ERRORS IN VARIABLES IN ECONOMETRICS: NEW DEVELOPMENTS AND RECURRENT THEMES *

by P.A. BEKKER, T.J. WANSBEEK and A. KAPTEYN

Abstract The paper reviews some old and new approaches to the analysis of linear models with errors in variables. The emphasis is on the identification problems that usually arise in errors-in-variables models and on the various types of additional information that econometricians have invoked to be able to estimate parameters consistently. The approaches discussed include instrumental variables, grouping, simultaneous equations, multiple equations and bounds on measurement error variances.

Key words: *Latent variables, linear models, identification*

1 Introduction

Most of econometrics is concerned with the estimation of relationships between variables. In the generic simple case, a linear relation between k, say, random variables is postulated and the question is considered how parameters can be estimated for a sample of observations.

If these observed values would all lie exactly in a $(k-1)$-dimensional hyperplane, estimation would be trivial. As they usually don't one has to become explicit about the stochastic process that disperses the data around the postulated hyperplane. The usual formulation is to split up the set of variables into a single "dependent" variable and $k-1$ "explanatory" variables, and to write in what is sometimes referred to as the Dutch notation $y = X\beta + u$, with X and u orthogonal. The "disturbance" term u may represent measurement error in y, the influence of omitted variables orthogonal to X, or both. X is assumed to be measured without error.

This view of the single-equation model and the analogous treatment of the more complicated simultaneous equations model has not always been the paradigm of econometrics. In the thirties, most notably in the work of FRISCH and KOOPMANS, measurement errors in *all* variables were often considered to be the reason for an imperfect fit. This may be a plausible assumption in many cases, but it also spells trouble, as it leaves us with an underidentified model. For consistent estimation we need additional information, like knowledge of the relative sizes of the measurement error variances. Such information can come from informed guesses by the researcher or from outside sources, but cannot be inferred from the data at hand.

In the course of the fourties this way of modelling the imperfect fit was superseded by the development of the now standard linear simultaneous equations

* Financial support by the Netherlands Organization for the Advancement of Pure Research (Z.W.O.) is gratefully acknowledged.

model, in which errors in equations are the only source of random deviations from the hypothesized linear relations between observed variables. Attention for measurement error only re-emerged about ten to fifteen years ago. This cyclical pattern in interest is sketched lucidly by GOLDBERGER (1972) and GRILICHES (1974). GRILICHES (1985) DISCUSSES THE SAME POINT AND GIVES AN OVERVIEW OF MANY ASPECTS OF THE QUALITY OF DATA IN ECONOMETRICS.

In this essay we briefly discuss the main issues in the field of measurement error and describe some recent developments. In Section 2 we discuss the simple linear model with measurement error and state some standard results. Some aspects of the re-emergence of interest in measurement errors over the last decade are mentioned in Section 3. The central theme there is the embedding of equations with error-ridden variables in a more elaborate model. This procedure often allows for identification of parameters.

In this approach, sufficient structure is imposed upon the possible relationships between variables, so that consistent estimators of parameters can be constructed. Recently it has been investigated to what extent inference about regression parameters can be made when there is non-exact information, in the sense that upper bounds are imposed on measurement error variances. In that case, one cannot construct consistent estimators, but one can derive "asymptotic" bounds for the true parameters. This line of research is reviewed in Section 4.

In Section 5 we take one further step back regarding the amount of information one can invoke and consider the question of what can be said when, apart from the standard assumptions of the errors-in-variables model, no extra information is available. In a sense this is also a step back in time as we reconsider a classical result due to KOOPMANS (1937). In his thesis, he showed that under certain conditions on the matrix of second-order sample moments the vector of regression parameters lies in a polyhedron. We give a new and shorter proof of this result. This proof can be adapted to obtain a new result that is complementary, in a way, to KOOPMANS' result.

2 The basic problem

To fix ideas, consider the simple regression model

$$y = \xi\beta + u, \tag{1}$$

with $y(n\times 1)$ a vector of dependent variables, $\xi(n\times 1)$ a vector of explanatory variables, $u(n\times 1)$ a vector of disturbance terms. The elements of ξ and u are assumed to be independently normally distributed, with means zero and variance $\sigma_{\xi\xi}$ and σ_{uu}, respectively. It is well- known that the ordinary least-squares (OLS) estimator $\hat{\beta}$ of β has many desirable properties including consistency.

