Inference in a linear functional relationship with replications

Julio Hokama¹, Pedro Morettin², Heleno Bolfarine², Manuel Galea³

¹Universidade Federal Rural do Rio de Janeiro, Brazil
²Universidade de São Paulo, Brazil
³Pontificia Universidad Católica de Chile, Chile

Abstract

In this paper, we consider a model for data analysis with measurement errors. The main objective of this work is to develop statistical inference tools, such as parameter estimation and hypothesis tests in a linear functional relationship with replicated observations. For this purpose, we use the maximum likelihood method in the presence of incidental parameters, and the unbiased estimating equations approach. Both approaches lead to explicit expressions for the asymptotic covariance matrices of the estimators of the model parameters. A simulation study is performed to assess the empirical behavior of estimators and of a Wald statistic. The methodology is illustrated with a real data set.

Key words: Linear functional relationship; Replicated observations; Maximum likelihood estimation; Unbiased estimating equations; Asymptotic normality; Wald statistic.

1 Introduction

Measurement error models (MEM), also known as errors-in-variables models, are useful for describing different phenomena in many disciplines. MEM establish functional relationships among the observed variables that are subject to random errors of measurement. Examples include linear and nonlinear errors-in-variables regression models (when the predictors are contaminated with measurement error), factor analysis models, latent structure models, and simultaneous equation models. These models are used in medicine, life sciences, econometrics, chemometrics, geology, among other disciplines. Ignoring the measurement error may lead to biased or misleading results. Diverse aspects and applications of MEM can be found in Seber and Wild (1989), Cheng and van Ness (1999), Fuller (2006), Carroll et al. (2006), Buonaccorsi (2010), Wu (2010) and Yi (2017).

In this paper we consider a subclass of MEM commonly used in geology and analytical chemistry, and in many other areas to compare two measurement systems (Ripley and Thompson, 1987; Galea et al., 2003; Cheng and Riu, 2006; Fuller, 2006; Carstensen, 2010; Stevens et al., 2017). This class is an alternative to the simple linear regression model that assumes that the
predictor is subject to random errors of measurement. When $x$ is the predictor a distinction is usually made between the \textit{functional case} where the $x$ values are treated as fixed, and the \textit{structural case} where $x$ is considered as a random variable. In this article we consider the functional case. Estimation in functional relationships has been the subject of several papers in the statistical literature. As shown in Neyman and Scott (1948), maximum likelihood (ML) estimators are typically inconsistent and as considered in Patefield (1977) and Gleser (1981, 1985), the asymptotic covariance matrix is not given by the inverse of the Fisher information matrix. As shown in Solari (1969), the solution of the likelihood equations in the case of the functional relationship is not a maximum but a saddle point of the likelihood equation. Further, the likelihood function is unbounded.

ML estimation in the unreplicated case with the ratio of variance known is considered in Mak (1982), where a general treatment is presented for ML estimation in the presence of incidental parameters. Consistent estimation for the case of one variance known is considered in Cheng and van Ness (1991). Both papers derive their main results under the normality assumption. Extensions for the case of elliptical models are considered in Arellano-Valle et al. (1996) and Vilca-Labra et al. (1998). More recently, Giménez and Galea (2013) propose influence measures on corrected score estimators in a functional heteroscedastic model, while Riquelme et al. (2015) discussed inference and local influence measures in a linear functional mixed model with normal random effects and elliptical errors.

In this paper we propose methods for statistical inference considering replicated observations, which is a way of guaranteeing that no assumptions about the model variances are required to make the approach feasible. For this purpose, we use the ML method in the presence of incidental parameters, and the unbiased estimating equations (UEE) approach. Both approaches lead to explicit expressions for the asymptotic covariance matrices of the estimators of the model parameters.

The remaining of this article is organized as follows. In Section 2 we specify the functional model with replicas. In Section 3 we discussed the problem of statistical inference in the functional model with replicated observations, using the ML methodology and UEE approach. To obtain the ML estimators a simple EM type algorithm is implemented. The asymptotic covariance matrices of the estimators of the model parameters are derived and a Wald test statistic is presented. Model assessment is discussed in Section 4. A simulation study and an application to real data set are reported in Section 5. Some concluding remarks are given in Section 6. Some technical details and calculations are presented in the Appendices.

## 2 The Model

In the classical linear errors-in-variables regression model the relationship between the unobservable variables $x$ and $y$ is given by $y = \alpha + \beta x$, where $\alpha$ and $\beta$ denote unknown parameters. We observe instead $X = x + u$ and $Y = y + e$. We consider $r_i$ and $s_i$ replications on $x_i$ and $y_i$, respectively, so that we observe $Y_{ij}$ and $X_{ik}$, where

\begin{align}
Y_{ij} &= \alpha + \beta x_i + e_{ij} \\
X_{ik} &= x_i + u_{ik},
\end{align}

$j = 1, \ldots, s_i$, $k = 1, \ldots, r_i$ and $i = 1, \ldots, n$. Moreover, it is also assumed that $\{e_{ij}\}$ and $\{u_{ik}\}$
are independent and
\[ e_{ij} \overset{iid}{\sim} N(0, \sigma_{ee}) \quad \text{and} \quad u_{ik} \overset{iid}{\sim} N(0, \sigma_{uu}). \]

The above formulation allows for two versions of the model (1). The first, the functional model, considers that the \( x_1, ..., x_n \), are fixed and treated as parameters which are usually known as incidental parameters. The other formulation, the structural model, considers that \( x_1, ..., x_n \) are independent and identically distributed (iid) random variables.

Motivated by its applications in geology, analytical chemistry, medicine, among other areas, we consider the functional version of the measurement error model (Ripley and Thompson, 1987; Galea et al., 2003; Cheng and Riu, 2006; Fuller, 2006; Buonaccorsi, 2010). Note that this functional model has incidental parameters, that is, the number of parameters increases with the number of observations, and therefore statistical inference cannot be performed using the standard likelihood function. We use the ML method in the presence of incidental parameters (Mak, 1982), and using a modified likelihood we propose the unbiased estimating equations approach (Yi, 2017).

Regarding to the replicas, the design is balanced if each variable has the same number of measurements on every experimental unit, otherwise it is unbalanced. In many areas it is common to consider balanced designs, this is \( r_i = s_i = r, \) for \( i = 1, ..., n \). Other designs include i) \( r_i = r, \) and \( s_i = s; ii) r_i = r \) and \( s_i = 1, \) for \( i = 1, ..., n. \) In all these cases the number of replicates does not depend \( i = 1, ..., n. \) For the unbalanced case, a common design is \( r_i = s_i, \) for \( i = 1, ..., n. \) For an unbalanced design, we assume that the number of replicates in both \( x_i \) and \( y_i \) is bounded.

Let \( Z_i = (X_i^T, Y_i^T)^T \) with \( X_i = (X_{i1}, ..., X_{ir_i})^T \) and \( Y_i = (Y_{i1}, ..., Y_{is_i})^T, \) are independent for \( i = 1, ..., n. \) Then \( Z_1, ..., Z_n \) are random vectors such that
\[ Z_i \overset{ind}{\sim} N_{d_i}(\mu_i, \Sigma_i) \text{ with } \mu_i = \left( \frac{x_i1_{r_i}}{(\alpha + \beta x_i)1_{s_i}} \right) \text{ and } \Sigma_i = \begin{pmatrix} \sigma_{uu}I_{r_i} & 0 \\ 0 & \sigma_{ee}I_{s_i} \end{pmatrix}, \quad (2) \]
where \( d_i = r_i + s_i, \) for \( i = 1, ..., n. \) The parameter \( \theta = (\alpha, \beta, \sigma_{ee}, \sigma_{uu})^T \) is the parameter of interest or structural parameter and \( x = (x_1, ..., x_n)^T \) is the vector of incidental parameters.

