curve is relatively flat, as it is for Placement Groups 4 and 5, reasonably accurate estimates of $\hat{P}$ can still be obtained (see Figures 3.4 and 3.5). The logistic curves in Figure 3.4 display a somewhat atypical pattern; the baseline and 20% soft truncation conditions tend to underestimate $\hat{P}$, whereas the remaining soft truncation conditions tend to overestimate $\hat{P}$.

FIGURE 3.4. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 4: Flat slope, high skewness, n=500)

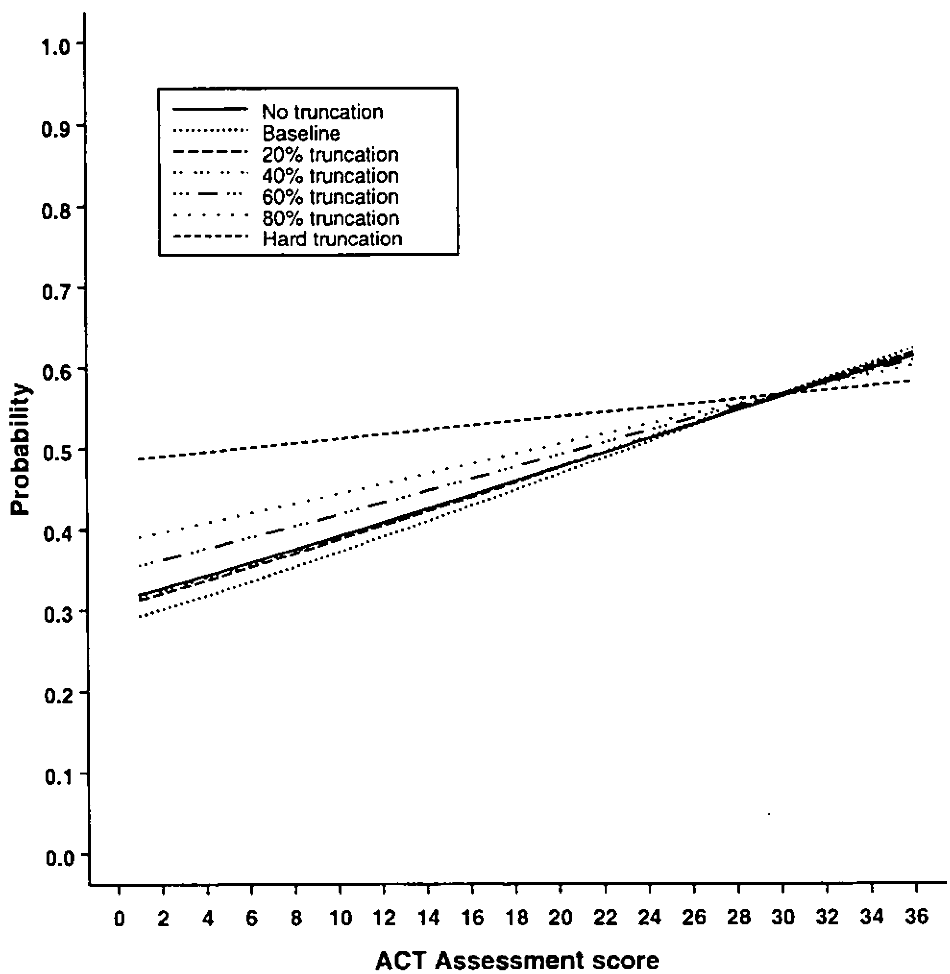
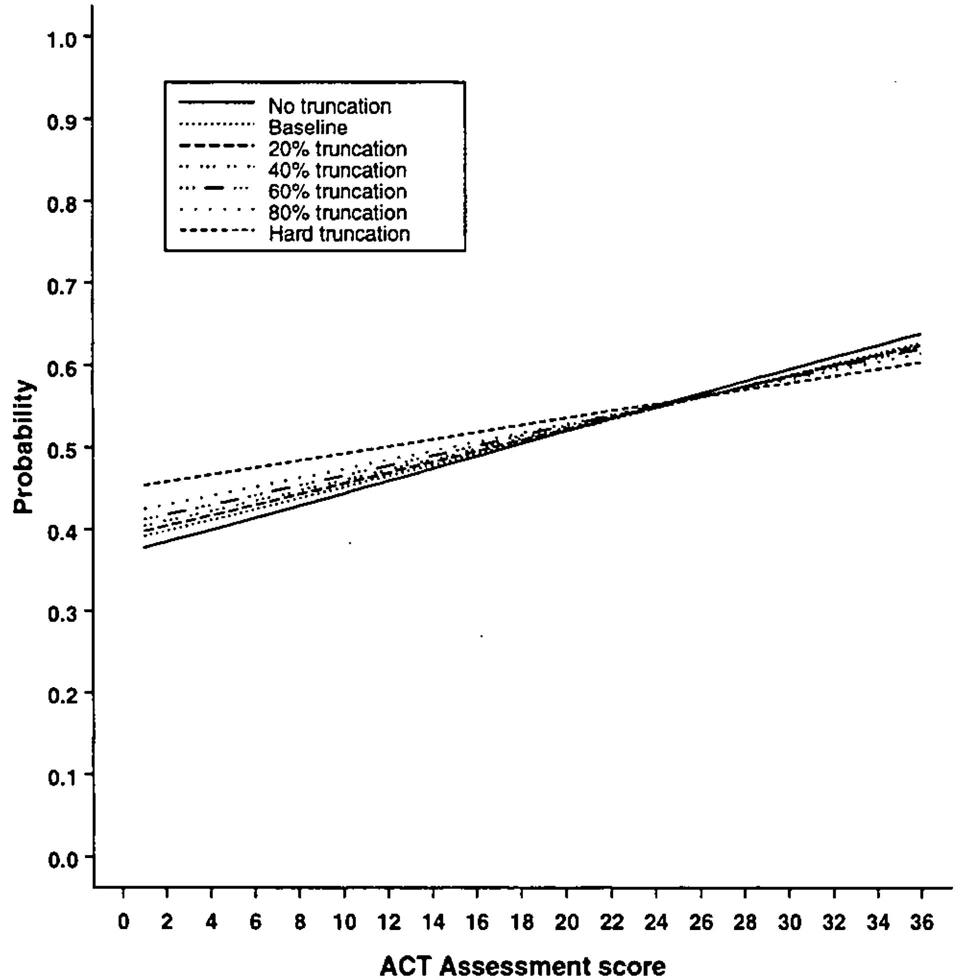


FIGURE 3.5. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 5: Flat slope, zero skewness, n=500)



Figures 3.6-3.8 illustrate that the estimates of $\hat{P}$ obtained from relatively small placement groups of size n=100 are comparable in accuracy to those of the size n=500 placement groups. Given the previous results, one would expect the logistic curves in Figure 3.6 that are associated with greater degrees of soft truncation to be progressively farther away

from the placement group logistic curve. This is not always the case, however. For example, the baseline truncation condition resulted in estimates of $\hat{P}_N$ that were slightly less accurate than those of the 40% truncation condition. Somewhat similar findings were observed by Schiel and Noble (1992), who reported that a hard truncation condition in which 78% of the observations remained in the sample yielded somewhat more accurate estimates of $\hat{P}_N$ for lower scores than did a hard truncation condition in which 89% of the observations remained. It is possible that small sample sizes are related to these particular findings. In the Schiel and Noble study, for example, respective sample sizes of 76 and 69 were analyzed for the two different hard truncation conditions. In the present study, data for the baseline and 40% truncation conditions for Group 6 consisted of 85 and 79 observations, respectively, for each of the 500 simulated truncation samples.

Figure 3.6. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 6: Steep slope, high skewness, N=100)

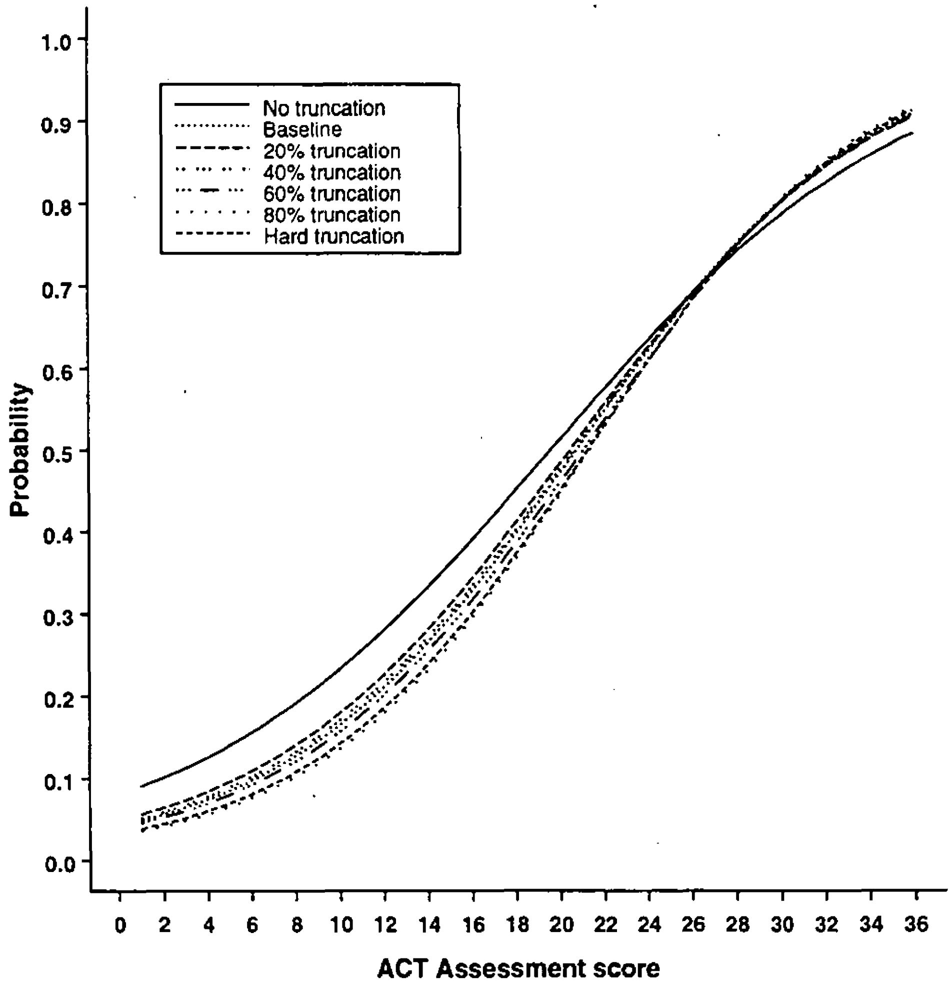


Figure 3.7. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 7: Steep slope, medium skewness, N=100)

