A COMPARISON OF ROBUST ESTIMATORS IN SIMPLE LINEAR REGRESSION

E. Jacquelin Dietz

North Carolina State University
Raleigh, North Carolina

Key Words and Phrases: deficiency; intercept; point estimation; slope; Theil estimator.

ABSTRACT

The mean squared errors of various estimators of slope, intercept, and mean response in the simple linear regression problem are compared in a simulation study. A weighted median estimator of slope proposed by Sievers (1978) and Scholz (1978) and two intercept estimators based upon it are found to perform well for most error distributions studied. Theil's (1950) estimator of slope and two intercept estimators based on it are preferable for certain heavily contaminated error distributions.

1. INTRODUCTION

The topic of simple linear regression is either omitted entirely or covered incompletely in most popular nonparametric statistics texts. Many such books include a distribution-free test or confidence interval for the slope of the regression line, but omit any discussion of the intercept or the mean response at a given x.

In Dietz (1986), I compared several previously proposed estimators of slope, intercept, and mean response with regard to unbiasedness, efficiency, and breakdown properties. In this

paper, I investigate the mean squared erors of those same estimators in a simulation study.

2. THE ESTIMATORS

I assume the model $Y_i = \alpha + \beta x_i + e_i$, $i = 1,2,...,n$, where α and β are unknown parameters, $x_1 \leq x_2 \leq \ldots \leq x_n$ are known constants (not all equal), and the e_i's are independent and identically distributed continuous random variables with mean zero.

2.1 Slope Estimators

The estimators of β considered in this paper can all be viewed as functions of the N sample slopes

$$S_{ij} = (Y_j - Y_i)/(x_j - x_i) \text{ , } i < j \text{ , } x_i \neq x_j \text{ .}$$

If the x_i's are all distinct, $N = \binom{n}{2}$. The following estimators of β are considered:

1. $\hat{\beta}_{LS} = \sum_{i<j} w_{ij}S_{ij} / \sum_{i<j} w_{ij}$, where $w_{ij} = (x_j - x_i)^2$.
2. $\hat{\beta}_A = \sum_{i<j} S_{ij}/N$ (Randles and Wolfe, 1979, Problem 3.1.6).
3. $\hat{\beta}_M$ = median of S_{ij}'s (Theil, 1950).
4. $\hat{\beta}_{W1}$ = weighted median of S_{ij}'s, where $w_{ij} = j - i$.
5. $\hat{\beta}_{W2}$ = weighted median of S_{ij}'s, where $w_{ij} = x_j - x_i$.

A weighted median estimator (Sievers, 1978; Scholz, 1978) equals the median of the probability distribution obtained by assigning probability $w_{ij} / \sum_{i<j} w_{ij}$ to S_{ij}.

2.2 Intercept Estimators

The estimators of α considered in this paper can be divided into two groups, those that are functions of the N sample intercepts

$$A_{ij} = (x_jY_i - x_iY_j)/(x_j - x_i) \text{ , } i < j, \; x_i \neq x_j \text{ ,}$$

and those based on residuals associated with a particular slope estimate. The following estimators of α are considered:

1. $\hat{\alpha}_{LS} = \sum_{i<j} w_{ij}A_{ij} / \sum_{i<j} w_{ij}$, where $w_{ij} = (x_j - x_i)^2$.

2. $\hat{\alpha}_A = \sum_{i<j} A_{ij}/N$ (Randles and Wolfe, 1979, Problem 3.1.6).

3. $\hat{\alpha}_M$ = median of A_{ij}'s.

4. $\hat{\alpha}_{1,M}$ = median of $Y_i - \hat{\beta}_M x_i$, i = 1,2,...,n.

5. $\hat{\alpha}_{1,W1}$ = median of $Y_i - \hat{\beta}_{W1} x_i$, i = 1,2,...,n.

6. $\hat{\alpha}_{1,W2}$ = median of $Y_i - \hat{\beta}_{W2} x_i$, i = 1,2,...,n.

7. $\hat{\alpha}_{2,M}$ = median of pairwise averages of $Y_i - \hat{\beta}_M x_i$.

8. $\hat{\alpha}_{2,W1}$ = median of pairwise averages of $Y_i - \hat{\beta}_{W1} x_i$.

9. $\hat{\alpha}_{2,W2}$ = median of pairwise averages of $Y_i - \hat{\beta}_{W2} x_i$.

10. $\hat{\alpha}_C = Y_{.50} - \hat{\beta}_M x_{.50}$, where $Y_{.50}$ and $x_{.50}$ are the sample medians of the Y's and x's (Conover, 1980, p.267).

Bhattacharyya (1968) considers estimators 7 and 8; Hettmansperger (1984) estimators 6 and 9.

2.3 Estimators of Mean Response

The mean response at a given x value, $E(Y) = \alpha + \beta x$, is often a more interesting parameter than the intercept α. (Of course, if $x = 0$, then $E(Y) = \alpha$.) The estimators of E(Y) considered here are of the form $\hat{\alpha} + \hat{\beta} x$, where each $\hat{\alpha}$ is associated with the $\hat{\beta}$ with the same subscript, except that $\hat{\alpha}_C$ is associated with $\hat{\beta}_M$.

3. UNBIASEDNESS AND SYMMETRY

Certain unbiasedness and symmetry properties possessed by the estimators are important in the simulation study. Since $E(S_{ij}) = \beta$, it follows that $\hat{\beta}_{LS}$ and $\hat{\beta}_A$ are unbiased estimators of β. Sen (1968, Section 5) shows that the distribution of $\hat{\beta}_M$ is symmetric about β, and Theorem 5 of Sievers (1978) implies that $\hat{\beta}_{W1}$ and $\hat{\beta}_{W2}$ are asymptotically unbiased, under certain conditions on the x's.

The assumption that the e_i's have mean zero implies that $E(A_{ij}) = \alpha$ and thus that $\hat{\alpha}_{LS}$ and $\hat{\alpha}_A$ are unbiased estimators of α. The alternative assumption that the e_i's have median zero leads to a reparameterization more appropriate for certain of the other intercept estimators. This issue is discussed further in Sections 4 and 5 in connection with the simulation study. For

symmetric error distributions, the two parameterizations coincide and all of the intercept estimators estimate the same parameter.

For symmetric error distributions all of the estimators of α and β in Section 2 except $\hat{\alpha}_C$ have distributions that are symmetric about the corresponding parameter. The distribution of $\hat{\alpha}_C$ is symmetric about α if $\beta = 0$; however, it does not seem possible to demonstrate any unbiasedness property for $\hat{\alpha}_C$ for general β.

4. SIMULATION STUDY: METHODS

The bias and mean squared error (MSE) of the various estimators of slope, intercept, and mean response were estimated and compared in a simulation study. All computing was done in Fortran on the IBM 3081 at the Triangle Universities Computation Center. Medians of pairwise averages were computed using the algorithm of Monahan (1984), which in turn uses the uniform random number generator of Schrage (1979). The IMSL (1984) routines GGNML, GGUBS, and GGCHS were used to generate normal, uniform, and chi square random numbers, respectively. The IMSL (1984) routine MSENO was used to calculate expected order statistics from the normal distribution.

