ACT Research Report Series 91-4

Standard Errors of Linear Equating for the Single-Group Design

Lingjia Zeng

August 1991

ACT

For additional copies write:
ACT Research Report Series
P.O. Box 168
Iowa City, Iowa 52243

Standard Errors of Linear Equating for the Single-Group Design

Lingjia Zeng

American College Testing

August, 1991

Abstract

Large sample standard errors of linear equating for the single-group design are derived without the normality assumption. Two general methods based on the delta method are described. One method uses the exact partial derivatives, and the other uses numerical derivatives. Simulation and real test data are used to evaluate the behavior of the estimated standard errors. Comparisons with standard errors derived with the normality assumption and bootstrap method are also conducted. The results indicate that the standard errors derived in this paper agree with those computed by the bootstrap method and are more accurate than the standard errors based on the normality assumption, especially in the situation in which the score distributions are skewed.

Acknowledgement

I would like to thank Michael J. Kolen and Bradley A. Hanson for their help in preparing this document.

Standard Errors of Linear Equating for the Single-Group Design

In linear equating, scores on one test form are transformed linearly to scores on the scale of another form. The purpose of linear equating is to adjust for presumably small differences in test difficulty between the two forms of the same test. The single-group design is one of the basic data collection schemes. In this design, examinees are administered both forms of a test to be equated. An advantage of this design is that the equating errors are relatively small compared to that of some other designs. However, the order of administering the two forms may have an influence on the examinees' performance. For example, if familiarity with the test increases performance, then the form administered last would appear easier than the form administered first, supposing the two forms are equally difficult. Such an effect is usually referred to as a practice effect. One way to minimize the practice effect is to administer the two forms, say, A and B, in a way such that a random half of the examinees take Form A first and another half of the examinees take Form B first. Thus, the order of administration of the two forms are counterbalanced. Lord (1950), Angoff (1984), and Petersen, Kolen and Hoover (1985) have presented descriptions of this design. Holland and Thayer (1990) addressed the issue of counterbalancing in detail.

Because equating is usually conducted using a sample of examinees drawn from a population, the results are subject to sampling error. The errors of equating can be quantified using standard errors. Standard errors of linear equating for the single-group design were derived by Lord (1950) under the assumption that the two test scores have a normal bivariate distribution in the population from which the sample is drawn. Because skewed score distributions are typical in many testing programs (Kolen, 1985), the normality assumption is seldom met. However, the standard error of equating for the single-group design with less restrictive assumptions has not been derived. Braun and Holland (1982) derived standard errors for the random groups design, and Kolen (1985) derived standard errors for the common-item nonequivalent-groups design, without making the restrictive normality assumption. Their results suggested that standard errors of linear equating based on the normality assumption might produce misleading results when score distributions are skewed or more peaked than a normal distribution. The purpose of this paper is to derive large sample standard errors of linear equating for the single-group design without making the normality assumption. Two general methods based on the delta method (Kendall and Stuart, 1977) are described. In one method the exact derivatives are used and in the other the numerical derivatives are used. Simulation and real test data are used to evaluate the behavior of the estimated standard errors. A comparison with standard errors derived with the normality assumption is also conducted.

Large Sample Standard Errors

Kendall and Stuart (1977) described a general method for estimating standard errors of functions of random variables by means of a Taylor expansion. This method is usually referred to as the delta method. According to Lord (1950), the linear equating function for equating two test forms, X and Y, under the single-group design can be written as

$$l[x|\mu(X), \sigma^2(X), \mu(Y), \sigma^2(Y)] = \frac{\sigma(Y)}{\sigma(X)}[x - \mu(X)] + \mu(Y) . \qquad (1)$$

It is assumed here that the form taken first has no effect on the performance on the form taken last. Let θ_1, θ_2, θ_3, and θ_4 be alternative names for the four parameter $\mu(X)$, $\sigma^2(X)$, $\mu(Y)$, $\sigma^2(Y)$ in function l, and let their estimates be $\hat{\theta}_1$, $\hat{\theta}_2$, $\hat{\theta}_3$, and $\hat{\theta}_4$. Define $\hat{l}$ as an estimate of l that uses the estimates of the parameters in Equation 1. According to the delta method described by Kendall and Stuart (1977) the sampling variance of $\hat{l}$ can be expressed as follows:

$$\text{var}[\hat{l}(x)] = \sum_{i=1}^{4} \left(\frac{\partial l}{\partial \theta_i}\right)^2 \text{var}(\hat{\theta}_i) + \sum_{i=1}^{4} \sum_{j \neq i=1}^{4} \frac{\partial l}{\partial \theta_i} \frac{\partial l}{\partial \theta_j} \text{cov}(\hat{\theta}_i, \hat{\theta}_j) + \text{remainder}. \qquad (2)$$