So, the log-likelihood function corresponding to the observed sample \( Z_1, ..., Z_n \) is given by
\[ L(\theta, x) = \sum_{i=1}^{n} \{l_1(\theta, x) + l_2(\theta, x)\}, \quad (3) \]
where
\[ l_1(\theta, x) = -\frac{r_i}{2} \log 2\pi - \frac{r_i}{2} \log \sigma_{uu} - \frac{\delta_{i1}}{2}, \]
\[ l_2(\theta, x) = -\frac{s_i}{2} \log 2\pi - \frac{s_i}{2} \log \sigma_{ee} - \frac{\delta_{i2}}{2}, \]
with \( \delta_{i1} = (1/\sigma_{uu}) \sum_{k=1}^{r_i} (X_{ik} - x_i)^2 \) and \( \delta_{i2} = (1/\sigma_{ee}) \sum_{j=1}^{s_i} (Y_{ij} - (\alpha + \beta x_i))^2, \) for \( i = 1, ..., n. \)

Notice that our model is similar to the model considered by Dorff and Gurland (1961), where different number of replications are considered for each \( x_i \) and \( y_i, i = 1, ..., n. \) However,
Dorff and Gurland (1961) do not consider ML estimation. A less general replication structure is considered in the structural model of Chan and Mak (1979) where the likelihood approach is implemented. The assumption of equal number of replications on $x$ and $y$ is required for studying the asymptotic behavior of the ML estimators. See also Stevens et al. (2017).

It is shown that the ML estimators of $\alpha$ and $\beta$ are consistent, but the ML estimators of $\sigma_{ee}$ and $\sigma_{uu}$ are not. However, consistent estimators can be obtained by slightly modifying the ML estimators of those parameters. In the most general situation the ML estimator of $\theta$ has to be obtained numerically. An EM-type algorithm can be implemented in that situation. In some less general situation, the ML estimator of $\beta$ can be obtained as the solution of a fourth degree equation. The problem of multiple roots can be circumvented by using the likelihood function or by picking the solution that follows from the solution of the EM-type algorithm when it converges. By using results in Mak (1982) on incidental parameter estimation, the asymptotic distribution of the ML estimator is obtained, a result so far not available in the literature. More recently, Rasekh and Fieller (2003) and Giménez and Patat (2014) discussed influence diagnostics in a functional measurement error model with replicated data, while Lin and Cao (2013) discuss estimation in a replicated structural measurement error model in which the replicated observations follow a scale mixtures of normal distributions.

### 3 Statistical Inference

In this section we discuss parameter estimation and hypothesis testing for model (1), namely $\alpha, \beta, \sigma_{ee}$ and $\sigma_{uu}$. We start with the ML method in presence of incidental parameters. Next we also consider of estimating equations method.

#### 3.1 Maximum Likelihood Method with Incidental Parameters

In this section the ML estimator of $x_i$ is obtained and it is shown that the ML estimators $\hat{\theta}$ of the true parameter value $\theta^0 = (\alpha^0, \beta^0, \sigma^0_{ee}, \sigma^0_{uu})^T$ are not consistent that is, $\hat{\theta} \overset{p}{\to} \theta^1 \neq \theta^0$. Specifically, we show that $\hat{\theta} \overset{p}{\to} \theta^1 = G \theta^0$ with $G$ a $4 \times 4$ full rank matrix. Then a consistent estimator, $\hat{\theta}_c$, is given by $\hat{\theta}_c = G^{-1} \hat{\theta}$.

**Lemma 3.1.** Given $\theta = (\alpha, \beta, \sigma_{ee}, \sigma_{uu})^T$ the ML estimator of $x_i$ is given by

$$\hat{x}_i = \hat{x}_i(Z_i, \theta) = \left\{ r_i \sigma_{ee} \bar{X}_i + s_i \beta \sigma_{uu} (\bar{Y}_i - \alpha) \right\} / c_i, \quad (4)$$

where $\bar{X}_i = \sum_{j=1}^n X_{ij} / r_i$, $\bar{Y}_i = \sum_{j=1}^n Y_{ij} / s_i$, and $c_i = r_i \sigma_{ee} + s_i \beta^2 \sigma_{uu}$ for $i = 1, \ldots, n$.

The proof of the above lemma follows by directly maximizing the log-likelihood (3) with respect to $x_i$ and keeping $\theta$ fixed. Also, a direct calculation leads to the following result.

**Lemma 3.2.** The expected value and variance of the ML estimator of $x_i$, $\hat{x}_i$, are given by

$$E(\hat{x}_i) = x_i \text{ and } Var(\hat{x}_i) = \sigma_{ee} \sigma_{uu} / c_i, \quad i = 1, \ldots, n. \quad (5)$$

That is, $\hat{x}_i$ is an unbiased estimator of $x_i$, for $i = 1, \ldots, n$. 

4
Replacing $\hat{x}_i$ given in (4) in the expression for the likelihood given in (3) we obtain the function

$$h_i(Z_i, \theta) = \frac{-r_i}{2} \log \sigma_{uu} - \frac{s_i}{2} \log \sigma_{ee} - \frac{1}{2} \left\{ \sigma_{uu}^{-1} \sum_{j=1}^{r_i} (X_{ij} - \hat{X}_i)^2 + \sigma_{ee}^{-1} \sum_{j=1}^{s_i} (Y_{ij} - \alpha - \beta \hat{x}_i)^2 \right\}, \quad (6)$$

from which we obtain

$$L_h(\theta) = \sum_{i=1}^{n} h_i(Z_i, \theta), \quad (7)$$

which upon maximized yields the ML estimator of $\theta, \hat{\theta}$.

In the following result we assume that the number of replicas does not depend on $i = 1,...,n$. For an unbalanced design, we assume that the number of replicates in both $x_i$ and $y_i$ is bounded.

**Theorem 3.1.** Given $\theta^0 = (\alpha^0, \beta^0, \sigma_{ee}^0, \sigma_{uu}^0)^T$ the true parameter value, the ML estimator $\hat{\theta}$ converges in probability to

$$\theta^1 = (\alpha^0, \beta^0, w\sigma_{ee}^0, w\sigma_{uu}^0)^T,$$

where $(\bar{r} + \bar{s} - 1)/(\bar{r} + \bar{s}) \uparrow w$ with $\frac{1}{2} < w < 1$, $w \uparrow 1$ when $(\bar{r} + \bar{s}) \to \infty$, and $\bar{r} = \sum_{k=1}^{n} r_k/n$, $\bar{s} = \sum_{j=1}^{s} s_j/n$.

**Proof.** After lengthy algebraic manipulations and disregarding unimportant constants, we obtain replacing (4) in (6) the following expression:

$$h_i(Z_i, \theta) = \frac{-r_i}{2} \log \sigma_{uu} - \frac{s_i}{2} \log \sigma_{ee} - \frac{1}{2} \left\{ \sigma_{uu}^{-1} \sum_{j=1}^{r_i} (X_{ij} - \hat{X}_i)^2 + \sigma_{ee}^{-1} \sum_{j=1}^{s_i} (Y_{ij} - \hat{Y}_i)^2 \right\} + \left( r_i s_i/c_i \right) (\hat{Y}_i - \alpha - \beta \hat{X}_i)^2,$$

where $\hat{X}_i$ and $\hat{Y}_i$ are given in (4). Now, taking expectation with respect to the true $\theta^0 = (\alpha^0, \beta^0, \sigma_{ee}^0, \sigma_{uu}^0)^T$ it follows that

$$\sum_{j=1}^{r_i} E_0(X_{ij} - \hat{X}_i)^2 = (r_i - 1)\sigma_{uu}^0, \quad \sum_{j=1}^{s_i} E_0(Y_{ij} - \hat{Y}_i)^2 = (s_i - 1)\sigma_{ee}^0$$

and

$$E_0(\hat{Y}_i - \alpha - \beta \hat{X}_i)^2 = \frac{\sigma_{ee}^0}{s_i} + \frac{\beta^2}{r_i s_i} \sigma_{uu}^0 + ((\alpha - \alpha^0) + x_i^0(\beta - \beta^0))^2,$$

which leads to
\[\psi(\theta) = \frac{1}{n} \sum_{i=1}^{n} E_0\{h_i(Z_i, \theta)\} = -\frac{\bar{r}}{2} \log \sigma_{uu} - \frac{\bar{s}}{2} \log \sigma_{ee}\]

\[= -\frac{1}{2n} \sum_{i=1}^{n} \left( \frac{r_i \sigma_{ee}^0 + s_i \beta_0^2 \sigma_{uu}^0}{c_i} \right) \]

\[= -\frac{1}{2} \sum_{i=1}^{n} \frac{r_is_i}{c_i} \left\{ (\alpha - \alpha_0^0) + (\beta - \beta_0^0) x_i^0 \right\}^2 \].

