DOCUMENT RESUME

ED 455 246 TM 030 677

AUTHOR	Tsai, Tsung-Hsun
TITLE	Estimating Minimum Sample Sizes in Random Groups Equating.
PUB DATE	1997-03-00
NOTE	67p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Chicago, IL, March 24-27, 1997).
PUB TYPE	Numerical/Quantitative Data (110) -- Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Equated Scores; Error of Measurement; *Estimation (Mathematics); *Sample Size; Science Tests
IDENTIFIERS	ACT Assessment; Equipercentile Equating; Linear Equating Method; *Randomized Experiments

ABSTRACT

The primary objective of this study was to find the smallest sample size for which equating based on a random groups design could be expected to result in less overall equating error than had no equating been conducted. Mean, linear, and equipercentile equating methods were considered. Some of the analyses presented in this paper assumed that the test scores were normally distributed. Other analyses were not based on this assumption. Real test data were used to check whether the theoretical methods provide reasonably accurate results for use in estimating sample size requirements. The science subtest of the ACT assessment provided the basic data for investigating the standard errors of equating and the minimum sample sizes needed to obtain less equating error than the identity equating. In general, as the sample size increased, the magnitude of the standard errors decreased for both forms of the test considered. In linear equating, the standard error becomes less as the raw score value approaches the mean score. In equipercentile equating, with nonnormality assumptions, raw scores less than or equal to 10 are associated with greater standard errors but the standard errors become smaller as the raw score approaches the middle score. Based on these results, it is reasonable to conclude that standard errors become less as sample size increases, and that they tend to be less for middle scores than the extreme scores for both the linear and equipercentile methods. (Contains 14 tables, 11 figures, and 12 references.) (Author/SLD)

Estimating Minimum Sample Sizes
in
Random Groups Equating

Tsung-Hsun Tsai
The University of Iowa

A Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, Illinois (USA) March, 1997. The author gratefully acknowledge helpful comments and advice from Professor Robert A. Forsyth, Dr. Michael J. Kolen, and Professor Hiram D. Hoover. Copies of this paper may be obtained by writing to the author at 218 A S. Lindquist Center, The University of Iowa, Iowa City, IA 52242

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract

The primary objective of this study was to find the smallest sample size for which equating based on a random groups design is expected to result in less overall equating error than had no equating been conducted. Mean, linear, and equipercentile equating methods were considered. Some of the analyses presented in this paper assumed that the test scores are normally distributed. Other analyses are not based on this assumption. Real test data were used to check whether the theoretical methods provide reasonably accurate results for use in estimating sample size requirements.

I.Introduction

Background

The comparability of scores derived from different tests measuring the same achievement trait, or ability, is an important concern in educational measurement. The primary purpose of equating is to achieve this comparability. Alternative forms of a test are usually constructed from the same content specifications and statistical specifications so that the item statistics for the forms are relatively similiar. Equating procedures are used to adjust test scores on alternative test forms that are somewhat different in difficulty so that the test scores on these forms can be used interchangeably for a specified population. The examinees used to accomplish the equating should be representative of the population that will be using the tests.

Equating studies frequently use three types of designs: (1) random groups; (2) single group with counterbalancing; and (3) common-item nonequivalent groups. In the random groups design, examinees are randomly assigned the form by a spiraling process. When this process is used, the first examinee takes Form X, the second examinee takes Form Y, the third examinee takes Form X, and so on. This spiraling process leads to two randomly equivalent groups. In this design, differences in the test scores on the alternative forms are attributed to the differences in difficulty of the alternative forms.

If practice or fatigue factors and order effects which may confound the differences between test scores on alternative forms can be controlled, then the single group design with conterbalancing can be used. In this design, every examinee takes both Form X and Form Y. In one method of counterbalancing, one-half of the booklets are printed with Form X following Form Y and one -half of the booklets are printed with Form Y following Form X. When the test booklets are handed out, the first examinee receives Form X first, the second examinee receives Form Y first, the third receives Form X first, and so on. The first form and second form are administered using separate time limits. In this way, the examinee group receiving Form X first is randomly equivalent to the examinee group receiving Form Y first (Kolen & Brennan, 1995). Hence, the spiraling process in the single group design with counterbalancing is similar to that used with the random groups design, except that in the random groups design each examinee takes only one form of the test. However, if order effects occur in which there are differences between the equating relationships for examinees taking Form X first compared to those taking Form Y first, then the data taken for the second test may need to be discarded.

If only one form can be used per test date for test security reasons, the common-item nonequivalent groups design might be considered. In this design, examinees from a specified group take one form on a specified test date and examinees from the other group take the other alternative form on the other specified test date. Thus the forms are taken by two nonequivalent groups. In this situation, a common item set in each test form is used to accomplish the equating.To accurately reflect group differences and to effectively separate group differences from form differences, the common item set in each test form should conform to several conditions. First, the common item set should be proportionally representative of Form X and Form Y in content and statistical characteristics. Second, each common item should have approximately the same item number in both forms. Third, the common items should be exactly the same in both forms.

A variety of statistical procedures or equating methods can be used to equate scores on Form X (new form) and Form Y (old form). Three statistical estimation methods are frequently used in observed score equating for the random groups design: (1) mean equating; (2) linear equating; and (3) equipercentile equating. In mean equating, Form X differs in difficulty from Form Y by a constant amount along the score scale (Kolen and Brennan, 1995). In mean equating, scores on Form X and Form Y that are an equal distance away from their respective means are set equal: $x-\mu(X) = y-\mu(Y)$. Then,

$m_Y(x) = y = x - \mu(X) + \mu(Y)$, (1) (2.2)[1]

where x is a particular score on Form X

X is the random variable score on Form X.

y is a particular score on Form Y.

Y is the random variable score on Form Y.

$\mu(X)$ is the mean on Form X of a population of examinees.

$\mu(Y)$ is the mean on Form Y of a population of examinees.

$m_Y(x)$ is a score x on Form X transformed to the scale of Form Y using mean equating.

For example, assume the mean on Form X is 40 and the mean of Form Y is 47. Then, by equation (1), 7 points must be added to each Form X score to transform the X-score to the Y scale. That is, a score of 40 on Form X indicates the same level of achievement as a score of 47 on Form Y and a score of 50 on Form X indicates the same level of achievement as a score of 57 on Form Y. When mean equating is used, the mean of the converted Form X scores is equal to the mean of the Form Y scores.

With the linear equating method, the differences in difficulty between two forms are allowed to vary along the score scale. Linear equating is accomplished by setting z-scores on X (test score on Form X minus the Form X mean divided by the standard deviation of the Form X test scores) equal to z-scores on Y (test score on Form Y minus the Form Y mean divided by the standard deviation of the Form X test scores) such that

$[x-\mu(X)] / \sigma(X) = [y-\mu(Y)] / \sigma(Y)$. Then, (2) (2.3)

$l_Y(x) = y = [\sigma(Y) / \sigma(X)]\, x + \{\mu(Y) - [\sigma(Y) / \sigma(X)]\, \mu(X)\}$, (3) (2.5)

[1] Throughout this paper a number of formulas reported in Kolen and Brennan (1995) are used. The second equation number provided for some equations is the equation number in Kolen and Brennan. Also, the notation used in this paper is consistent with the notation in Kolen and Brennan.

where $\sigma(X)$ and $\sigma(Y)$ are the standard deviations of Form X and Form Y scores, respectively.

$l_Y(x)$ is the linear conversion equation for converting x on Form X to the scale of Form Y.

Obviously, if the standard deviations of both forms are equal, then the linear equating method produces the same result as the mean equating method such that equation (1) is equal to equation (3). For a linear equating, the mean of the converted scores on Form X is equal to the mean of the Form Y scores and the standard deviation of the converted scores on Form X is the same as the standard deviation of the test scores on Form Y. To illustrate that the difference in test form difficulty varies with score level consider the following example. Supposed $\sigma(X)=5$, $\sigma(Y)=3$, $\mu(X)=40$, and $\mu(Y)=47$. Then, the resulting linear conversion equation is $l_Y(x) = 0.6x + 23$ [equation (3)]. Once x is known the equated value is known. For example, if x=44, then $l_Y(44)=49.4$. Alternately, if x=47, then $l_Y(47)=51.2$. If x=49, then $l_Y(50)=53$. The difference in difficuly between Form X and Form Y for a Form X score of 44 is 5.4 (49.4 - 44) but the difference for a Form X score of 50 is 3 (53 - 50).

Sometimes, the differences in difficulty between two test forms can be displayed by a curve rather than a straight line. In this situation, equipercentile equating may be appropriate. In equipercentile equating the cumulative distribution function of converted scores on Form X is equal to that of scores on Form Y. However, the estimated test score distributions or equipercentile relationships are somewhat irregular and produce random error. In this situation, smoothing methods may be used to reduce random error where the random error arises from the estimation of test score distributions and equipercentile relationship. Two general types of smoothing are used -- presmoothing and postsmoothing. In presmoothing, the test score distributions are smoothed and in postsmoothing, the equipercentile relationship is smoothed. Log-linear and strong-true score methods are often used in presmoothing and the cubic splines method is frequently used in postsmoothing. The equating results using smoothing are expected to have better precision than those based on unsmoothed relationships (Kolen & Brennan, 1995).

Purpose

The primary purpose of this study was to estimate the sample size required for equating using the random groups design. The approach taken in this study was to find the smallest sample size for which equating is expected to result in less overall equating error than had no equating been conducted. Mean, linear, and equipercentile equating methods were considered. Some of the analyses presented in this paper assumed that the test scores are normally distributed. Other analyses are not based on this assumption. Real test data from the ACT testing program were used to check whether the theoretical methods which will be discussed in the following sections provide reasonably accurate results for use in estimating sample size requirements.

The remainder of this paper is divided into four parts. The next section provides a conceptual framework for analyzing equating error. Then, previous investigations concerned with sample size issues in equating are discussed. The procedures used in this study are described next. Finally, the results are presented and the implications of these results are discussed.

II.A Conceptual framework for analyzing equating error

In this paper, equating is considered to be potentially worthwhile, if application of an equating method is expected to provide less overall error than does the use of the identity equating. In the identity equating, it is assumed that a score of 0 on Form X is equivalent to a 0 on Form Y, that a 1 on Form X is equivalent to a 1 on Form Y, etc. That is, in the identity equating it is assumed that the two forms are equal in difficulty at all points along the score scale. Consider a situation in which very few examinees took an examination. For example, assume we could randomly assign only four examinees to take each of the two forms. In this situation, would it be better to use the results from linear equating or would it be better to use the identity equating ? It could be argued that the identity equating would be preferred if it would be expected to result in less error than equating with these four examinees. The following discussion attempts to make these concepts more explicit.

Equating error arises from the difference between the Form Y equivalent estimated from the sample and the population Form Y equivalent. To show the components of total error, define $T_Y(x_i)$ as the population Form Y equivalent at a particular score on Form X. Define $\hat{t}_Y(x_i)$ as an estimator of $T_Y(x_i)$ that results from using an equating method. Define $t_Y(x_i)$ as equal to the expected value of $\hat{t}_Y(x_i)$ over replications of an equating method. Define total error at a particular x_i as $\hat{t}_Y(x_i)-T_Y(x_i)$. Define mean-squared error in equating at x_i using an equating method as

$$MSE[\hat{t}_Y(x_i)]=E[\hat{t}_Y(x_i)-T_Y(x_i)]^2. \qquad (4)\ (3.5)$$

Define variance of random error in using an equating method as

$$Var[\hat{t}_Y(x_i)]= E[\hat{t}_Y(x_i)-t_Y(x_i)]^2. \qquad (5)\ (3.6)$$

Define squared systematic error, squared bias, in equating using an equating method as

$$\{\text{Bias}_i[t_Y(x_i)]\}^2 = [t_Y(x_i) - T_Y(x_i)]^2. \qquad (6)\ (3.7)$$

Total error is comprised of random error and systematic error as expressed in the following equation

$$\hat{t}_Y(x_i) - T_Y(x_i) = [\hat{t}_Y(x_i) - t_Y(x_i)] + [t_Y(x_i) - T_Y(x_i)],$$

where $\hat{t}_Y(x_i) - t_Y(x_i)$ is random error and $t_Y(x_i) - T_Y(x_i)$ is systematic error.

It can be shown that mean-squared error in equating at x_i is

$$E[\hat{t}_Y(x_i) - T_Y(x_i)]^2 = E[\hat{t}_Y(x_i) - t_Y(x_i)]^2 + [t_Y(x_i) - T_Y(x_i)]^2$$

That is, $\text{MSE}[\hat{t}_Y(x_i)] = \text{Var}[\hat{t}_Y(x_i)] + \{\text{Bias}_i[t_Y(x_i)]\}^2$ (7) (3.8)

If the identity equating is used, then the Form Y equivalent of a Form X score is just the Form X score. That is, $\hat{t}_Y(x_i) = t_Y(x_i) = x_i$. The estimate, $\hat{t}_Y(x_i)$, for identity equating does not depend on the data. Thus, when using the identity equating, random error variance, $\text{Var}[\hat{t}_Y(x_i)]$ equals zero. However, the identity equating can have a large bias component. That is, the bias of identity equating is, $\text{Bias}_i = t_Y(x_i) - T_Y(x_i) = x_i - T_Y(x_i)$, and $\text{Bias}_i^2 = [x_i - T_Y(x_i)]^2$.

Hence, in deciding whether to equate or not, mean-squared error can be used as the index. If the mean-squared error for the equating method is less than the mean-squared error for the identity equating under the assumption that the equating method does not have any bias, then the equating method will produce less mean-squared error than the identity equating when the random error variance for that equating method is less than the squared bias for the identity equating. That is, when

$$\text{Var}[\hat{t}_Y(x_i)] < [x_i - T_Y(x_i)]^2,$$

where $Var[\hat{t}_Y(x_i)]$ is the random error variance for the equating method.

To provide a single index for equating error, error variance and equating error can be averaged over score points. Define f_i as the i^{th} raw score relative frequency at Form X and $\hat{eq}_Y(x_i)$ is an estimated equating function that is used to convert test score x_i to the scale of Form Y which is Form Y equivalent, y_j. Hence, the average squared bias over score points is $\Sigma_i f_i [x_i - eq_Y(x_i)]^2$ and the average error variance over score points is equal to $\Sigma_i f_i Var[\hat{eq}_Y(x_i)]$. Thus, if

$$\Sigma_i f_i Var[\hat{eq}_Y(x_i)] < \Sigma_i f_i [x_i - eq_Y(x_i)]^2, \qquad (8)$$

then the mean-squared error for the equating method is less than the mean-squared error for the identity equating. In this case, we might decide to use an equating method because application of the equating method results in less total error than application of the identity equating.

<u>Standard errors of mean, linear, and equipercentile equating using the random groups design</u>

In order to report the amount of equating error due to sampling from a specified population, standard errors of equating are needed. The delta method (Kendall & Stuart, 1977) can be used to derive the estimated standard errors with and without normality assumptions for various equating methods using a random groups design.

The delta method is based on a Taylor series expansion. Using this method, the appropoximate sampling variance of an equating function is (Kolen and Brennan, 1995):

$$Var[\hat{eq}_Y(x_i;\theta_1,\theta_2,\ldots,\theta_s)] \cong \Sigma_j eq_{Yj}'^{\,2}\, Var(\hat{\theta}_j) + \Sigma\Sigma_{j\neq k} eq_{Yj}' eq_{Yk}'\, Cov(\hat{\theta}_j,\hat{\theta}_k), \quad (9)\ (7.6)$$

where $\hat{\theta}_j$ and $\hat{\theta}_k$ are sample estimates of parameters θ_j and θ_k.

eq_{Yj}' is the first partial derivative of eq_{Yj} with respect to θ_j .

$eq_{Yj}'^{2}$ is the second partial derivative of eq_{Yj} with respect to θ_j .

$\hat{eq}_Y(x_i; \theta_1, \theta_2, ..., \theta_s)$ is an estimated equating function (mean, linear, or equipercentile) that is used to convert test score x_i to the scale of Form Y which is Form Y equivalent, y_j with estimated parameters $\theta_1, \theta_2, ..., \theta_s$.

In random groups mean equating, $\theta_1, \theta_2, ..., \theta_s$ reduces to only one θ, which is the mean. In linear equating, $\theta_1, \theta_2, ..., \theta_s$ are two moments and in equipercentile equating, $\theta_1, \theta_2, ..., \theta_s$ are cumulative probabilities.