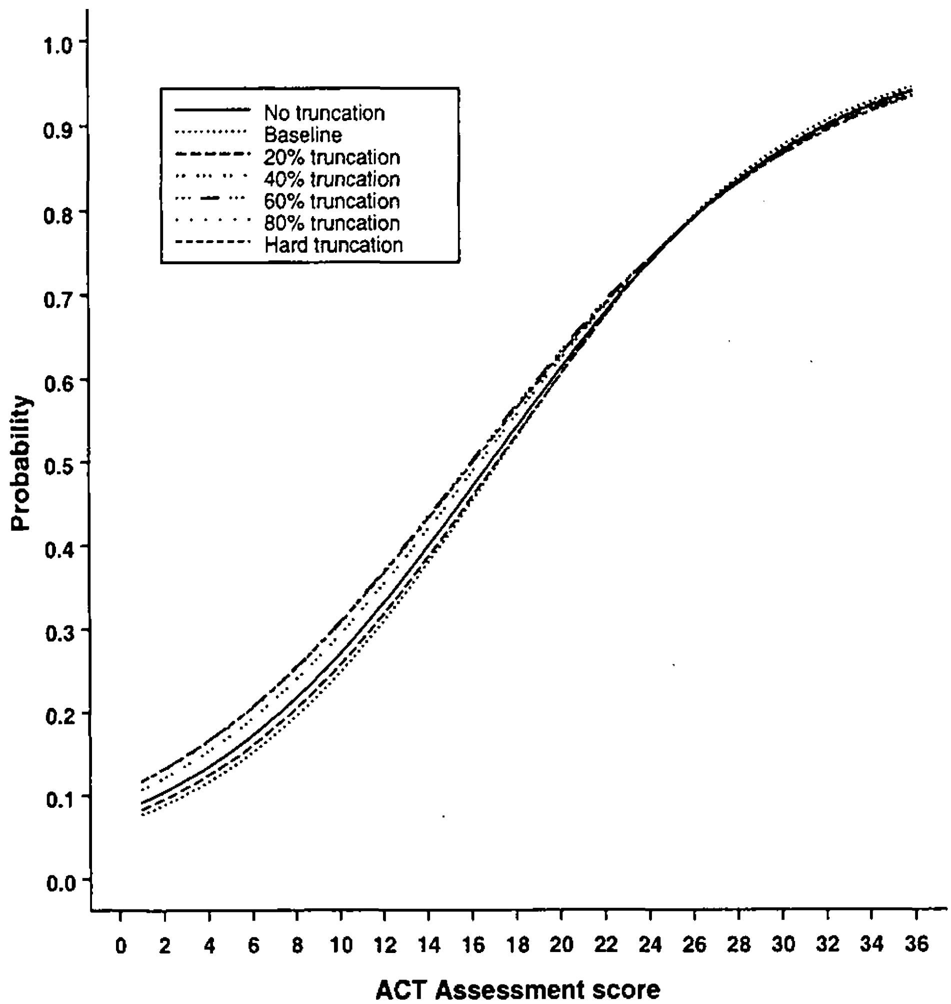
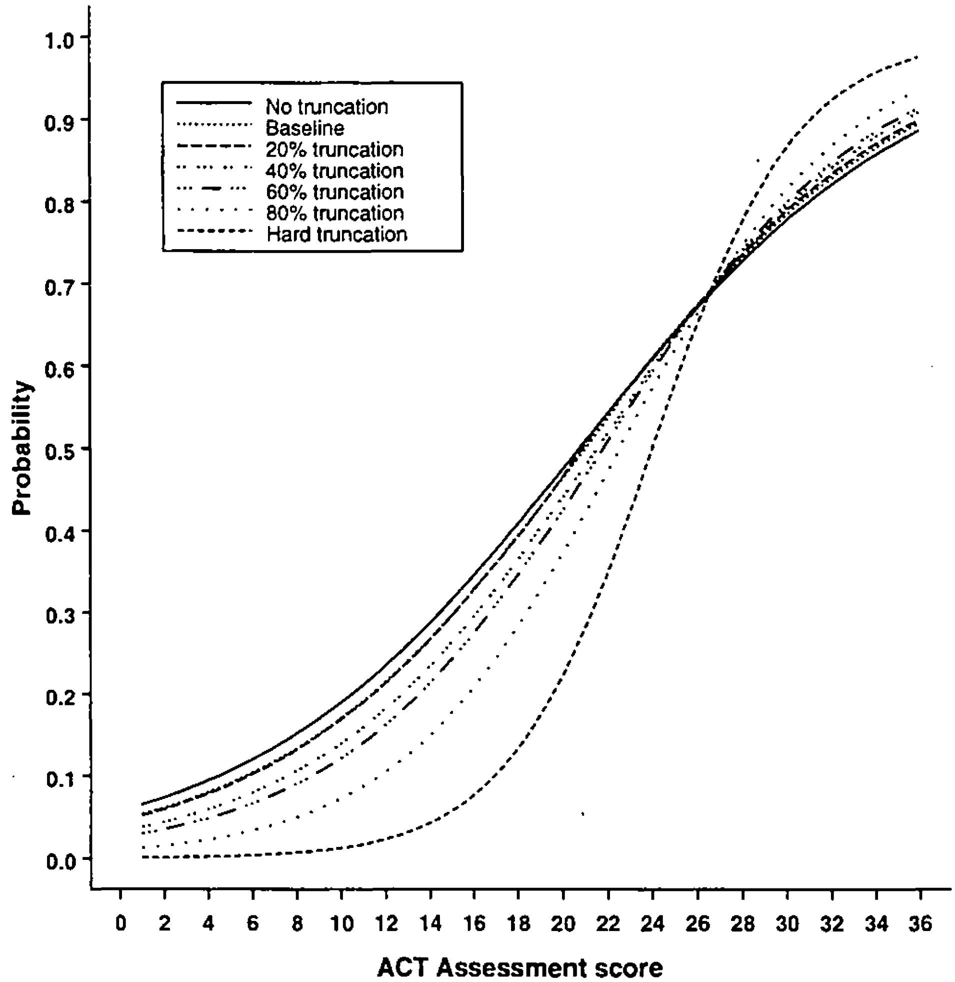


Figure 3.8. Effects of Soft Truncation on Estimated Conditional Probability of Success

(Placement Group 8: Steep slope, zero skewness, N=100)



The results shown in Figures 3.1-3.8 indicate that although reasonably accurate estimates of $\hat{P}_N$ were obtained under soft truncation, relatively less accurate estimates occurred when the sample size was fairly small and/or the distribution of the predictor variable was nearly symmetrical (e.g., Placement Groups 3, 6, and 8). There are exceptions to this, however; for example, the marginal ACT score distribution of Placement Group 5 was nearly symmetrical, yet relatively accurate estimates of $\hat{P}_N$ were obtained under all soft truncation conditions. In addition, these findings differ somewhat from those described in previous research in which soft truncation was defined differently; for example, Schiel (1998) found that relatively less accurate estimates of $\hat{P}_N$ were associated with placement groups having flat logistic curves. The estimates obtained for flat logistic curve placement groups (Groups 4 and 5) in the present study, in contrast, are among the most accurate associated with any placement group.

The results for placement groups of size n=150 were very similar to those of Placement Groups 1-8, with one exception, and are therefore not displayed in graphic or tabular form. These results are, however, described at the end of this section.

The mean values of $\Delta\hat{P}$ shown in Table 3 summarize the effect of truncation on $\hat{P}$. The column labeled $\hat{P}_N$ shows the mean $\hat{P}$, over 36 scale score points, for each placement group. The remaining columns show mean $\Delta\hat{P}$ and mean $|\Delta\hat{P}|$ for each truncation condition. For example, the first number beneath the column heading of "baseline" (.0187) is the mean $\Delta\hat{P}$ for the baseline truncation condition in Placement Group 1. This result indicates that the average difference between $\hat{P}_B$ and $\hat{P}_N$, over all scale score points, was .0187.

TABLE 3

Effects of Soft Truncation on Estimated Probability of Success, by Placement Group and Truncation Condition

Placement group	Mean $\hat{P}_N$	Mean Difference	Truncation Baseline	20%	40%	60%	80%	Hard
1: Steep slope, high skewness, n=500	.4880	$\Delta\hat{P}$	.0187	.0188	.0207	.0229	.0186	.0208
		$\lvert\Delta\hat{P}\rvert$	.0237	.0238	.0257	.0280	.0233	.0256
2: Steep slope, medium skewness, n=500	.4668	$\Delta\hat{P}$	-.0194	-.0159	-.0131	-.0052	.0069	.0286
		$\lvert\Delta\hat{P}\rvert$	.0325	.0279	.0251	.0159	.0072	.0292
3: Steep slope, zero skewness, n=500	.4566	$\Delta\hat{P}$	.0095	.0094	.0101	.0101	.0166	.0300
		$\lvert\Delta\hat{P}\rvert$	.0276	.0273	.0290	.0289	.0384	.0570
4: Flat slope, high skewness, n=500	.4648	$\Delta\hat{P}$	-.0098	-.0014	.0004	.0166	.0315	.0706
		$\lvert\Delta\hat{P}\rvert$	.0124	.0032	.0018	.0173	.0335	.0763
5: Flat slope, zero skewness, n=500	.5084	$\Delta\hat{P}$	.0013	.0035	.0057	.0085	.0119	.0209
		$\lvert\Delta\hat{P}\rvert$	.0067	.0093	.0118	.0153	.0205	.0328
6: Steep slope, high skewness, n=100	.4752	$\Delta\hat{P}$	-.0302	-.0230	-.0278	-.0370	-.0464	-.0441
		$\lvert\Delta\hat{P}\rvert$	.0401	.0313	.0364	.0457	.0579	.0542
7: Steep slope, medium skewness, n=100	.5382	$\Delta\hat{P}$	-.0085	-.0062	.0108	.0183	.0167	.0168
		$\lvert\Delta\hat{P}\rvert$	.0118	.0070	.0121	.0202	.0189	.0195
8: Steep slope, zero skewness, n=100	.4476	$\Delta\hat{P}$	-.0075	-.0067	-.0217	-.0303	-.0558	-.0995
		$\lvert\Delta\hat{P}\rvert$	.0116	.0130	.0313	.0426	.0770	.1445

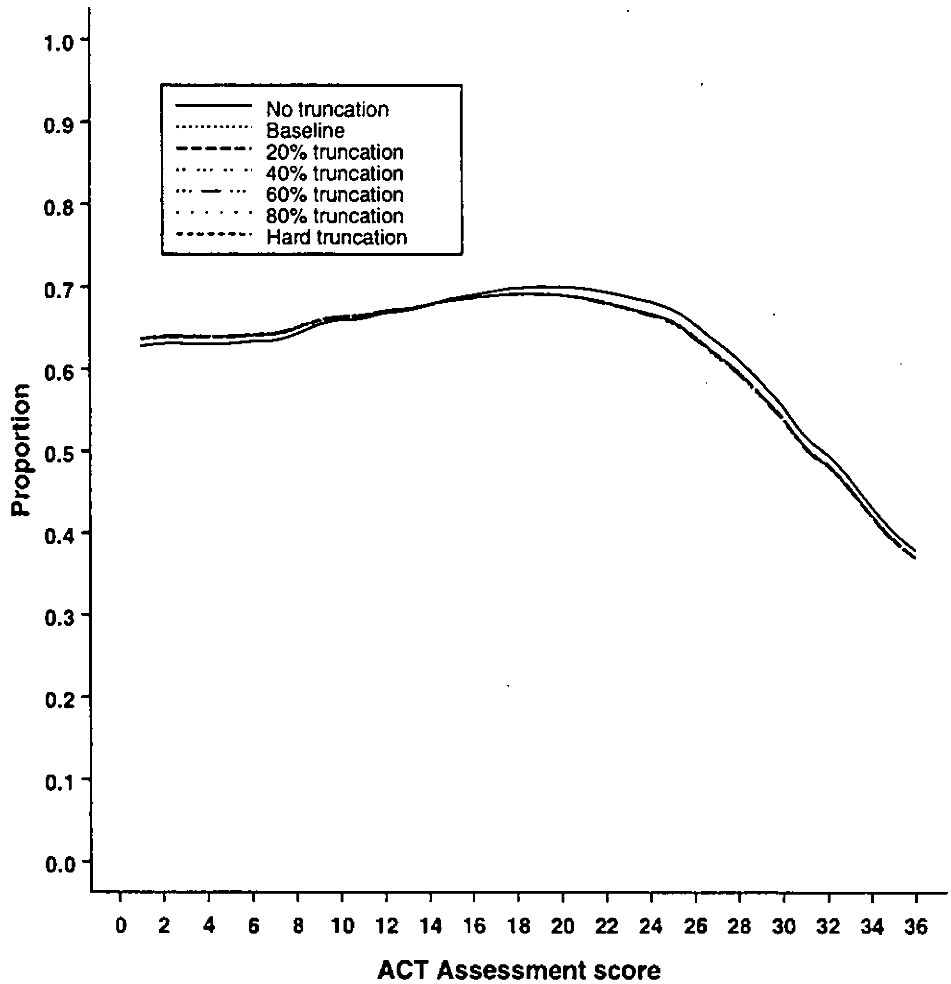
Results for the remaining truncation conditions are shown in the last five columns of Table 3. Note that the signs, positive or negative, of the $\Delta\hat{P}$ reflect whether the probabilities obtained under different truncation conditions over- or underestimated $\hat{P}_N$, respectively. The results in Table 3 show that, with respect to $\hat{P}$, the placement groups most affected by soft truncation were Groups 3 (steep slope, zero skewness, n=500), 6 (steep slope, high skewness, n=100), and 8 (steep slope, zero skewness, n=100). Mean $\lvert\Delta\hat{P}\rvert$ for these placement groups, across all soft truncation conditions, were higher than those of the other groups, ranging from .0116 (baseline truncation, Group 8) to .077 (80% truncation, Group 8).

