1987

Maximum Likelihood Fuzzy Range for Errors in Variables Model

H D. Vinod

ISSN:0318-725X
ISBN:0-7714-0825-0

RESEARCH REPORT 8704

MAXIMUM LIKELIHOOD FUZZY RANGE FOR ERRORS IN VARIABLES MODEL

H.D. Vinod

Department of Economics

University of Western Ontario

London, Ontario, Canada

January 1987

MAXIMUM LIKELIHOOD FUZZY RANGE FOR ERRORS IN VARIABLES MODEL

H.D. Vinod

Fordham University, New York.

ABSTRACT

The use of ordinary least squares(OLS) estimator in the presence of measurement errors (fuzziness) in many variables can lead to logical difficulties and inconsistencies (e.g. "second best" problem). We illustrate the feasibility of the well known, but rarely used maximum likelihood (ML) estimator. From the ML estimates of unobservable regressors we suggest a fuzzy symmetric range of their values near the observable counterparts. A simple estimator called AV from certain sufficient statistics based on the fuzzy range is suggested, and its properties are investigated. It is shown to yield as good as, or better estimator than one based on randomization within the fuzzy range. A textbook example is used for illustration and simulation.

May 1986

Address for Correspondence Prof. H. D. Vinod

92 Hillside Ave., Tenafly, NJ 07670

Tel. (201) 568-5976

1. INTRODUCTION

Social scientists have long recognized that their problems when the data are measured with errors can be different from those of natural scientists in following respects. (i) National Bureau of Standards or a similar institution maintains standard reference measurements against which length, weight, pressure, temperature, etc. are measured. Such precise and fixed reference measurements are often absent in social sciences. (ii) The activity of measurement has an effect on the measured values (Hawthorn effect). (iii) The reaction of economic or social agents is sensitive to the measured values even if they may be incorrect. For example, the reported cost of living index affects some wages and rents automatically. (iv) There is greater reliance on sampling and aggregation leading to unknown imprecision. The decision makers who use the socioeconomic data have a healty skepticism about the data and generally assume that there is a fuzziness range near the estimated relations. The traditional confidence intervals do not allow for measurement errors, but are informally used by some practitioners to express their skepticism about the estimated relationships. We explicitly recognize that when measurement errors are important, the relations between economic variables are not properly represented by "thin" regression lines (surfaces) as in natural sciences, but should be "thick" ,i.e., having a loss function with a flat bottom. The conventional errors in variables model (EVM) does not seem to properly allow for the special situation faced by the social sciences, even though Kendall and Stuart(1979, sec 29.15) do plot confidence ellipsoids around each measured values, and seem to come heuristically close.

The theory of maximum likelihood(ML) estimation for the EVM is well developed and appears in textbooks including Dhrymes(1978, pp 242-260), usually without numerical examples, possibly because computer software is not readily available. An important limitation of the ML estimator for the EVM is that it is necessary to specify the covariance matrix of measurement errors. Hence the practitioners end up using OLS. The logical difficulties with OLS in this case are illustrated by the following two "second best" results due to Garber and Klepper(1980) when a model contains some correctly and some incorrectly (mis) measured variables: (a) The bias of the regression coefficients associated with the correctly measured variables may not decrease even though the measurement error decreases. (b) The above bias may not increase if a proxy variable is omitted. Klepper and Leamer(1984) propose potentially useful diagnostics for this problem, but do not provide ML estimates. Our discussion is based on a general model where only some variables are mismeasured, and we allow for more general loss functions.

Varian(1985, p456) has proposed a nonparametric approach to testing the null hypothesis of cost minimization by economic agents, despite measurement errors. He suggests that specification of the "likely magnitude of measurement errors" should be much "less difficult" than specifying functional forms. He suggests that a certain small (e.g. 0.5) percent of the standard deviation of the observed variable may be used to approximate the standard deviation of its measurement errors in appropriate units. Accordingly we specify a diagonal matrix of measurement error variances, making the ML estimator consistent. We propose an ML estimator (See Result 2 below.) of the unobservables, using the symmetry of measurement errors to justify reversing the sign of the difference between the observed counterpart and its ML estimate. The resulting fuzzy range of regressor values arises from measurement errors, and is used to estimate likelihoods that provide equivalent "statistical evidence", Birnbaum(1972).