The term "remainder" in Equation 2 refers to higher order terms in the Taylor expansion that are small and thus can be ignored.

The standard error of equating $SE[\hat{l}(x)]$ is the square root of $\text{var}[\hat{l}(x)]$. To compute $\text{var}[\hat{l}(x)]$, we need to find the four first partial derivatives with respect to each of the four parameters, the four sampling variances and 12 covariances of the four parameters. These sampling variances and covariances are listed in Table 1. Because $\text{cov}(\hat{\theta}_i,\hat{\theta}_j) = \text{cov}(\hat{\theta}_j,\hat{\theta}_i)$, only six different covariances are listed. The calculation of the sampling variance of $\hat{l}$ (x), without assuming a normal distribution, involves estimating higher order central moments and cross-product moments. Because the higher order moments are very sensitive to sampling variation, a large sample size might be required to ensure accurate estimates of standard errors. So if the sample size is not large enough the standard errors computed from Equation 2 might not be accurate.

Insert Table 1 about here

The first partial derivatives of l with respect to each of the four parameters in the equating function are derived as follows:

$$\frac{\partial l}{\partial \theta_1} = -\frac{\sigma(Y)}{\sigma(X)}, \tag{3}$$

$$\frac{\partial l}{\partial \theta_2} = -\frac{\sigma(Y)}{2\sigma^3(X)}[x - \mu(X)] = -\frac{\sigma(Y)}{2\sigma^2(X)}Z, \tag{4}$$

$$\frac{\partial l}{\partial \theta_3} = 1, \tag{5}$$

and

$$\frac{\partial l}{\partial \theta_4} = \frac{1}{2\sigma(X)\sigma(Y)}[x - \mu(X)] = \frac{1}{2\sigma(Y)}Z, \tag{6}$$

where $Z = \dfrac{x - \mu(X)}{\sigma(X)}$.

Substituting the four partial derivatives into Equation 2, a formula for computing the sampling variance of linear equating for the single-group design is obtained as follows:

$$\mathrm{var}[\hat{l}(x)] = Z^2 \{ \frac{\sigma^2(Y)}{4\sigma^4(X)} \mathrm{var}[\hat{\sigma}^2(X)] - \frac{1}{2\sigma^2(X)} \mathrm{cov}[\hat{\sigma}^2(X),\hat{\sigma}^2(Y)] + \frac{1}{4\sigma^2(Y)} \mathrm{var}[\hat{\sigma}^2(Y)] \}$$

$$+ Z \{ \frac{\sigma^2(Y)}{\sigma^3(X)} \mathrm{cov}[\hat{\mu}(X),\hat{\sigma}^2(X)] - \frac{1}{\sigma(X)} \mathrm{cov}[\hat{\mu}(X),\hat{\sigma}^2(Y)] - \frac{\sigma(Y)}{\sigma^2(X)} \mathrm{cov}[\hat{\mu}(Y),\hat{\sigma}^2(X)]$$

$$+ \frac{1}{\sigma(Y)} \mathrm{cov}[\hat{\mu}(Y),\hat{\sigma}^2(Y)] \} + \frac{\sigma^2(Y)}{\sigma^2(X)} \mathrm{var}[\hat{\mu}(X)] - \frac{2\sigma(Y)}{\sigma(X)} \mathrm{cov}[\hat{\mu}(X),\hat{\mu}(Y)] + \mathrm{var}[\hat{\mu}(Y)], \quad (7)$$

where Z is the same as defined in Equations 4 and 6, the var's and cov's are defined in Table 1 under the label of "general".

