Consistent sets of estimates for regressions with correlated or uncorrelated measurement errors in arbitrary subsets of all variables
Bekker, P.; Kapteyn, A.J.; Wansbeek, T.J.

Publication date:
1984

Link to publication

Citation for published version (APA):
Bekker, P., Kapteyn, A. J., & Wansbeek, T. J. (1984). Consistent sets of estimates for regressions with correlated or uncorrelated measurement errors in arbitrary subsets of all variables. (Ter discussie FEW; Vol. 84.18). Unknown Publisher.
subfaculteit der econometrie

REEKS "TER DISCUSSIE"

Bestemming

<table>
<thead>
<tr>
<th></th>
<th>TIJDSSCHRIFTENBUREAU</th>
<th>BIBLIOTHEEK</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>KATHOLIEKE</td>
<td>HOGESCHOOL</td>
</tr>
<tr>
<td></td>
<td>HOGESCHOOL</td>
<td>TILBURG</td>
</tr>
</tbody>
</table>

Katholieke Hogeschool Tilburg
CONSISTENT SETS OF ESTIMATES FOR REGRESSIONS WITH CORRELATED OR UNCORRELATED MEASUREMENT ERRORS IN ARBITRARY SUBSETS OF ALL VARIABLES

Paul Bekker*)
Arie Kapteyn*)
Tom Wansbeek**)

*) Tilburg University. Financial Support by the Netherlands Organization for the Advancement of Pure Research (ZWO) is gratefully acknowledged.

**) Netherlands Central Bureau of Statistics.
Abstract

We consider the single equation errors-in-variables model and assume that a researcher is willing to specify an upper bound on the variance covariance matrix of measurement errors in the endogenous and exogenous variables. The measurement errors may show any pattern of correlations. It is shown that as a result the set of ML estimates is bounded by an ellipsoid. When, in addition, the variance covariance matrix of the errors is constrained to be diagonal, the set of ML estimates is shown to be bounded by the convex hull of \(2^\ell\) points (\(\ell\) being the number of error-ridden exogenous variables), lying on the surface of the ellipsoid. The results are applied to an empirical example and extensions to a simultaneous equations system are briefly discussed.
1. Introduction

Over the last decade the problem of measurement errors in the independent variables of a regression equation has attracted renewed interest among econometricians. In the fifties and sixties, the problem was considered to be more or less hopeless due to its inherent underidentification (e.g., Theil, 1971). Apart from instrumental variables, the most frequently cited textbook solution was Wald's method of grouping (Wald, 1940). Recent insight into the properties of the method of grouping can be interpreted as making this method worthless (Pakes, 1982). Since about 1970, new approaches to the problem have been explored, basically along three lines, viz. embedding the error-ridden equation into a set of multiple equations (e.g., Zellner, 1970, Goldberger, 1972), into a set of simultaneous equations (e.g., Hsiao, 1976, Geraci, 1976), and using the dynamics of the equation, if present (e.g., Maravall and Aigner, 1977). In view of the underidentification of the basic model, it is clear that all these methods invoke additional information of some kind. If this information takes the form of exact or stochastic knowledge of certain parameters in the model, the construction of consistent estimators is fairly straightforward (e.g. Fuller, 1980, Kapteyn and Wansbeek, 1984). For an overview of the state of the art, see Aigner et al. (1984).

An approach somewhat orthogonal to the ones described above has been to take the model as it is and to use prior ideas about the size of the measurement errors to diagnose how serious the problem is. Examples are Blomqvist (1972), Hodges and Moore (1972) and Davies and Hutton (1975). Leamer (1983) starts from the opposite direction by asking how serious the measurement error problem has to be in order to render the data useless for inference. In an empirical example, he shows that even very small measurement errors in some explanatory variables would open up the possibility of perfectly collinear explanatory variables and hence make the data useless for statistical inference (at least without additional prior information).

The most systematic analysis of the information loss caused by measurement error is due to Klepper and Leamer (1984). They start out by invoking a minimal amount of prior information and then ask the question under what conditions it is still possible to make some inferences regarding the vector of unknown regression parameters \( \beta \). In the special case where the measurement errors are assumed uncorrelated and the \( k+1 \) estimates of \( \beta \) obtained by re-
gressing each of the k+1 variables involved (i.e. the one dependent variable and the k independent variables) on the remaining k variables, are all in the same orthant, one can bound the ML estimates of β. In that case, the convex hull of the k+1 regressions contains all possible ML estimates and any point in the hull is a possible ML-estimate. If the k+1 regressions are not all in the same orthant then the set of ML estimates is unbounded.

In that case Klepper and Leamer (1984) introduce extra prior information which allows them to bound the set of maximum likelihood estimates. The prior information comes in two forms. Firstly, a researcher is supposed to be able to specify a maximum value of R² if all exogenous variables were measured accurately. It is shown that if this maximum is low enough, one can again bound the set of ML estimates by a convex hull. Secondly, if the R² bound does not help in bounding the estimates a researcher is assumed to be able to give upper and lower bounds for the measurement error variances. If the upper bound is tight enough, so that the true explanatory variables cannot be perfectly collinear, the set of maximum likelihood estimates is shown to be bounded by an ellipsoid. In the derivation of the ellipsoid, based on a result in Leamer (1982), it is assumed that all exogenous variables are measured with error. Obviously, this is restrictive.

Bekker, Kapteyn, Wansbeek (1984) have generalized Klepper and Leamer's result to the case where the variance covariance matrix of the measurement errors may be singular, but they still assumed, as did Klepper and Leamer, that the endogenous variable is measured without error or that the measurement error in the endogenous variables is uncorrelated with the errors in the exogenous variables. In this paper we relax this assumption, which turns out to be a non-trivial exercise. Apart from its intrinsic interest, the importance of the generalization also lies in the possibility to extend the analysis to more complicated models than just the linear regression model. Section 2 presents this result, after which Sections 3 and 4 are devoted to a discussion and a proof.

Although Klepper and Leamer (1984) assume throughout their paper that all measurement errors are uncorrelated, they do not exploit that information in the derivation of the ellipsoid. For any point in the ellipsoid we can find an \( \Omega \) (the variance covariance matrix of the errors in the explanatory variables) that yields this point as an ML estimate, but such an \( \Omega \) need not be diagonal. In Section 5 we investigate the consequences of the extra requirement that \( \Omega \) is diagonal. In that case the ML estimates are bounded by a po-
lyhedron, which need not be convex. Of course, the polyhedron lies within the ellipsoid. The convex hull of the polyhedron is determined by \(2^\ell\) vertex points that all lie on the ellipsoid, where \(\ell\) is the number of nonzero measurement error variances. These points can be computed easily and then used to find, for all elements of \(\beta\), intervals that bound the ML estimates. Generally, these intervals are tighter than the ones obtained from the ellipsoid.

In Section 6, an empirical example illustrates how the various types of prior restrictions affect the bounds on the ML estimates. Section 7 concludes by briefly discussing extensions to simultaneous equations models.
2. The Model and the Ellipsoid

Throughout we deal with the following model:

\[ \eta = \Xi \beta_0 + \epsilon \]  
\[ y = \eta + u \]  
\[ X = \Xi + V \; ; \]  

(2.1)
(2.2)
(2.3)

(2.1) is the classical linear model, which relates the n-vector of dependent variables \( \eta \) to the nxk-matrix of explanatory variables \( \Xi \) and the n-vector of disturbances \( \epsilon \). We assume that the distribution of \( \epsilon \) is independent of \( \Xi \) and satisfies \( E\epsilon = 0, \; E\epsilon \epsilon' = \sigma_0^2 I \). The k-vector of parameters \( \beta_0 \) and \( \sigma_0^2 \) are unknown and have to be estimated.