Five hundred samples were generated for each combination of n = 20 and 40, three x designs, and nine error distributions. Preliminary work indicated that this number of samples was sufficient to demonstrate highly significant differences among estimators.

All estimators of α and β considered here except $\hat{\alpha}_C$ have MSE's that do not depend on the values of α and β. (The MSE of $\hat{\alpha}_C$ does not depend on the value of α, but does depend on β.) Thus, I set $\alpha = \beta = 0$.

The value of each estimator of α and β defined in Section 2 was computed for each sample. For n = 20 and the eight symmetric error distributions, the value of each estimator of E(Y) was also computed, for each of two x values. (Note that E(Y) is not the natural parameter to estimate for an asymmetric error

distribution.) The five hundred values of an estimator thus obtained were used to estimate the bias and MSE of that estimator for that n, x design, and error distribution. The sample variances of the estimates and of the squared estimates over the 500 samples were used to estimate the variances of the bias and MSE estimates, respectively. An estimator will be considered to show significant bias if the absolute value of its estimated bias exceeds two times the estimated standard error of the bias.

The first two x designs consisted of the expected order statistics from samples of size n from the uniform and normal distributions, respectively. The third x design, referred to as the double exponential design, was generated by calculating $\pm a_r$, $r = 1,2,\ldots,n/2$, where

$$a_r = \sum_{t=1}^{r} (n/2 - t + 1)^{-1} ;$$

a_r is the expected value of the rth order statistic from a sample of size n/2 from the exponential distribution. Each set of x's was then standardized so that $\sum_{i=1}^{n} x_i = 0$ and $\sum_{i=1}^{n} x_i^2 = 1$. For n = 20, $E(Y) = \alpha + \beta x$ was estimated for each of $x = 20^{-.5} = .2236$ and $x = 2(20^{-.5}) = .4472$, the x values one and two standard deviations from the center of the x's. These two x values and the values of x for each design for n = 20 are shown in Figure 1.

Nine error distributions were considered -- the standard normal, six contaminated normal distributions, the heavy-tailed t distribution with three degrees of freedom, and the asymmetric lognormal distribution. A single sample of n standard normal variates was used to generate samples from all nine error distributions. Specifically, for each of the 500 samples for a particular n value and x design, triples (Z_i, U_i, V_i), $i = 1,2,\ldots,n$, were generated, where Z_i is standard normal, U_i is Uniform (0,1), and V_i is chi-square with three degrees of freedom. Then the t_3 and lognormal variates, standardized to have mean zero and variance one, were obtained by taking $Z_i V_i^{-1/2}$ and $(e^{Zi} - e^{1/2})(e^2 - e)^{-1/2}$, respectively, for $i = 1,2,\ldots,n$.

Contaminated normal (CN(k,α)) samples, k = 3,10 and α = .05, .10, .25, were generated by multiplying each Z_i by k with probability α, that is, if U_i fell in an appropriate interval:

k	α	Interval
3	.05	(.10,.15]
	.10	(.15,.25]
	.25	(.25,.50]
10	.05	(.50,.55]
	.10	(.55,.65]
	.25	(.65,.90]

As mentioned in Section 3, for asymmetric error distributions the intercept parameter can be defined in different ways. The lognormal variates defined above have mean zero, but median $(1 - e^{1/2})(e^2 - e)^{-1/2} = -.3001675$. For this error distribution only, the bias and MSE of the intercept estimators were estimated in two ways -- with $\alpha = 0$ and $\alpha = -.3001675$ as the target parameter values. In addition, for the case of lognormal errors, normal x design, and n = 20, I calculated the median of the 500 values of each intercept estimator, as well as the mean.

5. SIMULATION STUDY: RESULTS

5.1 Slope Estimators

The estimated MSE's of the slope estimators and the associated standard error estimates are displayed in Table I for n=20. Results for n=40 are similar and are not shown. Simulation results for CN(k, .10) errors are intermediate to those for CN(k, .05) and CN(k, .25) errors, k = 3 and 10, and are omitted from the table. Because of the unbiasedness results in Section 3, the estimated bias of the slope estimators is not shown. As expected, the simulation results showed little evidence of significant bias in the slope estimators. Note that for the uniform x design, $\hat{\beta}_{W1}$ and $\hat{\beta}_{W2}$ are identical.

The simulation results for normal and double exponential x's are shown graphically in Figures 2 and 3, respectively. These figures show what Andrews et al. (1972) refer to as relative

TABLE I

Mean Squared Error of Slope Estimates, n=20
(Estimated Standard Errors in Parentheses)

Error Distribution	$\hat{\beta}$	x Design: Uniform	Normal	Double Exponential
N(0,1)	A	1.0(.06)	1.1(.06)	9.1(.57)
	LS	1.0(.06)	1.0(.06)	1.0(.06)
	M	1.0(.07)	1.1(.07)	1.3(.08)
	W1	1.0(.07)	1.1(.06)	1.3(.08)
	W2	1.0(.07)	1.1(.06)	1.1(.07)
t_3	A	.9(.08)	1.2(.11)	7.6(.60)
	LS	.9(.07)	1.1(.10)	1.0(.12)
	M	.6(.04)	.6(.04)	.8(.06)
	W1	.6(.04)	.6(.04)	.8(.06)
	W2	.6(.04)	.6(.04)	.7(.05)
CN(3,.05)	A	1.6(.11)	1.5(.10)	12.8(.93)
	LS	1.5(.10)	1.4(.10)	1.4(.12)
	M	1.2(.08)	1.3(.08)	1.5(.10)
	W1	1.2(.08)	1.2(.07)	1.5(.10)
	W2	1.2(.08)	1.2(.08)	1.3(.09)
CN(3,.25)	A	3.0(.20)	3.2(.20)	26.8(1.95)
	LS	2.8(.17)	2.9(.19)	2.8(.22)
	M	1.8(.12)	2.1(.14)	2.6(.18)
	W1	1.8(.12)	2.1(.13)	2.5(.17)
	W2	1.8(.12)	2.0(.14)	2.3(.19)
CN(10,.05)	A	6.0(.61)	6.7(.59)	59.0(8.47)
	LS	5.8(.54)	5.8(.59)	6.5(.84)
	M	1.2(.08)	1.4(.09)	1.6(.11)
	W1	1.3(.09)	1.3(.08)	1.6(.11)
	W2	1.3(.09)	1.3(.08)	1.5(.10)
CN(10,.25)	A	25.0(2.00)	27.9(1.99)	235.4(20.81)
	LS	23.7(1.78)	25.1(1.95)	24.3(1.78)
	M	3.5(.30)	4.1(.39)	5.1(.51)
	W1	3.9(.53)	4.4(.45)	.4(.57)
	W2	3.9(.53)	4.9(.58)	5.9(.71)
Lognormal	A	1.0(.16)	1.1(.11)	10.4(1.40)
	LS	1.0(.13)	.9(.09)	1.4(.52)
	M	.2(.02)	.2(.02)	.2(.02)
	W1	.2(.02)	.2(.02)	.3(.02)
	W2	.2(.02)	.2(.02)	.3(.02)

deficiency:

$$\text{relative deficiency} = 1 - \frac{\text{minimum variance}}{\text{variance}} .$$

The minimum is taken over the estimators considered here. Deficiencies are all between zero and one, with one estimator assured a deficiency of zero. The graph of deficiencies for uniform x's is very similar to that for normal x's and is not shown. Note that since these estimators are unbiased, their variances are very similar to their MSE's, and thus the estimators are ordered very similarly by MSE and deficiency.