Now consider the situation where ξ is not directly observable, because of the presence of measurement error, or because ξ is a mental construct ("permanent income", "intelligence"), or whatever. In that case the elements of ξ are often called "latent variables." Assume that instead of ξ we observe $x(n\times 1)$ such that

$$x = \xi + \epsilon \tag{2}$$

with $\epsilon(n \times 1)$ a vector of measurement errors, with each element independently normally distributed with variance $\sigma_{\epsilon\epsilon}$. The elements of ϵ are distributed independently of the elements of ξ and u. We denote the variance of the elements of x by $\sigma_{xx}' = \sigma_{\xi\xi} + \sigma_{\epsilon\epsilon}$. It is simple to show that when we regress y on x the OLS estimator

$$b = (x'x)^{-1}x'y \tag{3}$$

is inconsistent. In particular, there holds

$$p \lim b = \beta(1 - \sigma_{\epsilon\epsilon} / \sigma_{xx}).$$

So b is biased towards zero. In the model as it stands there does not exist a consistent estimator for β, since the model is underidentified: there are three consistently estimable sample moments: σ_{yy}, σ_{yx}, σ_{xx}, which are related to four unknown parameters, $\sigma_{\xi\xi}$ σ_{uu}, $\sigma_{\epsilon\epsilon}$ and β. This underidentification lends the errors-in-variables problem a touch of hopelessness, as is apparent from a look in even the most recent econometric textbooks.

Up til now, we assumed all variables to be normally distributed, including ξ. A model with ξ stochastic is called "structural", and it is called "functional" when the elements of ξ are considered unknown fixed parameters. A fundamental result due to WALD (1949) implies that identifiability of β in the structural model under the normality assumptions above is equivalent to the existence of a consistent estimator for β in the functional model. So reinterpreting the model in functional terms does not offer a way out of the underidentification (cf. AIGNER et al. (1984)).

However, if in the structural model one may assume that the elements of ξ are not normally distributed, then the identification problem need not arise. REIERSÖL (1950) has shown that β is identified if and only if ξ is non-normal (given normality of u and ϵ). This result pertains to the simple regression model. The question is how this result extends to the multiple regression case, and it has only recently been settled by BEKKER (1985), as follows.

Let $\Xi(n \times k)$ be the matrix of regressors in the multiple regression case, so that (1) is replaced by

$$y = \Xi\beta + u, \tag{4}$$

and let ζ' be a row of Ξ. Then the $k \times 1$-vector β of regression coefficients is identified if and only if there does not exist a non-singular $k \times k$-matrix $A = (a_1, A_2)$ such that $\zeta' a_1$ is distributed normally and independently of $\zeta' A_2$. * So, non-normality of ξ (or Ξ) keeps identification, and consistent estimators of β based

* This corrects an earlier result by KAPTEYN and WANSBEEK (1983). In their paper they claim that the existence of a normally distributed linear combination of ζ is equivalent with non-identification of β.

on second and higher-order moments, for example, can be constructed (e.g. PAL (1980)). In practical work this approach has found little application, as it is intuitively clear that unless the "distance to normality" of the distribution of the elements of Ξ is quite large, the estimators will be unreliable.

Let us return to the simple, structural, normal model introduced in the beginning of this section. Usually, textbooks cite two ways out of the underidentification problem: the method of "instrumental variables" in general, and its particular instance of "grouping of data." In its simplest version, grouping amounts to splitting up the data in two sets, according to x-values above or below the median. The mid-points of both point-scatters in the (x,y)-plane are joined by a straight line, and the slope of this line is then defined as the "grouping" estimator of β. *If* the grouping according to x-values could be guaranteed to be the same as the one according to ξ-values, this estimator is consistent for β. This is a big *"if"*, and does not hold under normality of ξ and ϵ, or in fact under most other plausible distributional assumptions.

Recently, PAKES (1982) gave a definitive treatment of the grouping estimator. Under non-normality we can generally do better than grouping, and, more interestingly, under normality the inconsistency of the grouping estimator is the same as that of the OLS estimator. As PAKES notes, this is as it should be — if the probability limit of the grouping estimator would be a different function of β and $\sigma_{\epsilon\epsilon} / \sigma_{xx}$ than the probability limit of the OLS estimator, both estimators could be combined to construct a consistent estimator of β, which is impossible, due to the underidentification of the model.