Both \( \eta \) and \( \Xi \) are unobservable. Instead, \( y \) and \( X \) are observed and \( u \) and \( V \) therefore are the errors of measurement in \( y \) and \( X \). We assume that \( u \) and \( V \) are distributed independently of \( \Xi, \eta \) and \( \epsilon \) and that \( Eu = 0, EV = 0 \). Moreover, letting \( u_i \) be the i-th element of \( u \) and \( v_i' \) the i-th row of \( V \), we assume that

\[ \mathbb{E} \left[ \begin{array}{c} u_i \\ v_i' \end{array} \right] = \phi \equiv \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \Omega \end{bmatrix} \]

for all \( i \) and that \( (u_i', v_i') \) is stochastically independent of \( (u_j', v_j') \) for \( i \neq j \). Let \( \phi \) be known and define \( \beta \) and \( \sigma^2 \) by

\[ \beta \equiv (A-\Omega)^{-1}(Ab-\phi_{21}) \]  
\[ \sigma^2 \equiv \frac{1}{n} y'y - \phi_{11} - \beta'(A-\Omega)\beta \; , \]  

(2.4)
(2.5)

where \( A \equiv \frac{1}{n} X'X \), \( b \equiv (X'X)^{-1}X'y \). Under a variety of assumptions, \( (\beta, \sigma^2) \) will be a consistent estimator of \( (\beta_0, \sigma_0^2) \). For example, if \( \Xi, \epsilon, u, V \) are jointly normally distributed and \( E\epsilon = 0 \), \( (\beta, \sigma^2) \) is the ML-estimator of \( (\beta_0, \sigma_0^2) \). If \( \epsilon, u \) and \( V \) are jointly normally distributed and \( \Xi \) is non-stochastic with \( \lim \frac{1}{n} \Xi'\Xi = K \), say, then \( (\beta, \sigma^2) \) is not the ML-estimator but still consistent.
Of course, if $\phi = 0$, $(\beta, \sigma^2)$ reduces to the OLS-estimator $(\hat{b}, \hat{s}^2)$, where $s^2 \equiv \frac{1}{n} y'y - b'Ab$.

If $\phi$ is unknown, the model is not identified. Sometimes identifying information comes from extraneous sources, like instrumental variables, but the more common "solution" of the identification problem is to set $\phi$ equal to zero and hence estimate $(\hat{\beta}_0, \hat{\sigma}^2_0)$ by OLS. If, in reality, $\phi \neq 0$, this estimator will be inconsistent.

Although $\phi$ will usually be unknown, it seems reasonable to assume that a researcher will be able to specify bounds for $\phi$, i.e.,

$$0 \leq \phi \leq \phi^* \equiv \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix},$$

(2.6)

where $\phi^*$ is specified by the researcher.\(^1\) This bound on $\phi$ will be used to derive bounds on the estimates $\beta$ defined by (2.4).\(^2\) We assume that $\phi^*$ is symmetric and that

$$0 \leq \phi^* < B \equiv \frac{1}{n} \begin{bmatrix} y'y & y'X \\ X'y & X'X \end{bmatrix},$$

(2.7)

thereby guaranteeing the existence of the estimate $\beta$ and also the positiveness of the estimate $\sigma^2$ for any choice of $\phi$ satisfying (2.6). The latter can be shown easily by writing the positive definite matrix $(B-\phi)^{-1}$ as

$$(B-\phi)^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & (A-\Omega)^{-1} \end{bmatrix} + \sigma^{-2}(-1, \beta^t)(-1, \beta^t)^{-1},$$

(2.8)

so that

$$\sigma^2 = |e_1'(B-\phi)^{-1}e_1|^{-1} > 0,$$

(2.9)

1) The notation $C \preceq D$ means that $D-C$ is a positive semidefinite matrix; $C \prec D$ means $D-C$ is positive definite.

2) For simplicity, no notational difference is made between random variables and their realization. From now on $y$, $X$, $\sigma^2$ and $\beta$ refer to realizations of random variables. Hence $\sigma^2$ and $\beta$ are estimates rather then estimators.
where $e_1$ is the first unit vector. Furthermore, if we denote the estimate $(\beta, \sigma^2)$ by $(b^*, s^2)$ if $\phi = \Phi^*$, it is readily established that, as a consequence of the boundedness of $\phi$, also $\sigma^2$ is bounded:

$$s^2 > \sigma^2 > s^* > 0.$$  

(2.10)

We may now ask the question whether we can also delimit the set of estimates $\beta$ given that $\phi$ satisfies (2.6). To answer that question the following lemma is useful.

Lemma 1: Let $C$ and $C^*$ be symmetric matrices and let $C^-$ be an arbitrary $g$-inverse of $C^*$. Then

$$0 \leq C \leq C^*$$  

(2.11)

is equivalent with the following three conditions jointly:

(i) $\quad C^* \geq 0$

(ii) $\quad C^* = C$

(iii) $\quad C \geq C^* C^*$

(2.12)

Proof: The lemma has been proven by Bekker, Kapteyn and Wansbeek (1984) for the case where $C^*$ is the Moore-Penrose inverse of $C^*$. Obviously, (ii) is invariant under the choice of the $g$-inverse as it simply states that $C$ lies in the space spanned by $C^*$. Given (ii) it follows immediately from Lemma 2.2.4 in Rao and Mitra (1971) that $C^* C^*$ is invariant under the choice of $C^*$. Q.E.D.

Corollary: Let $x$ be a vector and $C^*$ a symmetric matrix, then

$$0 \leq xx' \leq C^*$$  

(2.13)

is equivalent with the following three conditions jointly:
(i) \( C^* \geq 0 \)

(ii) \( C^* x = x \) \hspace{1cm} (2.14)

(iii) \( x'C^* x \leq 1 \)

Proof: For the case that \( C^* \) is the Moore-Penrose inverse of \( C^* \), Bekker, Kapteyn and Wansbeek (1984) have given a proof. Obviously (i)-(iii) are invariant under the choice of the g-inverse. Q.E.D.

Define

\[
F^* = (A^{-1} - A^*)^{-1} - A^{-1}
\] \hspace{1cm} (2.15)

Using the lemma and its corollary, Sections 3 and 4 will be devoted to the proof of the following proposition.

Proposition 1: The set of solutions \( \beta \) satisfying (2.4), with \( \phi \) satisfying (2.6), is given by:

(i) \( (\beta - \frac{1}{2}(b+b^*))'F^* (\beta - \frac{1}{2}(b+b^*)) \leq \frac{1}{4}(s^2 - s^2) \) \hspace{1cm} (2.16)

(ii) \( F^* (\beta - \frac{1}{2}(b+b^*)) = \beta - \frac{1}{2}(b+b^*) \). \hspace{1cm} (2.17)

This bound is minimal, i.e., for any \( \beta \) satisfying (2.16) and (2.17) there exists a \( \phi \) such that (2.4) and (2.6) hold true.

Although a more extensive interpretation of this result will be given in the sections to come, it is worth noticing at this place that (2.16) and (2.17) represent an ellipsoid in the space spanned by \( F^* \).

If the number of regressors exceeds two, it will be hard in practice to represent the ellipsoid given by (2.16) and (2.17) in a transparent way. For that reason it is useful to derive bounds for linear functions of \( \beta \). Let \( \psi \) be a known vector, then bounds for \( \psi'\beta \) are implied by the following proposition.
Proposition 2: The maximum and minimum of \( \psi'\beta \), with \( \psi \) fixed and \( \beta \) satisfying (2.16) and (2.17), are given by

\[
\psi'\beta = \frac{1}{2} \psi'(b+b^*) \pm \frac{1}{2} \sqrt{(s^2-s^*)^2}, \psi'F\psi.
\] (2.18)

If \( \psi'F^* \neq 0 \), the corresponding vectors \( \hat{\beta} \) are given by

\[
\hat{\beta} = \frac{1}{2}(b+b^*) \pm \frac{1}{2} F^* \sqrt{\frac{(s^2-s^*)^2}{\psi'F\psi}}.
\] (2.19)

Proof: In view of the corollary, (2.16) and (2.17) are equivalent with

\[
[\beta - \frac{1}{2}(b+b^*)][\beta - \frac{1}{2}(b+b^*)]' \leq \frac{1}{4}(s^2-s^*)^2 F^*.
\] (2.20)

This implies

\[
[\psi'\beta - \frac{1}{2}\psi'(b+b^*)]^2 \leq \frac{1}{4}(s^2-s^*)^2 \psi'F\psi,
\] (2.21)

for any given vector \( \psi \). This makes it clear that (2.18) gives the extreme values of \( \psi'\beta \) and (2.19) gives the corresponding values of \( \hat{\beta} \). Q.E.D.