The unweighted average estimator $\hat{\beta}_A$ is clearly unacceptable in terms of MSE or deficiency, especially for the double exponential x's and heavily contaminated error distributions. (In fairness to Randles and Wolfe (1979), they do not recommend this estimator or the intercept estimator $\hat{\alpha}_A$. These estimators arise merely as answers to a textbook exercise.) Not surprisingly, the least-squares estimator, $\hat{\beta}_{LS}$, has the smallest MSE of any slope estimator for normally distributed errors. In almost every other case, however, the MSE of $\hat{\beta}_{LS}$ is exceeded only by that of $\hat{\beta}_A$.

The estimator $\hat{\beta}_{W2}$ performs well until the errors are heavily contaminated or asymmetric, at which point $\hat{\beta}_M$ is preferable. The MSE of $\hat{\beta}_{W1}$ is usually between those of $\hat{\beta}_M$ and $\hat{\beta}_{W2}$.

5.2 Estimators of Mean Response: Symmetric Errors

Estimation of α is equivalent to estimation of $E(Y) = \alpha + \beta x$ for $x = 0$, and is discussed here in that context. The estimated MSE's of the estimators of mean response, along with their estimated standard errors, are shown in Tables II and III for $x = 0$ and $x = .4472$, respectively, and $n = 20$. Results for $x = .2236$ are intermediate to those for the other x's and are not shown. Also, results for $n = 40$ for $x = 0$ are similar to those for $n = 20$ and are not shown. The MSE's for CN(k, .10) errors are intermediate to those for CN(k, .05) and CN(k, .25) errors, $k = 3$ and 10, and are omitted from the tables. Results for the asymmetric lognormal errors are shown and discussed separately. Because of the unbiasedness results in Section 3, the estimated

TABLE II

Mean Squared Error of Intercept Estimates,
n=20, Symmetric Error Distributions
(Estimated Standard Errors in Parentheses)

Error Distribution	$\hat{\alpha}$	x Design: Uniform	Normal	Double Exponential
N(0,1)	A	.11(.007)	.10(.006)	.12(.007)
	LS	.05(.003)	.06(.003)	.05(.003)
	M	.10(.007)	.11(.007)	.08(.005)
	1,M	.07(.005)	.09(.005)	.07(.004)
	C	.07(.004)	.08(.005)	.07(.004)
	2,M	.05(.003)	.06(.003)	.05(.003)
t_3	A	.09(.007)	.11(.024)	.12(.013)
	LS	.04(.003)	.06(.007)	.05(.003)
	M	.04(.003)	.05(.003)	.04(.003)
	1,M	.03(.002)	.04(.002)	.03(.002)
	C	.03(.002)	.03(.002)	.03(.002)
	2,M	.03(.002)	.03(.002)	.03(.002)
CN(3,.05)	A	.15(.010)	.15(.011)	.17(.013)
	LS	.07(.005)	.08(.006)	.07(.005)
	M	.10(.007)	.12(.007)	.09(.005)
	1,M	.08(.005)	.09(.005)	.08(.005)
	C	.07(.005)	.09(.005)	.08(.005)
	2,M	.06(.004)	.07(.004)	.06(.004)
CN(3,.25)	A	.30(.022)	.29(.019)	.35(.025)
	LS	.13(.010)	.15(.009)	.15(.010)
	M	.16(.012)	.17(.011)	.14(.009)
	1,M	.10(.007)	.13(.007)	.12(.008)
	C	.10(.007)	.12(.007)	.11(.007)
	2,M	.09(.007)	.11(.006)	.11(.007)
CN(10,.05)	A	.66(.058)	.50(.054)	.75(.108)
	LS	.30(.024)	.30(.022)	.30(.026)
	M	.11(.008)	.12(.008)	.10(.006)
	1,M	.08(.005)	.09(.006)	.08(.005)
	C	.07(.005)	.09(.005)	.08(.005)
	2,M	.06(.004)	.07(.004)	.07(.004)
CN(10,.25)	A	2.62(.193)	2.35(.166)	3.14(.226)
	LS	1.20(.086)	1.32(.086)	1.30(.082)
	M	.23(.019)	.26(.020)	.21(.016)
	1,M	.15(.011)	.17(.012)	.14(.009)
	C	.11(.008)	.14(.009)	.13(.008)
	2,M	.18(.017)	.21(.016)	.18(.012)

TABLE III

Mean Squared Error of Estimates of Mean Response
n = 20, x = .4472, Symmetric Error Distributions
(Estimated Standard Errors in Parentheses)

Error Distribution	$\hat{\alpha}$	x Design: Uniform	Normal	Double Exponential
N(0,1)	A	.31(.02)	.33(.02)	1.96(.12)
	LS	.24(.01)	.26(.02)	.26(.02)
	M	.32(.02)	.36(.02)	.36(.02)
	1,M	.29(.02)	.35(.02)	.35(.02)
	C	.27(.02)	.34(.02)	.34(.02)
	2,M	.27(.02)	.30(.02)	.33(.02)
t_3	A	.27(.02)	.33(.03)	1.56(.12)
	LS	.23(.02)	.27(.03)	.27(.03)
	M	.17(.01)	.19(.01)	.21(.02)
	1,M	.16(.01)	.18(.01)	.20(.02)
	C	.14(.01)	.17(.01)	.19(.01)
	2,M	.15(.01)	.16(.01)	.20(.02)
CN(3,.05)	A	.47(.03)	.49(.03)	2.82(.22)
	LS	.37(.03)	.38(.03)	.37(.03)
	M	.35(.02)	.40(.03)	.41(.03)
	1,M	.32(.02)	.38(.03)	.39(.03)
	C	.31(.02)	.37(.02)	.39(.03)
	2,M	.30(.02)	.34(.02)	.37(.03)
CN(3,.25)	A	.89(.06)	.96(.06)	5.73(.45)
	LS	.71(.05)	.76(.05)	.73(.05)
	M	.57(.04)	.64(.04)	.70(.05)
	1,M	.49(.03)	.59(.04)	.68(.05)
	C	.46(.03)	.58(.04)	.65(.04)
	2,M	.48(.03)	.57(.04)	.66(.05)
CN(10,.05)	A	1.72(.16)	1.80(.15)	12.89(1.92)
	LS	1.54(.15)	1.51(.15)	1.52(.19)
	M	.39(.03)	.43(.03)	.45(.03)
	1,M	.36(.03)	.41(.03)	.43(.03)
	C	.32(.02)	.40(.03)	.43(.03)
	2,M	.33(.02)	.38(.03)	.42(.03)
CN(10,.25)	A	8.06(.59)	9.01(.66)	51.02(4.35)
	LS	6.30(.49)	6.59(.55)	6.43(.48)
	M	.96(.07)	1.20(.13)	1.28(.11)
	1,M	.83(.07)	1.04(.10)	1.17(.10)
	C	.79(.06)	1.01(.09)	1.20(.10)
	2,M	.90(.07)	1.12(.11)	1.21(.11)

bias of the estimators is not shown. For n = 20, no estimator of mean response showed significant bias for any symmetric error distribution.