Clearly, expression (8) is maximized when

\[\alpha = \alpha_0^0, \beta = \beta_0^0, \sigma_{ee} = w \sigma_{ee}^0 \text{ and } \sigma_{uu} = w \sigma_{uu}^0,\]

independent of the value of \(x_i^0\), for \(i = 1, \ldots, n\). For details see Hokama (2001).

Thus, given the ML estimator \(\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}_{ee}, \hat{\sigma}_{uu})^\top\) of \(\theta^0\), which follows by maximizing the likelihood function (7), a consistent estimator of \(\theta^0\) is given by

\[\hat{\theta}_c = G^{-1} \hat{\theta} = (\hat{\alpha}, \hat{\beta}, w^{-1} \sigma_{ee}, w^{-1} \hat{\sigma}_{uu})^\top,\]

with \(G\) a 4 \times 4 known matrix given by \(G = \text{diag}(1, 1, w, w)\). Note also that this matrix satisfies the equation \(G^{-1} \theta^1 = \theta^0\).

As we already mentioned in some important cases in practice \(w\) does not depend on \(n\). These are summarized in the following corollary.

**Corollary 3.1.** Under the assumptions of Theorem 3.1, we have that,

i) If \(r_i = r\), and \(s_i = s\), \(i = 1, \ldots, n\), it follows that

\[\theta^1 = \left( \alpha_0^0, \beta_0^0, \frac{r + s - 1}{r + s} \sigma_{ee}^0, \frac{r + s - 1}{r + s} \sigma_{uu}^0 \right)^\top;\]

ii) If \(r_i = s_i = r\), \(i = 1, \ldots, n\), it follows that

\[\theta^1 = \left( \alpha_0^0, \beta_0^0, \frac{2r - 1}{2r} \sigma_{ee}^0, \frac{2r - 1}{2r} \sigma_{uu}^0 \right)^\top;\]

iii) If \(r_i = r\) and \(s_i = 1\), \(i = 1, \ldots, n\), it follows that

\[\theta^1 = \left( \alpha_0^0, \beta_0^0, \frac{r}{r + 1} \sigma_{ee}^0, \frac{r}{r + 1} \sigma_{uu}^0 \right)^\top.\]

iv) If \(r_i = s_i\), \(i = 1, \ldots, n\), then the ML estimator \(\hat{\theta}\) of \(\theta\) converges in probability to

\[\theta^1 = \left( \alpha_0^0, \beta_0^0, \frac{\bar{r} + 1 - 1/\bar{r}}{2} \sigma_{ee}^0, \frac{\bar{r} + 1 - 1/\bar{r}}{2} \sigma_{uu}^0 \right)^\top.\]
More general replication schemes, as the one considered in Fuller (1995), for example, may also be entertained. We discuss next the asymptotic distribution of the ML estimators and then derive the asymptotic distribution of the consistent estimator $\hat{\theta}_c$.

The ML estimator of $\theta$ follows by maximizing the function $L_h(\theta)$ in (7). Using the same notation as in Mak (1982), the following expressions are obtained after differentiating the function $h_i = h_i(Z_i, \theta)$ given in (6)

\[ q_{i\alpha} = \frac{\partial h_i}{\partial \alpha} = \frac{r_is_i}{c_i} (\bar{Y}_i - \alpha - \beta \bar{X}_i), \quad q_{i\beta} = \frac{\partial h_i}{\partial \beta} = q_{i\alpha} \bar{x}_i, \]  \hfill (11)

\[ q_{i\sigma_{ee}} = \frac{\partial h_i}{\partial \sigma_{ee}} = -\frac{s_i}{2\sigma_{ee}} + \frac{1}{2} \left\{ \frac{1}{\sigma_{ee}} \sum_{j=1}^{s_i} (Y_{ij} - \bar{Y}_i) \right\}^2 + \frac{r_i^2 s_i}{c_i^2} (\bar{Y}_i - \alpha - \beta \bar{X}_i)^2, \]  and

\[ q_{i\sigma_{uu}} = \frac{\partial h_i}{\partial \sigma_{uu}} = -\frac{r_i}{2\sigma_{uu}} + \frac{1}{2} \left\{ \frac{1}{\sigma_{uu}} \sum_{j=1}^{s_i} (X_{ij} - \bar{X}_i) \right\}^2 + \frac{r_i^2 \sigma^2}{c_i^2} (\bar{Y}_i - \alpha - \beta \bar{X}_i)^2. \]

Let $q_i(\theta) = (q_{i\alpha}, q_{i\beta}, q_{i\sigma_{ee}}, q_{i\sigma_{uu}})^T$, for $i = 1, ..., n$. To obtain the ML estimator of $\theta$ we must solve the following system of equations,

\[ \frac{\partial L_h(\theta)}{\partial \theta} = \sum_{i=1}^{n} q_i(\theta) = 0. \]  \hfill (12)

To solve the system (12) we can use standard iterative methods. Here we adapt an EM type algorithm proposed by Kimura (1992), to obtain the ML estimators of the structural parameters in a functional comparative calibration model without replicated observations. In effect, from (12) we have,

\[ \hat{\alpha} = \bar{Y}_s - \beta \bar{x}_s, \quad \hat{\beta} = \frac{\sum_{i=1}^{n} \sum_{j=1}^{s_i} (Y_{ij} - \bar{Y}_i) (\hat{x}_i - \bar{x}_s)}{\sum_{i=1}^{n} s_i (\hat{x}_i - \bar{x}_s)^2}, \]  \hfill (13)

\[ \hat{\sigma}_{ee} = \frac{1}{S} \sum_{i=1}^{n} \sum_{j=1}^{s_i} (Y_{ij} - (\hat{\alpha} + \hat{\beta} \hat{x}_i))^2, \quad \text{and} \quad \hat{\sigma}_{uu} = \frac{1}{R} \sum_{i=1}^{n} \sum_{j=1}^{r_i} (X_{ij} - \hat{x}_i)^2, \]

where $\bar{x}_s = S^{-1} \sum_{i=1}^{n} s_i \hat{x}_i$, $\bar{Y}_s = S^{-1} \sum_{i=1}^{n} s_i \bar{Y}_i$, $S = \sum_{i=1}^{n} s_i$ and $R = \sum_{i=1}^{n} r_i$.

Let $\hat{\theta}^{(m)} = (\hat{\alpha}^{(m)}, \hat{\beta}^{(m)}, \hat{\sigma}_{ee}^{(m)}, \hat{\sigma}_{uu}^{(m)})^T$ be the solution of (13) in the $m$–th iteration. In the $(m+1)$–th iteration, the algorithm proceeds as follows,

**Step 1** Calculate $\hat{x}_i = \hat{x}_i(\hat{\theta}^{(m)})$ given by (4), for $i = 1, ..., n$.

**Step 2** Calculate $\hat{\theta}^{(m+1)} = (\hat{\alpha}^{(m+1)}, \hat{\beta}^{(m+1)}, \hat{\sigma}_{ee}^{(m+1)}, \hat{\sigma}_{uu}^{(m+1)})^T$, from (13). Repeat steps 1–2 until convergence.

As initial estimates we can use $\theta^{(0)} = (0, 1, W_{YY}, W_{XX})^T$, where $W_{YY} = (S - n)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{s_i} (Y_{ij} - \bar{Y}_i)^2$ and $W_{XX} = (R - n)^{-1} \sum_{i=1}^{n} \sum_{j=1}^{r_i} (X_{ij} - \bar{X}_i)^2$, see Dorff and Gurland (1961). As a convergence criterion
Lemma 3.3. For model (1)-(2) an unbiased estimating equation for \( \theta = (\alpha, \beta, \sigma_{ee}, \sigma_{uu})^{\top} \) is given by

\[
\sum_{i=1}^{n} \Psi_i(Z_i, \theta) = 0,
\]

where, \( \Psi_i(Z_i, \theta) \in \mathbb{R}^p \) is a function of \((Z_i, \theta)\), with \( p = \text{dim}(\Theta) \), for \( i = 1, \ldots, n \). In our case \( p = 4 \). The solution, \( \hat{\theta}_M \), of (16) is called a M-estimator of \( \theta \). The estimating equation (16) is said to be unbiased if it has mean 0 when evaluated at the true parameter, that is,

\[
E_0(\Psi_i(Z_i, \theta)) = 0,
\]

for \( i = 1, \ldots, n \). For more details see Appendix A.6 in Carroll et al. (2006). See also Yi (2017).