The plan of the paper is as follows. In Section 2 we present the model, the ML and our AV estimators and indicate their properties. In Section 3 we use Chow's(1983,p.34) example to illustrate the feasibility of ML estimation, and also include a simulation to evaluate potential gains from our proposal. Section 4 gives our conclusions.

2. THE PROPERTIES OF ML ESTIMATOR AND EXTENSIONS.

Consider the usual regression model for errors in variables:

$$y = Z\delta + u, \qquad Z=(X,W), \qquad \delta =(\alpha', \beta')' \tag{2.1}$$

where y is a Tx1 vector of the dependent variable, X is a Txs matrix of s well measured regressors whose t-th row will be denoted by $x_{t.}$, W is a Txr matrix of unobservable variables measured with errors (mismeasured) with t-th row denoted by $w_{t.}$, δ is a (s+r)x1 vector of unknown regression coefficients partitioned into column vectors α and β of dimension s and r respectively, and u is a Tx1 vector of disturbances.

Instead of W we observe W* with measurement error U* whose t-th rows are denoted by $w^*_{t.}$ and $u^*_{t.}$ respectively.

$$W^* = W + U^* \tag{2.2}$$

We assume that the 1 x (r+1) vector $u_{t.}=(u_t, u^*_{t.})$ is normally distributed with mean zero and a diagonal covariance matrix:

$$\mathrm{cov}(u'_{t.} \mid X) = \mathrm{Diag}(\sigma_{11}, \sigma_{22}, \ldots, \sigma_{TT}) = \Sigma$$

written as

$$\Sigma = \mathrm{diag}(\sigma_{00}, \Sigma_{22}) = \sigma_{00}\, \mathrm{diag}(1, \Sigma_0) = \sigma_{00}\, \Sigma_0^* \qquad (2.3)$$

which defines the notation, and where the r diagonal elements of Σ_{22} are measurement error variances, which are assumed to be known up to a scalar multiple (i.e. Σ_0^* is known). It is well known that without such an assumption the ML estimator is inconsistent, Kiefer and Wolfowitz(1956).

The log likelihood function in terms of the observables y_t and $w^*_{t.}$, the parameters and the unobservables is:

$$L = (-T/2)(r+1)\ln(2\pi) - (T/2)\ln|\Sigma| - (1/2)\sum_{t=1}^{T} (K\, \Sigma^{-1}\, K') \qquad (2.4)$$

where

$$K = (K_1, K_2) = (y_t - x_{t.}\alpha - w_{t.}\beta\ ,\ \ w^*_{t.} - w_{t.}) \qquad (2.5)$$

which defines K as a 1 x (r+1) row vector with two components: K_1 a scalar; and K_2 a 1 x r vector representing unobservable errors from (2.1) and (2.2) respectively. Differentiating L with respect to the unobservable $w_{t.}$, setting the result equal to zero, and considerable manipulation following Dhrymes (1978 p. 252) we have:

$$K_2 = w^*_{t.} - w_{t.} = (-1/\mu)(K^*_1 \beta' \Sigma_{22}) \qquad (2.6)$$

where

$$\mu = \sigma_{00} + \beta' \Sigma_{22} \beta \qquad (2.7)$$

and where

$$K^*_1 = (y_t - x_{t.}\alpha - w^*_{t.}\beta)$$

and

$$K = (1/\mu)\ K^*_1\ (\ \sigma_{00}\ ,\ \ -\ \beta'\Sigma_{22}) \tag{2.8}$$

for substitution in (2.4) so as to yield a "concentrated" likelihood function. This eliminates the unobservable nuisance (incidental) parameters W by a direct method, listed as the 9-th method by Basu(1977).