As x increases from 0 to .4472, so that estimation of E(Y) increasingly involves extrapolation, the MSE's of the estimators and their standard errors increase.

Certain conclusions hold for all combinations of x design, sample size, and x value included in the study. The unweighted average estimator $\hat{\alpha}_A + \hat{\beta}_A x$ can be eliminated from consideration on the basis of its large MSE. The median estimator $\hat{\alpha}_M + \hat{\beta}_M x$ is also noncompetitive; its MSE is usually one of the two or three largest. As expected, $\hat{\alpha}_{LS} + \hat{\beta}_{LS}x$ is the best estimator for N(0,1) errors. However, for heavily contaminated normal errors and t_3 errors, the MSE of $\hat{\alpha}_{LS} + \hat{\beta}_{LS}x$ is usually exceeded only by that of $\hat{\alpha}_A + \hat{\beta}_A x$.

For all combinations of sample size, x design, x value, and error distribution studied except the CN(10, .25) errors and the double exponential x design with x = .4472, the following results were found:

1. The estimators based on the $\hat{\alpha}_1$'s ($\hat{\alpha}_{1,W1}$, $\hat{\alpha}_{1,W2}$, $\hat{\alpha}_{1,M}$) are very similar in MSE; the same is true of the estimators based on the $\hat{\alpha}_2$'s ($\hat{\alpha}_{2,W1}$, $\hat{\alpha}_{2,W2}$, $\hat{\alpha}_{2,M}$). In fact, for x = 0, the MSE's of the three estimators within one of these groups rarely differ by more than .001; for x = .2236, by more than .01; and for x = .4472, by more than .02. Because of these similarities within the $\hat{\alpha}_1$ and $\hat{\alpha}_2$ groups, results for estimators based on $\hat{\alpha}_{1,W1}$, $\hat{\alpha}_{1,W2}$, $\hat{\alpha}_{2,W1}$, and $\hat{\alpha}_{2,W2}$ are omitted from Tables II and III.

2. The estimators based on the $\hat{\alpha}_2$'s have smaller MSE's than any other estimators for all symmetric error distributions studied except the N(0,1), for which they are second only to the least-squares estimator.

3. The MSE of the Conover estimator $\hat{\alpha}_C + \hat{\beta}_M x$ is usually between those of the $\hat{\alpha}_1$ estimators and the $\hat{\alpha}_2$ estimators.

For the heavily contaminated CN(10, .25) errors, the

following results were found:

1. The Conover estimator has the smallest MSE of all estimators considered.

2. Each $\hat{\alpha}_1$ estimator has smaller MSE than the corresponding $\hat{\alpha}_2$ estimator (that is, the $\hat{\alpha}_2$ estimator based on the same $\hat{\beta}$).

3. Within the $\hat{\alpha}_1$ and $\hat{\alpha}_2$ groups, the estimators based on $\hat{\beta}_M$ have the smallest MSE's and those based on $\hat{\beta}_{W2}$ the largest.

These patterns are shown graphically in Figure 4, which shows deficiencies of the estimators of mean response for uniform x's, n = 20, x = 0. Results for estimators based on $\hat{\alpha}_{1,W1} = \hat{\alpha}_{1,W2}$ and $\hat{\alpha}_{2,W1} = \hat{\alpha}_{2,W2}$ are omitted from the figure. Deficiencies were calculated by taking the minimum variance over all intercept estimators, including those omitted from Figure 4.

For the double exponential x design with x = .4472 (excluding the CN(10, .25) errors already discussed), the choice of $\hat{\beta}$ is more important than the choice between the $\hat{\alpha}_1$ and $\hat{\alpha}_2$ groups. The estimators based on $\hat{\beta}_{W2}$ have the smallest MSE's; the MSE's of the estimators based on $\hat{\beta}_{W1}$ and $\hat{\beta}_M$ (including $\hat{\alpha}_C + \hat{\beta}_M$ x but not $\hat{\alpha}_M + \hat{\beta}_M$ x) are almost always within .02 of each other. Deficiencies of estimators based on $\hat{\alpha}_{1,W2}$, $\hat{\alpha}_{1,M}$, $\hat{\alpha}_{2,W1}$, and $\hat{\alpha}_{2,W2}$ are shown graphically in Figure 5. Deficiencies of estimators based on $\hat{\alpha}_{1,W1}$, $\hat{\alpha}_{2,M}$, and $\hat{\alpha}_C$ are between those of $\hat{\alpha}_{1,M}$ and $\hat{\alpha}_{2,W1}$ for all symmetric error distributions. Note that x = .4472 represents greater extrapolation for the double exponential x design than for the other x designs (See Figure 1.).

5.3 Intercept Estimators: Lognormal Errors

Simulation results for the lognormal error distribution are shown in Table IV and, for uniform x's, in Figure 4. As seen for symmetric error distributions, within the $\hat{\alpha}_1$ group and $\hat{\alpha}_2$ group of estimators, there are no significant differences in MSE. In fact, for a particular n and x design, the three MSE's within a group are within .001 of each other. Therefore, only $\hat{\alpha}_{1,M}$ and $\hat{\alpha}_{2,M}$ are shown in the table and figure.

TABLE IV

Bias and Mean Squared Error (MSE) of Intercept Estimates, Lognormal Errors (Estimated Standard Errors in Parentheses)

		n = 20			n = 40	
$\hat{\alpha}$	x	Bias	MSE	Median[b]	Bias	MSE
		(a) Uniform x Design				
A	0	-.004(.015)	.105(.011)		-.008(.010)	.051(.004)
LS	0	-.005(.010)	.045(.005)		-.004(.007)	.024(.002)
2,M	-.3[a]	.113(.006)	.033(.002)		.113(.005)	.024(.002)
M	-.3	.058(.007)	.030(.003)		.053(.005)	.015(.001)
1,M	-.3	.022(.006)	.018(.001)		.017(.004)	.010(.001)
C	-.3	.008(.005)	.015(.001)		.014(.004)	.010(.001)
		(b) Normal x Design				
A	0	.023(.014)	.103(.009)	.033	-.006(.009)	.044(.004)
LS	0	.018(.011)	.056(.004)	-.031	.003(.007)	.025(.002)
2,M	-.3	.133(.007)	.043(.003)	.107	.109(.005)	.022(.001)
M	-.3	.075(.008)	.037(.004)	.048	.043(.005)	.013(.001)
1,M	-.3	.034(.007)	.023(.002)	.012	.012(.004)	.009(.001)
C	-.3	.023(.006)	.020(.001)	.010	.008(.004)	.008(.001)
		(c) Double Exponential x Design				
A	0	.023(.016)	.128(.014)		-.013(.011)	.056(.008)
LS	0	.017(.011)	.057(.006)		-.010(.007)	.024(.002)
2,M	-.3	.128(.007)	.040(.003)		.105(.005)	.022(.001)
M	-.3	.077(.007)	.032(.003)		.045(.005)	.013(.001)
1,M	-.3	.035(.006)	.021(.002)		.008(.004)	.009(.001)
C	-.3	.025(.006)	.018(.001)		-.000(.004)	.009(.001)

[a] Median of lognormal variates, actually -.3001675.

