PONTIFICIA
UNIVERSIDAD
CATÓLICA
DE CHILE

FACULTAD DE MATEMÁTICAS
DEPARTAMENTO DE ESTADÍSTICA

VALUE ADDED IN HIERARCHICAL LINEAR MIXED MODEL WITH ERROR IN VARIABLES

By

MAYO LUZ POLO GONZÁLEZ

Tesis presentada al Departamento de Estadística, Facultad de Matemáticas de la Pontificia Universidad Católica de Chile, para optar al título de Doctor en Estadística

Advisor
Ernesto Javier San Martín Gutiérrez, Ph.D.
DOCTOR EN ESTADÍSTICA

July
Santiago, Chile

Dedication

To my loved daughter

Valentina

Acknowledgements

This work would not have been possible without the support of all the people who accompanied me on this journey with many challenges and at the same time with the satisfaction that only the world of knowledge can offer.

I am grateful to my advisor Dr. Ernesto San Martín for his dedication, valuable theoretical contributions and useful comments regarding my research.

I also want to thank my mother Libia Rosa, my brothers Marco and Jairo for their constant unconditional support, especially in the most difficult moments of this stage of my life and I also thank to all the members of my family.

In memory of my brother Isma who was and will be a very special person in my life, although physically he is not with me, he will be always be present at every moment of my life.

I would like to thank Julio Montoya, my beloved companion for his great patience and support in this stage of my life.

Many thanks to my friends Martha Patricia Bohórquez, Luis Fernando Grajales and my Chilean friend Carla Leal for their unconditional support.

I would especially like to thank María Soledad Alcaino, Secretary of Postgraduate studies at the Mathematics faculty of the Pontificia Universidad Católica de Chile for her unconditional support during the time I was a graduate student in the country of Chile.

Finally, I would like to thank the professors and administrators of the Departments of Statistics of the Universities Pontificia Universidad Católica de Chile and the Universidad Nacional de Colombia for their support and for believing in my capacities to carry out this life project.

Contents

List of Tables

List of Figures

CHAPTER 1

Introduction

Higher education has to be a relevant and necessary issue in the agenda of all governments because it permits life transformation, poverty eradication, peace and social justice (UNESCO, 2015) [44]. In the last two decades higher education has undergone major changes such as: First, a substantial increase in student enrollment which means that higher education has gone from being an elitist education to an education for the masses (OECD, 2012, pp. 141) [32]. Second, the new knowledge economy which includes new fields of study in sciences, communication and technology. Third, the international mobility of students and professionals. Fourth, the new demands in the labor market (UNESCO, 2009, pp. 1-8) [33].

Some Latin American governments have recognized the importance of higher education in the progress of their countries and have made efforts to transform their education systems so they can respond to the challenges of social, cultural, economic and scientific development.

Discussion regarding higher education has focused on quality because this provides skilled professionals with significant contributions to the development of the economy and of society itself. Therefore, new quality assessment procedures are required to obtain information regarding the performance of institutions and the learning progress of their students.

It is important to mention the role played by the Organization for Economic Cooperation and Development (OECD) in encouraging, the use of measurements of Value Added for evaluating the quality of Higher Education in Colombia. The OECD and the World Bank's report represents a rigorous study regarding the topic of the Colombian Higher Educational System. This study provides an analysis pertaining to the achievements of the educational system over the last decade and the challenges that Colombia faces for educating its citizens with high professional skills so as to be able to resolve the economic and social problems of this country. As a result of this research, these agencies set forth some recommendations regarding the access, the equity, the research and the funding resources of Higher Educational system in Colombia.

Moreover, this study concludes that Value Added is an important indicator for measuring the quality of academic programs and the contribution made by the educational institutions to the students' academic improvement. Also, the estimation of Value Added permits knowing how the institutions have invested their resources and thus they will improve their level of responsibility. The OECD and the World Bank's report strongly recommends the Value Added as a measurement for the evaluation of the education system and for policy making decisions. (OECD, 2012, pp. 312) [32]

Goldstein(1997) [20] asserts that "School Effectiveness is a term used in educational research for exploring the differences between and within schools". The main objective of School Effectiveness is to study the relationship between response and explanatory variables using appropriate statistical models. Since the 1990s, researchers have had an increasing interest regarding performance indicators in terms of students' learning results, characteristics related to institutions and individual students.

In addition, School Effectiveness attempts to identify the factors explain the differences among institutions. One of these factors is known as Value Added which measures the contribution made by the institutions to the students' academic progress. Value Added is estimated using Linear Hierarchical (or Multilevel) Models due to hierarchical structure of educational data. These statistical models analyze the relationship between the response and explanatory variables taking into account that students are grouped by institutions. See Goldstein (2011) [21] and Snijders and Bosker (1999) [43] for broad revision of Multilevel Models.

Finally, it is important to mention that there is a relevant issue concerning measurement errors in the variables of the model. It is well known in the literature that if we ignore the measurement error of the variables in the model then the inferences made from these models can be misleading and not reliable. The two main references regarding measurement error are: Carroll et.al (2006) [6] , measurement error in generalized linear models or nonlinear models and Fuller (1987) [17], measurement error in linear models.

1.1 The Colombian Higher Education System

The higher education system in Colombia is regulated by Law 30 of 1992. This law focuses mainly on expanding higher education and assuring the quality of education. The main governmental agency in charge of higher education in Colombia is the Ministry of Education. This Ministry organizes and manages the plans, projects and policies at this educational level.

The Colombian tertiary education system comprises all post-secondary level of education. Also, there are two types of post-secondary education: university and professional technical education. The higher education institutions can be grouped into four types of institutions:

- The universities that offer undergraduate and graduate academic programs, master's degrees and doctorates.
- The university institutions that offer undergraduate programs and a postgraduate program called specialization.
- The technological institutions that offer programs at professional technical level (on a scientific basis).

- The professional technical institutions that offer undergraduate programs at technical and professional levels. (See OECD, 2012, pp. 32) [32]

In Colombia, two agencies are in charge of the process of quality assurance of education: The National Intersectoral Commission for Quality Assurance in Higher Education (CONACES, Comisión Nacional Intersectorial de Aseguramiento de Calidad en la Educación Superior), whose main function is to advise the Ministry of Education on what programs deserve to be included in the Qualified Registry; and the National Accreditation Commission (CNA, Comisión Nacional de Acreditación) which advises the Ministry of Education as to which higher education institutions and their programs must be awarded the high-quality accreditation (See OECD et al., 2012, pp. 177-178) [32].

The Colombian Institute for the Evaluation of Education (ICFES, Instituto Colombiano para la Evaluación de la Educación) is another agency that supports the process of the quality of education. This agency organizes the evaluations at all levels of education. It also offers economic support for investigations regarding the subject concerning improvement of the quality of education. In addition, the ICFES designs two important tests: The SABER 11 and the SABER PRO tests that permit measuring the students' progress at several levels of education.

The SABER 11 test evaluates the following knowledge areas: Biology, Social Sciences, Philosophy, Physics, English, Language, Mathematics and Chemistry. This test attempts to measure the students' competences in 11th grade of secondary education. The SABER 11 test is compulsory for students to be admitted to a university. Moreover, this test gives relevant information concerning Value Added indicators for higher education.

The SABER PRO test is an examination that evaluates students' competences in the last two years of higher education. This test focuses on two competences: Generic and specific. The generic competences are the minimun skills that all students have to develop while the specific competences are related to their field of training. The SABER PRO test also provides information for Value Added studies.

In Chapter 4, we present a more detailed revision concerning the higher education system in Colombia.

1.2 Motivation

The main motivation for this dissertation is to propose a statistical methodology that permits studying, analyzing and discussing the effects on Value Added for higher education institutions in the models with variables presenting low reliability (measurement error).

It is well known in the literature, that a good Value Added analysis has to include two students' examinations: Prior to the students' enrollment in an educational institution (Pre-Test) and after their enrollment in an institution, that is, when they are finishing their studies at that institution (Post-Test) (Kim and Lanlacette, 2013) [23].

The Colombian case is the only one in the world that presents the two tests mentioned above. Thus, the Colombian Higher Education System database provides us valuable information for the study concerning Value Added for educational institutions.

However, the Pre-Test (SABER 11 test) presents reliability problems, that is, it presents measurement errors. This problem was identified in the study undertaken by the OECD and it was an important conclusion put forth in this study. This situation is well argued and analyzed in the report presented by the OECD and the World Bank. Furthermore, it is important to mention that the situation was fundamental for proposing the research problem of this thesis. (See OECD, 2012, pp 189-191) [32].

The SABER 11 test is an examination that Colombian students take at the end of their secondary education. This exam is compulsory.

Table 1.1 summarizes the average reliability for each knowledge area of the SABER 11 test (See OECD, 2012, pp 190) [32]. According to the educational experts, these reliability values are lower because the SABER 11 test is a cumulative test and it is the only criterion for admission to a university. Also, the low reliability of the SABER 11 test implies that students can be incorrectly classified with respect to admission to a specific university because the students do not obtain their true score but rather have an observed score with measurement error.

TABLE 1.1. Average Reliability of the SABER 11 Test

Knowledge Area	Average Reliability
Biology	0.715
Social Sciences	0.783
Philosophy	0.663
Physics	0.555
English	0.795
Language	0.715
Mathematics	0.725
Chemistry	0.640

In Value Added models, if we do not correct the measurement error, we may incorrectly classify educational institutions because the Value Added estimators will be biased and therefore we might assert that an institution is not contributing to the students' progress when it is actually is, or on the contrary that an institution is contributing to the progress of students when it is actually not.

Finally, we stress the importance of adjusting or correcting measurement errors in the statistical model to solve the problems of low reliability of variables included in the model. If we do not correct the measurement error, we can get biased estimators and misleading conclusions which can have serious implications.

1.3 Measurement Error in Educational Data

The first authors who worked on the subject concerning measurement error in Linear Hierarchical Models or Multilevel Models were Woodhouse, Yang, Goldstein and Rasbash (1996) [47], assumed a two level model with the students nested within the schools. They also considered that the explanatory variables and response variable are measured with error. Fixed parameters and the variances are estimated using a method proposed by Woodhouse (1996) [46]. In this paper, the method requires prior estimates of measurement error variances and measurement error covariance between two variables. Whereas in the work of Woodhouse et al. (1996) [47], the expressions for variance and covariance of measurement error are derived. Moreover, the effects on parameters estimates for several values of measurement error are shown. These values of measurement error are expressed in terms of reliability of explanatory and response variables.

An important and complete reference regarding estimation in longitudinal random effects models with measurement error is Buonaccorsi et.al (2000) [4]. These authors consider a longitudinal mixed model with measurement error in a variable changing with time. They propose estimation methods that achieve the efficiencies of the estimators. Also, they present the mathematical developments concerning the statistical model and the model for measurement error. Also, they study extensions of the model and the asymptotic properties of the proposed estimators.

Another interesting work regarding estimation in the errors in variables models is described in Cui et.al (2004) [8]. In this paper, the authors assume a linear mixed effects model with measurement error in the variables. The estimators for the fixed parameters, the covariance matrix of random effects and the variance for model error are found by the method of moments. These authors also proved from the theoretical point of view the asymptotic properties of the obtained estimators in their research. They do not discuss the prediction of random effect. The authors assert that the maximum likelihood estimators and the estimators of moments may lose the consistency property if they do not take into account in the model the measurement error.

A most extensive and complete work regarding adjusting measurement error in multilevel models is discussed in Goldstein et.al (2008) [22]. They assume a multilevel model with random coefficients. They also consider two types of measurement error: measurement error in continuous variables and misclassification in categories. (That is, the individual can be classified into one or several categories with correct or incorrect probabilities). The parameters are estimated by the Bayesian estimation using Markow Chain Monte Carlo (MCMC Algorithm). For more details concerning this algorithm, see Goldstein (2011, Chapter 2) [21] Moreover, they study several extensions of the general model as such as correlated measurement error and binary variable and ordered category variables.

A more recent reference concerning measurement error adjustment for multilevel models having application to educational data is Battauz et.al (2011) [2]. These authors consider a two level model that includes class random effects for the intercept, the explanatory variable is the intake achievement and the response variable is the achievement in subsequent periods. Moreover, they propose a likelihood based measurement error adjustment for multilevel models.

1.4 Organization of the Thesis

This dissertation is organized as follows:

In the first chapter, we review some changes that higher education in Latin America has undergone in the last two decades and we refer to some efforts made in education by Latin American governments. This chapter also examines the concepts of school effectiveness and Value Added models. It ends with an overview concerning the methods used in the estimation in multilevel models with measurement error.

In chapter two, we propose a methodology for evaluating the effects on the Value Added estimates in Hierarchical Linear Mixed or Multilivel Models with measurement error in the variables. We begin with the model specification that includes the hypotheses regarding the variables and measurement error. We also discuss parameter identification taking into account measurement error.

In chapter three, we present several theoretical developments for estimating the parameters by means of maximum likelihood and we adapt the Bootstrap procedure to the multilevel model with measurement error. We illustrate the proposed methodology with a simulation study and discuss the effects on Value Added estimates for several values of reliability.

In chapter four, we illustrate the proposed methodology with an application to the Colombian case. We use a database of Colombian universities and we analyze the implications on the Value Added estimates when the measurement error is considered or when not.

In chapter five, we discuss the more relevant conclusions regarding the proposed methodology which deal with the issue of measurement error in multilevel models. We also propose some extensions and improvements for future works.

CHAPTER 2

Model Specification and Parameter Identification

2.1 Introduction

In this chapter, we focus on the model specification, parameter identification and definition of Value Added in a hierarchical linear mixed model with error in explanatory variables. In particular, we work with a two level model: Level one units correspond to the students and level two units correspond to the universities.

Moreover, we present the technical developments in model specification, the structural model and the definition of Value Added in a hierarchical linear mixed model with error in variables. We propose several hypotheses that describe the relationship between the observable and the unobservable variables included in the model. We build the structural model based on these hypotheses as well as the assumptions concerning the parameters and variables involved in the model. Also, we discuss the parameter identification problem in terms of the reliability of explanatory variables measured with error.

Finally, we obtain a mathematical expression for the Value Added for a school or university in a hierarchical linear mixed model with error in variables using the definition of Value Added proposed in Manzi et.al (2014) [28].

2.2 Notation

We introduce the notation that will be used in this and the following chapters. Let Y_{ij} be the response variable of the *i*-th level one unit that belongs to the *j*-th level two unit, e.g., the mathematics score of a student i that belongs to a school j. With $i = 1, 2, ...n_j$ (n_j is the total number of students at the school or university j) and $j = 1, 2, ..., J$ (J is the total number of universities or schools).

η_{ij} is a $K \times 1$ unobservable true variables vector, i.e., variables which cannot be directly observed such as students' abilities or aptitudes. X_{ij} is a $K \times 1$ explanatory variables vector of the *i*-th level one unit that belongs to the *j*-th level two unit where each component is a variable measured with error and is related to the unobservable variable η_{ij}.

The effect of the *j*-th level two unit is denoted by θ_j, in this case, it is a scalar. At the school level, we write the following notation: $Y_j = (Y_{1j}, Y_{2j}, ..., Y_{n_j j})^T$ is an $n_j \times 1$ vector, $X_j = (X_{1j}, X_{2j}, ..., X_{n_j j})^T$ is an $n_j \times K$ matrix and $\eta_j = (\eta_{1j}, \eta_{2j}, ..., \eta_{n_j j})^T$ is an $n_j \times K$ matrix. Also, at the population level, we can extend the notation as $Y = (Y_1^T, Y_2^T, ..., Y_J^T)^T$ is an $N \times 1$ vector with $N = \sum_{j=1}^{J} n_j$, $X = (X_1^T, X_2^T, ..., X_J^T)^T$ is an $N \times K$ matrix,$\theta = (\theta_1, \theta_2, \ldots \theta_J)^T$ is a $J \times 1$ vector and $\eta = (\eta_1^T, \eta_2^T, \ldots \eta_J^T)^T$ is an $N \times K$ matrix.

2.3 Hierarchical Linear Mixed Model with Error in Variables

2.3.1 Model Specification

Based on the theory regarding school effectiveness, the classical test theory and the above notation, we denote Y_{ij} as the score of student i at university j, X_{ij} is a K-dimensional vector of explanatory variables measured with error, these variables are measured at the student level or at the university level. η_{ij} is a K-dimensional vector of unobservable true covariables, that is, each component (η_{kij}) of the vector η_{ij} is a true variable, $k = 1, 2, ..., K$.