Estimated Accuracy Rates and Optimal Cutoff Scores

Figure 4.1 illustrates that the effects of soft truncation on estimated $\hat{A}$ for Placement Group 1 were minimal. As expected, the maximum estimated $\hat{A}_N$ corresponded to an ACT Assessment score of 20, indicating that this was the optimal cutoff score. The maximum $\hat{A}$ for nearly all of the soft truncation conditions and the hard truncation condition corresponded to a score of 19. For the 60% truncation condition, the maximum $\hat{A}$ corresponded to a score of 18. For all truncation conditions except the 60% condition, the "true" optimal cutoff score (20; corresponding to the maximum $\hat{A}_N$) was therefore underestimated by one ACT scale score point.

FIGURE 4.1. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 1: Steep slope, high skewness, n=500)



Figures 4.2-4.8 illustrate the effects of soft truncation on estimated $\hat{A}$ for the remaining placement groups. As occurred for the estimates of $\hat{P}$, relatively less accurate estimates of $\hat{A}$ were obtained for Placement Groups 3, 6, and 8, which had in common either small sample sizes or virtually no skewness of the marginal distribution of the predictor

variable. Mean $\Delta\hat{A}$ and $\left|\Delta\hat{A}\right|$, which are reported in Table 4, confirm these findings. Mean $\left|\Delta\hat{A}\right|$ for these placement groups ranged from .0087 (baseline truncation, Group 8) to .0665 (80% truncation, Group 8) across soft truncation conditions. Groups 1 (steep slope, high skewness, n=500), 4 (flat slope, high skewness, n=500) and 7 (steep slope, medium skewness, n=100) yielded relatively accurate estimates of $\hat{A}_N$; $\left|\Delta\hat{A}\right|$ ranged from .0011 (20% truncation, Group 4) to .0135 (60% truncation, Group 7) across soft truncation conditions. Note that mean $\Delta\hat{S}$ and $\left|\Delta\hat{S}\right|$ are also reported in Table 4; these statistics are discussed in the following section on estimated success rates.

FIGURE 4.2. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 2: Steep slope, medium skewness, n=500)

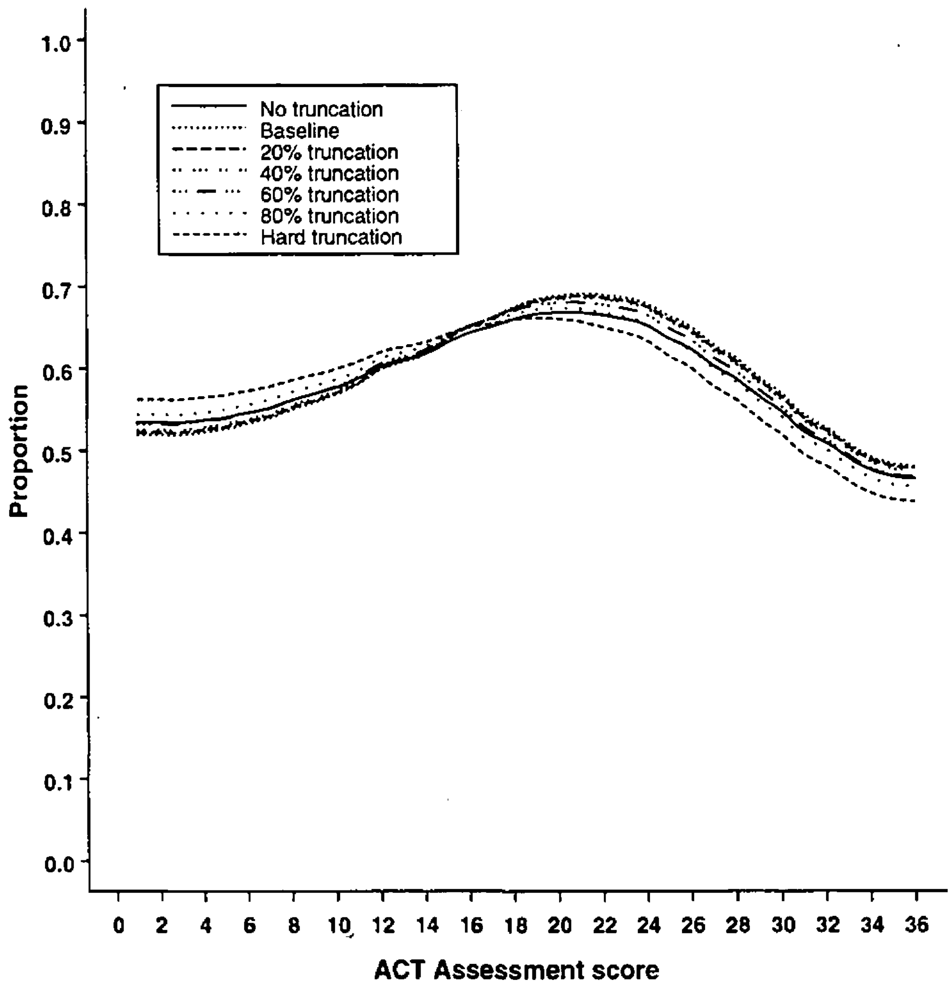


FIGURE 4.3. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 3: Steep slope, zero skewness, n=500)

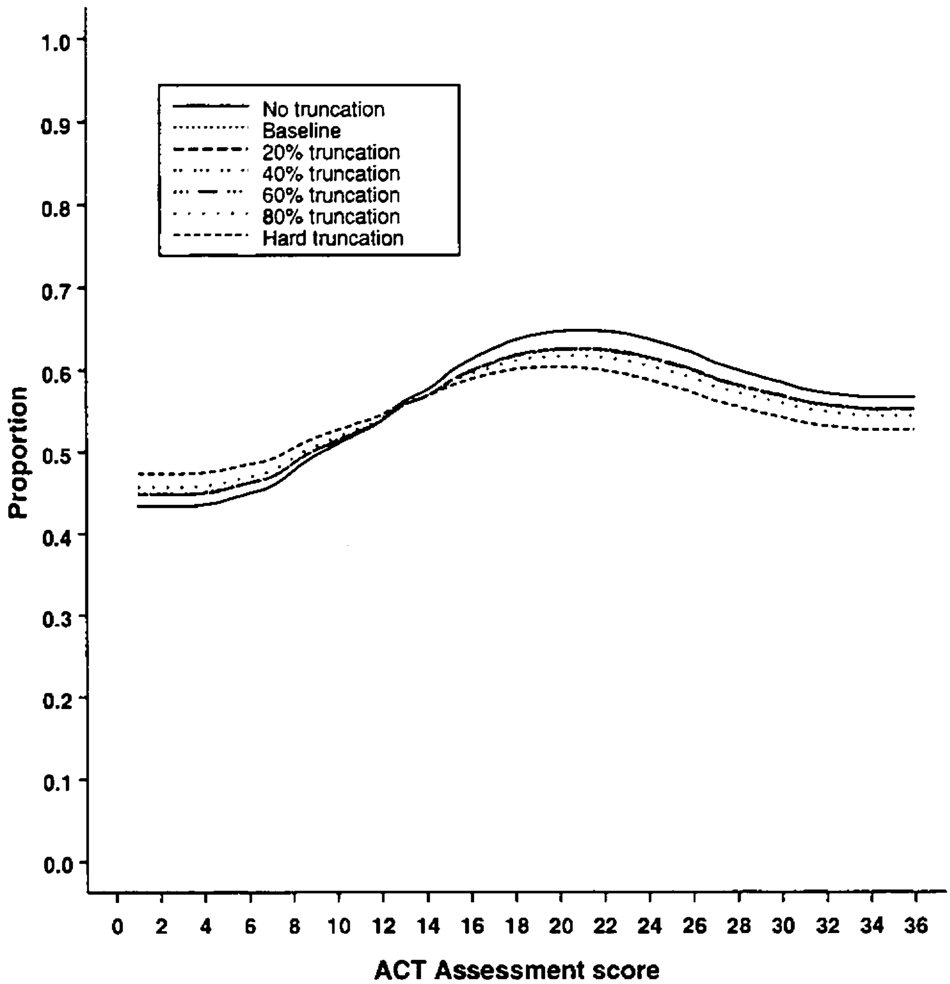


FIGURE 4.4. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 4: Flat slope, high skewness, n=500)

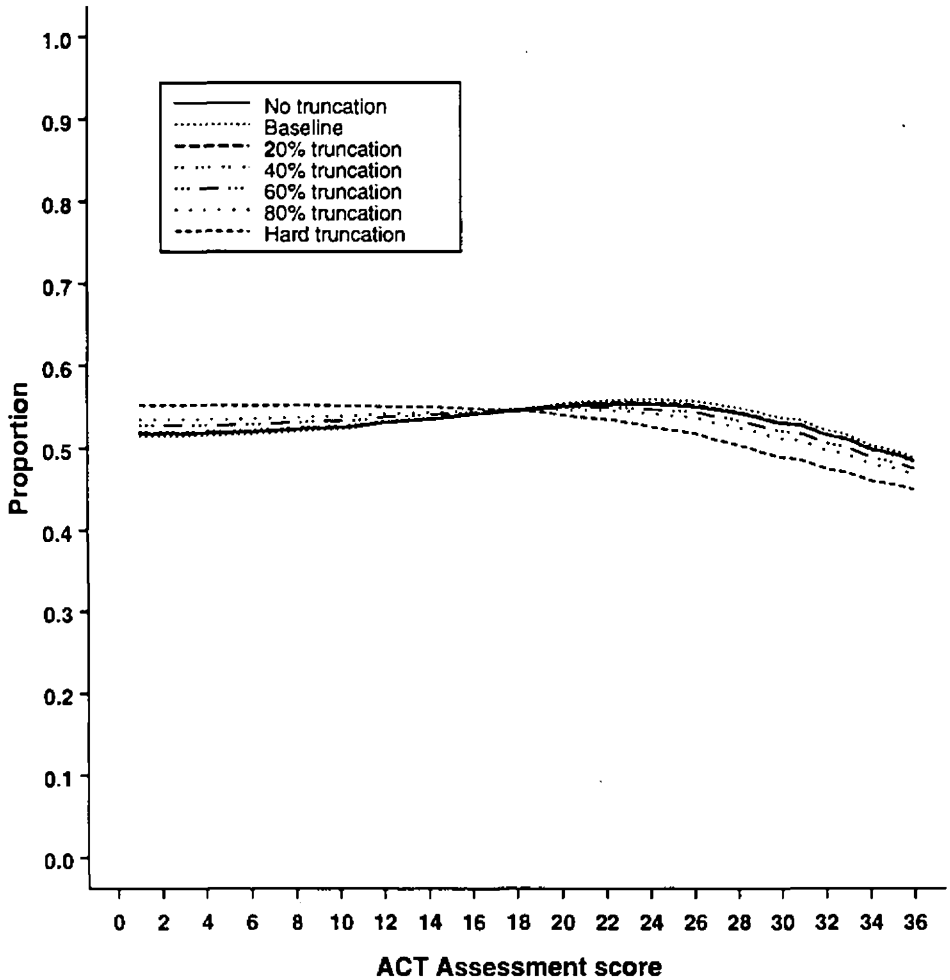


FIGURE 4.5. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 5: Flat slope, zero skewness, n=500)