For the proof of Proposition 1, it turns out to be convenient to distinguish two cases that differ with respect to the structure of \( \phi^* \) that is assumed. In Section 3 it is assumed that there exists a vector \( \lambda^* \) such that \((-1,\lambda^*) \phi^* = 0\), i.e. the first row of \( \phi^* \) is linearly dependent on the other rows of \( \phi^* \). In that case the model can be reformulated in such a way that only the regressors are subject to measurement error. In Section 4 it is assumed that there is no such vector \( \lambda^* \), i.e. the first row of \( \phi^* \) is linearly independent of the other rows of \( \phi^* \).
3. Case I: No Measurement Error in the Regressand or \( u_1 = v_1 \lambda^* \).

We first give a statistical interpretation of this case. Suppose there exists a vector \( \lambda^* \) such that

\[
(-1, \lambda^*)(\Phi^* = 0, \tag{3.1}
\]

then

\[
0 \leq (-1, \lambda^*) \Phi (-1)^{-1} \leq (-1, \lambda^*) \Phi (-1)^{-1} = 0.
\]

Consequently, also \((-1, \lambda^*) \Phi = 0\), which is equivalent with \( u_1 = v_1 \lambda^* \), the measurement error in \( y \) is a fixed linear combination of the measurement errors in \( X \). Clearly, there is no measurement error in \( y \) if \( \lambda^* = 0 \), but even if \( \lambda^* \neq 0 \), the model (2.1)-(2.3) may be reformulated as

\[
y - X \lambda^* = \Xi (\beta_0 - \lambda^*) + \varepsilon \tag{3.2}
\]

\[
X = \Xi + V, \tag{3.3}
\]

in which the observable regressand \( y - X \lambda^* \) is no longer subject to measurement error.

Proof of Proposition 1 (Case I): Note that in this case the consistent estimate \( \beta \) is given by

\[
\beta = (A - \Omega)^{-1} (Ab - \Omega \lambda^*). \tag{3.4}
\]

Let

\[
F = (A - \Omega)^{-1} A^{-1} = (A - \Omega)^{-1} \Omega A^{-1},
\]

So that \( \Omega = A - (A^{-1} + F)^{-1} \). Then (3.4) implies

\[
\beta - b = FA(b - \lambda^*), \tag{3.5}
\]

so that

\[
b^* - b = FA(b - \lambda^*) \tag{3.6}
\]

and
\[ \beta - b^* = (F^*F^*)A(b^*-\lambda^*). \]  
(3.7)

In model (3.2)-(3.3) the matrix \( \Omega \) is bounded: \( 0 \leq \Omega \leq \Omega^* \). Equivalently, the matrix \( F \) is bounded:

\[ 0 \leq F \leq F^*, \]  
(3.8)

where \( F^* \) is given in (2.15). Lemma 1 implies that

\[ (F-F^*)F^*F \leq 0, \quad \text{so} \quad (b^*-\lambda^*)'A(F-F^*)F^*F A(b^*-\lambda^*) \leq 0 \]  
(3.9)

and

\[ F^*F^*F = F, \quad \text{so} \quad F^*F^*F A(b^*-\lambda^*) = FA(b^*-\lambda^*). \]  
(3.10)

Consequently, the bounds on the set of solutions \( \beta \) satisfying (3.8), with \( F \) satisfying (3.8), can be found by substituting (3.5) and (3.7) in (3.9) and (3.10). We obtain:

\[ (\beta - b^*)'F^*F (\beta - b) \leq 0 \]  
(3.11)

\[ F^*F^*F (\beta - b) = \beta - b. \]  
(3.12)

This bound (3.11)-(3.12) is minimal: for each \( \beta \) satisfying this bound, we can find a matrix \( F \) (hence an \( \Omega \), hence a \( \Phi \)) for which both (3.5) and (3.8) hold true. If \( \beta = b \), we choose \( F = 0 \); if \( \beta \neq b \) we choose for \( F \):

\[ \tilde{F} = \frac{(\beta - b)(\beta - b)'}{\beta - b)'A(b^*-\lambda^*), \]  
(3.13)

and we will show that \( 0 \leq \tilde{F} \leq F^* \). First note that

\[ (\beta - b)'A(b^*-\lambda^*) = (\beta - b)'F^*F A(b^*-\lambda^*) = (\beta - b)'F^* (b^*-b) \geq (\beta - b)'F^* (\beta - b) > 0 \]  
(3.14)

(The first equality is based on (3.12), the second one on (3.6); the first inequality sign is based on (3.11).) So,
Since \( F^* \geq 0 \), and (3.12) hold, we may use the corollary of Lemma 1 to arrive at

\[
0 \leq \frac{(\beta-b)'(\beta-b)'}{(\beta-b)'A(b-\lambda')} \leq F^*
\]

(3.16)

Therefore, the bound is minimal. \(^{1)}\)

If condition (3.1) is satisfied, the consistent estimate \( \sigma^2 \), (2.5), of the error variance in the equation may be rewritten as follows. From (3.4),

\[
\Omega^* = A^*(\Omega - A^{-1}B), \text{ so}
\]

\[
\Omega^* = \lambda^*\Omega^* = \lambda^*A(b-B) + \lambda^*\Omega
\]

(3.17)

and, also from (3.4),

\[
\beta'(A^{-1}B) = \beta'Ab - \lambda^*\Omega
\]

(3.18)

We then have

\[
\sigma^2 = \frac{1}{n} y'y - \beta'(A^{-1}B) =
\]

\[
= \frac{1}{n} y'y - \lambda^*A(b-B) - \beta'Ab =
\]

\[
= \frac{1}{n} y'y - b'Ab - (\beta-b)'A(b-\lambda^*)
\]

\[
= s^2 - (\beta-b)'F^*(b-b)
\]

(3.19)

The last step is based on the definition of \( s^2 \) given between (2.5) and (2.6), and on (3.14). This shows that to each estimate \( \beta \) there corresponds only one estimate \( \sigma^2 \). By using (3.19) the translation of (3.11)-(3.12) into the minimal bound given in Proposition 1 is straightforward. In (3.19) take \( \beta = b^* \), so that \( \sigma^2 \) becomes \( s^2 \) and hence

\(^1\) Note that, to prove the minimality of the bound, we have constructed an F of rank 1. Of course, each \( \beta \) in the ellipsoid can in general be produced by an infinity of \( F \)'s of higher rank, but that is of no relevance here.
\[
\frac{1}{4}(b^*-b)'F^*(b^*-b) = \frac{1}{4}(\sigma^2-s^*^2) \tag{3.20}
\]

Add this to (3.11) and rearrange terms. \( \text{Q.E.D.} \)

Thus, when \( \Phi \) is bounded and condition (3.1) is satisfied, the consistent estimate \( \beta \) must lie on, or within the ellipsoid (2.16)-(2.17). Conversely, any point on or within this ellipsoid is a consistent estimate of \( \beta_0 \) for some \( \Phi \) satisfying (2.6). Both \( b \) and \( b^* \) lie on the surface of the ellipsoid and the centre of the ellipsoid is located at the midpoint of the line segment joining \( b \) and \( b^* \). See Figure 1. For the special case that \( \Omega^* \) is nonsingular, i.e. all regressors are subject to measurement error, and \( \lambda^* = 0 \), i.e. there is no measurement error in the regressand, this result was found before by Klepper and Leamer (1984), using a result obtained in Leamer (1982).\(^1\)

![Figure 1: The ellipsoid for Case I](image)

\(^1\) The result of Klepper and Leamer is seemingly of greater generality in a different respect, as they also consider the possibility of a non-zero lower bound on \( \Omega \). However, such a lower bound can simply be subtracted from \( A \) and for the rest the analysis is unaffected.
4. Case II: Measurement Error in the Regressand, and \( u_1 \neq v_1^t \lambda^* \)

The derivation of the ellipsoid, under the condition that there exists no \( \lambda^* \) such that (3.1) holds, rests on two steps. First we derive a "contaminated" bound for the vector \( \gamma \) defined by

\[
\gamma = \sigma^{-2} \beta^{-1} = -(B - \Phi)^{-1} e_1 ,
\]

where we used (2.8). In the second step we derive from this contaminated bound for \( \gamma \) a minimal bound for \( \beta \).