Once the sampling variances (Var) of θ_j and sampling covariances (Cov) of the θ_j and θ_k are known, the estimated variance of a equating function and the estimated standard error of a equating function are known.

<u>Mean Equating</u>

The derivation of the standard errors for mean equating follows that of Kolen and Brennan (1995), except that the final equations are modified such that $N=N_X=N_Y$. The parameters for the sampling variance for mean equating that need to be estimated are $\mu(X)$ and $\mu(Y)$. The estimate of mean equating on Form X is

$$\hat{m}_Y(x_i) = x_i - \hat{\mu}(X) + \hat{\mu}(Y). \tag{10}$$

The error variances for $\hat{\mu}(X)$ and $\hat{\mu}(Y)$ are

$$Var[\hat{\mu}(X)] = \sigma^2(X)/N \text{ and } Var[\hat{\mu}(Y)] = \sigma^2(Y)/N. \tag{11}$$

The partial derivatives are

$$\partial\hat{m}/\partial\hat{\mu}(X) = -1 \text{ and } \partial\hat{m}/\partial\hat{\mu}(Y) = 1. \qquad (12)$$

If equations (11) and (12) are substituted into equation (9), the general form of the sampling variance for mean equating is obtained:

$$\mathrm{Var}[\hat{m}_Y(x_i)] \cong [\sigma^2(X) + \sigma^2(Y)]/N. \qquad (13)(7.8)$$

Equation (13) shows that as sample size increases the standard error of mean equating, which is just the square root of the sampling variance, decreases.

Linear Equating

The parameters of the sampling variance of linear equating that need to be estimated are $\mu(X)$, $\mu(Y)$, $\sigma(X)$, and $\sigma(Y)$. Based on Braun and Holland (1982), the general form of the sampling variance without the normality assumption for linear equating is

$$\mathrm{Var}[\hat{l}_Y(x_i)] \cong [\sigma^2(Y)/N]\{\, 2+[sk(X) + sk(Y)][(x_i - \mu(X))/\sigma(X)] +[(ku(X) + ku(Y) -2)/4][(x_i - \mu(X))/\sigma(X)]^2\}. \qquad (14)\ (7.9)$$

With the normality assumption for both X and Y, $sk(X)=sk(Y)=0$ and $ku(X)=ku(Y)=3$, then equation (14) can be simplified to

$$\mathrm{Var}[\hat{l}_Y(x_i)] \cong [\sigma^2(Y)/N]\ \{2 + [(x_i - \mu(X))/\sigma(X)]^2\}. \qquad (15)\ (7.10)$$

From equations (14) and (15), the error variance for linear equating tends to become smaller as sample size increases and as the test score on Form X becomes closer to the test score mean of that form. For nonnormal distributions, the error variance tends to increase as the indexes of skewness and kurtosis of the X score distribution depart from the values of these indexes for a normal distribution.

Equipercentile Equating

In the case of equipercentile equating, the parameters that need to be estimated are cumulative probabilities. Combining equation (10) in Lord (1982) and the notation of chapter 2 in Kolen and Brennan (1995), the general form of error variance without normality assumptions is

$$\hat{Var}[e_Y(x_i)] \cong (1/N)\{1/[G(y^*_U)-G(y^*_U-1)]^2\}\{\{[P(x_i)/100][1-P(x_i)/100]\ 2\}$$

$$-\{[G(y^*_U) - P(x_i)/100][P(x_i)/100 - G(y^*_U-1)]\ /[G(y^*_U) - G(y^*_U-1)]\}\},\quad (16)\ (7.12)$$

where $P(x_i)$ is the percentile rank for a score of x_i; y^*_U is the smallest integer score with a cumulative percent $100G(y)$ that is greater than a given percentile rank; and $G(y^*_U)-G(y^*_U-1)$ represents the proportion of examinees at the scores on Form Y.

The amount of error variance for equipercentile equating without normality assumptions increases as the proportion of examinees at the scores on Form Y, as symbolized by $G(y^*_U) - G(y^*_U-1)$ decreases.

Under normality assumptions, Petersen, et al.(1989) used the two-group case and the continuous case in Lord's study (1982b) to develop the following equation:

$$\hat{Var}[e_Y(x_i)] \cong [2\sigma^2(Y)/N]\{[P(x_i)/100][1-P(x_i)/100]\ /\phi^2\},\qquad (17)\ (7.14)$$

where ϕ is the ordinate of the standard normal density at the unit-normal score, z, below which $P(x_i)/100$ of the case fall (Kolen & Brennan, 1995).

<u>Minimum sample size require for equating to result in less error than the identity equating</u>

Recall from equation (8), that the application of an equating method was considered to be preferable to the identity equating if the average error variance for that method was less than the bias due to the application of the identity equating. To find the minimum sample size necessary for this property to hold, we start by setting the average error variance for the method equal to the average squared bias associated with the identity equating as follows:

$$\Sigma_i f_i \hat{Var}\ [eq_Y(x_i)] = \Sigma_i f_i [x_i - eq_Y(x_i)]^2,$$

where f_i is the i^{th} raw score relative frequency at Form X.

Then, note that all of the error variance expressions in Equations 13 through 17 can be expressed as 1/N times a quantity made up of parameters (e.g., means, standard deviations, percentiles). Refer to the quantity apart from the 1/N term as NVAR. In this case, Var=NVAR/N, so that the preceding equation is the same as

$$(1/N)\Sigma_i f_i \text{NVAR} [\hat{eq}_Y(x_i)] = \Sigma_i f_i [x_i - eq_Y(x_i)]^2 .$$

Solving for N, we have

$$N=\Sigma_i f_i \text{NVAR} [\hat{eq}_Y(x_i)] / \Sigma_i f_i [x_i - eq_Y(x_i)]^2 ,$$

which is the sample size required for average error variance for an equating method to be equal to the average squared bias for the identity equating. A sample size larger than N will result in the equating method having less average error variance than the identity equating. In this way, we have a method for finding the minimum sample size required for an equating method to result in less error than using the identity.

Following this logic, we can obtain sample size estimates for each of the equating methods as follows,

<u>Mean Equating</u>

$$N = \Sigma_i f_i [\sigma^2(X)+\sigma^2(Y)] / \Sigma_i f_i [x_i - \hat{m}_Y(x_i)]^2 . \tag{18}$$

<u>Linear Equating</u>

A.Under nonnormality assumptions:

$$N = \Sigma_i f_i \{\sigma^2(Y) \{8 + 4z_x [sk(X) + sk(Y)] + z_x^2 [ku(X) + ku(Y) -2] \}/4\}$$
$$/ \Sigma_i f_i [x_i - \hat{l}_Y(x_i)]^2 . \tag{19}$$

Because $z_x =[x_i-\mu(X)/\sigma(X)]$, sk(X), sk(Y), ku(X), and ku(Y) are all constants N can be computed from equation (19). In the case of linear equating, sample size per form depends on the characteristics of the test score distributions such as sk(X), sk(Y), ku(X), and ku(Y).

B. Under normality assumptions:

$$N = \Sigma_i f_i [\sigma^2(Y) (2 + z_x^2)] / \Sigma_i f_i [x_i - \hat{l}_Y(x_i)]^2 . \qquad (20)$$

Equipercentile Equating

A.Under nonnormality assumptions:

$$N=\Sigma_i f_i \{ 1/ [G(y_U^*)-G(y_U^*-1)]^3 \{[P(x_i)/100][1-(P(x_i)/100)] 2 [G(y_U^*)- G(y_U^*-1)] \}-$$

$$\{[G(y_U^*) - (P(x_i)/100)] [(P(x_i)/100) - G(y_U^*-1)] \} \} / \Sigma_i f_i [x_i - \hat{e}_Y(x_i)]^2 . \qquad (21)$$

Because $G(y^*_U) - G(y^*_U -1)$, and $P(x_i)$ are all constants, N can be computed from equation (21). The sample size per form depends on $G(y^*_U) - G(y^*_U -1)$ which is the proportion of examinees at scores on Form Y.

B. under normality assumptions

$$N = \Sigma_i f_i \{ 2 \sigma^2(Y) [P(x_i)/100][1-P(x_i)/100]/ \phi^2 \} / \Sigma_i f_i [x_i - \hat{e}_Y(x_i)]^2. \qquad (22)$$

Theoretical Methods

The equations that follow give the minimum sample size requirements so that the average error variance for equating is less than mean-squared error for the identity equating. These expressions are given for various equating methods under normality assumptions using the random groups design.

Mean Equating

$$\text{Average of Variance} = \int_{-\infty}^{\infty} \{[\sigma^2(X) + \sigma^2(Y)] / N\} \, g(z) \, dz$$

$$= \{[\sigma^2(X) + \sigma^2(Y)] / N\} \int_{-\infty}^{\infty} g(z) \, dz$$

$$= [\sigma^2(X) + \sigma^2(Y)] / N$$

$$= 2 \, \sigma^2(Y) / N ,$$

where z is a unit-normal variable

g(z) is the probability density function of z.

$z_x = [x_i - \mu_x]/ \sigma(X)$

In this equality, $\sigma^2(X)$ is set to equal to $\sigma^2(Y)$. Moreover, the integral equals one because g(z) is the probability density function of z integrated over its range. If u is taken to be the maximum proportion of standard deviation units of Form Y that is judged to be appropriate for the standard error of equating, then $u^2 = 2/N$ and $N= 2/u^2$. If the two sample sizes are equal then

$$N_{tot} = 4/u^2 \tag{23}$$

Linear Equating

$$\text{Average of Variance} = \int_{-\infty}^{\infty} [\sigma^2(Y)/2]\,[\,1/N_X + 1/N_Y]\,[2+z_x^2]\,g(z)\,dz$$

$$= [\sigma^2(Y)/2]\,[\,2/N] \int_{-\infty}^{\infty} [2+z_x^2]\,g(z)\,dz$$

$$= [\sigma^2(Y)/N]\,[\,2 \int_{-\infty}^{\infty} g(z)\,dz + \int_{-\infty}^{\infty} z_x^2\,g(z)\,dz\,]$$

$$= 3\,\sigma^2(Y)/N$$

In this equality, the first integral equals one because g(z) is the probability density function of z integrated over its range. The second integral equals one because the variance of a unit-normal variable z equals one, and this integral represents the variance. If u is taken to be the maximum proportion of standard deviation units of Form Y that is judged to be appropriate for the standard error of equating, then $u^2 = 3/N$ and $N= 3/u^2$. If the two sample sizes are equal then

$$N_{tot} = 6/u^2 \tag{24}$$

Equipercentile Equating

To set up average variance and minimum sample sizes required in the equipercentile equating, first we divide the test score distribution into several parts at i = -3, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 3. Second we let ϕ_i be the normal density at i = -3, -2.5, -2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5, 3 and then standardized ϕ_i such that the densities sum to 1(i.e.; $f_i' = \phi_i/\Sigma_i \phi_i$ and $\Sigma_i f_i' = \Sigma_i [\phi_i / \Sigma_i \phi_i] = 1$.). Hence, from equation (17) we get

Average of Variance = $[(2\sigma^2(Y))/N]\ \Sigma_i\ f_i'\ [\ P_i\ (1-P_i)/\ \phi_i^{\ 2}]$. If u is taken to be the maximum proportion of standard deviation units of Form Y that is judged to be appropriate for the standard error of equating, then $u^2 = [2/N]\ \{\Sigma_i\ f_i'\ [\ P_i\ (1-P_i)/\ \phi_i^{\ 2}]\}$ and

$N= [2/u^2]\ \{\Sigma_i\ f_i'\ [\ P_i\ (1-P_i)/\ \phi_i^{\ 2}]\}$. If the two sample sizes are equal then

$N_{tot} = 4\ \Sigma_i\ f_i'\ [\ P_i\ (1-P_i)/\ \phi_i^{\ 2}\ ]\ /u^2$, where P_i is the probability at i (25)

An alternative theoretical approach is to derive the minimum sample size requirements based on z-scores of -2 and +2. In this approach, the minimum sample size is taken as that sample size for which equating produces less overall equating error than the identity equating at z= -2 and +2.

Mean equating

$$\mathrm{var}[\hat{m}_Y(x_i)] \cong [\sigma^2(X)+\sigma^2(Y)]/N$$

$$u^2\sigma^2(Y) \cong [\sigma^2(X)+\sigma^2(Y)]/N$$

$$N \cong [\sigma^2(X)+\sigma^2(Y)]/\ [u^2\sigma^2(Y)]= (1/u^2)[1+ (\sigma^2(X)/\sigma^2(Y))]$$

since $z_X=2$ $x-\mu(X)=2\sigma(X)$
$y-\mu(Y)=2\sigma(Y)$
set $x-\mu(X)=2\sigma(X)=y-\mu(Y)=2\sigma(Y);\ \sigma(X)=\sigma(Y)$

$$N_{tot} \cong 4/u^2 \qquad (26)$$

Linear equating

$N_{tot} \cong (2/u^2\ )(2+z^2_X\ )$, where $z_X=2$

$$N_{tot} \cong 12/u^2 \qquad (27)$$

Equipercentile equating

$N_{tot} \cong 4[P(x_i)/100][1-P(x_i)/100]/[u^2\ \phi^2\ ] \cong 30.5626/\ u^2$,

where $z_X=2$, $P(x_i)=.9772$, $\phi = .0540$ (28)

In order to estimate the equating relationship for a specified population, sampling from that specified population is required, and random error is present. Hence, it is obvious that the standard errors of equating decrease as the sample sizes increases. However, systematic errors are not necessarily related to sample size. That is, systematic errors arise from the differences between the equating using the identity and the equating using other equating methods in the random groups design.

III.Review of the literature

In this section, several studies that investigated the standard errors for various equating designs are reviewed. These studies focused on the use of small samples. At the end of this review, the importance of the results obtained from these studies for the current investigation is considered.

Livingston (1993)

The purpose of the Livingston (1993) study was to investigate how log-linear presmoothing procedures can be used to improve the overall accuracy of equating using the common-item nonequivalent groups design. Form X and Form Y were selected from a 100-item test with 58 noncommon items and 24 common items on each form. Form X was to be equated to Form Y by a chain equipercentile equating. The 24 common items on both forms mirrored the content distribution as similarly as possible, while differing systematically in difficulty. The data were taken from the responses of 93,283 test takers. Each of those 93,283 test takers had a score on each form and a score on the 24 common items. Then, a direct equipercentile equating of test score distributions on both forms in the population of 93,283 examinees was conducted as a criterion equating for assessing the results of the equatings based on samples from the population. Anchor equatings were performed with samples of 25, 50, 100, and 200 test takers. For each sample size, a pair of samples was used to create unsmoothed distributions. Two-moment, three-moment, and four-moment smoothing methods were used with the joint distribution of scores on the 58 noncommon-items and the 24 common-items over 50 replications of the equating procedure. The accuracy of equating was assessed in terms of the root-mean-squared deviation (RMSD) defined as follows: Let x represent a score on Form X. $\hat{e}_Y(x_i)$ represent a score x on Form X transformed to the scale of Form Y in the direct equipercentile equating in the population. Let $e_{Y_r}(x_i)$ represent the score on Form Y that equates to x in the jth replication of the anchor equipercentile equating. Then, RMSD(x) is equal to:

$$\text{RMSD}(x) = \text{square root } \{ [\Sigma_j (\hat{e}_{Y_r}(x_i) - e_Y(x_i)) / R] \},$$

where j=1,2,3,...,50 and R=50

Based on Livingston's definition and equation, RMSD is very similar to the estimated bootstrap standard errors of equating.

The results from Livingston's study show that the two-moment, three-moment, and four-moment smoothing methods significantly improved the equating as compared to the unsmoothed method. Regardless of sample size, the three-moment smoothing method had the smallest RMSD. It produced the most accurate results relative to the unsmoothing method.