The partial derivatives can also be computed numerically. Lord (1950) used numerical derivatives in the computation of asymptotic sampling variance. Let $\underline{\theta}$ denote the entire vector of values θ_i, i=1 to 4. Then the first partial derivatives of function l with respect to θ_i can be approximated by

$$\frac{\partial l}{\partial \theta_i}(x|\underline{\theta}) = \frac{l(x|\underline{\theta} + \underline{\delta}_i) - l(x|\underline{\theta} - \underline{\delta}_i)}{2h_i} - O(h^2), \quad (8)$$

where $\underline{\delta}_i$ is the ith row of the diagonal matrix $\underline{\Delta}$, where

$$\underline{\Delta} = \begin{bmatrix} h_1 & & & \\ & h_2 & & \\ & & \ddots & \\ & & & h_4 \end{bmatrix},$$

and $O(h^2)$ is the error of approximation. Because Equation 8 is derived by expanding $l(x|\underline{\theta})$ at two neighboring points $(\underline{\theta} + \underline{\delta}_i)$ and $(\underline{\theta} - \underline{\delta}_i)$ with a second order Taylor's series, the magnitude of the error is based on the magnitude of the third partial derivative. If the third partial derivative with respect to θ_i is zero, then the first partial derivative with respect to θ_i approximated by Equation 8 is the exact value. For example, approximate $\partial l \,/\, \partial \theta_3$ as

$$\frac{\partial l}{\partial \theta_3} = \frac{\partial l}{\partial \mu(Y)} = \frac{\{\frac{\sigma(Y)}{\sigma(X)}[x - \mu(X)] + \mu(Y) + h_3\} - \{\frac{\sigma(Y)}{\sigma(X)}[x - \mu(X)] + \mu(Y) - h_3\}}{2h_3}$$

$$= \frac{2h_3}{2h_3} = 1.0,$$

which is the exact derivative defined by Equation 5. Alternatively, if the function has a nonzero third partial derivative, then error will be involved in the approximation. The error of approximation is bounded by C h^2_i, where C is the maximum absolute value of the third partial derivative with respect to θ_i. Equation 8 suggests that the numerical derivative approaches the exact derivative as h_i approaches zero. But in actual computation, a too small h_i cannot be used, because h_i is used as a denominator. If the denominator is too small, the computer rounding error will become significant. As a result, the obtained numerical derivative may be incorrect. In present paper, h_i is set to θ_i /1000. This value was selected to yield desirable accuracy. More detailed discussion of numerical derivative with more than one variable can be found in many advanced calculus textbooks (e.g., Taylor and Mann, 1983).

In the present paper, the delta method is implemented using Equation 7, which uses the exact derivatives, as well as using Equation 2 with numerical derivatives approximated by Equation 8. Note that many of the expressions are presented with population parameters. In actual computation, the sample estimates for the parameters are substituted in the formulas.

SE Based on the Normality Assumption

If the score distributions of Forms X and Y are assumed to be bivariate normal, then the sampling covariances $cov[\hat{\mu}(X),\hat{\sigma}^2(X)]$, $cov[\hat{\mu}(X),\hat{\sigma}^2(Y)]$, $cov[\hat{\mu}(Y),\hat{\sigma}^2(X)]$ and $cov[\hat{\mu}(Y),\hat{\sigma}^2(Y)]$ have zero values (Kendall, and Stuart, 1977, p. 85). The variances and covariances based on the normality assumption are listed in Table 1 under the label of "Normal". Substitute these variances and covariance and the exact derivatives into Equation 2 to obtain

$$\mathrm{var}[\hat{l}(x)] = \frac{\sigma^2(Y)}{\sigma^2(X)} \frac{\sigma^2(X)}{n} + \frac{\sigma^2(Y)}{4\sigma^4(X)} Z^2 \frac{2\sigma^4(X)}{n} + \frac{\sigma^2(Y)}{n} + \frac{1}{4\sigma^2(Y)} Z^2 \frac{2\sigma^4(Y)}{n}$$

$$-2 \frac{\sigma(Y)}{\sigma(X)} \frac{\sigma(X,Y)}{n} - 2 \frac{\sigma(Y)}{2\sigma^2(X)} Z \frac{1}{2\sigma(Y)} Z \frac{2\sigma^2(X,Y)}{n}$$

$$= \frac{1}{n} [2\sigma^2(Y) + \sigma^2(Y) Z^2 - 2 \frac{\sigma(Y)\,\sigma(X,Y)}{\sigma(X)} - \frac{\sigma^2(X,Y)\, Z^2}{\sigma^2(X)}] \qquad (9)$$