4.1. A contaminated bound

Analogous to (4.1) we define

\[
c \equiv s^{-2} \beta^{-1} = -B^{-1} e_1 \quad \text{and} \quad c^* \equiv s^{-2} \beta^{-*} = -(B - \Phi^*)^{-1} e_1
\]

so that

\[
\gamma - c = -He_1 ,
\]

where \( H \equiv (B - \Phi)^{-1} - B^{-1} = B^{-1} \Phi (B - \Phi)^{-1} \) or \( \Phi = B - (B^{-1} + H)^{-1} \). Equivalent to the bounds for \( \Phi \), we have

\[
0 \leq H \leq H^* ,
\]

where \( H^* \equiv (B - \Phi^*)^{-1} - B^{-1} \). Thus, analogous to the derivation of (3.11)-(3.12), Lemma 1 implies

\[
(\gamma - c)^* H^* (\gamma - c) \leq 0
\]

\[
H^* H^* (\gamma - c) = \gamma - c .
\]

Again, this contaminated bound is minimal; for \( \gamma = c \) we take \( H = 0 \) and for each \( \gamma \neq c \) satisfying (4.5)-(4.6), we may take
\[
\tilde{H} = -\frac{(y-c)(y-c)'}{(y-c)'e_1},
\]

which satisfies both (4.3) and (4.4).

4.2. The minimal bound for \( \beta \)

In order to derive a minimal bound for \( \beta \), we reason as follows. For a certain \( \beta \) there exists a \( \tilde{\phi} \) satisfying both (2.4) and (2.6) if and only if there exists an arbitrary \( \tilde{\sigma}^{-2} \) such that \( \tilde{\gamma}' \equiv \tilde{\sigma}^{-2}(-1,\beta') \) satisfies (4.5)-(4.6). For only then there exists a \( \tilde{\phi} \) which generates \( \tilde{\gamma} \), and thus a \( \tilde{\phi} \) which generates \( \beta \).

If there exists no \( \lambda^* \) such that (3.1) holds, then the first row of \( \lambda^* \) is linearly independent of the other rows of \( \lambda^* \), and hence there exists a \( \mu^* \) such that \( e_1 = \lambda^* \mu^* \), or

\[
c = -B^{-1}e_1 = -B^{-1}\lambda^*(B-\lambda^*)^{-1}(B-\lambda^*)\mu^* = -H(B-\lambda^*)\mu^*.
\]

Hence,

\[
H^*Hc = -H^*H^{-1}(B-\lambda^*)\mu^* = -H(B-\lambda^*)\mu^* = c,
\]

and if we substitute \( \tilde{\sigma}^{-2}(-1,\beta') \) for \( \gamma' \), (4.6) translates into

\[
H^*H^{-1} = (-1)_{\beta}.
\]

For any choice of \( \beta \) satisfying (4.10), the bound (4.5) becomes a quadratic inequality in \( \tilde{\sigma}^{-2} \). A necessary and sufficient condition for this inequality to hold for at least one \( \tilde{\sigma}^{-2} \) is that the quadratic expression has a non-negative discriminant, i.e.

\[
\{(\beta_1^{-1})H^*c + \frac{1}{2}(c_1)_c^*\}^2 - \{(\beta_1^{-1})H^*^{-1}\}\{c^*H^*c\} \geq 0,
\]

and it remains to rewrite (4.10) and (4.11) as an ellipsoid in \( \beta \). We first note that \( c^*H^*c > 0 \): from (4.3), \( c^*-c = -H^*e_1 \), or \( c^* = c-H^*e_1 \), so using (4.9) we have
\[ c'Hc = c'Hc - e_1'c = c'Hc - e_1'H'Hc = c'Hc + s^{-2} > 0. \] (4.12)

It is convenient in the sequel to have a different expression for \( c'Hc \).
Substitute \( c = c' + He_1 \) to obtain

\[ c'Hc = c'Hc + c'H'e_1 = c'Hc + c'e_1, \] (4.13)

where the last step is based on \( H'Hc = c \), which follows from (4.6) (with \( \gamma = c' \)) and (4.9). From (4.12) and (4.13) it follows that

\[ c'Hc = \frac{1}{4}(c+c)'H(c+c) + \frac{1}{4}(c-c)'e_1. \] (4.14)

So when we define

\[ G = H' - \frac{H'(c+c')(c+c)'H}{(c-c)'e_1 + (c+c)'H(c+c')}, \] (4.15)

the minimal bound for \( \beta \), (4.10)-(4.11), can be written as

\[ (\beta^{-1})' G(\beta^{-1}) \leq 0, \text{ and } H'H(\beta^{-1}) = (\beta^{-1}). \] (4.16)

We complete the derivation by replacing \( G \) by a more convenient matrix which leaves the first part of (4.16) unaffected. We define

\[ M = H' + \frac{(c+c')(c+c)'}{(c-c)'e_1}, \] (4.17)

which always exists since \( (c-c)'e_1 < 0 \). As \( H'H(c+c') = c + c' \), it is easily verified that \( MG = H'H \) and that \( G \) is a g-inverse of \( M \). Thus (4.16) may be written as:

\[ (\beta^{-1})' M(\beta^{-1}) \leq 0, \text{ and } MM(\beta^{-1}) = (\beta^{-1}). \] (4.18)

It follows from Lemma 2.2.4 in Rao and Mitra (1971) that (4.18) is invariant under the choice of \( M^{-1} \). In order to take a more convenient g-inverse of \( M \) than
G, we proceed as follows. When we evaluate (2.8) at \( \phi = 0 \), and substitute the definition of \( c \) from (4.2), we obtain

\[
B^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & A^{-1} \end{bmatrix} + s^2 c c'.
\]  
(4.19)

When we evaluate (2.8) at \( \phi = \phi^* \) substitute the definition of \( c^* \) from (4.2), and use \( (A-\Omega)^{-1} = F^*A^{-1} \), we obtain

\[
(B-\Phi^*)^{-1} = \begin{bmatrix} 0 & 0 \\ 0 & F^*A^{-1} \end{bmatrix} + s^2 c c'.
\]  
(4.20)

So

\[
H^* = (B-\Phi^*)^{-1} - B^{-1} =
\begin{bmatrix} 0 & 0 \\ 0 & F^* \end{bmatrix} + s^2 c c' - s^2 c c'.
\]  
(4.21)

Next, substitute \( H^* \) into (4.17) and use \((c - c)^e_1 = s^2 - s^* \) to obtain

\[
M = \begin{bmatrix} 0 & 0 \\ 0 & F^* \end{bmatrix} + \frac{(-1)}{(b^* + b)^2} \frac{1}{4} (s^2 - s^*)
\]  
(4.22)

Applying the well-known formula for the generalized inverse of a partitioned symmetric matrix yields

\[
M = \begin{bmatrix} s^2 - s^2 \frac{2}{4} + (b^* + b)^2, F^* - (b^* + b)^2, (b^* + b)^2, F^* \\ F^* & \frac{b^* + b}{2} \end{bmatrix}
\]  
(4.23)

as a g-inverse of \( M \). There holds

\[
MM = \begin{bmatrix} 1 & 0 \\ -(I-F^*F^* - (b^* + b)^2, F^*F^* \end{bmatrix}
\]  
(4.24)
Thus, we may substitute $M^*$ for $M$ in (4.18), which results in the following minimal bound for $\beta$:

$$\frac{(\beta - \left(\frac{b+b^*}{2}\right))^2 F^* (\beta - \left(\frac{b+b^*}{2}\right))}{2} \leq \frac{s^2 - s'^2}{4}$$

(4.25)

and

$$F^* F^* (\beta - \left(\frac{b+b^*}{2}\right)) = \beta - \left(\frac{b+b^*}{2}\right)$$

(4.26)

which is exactly the bound given in Proposition 1. Q.E.D.