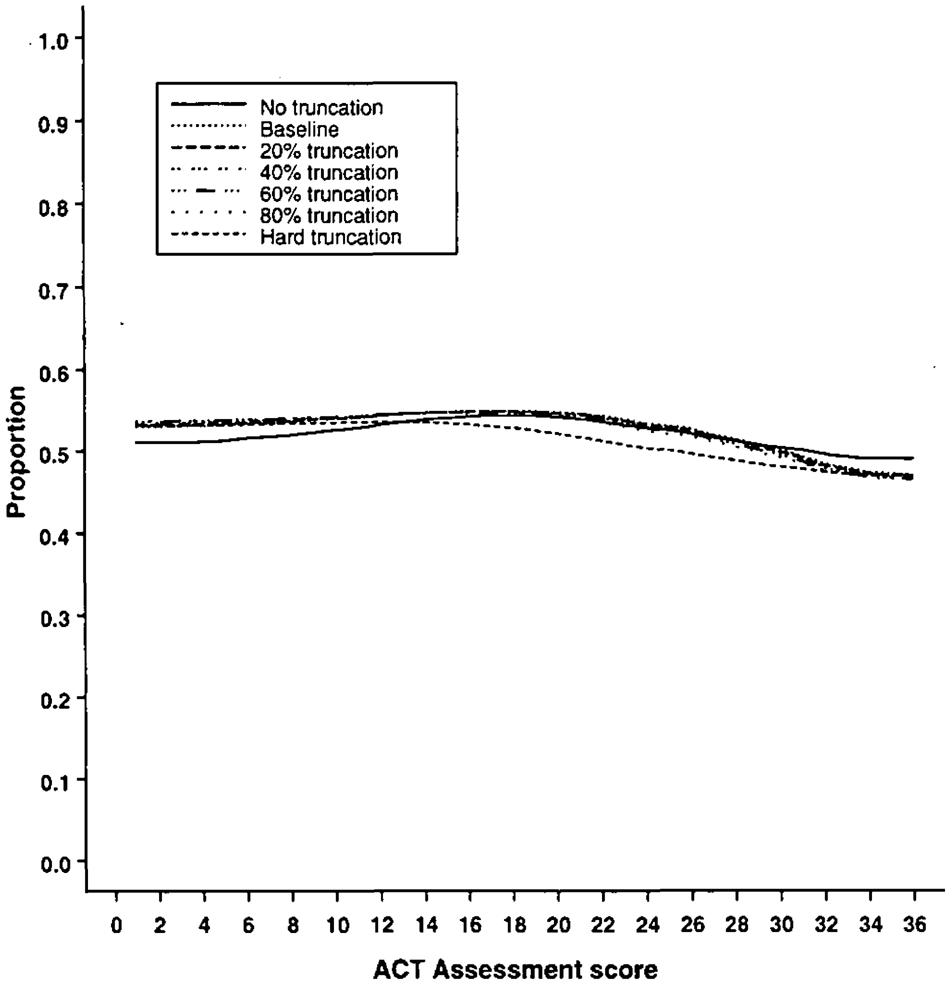


FIGURE 4.6. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 6: Steep slope, high skewness, n=100)

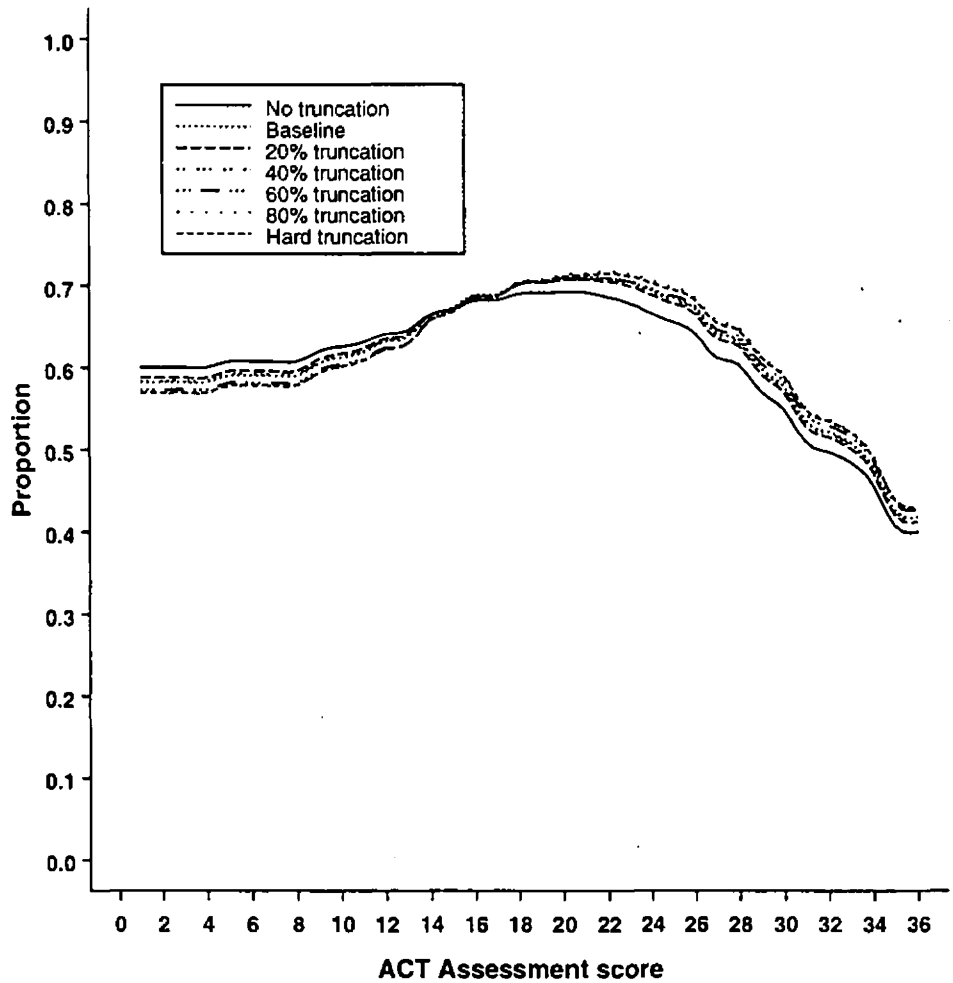


FIGURE 4.7. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 7: Steep slope, medium skewness, n=100)

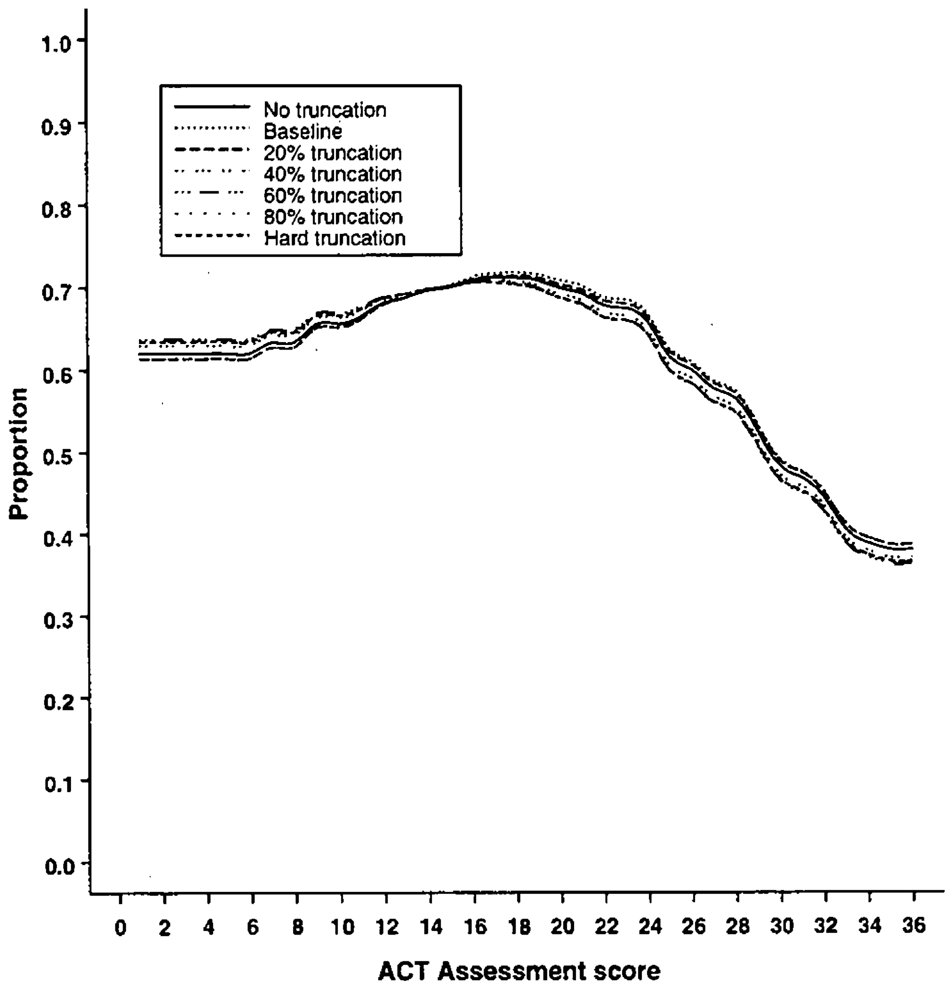


FIGURE 4.8. Effects of Soft Truncation on Estimated Accuracy Rate

(Placement Group 8: Steep slope, zero skewness, n=100)