Substituting $\rho_{xy} = \dfrac{\sigma(X,Y)}{\sigma(X)\sigma(Y)}$ into the above expression gives

$$\mathrm{var}[\hat{l}(x)] = \frac{\sigma^2(Y)}{n} [2 + Z^2 - 2\rho_{xy} - \rho^2_{xy} Z^2] = \frac{\sigma^2(Y)}{n} (1 - \rho_{xy}) [Z^2 (1 + \rho_{xy}) + 2] \qquad (10)$$

Equation 10 is the same as the standard error formula derived by Lord (1950) with the normality assumption. The standard error of equating is the square root of $\mathrm{var}[\hat{l}(x)]$. Equation 10 is valid only in the situation in which the distributions of X and Y are bivariate normal.

Computer Simulation

A computer simulation was conducted to study the behavior of the standard errors of linear equating for the single-group design. Simulated scores were generated to reflect two kinds of testing situations. In the first situation, the score distribution is nearly symmetric and the simulation is referred to as the *nearly symmetric simulation.* In the second situation, the score distribution is negatively skewed and the simulation is referred to as the *nonsymmetric simulation*. The beta-binomial model (Lord & Novick, 1968, chap. 23) was selected to generate observed scores. For the nonsymmetric simulation, the beta true score distributions were assigned parameters 20.7 and 7 to simulate the score distributions similar to those of a real professional certification examination (see the example used in next section of this paper). These parameters were selected through a trial-and-error procedure. For the nearly symmetric simulation the beta true score distributions were assigned parameters 15 and 14.5. Both tests were simulated to have 75 items. The simulation was conducted using the following steps:

1) Randomly generate a beta variate using the parameters associated with the desired distribution. This beta variate, p, represents an examinees proportion-correct true score. An algorithm described by Cheng (1978) was used to generate the beta variate.

2) Given the proportion-correct true score, p, in step 1, randomly generate two binomial variates with the number of trial parameter equal to 75. These two binomial variates represent observed scores on two 75-item Forms, X and Y, respectively. A function called BNLDEV described in Numerical Recipes (Press, Flannery. Teukolsky & Vetterlin, 1990, p. 218) was used to generate the binomial variates.

3) Repeat steps 1 and 2 n times, where n represents the sample size used in the simulation. Thus, a set of n pairs of observed scores for Forms X and Y were obtained.

4) Equate Forms X and Y using the data resulting from step 3. Compute the Y equivalent of X at the selected X levels, and compute the standard errors using the following three methods: (a) with the normality assumption; (b) the delta method with numerical derivatives; and (c) the delta method with exact derivatives.

This process was replicated 500 times. The Form Y equivalents and the three standard errors at the selected X levels were averaged over the 500 replications. The "true" standard errors of equating were computed. The "true" standard error of equating for a given score on Form X was defined here as the standard deviation of Form Y equivalents of that score over the 500 replications. The simulation was conducted using sample size of 100 and 500 examinees.

The descriptive statistics for the simulated observed score distributions are listed in Table 2. These statistics are the averages computed over the 500 replications. The means for Form X are slightly higher than those for Form Y. In the nonsymmetric simulations, the score distributions are negatively skewed.

Insert Table 2 about here

The results of the simulation are summarized in Table 3 at the selected Form X score levels. The standard errors estimated by the three methods are the average values over the 500 replications. The accuracy of the three standard errors can be evaluated by comparing the bias which is the difference between the "true" standard error and the average standard errors estimated by the three methods. The standard deviations of estimations are also listed in Table 3.

All the standard errors listed in Table 3 show a general pattern of the standard errors being smaller near the mean of the score distribution than at the extremes. Also, the standard errors become smaller as the sample size gets larger.

Insert Table 3 about here

The standard errors computed by the delta method with numerical and exact derivatives are almost identical at all selected score levels in both the nearly symmetric and nonsymmetric simulations. In the nearly symmetric simulations, the simulated score distributions are very close to the normal distribution. The standard errors based on the normality assumption are very close to those calculated without normality assumption (see Table 3). In the simulation with the smaller sample size (n=100), the standard errors based on the normality assumption are closer to the "true" values than those without the normality assumption. But in the simulations with a larger sample size (n=500), the standard errors computed by the delta method are closer to the "true" values than those based on the normality assumption at most of the selected score levels.