It is worth noticing that in Case I, where condition (3.1) does hold, (4.6) can be used to find an expression for $\sigma^2$ in terms of $\beta$:

$$\sigma^2 = \frac{c' (I - \hat{H} H^*)^{-1} \beta}{c' (I - \hat{H} H^*)c}$$

(4.27)

showing again that in case I to each $\beta$ there corresponds only one $\sigma^2$. This expression, (4.27), can then be substituted in (4.5) and (4.6) to yield a minimal bound for $\beta$. However, the resulting expressions are rather complicated and the derivation given in Section 3 is much simpler. In Case II, (4.27) is not defined as the denominator is equal to zero, by virtue of (4.9).

The centre of the ellipsoid (4.22)-(4.23) is again the midpoint of the line segment joining $b$ and $b^*$; see Figure 2. However, contrary to the ellipsoid which was derived in Section 3, where the first row of $\phi^*$ was linearly dependent on the other rows of $\phi^*$, the points $b$ and $b^*$ lie within the ellipsoid if the first row of $\phi^*$ is linearly independent of the other rows of $\phi^*$.

An illuminating way of looking at this phenomenon is the following one. Choose a parameter vector $\lambda$ such that

$$u_1 = \nu_1' \lambda + \tilde{u}_1,$$

(4.28)

where $\tilde{u}_1$ has zero mean and is uncorrelated with $\nu_1$. Therefore we may write:

$$\phi^* = \begin{bmatrix} 
\lambda' & \tilde{\phi} & \lambda' \\
\lambda & \Omega & \lambda' \\
\Omega & \lambda & \Omega
\end{bmatrix},$$

(4.29)
where $\tilde{\phi}_{11}^* \equiv \phi_{11}^* - \lambda^* \Omega^* \lambda^*$. In Section 3 we had the condition that $\phi_{11}^* = 0$, and in this section we have that $\phi_{11}^* > 0$. Now consider the alternative upper bound

$$
\tilde{\phi}^* = \phi^* - \begin{bmatrix} \phi_{11}^* & 0 \\ 0 & 0 \end{bmatrix}
$$

Clearly, $0 \leq \tilde{\phi}^* < \phi^*$ if $\phi_{11}^* > 0$. If we now consider the minimal bound for $\beta$ if $0 \leq \tilde{\phi} \leq \phi^*$, then $F^*$, which is only a function of $\Omega^*$, is unchanged; also $b$ and $b^*$ are unchanged. The only change is in $s^*$, the estimate of the error variance in the equation:

$$
s^* = s^2 + \phi_{11}^*
$$

which is a reflection of the fact that $\tilde{u}_1$ is indistinguishable from $e_1$, the error in the equation.\(^1\)

As has been shown in the previous section, $b$ and $b^*$ lie on the surface of the ellipsoid if $0 \leq \phi \leq \phi^*$. Consequently, $b$ and $b^*$ lie within the ellipsoid if $\phi_{11}^* > 0$, and $0 \leq \phi \leq \phi^*$. The larger $\phi_{11}^*$ becomes, the more the ellipsoid expands. The intuitive explanation for this is that if $\phi_{11}^*$ increases, we do not only allow more measurement error in $y$ (which is indistinguishable from errors in the equation anyway) but also more covariance between the measurement errors in $y$ and $X$. Thus, the bound on $\phi$ becomes less tight and the ellipsoid expands.

Figure 2: The ellipsoid for Case II

1) Remember from (2.10) that $s^2 \geq \sigma^2 > s^2 > 0$. If we replace $s^2$ by $s^2$ we increase the lower bound on $\sigma^2$. That is, a larger share of the random component in $y$ is ascribed to errors in the equations and less to measurement error.
5. Uncorrelated measurement errors

In this section we assume that, in addition to the bounds on $\phi$ as given in (2.6), a researcher is also willing to assume that $\phi^*$ and $\phi$ are diagonal. That is, measurement errors in different variables are uncorrelated.

The first thing to notice is that in this case the measurement error in the regressand is completely indistinguishable from the error in the equation. Therefore it is of no consequence for the set of estimates $\beta$. Since $\phi$ is diagonal, $\phi_{21} = 0$ and the estimator $\beta$ is simply given by

$$\beta = (A-\Omega)^{-1}Ab,$$

(5.1)

where $\Omega$ is diagonal and bounded by

$$0 \leq \Omega \leq \Omega^* < A$$

(5.2)

Clearly, the set of estimates is unchanged if we choose $\phi^*_{11} = \phi^*_{11} = 0$. Consequently the set of estimates $\beta$ is bounded by the ellipsoid (3.11)-(3.12), which we will refer to as "the ellipsoid spawned by $\Omega^*$". Of course, this ellipsoid is no longer a minimal bound if $\Omega$ and $\Omega^*$ are restricted to be diagonal.

In order to derive a more satisfactory bound we define the following points

$$\beta^*_\delta = (A-\Omega^*_\delta)^{-1}Ab,$$

(5.3)

where $\Omega^*_\delta = \Delta^* \Delta = \Delta^* \Delta^*$, with $\Delta = \text{diag}(\delta)$ and $\delta$ a vector with units and zeros as elements. If $\Omega^*$ has $\ell$ non-zero diagonal elements then there are $2^\ell$ different matrices $\Omega^*_\delta$, which all satisfy (5.2). Clearly the $2^\ell$ solutions $\beta^*_\delta$ are bounded by the ellipsoid spawned by $\Omega^*$.

Proposition 3: All $\beta^*_\delta$ lie on the surface of the ellipsoid spawned by $\Omega^*$

Proof: Clearly $\Omega^*_\delta = \Delta^* \Omega^* \Delta^*$, and $\Omega^* = \Delta^* \Omega^* \Delta^*$, so that

$$(\Omega^*_\delta^* - \Omega^*)\Omega^* \Delta^* = 0$$

(5.4)

Thus
\[(A-\Omega^*)\{(A-\Omega^*)^{-1} - (A-\Omega^*)^{-1}\}[(A-\Omega^*)\Sigma^{-1}B]A\{A^{-1}\Omega^*(A-\Omega^*)^{-1}\}(A-\Omega^*) = 0,\]
or
\[(A-\Omega^*)\{(F^*_{\delta}F^*)^{-1}F^*_{\delta}(A-\Omega^*) = 0\]

with \(F^*_{\delta}\) defined implicitly. The expression simplifies to
\[(F^*_{\delta}F^*)^{-1}F^*_{\delta} = 0.\]  (5.5)

Hence, the inequalities (3.9) and (3.11) become equalities if \(F = F^*_{\delta}\), or \(\Omega = \Omega^*\). Q.E.D.

We now turn to the main result of this section. Having established that all \(\beta_{\delta}\) lie on the surface of the ellipsoid spawned by \(\Omega^*\), we now show that \(\beta\) lies in the convex hull of the \(L^2\) points \(\beta_{\delta}\) that are generated by \(\Omega^*\) ("in the convex hull of \(\Omega^*\), for short). We use the following notation and conventions. Let \(\Omega^*\) have diagonal elements \(w_i^*\), let \(W^*_i = w_i^*e_i^*\) (\(e_i^*\) being the \(i\)-th unit vector), let \(\Omega^*_m = \sum_{i=1}^{m} W^*_i\), and let \(w_i^*, W_i^*\) and \(\Omega^*\) be defined analogously. Without loss of generality we assume that the first \(L\) diagonal elements of \(\Omega^*\) are non-zero and the remaining \(k-L\) elements are zero.

Proposition 4: If \(\Omega\) and \(\Omega^*\) are diagonal and satisfy (5.2), then the set of estimates \(\beta\) satisfying (5.1) is contained in the convex hull of \(\Omega^*\).