[b] Medians of 500 values of $\hat{\alpha}_{1,W1}$, $\hat{\alpha}_{1,W2}$, $\hat{\alpha}_{2,W1}$, and $\hat{\alpha}_{2,W2}$ were .014, .018, .106, and .107, respectively.

The bias and MSE of $\hat{\alpha}_{LS}$ and $\hat{\alpha}_A$ are computed using $\alpha = 0$ as the target parameter value. Neither estimator shows evidence of bias; $\hat{\alpha}_{LS}$ is preferable to $\hat{\alpha}_A$ in terms of MSE.

For all other estimators, the bias and MSE are computed using the median of the lognormal variates as the target parameter value. This parameterization is appropriate for the three $\hat{\alpha}_1$ estimators, $\hat{\alpha}_C$, and $\hat{\alpha}_M$. All show significant bias; however, for n = 20 and the normal x design, the median of the 500 values of each of these estimators (minus -.3001675) is closer to zero than is the mean of the 500 values, suggesting that these estimators are median unbiased. Of these estimators, $\hat{\alpha}_M$ has the largest MSE.

The three $\hat{\alpha}_2$ estimators evidently estimate the pseudomedian of the lognormal distribution. The pseudomedian of a distribution F is the median of the distribution of $(Z_1 + Z_2)/2$, where Z_1 and Z_2 are independent, each with distribution F (Hollander and Wolfe, 1973, p. 458). To estimate the pseudomedian of the lognormal distribution, I generated ten independent random samples of 10,000 lognormal variates each and found the median of pairwise averages for each sample. The median of these ten sample medians (minus -.3001675) was .10. This agrees moderately well with the estimated bias and median of $\hat{\alpha}_{2,M}$ shown in Table IV.

It is interesting to note that $\hat{\alpha}_{LS}$, although unbiased, has larger MSE for estimating the mean than do the other estimators (excluding $\hat{\alpha}_A$) for estimating the median. This is true despite the bias of the other estimators and the fact that $\hat{\alpha}_{2,M}$ is not even estimating the "right" parameter.

6. SUMMARY

The main findings of the simulation study are:

1. The estimators $\hat{\beta}_A$, $\hat{\alpha}_A$, and $\hat{\alpha}_M$ can be eliminated from consideration because of their large MSE's.

2. The least-squares estimators, $\hat{\beta}_{LS}$ and $\hat{\alpha}_{LS}$, have smaller MSE's than competitors for normal errors, but have very large MSE's for contaminated or asymmetric error distributions.

3. The estimator $\hat{\beta}_{W2}$ performs well until the errors are

heavily contaminated or asymmetric, at which point $\hat{\beta}_M$ is preferable.

4. For most situations considered, the estimators of mean response based on $\hat{\alpha}_1$'s do not differ among themselves in MSE; neither do those based on $\hat{\alpha}_2$'s. The $\hat{\alpha}_2$ estimators perform well until the errors are heavily contaminated or asymmetric, at which point the $\hat{\alpha}_1$ estimators and $\hat{\alpha}_C$ are preferable.

5. When estimation of the mean response requires extrapolation, the choice of $\hat{\beta}$ estimator can be more important than the choice between the $\hat{\alpha}_1$ and $\hat{\alpha}_2$ groups of estimators. The estimators based on $\hat{\beta}_{W2}$ perform well until the errors are heavily contaminated. Then the estimators based on $\hat{\beta}_M$ are better.

BIBLIOGRAPHY

Andrews, D.F., Bickel, P.J., Hampel, F.R., Huber, P.J., Rogers, W.H., and Tukey, J.W. (1972). *Robust Estimates of Location.* Princeton, New Jersey: Princeton University Press.

Bhattacharyya, G.K. (1968). Robust estimates of linear trend in multivariate time series. *Ann. Inst. Statist. Math.* 20, 299-310.

Conover, W.J. (1980). *Practical Nonparametric Statistics* (2nd edition). New York: Wiley.

Dietz, E.J. (1986). On estimating a slope and intercept in a nonparametric statistics course. Institute of Statistics Mimeograph Series No. 1689R, North Carolina State University.

Hettmansperger, T.P. (1984). *Statistical Inference Based on Ranks.* New York: Wiley.

Hollander, M. and Wolfe, D.A. (1973). *Nonparametric Statistical Methods.* New York: Wiley.

International Mathematical and Statistical Libraries, Inc. (1984). *The IMSL Library* (9th edition). Houston, Texas: Author.

Monahan, J.F. (1984). Algorithm 616. Fast computation of the Hodges-Lehmann location estimator. *ACM Trans. Math. Software* 10, 265-270.

Randles, R.H. and Wolfe, D.A. (1979). *Introduction to the Theory of Nonparametric Statistics.* New York: Wiley.

Scholz, F.W. (1978). Weighted median regression estimates. *Ann. Statist.* 6, 603-609.

Schrage, L. (1979). A more portable Fortran random number generator. *ACM Trans. Math. Software* 5, 132-138.

Sen, P.K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journ. Amer. Statist. Assoc.* 63, 1379-1389.

Sievers, G.L. (1978). Weighted rank statistics for simple linear regression. *Journ. Amer. Statist. Assoc.* 73, 628-631.

Theil, H. (1950). A rank-invariant method of linear and polynomial regression analysis. *Indagationes Mathematicae* 12, 85-91.

Figure captions

FIG. 1 Positive values of x for n = 20. The negative of each number is included also. x designs: 1 - uniform; 2 - normal; 3 - double exponential. Vertical dotted lines indicate x values for which E(Y) is estimated.

FIG. 2 Deficiency of slope estimators for normal x design, n = 20. Error distributions: 1-N(0,1); 2-t_3; 3-CN(3,.05); 4-CN(3,.10); 5-CN(3,.25); 6-CN(10,.05); 7-CN(10,.10); 8-CN(10,.25); 9 - lognormal.

FIG. 3 Deficiency of slope estimators for double exponential x design, n = 20. Error distributions: 1-N(0,1); 2-t_3; 3-CN(3,.05); 4-CN(3,.10); 5-CN(3,.25); 6-CN(10,.05); 7-CN(10,.10); 8-CN(10,.25); 9 - lognormal.