In the nonsymmetric simulations, the differences between the standard errors based on the normality assumption and those without normality assumption are larger than those in the nearly symmetric simulations (see Table 3). The standard errors computed by the delta method are very close to the "true" values, and the standard errors based on the normality assumption are more biased. The method based on the normality assumption tends to underestimate the standard errors at lower scores and to overestimate them at higher scores for the negatively skewed score distribution.

The standard deviation of the estimated standard errors computed over the 500 replications is a measure of variability in estimating standard errors. The simulation results in Table 3 indicated that the standard errors estimated by the formula based on the normality assumption are generally less variable over the 500 replications than those based on the less restrictive assumption. An explanation given by Kolen (1985) is that the normal standard errors requires estimation of only means and variances, whereas the estimation of nonnormal standard errors requires the estimation of these parameters as well as high-order central moments and cross-product moments. Because high-order central moments and cross-product moments are very sensitive to sampling variation, the estimation of nonnormal standard errors are more variable over the replications.

A Real Data Example

Data from a 150-item multiple-choice professional licensure examination were used in this example. The 150-item test was divided by odd-even splits into two half tests. These two half tests were designated as Form X and Form Y. Each form consisted of 75 items. Data obtained from 500 examinees were used in this example. The descriptive statistics for the sample are listed in Table 4. The mean scores for both forms indicate that approximately 73 percent of the items were answered correctly on average. The score distributions of both forms are considerably skewed.

Insert Table 4 about here

Bootstrap standard errors were also computed from 1000 bootstrap replications using the procedure described by Kolen (1985). Efron (1982) presented a variety of examples in which standard errors computed from bootstrap method were more accurate for small sample situations than standard errors based on the delta method. In the present paper, the bootstrap standard errors are used for evaluating the accuracy of standard errors based on the normality assumption and standard errors without the normality assumption.

Results from linear equating and standard errors of equating computed in four different ways at the selected score levels are given in Table 5. The standard errors computed using the delta method with numerical derivatives are almost identical to those using the delta method with exact derivatives at all the selected score levels (the maximum difference is 0.001). The standard errors calculated without the normality assumption are very close to the bootstrap standard errors. In fact, the standard errors computed using the two delta methods agree with those computed by bootstrap method to two decimal places. In general, the standard errors are the smallest near the mean, and become larger farther away from the mean. The standard errors based on the normality assumption are smaller at the lower scores and larger at the higher scores than those for the other methods.

Insert Table 5 about here

Discussion and Conclusion

Three methods of estimating standard errors of linear equating for the single-group design were compared in this paper by using simulation and real test data. The results of the simulation suggest that when the score distributions are symmetric or nearly symmetric the standard errors computed based on the normality assumption and the delta method with numerical and exact derivatives are very similar to each other. When the score distributions are skewed, the results obtained from both the simulation and real data suggest that the standard errors derived without the normality assumption are less biased than those based on the normality assumption. In terms of variability in estimation, the standard errors based on the normality assumption are less variable than those derived without such an assumption.

The bootstrap method can yield accurate estimates of standard errors. However, the bootstrap method is very time consuming, because the number of resamplings must be large for the bootstrap standard errors to be accurate. The delta methods might be preferable because they yield accurate results with considerably less computation.

The standard errors computed by the delta method with numerical and exact partial derivatives were almost identical. The advantage of using the exact partial derivatives is that an equation can be provided for calculating the standard errors. But the delta method with numerical derivatives is often much simpler to compute and to program on a computer than the method with exact derivatives. This advantage is not that pronounced in computing the standard errors of linear equating under single-group design, because the linear function for this design is simple and involves only four simple first-order partial derivative. But for some more complicated equating designs, like the common-Item nonequivalent-group design (Kolen, 1985), using numerical derivatives has the capacity to make the computation dramatically less complicated than using the exact derivatives.