As it is known from the school effectiveness theory, the heterogeneity of the students' scores ($Y_{ij}'s$) is explained by the observable explanatory variables (or covariates) (X_{kij}) and an effect at the school level which is unobservable and is called the school effect (θ_j).

The structural model assumes the following hypotheses:

H1: Independence between Y_j and X_j conditional on θ_j and η_j. The response vector is determined by θ_j and η_j, i.e., $Y_j \perp\!\!\!\perp X_j | \theta_j, \eta_j$.

H2: The axiom of local independence associated with the response variable. This asserts that the students' responses are independent of each other conditionally on the true variable and the school effect, that is, $\underset{i}{\perp\!\!\!\perp} Y_{ij} | \eta_j, \theta_j \quad \forall\, i = 1, 2, 3 ... n_j$.

H3: Independence between X_j and θ_j conditional on η_j. The explanatory variable vector is determined by η_j,that is, $X_j \perp\!\!\!\perp \theta_j | \eta_j$.

H4: Independence between X_{ij} and η_j conditional on η_{ij}. The X_{ij} vector is characterized by η_{ij}.i.e.,$X_{ij} \perp\!\!\!\perp \eta_j | \eta_{ij}$

H5: Exogeneity. The school effect and the unobservable variables vector are mutually exogeneous, i.e., $\theta_j \perp\!\!\!\perp \eta_j$.

In this study, we assume that the school effect is not correlated with the observable covariates (X_{kij}). The model structure is linear, that is, the expectation of score Y_{ij} depends on true covariates and the school effect and this expectation of score is linearly dependent on the true covariates and the school effect, this relationship is represented by

$$E(Y_{ij} | \eta_{ij}, \theta_j) = \alpha + \eta_{ij}^T \beta + \theta_j \tag{2.1}$$

where $\alpha \in \mathbb{R}$

Structural Model

The multidimentional model for university j is written as

$$(Y_j|\eta_j, \theta_j) \sim \mathcal{N}(\alpha 1_{n_j} + \eta_j \beta + \theta_j 1_{n_j}, \sigma^2 I_{n_j}) \tag{2.2}$$

where $\alpha \in \mathbb{R}$, 1_{n_j} is an $n_j \times 1$ vector of ones, η_j is an $n_j \times k$ matrix, β is a $k \times 1$ vector, θ_j is a scalar, $\sigma^2 > 0$, I_{n_j} is the $n_j \times n_j$ identity matrix.

And the multidimentional model for all universities is given by

$$(Y|\eta, \theta) \sim \mathcal{N}(\alpha 1_N + \eta \beta + \theta 1_N, \sigma^2 I_N) \tag{2.3}$$

With

$Y = (Y_1^T, Y_2^T, ..., Y_J^T)^T$ is an $N \times 1$ vector with $N = \sum_{j=1}^{J} n_j$, $X = (X_1^T, X_2^T, ..., X_J^T)^T$ is an $N \times K$ matrix, $\theta = (\theta_1, \theta_2, \dots \theta_J)^T$is is a $J \times 1$ vector and $\eta = (\eta_1^T, \eta_2^T, \dots \eta_J^T)^T$ is an $N \times K$ matrix.

Additionally, we use several above hypotheses and we enunciate the following propositions:

Proposition 1 We assert that $X_j \perp\!\!\!\perp \theta_j$. This follows from hypotheses H3 and H4, i.e., $X_j \perp\!\!\!\perp \theta_j | \eta_j \wedge \theta_j \perp\!\!\!\perp \eta_j \Rightarrow (X_j, \eta_j) \perp\!\!\!\perp \theta_j \Rightarrow X_j \perp\!\!\!\perp \theta_j$.

Proposition 2 We state that $Y_j \perp\!\!\!\perp X_j | \eta_j$. It follows from hypotheses H1 and H3, that is, $Y_j \perp\!\!\!\perp X_j | \theta_j, \eta_j \wedge X_j \perp\!\!\!\perp \theta_j | \eta_j \Rightarrow (Y_j, \theta_j) \perp\!\!\!\perp X_j | \eta_j \Rightarrow Y_j \perp\!\!\!\perp X_j | \eta_j$.

On the other hand, based on the classical test theory, a model of measurement error is given by

$$X_j = \eta_j + u \tag{2.4}$$

Where $vec(\eta_j)$ is a $Kn_j \times 1$ true variables vector with distribution

$$vec(\eta_j) \sim \mathcal{N}_{Kn_j}(\mathbf{0}_{Kn_j \times 1}, \Psi_{K \times K} \otimes I_{n_j}) \tag{2.5}$$

With

$$\Psi = \begin{pmatrix} \psi_1^2 & \psi_{12}^* & \cdots & \psi_{1K}^* \\ \psi_{21}^* & \psi_2^2 & \cdots & \psi_{2K}^* \\ \vdots & \vdots & \ddots & \\ \psi_{K1}^* & \psi_{K2}^* & \cdots & \psi_K^2 \end{pmatrix}$$

and $vec(u)$ is a $Kn_j \times 1$ measurement error vector with distribution

$$u \sim \mathcal{N}(0, T_{K\times K} \otimes I_{n_j}) \tag{2.6}$$

where $T_{K\times K} = \text{diag}(\tau_1^2, \tau_2^2, ..., \tau_K^2)$

Based on the equations 2.4 to 2.6 we have

$$(X_j|\eta_j) \sim \mathcal{N}(\eta_j, \Psi_{K\times K} \otimes I_{n_j}) \tag{2.7}$$

Moreover, we assume that the school effects are independent of each other with normal distribution, that is,

$$\theta_j \sim \mathcal{N}(0, \omega^2)$$

Proposition 3 The correlation between a pair of true variables (η_{kij}) induces a correlation between the corresponding pair of observable variables (X_{kij}), that is,

$$Cov(X_{kij}, X_{k'ij}) = Cov(\eta_{kij}, \eta_{k'ij}) \neq 0 \; \forall \; k = 1, 2, ..., K$$

Proof of the above.

$$\begin{aligned} Cov(X_{kij}, X_{k'ij}) &= Cov[E(X_{kij} \mid \eta_{kij}), E(X_{k'ij} \mid \eta_{k'ij})] + E[Cov(X_{kij}, X_{k'ij} \mid \eta_{kij})] \\ &= Cov(\eta_{kij}, \eta_{k'ij}) \\ &\neq 0 \end{aligned}$$

Remark Based on Proposition 3, we can conclude that the covariance of the observable variables (X_{kij}) is derived from the covariance of the unobservable variables(η_{kij}). This important result implies that the linear correlation between two different observable variables will depend on the linear correlation between their correspondent two unobservable variables. In particular, if we assume on one side that there is measurement error of the variables (X_{kij}) $\forall k = 1,2,...,K$ and on the other side that the covariance of the unobservable variables is different from zero then the covariance of both of these observable variables and that of the measurement errors will play an important role for obtaining estimators. Therefore, the measurement error will be a relevant variable that has to take into account for obtaining good estimators and valid conclusions.

2.3.2 Parameter Identification

Parameter identification is important in the inference procedure because if the parameters are not identified then we cannot obtain consistent estimators. The parameters of interest in our model are $\alpha \in \mathbb{R}, \beta \in \mathbb{R}^K$ (fixed parameters), the variances $\sigma_k^2, \tau_k^2, \psi_k^2, \omega^2$ and the covariances $\psi_{kk'}$, $\forall\ k \neq k'$ and $k = 1,2,...,K$. Thus, the total number of parameters is equal to $\frac{6(K+1)+K(K-1)}{2}$, where K is the number of explanatory variables included in the model.

We examine the identification of the parameters mentioned above, based on the probability distribution of observable variables, in particular, the joint distribution of vector $(Y_j, \text{vec}(X_j))^T$. Based on the structural model given in Equation 2.2 and hypotheses H1, H2, H3, H4 and H5, we assert that the $(Y_j, \text{vec}(X_j))^T$ is distributed $n_j + Kn_j$ multivariate normal.

The following theorem provides the distribution of $(Y_j, \text{vec}(X_j))^T$.

Theorem 1: Given the structural model expressed in Equation 2.2 and the hypotheses H1,H2,H3,H4 and H5 of the structural model then the joint distribution of $(Y_j, \text{vec}(X_j))^T$

is $n_j + Kn_j$ multivariate normal, that is,

$$\begin{pmatrix} Y_j \\ \mathrm{vec}(X_j) \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \alpha 1_{n_j} \\ \mathbf{0}_{Kn_j\times 1} \end{pmatrix}; \begin{pmatrix} (\phi+\sigma^2)I_{n_j} + \omega^2 J_{n_j} & \beta^T\Psi \otimes I_{n_j} \\ \Psi^T\beta \otimes I_{n_j} & (T+\Psi)\otimes I_{n_j} \end{pmatrix}\right]$$

Where

1_{n_j} is an $n_j \times 1$ vector of ones, $\mathbf{0}_{Kn_j\times 1}$is a $Kn_j \times 1$ vector of zeros $J_{n_j} = 1_{n_j}1_{n_j}^T$, $\otimes$ is a Kronecker product, $\phi := \beta^T\Psi_{K\times K}\beta = \sum_{k=1}^{K}\beta_k^2\Psi_k^2 + \sum\sum_{k\neq k'}\beta_k\beta_{k'}\Psi_{k,k'}$ and the matrix $\Psi_{K\times K}$ is given by

$$\mathrm{V}(\eta_{ij}) = \Psi = \begin{pmatrix} \psi_1^2 & \psi_{12}^* & \cdots & \psi_{1K}^* \\ \psi_{21}^* & \psi_2^2 & \cdots & \psi_{2K}^* \\ \vdots & \vdots & \ddots & \\ \psi_{K1}^* & \psi_{K2}^* & \cdots & \psi_K^2 \end{pmatrix}$$

And

$V(X_{ij} \mid \eta_{ij}) = \mathrm{diag}(\tau_1^2, \tau_2^2, ..., \tau_K^2) = T_{K\times K}$, with $\psi_k^2 = V(\eta_{kij}), \psi_{kk'} = cov(\eta_{kij}, \eta_{k'ij}), \forall k \neq k'$ and $\tau_k^2 = V(X_{kij})$. As seen in previous sections that $\sigma^2 = V(Y_{ij}|\eta_{ij}, \theta_j)$, $\omega^2 = V(\theta_j)$, I_{n_j} is the $n_j \times n_j$ identity matrix and β is a $k \times 1$. Further details are described in Appendix A2, A3 and A4.

As well is known from the theory of parameter identification, the expectation and variance of a normal multivariate are identified and from Theorem 1 follows.

1. Based on $E(Y_j)$, parameter α is identified.

2. Based on $V(Y_{ij})$ and $Cov(Y_{ij}, Y_{i'j})$ for $i \neq i'$. It follows that $\beta^T\Psi_{k\times k}\beta + \omega^2 + \sigma^2$ and ω^2 are identified.

3. Based on $V(\mathrm{vec}(X_j))$ $T_{K\times K} + \Psi_{K\times K}$ is identified, therefore $T_{K\times K}$ and $\Psi_{K\times K}$ are identified because the first is a diagonal matrix and the second is a correlation matrix.

4. Based on $Cov(Y_j, \mathrm{vec}(X_j))$ $\beta^T\Psi_{K\times K}$ is identified and with the identification of $\Psi_{K\times K}$ it follows that β is identified provided that $r(\Psi_{K\times K}) = K$.

5. σ^2 is identified because ω^2 and $\beta^T\Psi_{K\times K}\beta$ are identified.

2.3.2.1 Further discussion regarding parameter identification

Continuing with the discussion regarding parameter identification, it is necessary, in some situations, to consider some constraints on these parameters. We examine these constraints in terms of the reliability of the variable measured with error. Remember that reliability (R) is defined as the quotient between the variance of the true variable(η_{kij}) and the variance of the observed variable (X_{kij}), that is, $R = \frac{\psi_k^2}{\psi_k^2+\tau_k^2}$.

In order to determine which constraint is the most appropiate, we analyze the following two cases that relate each parameter τ_k^2 and ψ_k^2 to the reliability of each variable measured with error.

1. We assume that $\tau_k^2 = 1$ then the reliability will be equal to $\frac{1}{1+\frac{1}{\psi_k^2}}$ $\forall k = 1,2,...,K$.The reliability will be close to 1 if $\psi_k^2 = \infty$, that is, the variance of the true variable is infinity which means that the variable takes large values, which are far from its mean and therefore true variable distribution will be a heavy tailed distribution; certainly, this disproves the assumption concerning the normality of the true variable which is stated in the previous sections. On the other hand, in practice, it is not possible to consider variables with infinite variance as this is invalid from the statistical point of view, the inferences on the parameters are established under the assumption of finite variance.

2. If $\psi_k^2 = 1$ then the reliability is given by $\frac{1}{1+\tau_k^2}$.The reliability is approximately equal to 1 if $\tau_k^2 = 0$, that is, $\tau_k^2 = V(X_{kij}|\eta_{kij}) = 0$ and that is true if and only if $X_{kij} = f(\eta_{kij})$ a.s. Indeed, in this study the variable X_{kij} is related with η_{kij}, this relationship is written in Equation 2.4.

We consider that result 2 assumes that $\psi_k^2 = 1$ is more appropriate with the concept of perfect reliability because the variance of the measurement error is zero, implying that there is no error. Therefore, the observed score will be equal to the true score and the reliability will be equal to 100%. Therefore, it is the strategy we use to improve the restriction regarding these parameters.

Having this new restriction concerning ψ_k^2, the variance-covariance matrix $\Psi_{K\times K}$ becomes a correlation matrix and can be re-written as:

$$\Psi_{K\times K}=\begin{pmatrix} 1 & \rho_{12} & \dots & \rho_{1K} \\ \rho_{21} & 1 & \dots & \rho_{2K} \\ \vdots & \vdots & \ddots & \\ \rho_{K1} & \rho_{K2} & \dots & 1 \end{pmatrix}$$

where $\rho_{kk'}$ is the correlation between η_{kij} and $\eta_{k'ij}$ $\forall\ k \neq k'$.

2.4 Value Added

2.4.1 Definition of Value Added

In order to introduce a definition of Value Added, first it is important to consider the following decomposition for the response variable Y_{ij}.

$$Y_{ij} = \underbrace{E(Y_{ij}|X_{ij})}_{component\ 1} + \underbrace{\{E(Y_{ij}|X_{ij},\theta_j) - E(Y_{ij}|X_{ij})\}}_{component\ 2} + \underbrace{\{Y_{ij} - E(Y_{ij}|X_{ij},\theta_j)\}}_{component\ 3} \tag{2.8}$$

Based on the decomposition 2.8, we observe that:

1. The three components are orthogonal by construction, this means, they are uncorrelated.

2. The first component corresponds to the part of the response variable that it is explained by the covariates X_{ij}.

3. The second component describes the part of the response variable that it is explained by the school (θ_j) after controlling by the covariates X_{ij}.

4. The third component represents the part of the response of variable that it is not explained neither by the school (θ_j) nor by the covariates X_{ij}.

Therefore, from decomposition 2.8 the Value Added is defined as an average of the part that depends on the school, that is, the average of the second component. This definition of Value Added is proposed in Manzi et.al (2014) [28] which is defined as the difference between the average of the conditional expectation for the response variable given both the school effect and the explanatory variables (covariates) and the average of the conditional expectation of the response variable conditional on the covariates only, that is,

$$VA_j = \frac{1}{n_j}\{\sum_{i=1}^{n_j}[E(Y_{ij}|X_{ij},\theta_j) - E(Y_{ij}|X_{ij})]\} \tag{2.9}$$

From definition 2.9 it is important to emphasize that:

- Value Added depends on the covariates through conditional expectations. Moreover, the Value Added can be different and it will depend on the covariates included in the model. Therefore, the selection of the covariates must take into account the objectives of public policy for which a Value Added analysis is required.
- Value Added measures the gain or loss of a given school regarding an average or reference school.
- The concept of "average or reference school" is determined by the second average conditional expectation because it is an average of the response variable regarding the school effect distribution, i.e,

$$E(Y_{ij}|X_{ij}) = E[E(Y_{ij}|X_{ij},\theta_j)|X_{ij}] \tag{2.10}$$

Thus, $E(Y_{ij}|X_{ij})$ is an average of the response variable in a reference school. For a more extensive discussion about Value Added see Raudenbush et.al (2004) [35] and Goldstein(1997) [20].