FIG. 4 Deficiency of intercept estimators for uniform x design, n = 20. Error distributions: 1-N(0,1); 2-t_3; 3-CN(3,.05); 4-CN(3,.10); 5-CN(3,.25); 6-CN(10,.05); 7-CN(10,.10); 8-CN(10,.25); 9 - lognormal.

FIG. 5 Deficiency of estimators of mean response for double exponential x design, n = 20, x = .4472. Error distributions: 1-N(0,1); 2-t_3; 3-CN(3,.05); 4-CN(3,.10); 5-CN(3,.25); 6-CN(10,.05); 7-CN(10,.10) 8-CN(10,.25). Notice the restricted range of the deficiencies in this figure.

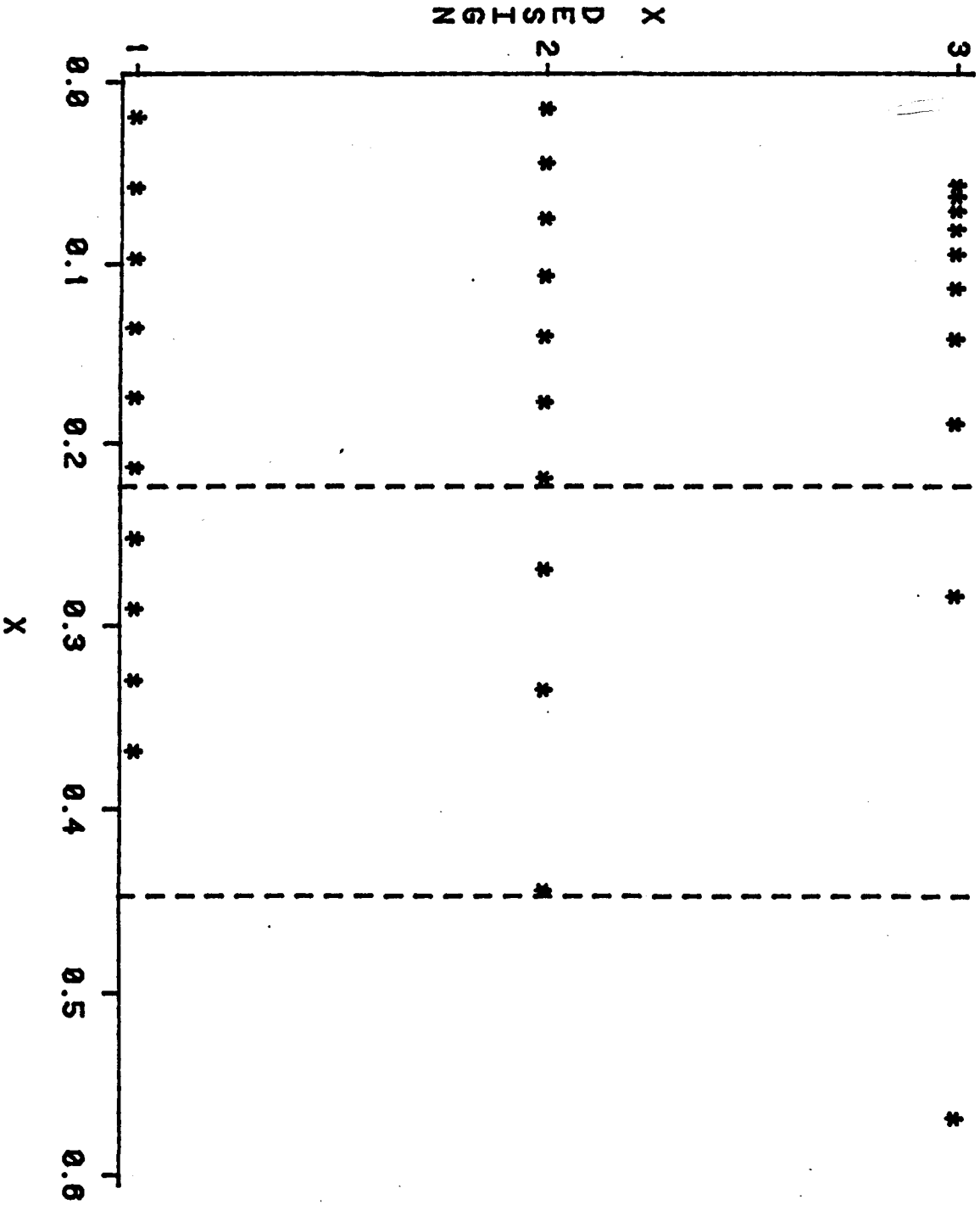

FIG. 1

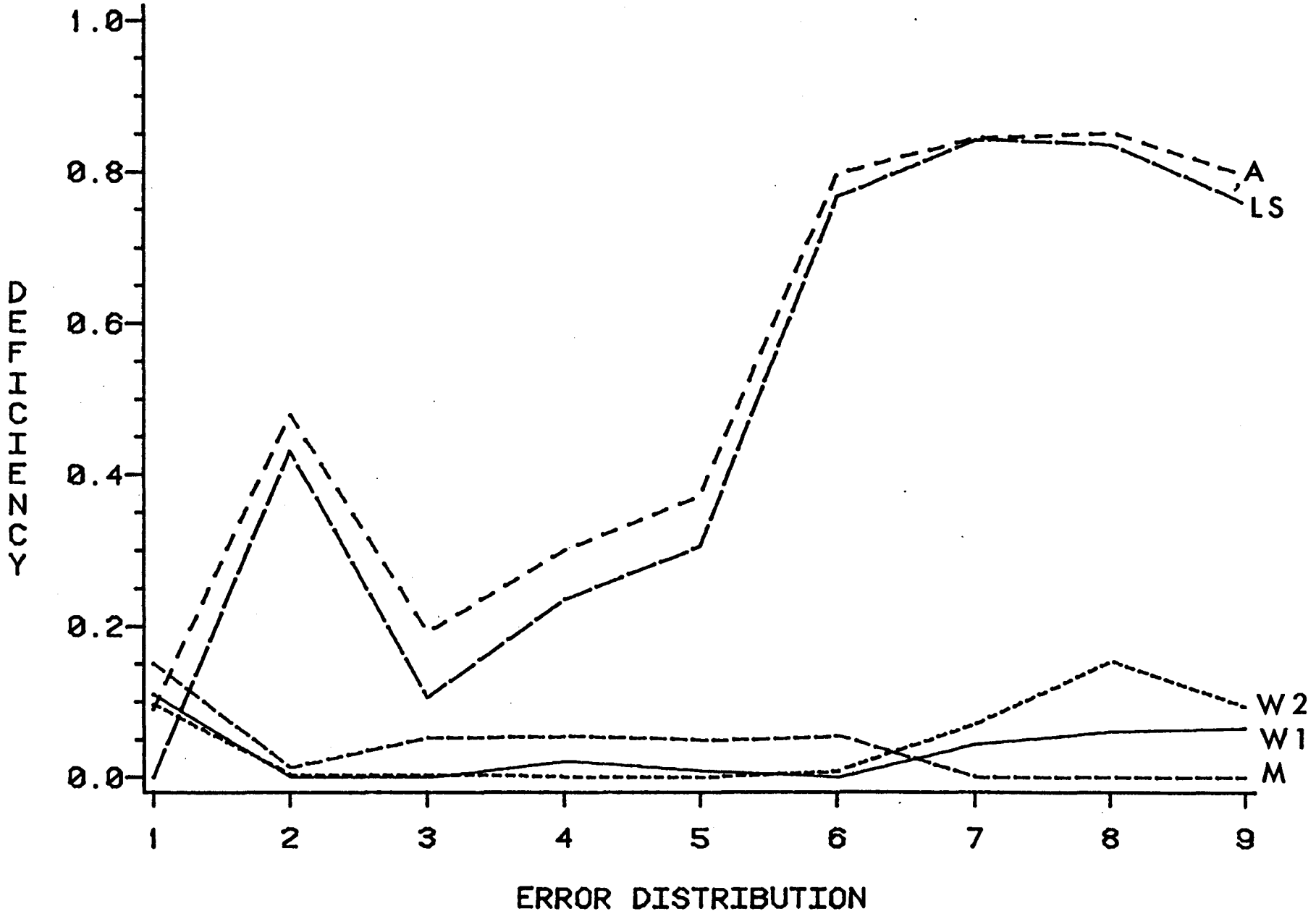
1.0
0.8
0.6
0.4
0.2
0.0
DEFICIENCY
A
LS
W2
W1
M
1
2
3
4
5
6
7
8
9
ERROR DISTRIBUTION