Proof: The proof is by induction. Let the proposition hold for \(\Omega^{-1}_{m-1}\) and \(\Omega^*_{m-1}\) with \(m-1 < k\). Then \(\alpha \equiv (A-\Omega^{-1}_{m-1})^{-1}Ab\) lies in the convex hull of \(\Omega^*_{m-1}\). By assumption, the proposition holds for \(\Omega^{-1}_{m-1}\) and \(\Omega^*_{m-1}\) and for any \(A\), for instance also for \(A - W^*_m\). Thus also \(\gamma \equiv (A-\Omega^{-1}_{m-1}-W^*_m)^{-1}Ab = ((A-W^*_m)^{-1}(A-\Omega^*_{m-1})^{-1}Ab = (A-W^*_m)^{-1}Ab\), lies in the convex hull of \(\Omega^*_{m-1}\). (This is of course a different convex hull since \(A\) has been replaced by \(A - W^*_m\).) So the points

\[\beta^*_m \equiv (A-\Sigma_{i=1}^{m} \delta_i W^*_i)^{-1}Ab,\]  (5.6)

with \(\delta_i = 0, 1\) if \(i < m\), generate the hull for \(\alpha\) when \(\delta_m = 0\), and the hull for \(\gamma\) when \(\delta_m = 1\). Clearly the assumption is true for \(m = 1\), where \(\Omega^0 = \Omega^*_{0} = 0\).
Now consider $\eta \equiv (A-\Omega_m^{-1})^{-1}Ab$. We will show that $\eta$ lies between $\alpha$ and $\gamma$, and hence in the hull of the $2^m - 1 = 2^m$ vertices corresponding with $\Omega_m^\ast$. This then establishes the induction step.

The first step is to show that

$$\eta = \alpha(1-\mu) + \gamma\mu$$ \hspace{1cm} (5.7)

with

$$\mu = \frac{\omega \eta_m}{\omega \gamma_m}\hspace{1cm} (\text{if } \gamma_m \neq 0)$$ \hspace{1cm} (5.8)

where $\eta_m$ and $\gamma_m$ are the $m$-th elements of $\eta$ and $\gamma$, respectively. This is so since $(A-\Omega_m)\eta = Ab$, so $(A-\Omega_m^{-1})\eta - Ab = e \omega \eta_m$. Moreover, $(A-\Omega_m^{-1})\alpha = Ab$, so when we denote the $m$-th column of $(A-\Omega_m^{-1})^{-1}$ by $x_m$, there holds

$$\eta - \alpha = x_m \omega \eta_m$$ \hspace{1cm} (5.9)

and analogously

$$\gamma - \alpha = x_m \omega \gamma_m^\ast$$ \hspace{1cm} (5.10)

If $\gamma_m = 0$, then $\gamma = \alpha = \eta$. If $\gamma_m \neq 0$, then we may solve (5.9) and (5.10) for $x_m$, and (5.8) follows immediately.

The second step is to show that $0 \leq \mu \leq 1$. Taking the $m$-th element in (5.9) and (5.10), we have,

$$\omega \eta_m = \frac{\omega \alpha_m}{1-x_m \omega_m}, \text{ and } \omega \gamma_m^\ast = \frac{\omega \alpha_m}{1-x_m \omega_m^\ast}$$ \hspace{1cm} (5.11)

So

$$\mu = \frac{\omega (1-x_m \omega_m)}{\omega (1-x_m \omega_m^\ast)}.$$ \hspace{1cm} (5.12)
Now since \( 0 < x_{mm} a_{mm} < 1 \) (\( a_{mm} \) denoting the \((m,m)\)-th element of \( A \)) and since 
\( 0 < \omega_m < \omega_{mm} < a_{mm} \), there holds \( 0 \leq \mu \leq 1 \).

We have thus shown that the diagonality of \( \Omega \) further reduces the region where \( \beta \) may lie when measurement error is present. In practical applications, the most obvious use of this result is to compute all \( 2^k \) points \( \beta_0 \) and to derive the interval in which each coefficient lies. These intervals will in general be smaller than the ones obtained from Proposition 2 by choosing for \( \psi \) the \( k \) unit vectors respectively.

An example shows that the convex polyhedron need not be a minimal bound for \( \beta \). Consider the case of two regressors, where all variables (including the regressand) are subject to measurement error and the 3x3 matrix \( \Omega \) is diagonal. If \( \phi \) is not restricted to be diagonal, the set of estimates \( \beta \) is bounded by the ellipsoid spawned by \( \phi^* \) given in Proposition 1. As has been shown in Section 4, the ellipsoid spawned by \( \Omega^* \) (or by \( \phi^* \) in which \( \phi_{11}^* \) is set equal to zero) is the same ellipsoid with a smaller radius. If \( \phi \) is restricted to be diagonal all \( 4 \) vertex points (\( k = 2 \)) lie on the surface of this latter ellipsoid.

Let \( a \) and \( c \) be the vertex points (besides \( b \) and \( b^* \)):

\[
a = (A^{-1} - W^*_1)^{-1} A b
\]

(5.13)

\[
c = (A^{-1} - W^*_2)^{-1} A b
\]

(5.14)

Assume without loss of generality that \( b > 0 \). Let us follow the path from \( b \) to \( a \). Note that

\[
\beta - b = A_1^{-1} \omega_1 \beta_1 + A_2^{-1} \omega_2 \beta_2,
\]

(5.15)

where \( A_1^{-1} \) and \( A_2^{-1} \) are the first and second column of \( A^{-1} \), respectively, and \( \beta_1 \) and \( \beta_2 \) are the two elements of \( \beta \). Going from \( b \) to \( a \), we set \( \omega_2 = 0 \) and let \( \omega_1 \) go from \( 0 \) to \( \omega^* \). So \( \beta - b = A_1^{-1} \omega_1 \beta_1 \). As \( b_1 > 0 \), \( \omega_1 \beta_1 > 0 \) and has as its maximum \( \omega_1 a_1 \). As \( (A_1^{-1})_{11} \) > 0 the line has a positive angle with \( e_1 \). Analo-

1) As \( a_{kk} \) is also the \((k,k)\)-th element of \( A - \Omega_{k-1} \), this inequality follows from the fact that the product of a diagonal element of a positive definite matrix with the corresponding element of its inverse does not exceed 1; moreover, both elements are of course positive.
gously, the line from $b$ to $c$ has a positive angle with $e_2$. A possible case is given in Figure 3, with $a > 0$, and $c > 0$. Going from $a$ to $b^*$ we have $\beta - a = (A-W_1^*)^{-1} \omega_2 \beta_2^*$. As $a_2 > 0$, $\omega_2 \beta_2 > 0$ with maximum $\omega_2^* \beta_2^*$. As $(A-W_1^*)^{-1} > 0$, the line has a positive angle with $e_2$. Analogously, the line from $c$ to $b^*$ has a positive angle with $e_1$. So if $\phi$ is restricted to be diagonal we end up with the shaded area in Figure 3 (the outer ellipsoid gives the bound for the estimates if $\phi$ is not restricted to be diagonal).

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure3.png}
\caption{The convex hull when $\Omega$ is diagonal and the vertices are in the same orthant}
\end{figure}

Now assume $c_1 < 0$. This line from $a$ to $b^*$ has again a positive angle with $e_2$, but the line from $c$ to $b^*$ has a negative angle with $e_1$. This is so since $\beta - c = (A-W_2^*)^{-1} \omega_1 \beta_1$, and $c_1 < 0$, so $\omega_1 \beta_1 < 0$. See Figure 4.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure4.png}
\caption{The convex hull when $\Omega$ is diagonal and the vertices are not in the same orthant}
\end{figure}
Now all $\beta$'s are within the shaded area, which is clearly not convex. The wasp waist is on $e_2$: in (5.15), choose $\omega_1$ and $\omega_2$ such that $\beta_1 = 0$, then we can next vary $\omega_1$ at will without affecting $\beta$, as $\beta_1 = 0$. 
6. An Empirical Example

In Van de Stadt, Kapteyn, Van de Geer (1983) a model of preference formation is constructed and estimated. The central relationship of the model is the following one:

\[ u_i = \beta_0 + \beta_1 u_{i(-1)} + \beta_2 f_{i1}(-1) + \beta_3 f_{i1} + \beta_4 y_{i1} + \beta_5 y_{i1}^* + \beta_6 f_{i1}^* + \epsilon_i \]  

(6.1)

The index \( i \) refers to the \( i \)-th household in the sample; \( u_i \) is a measure of the household's present wants (\( \exp(u_i) \) is the income the household head would consider just about "sufficient to make ends meet"); \( u_{i(-1)} \) is the same measure observed one year ago for the same household; \( f_{i1} \) is the log of the present number of household members ("log-family size") whereas \( f_{i1}(-1) \) is log-family size one year ago; \( y_{i1} \) is the present after tax household log-income. The starred variables are sample means of log-incomes and log-family sizes in the "social group" to which household \( i \) belongs. A social group is a set of households with identical characteristics (the age of the household head is in the same age bracket, the household heads have a similar education and they live in a town of similar size); \( \epsilon_i \) is a random disturbance term. See SKG for further details.