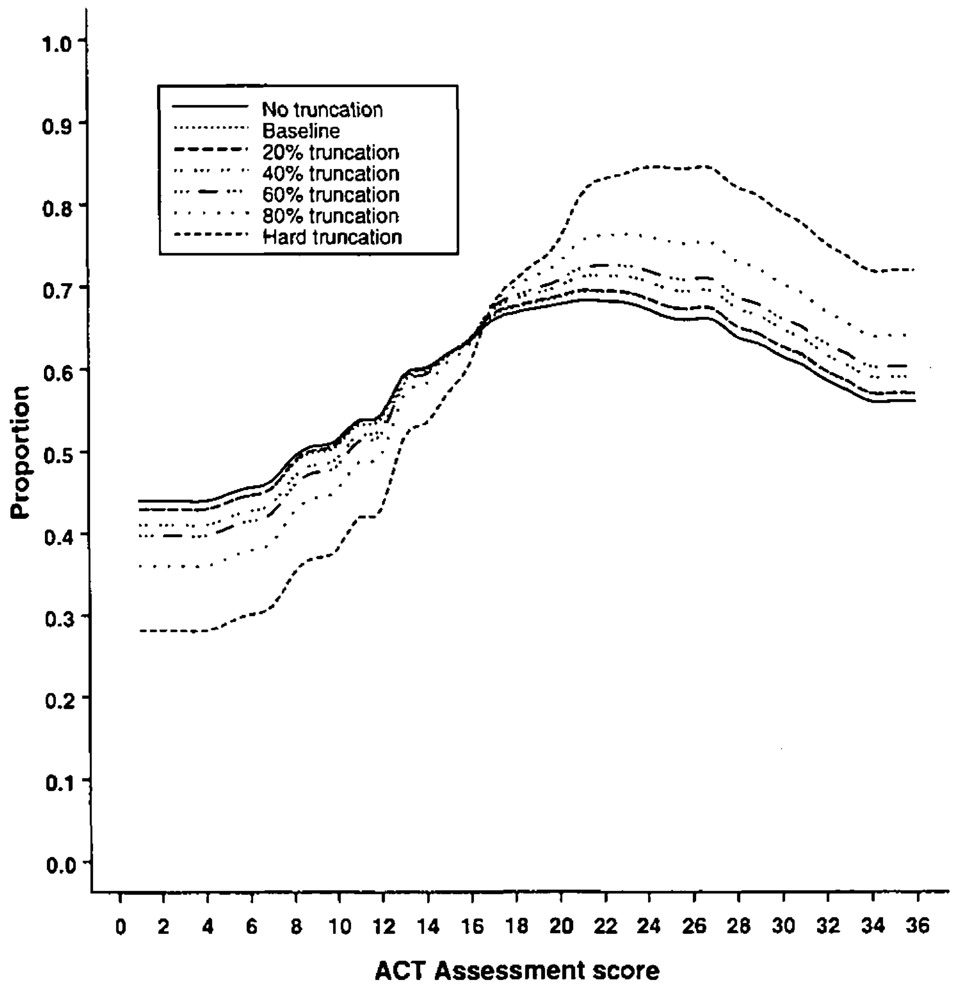


TABLE 4

Effects of Soft Truncation on Estimated Accuracy Rate and Success Rate, by Placement Group and Truncation Condition

	Mean			Truncation					
Placement group	$\hat{A}_N$	$\hat{S}_N$	Difference	Baseline	20%	40%	60%	80%	Hard
1: Steep slope, high skewness, n=500	.6225	.7241	$\Delta\hat{A}$	-.0046	-.0046	-.0046	-.0052	-.0047	-.0047
			$\lvert\Delta\hat{A}\rvert$	.0089	.0089	.0099	.0103	.0085	.0096
			$\Delta\hat{S}$	-.0010	-.0010	-.0003	-.0005	-.0012	-.0003
			$\lvert\Delta\hat{S}\rvert$	.0071	.0070	.0077	.0077	.0064	.0073
2: Steep slope, medium skewness, n=500	.5822	.6625	$\Delta\hat{A}$	.0079	.0069	.0068	.0052	.0031	-.0011
			$\lvert\Delta\hat{A}\rvert$	.0159	.0132	.0110	.0057	.0072	.0206
			$\Delta\hat{S}$	.0028	.0035	.0051	.0072	.0110	.0164
			$\lvert\Delta\hat{S}\rvert$	.0143	.0126	.0110	.0080	.0110	.0169
3: Steep slope, zero skewness, n=500	.5597	.5738	$\Delta\hat{A}$	-.0079	-.0079	-.0084	-.0084	-.0108	-.0155
			$\lvert\Delta\hat{A}\rvert$	.0145	.0143	.0154	.0148	.0217	.0349
			$\Delta\hat{S}$	-.0045	-.0045	-.0045	-.0050	-.0025	.0028
			$\lvert\Delta\hat{S}\rvert$	.0149	.0146	.0156	.0150	.0198	.0291
4: Flat slope, high skewness, n=500	.5307	.5495	$\Delta\hat{A}$	.0021	.0011	.0012	.0001	-.0008	-.0036
			$\lvert\Delta\hat{A}\rvert$	.0036	.0011	.0016	.0070	.0125	.0268
			$\Delta\hat{S}$	.0012	.0018	.0025	.0048	.0073	.0127
			$\lvert\Delta\hat{S}\rvert$	.0033	.0018	.0025	.0059	.0103	.0218
5: Flat slope, zero skewness, n=500	.5197	.5530	$\Delta\hat{A}$	.0058	.0052	.0049	.0043	.0030	-.0075
			$\lvert\Delta\hat{A}\rvert$	.0106	.0110	.0116	.0122	.0130	.0173
			$\Delta\hat{S}$	.0077	.0080	.0086	.0088	.0082	.0017
			$\lvert\Delta\hat{S}\rvert$	.0107	.0118	.0129	.0139	.0153	.0144
6: Steep slope, high skewness, n=100	.6078	.7047	$\Delta\hat{A}$	.0073	.0062	.0065	.0058	.0085	.0075
			$\lvert\Delta\hat{A}\rvert$	.0184	.0136	.0165	.0234	.0289	.0274
			$\Delta\hat{S}$	-.0008	.0006	-.0010	-.0066	-.0067	-.0074
			$\lvert\Delta\hat{S}\rvert$	.0143	.0107	.0127	.0184	.0222	.0209
7: Steep slope, medium skewness, n=100	.6056	.7504	$\Delta\hat{A}$	.0029	.0012	-.0021	-.0033	-.0034	-.0036
			$\lvert\Delta\hat{A}\rvert$	.0065	.0048	.0079	.0135	.0120	.0121
			$\Delta\hat{S}$	-.0005	-.0021	.0037	.0067	.0054	.0049
			$\lvert\Delta\hat{S}\rvert$	.0047	.0031	.0051	.0089	.0078	.0079
8: Steep slope, zero skewness, n=100	.5797	.6055	$\Delta\hat{A}$	.0031	.0039	.0072	.0097	.0167	.0247
			$\lvert\Delta\hat{A}\rvert$	.0087	.0088	.0251	.0352	.0665	.1336
			$\Delta\hat{S}$	-.0022	-.0001	-.0099	-.0149	-.0307	-.0670
			$\lvert\Delta\hat{S}\rvert$	.0066	.0072	.0202	.0280	.0526	.1153

Table 5 summarizes, for each placement group, the effect of truncation on estimated optimal cutoff scores. The estimated optimal cutoff score (corresponding to the maximum $\hat{A}$) is shown by placement group and truncation condition. Corresponding estimated $\hat{P}$ and $\hat{S}$ are also shown. For Placement Group 1, for example, the value of $\hat{A}$ was maximized at a score of 20 when no truncation was present. The maximum $\hat{A}_N$ was .70079; $\hat{P}_N$ and $\hat{S}_N$ were .52694 and .72336, respectively. Under the baseline truncation condition, the optimal cutoff score (19) was slightly underestimated and corresponded to a maximum $\hat{A}_B$ of .69196 (recall that this statistic is a median calculated across 500 truncation samples). Identical estimates of the optimal cutoff score occurred for Placement Group 1 under the remaining truncation conditions, except for the 60% condition.

TABLE 5

How Truncation Affects the Estimation of Optimal Cutoff Scores, by Placement Group

Placement group	Truncation	Optimal cutoff score	$\hat{P}$	$Max(\hat{A})$	$\hat{S}$
1: Steep slope, high skewness, n=500	None	20	.52694	.70079	.72336
	Baseline	19	.52152	.69196	.71502
	20%	19	.52147	.69193	.71516
	40%	19	.52360	.69167	.71605
	60%	18	.50098	.69141	.70746
	80%	19	.52104	.69225	.71484
	Hard	19	.52384	.69192	.71604
2: Steep slope, medium skewness, n=500	None	20	.50027	.66771	.66753
	Baseline	21	.51151	.69069	.68952
	20%	21	.51560	.68771	.69015
	40%	21	.51961	.68603	.69102
	60%	21	.52855	.68043	.69245
	80%	20	.51092	.67281	.67954
	Hard	19	.50781	.66076	.67338
3: Steep slope, zero skewness, n=500	None	21	.51100	.64808	.61274
	Baseline	21	.51623	.62611	.60198
	20%	21	.51552	.62596	.60193
	40%	21	.51631	.62484	.60212
	60%	21	.51482	.62497	.60121
	80%	21	.52091	.61705	.60240
	Hard	20	.51309	.60351	.59455
4: Flat slope, high skewness, n=500	None	23	.50284	.55357	.55422
	Baseline	24	.50765	.55911	.56127
	20%	23	.50269	.55567	.55640
	40%	23	.50456	.55513	.55694
	60%	21	.49921	.55012	.55355
	80%	20	.50541	.54730	.55382
	Hard	6	.50149	.55237	.55276
5: Flat slope, zero skewness, n=500	None	18	.50475	.54281	.54632
	Baseline	18	.50604	.54833	.55629
	20%	17	.50226	.54757	.55355
	40%	17	.50480	.54713	.55431
	60%	16	.50217	.54646	.55142
	80%	15	.50151	.54549	.55037
	Hard	12	.50133	.53478	.53842
6: Steep slope, high skewness, n=100	None	20	.51217	.69248	.70311
	Baseline	21	.50876	.71035	.70883
	20%	21	.51872	.70745	.70747
	40%	21	.51220	.70875	.70720
	60%	22	.53463	.70849	.72861
	80%	22	.52590	.71705	.72668
	Hard	22	.52816	.71425	.72513

(continued)

Placement group	Truncation	Optimal cutoff score	$\hat{P}$	$Max(\hat{A})$	$\hat{S}$
7: Steep slope, medium skewness, n=100	None	17	.50354	.71161	.73686
	Baseline	18	.52722	.71779	.75088
	20%	18	.52981	.71362	.75158
	40%	17	.52041	.70843	.74079
	60%	17	.53264	.70752	.74394
	80%	17	.52924	.70638	.74326
	Hard	17	.52936	.70575	.74286
8: Steep slope, zero skewness, n=100	None	21	.50816	.68288	.66605
	Baseline	22	.53552	.69360	.69291
	20%	21	.50458	.69505	.66713
	40%	22	.51697	.71253	.69278
	60%	22	.50753	.72454	.69262
	80%	23	.52195	.76302	.70114
	Hard	25	.57433	.84467	.81537

Placement Groups 1, 2, and 3 had relatively accurate estimated optimal cutoff scores across all soft truncation conditions. For example, all cutoffs estimated under soft truncation for Placement Group 3 (steep slope, zero skewness, n=500) were equivalent to the cutoff of the non-truncated placement group (21).

The largest difference between the optimal cutoff score estimated for a (non-truncated) placement group and one estimated for any soft truncation condition occurred for Placement Groups 4 (flat slope, high skewness, n=500) and 5 (flat slope, zero skewness, n=500). Three-point underestimates were obtained for both of these groups under the 80% truncation condition. Interestingly, optimal cutoff scores estimated under soft truncation for Placement Groups 3, 6, and 8 were, in most instances, reasonably accurate (within one scale score point of the optimal cutoff for the placement group) and comparable to those of the other placement groups, even though estimated $\hat{P}$ and $\hat{A}$ for these three groups were relatively inaccurate. Some exceptions occurred; for example, optimal cutoff scores were overestimated by two scale score points in the 60% and 80% soft truncation conditions for Group 6, and in the 80% condition for Group 8. Across all soft truncation conditions, the accuracy of estimated optimal cutoff scores was comparable to that observed by Schiel (1998), suggesting that reasonably accurate estimates can be obtained even under the alternative, more restrictive definition of soft truncation employed in the present study.