2.4.1.1 The role of the covariates in Value Added analysis

It is worth to emphasize the fundamental role of the covariates in the Value Added analysis. The interpretation and conclusion obtained from the Value Added analysis are determined by the set of covariates considered in a study. The covariates are characteristics measured or observed at the school level and at the student level. A covariate used in Value Added studies is the student's prior score which is expected to be highly related to contemporaneous score (Response variable). (Goldstein, 2011) [21]. Additionally, the reference school is characterized by these covariates which are allowing to compare the schools under the same observable conditions.

If the interest is to compare schools taking into account some selectivity process of the school, it is important to consider a covariate that describes the selection. For example, schools with high socioeconomic level usually attract students with high scores. Therefore, the comparison among schools cannot be fair. In this case, a variable related to the student's socioeconomic level must be included in the model.

Value Added under the Hierarchical Linear Mixed Model with Error in Variables

In the model with K-explanatory variables measured with error, we write the first conditional expectation as

$$E(Y_{ij}|X_{ij},\theta_j) = \alpha + E(\eta_{ij}^T\beta|X_{ij}) + \theta_j \tag{2.11}$$

and the second conditional expectation is given by

$$E(Y_{ij}|X_{ij}) = \alpha + E(\eta_{ij}^T\beta|X_{ij}) \tag{2.12}$$

Recall that $\alpha, \theta_j \in \mathbb{R},\ \beta \in \mathbb{R}^K$ and η_{ij} is an $K \times 1$ vector. Further calculations are shown in appendices A5 and A6.

Substituting (2.11) and (2.12) into (2.9), we have that the Value Added is equal to the school effect, that is,

$$\boxed{VA_j = \theta_j} \tag{2.13}$$

Remark Equation (2.12) is important because the measurement error is typically considered as a problem of endogeneity implying the Value Added has been corrected, that is to say, the Value Added is equal to the school effect plus an additional term. See

Manzi et.al (2014) [28]. Nevertheless, in this case, that it assumes a measurement error in the explanatory variables, the value added does not require this correction because it is only equal to the school effect.

2.4.2 Prediction of Value Added

As we have seen in the previous section, in order to estimate the Value Added of school j we need to predict the school effect θ_j. Moreover, it has been stablished in the literature that the school effect is an unobservable random variable and the empirical Bayes estimator is used for predicting it, that is, $\hat{\theta}_j = E(\theta_j|Y_j,X_j)$.
Where

$$E(\theta_j|Y_j,X_j) = E(\theta_j) + Cov(\theta_j,(Y_j,X_j))\left[V\begin{pmatrix}Y_j\\X_j\end{pmatrix}\right]^{-1}\left[\begin{pmatrix}Y_j\\X_j\end{pmatrix} - \begin{pmatrix}E(Y_j)\\E(X_j)\end{pmatrix}\right] \quad (2.14)$$

Prediction of Value Added under the Hierarchical Linear Mixed Model with Error in Variables

We use Equation 2.13 for obtaining the prediction of Value Added under the Hierarchical Linear Mixed Model with Error in Variables, this prediction is given by

$$\boxed{\hat{\theta}_j = \begin{pmatrix}\omega^2 1_{n_j}^T & 0 1_{n_j}^T\end{pmatrix}\left[V\begin{pmatrix}Y_j\\\mathrm{vec}(X_j)\end{pmatrix}\right]^{-1}\begin{pmatrix}Y_j-\alpha 1_{n_j}\\\mathrm{vec}(X_j)\end{pmatrix}} \quad (2.15)$$

where
$Cov(\theta_j,(Y_j,X_j)) = \begin{pmatrix}\omega^2 1_{n_j}^T & 0 1_{n_j}^T\end{pmatrix}$, $E(Y_j) = \alpha 1_{n_j}$ and $E(\mathrm{vec}(X_j)) = 0_{Kn_j\times 1}$

Note that we re-write the variance of $(Y_j, vec(Xj))^T$ of Equation 2.14 in terms of the reliability. As it is known, from Theorem 1, matrix $T+\Psi$ is a submatrix of that variance.

This submatrix is given by

$$\begin{pmatrix} \tau_1^2+1 & \rho_{12} & \dots & \rho_{1K} \\ \rho_{21} & \tau_2^2+1 & \dots & \rho_{2K} \\ \vdots & \vdots & \ddots & \\ \rho_{K1} & \rho_{K2} & \dots & \tau_K^2+1 \end{pmatrix}$$

It can be re-written as

$$\begin{pmatrix} 1/R_1 & \rho_{12} & \dots & \rho_{1K} \\ \rho_{21} & 1/R_2 & \dots & \rho_{2K} \\ \vdots & \vdots & \ddots & \\ \rho_{K1} & \rho_{K2} & \dots & 1/R_K \end{pmatrix}$$

Where the diagonal elements are obtained by using the definition of reliability. As seen before in this study the reliability is defined as $R_k = \frac{\psi_k^2}{\psi_k^2+\tau_k^2}$, with $\psi_k^2 = 1, \forall\, k = 1,2...,K$.

From the above, we conclude that the Value Added depends on the reliability of variables measured with error and at the same time, this reliability is associated with the measurement error. Therefore, if we do not correct or do not take into account the measurement error, the Value Added estimates are incorrect, especially for small reliability values . On the other hand, for reliability values nearly equal to one, the model with error in variables is similar to a model without error in variables.

Finally, we estimate the value added, using Equation 2.14 and we replace each parameter with its corresponding estimator which is obtained by Maximum Likelihood.

CHAPTER 3

Maximum Likelihood Estimation and Bootstrap Procedure

3.1 Introduction

In this chapter, we discuss the maximum likelihood estimation method in order to obtain the estimators of fixed parameters, variances and correlation coefficients under the Hierarchical Linear Mixed Model with Error in K explanatory variables. We also present a mathematical expression for the likelihood function associated with the conditional distribution of the response variable given the explanatory variables.

First we write the likelihood function in statistical package R. Next, we use the optim function from this package for optimizing the likelihood function and thus for finding the estimators of parameters. We also use some functions of the parameters to ensure that the estimates are not on the outside of the parameter space. For instance, for obtaining positive values for each variance parameter, we use the exponential function $\exp(x)$, that is, in the code, we write $\sigma^2 = \exp(x)$. For the correlation coefficients, we consider Fisher's transformation $\rho = \dfrac{\exp(2z)-1}{\exp(2z)+1}$, this transformation help's us to obtain values between -1 and 1 for these coefficients.

Moreover, we apply the Bootstrap procedure for a Hierarchical Linear Mixed Model following the methods proposed in Van der Leeden et.al (1997) [45]. In our study, we use the Bootstrap procedure for calculating the estimated variances for the prediction of Value Added. The Bootstrap procedure is used because we do not have any analytical expression for the variance of prediction of Value Added.

Finally, we evaluate the proposed methodology in a simulation study. In this study, we consider a model with four explanatory variables measured with error, that is, $K = 4$. We simulate the response and explanatory variables considering the assumptions of the model. In addition, we compare a model whose measurement error is taken into account with another model without measurement error.

3.2 Estimation by the Maximum Likelihood Method

In our model with K explanatory variables measured with error, the parameters $\alpha \in \mathbb{R}, \beta \in \mathbb{R}^K$ (fixed parameters), variances $\sigma_k^2, \tau_k^2, \omega^2$ and correlation coefficient $\rho_{kk'}$, $\forall\, k \neq k'$ and $k = 1,2,\ldots,K$, are estimated by maximum likelihood. The total number of parameters is $\frac{6(K+1)+K(K-1)}{2}$

As seen in Equation 2.2, Section 2.3, the multidimentional model for j-th school is given by

$$Y_j|\eta_j,\theta_j \sim N(\alpha 1_{n_j} + \eta_j\beta + \theta_j 1_{n_j}, \sigma^2 I_{n_j})$$

this model is equivalent to

$$Y_j = \alpha 1_{n_j} + \eta_j\beta + \theta_j 1_{n_j} + \varepsilon_j \tag{3.1}$$

with

$$\varepsilon_j \overset{ind}{\sim} N(0, \sigma^2 I_{n_j})$$

$$\theta_j \overset{ind}{\sim} N(0, \omega^2)$$

we also assume

$$\text{vec}(\eta_j) \overset{ind}{\sim} \mathcal{N}_{Kn_j}(\mathbf{0}_{Kn_j\times 1}, \Psi \otimes I_{n_j})$$

and

$$X_{kij} = \eta_{kij} + u_{kij}, \; \forall \; k = 1, 2, ..., K$$

with $u_{kij} \sim N(0, \tau_k^2)$.

As mentioned before variable X_{kij} is a variable measured with error.

Our aim is to estimate the parameters from the structural model given in Equation (2.2) which leads to a statistical model $p(Y|X,\theta)$ where $X_j \perp\!\!\!\perp \theta_j$ is a fundamental hypothesis (See Proposition 1, Chapter 2). Usually, in models with unobservable variables, we have to indirectly estimate the parameters based the conditional distribution of the response variable given the observable variable , that is, $p(Y_j|X_j,\theta_j)$.

We use the hypotheses proposed in Chapter 2, Section 2.3, and the properties of conditional expectation and conditional variance for determining the distribution of $Y_j|X_j,\theta_j$ which is an n_j multivariate normal, with the mean vector written as

$$E(Y_j|X_j,\theta_j) = \alpha 1_{n_j} + [\beta^T \Psi(\Psi + T)^{-1} \otimes I_{n_j}]\text{vec}(X_j) + \theta_j 1_{n_j}$$

And the covariance-variance matrix expressed as

$$V(Y_j|X_j,\theta_j) = \sigma^2 I_{n_j} + (\beta^T \Psi \beta \otimes I_{n_j}) - [\beta^T \Psi(\Psi + T)^{-1} \Psi^T \beta \otimes I_{n_j}]$$

In Appendices B1, B2 and B3, the detailed calculation of the above expressions is shown.

3.2.1 Likelihood Function

We introduce the statistical model associated with the observed data for each school j as

$$Y_j|\text{vec}(X_j) \sim \mathcal{N}_{n_j} \quad \forall j = 1, 2, ..., J$$

with

$$E(Y_j|\text{vec}(X_j)) = \alpha 1_{n_j} + (\beta^T \Psi \otimes I_{n_j})[(\Psi + T) \otimes I_{n_j}]^{-1}\text{vec}(X_j) \tag{3.2}$$

and

$$V(Y_j|\text{vec}(X_j)) = (\phi+\sigma^2)I_{n_j} + \omega^2 J_{n_j} - (\beta^T\Psi\otimes I_{n_j})[(\Psi+T)\otimes I_{n_j}]^{-1}(\Psi^T\beta\otimes I_{n_j}) \quad (3.3)$$

Let $f(y|x)$ be the density function of the student's scores of all the schools in the population and let $\mathit{\Gamma} = (\alpha,\sigma^2,\omega^2,\beta_k,\tau_k^2,\rho_{kk'})$ be the vector of parameters associated with this density function, $k = 1,2,...,K$.

Due to the independence of Y_j 's, the density function of $Y|X$ can be written as

$$f(Y|X;\Gamma) = \prod_{j=1}^{J} f(y_j|x_j;\Gamma)$$

The log-likelihood of $\mathit{\Gamma}$ denoted by $l(\mathit{\Gamma};y|x)$ is given by

$$\ell(\Gamma;(y|x) = -\frac{N}{2}\ln(2\pi) - \frac{1}{2}\sum_{j=1}^{J}\ln|V_j|$$
$$-\frac{1}{2}\sum_{j=1}^{J}[y_j-\alpha 1_{n_j}-(\beta^T\Psi(\Psi+T)^{-1}\otimes I_{n_j})\text{vec}(X_j)]^T[V_j]^{-1}[y_j-\alpha 1_{n_j}-(\beta^T\Psi(\Psi+T)^{-1}\otimes I_{n_j})\text{vec}(X_j)] \quad (3.4)$$

For simplicity, we denote $V(Y_j|\text{vec}(X_j)) := V_j$ which is given by Equation (3.3).

Note that the log-likelihood function (Equation 3.4) is a complex function which does not have an analytical solution. Therefore, it is necessary to use an optimization criterion for obtaining the estimators of parameters; in this case, we use the optim function of the statistical package R.

3.3 Bootstrap Procedure

The Bootstrap procedure introduced by Efron (1979) [12] enables us to obtain the variance of estimators and confidence intervals for parameters in situations in which an analytical expression of the variance does not exist or this variance has a complex expression which is not possible to solve analytically.

The main idea concerning this procedure is that the empirical distribution (F_n) associated with a sample which is denoted by $\{w_1, w_2, ...w_n\}$ is a consistent estimator of distribution F which is associated with the random variable W. See Rien Van der Leeden et.al (1997). That is, if $\hat{\theta}$ (Asocciated with F_n) is an estimator of θ (F distribution parameter) then, when the sample size increases to infinity, $\hat{\theta}$ will be a good approximation of θ. Thus, the Bootstrap procedure allows us to select B samples for F_n with replacement for calculating the different estimators denoted by θ_b^* $(\forall b = 1, 2, ..., B)$ which will be a good approximation of $\hat{\theta}$ given the consistency property of F_n. For a broader discussion, see Davison and Hinkley (1997) [9], Efron and Tibshirani(1994) [13] and Boos(2003) [3].

In the Bootstrap procedure, the variance estimator is given by

$$\hat{V}(\theta^*) = \frac{\sum_{b=1}^{B}(\theta_b^* - \bar{\theta}_b^*)^2}{B-1} \tag{3.5}$$

where $\bar{\theta}_b^* = \frac{\sum_{b=1}^{B}\theta_b^*}{B}$ $\forall b = 1, 2, ..., B$

The standard error estimator is obtained using the following expression:

$$\hat{SE}(\theta^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\theta_b^* - \bar{\theta}_b^*)^2} \tag{3.6}$$

and the bias estimator is computed as

$$\hat{\text{bias}}(\theta_b^*) = \bar{\theta}_b^* - \hat{\theta} \tag{3.7}$$

It is important to note that the Bootstrap procedure cannot be applied directly to the Hierarchical Linear Mixed Model (or Multilevel Models) because this procedure is based on independent observations and consequently, this procedure does not take into account intraclass dependency in the data following a hierarchical structure. Thus it is necessary to modify or adjust this procedure by considering the dependence between the different observations. Several authors have worked on the application of the Bootstrap Procedure in the estimation of parameters in Hierarchical Linear Mixed Models or Multilevel

Models, see for example Rien Van der Leeden et.al (1997) [45] and Carpenter et.al (2003) [5].

In our study, we follow the ideas presented in the paper by Rien Van der Leeden et.al (1997) [45] who proposed three methods in order to apply the Bootstrap procedure to multilevel models: 1) The parametric Bootstrap which assumes that explanatory variables are fixed, and both the model and distributions are correctly specified. 2) The residual Bootstrap they requires that the explanatory variables be fixed and the model specification be correct. 3) And the cases Bootstrap which only requires the hierarchically structured data. These methods are summarized below.

3.3.1 Parametric Bootstrap

The first method, Parametric Bootstrap generates level one residuals (ε_{ij}) from normal distribution $\mathcal{N}(0,\hat{\sigma}^2)$ with $\hat{\sigma}^2$ being the estimated variance obtained from the original sample. Level two residuals (ζ_j) are generated from $\mathcal{N}(0,\hat{\Theta})$ where $\hat{\Theta}$ is the covariance–variance matrix of level two residuals. The steps are illustrated as follows

1. Select J vectors (ζ_j) from $\mathcal{N}(0,\hat{\Theta})$.

2. Choose J vectors ε_j^* of size n_j from $\mathcal{N}(0,\hat{\sigma}^2)$.

3. The vector Y_j is generated according to the proposed model using level one and two residuals generated in steps 1 and 2. The vector of fixed parameters is estimated by maximum likelihood.

4. Calculate all estimators of parameters in the model and steps 1, 2, 3 and 4 repeat B times for calculating the standard errors or bias of estimators.

3.3.2 Residual Bootstrap

The second, the Residual Bootstrap method draws samples with replacement of level one and level two residuals. Rien Van der Leeden et.al (1997) [45] uses two types of residuals; ”raw and shrinkage”, here discuss only one of these residuals. Additionally, they use the centered residuals because the residuals in the multilevel models do not have an average equal to zero. The centered estimates of level one and level two residuals are denoted by $\{\hat{\varepsilon_{ij}}\}$ and $\{\hat{\zeta_j}\}$, respectively. This procedure is summarized as follows

1. Choose a sample of $\{\hat{\zeta_j}\}$ of size J with replacement.

2. Extract J samples of size n_j with replacement from $\{\hat{\varepsilon_{ij}}\}$.

3.The vector Y_j is generated according to the proposed model using level one and two residuals generated in steps 1 and 2. The vector of fixed parameters is estimated by maximum likelihood.