Further maximizing with respect to α, β and Σ one obtains the maximum likelihood (ML) estimators. The next step in ML estimation is to compute

$$A = (1/T)\ (y,\ W^*)'\ (I-X(X'X)^{-1}X')\ (\ y,\ W^*),$$

and the diagonal matrix Σ_0^* from (2.7) with known Σ_0. Now we can compute the eigenvalues λ_i and the corresponding eigenvectors γ_i based on the following determinantal polynomial equation:

$$|\lambda\ \Sigma_0^*\ -\ A|\ =\ 0 \tag{2.9}$$

If λ_1 is the smallest eigenvalue, a normalization of γ_1 such that its first element is unity yields $\gamma_1 = (1,\ -\ \hat{\beta})'$, hence the ML estimate of the β vector. This manipulation is similar to the limited information maximum likelihood (LIML) estimation in econometrics.

Next we estimate the coefficients of well measured regressors as:

$$\hat{\alpha} = (X'X)^{-1}\ X'\ (y\ -\ W^*\beta) \tag{2.10}$$

and the scale factor is estimated by the consistent estimator:

$$\hat{\sigma}_{00} = \hat{\lambda}_1 \tag{2.11}$$

To find standard errors note that $T^{1/2}(\hat{\delta}-\delta)$ is asymptotically normal $N(0,\mu\psi)$ where

$$\psi = Q^{-1}\begin{bmatrix} 0 & 0 \\ 0 & R \end{bmatrix} Q^{-1} + Q^{-1} \tag{2.12}$$

where we estimate R and Q by $\hat{R}$ and $\hat{Q}$ defined as follows

$$\hat{R} = \hat{\sigma}_{00}\left(\textstyle\sum_0 + (2\hat{\sigma}_{00}/\hat{\mu}^*) \sum_0 \hat{\beta}\hat{\beta}'\sum_0\right),$$

$$\hat{Q} = (1/T)\begin{bmatrix} X'X & X'W^* \\ W^{*\prime}X & W^{*\prime}W^* - T\hat{\sigma}_{00}\sum_0 \end{bmatrix}$$

$$\hat{\mu}^* = \hat{\sigma}_{00}\,(1 + \hat{\beta}'\textstyle\sum_0 \hat{\beta}), \text{ and } \sum_0 = (1/\hat{\sigma}_{00})\sum_{22} \tag{2.13}$$

This completes our review of the ML estimator $\hat{\delta}$ for (2.1). For the case when all regressors are subject to measurement error, the following Result has been given by Brown(1982), and provides an alternative representation of the ML estimation of $\hat{\alpha}$ and $\hat{\beta}$ obtained from (2.9) and (2.10). More generally we have:

<u>Result 1</u>

The maximum likelihood estimator of δ is also given by

$$\underset{\bar{\delta}\ \bar{Z}}{\min}^{-1}\left[(y - \bar{Z}\bar{\delta})'(y-\bar{Z}\bar{\delta})\sigma_{00}^{-1} + (W-W^*)' \textstyle\sum_{22}^{-1}(W-W^*) \right] \tag{2.14}$$

where the symbol to the left of the bracket means that value of δ for which the bracketed function of $\bar{\delta}$ and $\bar{Z}$ is a minimum for $-\infty < \bar{\delta} < \infty$ and $-\infty < \bar{Z} < \infty$. The maximum likelihood estimates of W are obtained by o [ ]/oW = 0 where [ ] is the bracketed quantity in (2.1). Upon substituting a4nd setting o [ ]/oδ = 0 we have:

$$\hat{\delta} = (\hat{Z}'\hat{Z})^{-1}\hat{Z}'y \qquad (2.15)$$

where $\hat{Z} = [X, \hat{W}]$ contains the maximum likelihood estimates $\hat{W}$ obtained as follows. Its t-th row is given by:

$$\hat{w}_{t.} = w^*_{t.} - \hat{K}_2 \qquad (2.16)$$

where the maximum likelihood estimate of the r x 1 vector of measurement errors $\hat{K}_2$ is obtained by substituting the ML estimates $\hat{\alpha}$, $\hat{\beta}$ and $\hat{\sigma}_{00}$ in (2.6). Equation (2.15) bears a convenient formal analogy with the OLS, which will be exploited below. Since the ML estimator for δ is consistent under our assumptions, Result 1 shows that in our context, eliminating the unobservable nuisance parameter W by "estimating it away" (Basu's(1977) 3-rd method) is equivalent to the "direct" method (Basu's 9-th) mentioned above. When we evaluate the process of estimating W by (2.16) we have:

<u>Observation 1:</u>

Note that the sign of $\hat{K}_2$ used in estimating W from (2.16) is arbitrary, and there is no reason to think that the unobservable $w_{.t} > w^*_{.t}$ any more than $w_{.t} < w^*_{.t}$ due to measurement errors.

This assumes symmetry of measurement errors, which has been made by many authors including Zellner(1970) and Vinod(1982 a,b). Let K_2^p denote a T x r matrix such that it contains the absolute values of $\hat{K}_2$ along the t-th row for each row of data. The superscript p reminds us that positive (absolute) values are involved. Now using Observation 1 and (2.14) we state:

<u>Result 2</u>

A Maximum Likelihood estimate of the symmetric fuzzy range in the mismeasured data is given by:

$$W_L = W^* - K_2^p \le W \le W^* + K_2^p = W_U \qquad (2.17)$$

where the subscripts L and U represent the lower and upper values respectively.

The result is true because each row of K_2^p is based on the ML estimate of the measurement error in the corresponding row of the T x r matrix W. The "statistical evidence," defined by Birnbaum(1972), and contained in any values inside the fuzzy range above, may be regarded as being equivalent or indistinguishable, from a practical viewpoint of the data user or (market) agent. Asssuming that the likelihood does not vanish for all values of the parameter space, Joshi(1976) shows that in a (fuzzy) range of values based on "accuracy" with which a random variable can be measured, the likelihood functions within the range are equal to each other, up to an "undetermined" positive constant. Thus Birnbaum's likelihood axiom of statistical evidence is satisfied and implies the conditionality axiom, according to Birnbaum's theorem 2. Hence there is an "ancillary" statistic having a specified probability distribution. Choosing the uniform distribution we note:

Observation 2

Once the fuzzy range (2.17) is determined, we may reasonably assume that the true values are uniformly distributed inside the range. Hence from a well known property of the uniform distribution, the limiting order statistics W_L and W_U are jointly sufficient and complete for estimating W.

In the following discussion we use the fuzzy range (2.17) to suggest a "thick" line (surface) connecting the variables, which seems to yield a realistic representation of the underlying relationship. Furthermore each point can yield ML estimation of δ with reference to other evidentially equivalent (Birnbaum-Joshi) likelihood functions within the fuzzy range (2.17). Our ML estimator is implemented by OLS-like computations thanks to the Result 1 above, which is numerically verified in an example of Section 3.

Among modifications of the ML estimator based on the fuzzy range, an attractive possibility is to randomize within the uniformly distributed range according to standard statistical methods. The main disadvantage of randomization methods appears to be the unfortunate fact that each randomized run yields a different estimate. If the assumptions implying "sufficiency" of W_L and W_U (Observation 2 above) are accepted, we may use the Rao-Blackwell theorem to justify the use of the following simple estimator, which avoids randomization. Thus we have:

<u>Result 3</u>

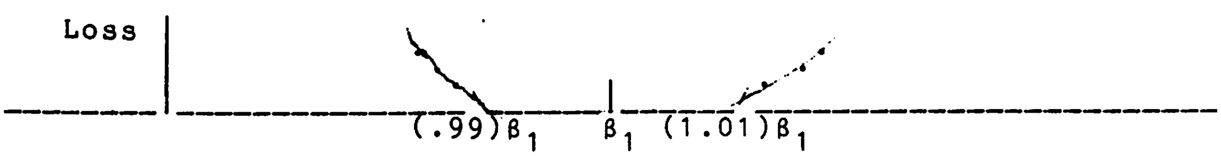
Define $\hat{Z}_U = (X, W_U)$ and $\hat{Z}_L = (X, W_L)$ as the two limiting estimates of Z using (2.15). They yield $\hat{\delta}_U$ and $\hat{\delta}_L$ respectively with appropriate substitutions in (2.15). Now the smoothed estimator:

$$\hat{\delta}_{AV} = (1/2)[\, \hat{\delta}_U + \hat{\delta}_L] \tag{2.18}$$

is a uniformly better estimator than a randomized estimator, provided the underlying loss function is strictly convex. If the loss function is (only) convex, (2.18) is no worse than a randomized estimator.

Lehmann's(1983, p51) proof of a corollary to the Rao-Blackwell theorem is directly applicable here, since the estimator of δ may be regarded as a transformation of W and X matrices through (2.15), using the sufficient statistics W_L and W_U. Detailed arguments justifying the averaging of the estimators based on two sufficient statistics of the uniform distribution may be omitted, since they are similar to those in a textbook, Ferguson(1967, p125).

The estimate (2.18) remains a "point estimate" of the regression coefficients, representing a thin line or a surface. In light of the discussion in the introduction to this paper, there may be an interest in a more fuzzy (thick) description of the estimated (surface) relationship for application in social sciences, where the agents are aware of the fuzziness of the data. For the purpose of discussion it is convenient to assume that the percent fuzziness of the regression coefficients, is the same as the percent fuzziness in the regressor, namely one percent. Of course, any known percentages (including zero) can be used in a particular application, making our treatment more general than what is found in the literature. The perceived loss function associated with a point estimate β_1 may be represented as follows.



The illustration suggests that the loss function has a flat bottom, representing the idea that the agents are skeptical about small departures from reported point estimates.

3. A NUMERICAL EXAMPLE AND SIMULATION

Chow (1983 p. 34) gives 1954-1965 time series data on demand for computers. The dependent variable is $y_t = \ln (q_t/q_{t-1})$, natural log of the ratio of the quantity index of computers at time t to the index at t-1. The well measured regressor is the column vector of ones (the intercept). The two mismeasured regressors are the two columns of $w^*_{t.}$ $= [\ln p_t, \ln q_{t-1}]$ where p_t is a price index appropriately deflated by the GNP deflator, and q_{t-1} is the quantity index mentioned above. Their respective measurement error standard deviations are assumed to be one percent of their own standard deviations in the observed data. This defines the 2x2 diagonal Σ_0 matrix with elements which are squares of 0.0081621 and 0.0190597. Table 1 reports the estimated coefficients by three methods: (i) OLS similar to Chow's. (There are minor discrepancies in the last decimal). (ii) Maximum Likelihood (ML) estimator from (2.9) and (2.10), and finally (iii) Our AV estimator defined by (2.18) and justified by the Result 3 above.

For the Chow data the numerical estimates for the OLS, ML, and AV estimator are reported in Table 1, and are all very close to each other. For the proposed AV estimator of (2.18) Table 1 reports $\hat{\alpha}$ = 2.948, $\hat{\beta}_1$ = -0.363, and $\hat{\beta}_2$ = -0.252, along with the one percent fuzzy range for the slope coefficients associated with the mismeasured regressors. Table 1 also reports the estimates of standard errors (SEs) based on the asymptotic variance covariance matrix of ML regression coefficients given in (2.16) and (2.19), suggesting improved asymptotic efficiency of ML estimators. The SE of the OLS coefficients is seen to be larger than the asymptotic SE of the ML coefficients. The asymptotic SE for ML may serve as an approximation to the asymptotic SE of the AV estimator, because it is based on "equivalent" likelihoods in terms of their statistical evidence (Birnbaum-Joshi) in the fuzzy range.