Hard truncation, as noted previously, can substantially affect the accuracy of estimated optimal cutoff scores. For example, the optimal cutoff estimated for Placement Group 4 (flat slope, high skewness, n=500) was 23; under hard truncation, this estimate decreased to a scale score of 6. Similarly, a six-point underestimate and a four-point overestimate of optimal cutoff scores were obtained for Placement Groups 5 (flat slope, zero skewness, n=500) and 8 (steep slope, zero skewness, n=100), respectively, under the hard truncation condition.

Estimated Success Rates

Table 4 contains mean $\Delta\hat{S}$ and mean $|\Delta\hat{S}|$, by placement group and truncation condition. The relative accuracy of $\hat{S}$, as measured by these statistics, followed a similar pattern to that of the $\hat{A}$; Placement Groups 1, 4, and 7 yielded the most accurate estimates of $\hat{S}$ across soft truncation conditions, whereas Groups 3, 6, and 8 yielded the least accurate estimates of this statistic. $|\Delta\hat{S}|$ ranged from .0018 (20% truncation, Group 4) to .0103 (80% truncation, Group 4) across soft truncation conditions for the former three groups, and ranged from .0066 (baseline truncation, Group 8) to .0526 (80% truncation, Group 8) across soft truncation conditions for the latter three groups. The effect of soft truncation on estimated $\hat{S}$ is displayed graphically, by placement group, in Figures A.1-A.8 in the appendix.

Placement Groups of Size N=150

Results for two of the placement groups (Groups 9 and 11) of size n=150 were very similar to the results observed for Placement Groups 1-8. For example, optimal cutoff scores for these two groups were overestimated by no more than one scale score point across all truncation conditions, including hard truncation. Results for Placement Group 10 (steep slope, medium skewness), however, differed considerably from those of all other placement groups. For example, the estimated optimal cutoff score for the no truncation condition was 18, but optimal cutoff scores for the baseline, 20%, and 40% soft truncation conditions were all considerably lower (14).

These puzzling findings occurred for two reasons: 1) a relatively small placement group was simulated, and 2) the design of this study required that this placement group be simulated only one time, as opposed to the 500 simulations performed for each truncation condition. What this means is that it is possible to simulate, on occasion, a relatively unusual joint distribution of ACT scores and course outcomes. Simulating only one placement group is not a particular concern when 500 observations are simulated, because the accuracy of estimated validity statistics, at least under hard truncation, tends to increase as sample size increases, and large-sample results tend to be relatively stable.

When an entirely new placement group meeting the same specifications of Placement Group 10 (steep slope, medium skewness, n=150) was simulated, the results of soft truncation were very different from those of the original Placement Group 10, and were more similar to those of the other placement groups. For example, estimated optimal cutoff scores under the various soft truncation conditions were all within one or two points of the optimal cutoff estimated under the non-truncated condition.

Discussion

This study extended Schiel's (1998) research by redefining soft truncation to allow for observations that are below the cutoff score, but nearer to it, to have a higher probability of being retained in the truncation sample than those observations that are below the cutoff, but farther from it. In addition, both small (n=100) and large (n=500) placement group sample sizes were investigated. The results were similar in several ways to those of previous soft truncation research. For example, they suggest that although greater degrees of soft truncation are associated with less accurate estimates of conditional probabilities of success and accuracy rates, these estimates are nonetheless acceptable and result in estimated optimal cutoff scores that typically vary by no more than one scale score point from the so-called true cutoff score (i.e., the cutoff score corresponding to the maximum accuracy rate for the non-truncated placement group).

There are, however, some ways in which the results of this study differ from those of previous soft truncation research. For example, fairly accurate estimated conditional probabilities of success obtained under soft truncation are associated in this study with placement groups having flat logistic regression curves. Schiel (1998) found, in contrast, that these particular placement groups tend to yield relatively less accurate estimates of this statistic. In addition, optimal cutoff scores were overestimated more frequently in this study than they were in previous truncation research.

Given the definition of soft truncation used in this study, reasonably accurate estimates of the conditional probability of success, accuracy rate, success rate, and optimal cutoff score can likely be obtained under 40% (beyond baseline) soft truncation and, in certain instances, even under 60% and 80% soft truncation. Furthermore, it is likely that the optimal cutoff score can be estimated to within one ACT Assessment scale score point under most soft truncation conditions, irrespective of the steepness of the logistic regression curve and the skewness of the marginal distribution of the predictor variable.

Postsecondary institutions that experience a moderate degree of soft truncation, say 20% to 60% (beyond baseline) of their respective placement groups can expect, when using logistic regression and decision theory to evaluate their course placement systems, to obtain acceptably accurate estimates of optimal cutoff scores. In comparison, institutions whose data exhibit extreme soft truncation (e.g., 80% beyond baseline) and produce a flat logistic regression curve would be well advised to exercise caution in interpreting and using their results.

Institutions whose placement group data become softly truncated in ways other than those defined in this study could find that their estimated validity statistics differ somewhat from those described here. Nevertheless, unless institutions are calculating these statistics under a fairly high degree of soft truncation or under hard truncation, any differences in accuracy that they experience are likely to be quite small and will probably have little practical effect on placement decisions.

The somewhat aberrant results exhibited by one placement group (Group 10; steep slope, medium skewness, n=150) suggest that caution should be used when interpreting results for the small placement groups. The small placement group results, while encouraging in the fact that their respective estimated validity statistics are comparable in accuracy to those of the large placement groups, should be considered preliminary until alternative methods can be used for examining the effects of soft truncation under small sample size conditions. For example, the simulation procedure in this study could be revised so that 500 placement groups are simulated for each different combination of sample size, slope, and skewness. Then, each soft truncation condition could be simulated one time for each of these 500 placement groups. Such a design would, in other words, use a large number of placement groups and a small number of simulated truncation samples; previous soft truncation research, in comparison, including that described in this study, used a small number of placement groups and a large number of truncation samples. An even stronger design would have 500-1000 placement group simulations and, for each of these, 500-1000 soft truncation simulations for each soft truncation condition. This design would, of course, require considerable computing resources. Regardless of which of these alternative designs is used, increasing the number of simulated placement groups and then summarizing the results over all of them would reduce the amount of error attributable to the placement group simulation process itself when relatively small placement group samples are involved.

References

American College Testing (1994). *ACT Assessment Course Placement Service Interpretive Guide*. Iowa City, IA: Author.

Crouse, J. (1996). *Bootstrap estimation of confidence intervals for CPS results*. Unpublished manuscript.

Houston, W. M. (1993). *Accuracy of validity indices for course placement systems*. Paper presented at the Annual Meeting of the American Educational Research Association, Atlanta, GA.

Sawyer, R. L. (1989). *Validating the use of ACT Assessment scores and high school grades for remedial course placement in college* (Research Report No. 89-4). Iowa City, IA: ACT.

Sawyer, R. L. (1996). Decision theory models for validating course placement tests. *Journal of Educational Measurement, 33*, 271-290.

Schiel, J. (1998). *Estimating conditional probabilities of success and other course placement validity statistics under soft truncation* (Research Report No. 98-2). Iowa City, IA: ACT.

Schiel, J. & Noble, J. (1992). *The effects of data truncation on estimated validity indices for course placement* (Research Report No. 92-3). Iowa City, IA: ACT.

Appendix

Effects of Soft Truncation on Estimated Success Rate, by Placement Group

FIGURE A.1. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 1: Steep slope, high skewness, N=500)

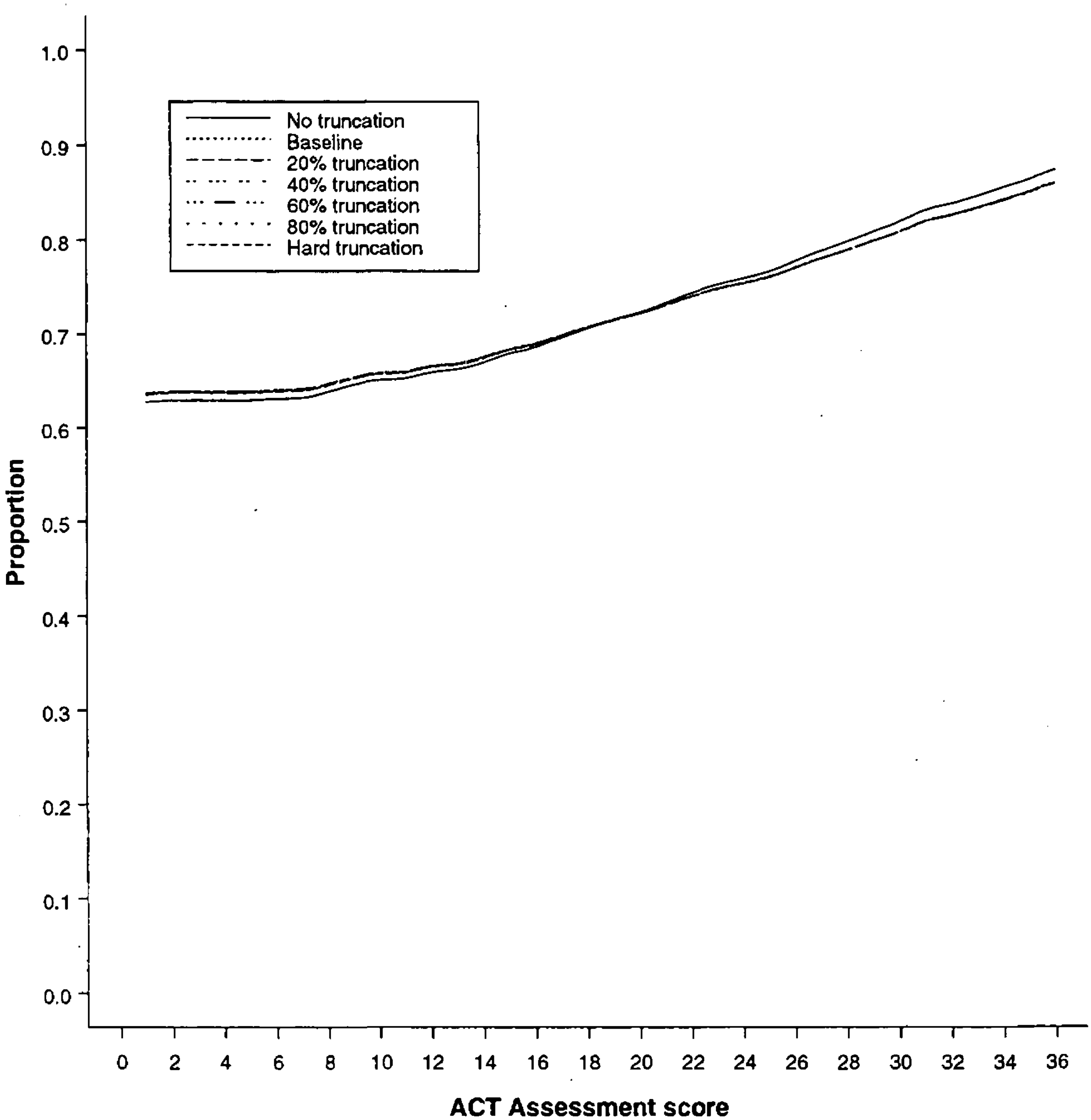


FIGURE A.2. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 2: Steep slope, medium skewness, N=500)