Another advantage of using numerical derivatives is that it is easier to develop a general computer algorithm for computing the standard errors of equating with different methods (Lord, 1975). A major task in deriving standard errors of equating with the delta method is to derive the partial derivatives with respect to each parameter involved in the equating function. For a different equating method a different set of partial derivatives need to be derived. If numerical derivatives are used there is no need to derive analytical formulas for all the partial derivatives. Thus, a general computation algorithm can estimate standard errors for different equating methods by just changing the equating function accordingly. The delta method with numerical derivatives might also prove useful for estimating standard errors under complicated designs such as chains of equatings.

References

Angoff, W. H. (1984). Scales, norms, and equivalent scores. Princeton, NJ: Educational Testing Service.

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

Cheng, R. C. H. (1978). Generating beta variates with nonintegral shape parameters. Communications of ACM, 21, 317-322.

Holland, W. P., & Thayer, D. (1990). Kernel equating and the counterbalanced design. Paper presented at the American Educational Research Association. (Boston, MA, April).

Efron, B. (1982). The jackknife, the bootstrap, and other resampling plans. Philadelphia, PA: Society for Industrial and Applied Mathematics.

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

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

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

Lord, F. M. (1975). Automated hypothesis tests and standard errors for nonstandard problems. The American Statistician, 29, 56-59.

Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. Reading, MA: Addison-Wesley.

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

Press, W. H., Flannery, B. P., Teukolsky, S. A., & Vetterling, W. T. (1990). Numerical Recipes in C. Cambridge, MA: Cambridge University Press.

Taylor, A. E., and Mann, W. R. (1983). Advanced calculus, 3rd edition New York: John Wiley & Sons, Inc.

Table 1. Sampling Variances and Covariances of Bivariate Moments

Statistic(s)	Sampling Variances and Covariances General	Normal
$var[\hat{\mu}(X)]$	$\sigma^2(X)/n$	$\sigma^2(X)/n$
$var[\hat{\sigma}^2(X)]$	$\{E[x-\mu(X)]^4 - \sigma^4(X)\}/n$	$2\sigma^4(X)/n$
$var[\hat{\mu}(Y)]$	$\sigma^2(Y)/n$	$\sigma^2(Y)/n$
$var[\hat{\sigma}^2(Y)]$	$\{E[y-\mu(Y)]^4 - \sigma^4(Y)\}/n$	$2\sigma^4(Y)/n$
$cov[\hat{\mu}(X),\hat{\mu}(Y)]$	$\sigma(X,Y)/n$	$\sigma(X,Y)/n$
$cov[\hat{\mu}(X),\hat{\sigma}^2(X)]$	$E[X-\mu(X)]^3/n$	0
$cov[\hat{\mu}(X),\hat{\sigma}^2(Y)]$	$E[x-\mu(X)][y-\mu(Y)]^2/n$	0
$cov[\hat{\sigma}^2(X),\hat{\sigma}^2(Y)]$	$\{E[x-\mu(X)]^2[y-\mu(Y)]^2 - \sigma^2(X)\sigma^2(Y)\}/n$	$2\sigma^2(X,Y)/n$
$cov[\hat{\sigma}^2(X),\hat{\mu}(Y)]$	$\{E[x-\mu(X)]^2[y-\mu(Y)]\}/n$	0
$cov[\hat{\mu}(Y),\hat{\sigma}^2(Y)]$	$E[y-\mu(Y)]^3/n$	0

Note: The terms in the body of the table were adapted from Kolen (1985), and are typically based on large sample theory. E refers to expected value. n is the sample size.

Table 2 Descriptive Statistics of Simulated Scores

Form	mean	s.d.	skewness	kurtosis
Nearly Symmetric, n=100				
X	39.46	7.95	-0.04	-0.23
Y	38.84	7.93	-0.03	-0.24
Nearly Symmetric, n=500				
X	39.41	8.00	-0.04	-0.21
Y	38.82	8.00	-0.04	-0.20
Nonsymmetric, n=100				
X	57.30	6.71	-0.42	0.04
Y	56.71	6.71	-0.41	0.02
Nonsymmetric, n=500				
X	57.30	6.78	-0.44	0.07
Y	56.72	6.80	-0.43	0.07