3.2..A Simulation

The purpose of our simulation is to evaluate the mean squared error (MSE) and assess the properties of the ML, and especially the AV estimator having a fuzzy range. Also, a simulation often reveals any unexpected undesirable properties of the estimators. We use the same data on the right hand side regressors as in the numerical example above, and assign the following positive known regression coefficients $\alpha = 4$ and $\beta = (1,2)$. A 1% measurement error is assumed in the two mismeasured regressors. The t-th element of y given by

$$y_t = 4 + w^*_{t.}(1,\ 2)' + 0.1\ N(0,1) \qquad (3.1)$$

where N(0,1) is a unit normal realization having zero mean and unit variance, and where the true coefficients are 1 and 2. Repeating (3.1) T times we obtain the T observations for an artificial dependent variable. Thus we have k=1,2,...,S (S=225) regression problems having S sets of simulated y vectors.

A crucial feature of measurement errors is that some columns of the regressor matrix may be observed subject to error (perturbation). Hence for an appropriate simulation we perturb the W* matrix by adding a T x r matrix of N(0,1) random numbers, scaled by post multiplying by an r x r diagonal matrix with elements (0.0081621, 0.0190597), used earlier in difining the matrix Σ_0 . For our k-th regression problem, (k=1,2,...,225) we create a separate 11 x 2 matrix of random perturbations to be added to the observed W* matrix.

The regression coefficients of the k-th regression problem are estimated by each of the three methods of the previous section (OLS, ML, and AV). These are compared to the assigned true values (4,1,2), which are themselves subject to one percent fuzziness. In light of the special loss function illustrated above, we need to modify the conventional definition of squared errors as follows. Consider the following interval in which the estimated value of β_1 may lie (say). Let one such estimate be at 1.03.

```
          0.99     1     1.01          1.03
|-----|.....|.....|-----|-----|-----|
     (.99)β1    β1  1.01β1         β̂1
```

where the fuzziness range is between 0.99 to 1.01 times β_1 (=1). Since we assume that the true value is 1, what is the absolute error in estimating it as 1.03, when the assumed loss function has a flat bottom with a one percent spread around the true value? Recall that we are allowing for the fact that in social science applications it may be realistic to assume that the true value itself is fuzzy by a "small" percent. If the estimate is 1.01 the (economic) agent will presumably regard it to have a negligible or zero error. Thus, all absolute errors need to be adjusted downward in absolute terms by $(0.01|\beta_1|)$, replacing the negative adjustments by zeros. For $\hat{\beta}_1$ = 1.03 alluding to the above illustration, the absolute error is only 0.02 and not 0.03. In general, a computer algorithm can reduce all absolute errors by the formula:

$$AE_k = \max[0, (|\hat{\beta}_i - \beta_i| - 0.01|\beta_i|)], \quad (i=1,..,r) \qquad (3.2)$$

for the mismeasured regressors, where k=1,...,S refers to the S regression problems. For the well-measured regressors we use $AE_j = |\hat{\alpha}_j - \alpha_j|$ for j = 1,..,s; on the grounds that the coefficients for well-measured regressors are not subject to the same fuzziness.

To simulate the proposed AV estimator of (2.18) we need to make the following adjustment to the squared errors due to our special loss function.

```
         0.99    1      1.01   1.02  1.03  1.04
|-----|:::::|:::::|-----|=====|=====|-----
             β1                 β̂1(AV)
```

where the estimator $\hat{\beta}_1(AV)$ is also surrounded by a fuzzy range from 1.02 to 1.04 (say) to represent their interpretation by appropriate agents. Clearly, the error AE_1 for this fuzzy estimator is given by the distance between 1.01 and 1.02 which equals 0.01 in the above illustration, since the lower bound of $\hat{\beta}_1$ (=1.02) is higher than the upper bound (=1.01) of the true value by 0.01. In general, we use the following formula to compute the absolute errors for the AV estimator

$$AE_k = \max\,[0,\ (0.99|\hat{\beta}_i| - 1.01|\beta_i|)],\ \text{if}\ |\hat{\beta}_i| > |\beta_i|, \quad (3.3)$$

and

$$AE_k = \max\,[0,\ (0.99|\beta_i| - 1.01\ |\hat{\beta}_i|)],\ \text{if}\ |\hat{\beta}_i| < |\beta_i| \quad (3.4)$$