Although since its introduction by WALD (1940), grouping has drawn a lot of attention (two versus three groups, multivariate extensions, optimal division in groups, asymptotic distribution, etc.), PAKES' short note can be interpreted as a postmortem.

If we substitute $x-\epsilon$ for ξ (see (2)) in (1), we obtain

$$y = x\beta+v, \tag{5}$$

with $v := u-\epsilon\beta$ a newly defined disturbance term. In (5), the disturbance term and the explanatory variable have non-zero correlation, and this is of course another way to explain the inconsistency of the OLS estimator (3). If we observe a vector $z(n \times 1)$, say, whose elements correlate with the elements of x but do not correlate with the elements of v, then the estimator

$$\tilde{\beta} := (z'x)^{-1}z'y \tag{6}$$

will be consistent for β under weak assumptions. Such a variable z is called an *instrumental variable* (IV), and $\tilde{\beta}$ is the IV estimator of β. IV estimation is the main textbook solution for the measurement error problem. Note that grouping is a particular kind of IV, by letting the elements of z be -1 or +1, according to the sub- or supra-median value of the corresponding element of x. Also note, however, that the elements of z are now dependent upon x and hence on ϵ, so that one of the conditions for consistency of $\tilde{\beta}$ is not met.

In the errors-in-variables context there is no obvious procedure to find instrumental variables; they may or may not be available. In econometrics instrumental variables are used for the estimation of a wide variety of models. The best known of these is the two-stage least-squares estimator for a single equation in a simultaneous system of equations. In such an equation the explanatory endogenous variables correlate with the disturbance term; such correlation was also present in (5). Instruments can in principle be constructed by projecting the explanatory endogenous variables in the equation under consideration onto the space spanned by all exogenous variables in the system. Estimation then takes place by means of least squares after having replaced the explanatory endogenous variables by their projection.

A general treatment of IV has recently been given by WHITE (1984). For the multiple regression equation $y=X\beta+u$, with X an $n\times k$-matrix of explanatory variables, he extensively discusses the properties of estimators of β that are obtained by minimizing the quadratic loss function

$$q = (y-X\beta)'ZPZ'(y-X\beta),$$

where $Z(n\times l)$ is a matrix with l instruments and P $(l\times l)$ is some, possibly stochastic, positive definite weighting matrix. Asymptotic properties of estimators under a variety of non-standard assumptions are derived.

These three elements, underidentification, grouping, and IV, constitute the core of the usual textbook treatment of the subject. Until recently, the errors-in-variables model has not been used frequently in applied work. Since about 1970, there has been a number of developments which have changed this situation considerably.

3 Some recent developments

Since the earliest days of econometrics, much work concentrated on macromodels. A major step forwards was the development of the simultaneous equations model, in the course of the forties, and of adequate estimation methods, in the fifties and early sixties. Although various estimation methods from statistics had been used before in econometric problems, for the first time econometrics had something of their own. In this development, scant attention was being paid to measurement error.

Over the last decade-or-so, the analysis of microdata has become a relatively prominent field of econometric activity, and many new results and insights have been obtained. Measurement errors, or more generally latent variables, are among the popular themes.

A seminal paper is ZELLNER (1970). Let, for example, (1) be a consumption function at the micro-level, relating consumption (y) to "permanent income" (ξ). We only observe actual income (x), cf. (2). Now assume that the researcher is willing to specify a relation that explains permanent income as a linear function of background variables:

$$\xi = \gamma_1 v_1 + \cdots + \gamma_k v_k, \tag{7}$$

where v_1 through v_k are n-vectors of explanatory variables, like age, education, and job-type. It is clear that the composite model (1), (2) and (7) is identified (just substitute (7) into (1) and (2) to get rid of ξ), and it is simple to derive consistent maximum-likelihood or two-stage least-squares estimators of the unknown parameters.