Parshall, Houghton, and Kromrey (1995)

Parshall, Houghton, and Kromrey (1995) compared standard errors of equating for five subject matter tests using linear equating and a common-item nonequivalent groups design. Each subject test had two parallel forms (Form X and Form Y). Samples of 15, 25, 50, 100, and 500 were randomly selected with replacement for 1000 replications. For each pair of Form X and Form Y tests at each size level, they estimated the linear equivalent at x_i and referred to this estimate as $\hat{l}_{Y_r}(x_i)$. This procedure was repeated 1000 times and bootstrap estimates $\hat{l}_{Y1}(x_i)$, $\hat{l}_{Y2}(x_i)$,..., $\hat{l}_{Y1000}(x_i)$, were obtained. The accuracy of equating was evaluated by computing bootstrap standard errors of equating. The standard errors of linear equating were the standard deviations of the obtained linear equated scores x_i in the bootstrap samples. Then, the estimated standard errors were defind as

$$\hat{SE}[\hat{l}_Y(x_i)] = \text{square root } \{\Sigma_r [\hat{l}_{Y_r}(x_i) - \hat{l}_{Y.}(x_i)]^2 / (R-1) \},$$

where $\hat{SE}[\hat{l}_Y(x_i)]$ is standard error for equated score x_i, $\hat{l}_{Yr}(x_i)$ is obtained equated score x_i in jth replication, $\hat{l}_{Y.}(x_i)$ is mean equated score x_i over the 1,000 replications (R=1000).

The results for each of the five tests showed that the standard errors decreased near the mean raw score and increased farther away from the mean raw score. As the sample size increased the standard error became smaller. There was a tendency in the larger samples for the standard error curves to rise less sharply than for smaller samples.

Lord (1982)

In the Lord (1982) study, he derived the standard error of equipercentile equating for four different situations: (1) scores x and y are continuous in the random groups design; (2) scores x and y are positive integers in the random groups design; (3) scores x and y are positive integers in the single group design; (4) scores x and y are continuous in the single group design. Situations (1) and (2), which deal with the random groups design, are related to this paper. In situation (1), one thousand students who had scores x and y on parallel Forms X and Y were randomly drawn from a population. From these 1000 students the equated score, $\hat{e}_Y(x_i)$, was computed at x_i = 0, 0.5, 1.0, 1.5, 2.0, 2.5 (repeated 1000 times for each x_i). Lord used an equation equivalent to equation (17) when $N=N_x=N_Y=1/2\ N_{tot}$ to compute "standard errors of equipercentile equating for normally distributed variables." The results showed that the size of the standard error along the X score scale was "acceptably" small.

Kolen (1985)

Kolen (1985) derived large sample standard errors for the Tucker method of linear equating with and without normality assumptions in the common-item nonequivalent groups design. He used a computer simulation and a real data example. The bootstrap method was used to verify the accuracy of the derived standard errors.

A computer simulation was conducted to study the estimated standard errors. To mirror the score distributions of test forms from two different testing programs, score distributions were simulated. The simulation included a nonsymmetric simulation and a nearly symmetric simulation. For the nonsymmetric simulation the score distributions for the two test forms modeled those of a particular professional certification testing program. For these simulations the score distributions were negatively skewed.The nearly symmetric simulation modeled score distributions that were symmetrical. Simulations were conducted for two sample sizes: 100 examinees per form and 250 examinees per form.The delta method with the normality assumption and without the normality assumption was used to estimate standard errors of equating for each X score. The standard deviation of Form Y equivalents of a given X score over the 500 replications was defined as the "true" standard error of equating. The mean delta method standard error based on the normality assumption over the 500 replications was defined as the normal delta method standard error associated with each X score. The nonnormal delta method was defined similarly.

The results indicated that the standard errors became larger at the extremes and smaller near the mean X score. The standard errors also decreased as sample size increased. For both simulations, the standard errors based on the nonnormality assumptions were more accurate than the standard errors based on the normality assumptions.

Root mean squared errors (RMSE), a measure of the variability in estimating standard errors, was defined as follows.

> "The variance of the estimated standard errors over the 500 replications was computed and added to the squared difference between the "true" standard error and the delta method standard error. The square root of this sum is RMSE."

The RMSE was smaller for the nonnormal standard errors than for the normal standard errors except for the nearly symmetric simulation with sample size of 100.

The bootstrap standard error was also used to evaluate the accuracy of the standard errors. The bootstrap standard error was defined as the standard deviation of the obtained equated scores over replications. Real data from Form X and Form Y (each form had 125 noncommon-item and 30 common-item) were used. 773 examinees from population 1 took Form X and 795 examinees from population 2 took Form Y.

The results from this real data example illustrated again that the standard errors became larger at the extremes and smaller near the mean X score. Standard errors derived from the delta method without the normality assumption were very similar to those derived from the bootstrap method. In addition, at the higher scores, the standard errors under the normality assumption were slightly larger than those derived with nonnormality assumption; whereas, at the lower scores, the standard errors under the normality assumption were slightly smaller than those derived with nonnormality assumption.

In summary, the results from the computer simulations illustrated that the standard errors based on nonnormality assumptions were more accurate than those based on normal assumptions, particularly for large samples. The results from the real data example provides evidence that the bootstrap standard errors are very similar to the delta method standard errors without the normality assumption.

Summary

Some important results from these related studies are discussed below.

Livingston (1993) and Parshall, Houghton, Kromery (1995) used the bootstrap method to deal with standard errors of equating while Lord (1982), and Kolen (1985) used both the delta and bootstrap methods to derive standard errors of equating. In the Kolen (1985) study the standard errors derived from the delta and bootstrap methods were very similar for both linear and equipercentile equating of number-correct scores when a large number of bootstrap replications are used (i.e., 1000 replications). Based on the foregoing empirical studies, the bootstrap method and the delta method standard errors can be expected to be similar.

As expected, as the sample size becomes larger, standard errors become smaller. This outcome was illustrated in the Livingston (1993), Parshall, Houghton, Kromery (1995), and Kolen (1985) studies for the common-item nonequlvalent groups design.

In the Livingston (1993) study, the log-linear presmoothing method did improve the overall accuracy of equating as compared to no smoothing. However, for presmoothings, the highest degree of presmoothed distributions did not always produce the most accurate equating.

Parshall, Houghton, Kromery (1995) and Kolen (1985) observed larger standard errors of equating at score points deviating from the mean

Only Kolen (1985) and Lord (1982) considered standard errors of equating under both normality and nonnormality assumptions. In the current study, both normality and nonnormality assumptions are considered. In practice, score distributions may not meet the normality assumption. In real situations, it is appropriate to derive the standard errors under the nonnormality assumption. Moreover, the studies reviewed considered the standard error of equating in the common-item nonequivalent groups design but they did not simultaneously use mean, linear, and equipercentile equatings to compute standard errors of equating for a random groups design. This study considers the magnitudes of standard errors and minimum sample sizes required for mean, linear, and equipercentile equatings in the random groups design.

IV.Procedures

The procedures used both to investigate the standard error at selected score points for various sample sizes under different equating methods and to investigate what sample sizes are needed to obtain equating error values less than the identity equating values are described below.

1. For selected raw score distributions, select an approptiate C value, the order of a polynomial log-linear model used to fit each raw score distribution.

2. Equate each of these distributions to the old form using the smoothed distributions and equipercentile equating.

3.Estimate what the standard error would be at selected score levels for various sample sizes per form-- 25, 50, 100, 200, 500, and 1,000 using the square root of equations (13), (14), and (16) for mean, linear, and equipercentile equating with nonnormality assumptions, respectively over the all Form X raw score points and using the square root of equations (13), (15) and (17) for mean, linear, and equipercentile equating with normality assumptions, respectively over all Form X raw score points.

4.Estimate what the minimum sample size needs to be to reduce equating error relative to the identity equating using equations (18), (19) and (21) for mean, linear and equipercentile equating with nonnormality assumptions.

5.Estimate what the minimum sample size needs to be to reduce equating error relative to the identity equating using equations (18), (20) and (22) for mean, linear and equipercentile equating with normality assumptions.

6. Estimate what the minimum sample size needs to be to reduce equating error relative to the identity equating using equations (18) through (22), for mean, linear and equipercentile equating with nonnormality and normality assumptions for real data.

7. Compare the results from Method 1 [The u values are derived from real data and these values are used in equations (23) through (25).] with those from Method 2 [equations (18), (20), and (22) are used and do not require a u-value.].[2]

[2] Method 1 and Method 2 are defined in the next section.

V.Results

The science subtest of the ACT Assessment provided the basic data for investigating the standard errors of equating and the minimum sample sizes needed to obtain less equating error than the identity equating. This test was selected because the score distributions tend to vary across test forms. Table 1 gives descriptive statistics for each of the three forms used in this study. Form Y is considered the old form and Forms X and Z are considered the new forms for investigating standard errors of equating and minimum sample size requirements under mean, linear, and equipercentile equating.

As can be seen in Table 1, Forms X and Z are both slightly more difficult and slightly less variable than Form Y. Also, the score distribution for Form X is positively skewed and the Form Z distribution is negatively skewed. The Form Y distribution is also negatively skewed. All three score distributions have kurtosis values less than 3.0.

Table 1 also shows the moments and fit statistics for presmoothing on Form Y, Form X, and Form Z. The next-to-last column shows the likelihood ratio chi-squared statistics (with degrees of freedom) for each C-value. This chi-squared test is an overall goodness-of-fit test that compares the fitted log-linear model to the empirical score distribution. The log-linear model is assumed to fit the empirical score distribution if this chi-squared test is not statistically significant at the given α. For example, $\chi^2(33) = 21.996$ at C =7 on Form Y means that the log-linear model with polynomial of degree 7 fits the empirical score distribution because 21.996 is smaller than the chi-squared table value at α=0.05 with df=30 (43.77). The last column at the right provides a difference statistic,

$\chi^2_C - \chi^2_{C+1}$, with one degree of freedom. $\chi^2_C - \chi^2_{C+1}$ is the difference between the overall χ^2 at C and the overall χ^2 at C+1. If $\chi^2_C - \chi^2_{C+1}$ is statistically significant, then the log-linear model with C+1 improves the fit over that of the model with C.

The C-value for Form Y is 7 because its chi-squared value with df=33 (21.996) was smaller than the chi-squared table value at α=0.05 with df=30 (43.77) and its difference chi-squared value with df=1 (2.23) is smaller than the chi-squared table value with df=1 at α=0.05 (3.841). C-values of 8 for Form X and 6 for Form Z were selected based on a similar analysis. Figure 1,2,and 3 show the fit of the smoothed distributions to the actual raw score distributions.

Standard Errors

Mean Equating

For mean equating, the standard error at all score points is the same. The results reported in Table 2 are based on equation (13). The values of the standard errors for the sample sizes of interest in this study are provided in Table 2. As expected, the standard errors decrease as sample size increases for both Form X and Form Z. However, the standard error for Form Z is greater than the standard error for Form X for all sample sizes.

Linear Equating

Tables 3 and 4 provide the standard error values for linear equating based on equation (14) [nonnormality assumption]. Tables 5 and 6 give the standard errors for linear equating based on equation (15) [normality assumption]. Figures 4 through 7 provide a graphical representation of these standard errors. Again, the standard errors decrease as sample sizes increase for both Form X and Form Z. However, unlike mean equating, for linear equating, the standard error becomes less as the raw score value approaches the mean score.

Four different situations are discussed below: (1) Table 3 vs Table 4 -- across forms within nonnormality assumptions; (2) Table 5 vs Table 6 -- across forms within normality assumptions; (3) Table 3 vs Table 5 -- across conditions within Form X; and (4) Table 4 vs Table 6 -- across conditions within Form Z.

In situation (1), Form X has the smaller standard errors relative to the standard errors for Form Z over the range of raw scores between 0 and 24 for all sample sizes. However, Form X has the larger standard errors relative to the standard errors for Form Z over the range of raw scores between 25 and 40 for all sample sizes.

In situation (2), Form X has the smaller standard errors relative to the standard errors for Form Z over the range of raw scores between 0 and 23 for all sample sizes. However, Form X has the larger standard errors relative to the standard errors for Form Z over the range of raw scores between 24 and 40 for all sample sizes.

For both situation (3) and (4), standard errors based on the nonnormality assumption are smaller than standard errors based on the normality assumption at all score points for all sample sizes.

Equipercentile Equating

Tables 7 and 8 give the standard error values based on equipercentile equating using equation (16) [nonnormality assumption]. The standard error values in Tables 9 and 10 are based on equation (17) [normality assumption]. Figures 8 through 11 provide standard error values for both Form X and Form Z and for both conditions[3]. Again, the standard errors decrease as sample size increases for both forms and under both with nonnormality and normality conditions.

As before, four different situations are discussed below: (1) Table 7 vs Table 8 -- across forms within nonnormality assumptions; (2) Table 9 vs Table 10 -- across forms

[3] No standard errors are given for raw scores 0 and 1 in Tables 7 through 10. These points have very small standard errors -- almost "0". Figures 8 through 11 delete raw scores with percentile ranks less than 0.5.

within normality assumptions (3) Table 7 vs Table 9 -- across conditions within Form X; and (4) Table 8 vs Table 10 -- across conditions within Form Z.

In situation (1), Form X has smaller standard errors relative to the standard errors for Form Z over the range of raw scores between 15 and 40 (except for raw scores of 35, 36, 38) for all sample sizes. Over the range of raw scores between 3 and 7 on Form X and over the range of raw scores between 4 and 8 on Form Z very large standard errors occur. [For Form X, over the range of raw scores between 3 and 7 the values of $G(y_U^*) - G(y_U^*-1)$ are smaller than 0.002 and their percentile ranks are smaller than 0.2. For Form Y, over the range of raw scores between 4 and 8 on Form Z, the values of $G(y_U^*) - G(y_U^*-1)$ are smaller than 0.0006 and their percentile ranks are smaller than 0.08.] For Form X, as the raw score values approaches 40 (except for raw scores of 35, 36, and 38), the standard error decreases. For Form Z, as the raw score values approaches 40 (except for raw scores of 34, 37, and 39), the standard error decreases.

In situation (2), Form X has smaller standard errors relative to the standard errors for Form Z both over the range of raw scores between 7 and 17 and over the range of raw scores between 19 and 23 for all sample sizes. However, Form X has larger standard errors relative to the standard errors for Form Z at other score points for all sample sizes.

In situation (3), the standard errors under the nonnormality assumption are smaller than the standard errors under the normality assumption over the range of raw scores between 27 and 40 for all sample sizes. On the other hand, nonnormality assumption has the larger standard errors relative to normality assumption over the range of raw scores between 2 and 26 for all sample sizes.

In situation (4), the nonnormality assumption results in smaller standard errors relative to the normality assumption over the range of raw scores between 29 and 40 for all sample sizes. However, the nonnormality assumption results in larger standard errors relative to normality assumption over the range of raw scores between 2 and 28 for all sample sizes.

Sample Sizes Required

Table 11 shows the sample size required so that a particular equating produces less overall equating error than the identity equating. The values in Table 11 are based on equations (26), (27), and (28) assuming normality. As an illustration of the interpretation of the data in Table 11 consider the row with standard deviation unit equal to 0.1. The values in this row represent the sample sizes required so that over the range of Form X z-scores between -2 and 2, the standard error of equating will be less than 0.1. Thus, for example, for mean equating a sample size of 400 is required in order for this procedure to provide less error than the identity equating when u=0.1. For u=0.1, equipercentile equating requires a sample size about three times as large as that required for linear equating and requires almost eight times as many examinees as mean equating.

Table 12 provides similiar information, except that the values in Table 12 were derived using equations (23), (24), and (25). Again, it can be seen that equipercentile equating requires considerably greater sample sizes than linear equating and mean equating.[4]

In mean equating, the sample size required remains the same for all standard units regardless of which method is used. (See Tables 11 and 12.) However, the sample size required with linear equating using the average variance criterion is only one-half that required for the +2/-2 criterion. Likewise, the sample size required in equipercentile equating for the average variance criterion is consistently less than the sample size based on

[4] The major difference between the two theoretical methods used to obtain the values in Tables 11 and 12 is described below. As mentioned, the rationale for deriving equations (26) through (28) was to identify the sample size needed at z = +2 and / or -2 for a specific u value. The rationale for the other theoretical method based on equations (23), (24), and (25) was to identify the sample size needed so that the average error variance for equating is less than mean-squared error for the identity equating for a specific value of u.

the +2/-2 criterion. As indicated in Tables 11 and 12, equipercentile equating requires larger samples than either linear equating or mean equating.

The minimum sample size requirements based on real data are summarized in Table 13. The values in Table 13 are based on the ACT science test scores and equations (18), (19), (20), (21), and (22). The minimum sample size requirements for mean equating to have less error than the identity equating on Form X and Form Z are 22 and 81, respectively [equation (18)]. The minimum sample size requirements for linear equating to have less error than identity equating for Form X and Form Z under nonnormality / normality assumptions are 29, 89, 32, and 99, respectively [equation (25)]. Finally, the minimum sample size requirement for equipercentile equating to be better than the identity equating for Form X and Form Z under nonnormality / normality assumptions are 44, 145, 51, and 161, respectively. For each equating method, Form Z requires larger sample sizes than Form X under both conditions. If normality assumptions are made, greater sample sizes are needed relative to the nonnormality assumptions for these tests.