4. Calculate all estimators of parameters in the model and steps 1, 2, 3 and 4 repeat B times for calculating the standard errors or bias of estimators.

3.3.3 Cases Bootstrap

The third, the Cases Bootstrap method considers entire cases of level one and two units, this is possible whatever the explanatory variables are assumed to be random variables. This procedure is described as follows

1. Select a sample of level two units of size J with replacement and their corresponding variables measured at level two.

2. For each level two unit selected in step 1, we select all the cases of level one units with replacement from the level two unit in the original sample with their corresponding responses and explanatory variables.

3. Calculate all estimators of parameters in the model and steps 1, 2, 3 and 4 repeat B times for calculating the standard errors or bias of estimators.

3.4 The Cases Bootstrap method applied to our study

Rien Van der Leeden et.al (1997) [45] discuss different versions of the cases Bootstrap method. They suggest selecting samples with replacement of level two units, or of level one units or both, depending on the nature of the data and the interest of the investigation.

In our work, we decide to use the Cases Bootstrap method for estimating the standard error of the prediction of Value Added for each university for the following reasons:

- The Parametric and Residual Bootstrap methods assume that the explanatory variables are fixed. In our case, these methods do not apply because the explanatory variables are random variables.
- The Cases Bootstrap presents several versions which permit adjusting to different research, according to the objectives of our study and the characteristics of the data.
- The Cases Bootstrap only require that data follows a hierarchical structure. In our study, the data, present this characteristic.
- The Case Bootstrap allows us to have information concerning all of the universities and we do not select a sample with replacement of universities. In our research, we need information regarding all of the universities that participated in our study.

We describe the procedure as follows

1. For each university j, we select n_j students with replacement along with their corresponding response variable and the explanatory variables measured at the student or university level.

2. For each sample of size n_j students, we calculate the prediction of Value Added given by Equation (2.14) replacing the parameters with their corresponding estimators obtained by maximum likelihood.

3. Steps 1 and 2 repeat B times and we calculate the Bootstrap estimates of the standard error of the predicted Value Added for each university using Equation (3.6).

3.5 Simulation Study

We evaluate the effect on the prediction of the Value Added for universities in the presence of measurement error in a simulation study. We simulate a population with the number of universities and students that are equal to the Colombian database of universities. We assume a model with four explanatory variables measured with error and we examine two scenarios: The first model with adjusted error in terms of reliability and the second without adjusted error. Next, we compute the prediction of Values Added in each scenario and their corresponding standard error estimates using the Cases Bootstrap method mentioned in the previous section.

All variables were simulated for each university in order to maintain the hierarchical structure of data. Also, it is important to note that the computational cost is high when the model includes many explanatory variables measured with error. We use 1,000 iterations in each scenario.

We fixed the following values of parameters $\alpha = 2, \beta_1 = 2, \beta_2 = 1.4, \beta_3 = 1.5, \beta_4 = 1.6, \sigma^2 = 1, \omega^2 = 2.5, \tau_1^2 = 1.4, \tau_2^2 = 1.6, \tau_3^2 = 1.8, \tau_4^2 = 1, \rho_{12} = 0.49, \rho_{13} = 0.47, \rho_{14} = 0.55, \rho_{23} = 0.42, \rho_{24} = 0.41, \rho_{34} = 0.51$.

We describe the simulation study as follows

1. The variables in the model were simulated as

- $\eta_{kij} \sim \mathcal{N}(0,1)$
- $\theta_j \sim \mathcal{N}(0,2.5)$
- $(Y_{ij}|\eta_{ij},\theta_j) \sim \mathcal{N}(2+\eta_{ij}^T\beta+\theta_j,1)$
- $(X_{kij}|\eta_{kij}) \sim \mathcal{N}(\eta_{kij},\tau_k^2)$

2. We estimate Value Added for each university in terms of reliability. We consider five scenarios with reliability values equal to $0.4, 0.5, 0.6, 0.8$ and 1. For simplicity, we use, in each case, the same reliability values for all explanatory variables.

3. We obtain the standard error estimate of Value Added for each university using the Bootstrap procedure with 1,000 iterations for each scenario.

Results

Figure 3.1 shows the model with adjusted error and several of the reliability values versus the model without error. We observe that when the reliability decreases, that is, the error measurement increases, the variability of Value Added estimates for each university is greater which means that Values Added estimates are more dispersed and less concentrated around their mean. Therefore, we assert that if we do not adjust the error measurement then it has an implication in the variability of the estimators and consequently in the variance estimator.

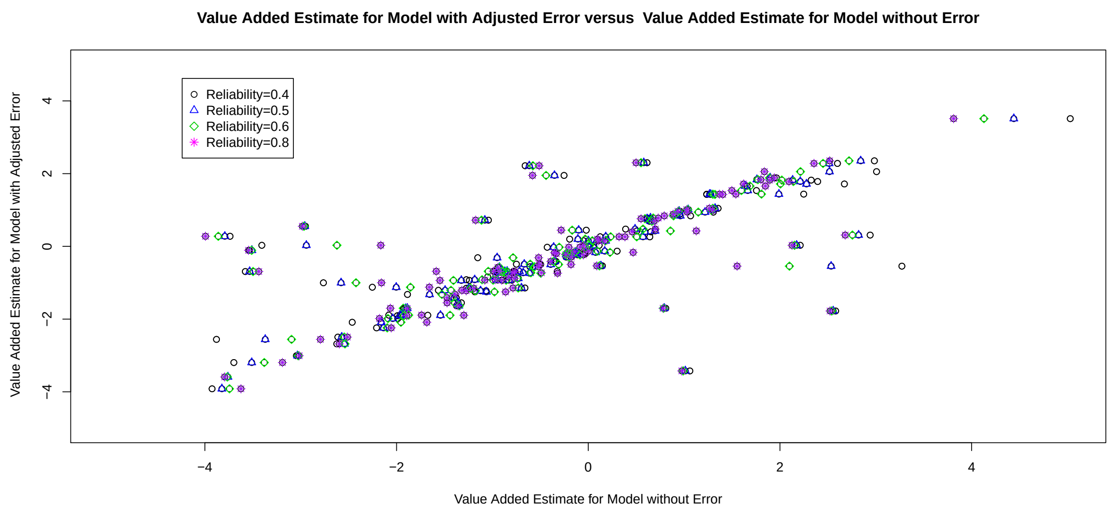

FIGURE 3.1. Models with Adjusted Error Versus Model without Error

Figures 3.2 and 3.3 illustrate two models with adjusted error and several reliability values. We note that the observed trend remains in the graphs, that is, greater dispersion is related to lower reliability. We also observed in the case of the universities that do not follow the trend of the data. The Value Added estimates are very similar for the models with adjusted measurement error and they are very different for the model that does not correct the measurement error. Moreover, Value Added estimates have opposite signs for the same university. This implies that we can obtain different conclusions for one same university, that is to say, we could assert that the university is contributing to the progress of the students when in fact it is not.

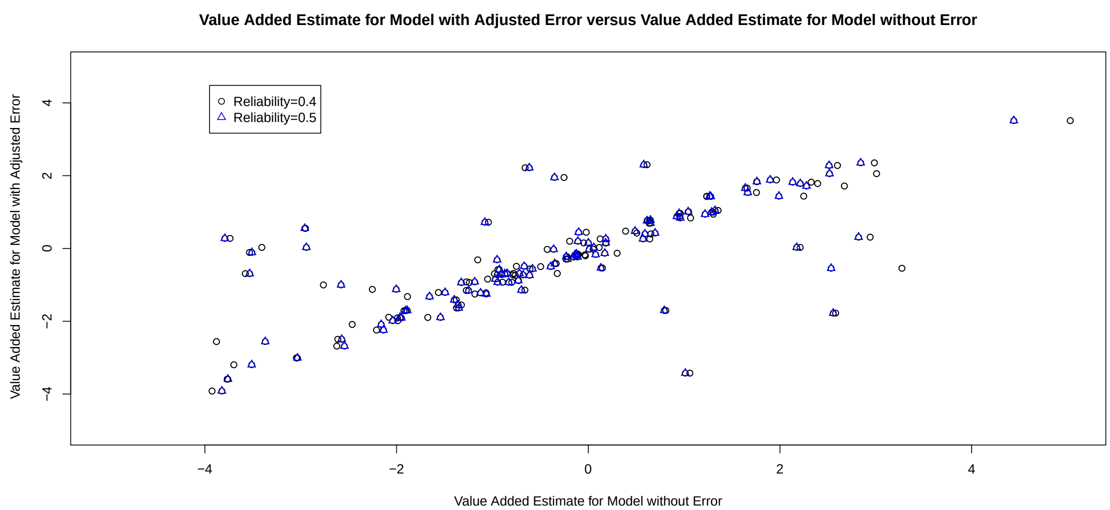

FIGURE 3.2. Models with Adjusted Error and R=0.4, 0.5 versus the Model without Error

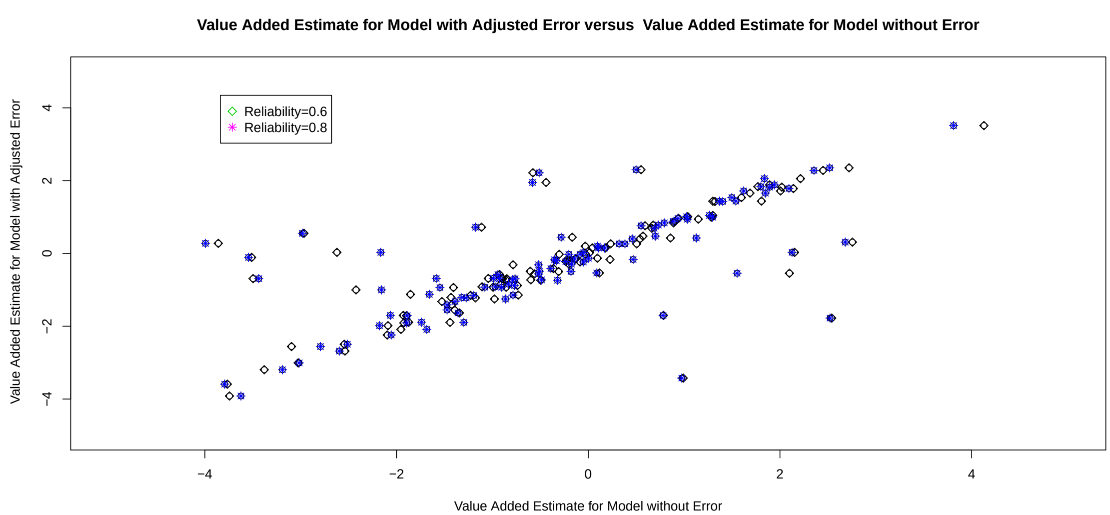

FIGURE 3.3. Models with Adjusted Error and R=0.6, 0.8 versus the Model without Error

3.6 Bootstrap Confidence Intervals for Value Added

It is well known that for estimating Value Added for a school or university, we need to predict the school effect, that is, $\hat{\theta}_j = E(\theta_j|Y_j, X_j)$. This expression depends on parameters which are estimated by maximum likelihood, this way, we obtain $\hat{\hat{\theta}}_j$ which is called the estimated best predictor for the school effect.

On the other hand, as well is known in the literature, the confidence interval for θ_j is characterized by

$$P(\theta_j - \hat{\theta}_j|X) = 1 - \alpha$$

In this work, we use $\alpha = 0.05$

Thus, we calculate the confidence intervals for Value Added using the variance of error of prediction, that is, $V(\theta_j - \hat{\hat{\theta}}_j|X)$ which is calculated by the Bootstrap procedure. We illustrate the more relevant results concerning Bootstrap confidence intervals for Value Added as is shown in the following figures.

Figures 3.4 to 3.8 show the graph of confidence intervals for the Value Added for each university with several reliability values, also, we observe that when the reliability is close to one, the confidence interval for Value Added associated with this reliability is similar to the confidence intervals for Value Added associated to the model without error.