This is all quite simple, but it is indicative of where the developments went: when confronted with an equation containing an unobservable variable, the researcher may try to specify additional equations containing the same variable. In ZELLNER's approach this additional equation specifies how ξ is "caused." Likewise, one may specify equations in which ξ itself is the cause of a number of so-called indicator variables, each of which is observable and of the form

$$z = \xi\delta + w, \tag{8}$$

with w $(n \times 1)$ a vector of disturbance terms independently distributed of the elements of u and ϵ. Each z that can be written in this form may be used as an instrumental variable for estimating β. Both approaches can be integrated into the so-called MIMIC-model (Multiple indicators and multiple causes of a single latent variable). An overview is given by GOLDBERGER (1974). An extension to a situation where more than one unobservable occurs is obvious. All these types of models have found numerous empirical applications. (For example: CHAMBERLAIN and GRILICHES ((1975), (1977)), CHAMBERLAIN ((1977a), (1977b), (1978)), ATTFIELD (1977), AVERY (1979) and SINGLETON (1980).)

Although these models consist of multiple equations, they are not simultaneous models in a strict sense. A simultaneous model with measurement error in one or more exogenous variables,, may be identified without adding equations to the model. Depending on the specific structure of the model and on which variables are subject to measurement error, it may be possible to use overidentification of certain equations, if present, to identify measurement error. This point was noted by GOLDBERGER (1972) and has been elaborated by HSIAO (1976), GERACI ((1976), (1983)) and HAUSMAN (1977).

Both ideas, multiple equations and simultaneous equations, have been integrated in JÖRESKOG's LISREL (cf. JÖRESKOG and SÖRBOM (1981), BOOMSMA (1983)). LISREL is both a model and the registered trademark of a computer program which can be employed to estimate the model. The model consists of two parts, a "structural" part that is a standard simultaneous equations model with all variables unobservable, and a "measurement" part which relates observable variables linearly to the unobservable ones in the structural part. LISREL has played an important role in popularizing latent variable modelling in empirical work in econometrics.

4 Inequality constraints

The previous section was concerned with the additional information that can be obtained by embedding an equation in a system of relations. The extra information is then used to solve the identification problem. In many cases, there are no obvious relations that can be added to the single equation, so that the identification problem remains.

In the present section we report on some recent results for the case where in a single equation setting there is no exact additional information, but we have inexact information in that we know (or are willing to guess) upper bounds on the measurement error variances. The basic result is due to LEAMER (1982).

Consider the single equation model

$$y = \Xi\beta + \epsilon, \tag{9}$$

$$X = \Xi + V \tag{10}$$

where Ξ is an $(n \times k)$-matrix of latent variables, X is its measured counterpart and V is a matrix of measurement errors. Let v'_i be a typical row of V, then we assume that $v_i \sim N(0,\Omega)$, for all i, while v_i and v_j are independently distributed for $i \neq j$ $(i,j = 1, \ldots, n)$ and the measurement errors are independent of the elements of Ξ and ϵ. The stochastic assumptions on ϵ and Ξ are as before. An upper bound on the measurement error variance is formulated as

$$0 \leqslant \Omega \leqslant \Omega^* < A := \frac{1}{n} X'X \tag{11}$$

where Ω^* has to be specified by the researcher and the notation $\Omega \leqslant \Omega^*$ means that $\Omega^* - \Omega$ is positive semi-definite. Finally we define

$$\hat{\beta} := (A - \Omega)^{-1} A b \tag{12}$$

$$b := (X'X)^{-1} X'y = A^{-1} \frac{1}{n} X'y \tag{13}$$

$$b^* := (A - \Omega^*)^{-1} A b \tag{14}$$

$$F^* := (A - \Omega^*)^{-1} - A^{-1} \tag{15}$$

Under our assumptions $\hat{\beta}$ is the maximum likelihood estimator of β if Ω is known. Under a variety of different assumptions $\hat{\beta}$ will still be a consistent estimator for β. Let us, for simplicity, use the same notation for estimators and their realizations. It is clear that the estimate $\hat{\beta}$ one gets in any practical application will depend on the choice of Ω. It turns out that if we let Ω vary between zero and Ω^*, cf. (11), $\hat{\beta}$ varies over an ellipsoid. In particular, LEAMER (1982) proves the following result (under the condition $\Omega^* > 0$):

$$(\hat{\beta} - b^*)' F^{*-1} (\hat{\beta} - b) \leqslant 0. \tag{16}$$

Thus, although we cannot estimate β consistently, (16) provides the set of all consistent estimates (12) that we obtain if we were to pick Ω such that (11) holds true.