Thus relation (6.1) explains the level of a household's present financial wants by its family size (both present and lagged one period), its present log income (habit formation), by present log income and log-family size in the household's social group (preference interdependence), and by the level of financial wants one year ago (habit formation).

Since \( \epsilon_i \) is allowed to show negative serial correlation\(^1\), \( u_{i(-1)} \) may correlate negatively with \( \epsilon_i \). This is equivalent to allowing a measurement error in \( u_{i(-1)} \). The variables \( y_{i1}^* \) and \( f_{i1}^* \) are proxies for reference group effects and may therefore be expected to suffer from measurement error; \( f_{i1} \) and \( f_{i1}(-1) \) are crude proxies of the effects of family composition on financial wants, which can therefore also be expected to suffer from measurement error. Finally, \( y_{i1} \) may be subject to measurement error as well.

\(^1\) As a matter of fact, \( \epsilon_i \) has the form \( u_i - au_{i(-1)} + v_i \), where \( u_i \) and \( u_{i(-1)} \) are uncorrelated with each other or with \( v_i \); \( v_i \) may be serially correlated.
Table 1. Sample means and covariances of the observed variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Covariance with</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$\mu_i$</td>
</tr>
<tr>
<td>$\mu_i$</td>
<td>10.11</td>
<td>.1260</td>
</tr>
<tr>
<td>$\mu_i(-1)$</td>
<td>10.07</td>
<td>.1123</td>
</tr>
<tr>
<td>$fs_i(-1)$</td>
<td>1.01</td>
<td>.0876</td>
</tr>
<tr>
<td>$fs_i$</td>
<td>1.00</td>
<td>.0887</td>
</tr>
<tr>
<td>$y_i$</td>
<td>10.31</td>
<td>.1238</td>
</tr>
<tr>
<td>$y_i^*$</td>
<td>10.30</td>
<td>.0606</td>
</tr>
<tr>
<td>$fs_i^*$</td>
<td>1.00</td>
<td>.0434</td>
</tr>
</tbody>
</table>

Table 2. Specification of $\phi^*$.

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\mu_i$</th>
<th>$\mu_i(-1)$</th>
<th>$fs_i(-1)$</th>
<th>$fs_i$</th>
<th>$y_i$</th>
<th>$y_i^*$</th>
<th>$fs_i^*$</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_i$</td>
<td>.0154</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>$\mu_i(-1)$</td>
<td>.0165</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>35</td>
</tr>
<tr>
<td>$fs_i(-1)$</td>
<td></td>
<td>.0061</td>
<td>.0061</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>$fs_i$</td>
<td></td>
<td>.0061</td>
<td>.0061</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>$y_i$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.0040</td>
<td></td>
<td>15</td>
</tr>
<tr>
<td>$y_i^*$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.0130</td>
<td>.0100</td>
<td>40</td>
</tr>
<tr>
<td>$fs_i^*$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>.0100</td>
<td>.0150</td>
<td>40</td>
</tr>
</tbody>
</table>

The sample means, standard deviations and correlations of all variables involved are given in Table 1. Our specification of $\phi^*$ is given in Table 2. The column headed "% error" indicates the standard deviation of the mea-
surement errors (the square root of the diagonal of $\Omega^*$) as a percentage of the sample standard deviation of the corresponding observed variables. So we allow for example a 40% measurement error in $y_1^*$ and $fs_1^*$. Furthermore, a perfect correlation is imposed between the measurement error in $fs_1(-1)$ and $fs_1$.

We present extreme values for the elements of $\beta$ (using Proposition 2 with $\psi$ equal to the successive unit vectors, or by using Proposition 4) for five cases.

(i) $\phi^*$ is as given in Table 2.

(ii) $\phi^*_{11} = 0$. For the rest $\phi^*$ is as given in Table 2. The intervals for $\beta$ should be tighter than in the previous case.

(iii) The off-diagonal elements in Table 1 are set equal to zero.

(iv) As Case (iii), with $\phi^*_{11} = 0$. The intervals for $\beta$ should be tighter than in the previous case.

(v) As Case (iii), but diagonality is imposed on $\phi$. Again, this should narrow the intervals relative to the previous case.

In Table 3 the values of $b$ and $b^*$ are presented, along with the extreme values of $\beta$ for the five cases considered.

For all specifications of $\phi^*$, $R-\phi^*$ is positive definite. As a result, $s_{\phi^*}^2$ is always positive, as it should be. The various columns in Table 3 are pretty much according to expectation. The intervals for $\beta_i$ are a great deal wider in Case (i) than in Case (ii). In Case (ii) we see that $\beta_5$ and $\beta_6$ can switch signs depending on the choice of $\phi$. In Case (i) the interval for $\beta_2$ becomes so wide that this parameter may reverse signs as well. Similarly, Case (iii) gives rise to wider intervals than Case (iv). Comparing (iii) and (iv) to (i) and (ii) makes it clear that, in this example, the diagonal $\phi^*$ generates wider intervals. Now, $\beta_3$ may reverse signs as well. Finally, imposing diagonality narrows the interval dramatically. No parameter estimate reverses signs.

The example illustrates two points. First, it is important to use prior information economically. If one "knows" that $\phi$ is diagonal, this knowledge should be used. Otherwise the computed intervals are much wider than the intervals that correspond to one's prior knowledge. Secondly, allowing for measurement error in the endogenous variable (and correlation between this error and the errors in the exogenous variables) has a non-trivial influence on the intervals for the $\beta_i$. 
Table 3. Extreme values of $\beta^1$

<table>
<thead>
<tr>
<th>$\phi = 0$</th>
<th>non-diagonal $\phi^*$</th>
<th>diagonal $\phi^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(i)</td>
<td>(ii)</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.509 (.026)</td>
<td>0.772</td>
</tr>
<tr>
<td></td>
<td>0.369</td>
<td>0.491</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>-0.013 (.032)</td>
<td>-0.079</td>
</tr>
<tr>
<td></td>
<td>-0.121</td>
<td>-0.087</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>0.066 (.031)</td>
<td>0.095</td>
</tr>
<tr>
<td></td>
<td>0.038</td>
<td>0.057</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>0.298 (.031)</td>
<td>0.149</td>
</tr>
<tr>
<td></td>
<td>0.029</td>
<td>0.117</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>0.071 (.029)</td>
<td>0.047</td>
</tr>
<tr>
<td></td>
<td>-0.152</td>
<td>-0.057</td>
</tr>
<tr>
<td>$\beta_6$</td>
<td>-0.031 (.025)</td>
<td>-0.025</td>
</tr>
<tr>
<td></td>
<td>-0.180</td>
<td>-0.112</td>
</tr>
<tr>
<td>$s^2$</td>
<td>0.0021</td>
<td>0.0175</td>
</tr>
</tbody>
</table>

1) Standard errors in parentheses. Each cell in the columns (i)-(v) contains the extreme values for the elements of $\beta$. 
7. Conclusion

As illustrated in Section 6, it is very simple to apply Propositions 2 and 4 to empirical problems, and the analysis could easily be incorporated in regression packages. Since the propositions cover a wide range of cases, the researcher has considerable freedom to express his prior ideas about \( \Omega \) as precisely or as vaguely as he wants. The result of the analysis will then succinctly summarize the sensitivity of estimation outcomes for assumptions about the quality of the data used.