Lemma 3.3. For model (1)-(2) an unbiased estimating equation for \( \theta = (\alpha, \beta, \sigma_{ee}, \sigma_{uu})^{\top} \) is given by

\[
\sum_{i=1}^{n} \Psi_i(Z_i, \theta) = \sum_{i=1}^{n} (\Psi_{i\alpha}, \Psi_{i\beta}, \Psi_{i\sigma_{ee}}, \Psi_{i\sigma_{uu}})^{\top} = 0,
\]
respectively. A consistent estimator of the asymptotic covariance matrix of \( \hat{\Omega} = \sigma \), see Dorf and Gurland (1961), in Appendix B.

The matrices \( A \) and \( \Psi \) are consistent estimators for \( Y \) and \( \sigma \).

From (11) it is easy to see that, \( E(q_{i\alpha}) = E(q_{i\beta}) = 0 \), \( E(q_{i\sigma_{ee}}) = \frac{1}{2} \left( \frac{r_i}{c_i} - \frac{1}{\sigma_{ee}} \right) \) and \( E(q_{i\sigma_{uu}}) = \frac{1}{2} \left( \frac{s_i \beta^2}{c_i} - \frac{1}{\sigma_{uu}} \right) \). Then, it follows that \( \Psi_{i\alpha} = q_{i\alpha} \), \( \Psi_{i\beta} = q_{i\beta} \), \( \Psi_{i\sigma_{ee}} = q_{i\sigma_{ee}} - E(q_{i\sigma_{ee}}) \) and \( \Psi_{i\sigma_{uu}} = q_{i\sigma_{uu}} - E(q_{i\sigma_{uu}}) \), as defined in (17), for \( i = 1, \ldots, n \).

Then \( \hat{\theta}_M \) is asymptotically normally distributed with mean \( \theta \) and covariance matrix \( n^{-1} \Omega \) with \( \Omega = A^{-1} V A^{-T} \), where

\[
A = \frac{1}{n} \sum_{i=1}^{n} E(\frac{\partial \Psi_{i}(Z_i, \theta)}{\partial \theta^T}) \quad \text{and} \quad V = \frac{1}{n} \sum_{i=1}^{n} \text{Cov}(\Psi_{i}(Z_i, \theta)).
\]

The matrices \( A \) and \( V \) can be estimated by

\[
\hat{A} = \frac{1}{n} \sum_{i=1}^{n} \left( \frac{\partial \Psi_{i}(Z_i, \hat{\theta}_M)}{\partial \theta^T} \right) \quad \text{and} \quad \hat{V} = \frac{1}{n} \sum_{i=1}^{n} \Psi_{i}(Z_i, \hat{\theta}_M) \Psi_{i}^T(Z_i, \hat{\theta}_M),
\]

respectively. A consistent estimator of the asymptotic covariance matrix of \( \hat{\theta}_M \) is given by

\[
\hat{\Omega} = \hat{A}^{-1} \hat{V} \hat{A}^{-T}.
\] (18)

This estimator is known as a sandwich estimator. The elements to obtain the matrix \( \hat{A} \) are presented in Appendix B.

Additionally, as suggested by a referee, we can estimate \( \sigma_{ee} \) and \( \sigma_{uu} \) using the replicates. In effect, see Dorf and Gurland (1961),

\[
\hat{\sigma}_{ee} = \sum_{i=1}^{n} v_{iy} V_{iY}, \quad V_{iY} = (s_i - 1)^{-1} \sum_{j=1}^{s_i} (Y_{ij} - \bar{Y}_{i})^2, \quad v_{iy} = (s_i - 1)/(S - n), \quad \text{and} \quad \hat{\sigma}_{uu} = \sum_{i=1}^{n} v_{ix} V_{iX}, \quad V_{iX} = (r_i - 1)^{-1} \sum_{k=1}^{r_i} (X_{ik} - \bar{X}_{i})^2, \quad v_{ix} = (r_i - 1)/(R - n),
\] (19)

are consistent estimators for \( \sigma_{ee} \) and \( \sigma_{uu} \), respectively. Notice that \( \sum_{i=1}^{n} v_{ix} = \sum_{i=1}^{n} v_{iy} = 1 \). That is, \( \hat{\sigma}_{ee} \) and \( \hat{\sigma}_{uu} \) correspond to the weighted averages of the variances of the replicates, in each experimental unit, in the response variable \( \bar{Y} \) and in the covariate \( \bar{X} \). Furthermore, \( \hat{\sigma}_{ee} \) and \( \hat{\sigma}_{uu} \) are unbiased estimators of \( \sigma_{ee} \) and \( \sigma_{uu} \), respectively. As suggested by a referee,

\[
\hat{\sigma}_{ee}^0 = (1/n) \sum_{i=1}^{n} V_{iY} \quad \text{and} \quad \hat{\sigma}_{uu}^0 = (1/n) \sum_{i=1}^{n} V_{iX},
\] (20)
are also unbiased and consistent estimators of $\sigma_{ee}$ and $\sigma_{uu}$. Then, using the replications, alternative unbiased estimating equations for $\sigma_{ee}$ and $\sigma_{uu}$ can be considered. We call this type of unbiased estimating equations Mr-estimation.

**Lemma 3.4.** For model (1)-(2) an alternative unbiased estimating equation for $\theta = (\alpha, \beta, \sigma_{ee}, \sigma_{uu})^\top$ is given by

$$
\sum_{i=1}^{n} q_i^r(\theta) = \sum_{i=1}^{n} (q_{i\alpha}, q_{i\beta}, q_{i\sigma_{ee}}, q_{i\sigma_{uu}})^\top = 0, \tag{21}
$$

where $q_{i\alpha}, q_{i\beta}$ are defined in the Equation (11), and

$$
q_{i\sigma_{ee}} = v_{iy}(\sigma_{ee} - V_iy), \quad q_{i\sigma_{uu}} = v_{ix}(\sigma_{uu} - V_ix).
$$

The solution of Equation (21) is an M-estimator, that we denote as $\hat{\theta}_{Mr}$. From (21), it follows that the estimators of $\alpha$ and $\beta$, satisfy the equations,

$$
\hat{\alpha} = \bar{Y}_w - \hat{\beta}\bar{X}_w \quad \text{and} \quad \hat{\beta} = \frac{\sum_{i=1}^{n} w_i(\bar{Y}_i - \bar{Y}_w)\bar{x}_i}{\sum_{i=1}^{n} w_i(X_i - X_w)\bar{x}_i}, \tag{22}
$$

where, $\bar{Y}_w = \frac{\sum_{i=1}^{n} w_i\bar{Y}_i}{\sum_{i=1}^{n} w_i}$, $\bar{X}_w = \frac{\sum_{i=1}^{n} w_i\bar{X}_i}{\sum_{i=1}^{n} w_i}$, $\bar{x}_i$, as defined in Equation (4), and $w_i = r_is_i/c_i$, $i = 1, \ldots, n$. So, applying the general theory of unbiased estimating equations we can obtain the asymptotic variance-covariance matrix of $\hat{\theta}_{Mr}$. In this case, $q_{i\sigma_{ee}\sigma_{ee}} = v_{iy}$, $q_{i\sigma_{ee}\sigma_{uu}} = 0$ and $q_{i\sigma_{uu}\sigma_{uu}} = v_{ix}$, $i = 1, \ldots, n$. The other derivatives to estimate the matrix $A$ can be found in the Appendix D. Similarly to Equation (18), the matrix $V$ can be estimated by

$$
\hat{V}^r = \frac{1}{n} \sum_{i=1}^{n} q_i^r(\hat{\theta}_{Mr})q_i^{r\top}(\hat{\theta}_{Mr}).
$$