To evaluate whether the theoretical methods (assuming normality assumption) provide reasonably accurate sample size estimates, equations (23), (24), and (25) were used. The u^2 values needed in these equations were estimated using the following formula;

$$\{\Sigma_i f_i [x_i - e_Y(x_i)]^2 \} / \sigma^2(Y) = u^2 . \qquad (29)$$

Table 14 shows both the theoretical sample size estimate (Method 1) and the sample size estimate based on the actual equating (Method 2). As can be seen in Table 14, the sample size estimates derived from the actual equating are approximately the same as the estimates based on the theoretical models.

VI. Discussion and Conclusions

Based on real data, in general, as the sample size increases, the magnitude of the standard errors decreases for both forms. (E.g., the smallest standard errors are found at n=1000; whereas, the largest standard errors are found at n=25.) In linear equating, the standard error becomes less as the raw score value approaches the mean score. In equipercentile equating, with nonnormality assumptions, raw scores less than or equal to 10 are associated with greater standard errors but the standard errors become smaller as the raw score approaches the middle score. For both forms, standard errors associated with normality assumptions have the smallest standard errors over the range of raw scores between 20 and 26 and have the largest standard errors at extreme score points. Based on these results, it is reasonable to conclude that standard errors become less as sample size increases and that they tend to be less for middle scores than the extreme scores for both the linear and equipercentile methods.

As shown in Tables 11 and 12, the two theoretical methods for deriving minimum sample size estimates do not provide the same estimates for either the linear or equipercentile method. As can be seen given equations (23) and (26), the minimum sample size estimates are the same in mean equating. However, for linear and equipercentile equating, equations (24) and (25) are derived so that the average error variance (over the z-range -3 to +3) for equating is equal to the average-squared bias. Whereas, equations (27) and (28) are derived so that the standard error of equating at a z value of -2 and +2 is equal to u standard deviation units on Form Y. Thus, it is not surprising that the sample size estimates based on equations (27) and (28) are greater than those based on equations (24) and (25).

The minimum sample size requirements for equating to result in less overall equating error than the identity equating for real data were provided in Table 13. For each equating method, Form Z requires a greater sample size than Form X under all conditions.

The reason why Form Z requires a larger sample size than Form X is because the mean, standard deviation, skewness, and kurtosis of raw score distributions, provided in Table 1, between Forms Z and Y (24.35 vs. 25.92; 6.56 vs. 7.55; -0.17 vs. -0.28; 2.44 vs. 2.30) are closer to one another than those for Forms X and Y (22.9 vs. 25.92; 6.25 vs. 7.55; 0.12 vs. -0.28; 2.45 vs. 2.30). That is, Forms Z and Y are more similar to one another than Forms X and Y. Moreover, a greater sample size is needed under normality assumptions than under nonnormality assumptions for linear and equipercentile equating methods. Comparing equation (19) to (20), the magnitude of $\{8+ 4z_x\,[sk(X)+sk(Y)]+ z_x^2\,[ku(X) + ku(Y) -2]\,\}/4$ in equation (19) is small relative to $\sigma^2(Y)\,(\,2 + z_x^2\,)$ in equation (20), so a greater sample size is needed under normality assumptions than under nonnormality assumptions for linear equating. That is, kurtosis plays an important role here. Similarly, the magnitude of

$\{1/\,[G(y_U^*)-G(y_U^*-1)]^3\{[P(x_i)/100][1-(\,P(x_i)/100\,)]\;2\;[G(y_U^*)-\,G(y_U^*-1)]\,\}-$

$\{[G(y_U^*) - (\,P(x_i)/100\,)]\,[\,(P(x_i)/100) - G(y_U^*-1)]\,\}\}$ in equation (21) is small relative to

$\{2\,\sigma^2(Y)\,[P(x_i)/100][1-P(x_i)/100]/\,\phi^2\,\}$in equation (22), so a greater sample size is needed under normality assumptions than under nonnormality assumptions for equipercentile equating. That is, $G(y^*_U) - G(y^*_U -1)$, which is the proportion of examinees at scores on Form Y, plays an important role here. As compared to mean and linear equating, the equipercentile equating method requires greater sample sizes.

Two different methods were used to estimate the minimum sample size using real data. Method 1 requires u-values to be estimated from the data; whereas, Method 2 does not require the use of a u-value. If the sample size estimates obtained from these two methods are similar, then setting the average error variance equal to $u^2 \times \sigma^2(Y)$ (Method 1) is consistent with setting the average error variance equal to the average squared bias (Method 2). As can be seen in Table 14, the minimum sample size requirements for both forms under Method 1 and Method 2 are similar . Thus, it seems reasonable to set the average error variance equal to the average squared bias or to set the average error variance equal to $u^2 \times \sigma^2(Y)$ for all three equating methods.

Two other approaches to sample size estimation that assume that the scores on Form X are normally distributed have been discussed in Kolen and Brennan (1995). One approach is to choose a sample size so that the standard error of equating is small relative to the standard deviation (Method 3). Another approach chooses a sample size so that the standard error of equating is small relative to the standard error of measurement (Method 4). The following paragraphs summary these alternative approaches from Kolen and Brennan (1995) and illustrate and discuss how these alternative approaches be used and contrasted with Method 1 and Method 2 for this study.

Method 3:

Linear Equatingwith the random groups design:

Consider equation (7.10). Let u is the maximum number of Form Y standard deviation units allowed for equating and equation (7.10) can be written as

$$u^2\, \sigma^2(Y) \cong [\sigma^2(Y)/N]\ \{2 + [(x_i - \mu(X))/\sigma(X)]^2\}.$$

Then, $N \cong \{2 + [(x_i - \mu(X))/\sigma(X)]^2\}/\ u^2$.

Therefore, $N_{tot} \cong 2\ \{2 + [(x_i - \mu(X))/\sigma(X)]^2\}/\ u^2$. (30)(7.18)

Thus, N_{tot} is the total sample size required for the standard error of equating to be equal to u standard deviation units on Form Y. To illustrate how this approach can be used on real data in this study, let u equal 0.44. Then, the total sample size required over the range of Form X z-scores between +2 and -2 is 62, based on equation (7.18), which is a larger sample size requirement than either Method 1 (31) or Method 2 (32) as shown in Table 14. Thus, at least 62 examinees would be needed over the range of Form X z-scores between +2 and -2 provided the standard error of equating will be less than 0.44 Form Y standard deviation units using linear equating based on real data under this approach. Similarly, let u equal 0.25. Then, the total sample size required over the range of Form Z z-scores between +2 and -2 is 192, which leads to larger sample size requirements than Method 1 (96) and Method 2 (99) as shown in Table 14. Thus, at least 192 examinees would be needed over the range of Form Z z-scores between +2 and -2 provided the standard error of equating will be less than 0.25 Form Y standard deviation units using linear equating based on real data under this approach.

Equipercentile Equating:

Consider equation (7.14) and use the same rationale that was used with linear equating then N can be written as

$$N \cong 2\{[P(x_i)/100][1-P(x_i)/100]\}/\ u^2\phi^2.$$

Therefore, $N_{tot} \cong 4\{[P(x_i)/100][1-P(x_i)/100]\}/\ u^2\phi^2.$ (31) (7.19)

For example, based on real data in this study, if u=0.45 then the total sample size required over the range of Form X z-scores between +2 and -2 is 151 based on equation (7.19), which leads to larger sample size requirements than Method 1 (60) and Method 2 (51) as shown in Table 14. Thus, at least 151examinees would be needed over the range of Form X z-scores between +2 and -2 provided the standard error of equating will be less than 0.45 Form Y standard deviation units using equipercentile equating based on real data under this approach. Similarly, if u=0.25 then the total sample size required over the range of Form Z z-scores between +2 and -2 is 489, which leads to larger sample size requirements than Method 1 (193) and Method 2 (161) as shown in Table 14. Thus, at least 489 examinees would be needed over the range of Form Z z-scores between +2 and -2 provided the standard error of equating will be less than 0.25 Form Y standard deviation units using equipercentile equating based on real data under this approach.

This approach which chooses a sample size so that the standard error of equating is small relative to the standard deviation (Method 3) will lead to larger sample size requirements than Method 1 and Method 2 for Forms X and Z provided in Table 14.

Method 4:

This approach chooses a sample size so that the standard error of equating is small relative to the standard error of measurement. Thus, the relationship between standard error of measurement units (u_{sem}) and the maximum number of Form Y standard deviation units allowed for equating (u) needs to be explored.

First, recall the relationship between the standard error of measurement (sem) and Form X score reliability. That is,

$$\text{sem} = \sigma(Y) \text{ x square root of } [1 - \rho(X,Y)], \quad (31)$$

where $\rho(X,Y)$ is alternate forms reliability

Then, equation (31) can be written as

$$u_{sem} \text{ x sem} = u_{sem} \text{ x } \{\sigma(Y) \text{ x square root of } [1 - \rho(X,Y)]\}$$

From the earlier definition of u,

$$u_{sem} \text{ x square root of } [1 - \rho(X,Y)] = u. \quad (32)$$

Thus, once $\rho(X,Y)$ and usem are known then u can be computed and equations (7.18) and (7.19) can be used for estimating the total sample size required using linear and equipercentile equating with the random groups design.

The other practical issue in sample size determination is to hypothesize extent to which two forms differ in terms of u. For example, in Table 11, the sample size required is determined by a given level of equating precision. That is, the investigator has to decide on the appropriate magnitude for u. In order to answer this question, the investigator can examine test forms from similar testing programs. For instance, if two forms differ by a maximum of 0.2 Form Y standard deviation units in a similar testing program, then the investigator might select u=0.2.

References

Angoff, W.A. (1971). Scales, norms, and equivalent scores. In R.L.Thorndike (Ed.), Educational measurement (2nd ed., pp. 508-600). Washington, DC: American Council on Education. (Reprinted as W.A. Angoff, Scales, norms, and equivalent scores. Princeton, NJ: Educational Testing Service, 1984.)

Braun, H.I., & Holland, P.W. (1982). Observed-score test equating: A mathematical analysis of some ETS equating procedures. In P.W. Holland and D.B.Rubin (Eds.). Test equating (pp. 9-49). New York: Academic.

Brennan, R.L., & Kolen, M.J. (1987). Some practical issues in equating. Applied Psychological Measurement, 11, 279-290.

Kendall, M., & Stuart, A. (1977). The advanced theory of statistics (4th ed., Vol. 1). New York: Macmillan.

Kolen, M.J., & Brennan, R.L. (1995). Test Equating: Methods and Practices.

Kolen, M.J. (1985). Standard errors of Tucker equating. Applied Psychological Measurement, 9, 209-223.

Kolen, M.J. (1988). An NCME instructional module on traditional equating methodology. Educational Measurement: Issues and Practice, 7, 29-36.

Livingston, S.A. (1993). Small-sample equating with log-linear smoothing. Journal of Educational Measurement , 30, 23-39.

Lord, F.M. (1950). Notes on comparable scales for test scores (Research Bulletin 50-48). Prtinceton, NJ: Educational Testing Service.

Lord, F.M. (1982b). The standard error of equipercentile equating. Journal of Educational Statistics, 7, 165-174.

Parshall, C.G., Houghton, P.D.B., & Kromrey, D.J. (1995, Spring). Equating error and statistical bias in small sample linear equating. Journal of Educational Measurement, 32, 37-54.

Petersen, N.S., Kolen, M.J., & Hoover, H.D. (1989). Scaling, norming, and equating. In R.L. Linn (Ed.), Educational measurement (3rd ed., pp. 221-262). New York: American Council on Education and Macmillan.

Table 1. Moments and Fit Statistics for Presmoothing

Form Y

	Mean	s.d.	skewness	kurtosis	χ^2 (*df*)	$\chi^2_c - \chi^2_{c+1}$
Raw	25.922498	7.552886	-0.284190	2.302154		
Log-linear						
C=10	25.922498	7.552886	-0.284190	2.302154	18.059(30)	0.000
C= 9	25.922498	7.552886	-0.284190	2.302154	18.370(31)	0.312
C= 8	25.922498	7.552886	-0.284190	2.302154	19.735(32)	1.365
C= 7	25.922498	7.552886	-0.284190	2.302154	21.966(33)	2.230
C= 6	25.922498	7.552886	-0.284190	2.302154	27.669(34)	5.703
C= 5	25.922498	7.552886	-0.284190	2.302154	36.905(35)	9.236
C= 4	25.922498	7.552886	-0.284190	2.302155	39.516(36)	2.611
C= 3	25.922498	7.552886	-0.284190	2.614418	113.213(37)	73.697
C= 2	25.922498	7.552886	-0.328621	2.690028	115.407(38)	2.193
C= 1	25.922498	10.924828	-0.624494	2.344978	1272.280(39)	1156.873

Form X

	Mean	s.d.	skewness	kurtosis	χ^2(*df*)	$\chi^2_c - \chi^2_{c+1}$
Raw	22.901472	6.248523	0.123761	2.451064		
Log-linear						
C=10	22.901472	6.248523	0.123761	2.451064	22.546(30)	0.000
C= 9	22.901472	6.248523	0.123761	2.451064	23.124(31)	0.578
C= 8	22.901472	6.248523	0.123761	2.451064	25.860(32)	2.737
C= 7	22.901472	6.248523	0.123761	2.451064	29.726(33)	3.865
C= 6	22.901472	6.248523	0.123761	2.451064	31.613(34)	1.887
C= 5	22.901472	6.248523	0.123761	2.451064	31.664(35)	0.051
C= 4	22.901472	6.248523	0.123761	2.451064	31.999(36)	0.334
C= 3	22.901472	6.248523	0.123761	2.806226	82.458(37)	50.460
C= 2	22.901472	6.248523	-0.054557	2.861598	105.399(38)	22.941
C= 1	22.901472	11.617498	-0.297264	1.920897	2823.196(39)	2717.797

Form Z

	Mean	s.d.	skewness	kurtosis	χ^2(*df*)	$\chi^2_c - \chi^2_{c+1}$
Raw	24.345199	6.558786	-0.165047	2.436220		
Log-linear						
C=10	24.345199	6.558786	-0.165047	2.436220	28.756(30)	0.000
C= 9	24.345199	6.558786	-0.165047	2.436220	29.393(31)	0.637
C= 8	24.345199	6.558786	-0.165047	2.436220	29.409(32)	0.016
C= 7	24.345199	6.558786	-0.165047	2.436220	29.910(33)	0.501
C= 6	24.345199	6.558786	-0.165047	2.436220	29.975(34)	0.065
C= 5	24.345199	6.558786	-0.165047	2.436220	34.883(35)	4.908
C= 4	24.345199	6.558786	-0.165047	2.436220	39.846(36)	4.963
C= 3	24.345200	6.558789	-0.165049	2.831210	99.521(37)	59.676
C= 2	24.345199	6.558786	-0.122869	2.785140	100.875(38)	1.353
C= 1	24.345199	11.348019	-0.450204	2.080416	2264.908(39)	2164.033

Table 2.Standard error of mean equating for Form X and Z

Sample size \ Form	Form X	Form Z
$N_{X(Z)} = N_Y = 25$	1.9605	2.0006
$N_{X(Z)} = N_Y = 50$	1.3863	1.4147
$N_{X(Z)} = N_Y = 100$	0.9803	1.0003
$N_{X(Z)} = N_Y = 200$	0.6931	0.7073
$N_{X(Z)} = N_Y = 500$	0.4384	0.4474
$N_{X(Z)} = N_Y = 1000$	0.3100	0.3163