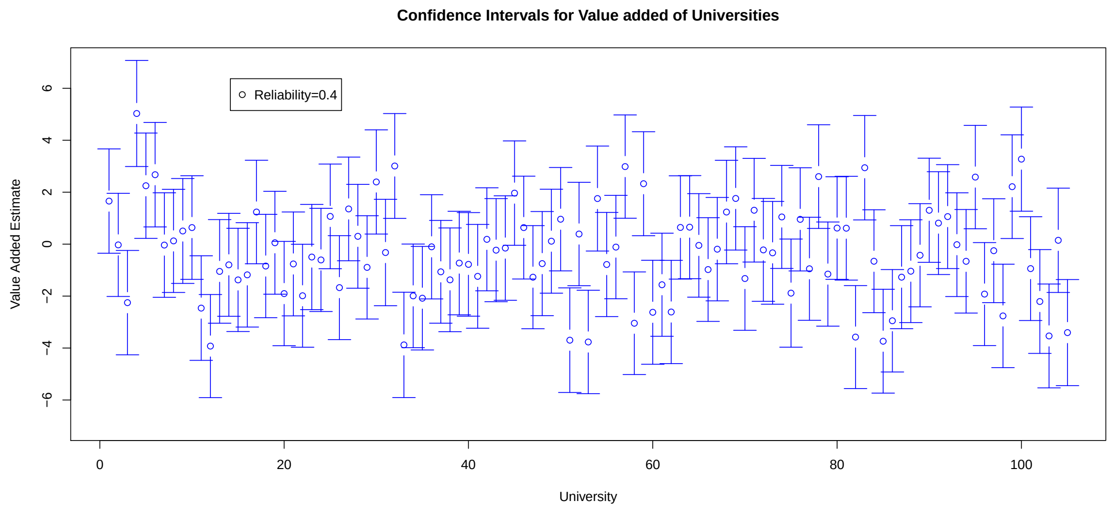

FIGURE 3.4. Confidence intervals for Value Added for 105 Universities with R=0.4

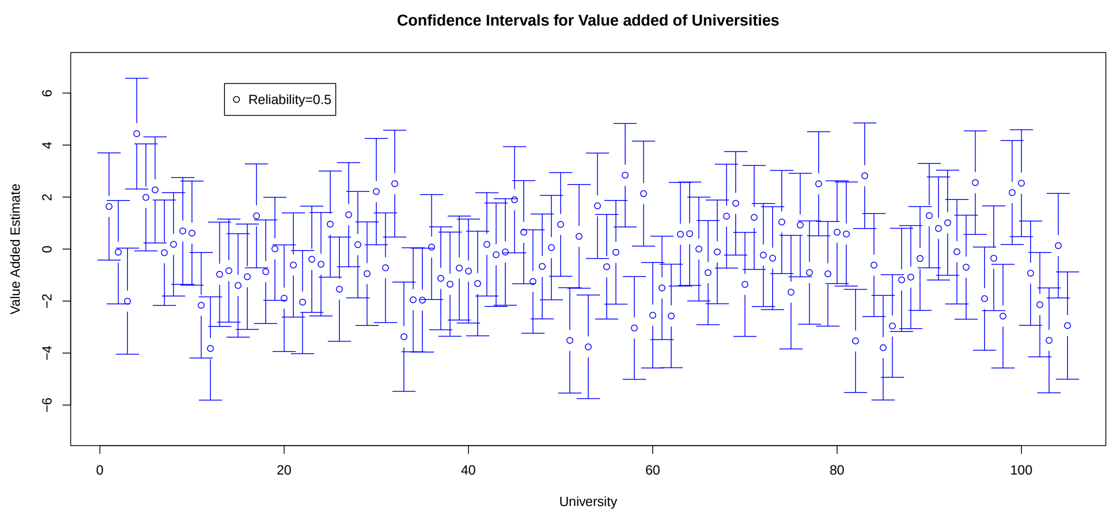

FIGURE 3.5. Confidence intervals for Value Added for 105 Universities with R=0.5

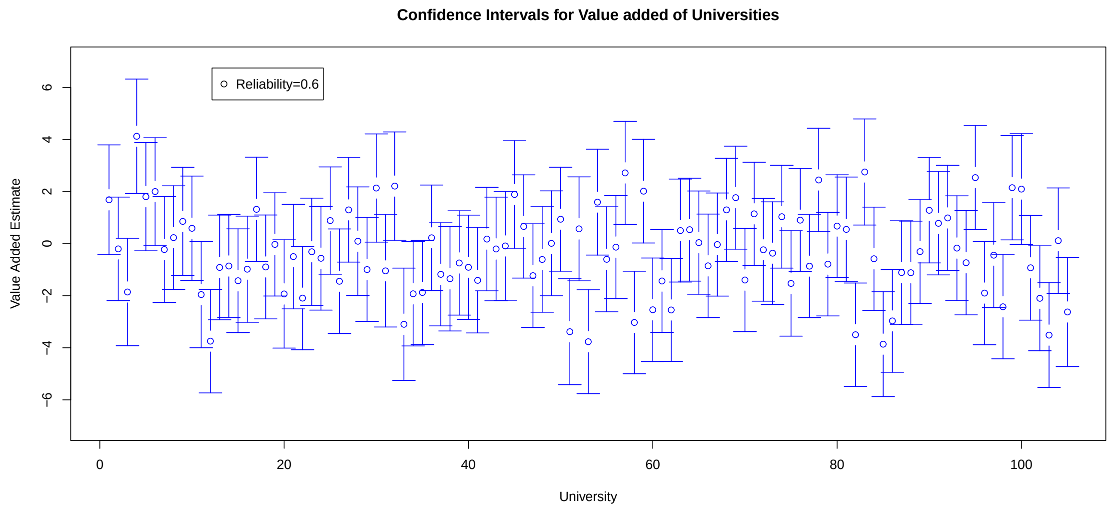

FIGURE 3.6. Confidence intervals for Value Added for 105 Universities with R=0.6

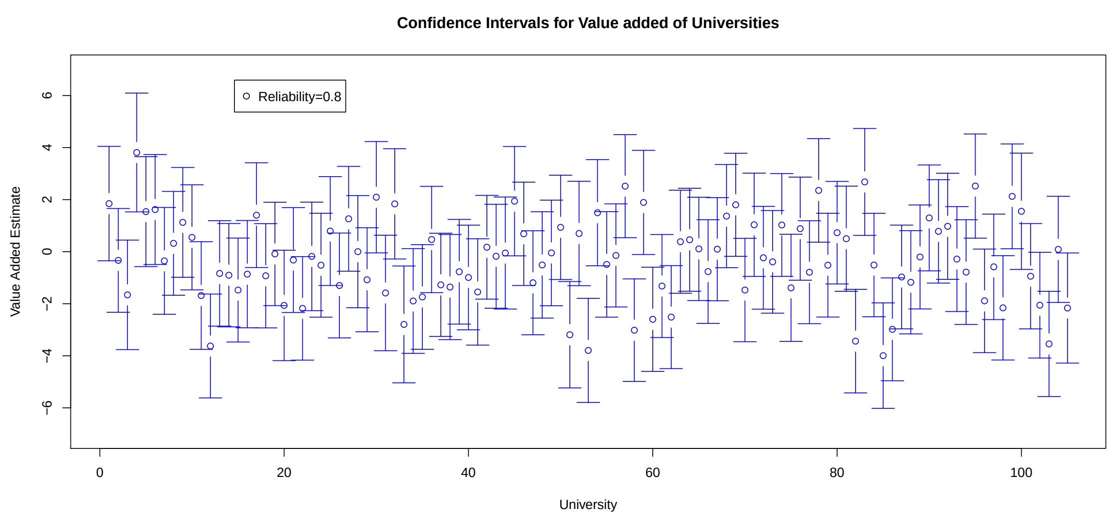

FIGURE 3.7. Confidence intervals for Value Added for 105 Universities with R=0.8

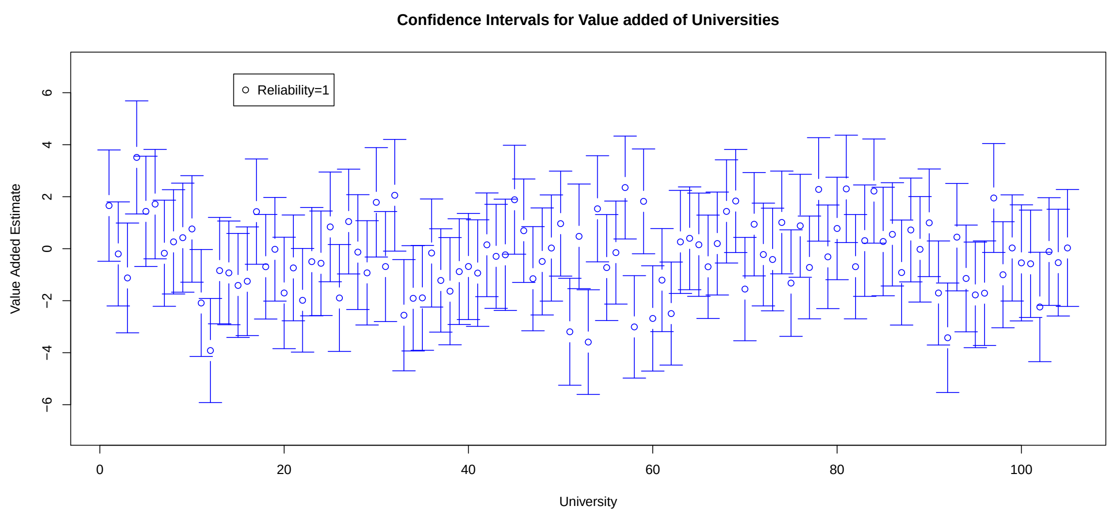

FIGURE 3.8. Confidence intervals for Value Added for 105 Universities with R=1

Finally, figures 3.9 to 3.12 illustrate the graphs of confidence intervals for Value Added with a specific reliability versus the model without error. When the reliability is equal to 0.4, the graph is very different from the graph of the model without error. It also is important to note that as reliability approaches one, the confidence intervals for Value Added parameters are very similar to the confidence intervals for Value Added parameters for the model without error. For reliability values less than 0.5 the confidence intervals are very different. For example, we observe that concernig university 10, the upper limit of the confidence intervals is greater than 6 for in the case of all models with reliability lower than 0.8 while the upper limit of the confidence intervals for the parameter of Value Added for the model without error is less than 6.

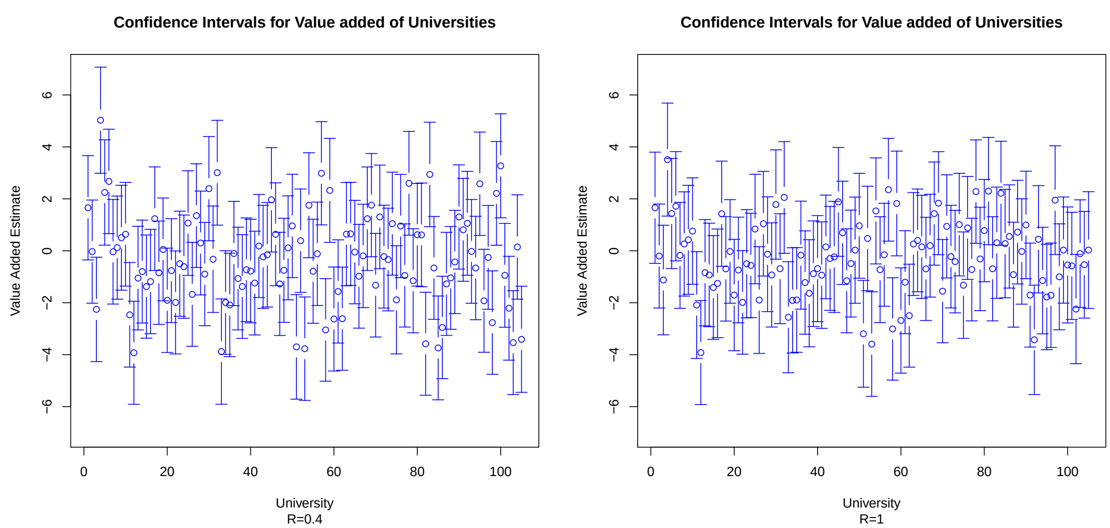

FIGURE 3.9. Confidence intervals for Value Added for 105 Universities with R=0.4 and R=1

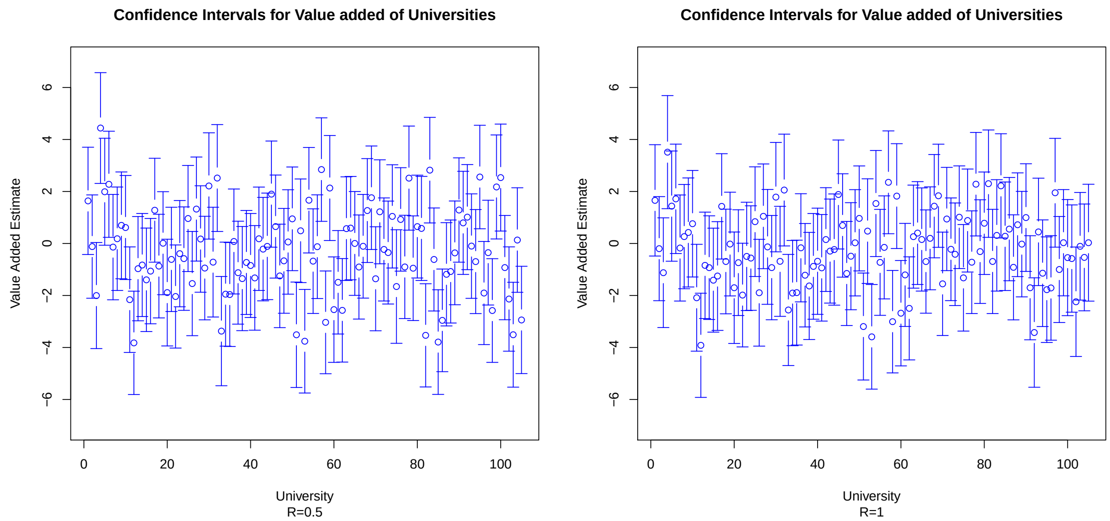

FIGURE 3.10. Confidence intervals for Value Added for 105 Universities with R=0.5 and R=1

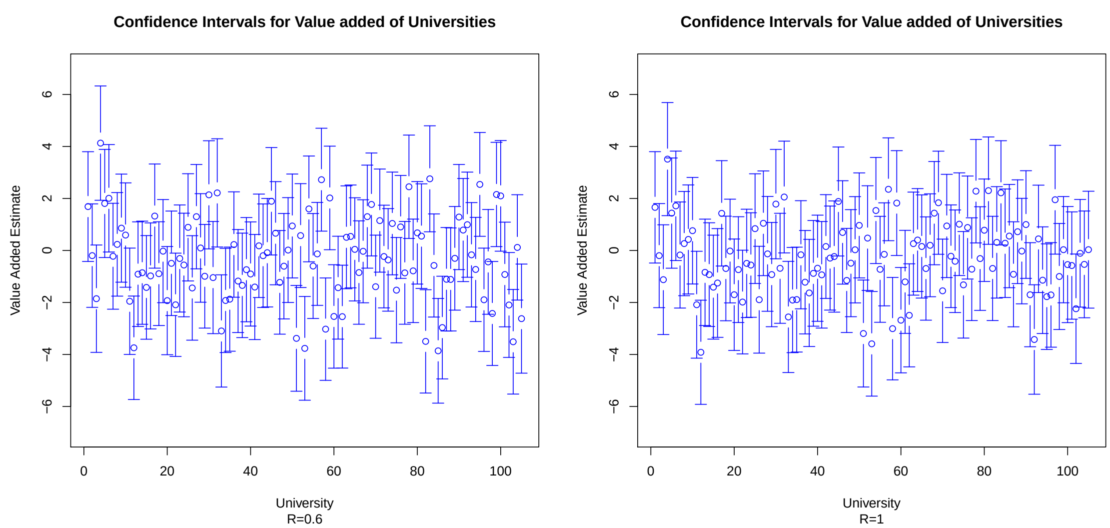

FIGURE 3.11. Confidence intervals for Value Added for 105 Universities with R=0.6 and R=1

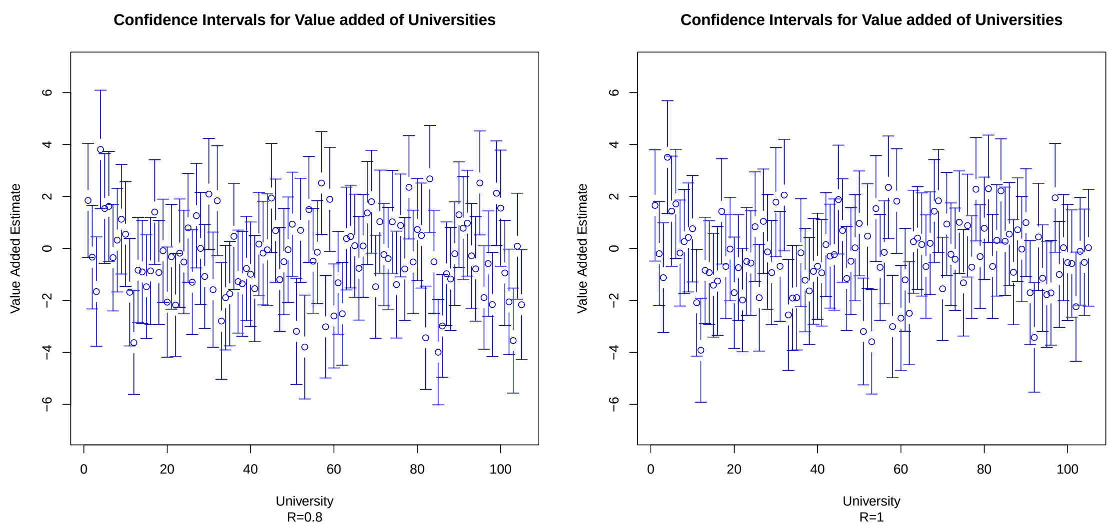

FIGURE 3.12. Confidence intervals for Value Added for 105 Universities with R=0.8 and R=1

3.7 Final Remarks

We summarize the more relevant results from the simulation study:

1. The variability of Value Added estimates for each university is greater when the variables measured with error present less reliability.

2. If we do not adjust the error measurement, this situation has an implication on the variability of the estimators and consequently on the variance estimator.

3. The confidence intervals for Value Added in a model with reliability approximately equal to one are similar to the confidence intervals for Value Added in a model without error measurement. The confidence intervals for Value Added in a model with reliability approximately equal to one, are similar to the confidence intervals for Value Added in a model without error measurement. Moreover, for small reliability values, for instance, of less than 40%, the confidence intervals for Value Added are wider.

4. The error measurement of variables affects both Value Added estimation and Value Added variance because they both depend on the variance of the variables measured with error.

5. If we do not take into account the measurement error of the variables for Value Added models, we will have misleading conclusions from the statistical point of view.

CHAPTER 4

Application to the Colombian Case

4.1 Introduction

In this chapter, we present an overview of the higher education system in Colombia. We also review some main aspects such as the agencies in charge of this educational level and their major functions, the types of educational institutions that offer both undergraduate and graduate programs; and the mechanisms used in the evaluation of the quality of higher education.

We also apply the methodology proposed in this research to a Colombian database which consists of students' scores for two standardized tests organized by government agencies. Additionally, we estimate the Value Added for Colombian higher education institutions by considering the following models under four scenarios: Firstly, Model I with Adjusted Error (includes four explanatory variables measured with error). Secondly, Model I without Adjusted Error (includes four explanatory variables measured with error). Third, Model II with Adjusted Error (includes eight explanatory variables measured with error). Fourth, Model II without Adjusted Error (includes eight explanatory variables measured with error).

Finally, we compare the Value Added estimates with the average accreditation per year for higher education institutions. We analyze the relationship between these estimates and the average accreditation per year. Also, the Value Added indicators have been taking into account in the accreditation process.

4.2 Higher Education in Colombia

The higher education system in Colombia is regulated by Law 30 of 1992 which establishes the fundamental principles that govern this educational level. This Law defines higher education as "a cultural public service, inherent in the social purpose of the State". In Colombia, this level of education has undergone important changes which has led to considering a reform to this law. So that it is closer to the real situation in the country and the world. Moreover, there is a new demand from the labour market, therefore, other competencies and skills are required of the students. In 2011, the Colombian government attempted to carry out a reform of this law but it did not have the support of the students and staff of the universities.

4.2.1 Colombian Higher Education Agencies

We mention bellow some agencies and institutions that are in charge of education:

- The National Ministry of Education (MEN, Ministerio de Educación Nacional) whose function is to manage and supervise the entire education system.

- The Vice-Ministry of Higher Education, created in 2003, manages the application of national higher education policies as well as the planning and monitoring of the sector. It is divided into two main agencies: Directory for the Promotion of Higher Education and the agency for Quality Management for higher education. The main functions of this first agency are to expand the supply and to improve the regional distribution of tertiary places. The second agency is responsible for everything related to quality assurance in higher education.