### 3.3 The Wald test

In many situations we are interested in testing linear hypotheses of the form

$$
H_0: C\theta = c, \tag{23}
$$

where $C$ is a known matrix of order $r \times p$, with $\text{rank}(C) = r$ ($r \leq p$) and $c \in \mathbb{R}^r$. For the hypothesis (23), the Wald test uses the test statistic (Boos and Stefanski, 2013),

$$
W = n(C\hat{\theta} - c)^\top(C\hat{\Omega}^{-1}C^\top)^{-1}(C\hat{\theta} - c), \tag{24}
$$

where $\hat{\theta}$ is a consistent estimator of $\theta$ (in our case $\hat{\theta}_c$ or $\hat{\theta}_{Mr}$) and $\hat{\Omega}$ is the asymptotic covariance matrix defined in (14) or (18). An asymptotic test can be constructed because $W \xrightarrow{D} \chi^2_r$, where $\xrightarrow{D}$ denotes convergence in distribution. The test rejects at the $\alpha$ level if $W > \chi^2_{r,1-\alpha}$, where $\chi^2_{r,1-\alpha}$ is the $1 - \alpha$ quantile of the chi-squared distribution with $r$ degrees of freedom.
4 Model assessment and outlier detection

Any statistical analysis should include a critical analysis of the model assumptions. The quadratic forms, \( \delta_1 \) and \( \delta_2 \) defined in (3) may be useful for evaluating the quality of fit of the model (1). In effect, \( \delta_1 \sim \chi^2(r_i) \) and \( \delta_2 \sim \chi^2(s_i) \) and as \( \delta_1 \) and \( \delta_2 \) are independent random variables, \( \delta_i = \delta_1 + \delta_2 \sim \chi^2(d_i) \), for \( i = 1, ..., n \). By applying the Wilson-Hilferty transformation we have

\[
z_i^* = \left\{ \left( \frac{\delta_i / d_i}{1/3} - \left\{ 1 - \left(2/9d_i\right) \right\} \right) \right\} \sim N(0,1),
\]

approximately. A Q-Q plot of the transformed distances \( \{z_i^*, i = 1, ..., n\} \) can be used to evaluate the fit of the normal model (1). In practice, however, we must replace \( \theta \) for a consistent estimator, in our case by \( \hat{\theta}_c \). Deviations from the 45-degree line suggest lack of normality. Larger than expected values of the scale Mahalanobis distance, \( D_i = \delta_i / d_i, i = 1, ..., n \), identify outlying cases.

We can also analyze how deletion of the \( i \)-th individuals influences \( \hat{\theta}_1 = (\hat{\alpha}, \hat{\beta})^T \). Using infinitesimal case deletion diagnostics, we can consider the following perturbation of the log-likelihood function (7), see Demidenko and Stukel (2005),

\[
L_{\omega}(\theta_1) = \sum_{j \neq i} h_j(Z_j, \theta_1) + \omega h_i(Z_i, \theta_1),
\]

and let \( \hat{\theta}_1(\omega) \) be the estimator of \( \theta_1 \) under this perturbed log-likelihood. Then the sensitivity of the estimator \( \hat{\theta}_1(\omega) \) to deletion of the \( i \)-th log-likelihood contribution can be assessed using the derivative as

\[
\frac{d\hat{\theta}_1(\omega)}{d\omega} \bigg|_{\omega=1} = -\left( \sum_{j=1}^{n} \frac{\partial q_j}{\partial \theta_1} \right)^{-1} q_i = \left( \begin{array}{c} S_{i\alpha} \\ S_{i\beta} \end{array} \right),
\]

where the right-hand side is evaluated at \( \hat{\theta}_1 = \hat{\theta}_1(1) \), \( q_i = (q_{i\alpha}, q_{i\beta})^T \) and

\[
\sum_{j=1}^{n} \frac{\partial q_j}{\partial \theta_1} = \sum_{j=1}^{n} \left( \begin{array}{cc} q_{j\alpha\alpha} & q_{j\alpha\beta} \\ q_{j\beta\alpha} & q_{j\beta\beta} \end{array} \right),
\]

where \( q_{i\alpha}, q_{i\beta} \) as defined in (11), \( i = 1, ..., n \), and the elements \( q_{j\alpha\alpha}, q_{j\alpha\beta} \) and \( q_{j\beta\beta} \) are given in the Appendix D. An index plot of \( S_{i\alpha} \) and \( S_{i\beta} \) is useful to analyze the sensitivity of the estimators \( \hat{\alpha} \) and \( \hat{\beta} \).

5 Applications

In this section, two examples are presented to illustrate the methodology proposed in this paper. First, a simulation study is performed to assess the empirical behavior of the ML, M and Mr estimators, and of a Wald statistic, \( W \), to test the hypothesis \( H_0: \alpha = 0, \beta = 1 \). Second, a real data set is studied, Blood glucose.

5.1 Simulation study

Monte Carlo simulation was used to validate the methodology proposed in this paper. The values assigned for the parameters are \( \alpha = 0.10, \beta = 0.80 \) and \( \sigma_{ee} = \sigma_{uu} = 0.05 \). The true variable values \( x \) are
Table 1: Results from simulations comparing the ML, M and Mr estimators, including average parameter estimates (AVE), asymptotic standard errors (ASE) and the empirical standard errors (ESE).

<table>
<thead>
<tr>
<th>Estimate type</th>
<th>AVE $\hat{\alpha}$</th>
<th>AVE $\hat{\beta}$</th>
<th>ASE $\hat{\alpha}$</th>
<th>ASE $\hat{\beta}$</th>
<th>ESE $\hat{\alpha}$</th>
<th>ESE $\hat{\beta}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>n = 60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>0.0992</td>
<td>0.7971</td>
<td>0.0217</td>
<td>0.0391</td>
<td>0.0216</td>
<td>0.0387</td>
</tr>
<tr>
<td>M</td>
<td>0.0995</td>
<td>0.8001</td>
<td>0.0213</td>
<td>0.0373</td>
<td>0.0216</td>
<td>0.0381</td>
</tr>
<tr>
<td>Mr</td>
<td>0.0995</td>
<td>0.7989</td>
<td>0.0209</td>
<td>0.0001</td>
<td>0.0214</td>
<td>0.0374</td>
</tr>
<tr>
<td>n = 120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>0.0998</td>
<td>0.7974</td>
<td>0.0152</td>
<td>0.0278</td>
<td>0.0151</td>
<td>0.0266</td>
</tr>
<tr>
<td>M</td>
<td>0.0998</td>
<td>0.8006</td>
<td>0.0150</td>
<td>0.0267</td>
<td>0.0151</td>
<td>0.0262</td>
</tr>
<tr>
<td>Mr</td>
<td>0.0997</td>
<td>0.8009</td>
<td>0.0150</td>
<td>0.00004</td>
<td>0.0152</td>
<td>0.0277</td>
</tr>
<tr>
<td>n = 240</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ML</td>
<td>0.0999</td>
<td>0.7975</td>
<td>0.0107</td>
<td>0.0198</td>
<td>0.0110</td>
<td>0.0189</td>
</tr>
<tr>
<td>M</td>
<td>0.0999</td>
<td>0.8008</td>
<td>0.0106</td>
<td>0.0190</td>
<td>0.0110</td>
<td>0.0186</td>
</tr>
<tr>
<td>Mr</td>
<td>0.1007</td>
<td>0.7996</td>
<td>0.0107</td>
<td>0.00001</td>
<td>0.0102</td>
<td>0.0187</td>
</tr>
</tbody>
</table>

generated from a uniform distribution in the interval $[-1, 1]$. We assume $n = 60, 120, 240$, $r_i = s_i = 3$, $i = 1, ..., n$. Also, other scenarios with unbalanced designs were simulated, however the results are similar to the balanced case and are not reported in the paper. We performed the calculation of the average (AVE), asymptotic standard errors (ASE) and empirical standard errors (ESE) of each estimate. The average of any estimate from $\theta_1 = (\alpha, \beta)^T$, namely $\hat{\theta}$, are estimated as $\text{AVE} = \frac{1}{R} \sum_{r=1}^{R} \hat{\theta}^{(r)}$, with $\hat{\theta}^{(r)}$ being the estimate of $\theta$ in the r-th replication, for $r = 1, 2, ..., R$. The asymptotic standard errors is estimated as $\text{ASE} = \sqrt{\frac{1}{R-1} \sum_{r=1}^{R} \text{Var}_{\text{asy}}^{(r)}(\hat{\theta})}$ and the empirical standard errors is calculated as $\text{ESE} = \sqrt{\frac{1}{R-1} \sum_{r=1}^{R} (\hat{\theta}^{(r)} - \text{AVE})^2}$, with $\text{Var}_{\text{asy}}^{(r)}(\hat{\theta})$ being the estimate from $\text{Var}(\hat{\theta})$ in the r-th replication, which is obtained from (14) for ML estimator and of (18) for the M estimator. We consider in each scenario $R = 1000$ replicates. A summary of the results of the simulations is presented in the Table 1 and Figure 1 reports comparisons of the empirical and theoretical distributions of the $W$ statistic under the null hypothesis $H$.