Table 3. Standard error of linear equating on Form X under nonnormality assumptions

x_i	N_X=N_Y=25	N_X=N_Y=50	N_X=N_Y=100	N_X=N_Y=200	N_X=N_Y=500	N_X=N_Y=1000
0	5.19646514	3.67445574	2.59823257	1.83722787	1.16196493	0.82163328
1	5.01422056	3.54558936	2.50711028	1.77279468	1.12121380	0.79281788
2	4.83342577	3.41774814	2.41671321	1.70887430	1.08078701	0.76423182
3	4.65425259	3.29105357	2.32712629	1.64552678	1.04072252	0.73590195
4	4.47689359	3.16564182	2.23844680	1.58282091	1.00106384	0.70785903
5	4.30157392	3.04167209	2.15078696	1.52083605	0.96186117	0.68013856
6	4.12855340	2.91932810	2.06427670	1.45966405	0.92317260	0.65278161
7	3.95813352	2.79882306	1.97906676	1.39941153	0.88506556	0.62583586
8	3.79066508	2.68040499	1.89533254	1.34020249	0.84761848	0.59935678
9	3.62655697	2.56436303	1.81327849	1.28218151	0.81092279	0.57340901
10	3.46628650	2.45103469	1.73314325	1.22551735	0.77508523	0.54806802
11	3.31041111	2.34081414	1.65520555	1.17040707	0.74023043	0.52342195
12	3.15958133	2.23416139	1.57979067	1.11708069	0.70650386	0.49957367
13	3.01455463	2.13161202	1.50727732	1.06580601	0.67407491	0.47664294
14	2.87620896	2.03378686	1.43810448	1.01689343	0.64313988	0.45476857
15	2.74555446	1.94140018	1.37277723	0.97070009	0.61392464	0.43411028
16	2.62374037	1.85526461	1.31187018	0.92763230	0.58668618	0.41484978
17	2.51205308	1.77628977	1.25602654	0.88814488	0.56171214	0.39719047
18	2.41189982	1.70547072	1.20594991	0.85273536	0.53931720	0.38135485
19	2.32477177	1.64386188	1.16238588	0.82193094	0.51983477	0.36757869
20	2.25218110	1.59253253	1.12609055	0.79626626	0.50360300	0.35610110
21	2.19557022	1.55250259	1.09778511	0.77625130	0.49094443	0.34715013
22	2.15619814	1.52466232	1.07809907	0.76233116	0.48214056	0.34092486
23	2.13501877	1.50968625	1.06750938	0.75484312	0.47740471	0.33757611
24	2.13257422	1.50795769	1.06628711	0.75397885	0.47685809	0.33718959
25	2.14892843	1.51952187	1.07446422	0.75976093	0.48051501	0.33977542
26	2.18365907	1.54408014	1.09182953	0.77204007	0.48828101	0.34526681
27	2.23590997	1.58102710	1.11795498	0.79051355	0.49996467	0.35352841
28	2.30448970	1.62952029	1.15224485	0.81476015	0.51529956	0.36437181
29	2.38799185	1.68856523	1.19399593	0.84428262	0.53397121	0.37757466
30	2.48491254	1.75709851	1.24245627	0.87854926	0.55564334	0.39289917
31	2.59374797	1.83405678	1.29687398	0.91702839	0.57997968	0.41010756
32	2.71306462	1.91842639	1.35653231	0.95921320	0.60665969	0.42897318
33	2.84154248	2.00927396	1.42077124	1.00463698	0.63538822	0.44928732
34	2.97799608	2.10576122	1.48899804	1.05288061	0.66590017	0.47086252
35	3.12137959	2.20714868	1.56068980	1.10357434	0.69796170	0.49353345
36	3.27078177	2.31279197	1.63539089	1.15639598	0.73136904	0.51715601
37	3.42541518	2.42213430	1.71270759	1.21106715	0.76594612	0.54160569
38	3.58460289	2.53469701	1.79230144	1.26734850	0.80154157	0.56677548
39	3.74776460	2.65006976	1.87388230	1.32503488	0.83802564	0.59257361
40	3.91440341	2.76790119	1.95720170	1.38395060	0.87528721	0.61892152

Table 4. Standard error of linear equating on Form Z under nonnormality assumptions

z_i	$N_Z=N_Y=25$	$N_Z=N_Y=50$	$N_Z=N_Y=100$	$N_Z=N_Y=200$	$N_Z=N_Y=500$	$N_Z=N_Y=1000$
0	5.46730615	3.86596925	2.73365307	1.93298463	1.22252682	0.86445700
1	5.07591677	3.58921517	2.53795838	1.79460758	1.13500949	0.80257291
2	4.89312129	3.45995925	2.45157756	1.73352712	1.09637881	0.77525689
3	4.73180204	3.34588931	2.36590102	1.67294465	1.05806310	0.74816359
4	4.56201625	3.22583263	2.28100812	1.61291631	1.02009784	0.72131810
5	4.39397944	3.10701266	2.19698972	1.55350633	0.98252367	0.69474915
6	4.22790015	2.98957687	2.11395008	1.49478843	0.94538721	0.66848971
7	4.06401838	2.87369496	2.03200919	1.43684748	0.90874214	0.64257773
8	3.90261098	2.75956269	1.95130549	1.37978134	0.87265034	0.61705698
9	3.74399797	2.64740635	1.87199898	1.32370318	0.83718340	0.59197806
10	3.58854991	2.53748798	1.79427496	1.26874399	0.80242415	0.56739956
11	3.43669630	2.43011126	1.71834815	1.21505563	0.76846865	0.54338940
12	3.28893505	2.32562828	1.64446753	1.16281414	0.73542824	0.52002629
13	3.14584288	2.22444683	1.57292144	1.11222342	0.70343185	0.49740143
14	3.00808616	2.12703813	1.50404308	1.06351906	0.67262851	0.47562018
15	2.87643158	2.03394428	1.43821579	1.01697214	0.64318966	0.45480377
16	2.75175512	1.94578471	1.37587756	0.97289235	0.61531115	0.43509069
17	2.63504747	1.86325994	1.31752374	0.93162997	0.58921453	0.41663759
18	2.52741279	1.78715073	1.26370640	0.89357536	0.56514668	0.39961905
19	2.43005700	1.71830978	1.21502850	0.85915489	0.54337726	0.38422575
20	2.34426106	1.65764289	1.17213053	0.82882144	0.52419271	0.37066022
21	2.27133531	1.60607660	1.13566766	0.80303830	0.50788602	0.35912965
22	2.21255274	1.56451105	1.10627637	0.78225552	0.49474183	0.34983531
23	2.16906351	1.53375952	1.08453175	0.76687976	0.48501735	0.34295905
24	2.14179940	1.51448088	1.07089970	0.75724044	0.47892091	0.33864822
25	2.13138316	1.50711549	1.06569158	0.75355774	0.47659176	0.33700127
26	2.13806104	1.51183746	1.06903052	0.75591873	0.47808498	0.33805713
27	2.16167462	1.52853478	1.08083731	0.76426739	0.48336514	0.34179077
28	2.20167905	1.55682218	1.10083952	0.77841109	0.49231040	0.34811602
29	2.25720301	1.59608355	1.12860150	0.79804178	0.50472594	0.35689513
30	2.32713590	1.64553358	1.16356795	0.82276679	0.52036341	0.36795249
31	2.41022384	1.70428562	1.20511192	0.85214281	0.53894243	0.38108985
32	2.50515822	1.77141437	1.25257911	0.88570718	0.56017041	0.39610029
33	2.61064701	1.84600620	1.30532350	0.92300310	0.58375842	0.41277954
34	2.72546495	1.92719475	1.36273247	0.96359737	0.60943249	0.43093385
35	2.84848414	2.01418245	1.42424207	1.00709123	0.63694042	0.45038489
36	2.97868863	2.10625093	1.48934431	1.05312546	0.66605503	0.47097203
37	3.11517757	2.20276319	1.55758879	1.10138159	0.69657488	0.49255282
38	3.25716103	2.30316065	1.62858052	1.15158033	0.72832335	0.51500238
39	3.40395153	2.40695721	1.70197576	1.20347860	0.76114670	0.53821199
40	3.55495363	2.51373182	1.77747682	1.25686591	0.79491180	0.56208752

Table 5. Standard error of linear equating on Form X under normality assumptions

x_i	$N_X=N_Y=25$	$N_X=N_Y=50$	$N_X=N_Y=100$	$N_X=N_Y=200$	$N_X=N_Y=500$	$N_X=N_Y=1000$
0	5.93427521	4.19616624	2.96713761	2.09808312	1.32694428	0.93829130
1	5.70939684	4.03715322	2.85469842	2.01857661	1.27665994	0.90273490
2	5.48595368	3.87915505	2.74297684	1.93957753	1.22669654	0.86740544
3	5.26412852	3.72230097	2.63206426	1.86115049	1.17709492	0.83233180
4	5.04413480	3.56674192	2.52206740	1.78337096	1.12790283	0.79754774
5	4.82622299	3.41265500	2.41311149	1.70632750	1.07917627	0.76309286
6	4.61068828	3.26024895	2.30534414	1.63012447	1.03098124	0.72901383
7	4.39788018	3.10977090	2.19894009	1.55488545	0.98339590	0.69536591
8	4.18821435	2.96151477	2.09410717	1.48075738	0.93651320	0.66221483
9	3.98218714	2.81583153	1.99109357	1.40791576	0.89044411	0.62963907
10	3.78039349	2.67314188	1.89019675	1.33657094	0.84532168	0.59773269
11	3.58354868	2.53395158	1.79177434	1.26697579	0.80130585	0.56660880
12	3.39251426	2.39886984	1.69625713	1.19943492	0.75858925	0.53640360
13	3.20832830	2.26863069	1.60416415	1.13431535	0.71740402	0.50728124
14	3.03223903	2.14411678	1.51611951	1.07205839	0.67802926	0.47943909
15	2.86573939	2.02638375	1.43286969	1.01319188	0.64079881	0.45311318
16	2.71059708	1.91668158	1.35529854	0.95834079	0.60610793	0.42858303
17	2.56887067	1.81646587	1.28443533	0.90823293	0.57441694	0.40617412
18	2.44289625	1.72738850	1.22144812	0.86369425	0.54624821	0.38625581
19	2.33522446	1.65125305	1.16761223	0.82562653	0.52217206	0.36923141
20	2.24848618	1.58991982	1.12424309	0.79495991	0.50277679	0.35551688
21	2.18517563	1.54515251	1.09258782	0.77257625	0.48862013	0.34550660
22	2.14736597	1.51841704	1.07368298	0.75920852	0.48016563	0.33952837
23	2.13641155	1.51067109	1.06820577	0.75533555	0.47771614	0.33779633
24	2.15272237	1.52220459	1.07636119	0.76110229	0.48136336	0.34037529
25	2.19569091	1.55258793	1.09784545	0.77629397	0.49097141	0.34716922
26	2.26379971	1.60074813	1.13189986	0.80037406	0.50620100	0.35793816
27	2.35486843	1.66514344	1.17743422	0.83257172	0.52656459	0.37233739
28	2.46635503	1.74397637	1.23317752	0.87198818	0.55149375	0.38996497
29	2.59562990	1.83538750	1.29781495	0.91769375	0.58040049	0.41040512
30	2.74017657	1.93759744	1.37008829	0.96879872	0.61272211	0.43325996
31	2.89771053	2.04899077	1.44885527	1.02449538	0.64794777	0.45816826
32	3.06623069	2.16815251	1.53311535	1.08407626	0.68563003	0.48481364
33	3.24402537	2.29387234	1.62201268	1.14693617	0.72538612	0.51292545
34	3.42965248	2.42513053	1.71482624	1.21256526	0.76689361	0.54227567
35	3.62190797	2.56107569	1.81095398	1.28053784	0.80988324	0.57267393
36	3.81979111	2.70100020	1.90989556	1.35050010	0.85413126	0.60396201
37	4.02247145	2.84431684	2.01123573	1.42215842	0.89945196	0.63600858
38	4.22925935	2.99053797	2.11462967	1.49526898	0.94569114	0.66870462
39	4.43958087	3.13925774	2.21979044	1.56962887	0.99272046	0.70195937
40	4.65295687	3.29013736	2.32647843	1.64506868	1.04043279	0.73569708

Table 6. Standard error of linear equating on Form Z under normality assumptions

z_i	$N_Z=N_Y=25$	$N_Z=N_Y=50$	$N_Z=N_Y=100$	$N_Z=N_Y=200$	$N_Z=N_Y=500$	$N_Z=N_Y=1000$
0	6.00020498	4.24278563	3.00010249	2.12139281	1.34168662	0.94871571
1	5.78556435	4.09101179	2.89278218	2.04550589	1.29369152	0.91477805
2	5.57217539	3.94012300	2.78608769	1.97006150	1.24597629	0.88103829
3	5.36018758	3.79022498	2.68009379	1.89511249	1.19857438	0.84752007
4	5.14977396	3.64144009	2.57488698	1.82072004	1.15152446	0.81425076
5	4.94113564	3.49391052	2.47056782	1.74695526	1.10487152	0.78126214
6	4.73450734	3.34780224	2.36725367	1.67390112	1.05866802	0.74859134
7	4.53016409	3.20330975	2.26508204	1.60165487	1.01297548	0.71628183
8	4.32842953	3.06066187	2.16421477	1.53033094	0.96786627	0.68438480
9	4.12968600	2.92012897	2.06484300	1.46006449	0.92342586	0.65296069
10	3.93438678	2.78203157	1.96719339	1.39101579	0.87975563	0.62208117
11	3.74307104	2.64675092	1.87153552	1.32337546	0.83697613	0.59183150
12	3.55638173	2.51474164	1.77819087	1.25737082	0.79523113	0.56231332
13	3.37508664	2.38654665	1.68754332	1.19327333	0.75469232	0.53364805
14	3.20010270	2.26281432	1.60005135	1.13140716	0.71556472	0.50598066
15	3.03252262	2.14431731	1.51626131	1.07215865	0.67809267	0.47948393
16	2.87364198	2.03197173	1.43682099	1.01598586	0.64256588	0.45436269
17	2.72498287	1.92685387	1.36249144	0.96342693	0.60932469	0.43085762
18	2.58830712	1.83020951	1.29415356	0.91510476	0.57876307	0.40924729
19	2.46560833	1.74344837	1.23280417	0.87172419	0.55132678	0.38984691
20	2.35906842	1.66811328	1.17953421	0.83405664	0.52750374	0.37300147
21	2.27096276	1.60581317	1.13548138	0.80290658	0.50780271	0.35907074
22	2.20350369	1.55811241	1.10175185	0.77905620	0.49271841	0.34840453
23	2.15862777	1.52638034	1.07931389	0.76319017	0.48268384	0.34130902
24	2.13775767	1.51162294	1.06887883	0.75581147	0.47801715	0.33800917
25	2.14159531	1.51433657	1.07079765	0.75716828	0.47887527	0.33861595
26	2.17000962	1.53442852	1.08500481	0.76721426	0.48522890	0.34310865
27	2.22205798	1.57123227	1.11102899	0.78561613	0.49686727	0.35133822
28	2.29613376	1.62361175	1.14806688	0.81180588	0.51343112	0.36305063
29	2.39018984	1.69011945	1.19509492	0.84505972	0.53446270	0.37792220
30	2.50197389	1.76916271	1.25098695	0.88458135	0.55945837	0.39559681
31	2.62922572	1.85914334	1.31461286	0.92957167	0.58791274	0.41571709
32	2.76981427	1.95855445	1.38490713	0.97927723	0.61934930	0.43794609
33	2.92181500	2.06603520	1.46090750	1.03301760	0.65333770	0.46197952
34	3.08354072	2.18039256	1.54177036	1.09019628	0.68950067	0.48755060
35	3.25354154	2.30060128	1.62677077	1.15030064	0.72751400	0.51443009
36	3.43058746	2.42579166	1.71529373	1.21289583	0.76710268	0.54242350
37	3.61364314	2.55523157	1.80682157	1.27761578	0.80803517	0.57136715
38	3.80184058	2.68830726	1.90092029	1.34415363	0.85011740	0.60112378
39	3.99445310	2.82450487	1.99722655	1.41225244	0.89318687	0.63157849
40	4.19087198	2.96339400	2.09543599	1.48169700	0.93710746	0.66263504