This bound is somewhat difficult to work with in practice, but it can be used to derive upper and lower bounds for any linear combination of $\hat{\beta}$. For example, the bounds for $\hat{\beta}_i$ are

$$\hat{\beta}_i(\max,\min) = \frac{1}{2}(b_i+b_i^*)\pm\frac{1}{2}\{(b^*-b)'AbF_{ii}^*\}^{\frac{1}{2}},$$

where subscripts denote the corresponding elements of vectors or matrices. LEAMER's result allows a researcher to translate his prior beliefs about the size of measurement error into bounds on regression coefficients.

LEAMER's result has been extended by BEKKER et al. (1984a). They allow for singular Ω^* (e.g., some variables may be measured without error). This extension is non-trivial as F^* now becomes singular. They also take a first step towards extending this result from the single equation model to the simultaneous equations case. It appears to be possible to derive an ellipsoid bound for the two-stage least squares estimator. This result was obtained under the stringent assumption that there are no exogenous variables in the equation concerned. What can be said when this assumption is relaxed remains an open question.

In a subsequent paper, BEKKER et al. (1984b) consider two more variations on LEAMER's result. First it is explored how the ellipsoid changes when also measurement error in y is considered *and* when this measurement error is allowed to correlate with the errors in X. It appears that the ellipsoid bound expands. The new bound is also an ellipsoid, with the same midpoint and the same principal axes as the ellipsoid for measurement error in X only. This result is remarkable, as usually the measurement error in y is indistinguishable from the disturbance term. Hence, the two play a different part.

The second result concerns measurement error that is uncorrelated between different regressors. So, Ω and Ω^* are diagonal. It is shown that β can now be bounded by a polyhedron with 2^k vertices, all lying on the ellipsoid; b and b^* are two of these vertices. So the bound is tighter, as it should be.

The advantage of the ellipsoid approach is that it permits the researcher to express prior ideas on measurement error in a simple form and next to translate these ideas into intervals for regression parameters. In particular, possible sign value changes due to variations in assumptions on measurement error can be detected.

5 Back to a result due to Koopmans

let us return for a moment to the simple model (1) and (2), and note that

$$p\lim(x'y)^{-1}y'y = \beta+\sigma_{uu}/(\sigma_{\xi\xi}\beta). \tag{17}$$

The left-hand side of (17) is the $p$lim of the inverse of the coefficient of the regression of x on y (the "reverse regression"). This regression also gives an inconsistent estimator of β, but with a bias away from zero. So (3) and (17) together can be taken as bounds between which a consistent estimate of β would lie. The bounds are obtained without making assumptions about the size of the measurement error.

Unfortunately, this classical result (FRISCH (1934)) does not carry over to the multiple regression case in general. Only under restrictive assumptions a generalization is possible. Here we present this generalization in the form of a theorem, due to KOOPMANS (1937). We provide a new proof of the theorem. After that we give an interpretation and a short discussion.

Theorem

Let Σ be a symmetric positive definite $(m \times m)$-matrix and let the elements σ^{ij} of Σ^{-1} be positive; let Φ be a diagonal $(m \times m)$-matrix and let the m-vector γ satisfy $(\Sigma - \Phi)\gamma = 0$.

(i) If $0 \leqslant \Phi \leqslant \Sigma$, then γ lies in the convex hull of Σ^{-1}.

(ii) For each γ in the convex hull of Σ^{-1} there exists one and only one Φ, such that $0 \leqslant \Phi \leqslant \Sigma$.

To prove this theorem we employ the following lemma.

Lemma

Let D be a symmetric $(m \times m)$-matrix with typical element d_{ij}, Λ a diagonal $(n \times m)$- matrix with i-th diagonal element λ_i and ι an m-vector of ones. If $d_{ij} > 0$ for all $i < j$, then

$$\mathrm{diag}(\Lambda D \Lambda \iota) \geqslant \Lambda D \Lambda \iff \lambda_i \lambda_j \geqslant 0 \quad \text{for all } i<j.$$

Proof

Let q be an m-vector. Then

$$\begin{aligned} q'\{\mathrm{diag}(\Lambda D \Lambda \iota) - \Lambda D \Lambda\} q &= \Sigma_i \Sigma_j d_{ij} \lambda_i \lambda_j (q_i^2 - q_i q_j) \\ &= \Sigma\Sigma_{i<j} d_{ij} \lambda_i \lambda_j (q_i - q_j)^2 \end{aligned}$$

(i) If $\lambda_i \lambda_j \geqslant 0$ for all $i<j$, then $d_{ij}\lambda_i\lambda_j(q_i - q_j)^2 \geqslant 0$ for all $i<j$.