The S(=225) regression problems give rise to S estimates of AE's for each coefficient by each of the three estimation methods. Mean squared error (MSE) is simply the average over the S squared errors (AE_k^2). Table 2 also reports relative MSE's as ratios with respect to the MSE for OLS. The standard error (SE) of our simulation's estimate of the MSE is the standard deviation over the S squared errors, divided by the square root of S, and is reported in Table 2. Of course, this SE can be reduced by choosing a larger S than 225. We also report the median of squared errors, since it is well known that their mean (=MSE) may not be a sufficiently robust indicator of the central tendency of the distribution of squared errors. The simulation suggests that the ML estimator does not reduce the MSE of OLS, especially when one considers the medians and extreme values of squared errors reported in Table 2. We conclude that ML may avoid the logical (e.g. second best) problems of OLS, but may not reduce the MSE of OLS. The simulation is encouraging for our AV estimator, even though the MSE for the intercept is no smaller than that of OLS. The median (the maximum value) of squared errors for the OLS intercept is 0.240896 (6.483036), which is slightly smaller (larger) than the corresponding median (the maximum) of the AV estimator 0.240897 (6.483034). We may conclude that the regression coefficients of well measured regressors are neither improved, nor worsened by our AV estimator. For the slope coefficients associated with the two mismeasured regressors AV estimator offers a statistically significant (see the SE of MSE) reduction in the MSE of OLS, which is confirmed by the maximum value of squared errors and medians. Reporting a fuzzy range of point estimates seems to be worthwhile, and perhaps less misleading to the user in the presence of measurement errors.

(Our computations were made on an IBM-PC-AT computer equipped with a 80287 math coprocessor using GAUSS language developed by Applied Technical Systems, Kent, Washington. A listing of the program is available upon request. Simultaneous reduction of two matrices of the eigenvalue problem was made after first computing a square root matrix by the Cholesky decomposition. The eigenvectors were adjusted by pre and post multiplying by the appropriate transform of the square root matrix. The simulation does not reveal any undesirable properties, and is not intended to confirm all the theoretical developments of section 2 above.)

4. FINAL REMARKS

We have discussed some special features of the measurement error problem in social sciences, since their measured values can have a reality of their own, and do influence actions of agents. The "statistical evidence" in Birnbaum's (1972) sense, contained in mismeasured observations should reflect the fact that they are treated as equivalent by skeptical agents. This leads to a recognition that the measured regression surfaces should be "thick" with one percent (say,) fuzzy range around the point estimates. This is characterized by allowing the relevant loss function for social sciences to have a flat bottom, representing an explicit fuzzy range in the parameter space.

Our first result is a minor generalization of Brown's(1982) characterization of the ML estimator as an application of OLS to corrected data. Our second result provides a maximum likelihood estimate of the fuzzy range based on measurement errors. Our third result develops the AV estimator (2.18) based on smoothing by a simple average of the upper and lower estimates of regression coefficients over the fuzzy range. Although the AV estimator is based only on evaluations at the two limits, rather than randomization over the entire fuzzy range, we show (using Rao-Blackwell theorem) that it will not be inferior to the latter, under appropriate assumptions.

The paper also indicates the practical feasibility of ML estimators along with asymptotic standard errors in a general setting, i.e. when only some of the regressors are subject to measurement errors, illustrated by a textbook example. A simulation suggests the potential advantages of ML and AV estimators over OLS, especially for coefficients of regressors subject to measurement errors. The suggestions in this paper seem to deserve further study.

REFERENCES

Basu, D.(1977) "On the Elimination of Nuisance Parametrs" Journal of the American Statistical Association 72, 355-366.

Birnbaum, A.(1972) "More on Concepts of Statistical Evidence," Journal of the American Statistical Association 67, 858-861.

Brown, M.L.(1982) "Robust Line Estimation with Errors in Both Variables," Journal of the American Statistical Association 77, 71-79.