Table 3. Results for Two Simulations at Two Sample Sizes

		Standard Errors				Standard Deviation of Estimates		
x	$\hat{f}$ (x)	'true'	Norm.	Num.	Exact	Norm.	Num.	Exact
				Nearly Symmetric Simulation n = 100				
10	9.35	2.173	2.097	2.048	2.049	0.2069	0.2778	0.2779
20	19.36	1.519	1.442	1.427	1.427	0.1422	0.1945	0.1946
30	29.37	0.933	0.858	0.869	0.870	0.0813	0.1127	0.1127
40	39.38	0.641	0.602	0.597	0.597	0.0481	0.0494	0.0494
50	49.39	0.963	0.980	0.912	0.912	0.0954	0.1166	0.1167
60	59.40	1.556	1.590	1.479	1.480	0.1563	0.1943	0.1943
70	69.41	2.212	2.252	2.104	2.105	0.2210	0.2761	0.2763
80	79.41	2.888	2.930	2.746	2.748	0.2874	0.3605	0.3607
				n = 500				
10	9.38	0.912	0.935	0.923	0.923	0.0414	0.0529	0.0529
20	19.38	0.633	0.643	0.643	0.643	0.0285	0.0365	0.0365
30	29.39	0.386	0.383	0.392	0.392	0.0164	0.0207	0.0207
40	39.40	0.278	0.271	0.269	0.269	0.0094	0.0092	0.0092
50	49.41	0.429	0.439	0.411	0.411	0.0189	0.0241	0.0241
60	59.42	0.686	0.711	0.666	0.666	0.0311	0.0403	0.0403
70	69.43	0.967	1.006	0.947	0.948	0.0440	0.0569	0.0569
80	79.44	1.257	1.308	1.236	1.237	0.0572	0.0738	0.0739
				Nonsymmetric Simulation n = 100				
10	9.25	3.562	3.315	3.423	3.424	0.3472	0.5527	0.5530
20	19.28	2.851	2.624	2.738	2.739	0.2757	0.4434	0.4436
30	29.32	2.147	1.941	2.060	2.061	0.2048	0.3340	0.3342
40	39.35	1.459	1.278	1.397	1.398	0.1352	0.2245	0.2246
50	49.38	0.829	0.694	0.789	0.790	0.0687	0.1130	0.1131
60	59.42	0.534	0.558	0.504	0.504	0.0522	0.0436	0.0436
70	69.45	0.960	1.061	0.921	0.921	0.1084	0.1209	0.1209
80	79.49	1.611	1.710	1.551	1.552	0.1747	0.2225	0.2226
				n = 500				
10	9.24	1.691	1.473	1.556	1.557	0.0667	0.1153	0.1154
20	19.28	1.356	1.167	1.245	1.246	0.0530	0.0928	0.0928
30	29.32	1.024	0.863	0.937	0.938	0.0393	0.0702	0.0702
40	39.36	0.699	0.569	0.636	0.636	0.0259	0.0475	0.0475
50	49.39	0.394	0.310	0.359	0.359	0.0130	0.0240	0.0240
60	59.43	0.227	0.250	0.226	0.226	0.0099	0.0091	0.0091
70	69.47	0.420	0.472	0.414	0.415	0.0203	0.0235	0.0235
80	79.51	0.728	0.760	0.700	0.701	0.0329	0.0440	0.0440

Table 4 Descriptive Statistics for
a Professional Certification Examination

Form	mean	S. D.	skewness	kurtosis
X	55.694	5.719	-0.657	0.493
Y	55.594	5.687	-0.527	-0.044

Table 5 Standard Errors of Linear Equatings for
a Professional Certification Examination

x	$\hat{f}(x)$	Boot.	Norm.	Numer.	Exact.
10	10.158	1.578	1.540	1.569	1.570
15	15.131	1.415	1.374	1.406	1.407
20	20.104	1.251	1.209	1.244	1.245
25	25.076	1.089	1.044	1.082	1.083
30	30.049	0.927	0.881	0.921	0.921
35	35.022	0.766	0.720	0.761	0.762
40	39.994	0.608	0.562	0.604	0.605
45	44.967	0.455	0.411	0.452	0.453
50	49.940	0.315	0.280	0.313	0.313
55	54.913	0.214	0.209	0.214	0.214
60	59.885	0.217	0.255	0.218	0.218
65	64.858	0.320	0.377	0.321	0.321
70	69.831	0.461	0.525	0.461	0.461
75	74.803	0.615	0.682	0.614	0.614