(ii) If for some $i<j$ $\lambda_i\lambda_j < 0$, choose $q_i = sign(\lambda_i)$, so that $\Sigma\Sigma_{i<j} d_{ij}\lambda_i\lambda_j(q_i - q_j)^2 < 0$.
Q.E.D.

Proof of Theorem

Define: $\lambda := \Sigma\gamma = \Phi\gamma$; $\Lambda := diag(\lambda)$; $C := diag(\gamma)$. Thus, $\Lambda = \Phi C$ and $C = diag(\Sigma^{-1}\lambda) = diag(\Sigma^{-1}\Lambda\iota)$.

(i)

$$0 \leqslant \Phi \leqslant \Sigma \iff \qquad (18)$$

(see BEKKER et al. (1984a))

$$\Phi \geqslant \Phi\Sigma^{-1}\Phi \;\Rightarrow \tag{19}$$

$$C\Phi C \geqslant C\Phi\Sigma^{-1}\Phi C \;\Leftrightarrow \tag{20}$$

$$\Lambda C \geqslant \Lambda\Sigma^{-1}\Lambda \;\Leftrightarrow \tag{21}$$

$$diag(\Lambda\Sigma^{-1}\Lambda\iota) \geqslant \Lambda\Sigma^{-1}\Lambda \;\Leftrightarrow \tag{22}$$

(according to the lemma)

$$\lambda_i\lambda_j \geqslant 0 \;\text{ for all } i<j \;\Leftrightarrow \tag{23}$$

$$\gamma = \Sigma^{-1}\lambda \text{ lies in the convex hull of } \Sigma^{-1} \tag{24}$$

(ii) If γ lies in the convex hull of Σ^{-1}, then $\gamma_i > 0$ for all i, so C is non-singular, i.e. $\Phi = \Lambda C^{-1}$ is unique. Furthermore if C is non-singular, (19) and (20) are equivalent and hence (24) implies (18). Q.E.D.

Apart from KOOPMANS' proof, later proofs have been given by many authors, including KALMAN (1982) and KLEPPER and LEAMER (1984). These authors invoke the PERRON-FROBENIUS theorem and usually assume non-singularity of Φ, although KALMAN notes that this condition is not necessary.

From the proof the following new fact emerges: Let σ^{ij} be a typical element of Σ^{-1}. If $\sigma^{ij}<0$ for all $i<j$ (so that all elements of Σ are positive), then, using a similar proof as has been given for the lemma, $\lambda_i\lambda_j \geqslant 0$ for all $i<j$ would imply $diag(\Lambda\Sigma^{-1}\Lambda\iota) \leqslant \Lambda\Sigma^{-1}\Lambda$. As (18) implies that $diag(\Lambda\Sigma^{-1}\Lambda\iota) \geqslant \Lambda\Sigma^{-1}\Lambda$, it cannot be true that $\lambda_i\lambda_j \geqslant 0$ for all $i<j$ and $\lambda_i\lambda_j \neq 0$ for some $i \neq j$. In this case γ is not in the convex hull of Σ^{-1}.

To see what bearing the theorem has on the subject of errors-in-variables, we make the following choices for Σ, Φ and γ.

$$\Sigma := \begin{bmatrix} \frac{1}{n}y'y & \frac{1}{n}y'X \\ \frac{1}{n}X'y & A \end{bmatrix}, \quad \Phi := \begin{bmatrix} \sigma_{\epsilon\epsilon} & 0 \\ 0 & \Omega \end{bmatrix}, \quad \gamma := \begin{bmatrix} -1 \\ \hat{\beta} \end{bmatrix}. \tag{25}$$

If $A - \Omega$ is nonsingular then equation $(\Sigma - \Phi)\gamma = 0$ corresponds to (12). Furthermore, let e_i be an m-vector with a one in position i and zeros elsewhere. It is then easy to see that the vector $\Sigma^{-1}e_1$ is proportional to $(-1,b)'$, with b defined by (13). Similarly $(e_1, \ldots, e_{i-1}, e_{i+1}, \ldots, e_m)'\Sigma^{-1}e_i$ is proportional to the vector of regression coefficients obtained by regressing the i-th variable on all remaining variables.