Chow, G.C. (1983) Econometrics, McGraw-Hill, N.Y.

Dhrymes, P.J.(1978) Introductory Econometrics, Springer Verlag, N.Y.

Ferguson, T.S.(1967)Mathematical Statistics A Decision Theoretic Approach, Academic Press, N.Y.

Garber, S. and S. Klepper (1980) "extending the Classical Normal Errors-in-Variables Model," Econometrica 48,1541-1546.

Joshi, V.M.(1976) "A Note on Birnbaum's Theory of the Likelihood Principle," Journal of the American Statistical Association 71, 345-346.

Kendall, M.G. and A. Stuart (1979), The Advanced Theory of Statistics, Vol. 2, 4th edition, Macmillan, N.Y..

Kiefer, J. and J. Wolfowitz(1956) "Consistency of the Maximum Likelihood Estimator in the Presence of Infinitely Many Incidental Parameters Annals of Mathematical Statistics 27, 887-906.

Klepper, S. and E. Leamer (1984) "Consistent Sets of Estimates for Regressions with Errors in All Variables," _Econometrica_ 52, 163-183.

Lehmann, E.L.(1983) _Theory of Point Estimation_ , Wiley, N.Y.

Varian, H.R.(1985) "Non-parametric Analysis of Optimizing Behavior With Measurement Error" _Journal of Econometrics_ 30, 445-458.

Vinod, H.D. (1982a), "Maximum Entropy Measurement Error Estimates of Singular Covariance Matrices in Undersized Samples," _Journal of Econometrics_, 20, 163-174.

Vinod, H.D. (1982b), "Enduring Regression Estimator," in O.D. Anderson, ed., _Time Series Analysis: Theory and Practice_ 1 (Amsterdam: North Holland), 397-416.

Zellner, A.(1970) "Estimation of Regression Relationships Containing Unobservable Independent Variables _International_ Economic Review 11,441-454.

Table 1

Regression of log change in Demand for Computers on log price deflated by GNP deflator and lagged quantity index

	o	8	8	SD(resid)
OLS	2.948232	-0.363298	-0.252442	10192.417439
SE OLS	0.701803	0.172499	0.073871	
ML	2.948292	-0.363313	-0.252448	10192.417444
SE ML	0.598499	0.147107	0.062997	
upper	2.948215	-0.363294	-0.25244	10192.417443
lower	2.948248	-0.363303	-0.252444	10192.363073
AV	2.948232	-0.363298	-0.252442	10192.417439
0.99AV	2.918749	-0.359665	-0.249917	
1.01AV	2.977714	-0.366931	-0.254966	
ML check	2.948292	-0.363313	-0.252448	10192.417444

Note: Last column contains 100,000 times the residual standard deviation. First column contains the intercept. The Standard Error(SE) of the ML is asymptotic. The 1% fuzzy range of the AV estimator is indicated. The last row agrees with ML and numerically verifies Result 1 in the text.

Table 2

Relative Mean Squared Errors and Medians for Simulation

		OLS	ML	AV
relative MSE alpha		1	1.000152	1
MSE		0.518401	0.51848	0.518401
standard error MSE		0.053362	0.053384	0.053362
relative MSE beta1		1	1.00017	0.906138
MSE		0.028376	0.028381	0.025713
standard error MSE		0.002977	0.002978	0.002818
relative MSE beta2		1	1.000219	0.617461
MSE		0.003676	0.003677	0.00227
standard error MSE		0.000471	0.000472	0.000366

Minimum, median and the maximum from 225 squared errors

alpha 1	Min:	2.099109E-005	2.843512E-005	2.098808E-005
	Median:	0.240896	0.240655	0.240897
	Max:	6.483036	6.488812	6.483034
beta1				
	Min:	0	0	0
	Median:	0.011275	0.011309	0.009108
	Max:	0.315224	0.315532	0.297825
beta2				
	Min:	0	0	0
	Median:	0.001018	0.001016	0.000138
	Max:	0.059884	0.059942	0.049313