- The National Council for Higher Education (CESU, Consejo Nacional de Educación Superior) has the functions of coordination, planning, recommendation and advice. It is composed of the Ministry of Education and members of the academic community.

- The National Intersectoral Commission for Quality Assurance of Higher Education (CONACES, Comisión Nacional Intersectorial de Aseguramiento de Calidad de la Educación Superior) provides advice to the Ministry of Education in the assuring of quality of higher education and also the evaluation of the requirements for the creation of new higher education institutions to acquire qualified registry.

- The National Accreditation Commission (CNA, Comisión Nacional de Acreditación) advises the Ministry of Education concerning the submission of applications from institutions for obtaining "high quality accreditation". That acreditation is for the institutions and their academic programs.

- The Administrative Department of Science, Technology and Innovation (COLCIENCIAS) works with higher education institutions and promotes policies that increase scientific research and knowledge production. It also finances many scientific investigations which are carried out by universities.

- The Colombian Institute for the Assessment of Education (ICFES, Instituto Colombiano para la Evaluación de la Educación) is the agency responsible for the design and management of four different national tests: Saber 5 is a test that students present at the end of primary education; Saber 9 is presented at the end of lower secondary education; Saber 11 is presented by 11th grade students and it is a requirement for accessing higher education; and Saber Pro (formerly known as ECAES) which is presented when university students have approved at least 75% of the credits of an undergraduate academic program.

- The Colombian Institute of Educational Loans and Technical Studies Abroad (ICETEX, Instituto Colombiano de Crédito Educativo y Estudios Técnicos en el Exterior) aims to promote the enrollment of students with lower incomes in higher education. It also provides loans to students for undertaking their studies in national and international universities. (See OECD et al., 2012, pp. 39-41) [32].

4.2.2 Types of Higher Education Institutions

As mentioned in Chapter 1, in Colombia, there are four types of higher education institutions, which are classified according to under and postgraduate academic programs and according to their time duration programs.

- Universities offering undergraduate and graduate academic programs, master and doctoral programs. They participate in scientific and technological research.
- University institutions offering undergraduate programs and possibly postgraduate program called specialization.
- Technological institutions offering programs at the professional technical level (on a scientific basis). The students at this level can continue their studies until recieving a university degree.
- Professional technical institutions offer undergraduate programs at the technical and vocational levels.(See OECD et al., 2012, pp. 32) [32].

4.2.3 Quality of Higher Education

In Colombia, the quality assurance process and the accreditation of institutions and their programs are carried out by two agencies: The National Intersectoral Commission for Quality Assurance in Higher Education (CONACES, Comisión Nacional Intersectorial de Aseguramiento de Calidad en la Educación Superior), whose main function is to advise the Ministry of Education on what programs deserve to be included in the Qualified Registry; and the National Accreditation Commission (CNA, Comisión Nacional de Acreditación) that advises the Ministry of Education concerning higher education institutions and their programs must be awarded high quality accreditation. (See OECD et al., 2012, pp. 177-178) [32].

The mechanisms for ensuring the quality of higher education are:

- The Qualified Registry allows higher education institutions to offer academic programs ensuring certain minimum quality requirements.
- The system of accreditation of institutions and their programs of high quality. The accreditation is voluntary and temporary.
- The students' scores of the SABER PRO test. (See OECD et al., 2012, pp. 41-42) [32]

4.3 Application to the Colombian Case

Colombia has very valuable information concerning education topics that allows us to develop a Value Added analysis. We have prior information from secondary education tests corresponding to a students' period prior to enrolling in the higher education institutions and we have additional information from university evaluations at the end of the university carrer.

In this research we apply the proposed methodology to a Colombian database for two main reasons. The Colombian database contains a very rich and quality information concerning students' tests prior to before these being enrolled in a higher education institution; and also another information regarding students' tests when these after being enrolled in a specific institution. Hovewer, it is known that Saber 11 has low reliability. This subject was discussed in detail in Chapter 1 and as we mentioned, it is a great motivation for this research.

The database considered here consists of the scores of two standardized tests: Saber 11 and Saber Pro tests, which are organized by the Colombian Institute for the Assessment of Education (ICFES). In the first test, the students are evaluated in the eleventh grade of upper secondary education and it is compulsory. However, we also know Saber 11 has reliability problems for measuring individual preformance.(OECD, 2012, pp. 59) [32]. The second test is an evaluation of Colombian students who have approved at least seventy five percent of a higher education program. It has been compulsory since the year 2010 and is a requirement for graduation.

The Saber Pro test is the Colombian state examination for evaluating the quality of higher education. The results of the Saber Pro test are useful in the construction of qualilty indicators for higher education institutions and their academic programs. The Saber Pro test evaluates some of the students' generic and specific skills. The generic competencies are composed of several modules such as written communication, quantitative reasoning, critical reading, citizenship competencies and English. See the National Report Saber Pro Results 2012-2015, ICFES, 2016. The Saber 11 exam is composed of six tests: Mathematics, Chemistry, Physics, Social Sciences, Spanish and English.

4.3.1 Actual Data

The original database consists of 32,582 observations and 90 variables, which include the students' Saber Pro scores from the engineering reference group who took the test in the years 2012 and 2013, additionally their Saber 11 test score have being taken in the years from 2006 to 2009 including other sociodemographic and demographic variables. We eliminated missing data and work with the educational institutions with less than 30 students. The educational institutions with 1 student or 4 students, were excluded because they are considred as small populations which need to be studied by other special statistical methods. Under these considerations, the original database was reduced to 31,535 observations, 105 universities and university institutions.

4.3.1.1 Validation of the assumptions in the real data

It is important to note that Colombian database used in this study fulfill with the assumptions of the statistical model provided for this research. A first fundamental assumption satisfied by data is the exogeneity which asserts that the covariates and the school effect are independent. Scores Saber 11 are assumed to be independent on university effects because the students take the test before they are enrolled in the university. Also, the averages of Saber 11 score in its several modules are independent

on the university effect due to these averages are calculated from independent covariates. Moreover, the true variables corresponding to the Saber 11 test score are independent on the university effect because the covariates (or variables measured with error) are the versions of these true variables.
Additionally, the conditional independence between the scores of the two tests given the university effect and the true covariate, is also fulfilled. Therefore, we can assure the most important assumptions of the model are verified.

4.3.1.2 Summary Statistics

In this section, we continue with summary statistics of the raw score in both Saber 11 and Saber Pro tests. Table 4.1 shows the averages and other statistical measurements of scores in several competencies of the Saber 11 test such as Mathematics, Language, Social Sciences and Chemistry. Also, it shows the average of Quantitative Reasoning of the Saber Pro test. It is important to note that the Mathematics scores have a standard deviation greater than the other scores, that is, a greater heterogeneity. The Language scores have a smaller standard deviation indicating a greater homogeneity.

TABLE 4.1. Summary Statistics

	Saber 11				Saber Pro
	Mathematics	Language	Social Sciences	Chemistry	Quantitative Reasoning
Mean	55.49	52.86	52	51.50	10.89
SD	11.55	6.8415	8.002	7.076	1.159
MIN	14.92	18.33	12.87	17.70	5.80
MAX	121.49	99.69	97.35	93.33	16.30

Table 4.2 shows the correlation coefficients between the raw Saber 11 test scores, we observe that all the correlations between scores are less than 50%, this suggests that competencies evaluated in several knowledge areas present a lower correlation. The Mathematics and Chemistry scores have a correlation greater than 50%.

Table 4.3 shows the correlation coefficients between the raw Saber 11 test and the raw Quantitative Reasoning of the Saber Pro test. Note that there is a relatively high correlation between Quantitative Reasoning and Mathematics and Chemistry. But the correlation with other knowledge areas is low, with values of less than 50%.

TABLE 4.2. Correlation coefficients between different areas of the raw Saber 11 Test Scores

	Mathematics	Language	Social Sciences	Chemistry
Mathematics	1	0.4562	0.4442	0.5370
Language	0.4562	1	0.4622	0.4271
Social Sciences	0.442	0.4622	1	0.47779
Chemistry	0.5370	0.4271	0.4779	1

TABLE 4.3. Correlation coefficients between the raw Saber 11 Test Scores and the raw Quantitative Reasoning

Saber Pro	Saber 11			
	Mathematics	Language	Social Sciences	Chemistry
Quantitative Reasoning	0.6042	0.4418	0.4739	0.5183

Finally, Figure 4.1 shows the distribution of Mathematics, Language, Social Sciences and Chemistry scores of the Saber 11 test. We observe that the Mathematics scores present greater dispersion or variability, that is, the Mathematics scores are more heterogeneous while the Chemistry scores present lower variability, i.e., these are more homogenous.

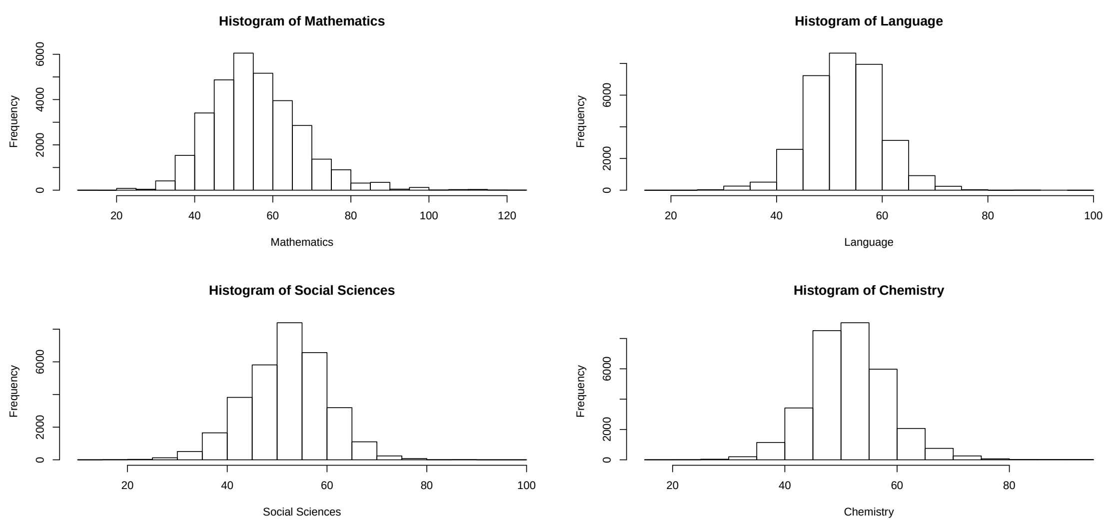

FIGURE 4.1. Histograms of the Raw Saber 11 Test Scores

4.3.2 Value Added Analysis

We analyze four models in different scenarios: First, is a model with four explanatory variables (measured with error), taking into account the measurement error from the model specification up to parameter estimation, we denominate this model, Model I with Adjusted Error. Second, is a model with the same explanatory variables as the first scenario, but without correction of the measurement error, we refer to this model as Model I without Adjusted Error. Third, is a model with eight explanatory variables, the same four variables used in the first scenario and we also include their corresponding university average, correcting the measurement error from the model specification up to parameter estimation, this model is Model II with Adjusted Error. Fourth, is a model with the same variables as the third scenario, but without correcting the error, this model is called Model II without Adjusted Error. In all these four scenarios, the model variables were standardized and centered with respect to the university average.

4.3.2.1 Model I

The Model I with adjusted error can be written as

$$Y_{ij} = \alpha + \beta_1 \eta_{1ij} + \beta_2 \eta_{2ij} + \beta_3 \eta_{3ij} + \beta_4 \eta_{4ij} + \theta_j + \xi_{ij}$$

Where

α, β_k are fixed parameters, η_{kij} are the nonobservable true variables, $(k = 1, 2, ..., K)$, $\xi_{ij} \perp \theta_j$, $\xi_{ij} \sim \mathcal{N}(0, \sigma^2)$ and $\theta_j \sim \mathcal{N}(0, \omega^2)$

In this model, we include the following observable variables:

Y_{ij}: is the observed Quantitative Reasoning score in the Saber Pro test for student i belonging to university j

X_{1ij}: is the observed Mathematics score in Saber 11 test for student i belonging to university j

X_{2ij}: is the observed Language score in Saber 11 test for student i belonging to university j

X_{3ij}: is the observed Chemistry score in Saber 11 test for student i belonging to university j

X_{4ij}: is the observed Social Sciences score in Saber 11 test for student i belonging to university j

Results

Figure 4.2 shows the Value Added estimates for Model I with Adjusted Error versus Value Added estimates for Model I without Adjusted Error. Note that Value Added estimates for Model I with Adjusted Error vary approximately between -0.003 and 0.004 while Value Added estimates for Model I without Adjusted Error vary between -1.5 and 0.5. We also observe that a given university has negative and positive Value Added estimates whether the error is adjusted or not; for instance, in this figure a university is shown whose Value Added estimate is equal to -0.003 for the model with adjusted error while the Value Added estimate is equal to 0.5 for the model without Adjusted Error. Thus, we can obtain invalid conclusions about the Value Added of a university if we do not adjust the measurement error.

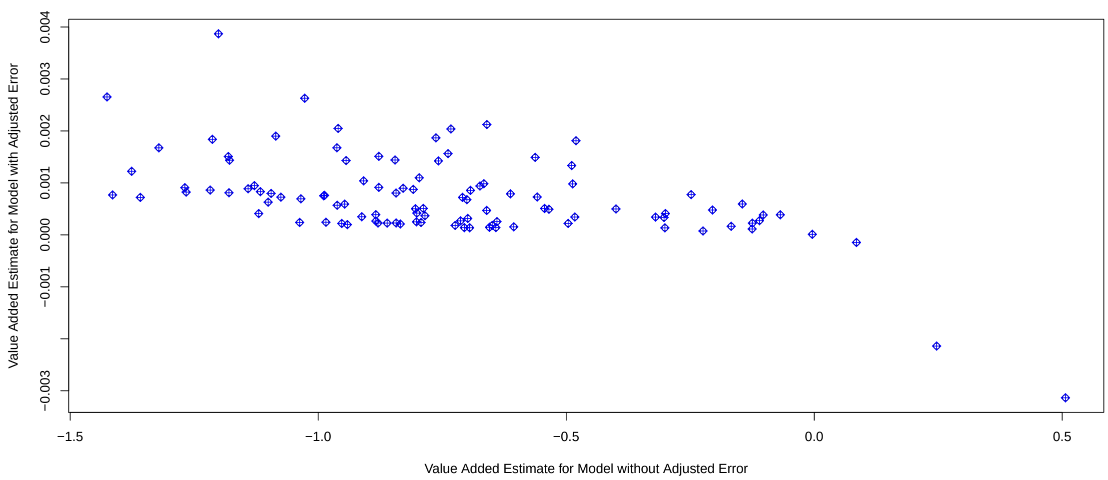

FIGURE 4.2. Model I with Adjusted Error versus Model I without Adjusted Error

Figures 4.3 and 4.4, the Value Added estimates for Model I without Adjusted Error present a greater variability than the model with Adjusted Error. It also is observed that the distribution of Value Added estimates for model with Adjusted Error is nearly symmetric while the distribution of Value Added estimates for model without adjusted error is not symmetric, in particular, skewed to the right.