Table 7. Standard error for equipercentile equating on Form X under nonnormality assumptions

x_i	$N_X=N_Y=25$	$N_X=N_Y=50$	$N_X=N_Y=100$	$N_X=N_Y=200$	$N_X=N_Y=500$	$N_X=N_Y=1000$
0	*	*	*	*	*	*
1	*	*	*	*	*	*
2	2.39324058	1.69227664	1.19662029	0.84613832	0.53514486	0.37840456
3	13.52836270	9.56599700	6.76418135	4.78299850	3.02503386	2.13902196
4	58.01630680	41.02372400	29.00815340	20.51186200	12.97284060	9.17318355
5	25.77125490	18.22302910	12.88562740	9.11151455	5.76262778	4.07479318
6	21.65386050	15.31159160	10.82693020	7.65579579	4.84195040	3.42377596
7	8.32320517	5.88539482	4.16160259	2.94269741	1.86112526	1.31601429
8	6.61307744	4.67615191	3.30653872	2.33807595	1.47872907	1.04561935
9	5.49375765	3.88467329	2.74687882	1.94233664	1.22844156	0.86863935
10	4.80389240	3.39686489	2.40194620	1.69843244	1.07418300	0.75956208
11	4.43082156	3.13306397	2.21541078	1.56653198	0.99076182	0.70057440
12	4.29997486	3.04054139	2.14998743	1.52027069	0.96150361	0.67988572
13	3.56753665	2.52262936	1.78376833	1.26131468	0.79772545	0.56407707
14	3.69656585	2.61386678	1.84828292	1.30693339	0.82657725	0.58447838
15	3.54092734	2.50381374	1.77046367	1.25190687	0.79177542	0.55986977
16	3.64988709	2.58085991	1.82494354	1.29042995	0.81613956	0.57709782
17	3.47066362	2.45412978	1.73533181	1.22706489	0.77606398	0.54876010
18	3.48386188	2.46346236	1.74193094	1.23173118	0.77901520	0.55084693
19	3.25763575	2.30349633	1.62881787	1.15174816	0.72842950	0.51507744
20	3.20451201	2.26593218	1.60225601	1.13296609	0.71655067	0.50667784
21	3.17170196	2.24273196	1.58585098	1.12136598	0.70921412	0.50149011
22	3.07468686	2.17413193	1.53734343	1.08706596	0.68752088	0.48615068
23	3.04241511	2.15131236	1.52120756	1.07565618	0.68030470	0.48104807
24	3.01833353	2.13428410	1.50916676	1.06714205	0.67491989	0.47724043
25	2.99155065	2.11534575	1.49577533	1.05767288	0.66893106	0.47300569
26	2.94986438	2.08586911	1.47493219	1.04293455	0.65960973	0.46641451
27	2.92297951	2.06685864	1.46148976	1.03342932	0.65359809	0.46216364
28	2.78142616	1.96676530	1.39071308	0.98338265	0.62194580	0.43978209
29	2.63527812	1.86342303	1.31763906	0.93171152	0.58926610	0.41667406
30	2.52389957	1.78466650	1.26194979	0.89233325	0.56436110	0.39906356
31	2.25530403	1.59474078	1.12765202	0.79737039	0.50430131	0.35659488
32	2.14094283	1.51387520	1.07047142	0.75693760	0.47872937	0.33851278
33	2.19807917	1.55427669	1.09903958	0.77713834	0.49150544	0.34754683
34	1.89641704	1.34096935	0.94820852	0.67048468	0.42405174	0.29984986
35	2.03275435	1.43737438	1.01637717	0.71868719	0.45453769	0.32140668
36	2.74684755	1.94231453	1.37342377	0.97115726	0.61421378	0.43431473
37	1.89024509	1.33660512	0.94512255	0.66830256	0.42267165	0.29887399
38	2.73075955	1.93093860	1.36537978	0.96546930	0.61061640	0.43177100
39	1.23565139	0.87373747	0.61782569	0.43686874	0.27630005	0.19537364
40	0.43972470	0.31093232	0.21986235	0.15546616	0.09832543	0.06952658

* represents almost zero

Table 8. Standard error for equipercentile equating on Form Z under nonnormality assumptions

z_i	$N_Z=N_Y=25$	$N_Z=N_Y=50$	$N_Z=N_Y=100$	$N_Z=N_Y=200$	$N_Z=N_Y=500$	$N_Z=N_Y=1000$
0	*	*	*	*	*	*
1	*	*	*	*	*	*
2	0.68994440	0.48786436	0.34497220	0.24393218	0.15427626	0.10908979
3	3.85547229	2.72623060	1.92773614	1.36311530	0.86210981	0.60960369
4	17.40424190	12.30665750	8.70212094	6.15332873	3.89170679	2.75185226
5	48.97018300	34.62714850	24.48509150	17.31357420	10.95006580	7.74286578
6	31.92833080	22.57673920	15.96416540	11.28836960	7.13939180	5.04831236
7	13.61445150	9.62687100	6.80722577	4.81343550	3.04428391	2.15263380
8	6.39482090	4.52182122	3.19741045	2.26091061	1.42992542	1.01110996
9	5.69884380	4.02969109	2.84942190	2.01484555	1.27430021	0.90106632
10	5.25378107	3.71498422	2.62689053	1.85749211	1.17478116	0.83069573
11	4.93477347	3.48941178	2.46738674	1.74470589	1.10344889	0.78025620
12	3.63633267	2.57127549	1.81816634	1.28563775	0.81310870	0.57495468
13	3.62754618	2.56506250	1.81377309	1.28253125	0.81114398	0.57356541
14	3.65082827	2.58152542	1.82541413	1.29076271	0.81635002	0.57724663
15	3.68666930	2.60686886	1.84333465	1.30343443	0.82436432	0.58291360
16	3.72034189	2.63067898	1.86017095	1.31533949	0.83189374	0.58823770
17	3.74161656	2.64572244	1.87080828	1.32286122	0.83665090	0.59160152
18	3.74502398	2.64813185	1.87251199	1.32406593	0.83741282	0.59214028
19	3.40598885	2.40839781	1.70299442	1.20419890	0.76160226	0.53853412
20	3.34359555	2.36427909	1.67179778	1.18213954	0.74765069	0.52866888
21	3.28068814	2.31979683	1.64034407	1.15989841	0.73358417	0.51872234
22	3.22471189	2.28021564	1.61235595	1.14010782	0.72106750	0.50987172
23	3.18181755	2.24988477	1.59090877	1.12494238	0.71147603	0.50308953
24	3.07280653	2.17280233	1.53640326	1.08640117	0.68710043	0.48585337
25	3.04101311	2.15032099	1.52050656	1.07516050	0.67999120	0.48082639
26	3.01645712	2.13295728	1.50822856	1.06647864	0.67450032	0.47694375
27	2.99133987	2.11519671	1.49566993	1.05759835	0.66888393	0.47297236
28	2.99696040	2.11917102	1.49848020	1.05958551	0.67014072	0.47386105
29	2.86553254	2.02623749	1.43276627	1.01311875	0.64075256	0.45308048
30	2.70764443	1.91459374	1.35382222	0.95729687	0.60544770	0.42811617
31	2.53394278	1.79176813	1.26697139	0.89588406	0.56660683	0.40065153
32	2.37067055	1.67631722	1.18533528	0.83815861	0.53009805	0.37483593
33	2.26637153	1.60256668	1.13318576	0.80128334	0.50677608	0.35834480
34	2.30706573	1.63134182	1.15353287	0.81567091	0.51587558	0.36477912
35	1.90452992	1.34670602	0.95226496	0.67335301	0.42586584	0.30113262
36	1.97665817	1.39770839	0.98832908	0.69885420	0.44199420	0.31253710
37	2.50600189	1.77201093	1.25300095	0.88600547	0.56035906	0.39623369
38	1.75453639	1.24064458	0.87726820	0.62032229	0.39232626	0.27741656
39	2.34088510	1.65525573	1.17044255	0.82762786	0.52343782	0.37012643
40	0.98294781	0.69504906	0.49147390	0.34752453	0.21979381	0.15541770

* represents almost zero

Table 9. Standard error for equipercentile equating on Form X under normality assumptions

x_i	N_X=N_Y=25	N_X=N_Y=50	N_X=N_Y=100	N_X=N_Y=200	N_X=N_Y=500	N_X=N_Y=1000
0	*	*	*	*	*	*
1	*	*	*	*	*	*
2	0.34885285	0.24667622	0.17442643	0.12333811	0.07800587	0.05515848
3	1.13685261	0.80387619	0.56842631	0.40193809	0.25420797	0.17975218
4	2.19125837	1.54945365	1.09562919	0.77472683	0.48998027	0.34646837
5	3.23429040	2.28698868	1.61714520	1.14349434	0.72320932	0.51138621
6	3.99643534	2.82590653	1.99821767	1.41295327	0.89363011	0.63189191
7	4.40855780	3.11732111	2.20427890	1.55866056	0.98578349	0.69705419
8	4.54106024	3.21101449	2.27053012	1.60550725	1.01541194	0.71800467
9	4.53679999	3.20800203	2.26839999	1.60400102	1.01445932	0.71733106
10	4.43707314	3.13748451	2.21853657	1.56874225	0.99215972	0.70156286
11	4.30516649	3.04421242	2.15258325	1.52210621	0.96266449	0.68070659
12	4.15219515	2.93604535	2.07609757	1.46802267	0.92845906	0.65651970
13	3.98865282	2.82040346	1.99432641	1.41020173	0.89188988	0.63066139
14	3.81336022	2.69645287	1.90668011	1.34822643	0.85269327	0.60294519
15	3.63166370	2.56797403	1.81583185	1.28398702	0.81206469	0.57421645
16	3.44668582	2.43717492	1.72334291	1.21858746	0.77070238	0.54496888
17	3.26708413	2.31017734	1.63354206	1.15508867	0.73054222	0.51657136
18	3.10079508	2.19259323	1.55039754	1.09629661	0.69335886	0.49027875
19	2.95495607	2.08946947	1.47747803	1.04473474	0.66074826	0.46721958
20	2.83651152	2.00571653	1.41825576	1.00285827	0.63426326	0.44849185
21	2.74919188	1.94397222	1.37459594	0.97198611	0.61473799	0.43468540
22	2.69518935	1.90578667	1.34759467	0.95289333	0.60266266	0.42614685
23	2.67617022	1.89233811	1.33808511	0.94616906	0.59840985	0.42313967
24	2.69262188	1.90397119	1.34631094	0.95198559	0.60208856	0.42574090
25	2.74393370	1.94025413	1.37196685	0.97012706	0.61356223	0.43385401
26	2.83034240	2.00135431	1.41517120	1.00067715	0.63288380	0.44751643
27	2.95414328	2.08889475	1.47707164	1.04444737	0.66056652	0.46709107
28	3.11824711	2.20493368	1.55912356	1.10246684	0.69726125	0.49303816
29	3.32562906	2.35157486	1.66281453	1.17578743	0.74363326	0.52582812
30	3.58497509	2.53496020	1.79248754	1.26748010	0.80162480	0.56683433
31	3.90497035	2.76123101	1.95248517	1.38061551	0.87317792	0.61743002
32	4.29891411	3.03979132	2.14945705	1.51989566	0.96126642	0.67971800
33	4.76920822	3.37233947	2.38460411	1.68616974	1.06642738	0.75407803
34	5.31021342	3.75488792	2.65510671	1.87744396	1.18739982	0.83961846
35	5.87000233	4.15071845	2.93500117	2.07535923	1.31257242	0.92812886
36	6.35192623	4.49149011	3.17596311	2.24574505	1.42033388	1.00432772
37	6.54827544	4.63032997	3.27413772	2.31516498	1.46423890	1.03537326
38	6.14703270	4.34660850	3.07351635	2.17330425	1.37451830	0.97193121
39	4.89175185	3.45899091	2.44587593	1.72949545	1.09382897	0.77345388
40	2.87543233	2.03323770	1.43771617	1.01661885	0.64296622	0.45464577

* represents almost zero

Table 10. Standard error for equipercentile equating on Form Z under normality assumptions

z_i	$N_Z=N_Y=25$	$N_Z=N_Y=50$	$N_Z=N_Y=100$	$N_Z=N_Y=200$	$N_Z=N_Y=500$	$N_Z=N_Y=1000$
0	*	*	*	*	*	*
1	*	*	*	*	*	*
2	0.12588143	0.08901161	0.06294072	0.04450581	0.02814794	0.01990360
3	0.42052686	0.29735740	0.21026343	0.14867870	0.09403267	0.06649114
4	1.08323427	0.76596230	0.54161713	0.38298115	0.24221855	0.17127438
5	2.20575189	1.55970212	1.10287595	0.77985106	0.49322112	0.34876000
6	3.53216154	2.49761538	1.76608077	1.24880769	0.78981533	0.55848378
7	4.70103362	3.32413275	2.35051681	1.66206638	1.05118307	0.74329868
8	5.64386078	3.99081223	2.82193039	1.99540611	1.26200564	0.89237274
9	6.20151262	4.38513163	3.10075631	2.19256581	1.38670038	0.98054524
10	6.32898957	4.47527144	3.16449478	2.23763572	1.41520509	1.00070112
11	6.04108235	4.27169030	3.02054118	2.13584515	1.35082708	0.95517899
12	5.69102975	4.02416573	2.84551487	2.01208286	1.27255294	0.89983081
13	5.24810205	3.71096855	2.62405103	1.85548427	1.17351129	0.82979779
14	4.78644772	3.38452964	2.39322386	1.69226482	1.07028225	0.75680384
15	4.28878529	3.03262916	2.14439264	1.51631458	0.95900154	0.67811650
16	3.91419618	2.76775466	1.95709809	1.38387733	0.87524087	0.61888876
17	3.59264931	2.54038669	1.79632466	1.27019335	0.80334081	0.56804773
18	2.48332600	1.75597666	1.24166300	0.87798833	0.55528858	0.39264832
19	3.08992507	2.18490697	1.54496253	1.09245348	0.69092825	0.48856005
20	2.93114165	2.07263014	1.46557082	1.03631507	0.65542320	0.46345419
21	2.81472259	1.99030943	1.40736130	0.99515472	0.62939111	0.44504672
22	2.73567580	1.93441491	1.36783790	0.96720746	0.61171571	0.43254832
23	2.69018708	1.90224953	1.34509354	0.95112476	0.60154412	0.42535593
24	2.67322832	1.89025787	1.33661416	0.94512894	0.59775202	0.42267451
25	2.68879556	1.90126558	1.34439778	0.95063279	0.60123297	0.42513591
26	2.73095967	1.93108010	1.36547984	0.96554005	0.61066115	0.43180264
27	2.79770960	1.97827943	1.39885480	0.98913972	0.62558689	0.44235673
28	2.90500035	2.05414545	1.45250018	1.02707272	0.64957783	0.45932089
29	3.02304806	2.13761778	1.51152403	1.06880889	0.67597410	0.47798587
30	3.16485393	2.23788968	1.58242697	1.11894484	0.70768285	0.50040734
31	3.32524846	2.35130574	1.66262423	1.17565287	0.74354816	0.52576795
32	3.54397029	2.50596543	1.77198515	1.25298271	0.79245585	0.56035090
33	3.73612946	2.64184248	1.86806473	1.32092124	0.83542395	0.59073394
34	3.92064183	2.77231242	1.96032091	1.38615621	0.87668216	0.61990790
35	4.07544714	2.88177631	2.03772357	1.44088815	0.91129768	0.64438477
36	4.24232948	2.99977995	2.12116474	1.49988997	0.94861371	0.67077119
37	4.22186782	2.98531137	2.11093391	1.49265568	0.94403834	0.66753592
38	4.02072428	2.84308141	2.01036214	1.42154070	0.89906128	0.63573233
39	3.52449628	2.49219522	1.76224814	1.24609761	0.78810133	0.55727179
40	2.49573903	1.76475399	1.24786951	0.88237700	0.55806421	0.39461099

* represents almost zero

Table 11. Sample size required using equations (26) through (28): z-score range -2 to +2
--- The alternative procedure of theoretical methods

Standard Unit	Mean	Linear	Equipercentile
0.05	1600	4800	12226
.10	400	1200	3057
.20	100	300	765
.30	45	134	340
.40	25	75	192
.50	16	48	123
.60	12	34	85
.70	9	25	63
.80	7	19	48
.90	5	15	38
1.0	4	12	31

Table 12. Sample size required using equations (23) through (25): z-score range -3 to +3
--- The first procedure of theoretical methods

Standard Unit	Mean	Linear	Equipercentile
0.05	1600	2400	4825
.10	400	600	1207
.20	100	150	302
.30	45	67	135
.40	25	38	76
.50	16	24	49
.60	12	17	34
.70	9	13	25
.80	7	10	19
.90	5	8	15
1.0	4	6	13

Table 13. Summary of the minimum sample size required based on real data

	Mean	Linear		Equipercentile	
Form		Nonnormality	Normality	Nonnormality	Normality
X	22	29	32	44	51
Z	81	89	99	145	161

Table 14. Sample size results based on Method 1 (M1) and Method 2 (M2)

	Mean			Linear			Equipercentile		
Form	M1		M2	M1		M2	M1		M2
X	u=.4	25	22	u=.44	31	32	u=.45	60	51
Z	u=.2	100	81	u=.25	96	99	u=.25	193	161

Note: In Method 1, the u values are derived from real data [equation (29)]. Then, these values are used in formulas (23), (24), and (25). In Method 2, formulas (18), (20), and (22) are used and do not require a u-value.