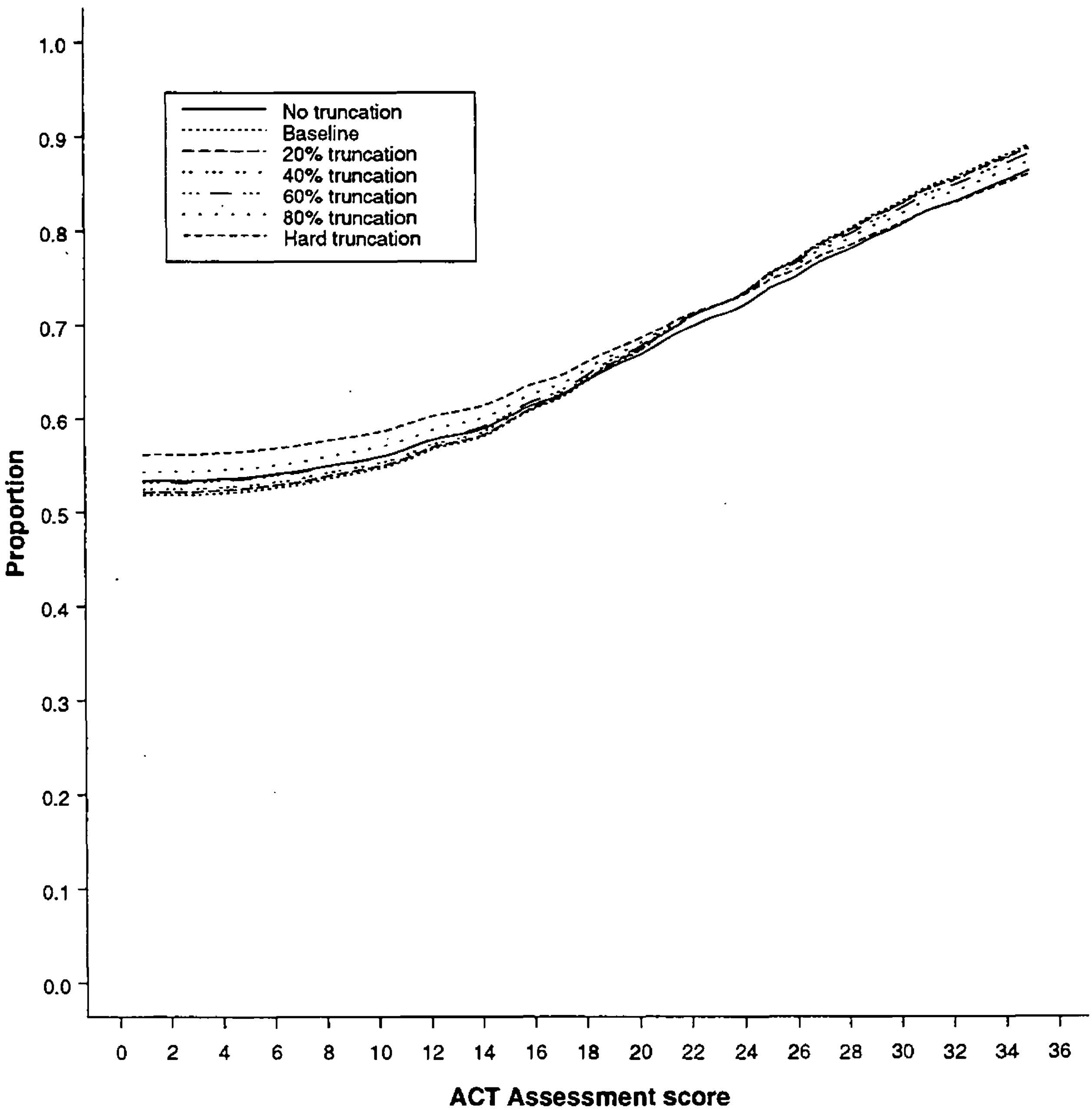


FIGURE A.3. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 3: Steep slope, zero skewness, N=500)

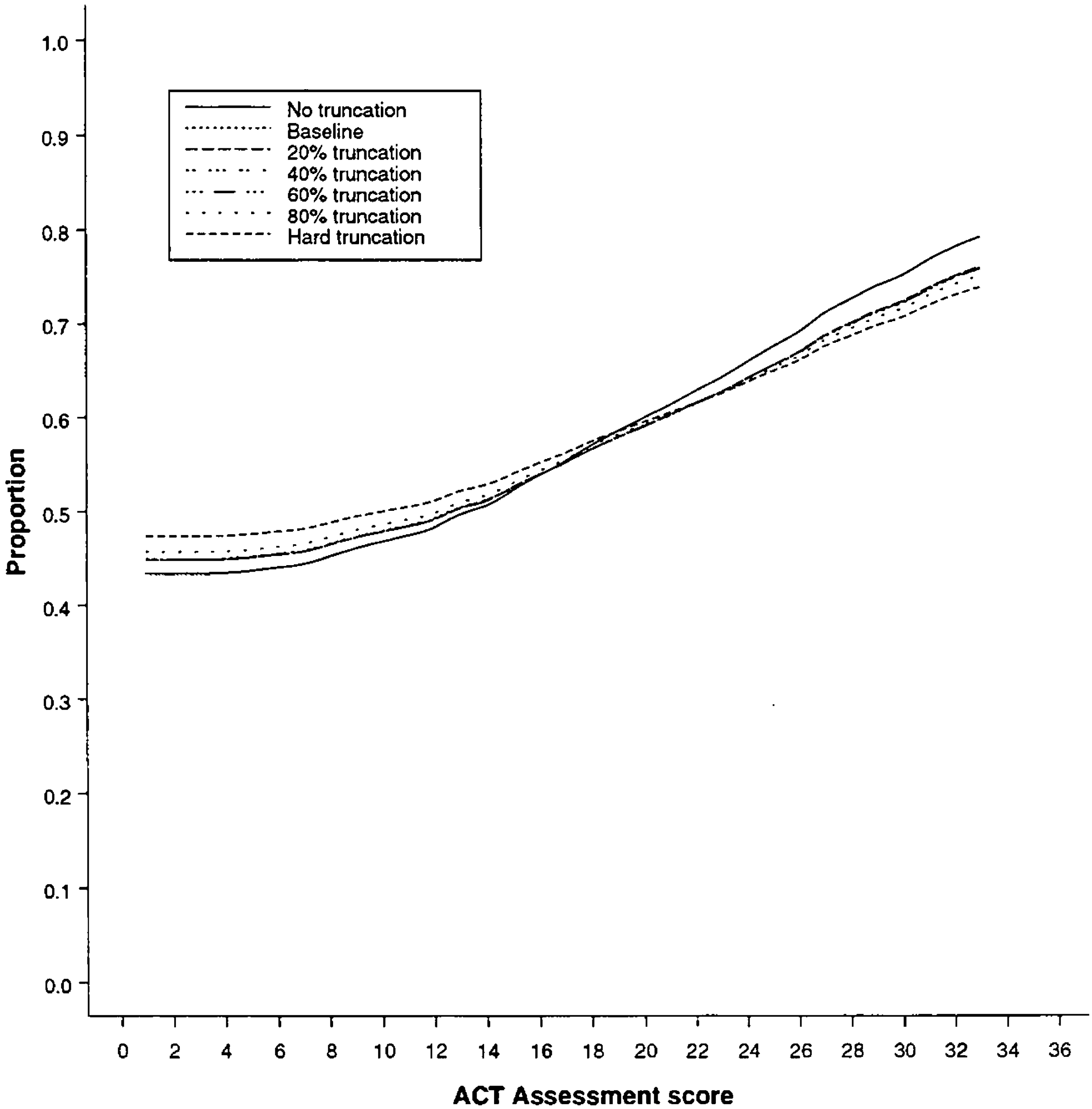


FIGURE A.4. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 4: Flat slope, high skewness, N=500)

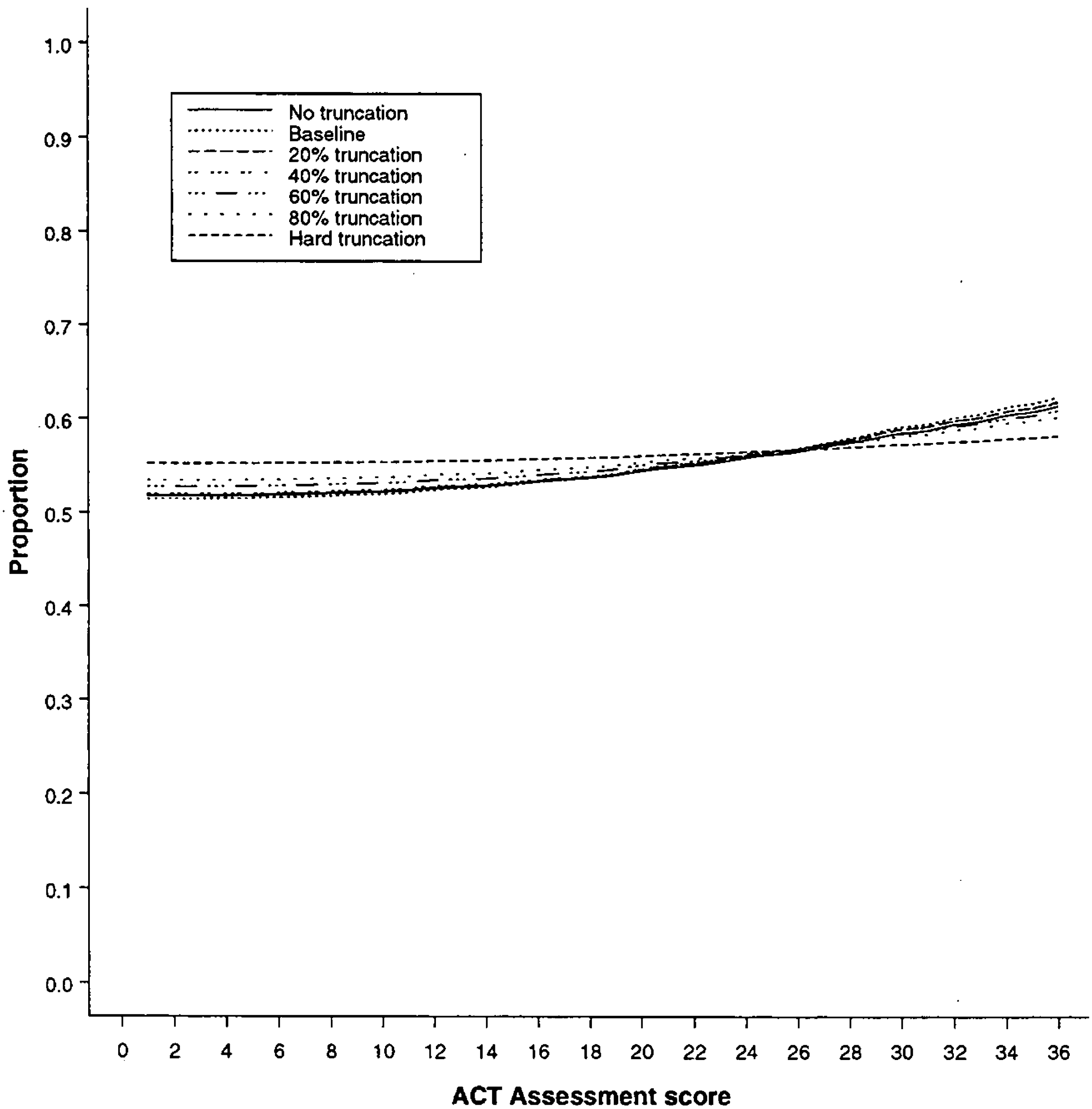


FIGURE A.5. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 5: Flat slope, zero skewness, N=500)