Three results appear evident from this simulation study. Firstly, ML, M and Mr estimates of regression coefficients are very similar. The relative bias of the three estimators (not shown here) is very low, less than 1%. Secondly, estimates of the standard error of parameter estimates of the three methods (ML, M and Mr estimation), appear to be very similar, with the exception of the ASE of $\hat{\beta}$ using Mr-estimation, which are surprisingly minor. And thirdly, Figure 1 indicate that the $W$ statistic for testing the hypothesis $H$, appears to have the appropriate $\chi^2$ distribution under the null hypothesis. Similar situations are observed for other values of $n$ and are not shown here.
5.2 Blood glucose data set

We consider the Blood glucose data set, see Everitt (2005). The data consists on measurements on blood glucose for 52 women. The responses $Y_1, Y_2$ and $Y_3$ represent fasting glucose measurements on three occasions and the $X_1, X_2$ and $X_3$ are glucose measurements one hour after sugar intake. In this case $r_i = s_i = 3, i = 1, ..., n$. We do not consider a woman that the blood glucose level was greater than 400(mg/100ml) in one replica. Since there is measurement error in the blood glucose level, the model (1) is fitted. As the regression coefficients $\theta_1 = (\alpha, \beta)^T$ are of primary interest, in what follows we focuses in some aspect of statistical inference on these parameters.

Table 2 presents the ML, M and Mr estimates for $\theta_1 = (\alpha, \beta)^T$, the standard errors $se(\cdot)$ and the values of the $Z$ statistic ($z$–stat) to test the hypotheses $\alpha = 0$ and $\beta = 0$. The results in Table 2 show that the estimates of the coefficients $\alpha$ and $\beta$ and their standard errors are very similar using the unbiased estimating equations approach (M and Mr estimation). Although not so much difference is observed, the standard errors of the ML estimators of $\alpha$ and $\beta$ are slightly lower. These results of the fitting the functional model, using all three approaches on Table 2, were expected, since the $\alpha$ and $\beta$ score functions ($\Psi_{i\alpha}$ and $\Psi_{i\beta}$) are the same in the three types of estimators.

Table 2: ML, M and Mr estimates for the Blood glucose data set.

<table>
<thead>
<tr>
<th>Estimator</th>
<th>$\hat{\alpha}$</th>
<th>$se(\hat{\alpha})$</th>
<th>$z$–stat</th>
<th>$\hat{\beta}$</th>
<th>$se(\hat{\beta})$</th>
<th>$z$–stat</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML</td>
<td>55.2812</td>
<td>5.7417</td>
<td>9.6280</td>
<td>0.1647</td>
<td>0.0553</td>
<td>2.9783</td>
</tr>
<tr>
<td>M</td>
<td>51.5802</td>
<td>8.3059</td>
<td>6.2101</td>
<td>0.1994</td>
<td>0.0796</td>
<td>2.5050</td>
</tr>
<tr>
<td>Mr</td>
<td>50.8720</td>
<td>6.9183</td>
<td>7.3533</td>
<td>0.2060</td>
<td>0.0665</td>
<td>3.0977</td>
</tr>
</tbody>
</table>
Figure 2 shows the scatter plot of the means of the replications and the adjusted regression lines using ML, M and Mr estimation. We can see that fitting a straight line seems reasonable for this data set.

Figure 2: Scatter plot of the means of the replications and lines adjusted for the glucose data set.

Figure 3 shows a Q-Q plot of the transformed distances and index plot of $D_i$ for the glucose data set. From Figure 3 top, we see that deviations from straight line are moderate and the assumption of normality seems reasonable, although there are some potential outliers. The $i$th observation is a potential outliers if $D_i > \overline{D} + 2\text{sd}(D)$, where $\overline{D} = \sum_{i=1}^{n}D_i/n$ and $\text{sd}(D)$ is the standard deviation of $D_1, ..., D_n$. Continuous line in Figure 3 bottom, corresponds to the value $\overline{D} + 2\text{sd}(D)$. This plot highlights slightly one potential outlier, the woman 40, who does not have a significant influence on the estimation process. Similar results are seen in Figure 4, which corresponds to the index plot of $S_{i\beta}$ for the M estimators. For ML and Mr estimators these graphics are very similar and they are not shown here. The dashed red lines correspond to plus and minus two standard deviations of $S_{i\beta}$, $i = 1, ..., n$.

5.3 Sensitivity analysis

In regression models it is of interest to determine if the estimates of the regression coefficients are sensitive to small changes in the response variable or in the predictor. Sensitivity aspects of the ML, M and Mr estimators can be illustrated perturbing some observations in the original data. For example, changes in the estimates of $\beta$ can be evaluated using the following procedure. First, an observation can be perturbed to create an outlier by $Z_m \leftarrow Z_m + \Delta_1p$, for $\Delta = -40.0, -20.0, 0.0, 20.0, 40.0$ and $m \in \{1, 2, ..., n\}$. Then, re-calculate the ML, M and Mr estimators of $\beta$. Let $\hat{\beta}_{\Delta j}$ be the estimator $j$ of $\beta$ using the perturbed data set, with $j = 1$ for the ML estimator, $j = 2$ for the M estimator and $j = 3$ for the Mr estimator. Finally, a graph of $\hat{\beta}_{\Delta j}$ versus $\Delta$, is useful to visualize changes in the estimators.
Figure 3: Q-Q plot of the transformed distances (top) and index plot of $D_i$ (below) for the glucose data set.

Figure 5 shows the curves of the estimates of $\hat{\beta}_{\Delta_j}$ versus $\Delta$ for each of the three estimators considered in this paper. We can see that the effect of this perturbation scheme is very similar in the M estimators of $\beta$. This perturbation scheme produces greater changes in the ML estimator.

We can also use the index plot of $S_i\beta$ to analyze the sensitivity of the estimator of the regression coefficient $\beta$. As an illustration we simultaneously replaced a replica in $X$ by $400$ (mg/100ml) in the women #3 and #14. In Figure 6, the index plot of $S_i\beta$ for this perturbation scheme is shown. We note that this sensitivity measure is effective in detecting outliers within replicates.
Figure 4: Index plot of $S_{i\beta}$ for the M-estimator of $\beta$, for glucose data set.

Figure 5: $\hat{\beta}$ in the perturbed glucose data set; ML(black line), M(blue line) and Mr(red line).

6 Conclusions

Measurement error models (MEM), also known as errors-in-variables models, are useful for describing different phenomena in many disciplines. MEM establish functional relationship among variables ob-
served subject to random errors of measurement. The modeling and treatment of measurement errors is a relevant topic in the statistical analysis of errors-in-variables models. In this paper we discussed ML and M estimation of the unknown parameters in a linear functional relationship with replications.

The approach proposed by Mak (1982) for ML estimation in the presence of incidental parameters was used, and the resulting estimators are shown to be consistent and asymptotically normal. To obtain the ML estimators of the structural parameters an EM type algorithm was implemented. A explicit expression was given for the asymptotic covariance matrix of the ML estimators. Additionally, we used the unbiased estimating equation approach, to obtain other consistent estimator, which also has asymptotically a multivariate normal distribution. A closed expression was given for the asymptotic covariance matrix of this estimator.

The methodology developed in this paper was illustrated with a glucose data set. For the assessment of the model adequacy we implemented Q-Q plots of the transformed distances, a topic that has been little discussed in the literature. A Wald test was proposed to test linear hypotheses of interest, which has a good behavior in finite samples. A scaled Mahalanobis distance was used to detect outliers and a diagnostic measure was proposed to analyze the sensitivity of the regression coefficients.