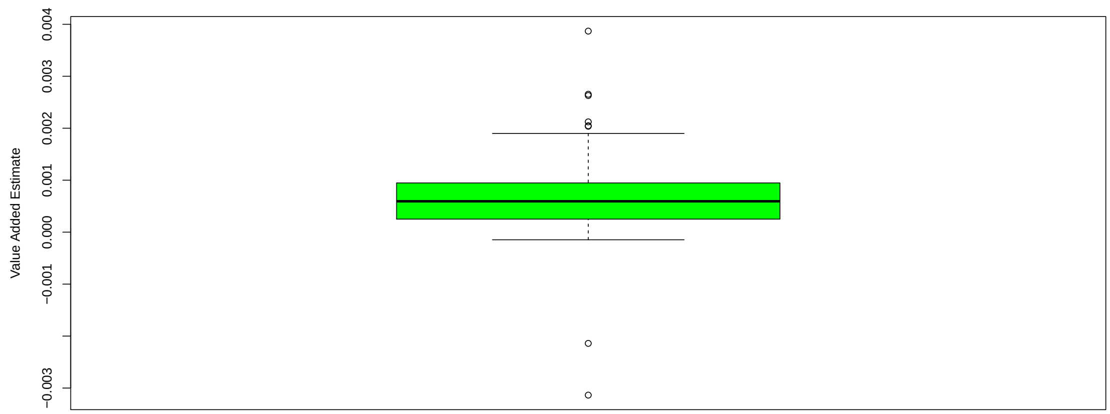

FIGURE 4.3. Boxplot of Value Added Estimate for Model I with Adjusted Error for 105 universities

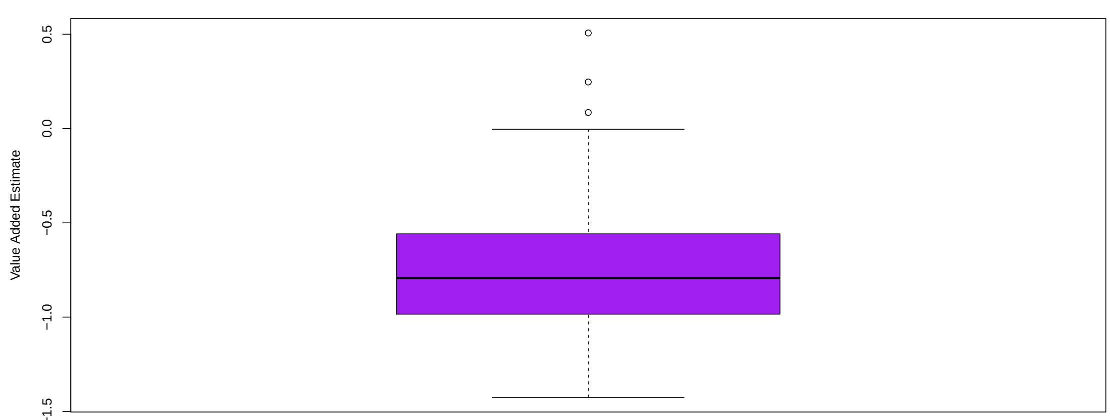

FIGURE 4.4. Boxplot of Value Added Estimate for Model I without Adjusted Error for 105 universities

Figure 4.5 presents the plot of confidence intervals for Value Added in each of the 105 universities for Model I with Adjusted Error and Model I without Adjusted Error, respectively. Note that, there is a difference between these plots which implies that Value Added estimates and their corresponding variances estimates have very different values in both models. However, it is worth mentioning that in Model I without adjustment of the measurement error, the confidence intervals are larger than in Model I with adjusted error. In Model I with Adjusted Error is there are no differences in Value Added estimates for all of the universities participating in our study. In Model II, this situation improves as we present below.

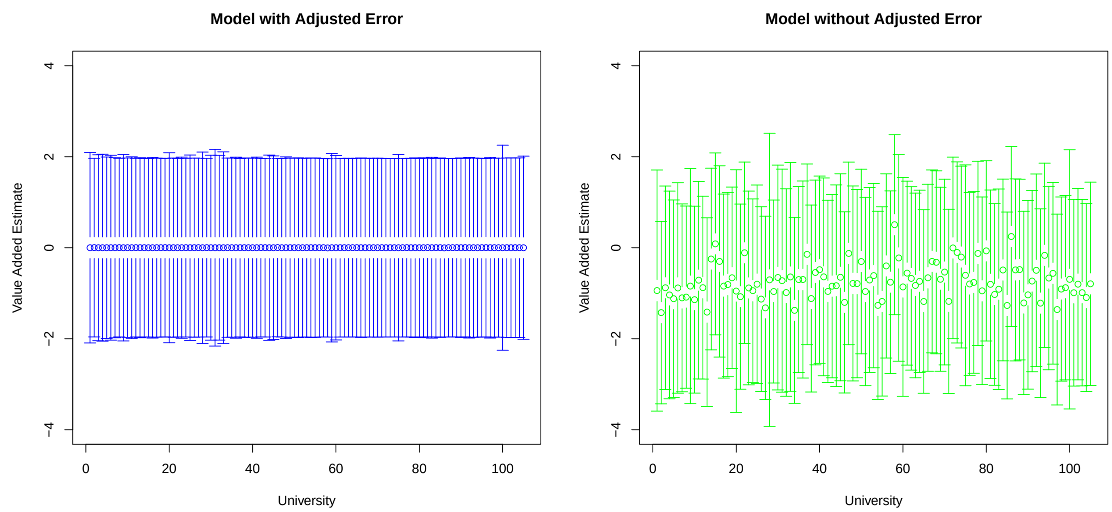

FIGURE 4.5. Confidence intervals for Value added for 105 universities for Model I with Adjusted Error and Model I without Adjusted Error

Table 4.4 shows the Value Added estimate and the standard error estimate for the first ten universities in both Models I; with and without Adjusted Error. We observe that the Value Added estimates of these universities are very different and they are of opposite sign. Also, the standard error estimates for Model I without Adjusted Error are less than their corresponding standard error estimates for Model I with Adjusted Error, but they are incorrect because we are not correcting the measurement error.

TABLE 4.4. Valued Added Estimates for the first ten universities

	Model I With Adjusted Error		Model I Without Adjusted Error	
	VAEst	SE VAEst	VAEst	SE VAEst
1	0.00019	1.91	-0.94	0.45
2	0.00265	6.56	-1.43	0.42
3	0.00022	2.12	-0.88	0.70
4	0.00024	2.40	-1.04	0.77
5	0.00041	3.45	-1.12	0.59
6	0.00027	2.15	-0.88	0.49
7	0.00063	3.50	-1.10	0.66
8	0.00189	8.03	-1.09	0.47
9	0.00022	2.00	-0.84	0.35
10	0.00089	4.12	-1.14	0.52

4.3.2.2 Model II

The Model II can be written as follows

$$Y_{ij} = \alpha + \eta_{ij}^T \beta + \theta_j + \xi_{ij}$$

Where

η_{ij} is a vector of true variables, $\xi_{ij} \perp \theta_j$, $\xi_{ij} \sim \mathcal{N}(0, \sigma^2)$ and $\theta_j \sim \mathcal{N}(0, \omega^2)$

In this model, we consider the following variables:

Y_{ij}: is the observed Quantitative Reasoning score in the Saber Pro test for student i belonging to university j

X_{1ij}: is the observed Mathematics score in the Saber 11 test for student i belonging to university j

X_{2ij}: is the observed Spanish score in the Saber 11 test for student i belonging to university j

X_{3ij}: is the observed Chemistry score in the Saber 11 test for student i belonging to university j

X_{4ij}: is the observed Social Sciences score in the Saber 11 test for student i belonging to university j

X_{5ij}: is the average Mathematics score in the Saber 11 test for student i belonging to university j

X_{6ij}: is the average Spanish score in the Saber 11 test for student i belonging to university j

X_{7ij}: is the average Chemistry score in the Saber 11 test for student i belonging to university j

X_{8ij}: is the average Science Social score in the Saber 11 test for student i belonging to university j

Results

Figure 4.6 is a plot of Value Added estimates for Model II with Adjusted Error versus the Value Added estimates for Model II without Adjusted Error. We observe that the Value Added estimates are different in both Models II, with and without Adjusted Error. In both Models II, there are many Value Added estimates of the same sign, unlike Model I there are many Value Added estimates of the opposite sign. The explanation for this fact is that Model II includes variables at the university level (compositional effect) while Model I does not. We have more information concerning educational institutions in the Model II, therefore we do not make any mistake with estimations.

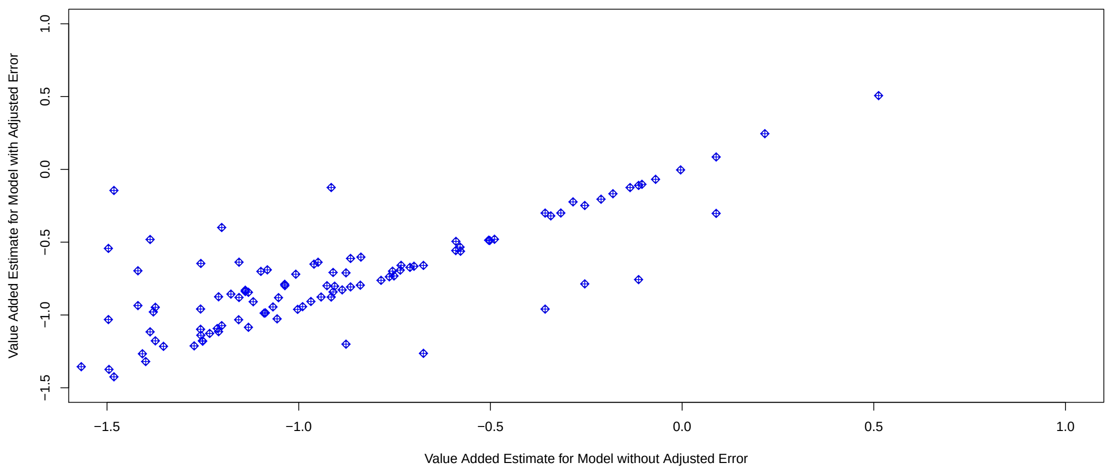

FIGURE 4.6. Model II with Adjusted Error versus Model II without Adjusted Error

Figure 4.7 shows the boxplot for both models; Model II with and without adjusted error. This boxplot is similar to the boxplot for Model I with and without adjusted error, in which the Value Added estimates present a greater variability for models without the correction of error than in models with corrected error.

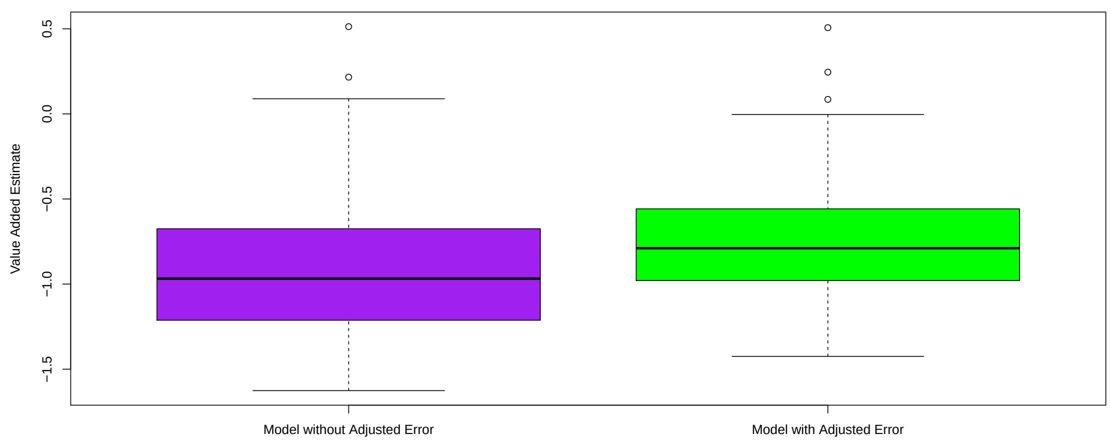

FIGURE 4.7. Boxplot of Value Added estimate for Model II with Adjusted Error and for Model II without Adjusted Error for 105 universities

Figure 4.8 shows the confidence intervals of Value Added for both models; Model II with and without Adjusted Error. The two plots have different shapes which suggest that the results obtained by adjusting the measurement error differ considerably from those obtained in the model which without adjusting measurement error. Also, it is important to note that if we do not adjust or correct the error, we will obtain incorrect interpretations concerning the Value Added for a university.

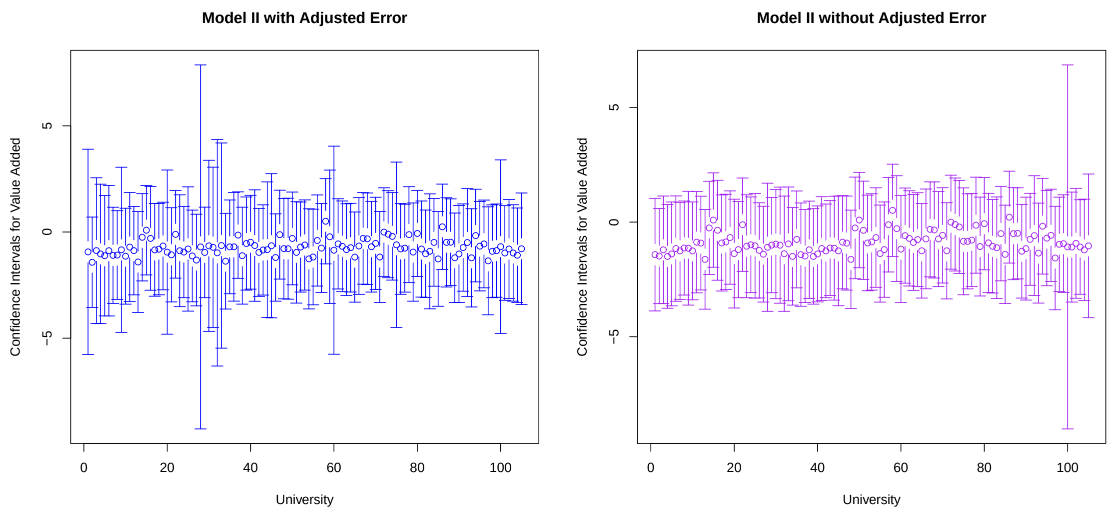

FIGURE 4.8. Confidence intervals for Value Added for 105 Universities for Model II with Adjusted Error and Model II without Adjusted Error

Table 4.5 shows the Value Added estimate and the standard error estimate for the first ten universities in both Models II; with and without Adjusted Error. It is important to note that the standard error estimates are lower for the Model II with Adjusted Error than for the Model II without Adjusted Error which means that we obtain the Value Added estimates with greater precision if we adjust or correct the measurement error. Finally, it is important to note that both the Value Added estimates and the standard error estimates for the Value Added were improved in the model, including the average scores for each knowledge area (Model II) in comparision with the model that do not include these averages (Model I). Moreover, in Model II with Adjusted Error, the standard error estimates for the Value Added are lower than in Model II without Adjusted Error. Therefore, we suggest that Model II with Adjusted Error is the best model which enables us to obtain Value Added estimates with greater precision.

TABLE 4.5. Valued Added Estimates for the first ten universities

	Model II With Adjusted Error		Model II Without Adjusted Error	
	VAEst	SE VAEst	VAEst	SE VAEst
1	-0.94	0.179	-1.42	0.592
2	-1.42	0.117	-1.48	0.166
3	-0.87	0.190	-1.21	0.521
4	-1.03	0.259	-1.50	0.816
5	-1.12	0.182	-1.39	0.425
6	-0.88	0.200	-1.16	0.510
7	-1.10	0.249	-1.26	0.456
8	-1.08	0.160	-1.13	0.224
9	-0.84	0.160	-1.14	0.406
10	-1.14	0.178	-1.26	0.289

4.3.3 Discussion on the Model Selection

It is important to emphasize that the selection of the model in a Value Added analysis is highly related to policy objectives and the characteristics of the educational system. Therefore, the statistical methods for selecting the models are not relevant when we choose the best model fitting the data in a study on Value Added indicators.

Depending on the public policy, the model that only includes Saber 11 scores, it can help to identify what academic programs are effective for the students. Whereas the model that includes both the Saber 11 scores and the average of Saber 11 scores in each university (or compositional effect as it is known in the literature concerning school effectiveness) allows us to compare pedagogical strategies among schools.

In the Colombian case, the Saber 11 and Saber Pro tests scores are used for obtaining Value Added estimates which give information about the quality of academic programs of higher educational institutions. Thus, the model that includes both the Saber 11 scores and the compositional effect is the model that fulfills the objectives concerning the quality of the educational institutions. Thus, if the objective is to compare the pedagogical strategies used in the universities, we consider Model II with Adjusted Error as the most properly.

4.3.4 Model validation and Diagnostic

In this section, we show several residual plots which are useful for checking on the validity or adequacy of the model. In particular, we verify the assumptions of homoskedasticity of the errors, normality of the errors and normality of the random effects.

Figure 4.9 presents a tendency in the observations that suggests a heteroscedastic model can be considered. Given the heterogeneity of data, the difference among the universities and the selectivity from universities, it is possible that un model that allows the variance of the error varies regarding some covariate, for example, a model whose variance of the error depends on the student's socioeconomic level or another variable that capture the variability in the Saber Pro scores. These topics mentioned above will be of interest in our future research that allows improving the apply methodology.