Figure1. Log-linear C=7

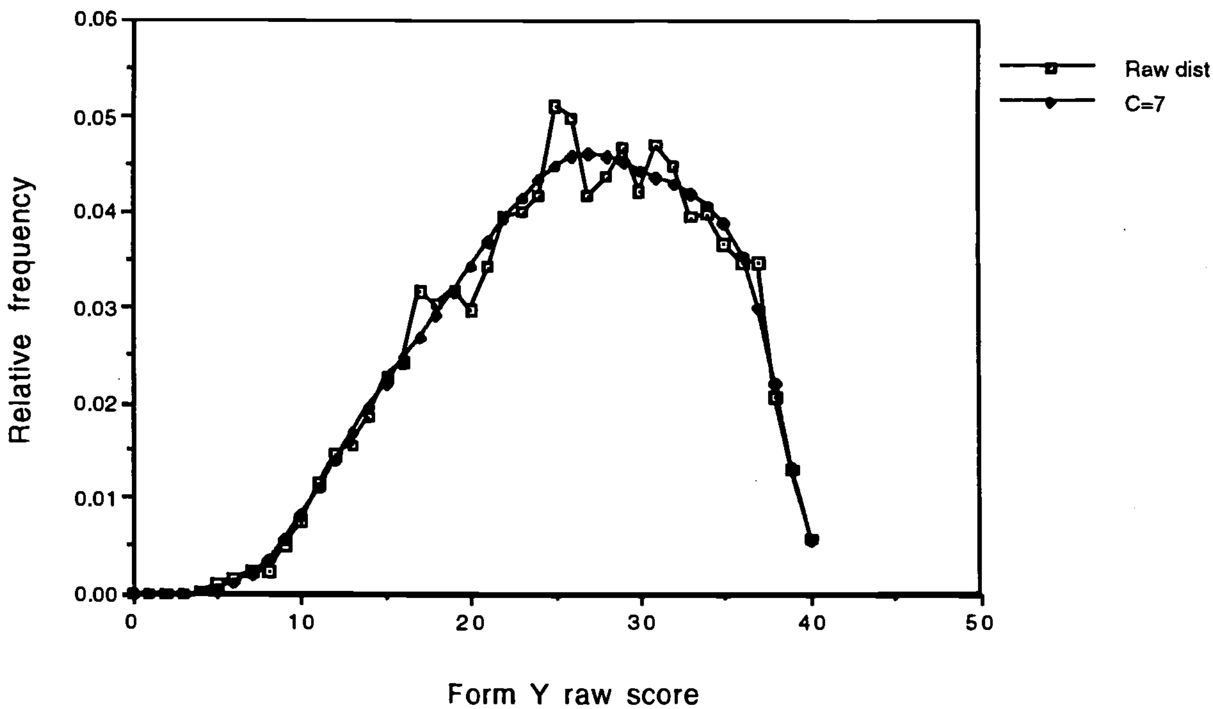


Figure 2. Log-linear C=8

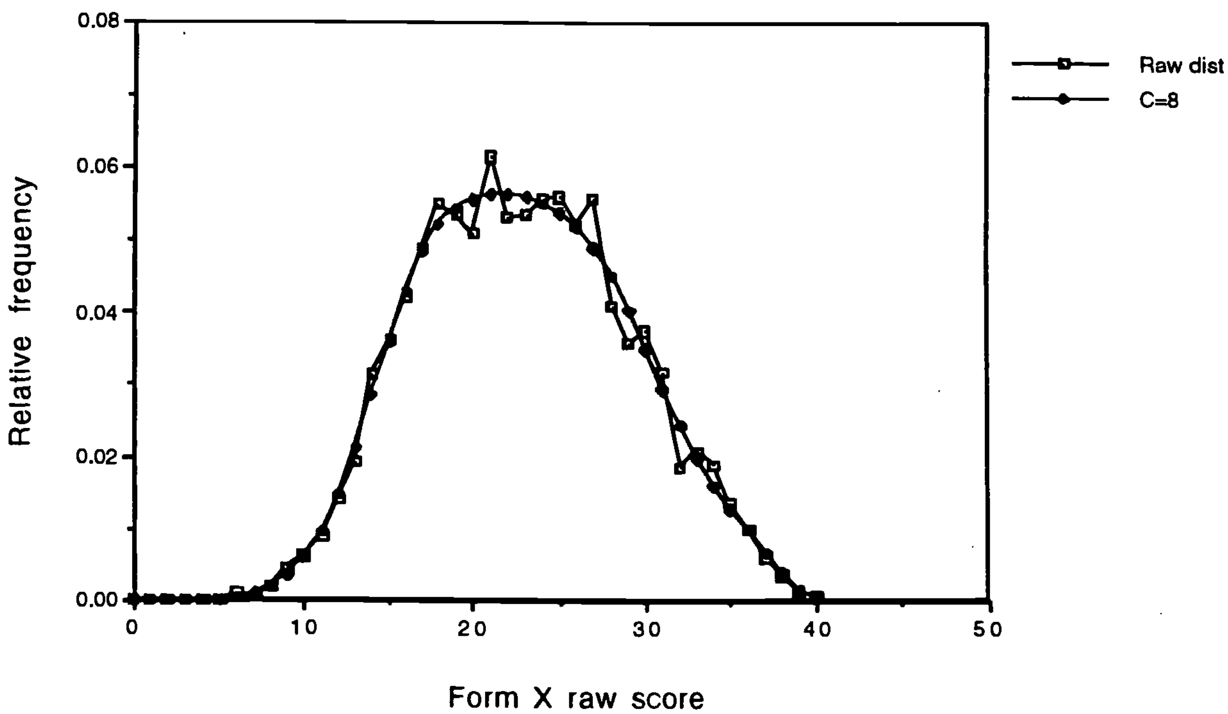


Figure 3. Log-linear C=6

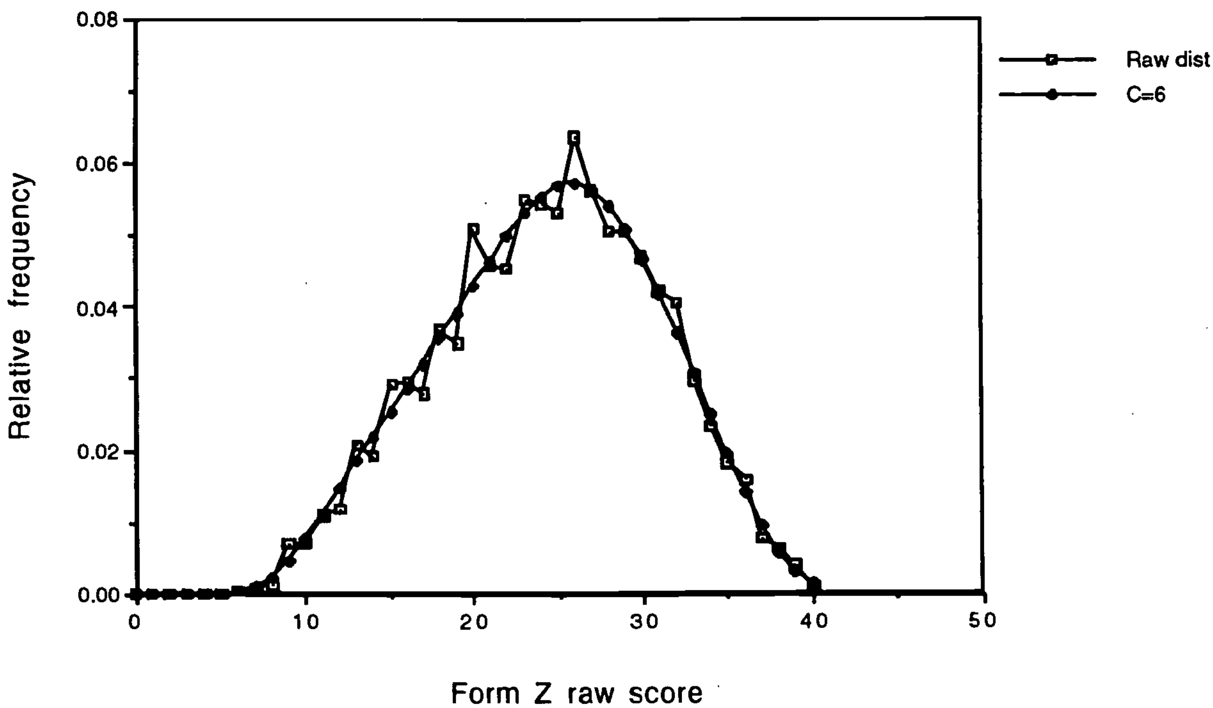


Figure 4. Standard error of linear equating under nonnormality assumptions

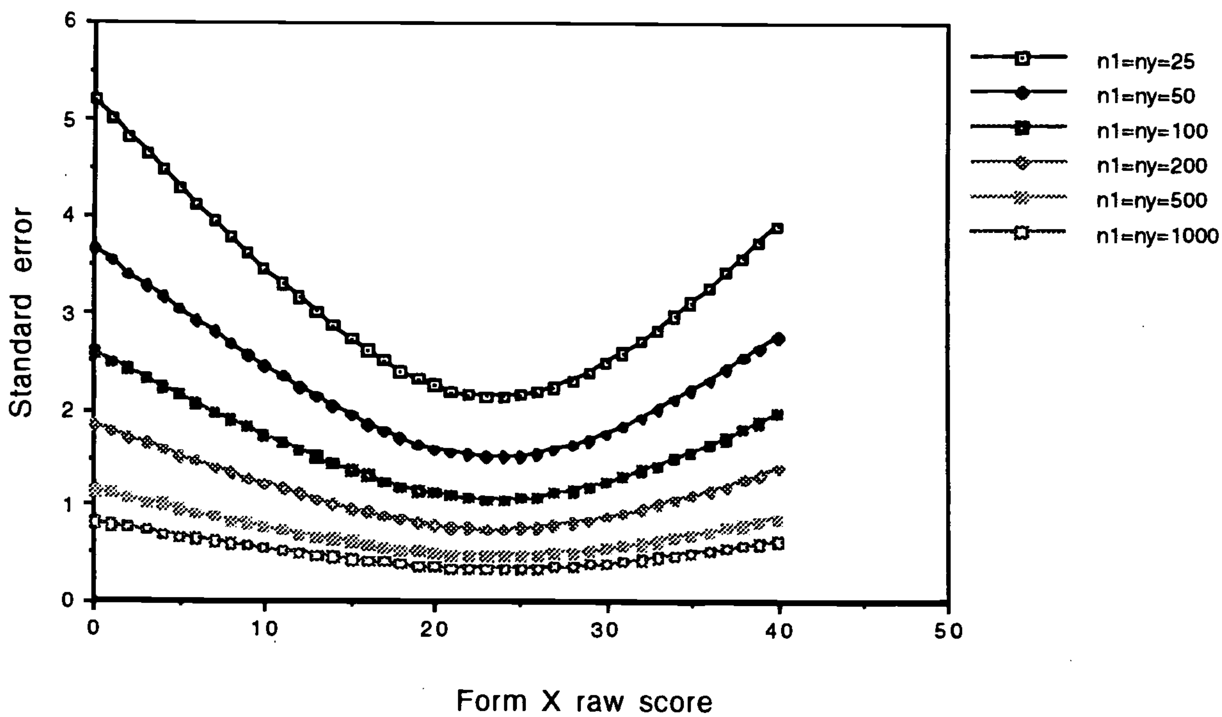


Figure 5. Standrad error of linear equating under nonnormality assumptions

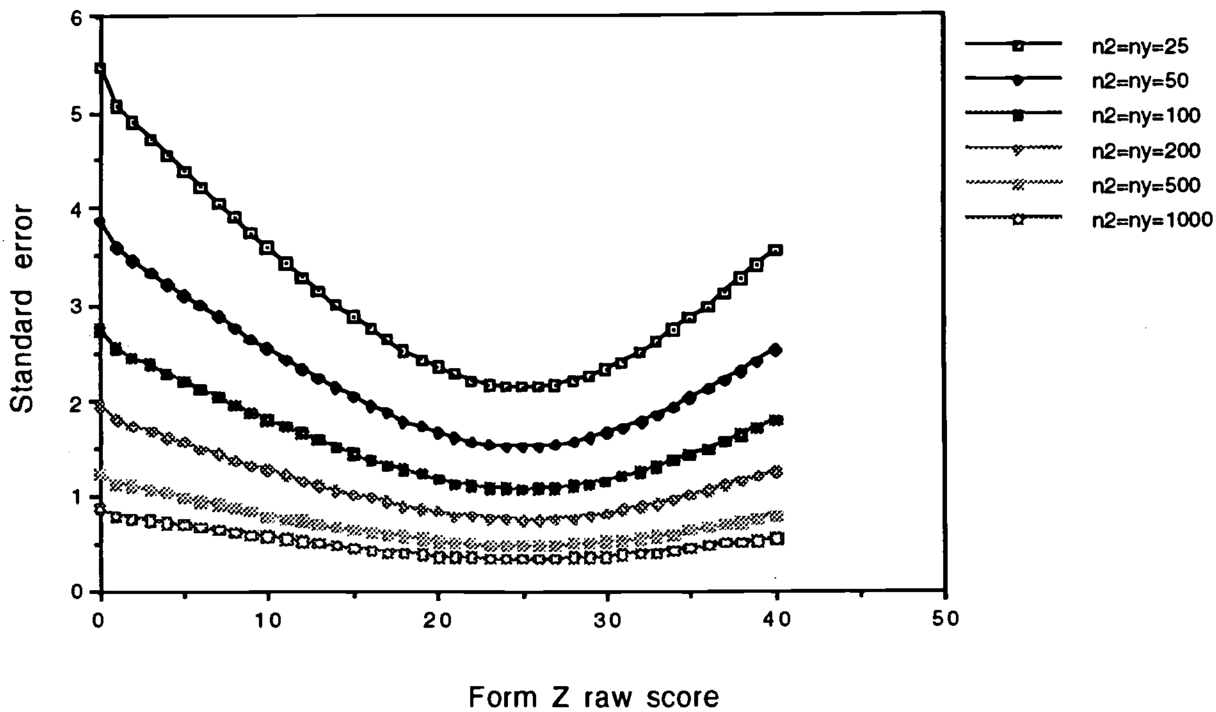


Figure 6. Standard error of linear equating under normality assumptions

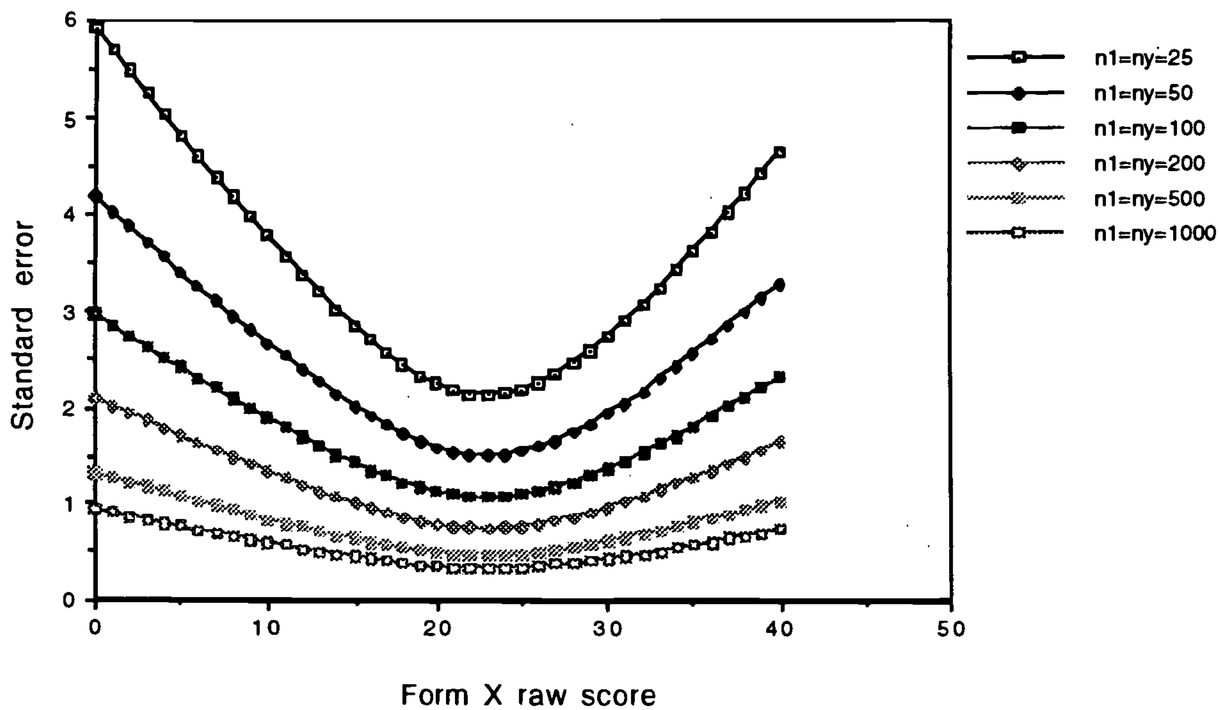


Figure 7. Standard error of linear equating under normality assumptions

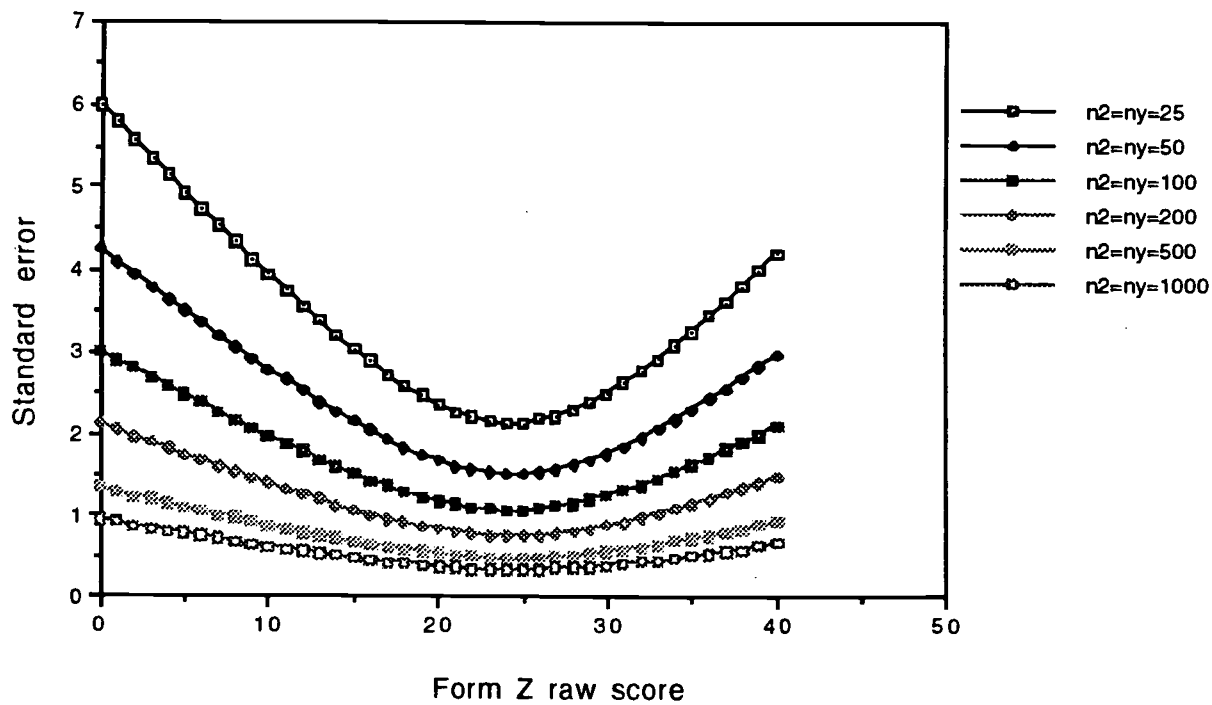


Figure 8. Standard error of equipercentile equating under nonnormality assumptions

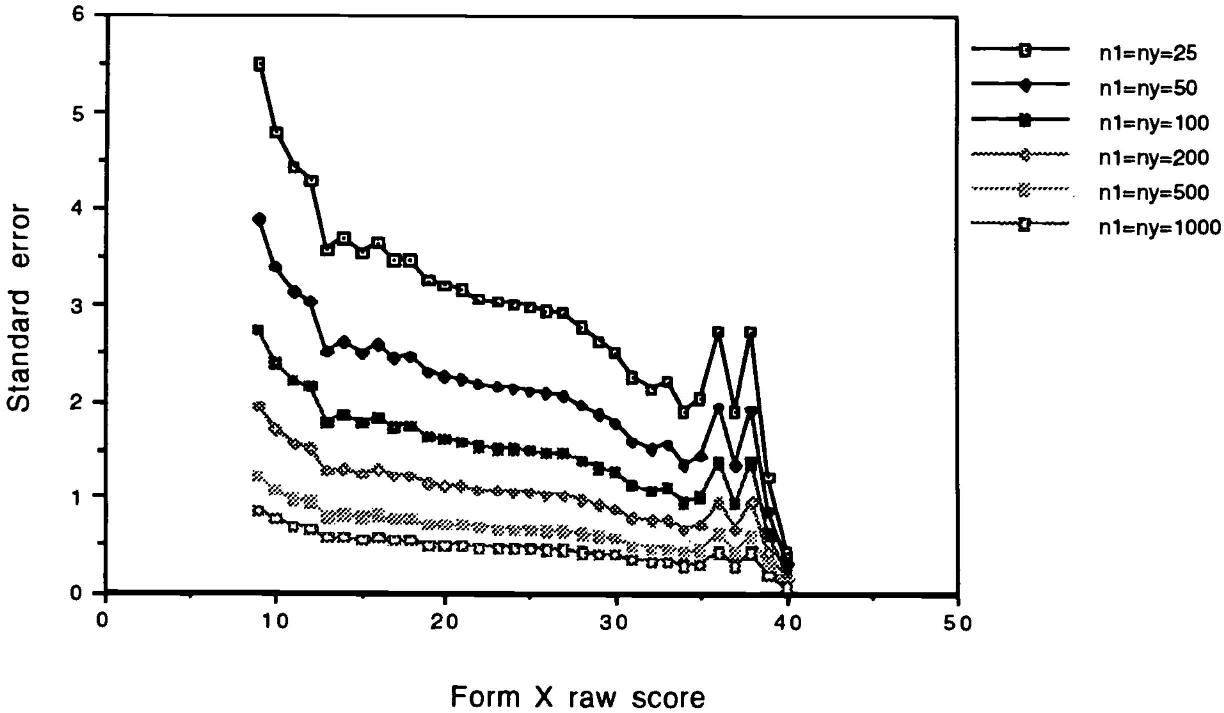


Figure 9. Standard error of equipercentile equating under nonnormality assumptions

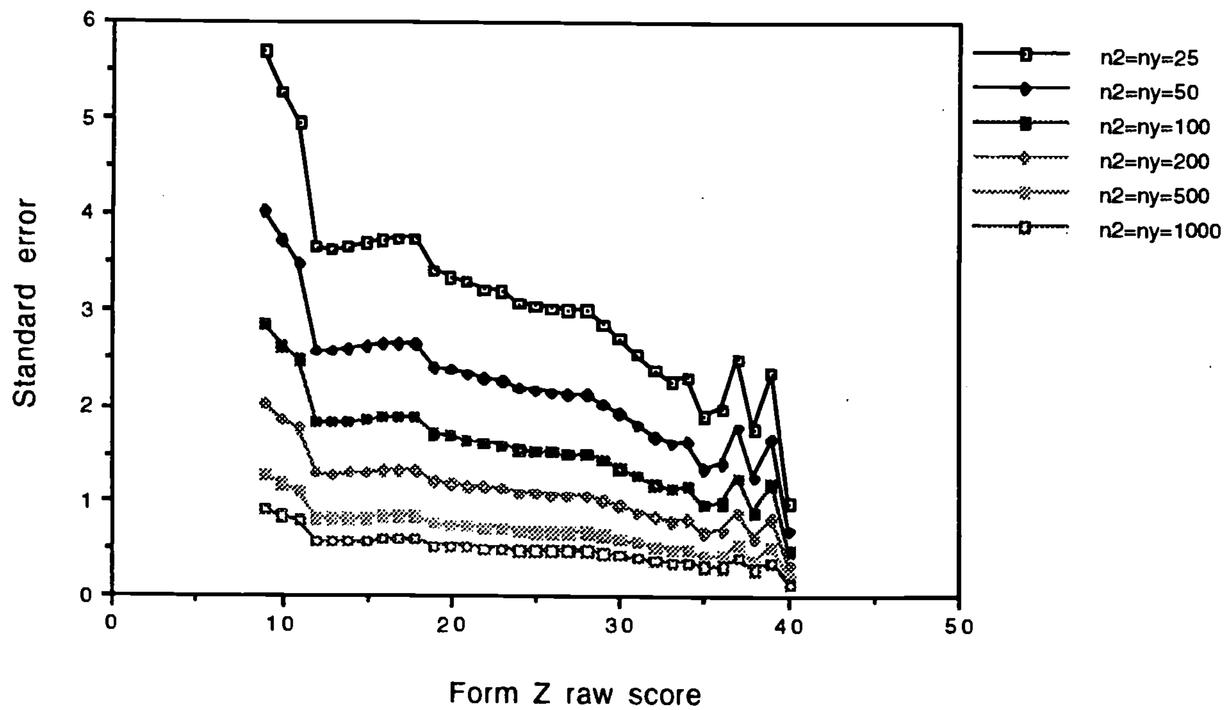


Figure 10. Standard error of equipercentile equating under normality assumptions

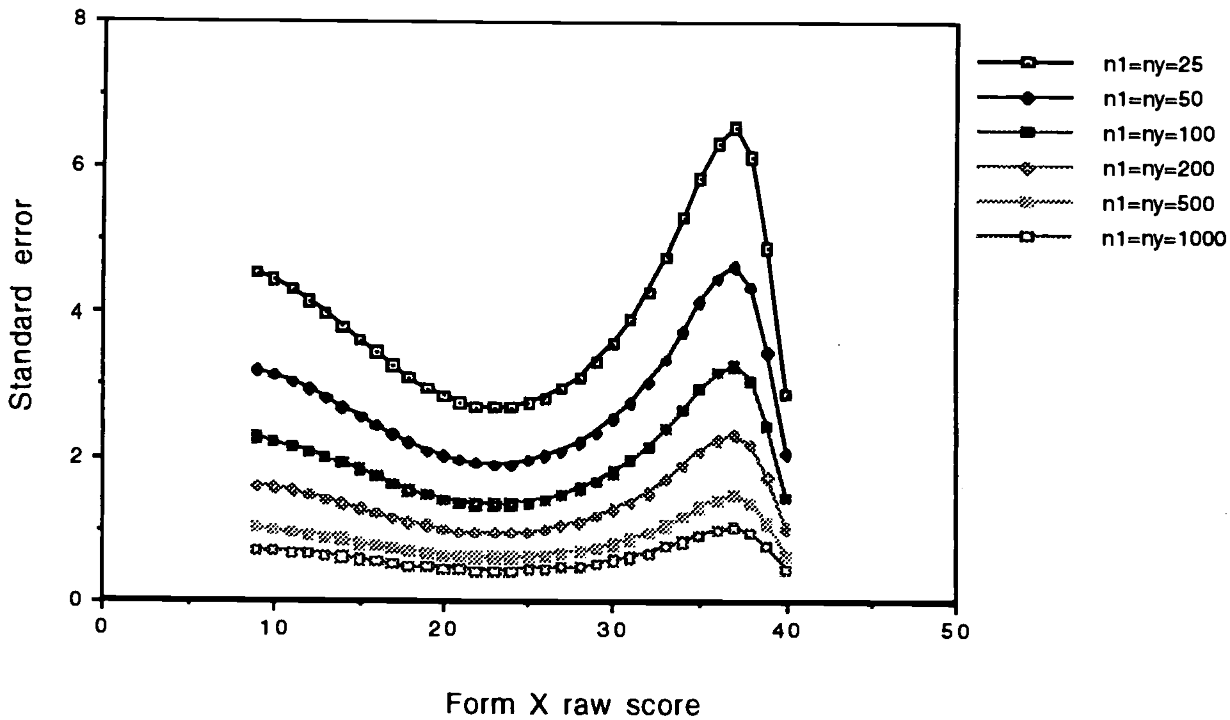