FIG. 2

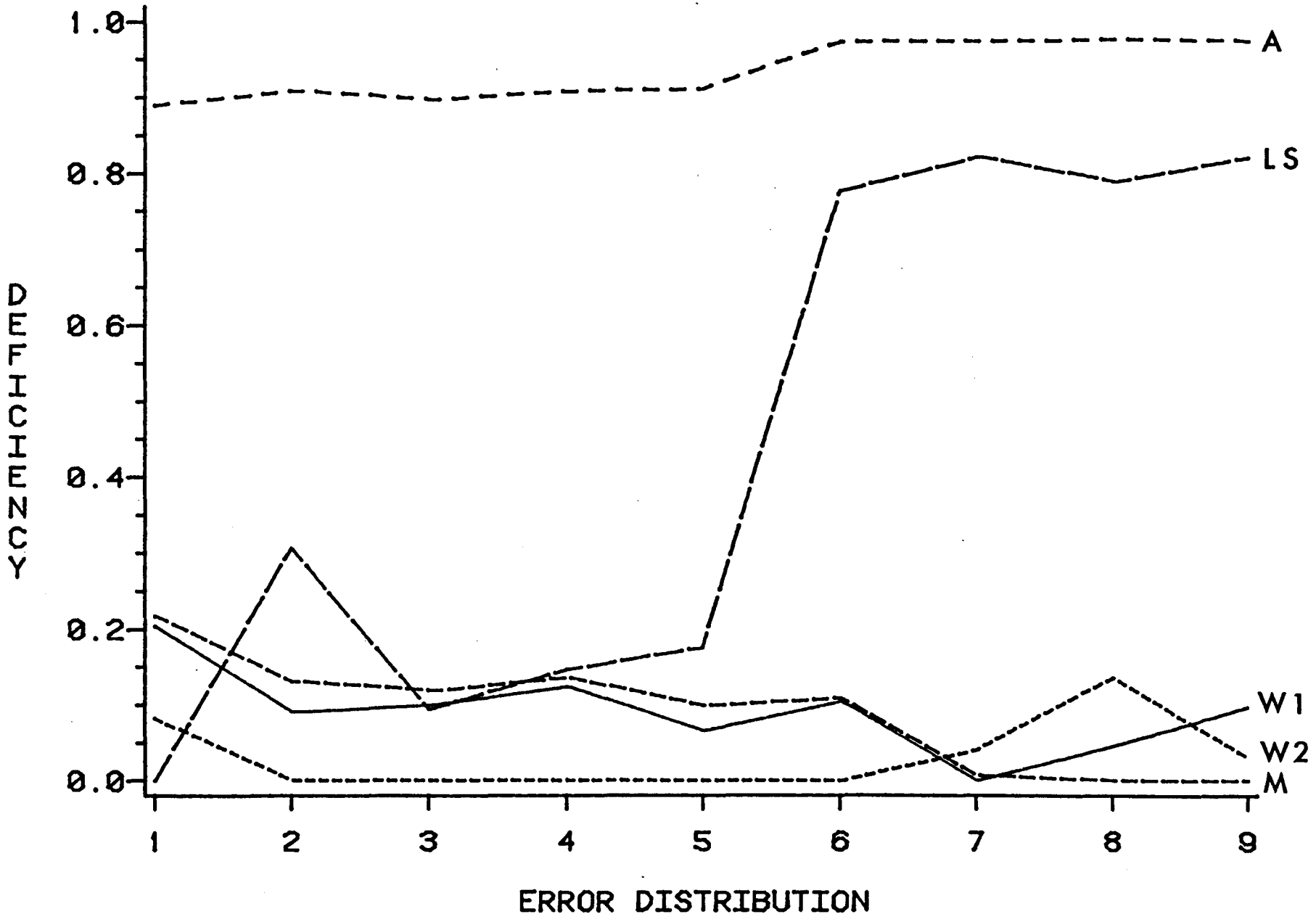
1.0
0.8
0.6
0.4
0.2
0.0
DEFICIENCY
A
LS
W1
W2
M
1
2
3
4
5
6
7
8
9
ERROR DISTRIBUTION