It appears that the framework developed in this paper will allow for extensions to more complicated models. Consider for example the \( j \)-th structural equation in a linear simultaneous equations system:

\[
y_j = Y_j\alpha_0 + \varepsilon_j\beta_0 + \varepsilon_j,
\]

where \( Y_j \) and \( \varepsilon_j \) are matrices of endogenous and exogenous variables respectively, included as explanatory variables in this equation; \( y_j \) is the vector of endogenous variables to be explained by this equation and \( \varepsilon_j \) is a vector of errors. Let \( \Xi \) be the matrix of all exogenous variables in the system. Then 2SLS amounts to GLS applied to

\[
\Xi'Y_j = \Xi'Y_j\alpha_0 + \Xi'\varepsilon_j\beta_0 + \Xi'\varepsilon_j.
\]

If \( \Xi \) is measured with error, this model becomes similar to (2.1)-(2.3). Since \( \Xi \) occurs on both sides of the equation, the measurement errors in the left and right hand side variables will in general be correlated. For the special case where \( \gamma_0 = 0 \), it is easy to show that Proposition 1 can be applied directly to derive an ellipsoid for a consistent estimate of \( \alpha_0 \), defined analogous to \( \beta \) (cf. (2.4)). (Bekker, Kapteyn and Wansbeek (1984) have derived the same ellipsoid without reference to Proposition 1, assuming that all exogenous variables are measured with error.) Proposition 1 is not applicable when \( \gamma_0 \neq 0 \). For that more general case further research is needed.
References


Leamer, E.E. (1983), Destructive Diagnostics for the Errors-in-Variable Model, Department of Economics, UCLA.


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

<table>
<thead>
<tr>
<th>No.</th>
<th>Authors/Co-authors</th>
<th>Title</th>
<th>Publication Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.</td>
<td>P. Meys</td>
<td>Het autoritair etatisme.</td>
<td>jan. 1983</td>
</tr>
<tr>
<td>5.</td>
<td>H.J. Klok</td>
<td>De klassieke politieke economie geherwaardeerd.</td>
<td>febr. 1983</td>
</tr>
<tr>
<td>6.</td>
<td>J. Glombowski, M. Krüger</td>
<td>Unemployment benefits and Goodwin's growth cycle model.</td>
<td>febr. 1983</td>
</tr>
<tr>
<td>7.</td>
<td>G.J.C.Th. van Schijndel</td>
<td>Inkomstenbelasting in een dynamisch model van de onderneming.</td>
<td>febr. 1983</td>
</tr>
<tr>
<td>8.</td>
<td>F. Boekema, L. Verhoef</td>
<td>Local initiatives: local enterprise agency/trust, business in the community.</td>
<td>febr. 1983</td>
</tr>
<tr>
<td>12.</td>
<td>W.J. Oomens</td>
<td>Economische interpretaties van de statistische resultaten van Lydia E. Pinkham.</td>
<td>maart 1983</td>
</tr>
</tbody>
</table>
IN 1983 REEDS VERSCHENEN (vervolg)

15. H. Gremmen
   Th. van Bergen
De universitaire economen over het
regeringsbeleid.
april

16. M.D. Merbis
On the compensator, Part III,
Stochastic Nash and Team Problems.
 april

17. H.J. Klok
Overheidstekort, rentestand en groei-
voet; terug naar een klassieke norm
voor de overheidsfinanciën?
 mei

18. D. Colasanto
   A. Kapteyn
   J. van der Gaag
Two Subjective Definitions of
Poverty: Results from the Wisconsin
 Basis Needs Study.
 mei

19. R.C.D. Berndsen
   H.P. Coenders
Is investeren onder slechte
omstandigheden en ondanks slechte
vooruitzichten zinvol?
 mei

20. B.B. v.d. Genugten
    J.L.M.J. Klijnen
Een Markovmodel ter beschrijving
van de ontwikkeling van de rundvee-
 stapel in Nederland.
 juni

21. M.F.C.M. Wijn
Enige fiscale-, juridische- en be-
drijfseconomische aspecten van
goodwill.
 juni

22. P.J.J. Donners
    R.M.J. Heuts
Een overzicht van tijdsvariërende
parametermodelspecificaties in
regressieanalyse.
 juni

23. J. Kriens
    R.H. Veenstra
Steekproefcontrole op ernstige
en niet-ernstige fouten.
 juli

24. M.F.C.M. Wijn
Mislukken van ondernemingen.
 juli

25. A.L. Hempenius
Relatieve Inkomenspositie,
Individuele en Sociale Inkomens-
bevrediging en Inkomensongelijkheid.
 aug.

26. B.R. Meijboom
Decomposition-based planning
procedures.
 sept.

27. P. Kooreman
   A. Kapteyn
The Systems Approach to Household
Labor Supply in The Netherlands
 sept.

28. B.B. v.d. Genugten
    K. v.d. Sloot
    M. Koren
    B. de Graad
Computergebruik bij propedeuse-
colleges econometrie
 sept.

29. W. de Lange
Korter werken of
Houden wat je hebt
Tendenzen, feiten, meningen
 okt.
30. A. Kapteyn
S. van de Geer
H. van de Stadt
The impact of changes in income
and family composition on
subjective measures of well-being
okt.

31. J. van Mier
Gewone differentievergelijkingen
met niet-constante coëfficiënten
en partiële differentievergelijkingen
nov.

32. A.B. Dorsman
J. van der Hilst
Een nieuwe marktindex voor de
Amsterdamse effectenbeurs
De Tam
nov.

33. W. van Hulst
Het vervangingsprobleem bij duurzame
productiemiddelen en de ondernemings-
doelstelling volgens J.L. Meij
dec.

34. M.D. Merbis
Large-Scale Systems Theory for the
Interplay Model
dec.

35. J.P.C. Kleijn
Statistische Analyse:
Wat is de kans op .....?
dec.
<table>
<thead>
<tr>
<th>No.</th>
<th>Author(s)</th>
<th>Title</th>
<th>Month</th>
</tr>
</thead>
<tbody>
<tr>
<td>02.</td>
<td>Frans Boekema, Leo Verhoef</td>
<td>Lokale initiatieven; Sleutel voor werkgelegenheidsontwikkeling op lokaal en regionaal niveau</td>
<td>febr.</td>
</tr>
<tr>
<td>03.</td>
<td>J.H.J. Roemen</td>
<td>In- en uitstroom van melkvee in de Nederlandse rundveesector geschat m.b.v. een &quot;Markov&quot;-model</td>
<td>febr.</td>
</tr>
<tr>
<td>04.</td>
<td>M.D. Merbis</td>
<td>From structural form to state-space representation</td>
<td>febr.</td>
</tr>
<tr>
<td>05.</td>
<td>R.H. Veenstra, J. Kriens</td>
<td>Steekproefcontrole op ernstige en niet-ernstige fouten (gecorrigeerde versie)</td>
<td>maart</td>
</tr>
<tr>
<td>06.</td>
<td>Th. Mertens</td>
<td>Kritiek op Habermas' communicatie-theorie: een evaluatie van het Gadamer-Habermas-debat en van Habermas' interpretatie van de taal-handelingstheorie. Een onderzoeksverslag</td>
<td>maart</td>
</tr>
<tr>
<td>08.</td>
<td>B.R. Meijboom</td>
<td>An input-output like corporate model including multiple technologies and &quot;make-or-buy&quot; decisions</td>
<td>april</td>
</tr>
<tr>
<td>09.</td>
<td>J.J.A. Moors</td>
<td>On the equivalence between cooperative games and superadditive functions</td>
<td>april</td>
</tr>
<tr>
<td>10.</td>
<td>J. van Mier</td>
<td>Gewone differentievergelijkingen met niet-constante coëfficiënten en partiële differentie vergelijkingen (vervolg R.T.D. 83.31)</td>
<td>april</td>
</tr>
<tr>
<td>11.</td>
<td>W.J. Oomens</td>
<td>Het optimale prijs- en reclamebeleid van een monopolist</td>
<td>april</td>
</tr>
<tr>
<td>12.</td>
<td>P.A. Verheyen</td>
<td>Een dynamische ondernemingstheorie en de reacties op de overheidspolitiek</td>
<td>mei</td>
</tr>
<tr>
<td>13.</td>
<td>G.J.C.Th. van Schijndel</td>
<td>Vermogensverschafferscliëntëles in statistische en dynamische ondernemingsmodellen</td>
<td>mei</td>
</tr>
</tbody>
</table>
14 P. Kooreman
   A. Kapteyn
   The effects of economic and demographic variables on the allocation of leisure within the household

15 L. Bosch
   Over flexibele produktie-automatisering

16 M. Janssens
   R. Heuts
   On distributions of ratios of dependent random variables

17 J. Plasmans
   Specification and estimation of the linkage block of Interplay II (1953-1980)