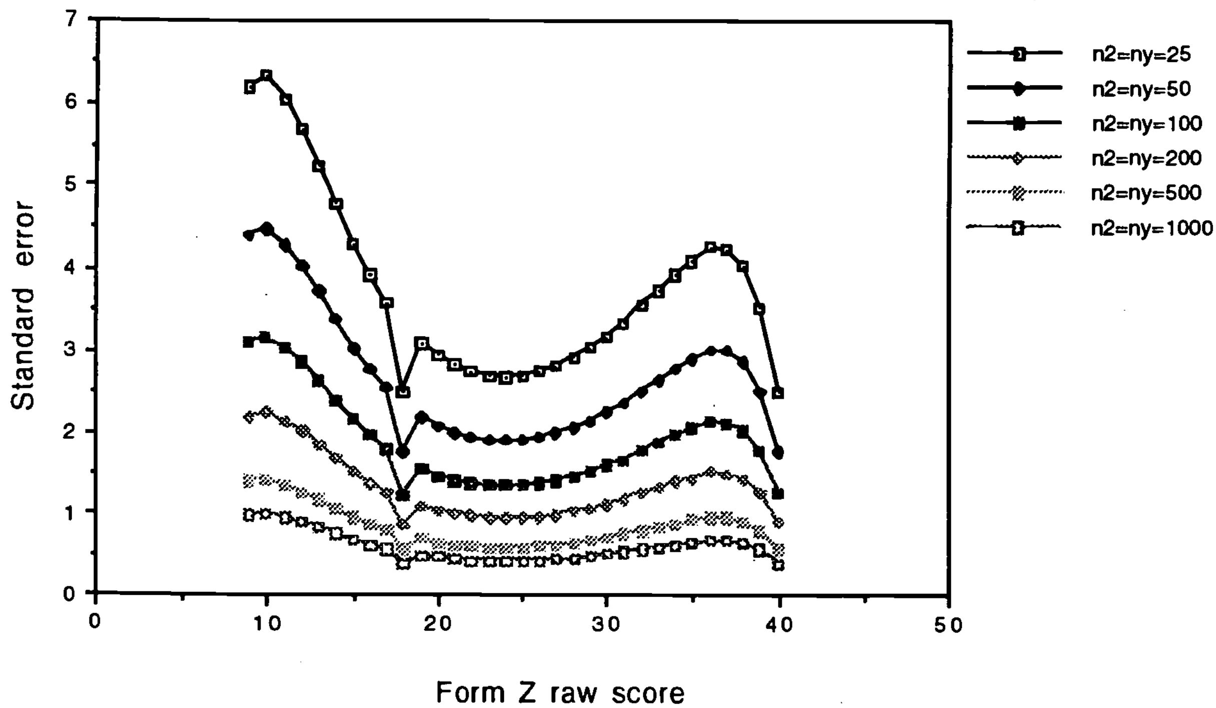


Figure 11.Standard error of equipercentile equating under normality assumptions



Reproduction Release

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Estimating minimum sample sizes in random groups equating	
Author(s): Tsung-Hsun Tsai	
Corporate Source:	Publication Date: March 1997

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	**Level 2A**	**Level 2B**
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: Tsung-Hsun Tsai	Printed Name/Position/Title: Tsung-Hsun Tsai/Research Scientist	
Organization/Address: CTB/McGraw-Hill 20 Ryan Ranch Rd Monterey, CA 93940	Telephone: 831-393-7651	Fax: 831-393-7016
	E-mail Address: ttsai@ctb.com	Date: 2/1/00

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:	
ERIC Clearinghouse on Assessment and Evaluation **1129 Shriver Laboratory (Bldg 075)** **College Park, Maryland 20742**	**Telephone: 301-405-7449** **Toll Free: 800-464-3742** **Fax: 301-405-8134** **ericae@ericae.net** **http://ericae.net**

EFF-088 (Rev. 9/97)