FIG. 3

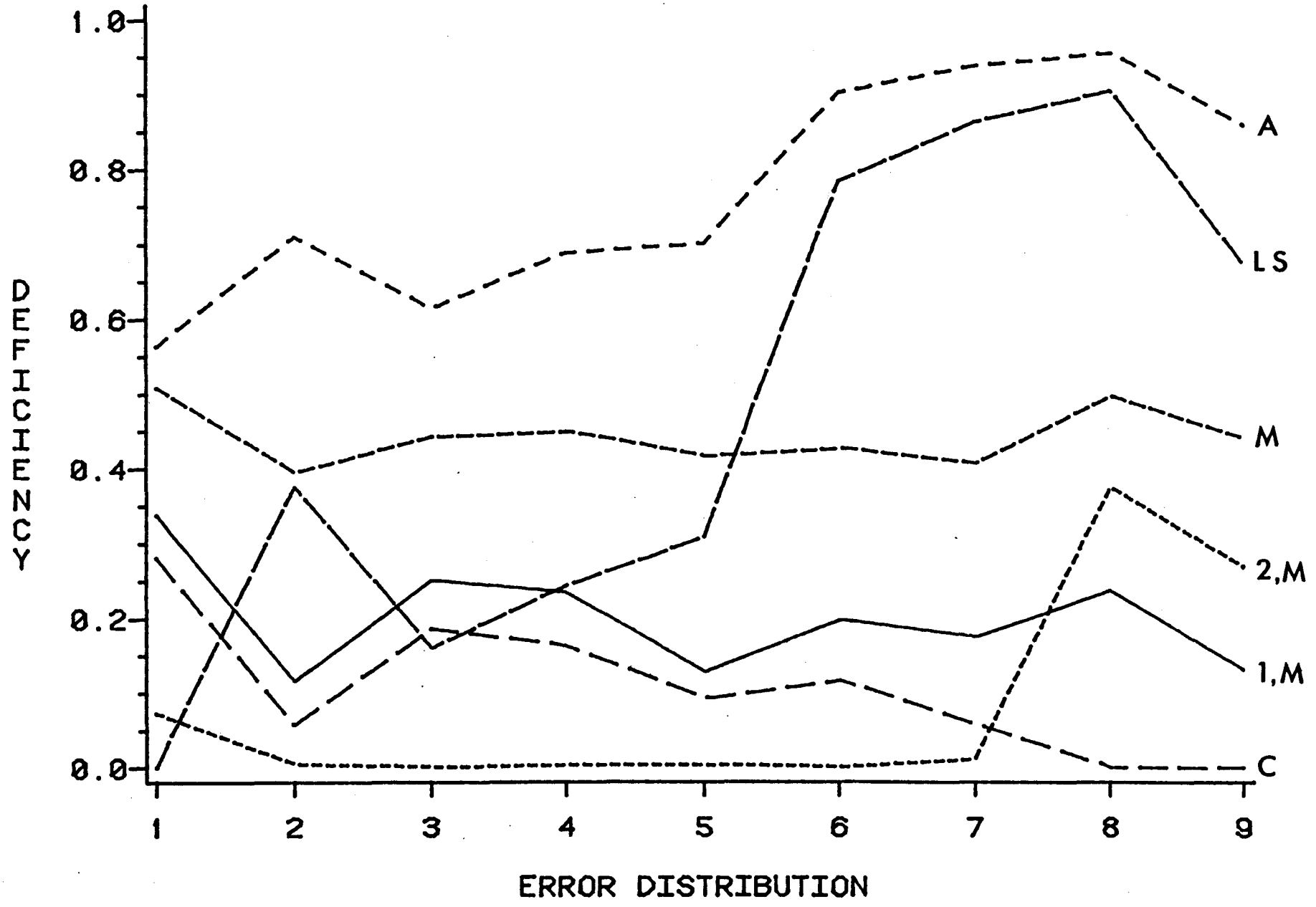
1.0
0.8
0.6
0.4
0.2
0.0
DEFICIENCY
A
LS
M
2,M
1,M
C
1
2
3
4
5
6
7
8
9
ERROR DISTRIBUTION

FIG. 4

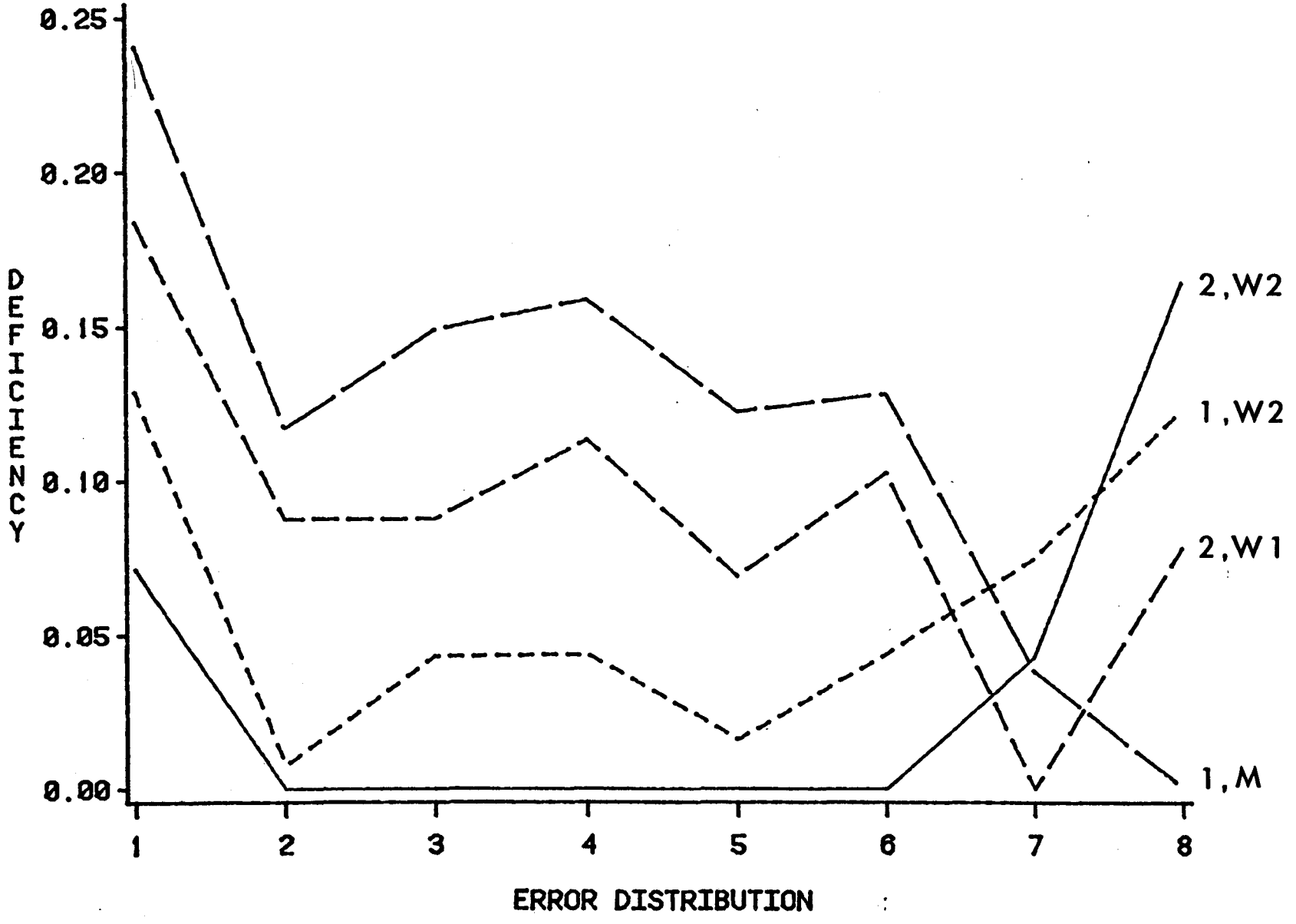
0.25
0.20
0.15
0.10
0.05
0.00
DEFICIENCY
1
2
3
4
5
6
7
8
ERROR DISTRIBUTION
2,W2
1,W2
2,W1
1,M

FIG. 5