So, if in (24) we take for λ the vectors $e_1, e_2, \ldots, e_m$, successively, we generate all possible regressions of one variable on all remaining variables. Thus, these regression vectors define the convex hull of Σ^{-1} and we may say that all *ML*-estimates γ are bounded by the hull of all possible regression vectors if all regression vectors are in the same orthant. In the case of a simple regression, this condition is trivially satisfied, cf. (3) and (17).

6 Conclusion

In many respects, the econometric approaches to errors-in-variables models mirror developments in other areas of econometric research. Gradually, the time-honored linear regression model becomes less prominent and is replaced by richer structures. These developments are triggered by a more critical attitude towards the quality of the data used and by the more stringent demands of economic theory. It is a sobering thought, however, that so much of the "progress" was already implicit in the work of the pioneers of econometrics, like T.C. KOOPMANS.

References

AIGNER, D.J., C. HSIAO, A. KAPTEYN, T.J. WANSBEEK (1984), Latent Variable Models in Econometrics, in: Z. GRILICHES and M.D. INTRILIGATOR (eds), *Handbook of Econometrics,* Vol. 2, North-Holland Publishing Co., Amsterdam.

ATTFIELD, C.L.F. (1977), Estimation of a Model Containing Unobservable Variables Using Grouped Observations: An Application to the Permanent Income Hypothesis, *Journal of Econometrics 6,* 51-63.

AVERY, R.B. (1979), Modelling Monetary Policy as an Unobserved Variable, *Journal of Econometrics 10,* 291-311.

BEKKER, P.A. (1985) Comment on: Identification in the Linear Errors in Variables Model, *Econometrica,* forthcoming.

BEKKER, P.A., A. KAPTEYN, T.J. WANSBEEK (1984a), Measurement Error and Endogeneity in Regression: Bounds for ML and IV Estimates, in: T.K. DIJKSTRA (ed.), *Misspecification Analysis,* Springer Verlag, Berlin.

BEKKER, P.A., A. KAPTEYN, T.J. WANSBEEK (1984b), Consistent Sets of Estimates for Regressions with Correlated of Uncorrelated Measurement Errors in Arbitrary Subsets of all Variables, Reeks "Ter Discussie", Tilburg University.

BOOMSMA, A. (1983), *On the Robustness of LISREL (Maximum Likelihood Estimation) against Small Sample Sizes and Non-Normality,* University of Groningen (Ph.D. dissertation).

CHAMBERLAIN, G. (1977A), EDUCATION, INCOME AND ABILITY REVISTED, CHAPTER 10 IN D.J. AIGNER AND A.S. GOLDBERGER (EDS), *Latent Variables in Socio-Economic Models,* (1977) North-Holland Publishing Co., Amsterdam.

CHAMBERLAIN, G. (1977b), An Instrumental Variables Interpretation of Identification in Variance-Components and MIMIC Models, Chapter 7 in P. TAUBMAN (ed.), *Kinometrics: The Determinants of Socio-Economic Success Within and Between Families,* North-Holland Publishing Co., Amsterdam.

CHAMBERLAIN, G. (1978), Omitted Variables Bias in Panel Data: Estimating the Returns to Schooling, *Annales de l'INSEE 30-31,* 49-82.

CHAMBERLAIN, G. and Z. GRILICHES (1975), Unobservables with a Variance-Components Structure: Ability, Schooling and the Economic Success of Brothers, *International Economic Review 16,* 422-429. Reprinted as Chapter 15 in D.J. AIGNER and A.S. GOLDBERGER (eds) (1977), *Latent Variables in Socio-Economic Models,* North-Holland Publishing Co., Amsterdam.

CHAMBERLAIN G. & Z. GRILICHES (1977), More on Brothers, Chapter 4 in P. TAUBMAN (ed.), *Kinometrics: The Determinants of Socio-Economic Success Within and Between Families,* North-Holland Publishing Co., Amsterdam.