From the simulation study and the application to glucose data set, we see that the three approaches discussed in this paper produce very similar results in terms of estimators, standard errors, model assessment and outlier detection. The implementation of ML estimation in the presence of incidental parameters in the linear functional models is more challenging (Bolfarine and Galea-Rojas, 1995; Galea et al., 2003, 2006) and it is not easy to extend to non-normal models (Vilca-Labra et al., 1998; de Castro and Galea, 2010). The unbiased estimating equations approach used in this article has computational

Figure 6: Index plot of $S_{i;\beta}$ for the perturbed glucose data set.
advantages, it is simpler to implement. Based on these finds, we think that the UEE approach proposed in this work is appropriate for the statistical modeling of this type of functional regression models.

Normality of the distribution of the measurement errors is a key assumption in this paper. Recently Galea and de Castro (2017) used Mr-estimation approach in the functional model with replicas under the $t$–distribution with finite second moment. More flexible distributions could be assumed for the measurement errors. This topic will be addressed elsewhere.

Acknowledgments

The authors acknowledge the partial financial support from Project Puente 001/2019 Direccióndé Investigación de la Vicerrectoría de Investigación de la Pontificia Universidad Católica de Chile, Chile (MG), CNPq (JH) and Fapesp 2018/04654-9 (PM).

References


Appendices

A Incidental Parameter Estimation

Let $Z_1, \ldots, Z_n$, independent $p$-dimensional random vectors with log-likelihood function given by

$$
\sum_{i=1}^{n} \log f_i(z_i; \theta, x_i),
$$

where $f_i(z_i; \theta, x_i)$ is the density of $Z_i$, $i = 1, \ldots, n$, $\theta = (\theta_1, \ldots, \theta_p)^\top \in \Theta \subset \mathbb{R}^p$ and $x_i \in X_i \subset \mathbb{R}$, $i = 1, \ldots, n$, are the incidental parameters. Suppose that $\theta^0 \in \Theta$ and $x_i^0 \in X_i$, $i = 1, \ldots, n$, where $\theta^0$ and $x_i^0$, $i = 1, \ldots, n$, denote the true parameter values. The expected values are taken with respect to $\theta^0$ and $x_i^0$, $i = 1, \ldots, n$, which will be denoted by $E_0(\cdot) = E(\cdot | \theta^0, x_1^0, \ldots, x_n^0)$. For each $i$ and given $\theta$, let $\tilde{\theta}_i = \tilde{\theta}_i(Z_i, \theta)$, be an estimator (possibly depending on $\theta$) of $x_i$, with a possibility of being the conditional ML estimator, obtained by maximizing (28) with respect to $x_i$ for fixed $\theta$. Thus, replacing $x_i$ by $\tilde{x}_i$ in (28) we obtain

$$
\sum_{i=1}^{n} \log f_i(z_i; \theta, \tilde{x}_i) = \sum_{i=1}^{n} h_i(z_i; \theta).
$$

We also define the following functions.

$$
q_{\theta_j}(z_i; \theta) = \frac{\partial h_i(z_i; \theta)}{\partial \theta_j}, \quad j = 1, \ldots, p,
$$

$$
q_{\theta_j \theta_k}(z_i; \theta) = \frac{\partial^2 h_i(z_i; \theta)}{\partial \theta_j \partial \theta_k}, \quad j, k = 1, \ldots, p,
$$

and

$$
q_{\theta_j \theta_k}(z_i; \theta) = q_{\theta_j}(z_i; \theta) q_{\theta_k}(z_i; \theta), \quad j, k = 1, \ldots, p.
$$

Moreover, let $E_0(A_n(\theta))$ be the $p \times p$ matrix with entry $(j, k)$ given by

$$
n^{-1} \sum_{i=1}^{n} E_0(q_{\theta_j \theta_k}(Z_i; \theta)), \quad j, k = 1, \ldots, p.
$$

In Mak (1982), Section 2, general conditions are established under which (29) has a maximum $\tilde{\theta}_n = \tilde{\theta}(z_1, \ldots, z_n)$, which converges in probability to some $\theta^1$ in the interior of $\Theta$, where $\theta^1$ maximizes the function $\overline{\psi}(\theta)$ for all large $n$, where

$$
\overline{\psi}(\theta) = n^{-1} \sum_{i=1}^{n} E_0(h_i(Z_i; \theta)),
$$

and

$$
\sqrt{n} (V_n(\theta^1))^{-1/2} \{ E_0(A_n(\theta^1)) \} (\hat{\theta}_n - \theta^1) \overset{D}{\rightarrow} N_p(0, I_p),
$$

where $\overset{D}{\rightarrow}$ means convergence in distribution, with the $(j, k)$-th element of the $p \times p$ matrices $A_n(\theta^1) = (a_{jk})$ and $V_n(\theta^1) = (v_{jk})$ given, respectively, by

$$
a_{jk} = \frac{1}{n} \sum_{i=1}^{n} q_{\theta_j \theta_k}(Z_i; \theta^1) \quad \text{and} \quad v_{jk} = \frac{1}{n} \sum_{i=1}^{n} \text{Cov}(q_{\theta_j}(Z_i; \theta^1), q_{\theta_k}(Z_i; \theta^1)),
$$

20
where, as pointed out before, the expected values are taken with respect to the true values $\theta^0$ and $x^0_i$, $i = 1, \ldots, n$. The above results are proved in Mak (1982) and Giménez and Bolfarine (1997).

Thus, from (31), the covariance matrix asymptotic of $\hat{\theta}_n$, is given by

$$
\frac{1}{n}\{E_0(A_n(\theta^1))\}^{-1}V_n(\theta^1)\{E_0(A_n(\theta^1))\}^{-1}.
$$

(32)

It is also noted in Mak (1982) that in some situations it is possible to obtain estimators $\hat{x}_i$ so that $\theta^1$ depends only on $\theta^0$ (is independent of $x^0_i$), that is, there exists a function $g(.)$ such that $\theta^1 = g(\theta^0)$. If $g$ is one to one then, a consistent estimator of $\theta^0$ is given by $\hat{\theta}_n = g^{-1}(\hat{\theta}_n)$.

## B Sandwich covariance matrix estimation

In this Appendix we presents the second derivatives of the functions $\Psi_{r\gamma}$, $\gamma = \alpha, \beta, \sigma_{ee}, \sigma_{uu}$, given in (17). In effect, after standard algebraic manipulations, it follows that, for $i = 1, \ldots, n$,

$$
\Psi_{i\alpha\alpha} = -\frac{r_is_i}{c_i^2}, \quad \Psi_{i\alpha\beta} = -\frac{r_is_i}{c_i^2}\{c_i\bar{x}_i + s_i\beta\sigma_{uu}(\bar{Y}_i - \alpha - \beta X_i)\},
$$

$$
\Psi_{i\alpha\sigma_{ee}} = -\frac{r_is_i}{c_i^2}(\bar{Y}_i - \alpha - \beta \bar{X}_i), \quad \Psi_{i\alpha\sigma_{uu}} = -\frac{r_is_i^2\beta^2}{c_i^2}(\bar{Y}_i - \alpha - \beta \bar{X}_i),
$$

$$
\Psi_{i\beta\beta} = \Psi_{i\alpha\beta}\bar{x}_i + \Psi_{i\alpha\sigma_{uu}}\{(\bar{Y}_i - \alpha) - 2\bar{x}_i\beta\}, \quad \Psi_{i\beta\sigma_{ee}} = \Psi_{i\alpha\sigma_{uu}}(\bar{Y}_i - \alpha),
$$

$$
\Psi_{i\sigma_{ee}\sigma_{ee}} = \frac{1}{2}\left\{\frac{s_i - 1}{\sigma_{ee}^2} + \frac{s_i^2}{c_i^2}\right\} - \frac{1}{\sigma_{ee}^3}\sum_{j=1}^{n}(Y_{ij} - \bar{Y}_i)^2 + c_i^{-3}s_i(Y_i - \alpha - \beta \bar{X}_i)^2, \quad \Psi_{i\sigma_{ee}\sigma_{uu}} = \frac{1}{2}\left\{\frac{s_i - 1}{\sigma_{uu}^2} + \frac{s_i^2\beta^4}{c_i^2}\right\} - \frac{1}{\sigma_{uu}^3}\sum_{j=1}^{n}(X_{ij} - \bar{X}_i)^2 + c_i^{-3}s_i^3\beta^4(Y_i - \alpha - \beta \bar{X}_i)^2.
$$