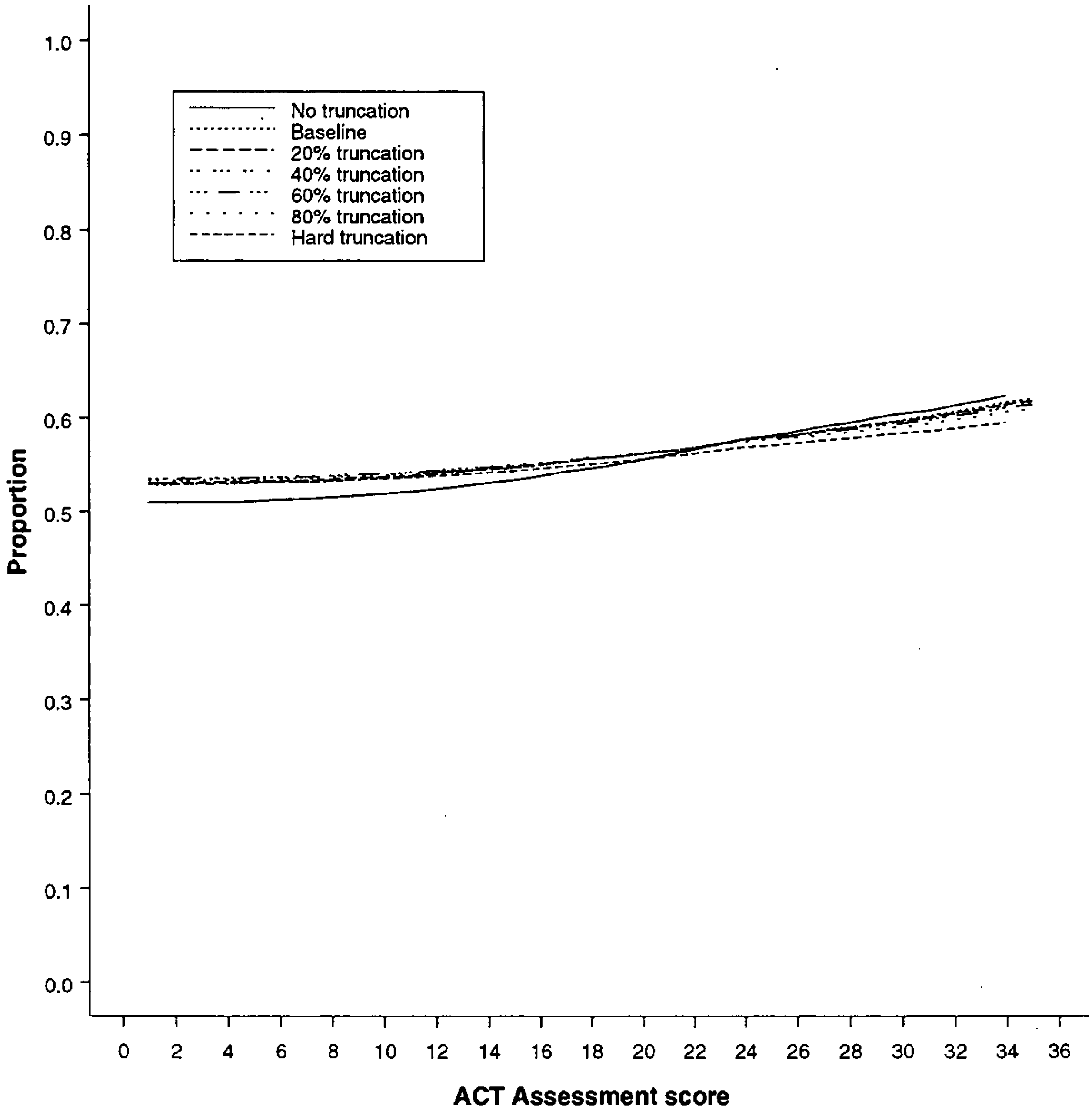


FIGURE A.6. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 6: Steep slope, high skewness, N=100)

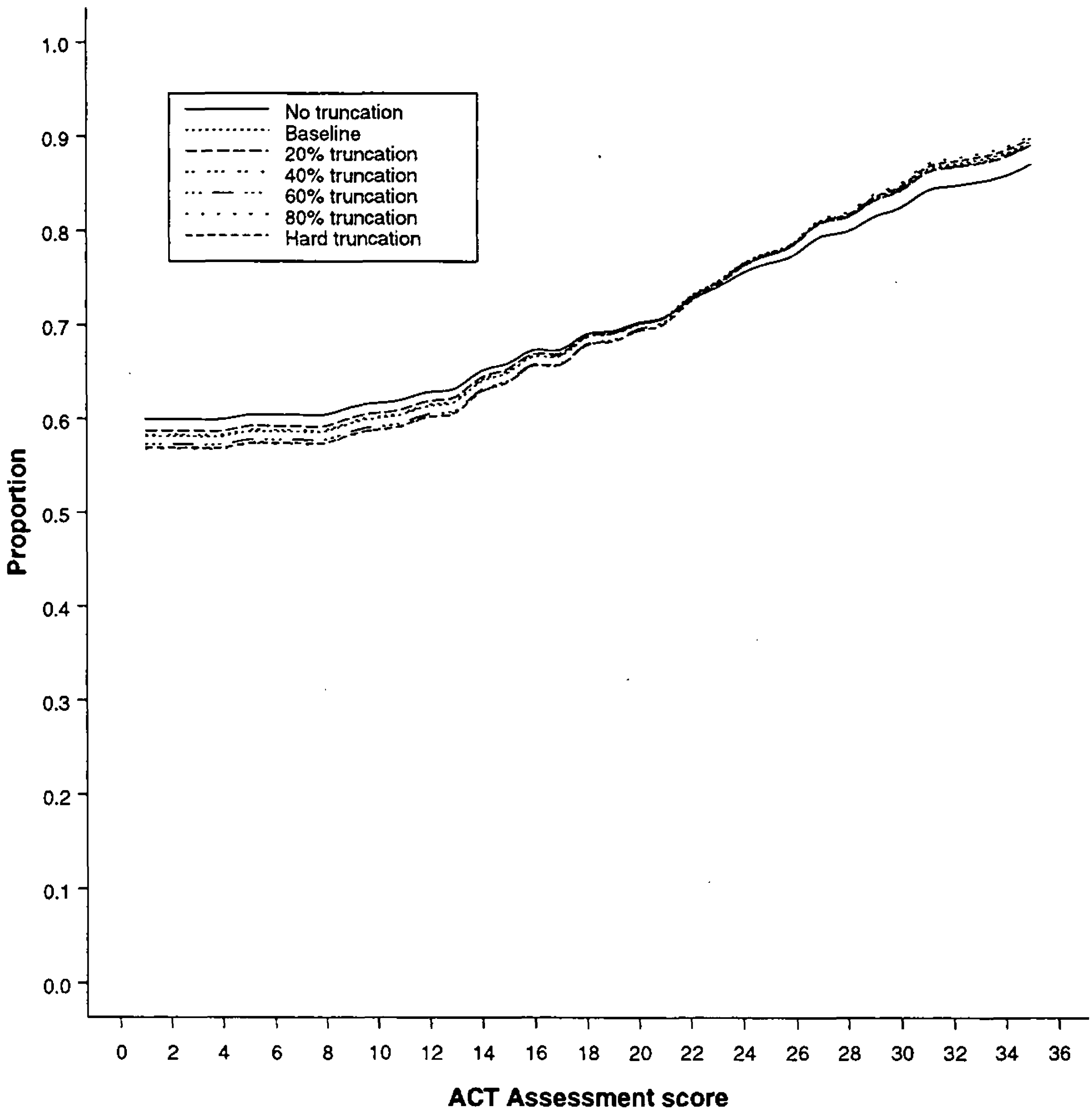


FIGURE A.7. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 7: Steep slope, medium skewness, N=100)

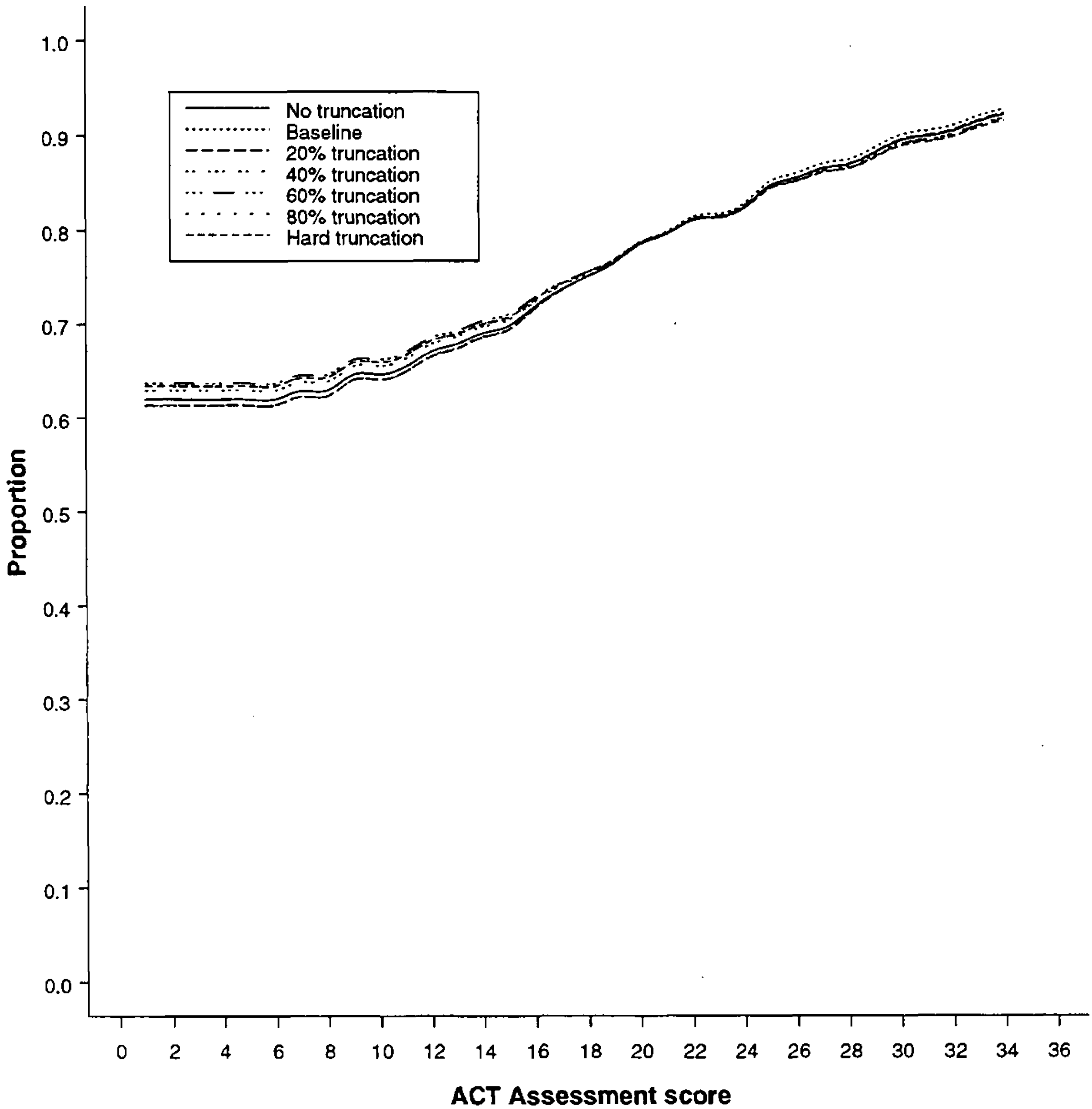


FIGURE A.8. Effects of Soft Truncation on Estimated Success Rate

(Placement Group 8: Steep slope, zero skewness, N=100)