FRISCH, R. (1934), Statistical Confluence Analysis by Means of Complete Regression Systems, Publication No. 5, University of Oslo, Economic Institute.

GERACI, V.J. (1976), Identifcation of Simultaneous Equation Models with Measurement Error, *Journal of Econometrics 4,* 263-283. Reprinted as Chapter 11 in D.J. AIGNER and A.S. GOLDBERGER (eds) (1977), *Latent Variables in Socio-Economic Models,* North-Holland Publishing Co., Amsterdam.

GERACI, V.J. (1983), Errors in Variables and the Individual Structural Equation, *International Economic Review 24,* 217-236.

GOLDBERGER, A.S. (1972), Structural Equations Methods in the Social Sciences, *Econometrica 40,* 979-1001.

GOLDBERGER, A.S. (1974), Unobservable Variables in Econometrics, in: P. ZAREMBKA (ed.), *Frontiers in Econometrics,* Academic Press, New York.

GRILICHES, Z (1974), Errors in Variables and Other Observables, *Econometrica 42,* 971-998. Reprinted as Chapter 1 in D.J. AIGNER and A.S. GOLDBERGER (eds), (1977), *Latent Variables in Socio-Economic Models,* North-Holland Publishing Co., Amsterdam.

GRILICHES, Z (1977), Estimating the Returns to Schooling: Some Econometric Problems, *Econometrica 45,* 1-22.

GRILICHES, Z. (1985), Economic Data Issues, in: Z. GRILICHES and M.D. INTRILIGATOR (eds), *Handbook of Econometrics,* Vol. 3, North-Holland Publishing Co., Amsterdam.

HAUSMAN, J.A. (1977), Errors in Variables in Simultaneous Equation Models, *Journal of Econometrics 5,* 389-401.

HSIAO, C. (1976), Identification and Estimation of Simultaneous Equation Models with Measurement Error, *International Economic Review,* 17, 319-339.

JÖRESKOG, K.G. and D. SÖRBOM (1981), *LISREL V User's Guide,* National Educational Resources, Chicago.

KALMAN, R.E. (1982), System Identification from Noisy Data, in: A. BEDNAREK (ed.), *Proc. International Symposium on Dynamical Systems* (Gainsville, FL February 1981), Academic Press, New York.

KAPTEYN and T.J. WANSBEEK (1983), Identification in the Linear Errors in Variables Model, *Econometrica 51,* 1847-1849.

KLEPPER, S. and E.E. LEAMER (1984), Consistent Sets of Estimates for Regressions with Errors in all Variables, *Econometrica 52,* 163-183.

KOOPMANS, T.C. (1937), *Linear Regression Analysis of Economic Time Series,* De Erven F. Bohn N.V., Haarlem.

LEAMER, E.E., (1982), Sets of Posterior Means with Bounded Variance Priors, *Econometrica 50,* 725-736.

PAKES, A. (1982), On the Asymptotic Bias of the Wald-Type Estimators of a Straight Line when Both Variables are Subject to Error, *International Economic Review 23,* 491-497.

PAL, M. (1980), Consistent Moment Estimators of Regression Coefficients in the Presence of Errors in Variables, *Journal of Econometrics 14,* 349-364.

REIERSÖL, O. (1950), Identifiability of a Linear Relation Between Variables which are Subject to Error, *Econometrica 18,* 375-382.

SINGLETON, K.J. (1980), A Latent Time Series Model of the Cyclical Behaviour of Interest Rates, *International Economic Review 21,* 559-576.

WALD A. (1940), The Fitting of Straight Lines if Both Variables are Subject to Error, *Annals of Mathematical Statistics 11,* 284-300.

WALD, A. (1949), Note on the Consistency of the Maximum Likelihood Estimate, *Annals of*

Mathematical Statistics 20, 595.

WHITE, H. (1984), *Asymptotic Theory For Econometricians*, Academic Press, Orlando.

ZELLNER, A. (1970), Estimation of Regression Relationships Containing Unobservable Independent Variables, *International Economic Review 11*, 441-454.

P.A. Bekker
A. Kapteyn
Dept. of Econometrics
Tilburg University
P.O. Box 90153
5000 LE Tilburg
The Netherlands

T.J. Wansbeek
University of Groningen
P.O. Box 800
9700 AV Groningen
The Netherlands