## C Matrices $E_0(A_n(\theta^1))$ and $V_n(\theta^1)$

In the following, $E_0(a_{jk})$ is used to denote the $(j,k)$–th element of the matrix $E_0(A_n(\theta^1))$, which are given by

$$
E_0(a_{jk}) = \left\{\frac{1}{n}\sum_{i=1}^{n}E_0(q_{i\theta_j\theta_k}(Z_i;\theta))\right\}_{\theta=\theta^1},
$$

where

$$
q_{i\theta_j\theta_k}(Z_i;\theta) = \frac{\partial q_{i\theta_j}}{\partial \theta_k}, \quad j, k = 1, \ldots, 4,
$$

for $i = 1, \ldots, n$, $q_{i\theta_j}$ as defined in (11) and $\theta_1 = \alpha$, $\theta_2 = \beta$, $\theta_3 = \sigma_{ee}$ and $\theta_4 = \sigma_{uu}$. Using the second derivatives, it follows that the matrix $E_0(A_n(\theta^1))$ takes the form

$$
E_0(A_n(\theta^1)) = \begin{pmatrix}
E_0(a_{11}) & E_0(a_{12}) & 0 & 0 \\
E_0(a_{12}) & E_0(a_{22}) & E_0(a_{23}) & E_0(a_{24}) \\
0 & E_0(a_{23}) & E_0(a_{33}) & E_0(a_{34}) \\
0 & E_0(a_{24}) & E_0(a_{34}) & E_0(a_{44})
\end{pmatrix},
$$

(33)
where

\[ E_0(a_{11}) = \frac{-w}{n} \sum_{i=1}^{n} (r_i s_i/c_i^0), \quad E_0(a_{12}) = \frac{-w}{n} \sum_{i=1}^{n} (r_i s_i x_i^0/c_i^0), \quad E_0(a_{22}) = \frac{-w}{n} \sum_{i=1}^{n} (r_i s_i x_i^{02}/c_i^0), \]

\[ E_0(a_{23}) = \frac{-w^2 \beta^0_\sigma^0 u_\sigma}{n} \sum_{i=1}^{n} (r_i s_i/c_i^0), \quad E_0(a_{24}) = \frac{w^2 \sigma^0_\theta}{n} \sum_{i=1}^{n} (r_i s_i/c_i^{02}), \]

\[ E_0(a_{33}) = \frac{w^3}{n} \left\{ \frac{n}{\sigma^0_{ee}} \left( \frac{s}{2w} - \bar{s} + 1 \right) - \sum_{i=1}^{n} (r_i/c_i^0)^2 \right\}, \quad E_0(a_{34}) = -\frac{w^3 \beta^0_\theta}{n} \sum_{i=1}^{n} (r_i s_i/c_i^{02}) \text{ and} \]

\[ E_0(a_{44}) = \frac{w^3}{n} \left\{ \frac{n}{\sigma^0_{uu}} \left( \frac{\bar{r}}{2w} - \bar{r} + 1 \right) - \beta^0_\theta \sum_{i=1}^{n} (s_i/c_i^0)^2 \right\}. \]

We know that this matrix is given by

\[ V_n(\theta^1) = \frac{1}{n} \sum_{i=1}^{n} \text{Cov}(q_i(Z_i, \theta^1)|\theta^0, x_i) = \frac{1}{n} \sum_{i=1}^{n} \text{Cov}_0(q_i(\theta^1)), \]

where \(q_i(\theta^1) = (q_i(\theta^1), q_i(\theta^1), q_{\sigma_{ee}}(\theta^1), q_{\sigma_{uu}}(\theta^1))^T.\) In the following, \(v_{jk}\) denotes the \((j,k)\)-th element of the matrix \(V_n(\theta^1)\) given by

\[ v_{jk} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \text{Cov}_0(q_i(Z_i, \theta), q_i(Z_i, \theta)) \right\} \bigg|_{\theta = \theta^1}, \]

for \(j, k = 1, 2, 3, 4\) where \(\theta_1 = \alpha, \theta_2 = \beta, \theta_3 = \sigma_{ee}\) and \(\theta_4 = \sigma_{uu}.\) After some algebraic manipulations, the matrix \(V_n(\theta^1)\) takes the form

\[ V_n(\theta^1) = \begin{pmatrix} v_{11} & v_{12} & 0 & 0 \\ v_{12} & v_{22} & 0 & 0 \\ 0 & 0 & v_{33} & v_{34} \\ 0 & 0 & v_{34} & v_{44} \end{pmatrix}, \quad (34) \]

where,

\[ v_{11} = \frac{w^2}{n} \sum_{i=1}^{n} (r_i s_i/c_i^0), \quad v_{12} = \frac{w^2}{n} \sum_{i=1}^{n} (r_i s_i x_i^0/c_i^0), \quad v_{22} = \frac{w^2}{n} \left\{ \sum_{i=1}^{n} (r_i s_i x_i^{02}/c_i^0) + \sigma^0_{\theta} \sigma^0_{\theta} \sum_{i=1}^{n} (r_i s_i/c_i^{02}) \right\}, \]

\[ v_{33} = \frac{w^4}{2n \sigma^0_{ee}} \left\{ S - n + \sigma^0_{ee} \sum_{i=1}^{n} (r_i^2/c_i^{02}) \right\}, \quad v_{34} = \frac{\beta^0_\theta w^4}{2n} \sum_{i=1}^{n} (r_i s_i/c_i^{02}) \text{ and} \]

\[ v_{44} = \frac{w^4}{2n \sigma^0_{uu}} \left\{ R - n + \beta^0_\theta \sigma^0_{uu} \sum_{i=1}^{n} (s_i^2/c_i^{02}) \right\}. \]
The elements of the matrices $A_i$

Let $q_{i\gamma\tau} = \partial q_{i\gamma}/\partial \tau$, for $\gamma, \tau = \alpha, \beta, \sigma_{ee}, \sigma_{uu}$, with $q_{i\gamma}$ as defined in (11). Then, it follows that, the elements of the matrices $A_i$, for $i = 1, \ldots, n$, are given by,

$$
\begin{align*}
q_{i\alpha\alpha} &= \Psi_{i\alpha\alpha}, \\
q_{i\alpha\beta} &= \Psi_{i\alpha\beta}, \\
q_{i\alpha\sigma_{ee}} &= \Psi_{i\alpha\sigma_{ee}}, \\
q_{i\alpha\sigma_{uu}} &= \Psi_{i\alpha\sigma_{uu}}, \\
q_{i\beta\beta} &= \Psi_{i\beta\beta}, \\
q_{i\beta\sigma_{ee}} &= \Psi_{i\beta\sigma_{ee}}, \\
q_{i\beta\sigma_{uu}} &= \Psi_{i\beta\sigma_{uu}}, \\
q_{i\sigma_{ee}\sigma_{ee}} &= \frac{s_i}{2\sigma_{ee}^2} - \left\{ \frac{1}{\sigma_{ee}^3} \sum_{j=1}^{s_i} (Y_{ij} - \bar{Y}_i)^2 + c_i^{-3} r_i^3 s_i (\bar{Y}_i - \alpha - \beta \bar{X}_i)^2 \right\}, \\
q_{i\sigma_{ee}\sigma_{uu}} &= \frac{\partial q_{i\sigma_{ee}}}{\partial \sigma_{uu}} = -r_i^2 s_i^2 c_i^{-3} \beta^2 (\bar{Y}_i - \alpha - \beta \bar{X}_i)^2, \\
q_{i\sigma_{uu}\sigma_{uu}} &= \frac{\partial q_{i\sigma_{uu}}}{\partial \sigma_{uu}} = \frac{\tau_i}{2\sigma_{uu}^2} - \left\{ \frac{1}{\sigma_{uu}^3} \sum_{j=1}^{r_i} (X_{ij} - \bar{X}_i)^2 + c_i^{-3} r_i^3 s_i^3 \beta^4 (\bar{Y}_i - \alpha - \beta \bar{X}_i)^2 \right\},
\end{align*}
$$

where $\Psi_{i\alpha\gamma}$ and $\Psi_{i\beta\gamma}$ as defined in Appendix B.