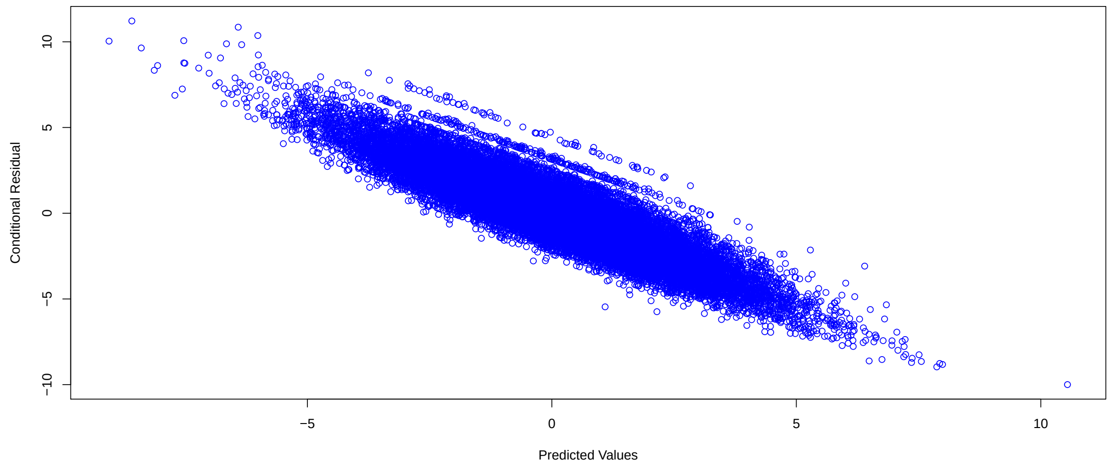

FIGURE 4.9. Conditional Residuals Versus fitted Values

Figures 4.10 and 4.11 do not indicate any deviations from a normality of both the errors and the random effects. We observe that almost all observations are on the straight line. Therefore the assumption of the normality is fulfilled.

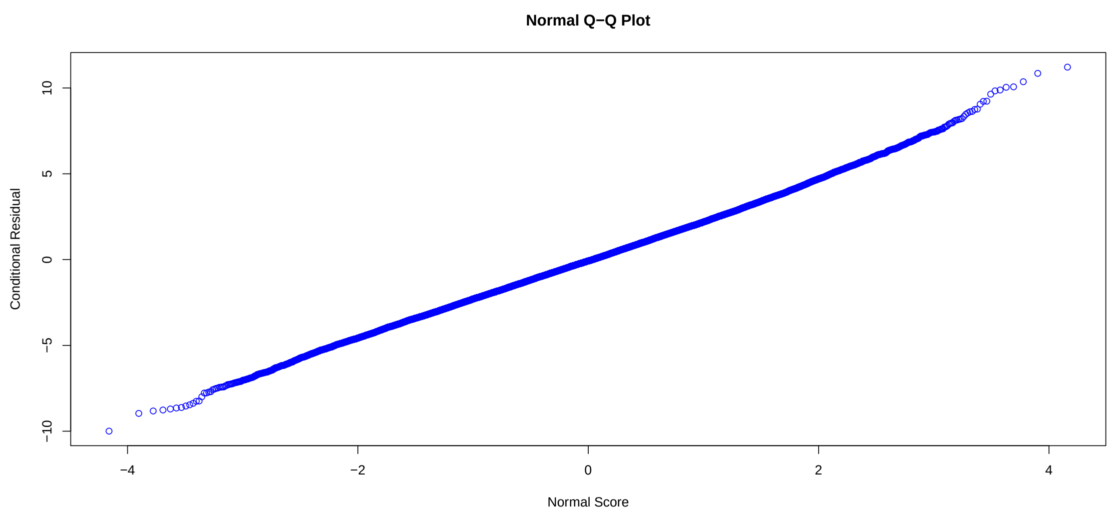

FIGURE 4.10. Normal Q-Q plot of conditional residuals

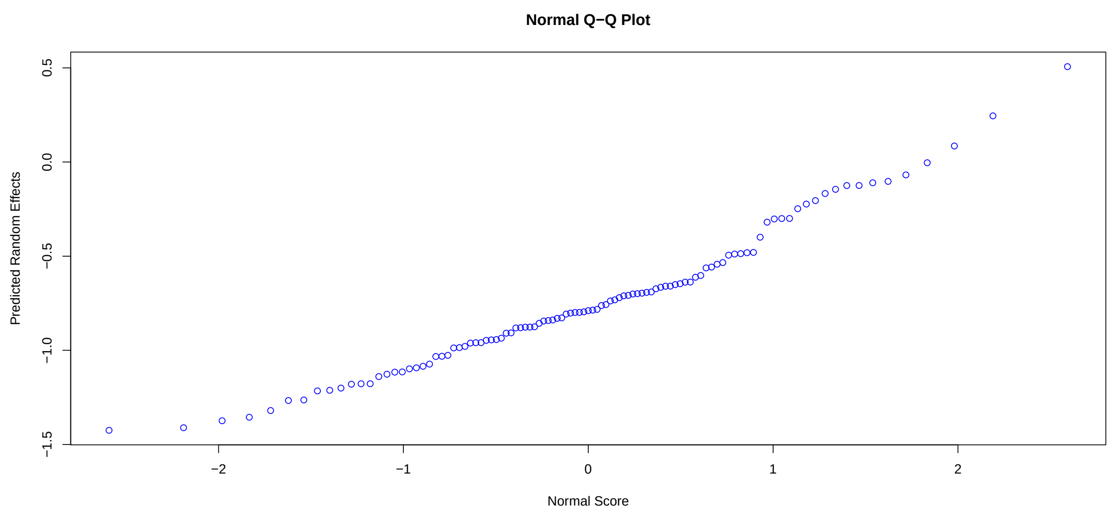

FIGURE 4.11. Normal Q-Q plot of random effects

4.4 Accreditation and Value Added for Higher Educational Institutions

Over the last two decades, the quality of higher education has been an important issue in most countries because this plays a relevant role in the social and economic development of a country. In practice, the quality of education is considered as a process which comprises of mechanisms and procedures of evaluation. One of these mechanisms mentioned above is the accreditation of institutions and their programs.

As we mentioned in previous sections, in Colombia the accreditation process is voluntary and takes into account the self-evaluation of institutions and a continued institutional improvement.

The accreditation also considers several factors mentioned as follows: (See OECD et al., 2012, pp. 183-184) [32]

- Institutional vision and goals
- Students
- Professors
- Academic processes
- Institutional Organization and administration
- Impact on society
- Institutional Infrastructure

On the other hand, there has been recent discussion concerning whether the accreditation process has to take into account student performance measurements as a result of enrollment in a specific institution as well as the Value Added indicators. And thus, we can obtain "fair measurements" for evaluating the quality of higher education so that tertiary education responds to the challenges faced in this modern society.

We present a brief analysis based on the results of this research which relates to the Value Added estimates for Colombian higher education institutions with the information regarding their accreditation in order to contribute to the discussion mentioned above. For us, it is worth mentioning that this analysis constitutes a further motivation for increasing the interest in this discussion.

We obtain the information concerning the accreditation of Colombian higher educational institutions from the official website of the National Accreditation Commission (CNA, Comisión Nacional de Acreditación). This information is public and of easy access to the general public. It also takes into account the information concerning academic programs and higher educational institutions. Due to the diversity of Colombian higher educational institutions, and Engineering programs; and the different accreditation periods for both programs and institutions, we consider: First, we select accredited Engineering programs from the year 2014 onwards because the data used in this research corresponds to students' results of the Saber Pro test from the years 2012 and 2013. Second, we calculated the average number of years of accreditation of the different engineering programs at any given university. Therefore, from 105 higher education institutions, 47 institutions fulfill the above requirements.

Figure 4.9 shows a relationship between the Accreditation Average per Year and Value Added estimate. We observe that in the Colombian higher educational institutions with negative Value Added estimates, their accreditation average per year ranged between 4 to 6 years. (The lowest accreditation per year is four) additionally the Colombian higher educational institutions with positive Value Added estimates have the accreditation per year greater than six. We observed that the universities which have better results in terms of Value Added obtain the highest number of years for their accreditation when compared with the universities that obtained lower estimation values. The above assertion clearly suggests there is a relationship between the Accreditation Average per year and the Value Added estimate and we should consider this relationship in the accreditation process for the higher educational institutions.

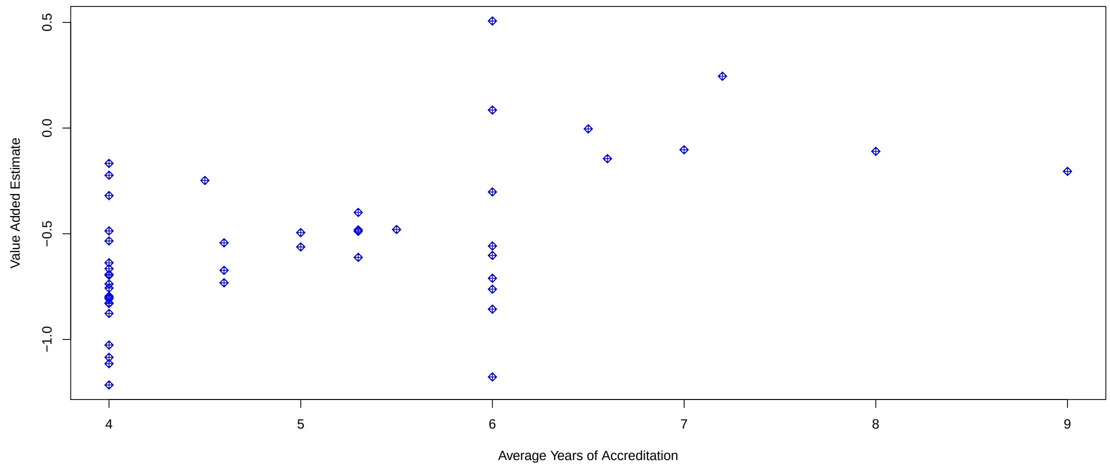

FIGURE 4.12. Value Added Estimate versus Accreditation Average per year

4.5 Final Remarks

We summarize the most relevant results as follows:

1. We find that any given university presents the negative and positive Value Added estimate whether the masurement error is adjusted or not. Thus, we can obtain wrong conclusions concerning the Value Added for a university if we do not adjust the measurement error.

2. Value Added estimates of several universities are very different and they are of an opposite sign. Also, standard error estimates for Model I without Adjusted Error are less than their corresponding standard error estimates for Model I with Adjusted Error, but standard error estimates for Model I without Adjusted Error are incorrect because we are not correcting the measurement error.

3. Model II includes explanatory variables measure at the group level (These variables are called the Compositional effect), the Value Added estimates are of the same sign in both cases with and without adjusted error.

4. For Model II when the measurement error is taken into account, standard error estimates for Value Added are less than those for Model II into Adjusted Error.

5. We obtain Value Added estimates more precise in the model that considers a compositional effect and corrects the measurement error. Therefore, this model is more appropriate for estimating the Value Added for universities.

6. If we do not adjust the measurement error, we can obtain misleading conclusions regarding the Value Added for any given university. Consequently, we cannot correctly classify universities.

7. If we have invalid inferences concerning the Value Added, we may incorrectly conclude that any given university is contributing to the progress of a student when in reality this is not so.

8. It is important to note that a heteroscedastic model can be considered with the variance of the error depending some covariate. We suggest that a model whose variance of the error varies concerning the student's socioeconomic level can be plausible. As mentioned in previous sections our new proposal will be discussed and analyzed in a future research.

9. We emphasize that the selection of the models in Value Added analysis is highly related to the policy objectives under which the Value Added estimates are required. The statistical methods are important for improving the model but these are not decisive for selecting the most proper model fitting data.

CHAPTER 5

Conclusions and Future work

This thesis is a contribution towards research studies regarding school effectiveness. In this work, we propose a methodology to resolve the measurement error problem when estimating the Value Added in multilevel models. In Chapter 2, we discuss model specification and parameter identification taking into account the measurement error of explanatory variables in the model. In Chapter 3, we present several mathematical developments for estimating the fixed parameters, the variances, and the covariances. We conclude this research presenting the main conclusions, some final remarks regarding school effectiveness and future work.

5.1 Conclusions

School effectiveness explores the differences within and between schools and it studies the relationship between explanatory and response variables using appropriate models. Generally, effectiveness studies consider the students' examination achievement as a response variable and the students' prior achievement and other socio-demographic variables as explanatory variables. See Goldstein (1997) [20].

School effectiveness attempts to identify factors which explain the differences between schools and Value Added is one of these factors. It is well known in the literature that the most appropriate models for studying Value Added are the Hierarchical Linear Mixed Models (or Multilevel Models) because these models take into account both the data structures and the correlation within groups or schools.

Value Added measures the contribution made by schools to the students' academic progress. It also can be defined as the difference between two conditional expectations: The first is the conditional expectation of the response variable given both the school effect and the explanatory variables (covariates). The second is the conditional expectation of the response variable given the covariates. See Manzi et.al (2014) [28]. Value Added is interpreted related to a reference school which is characterized by explanatory variables. This means that if we change the explanatory variable, the reference school changes and hence the interpretation of Value Added changes, too.

The explanatory variables play an important role in the defintion and interpretation of Value Added. Nevertheless, these variables are often measured with error in educational and social studies; these errors can arise from many situations, for instance, problems with the measuring instruments, incorrect definition of the variable, etc.(Goldstein, 1997) [20]. It is well known, that if the measurement errors are not properly adjusted, we can obtain biased estimators and incorrect conclusions having serious implications.

The proposals revised worked on the issue of Value Added estimation in multilevel models with measurement error in the variables and focused on several aspects: First, several authors began based on a general model with k variables, their interest focused on estimating the fixed parameters, the variances of residual error and also the variance of measurement error. They also made further research into measurement errors correlating and working with categorical independent variables. Goldstein et al (2008) [22]. Second, other authors proposed methods based on particular models of variance components with two or three explanatory variables. They estimated the fixed parameters and the variances of both the error of the residual and the random effect using estimation methods such as Bayesian estimation using the Markov Chain Montecarlo method (MCMC) Ferrao et.al (2009) [14]. Third, there were authors with a different approach to that of the first two, in which a structural approach was proposed in statistical modeling, focused on the specification of the model and they considered observations measured at two different times. Battauz et.al (2011) [2]. However, these proposals did not focus on an estimator of the Value Added of a school but rather on the associated parameters of the model.

In this thesis, we propose a novel method to address the problem of Value Added estimation in Hierarchical Linear Mixed Models with measurement error in the variables. Unlike the other proposals, we begin from a general two level model with K explanatory variables measured with error, a model of measurement error and a specific formula for Value Added for a school as suggested by Manzi et.al (2014) [28]. Also, it is important to mention that our proposal considers a general model instead of a particular model and additionally we include the specific expression for Value Added which was not taken into account by previous authors. Moreover, we consider the measurement error based on the specification of the model reaching the stage of the estimation of the parameters and thus our new method not only permits estimating the fixed parameters and variances of the residual error and the random effect, but also the Value Added of the schools. The Value Added estimator is obtained using Empirical Bayes estimators and under the new specification of the model given in our proposal. Another important contribution is the adjustment of a resampling method to our method for obtaining an uncertainty measurement for the Value Added estimator.

We illustrate our new proposed method with an application to a Colombian database. We observe that the Value Added estimators and all the estimators of the models are affected whether the measurement error is adjusted or not. We also note that both the estimates and the standard errors of the estimators change considerably in the models where the measurement error is corrected. Additionally, in the model that includes the explanatory variable at the school level with measurement error, the estimates improve significantly than in the model that did not include this variable. That is, we obtain lower standard errors of the estimators in the model with corrected measurement error than in the model without the corrected measurement error.

Finally, we consider that in educational studies, it is necessary to correct the measurement error in order to obtain more reliable and more precise Value Added estimators so as to lead to valid conclusions that permit making correct decisions in a given situation. It is also important that educational studies report measurements associated with the measurement error such as the reliability values.

5.2 Future work

We present some possible extensions of this research as follows:

- **Application to Small Populations**: To estimate the Value Added in models with measurement error when the number of students is small. We should have some special considerations in this case, in the estimations of the Value Added because the theory of larges samples will not apply in this case.
- **Study of estimator properties** : To compare several estimation strategies for the fixed parameters, the variance and the covariance in the models with error and for obtaining the standard error of these estimations.
- **Development of optimization algorithms**: To develop of optimization algorithms for obtaining the estimators of the fixed parameters, the variance and the covariance which are required for the estimation of Value Added.
- **Complex model for measurement error**: To assume a complex model for measurement error different from the classical model of measurement error and to compare both the classical and complex models.

APPENDIX A

Appendix of Chapter 2

A.1 Notation

Y_{ij} be the response variable of the i-th level one unit that belongs to the j-th level two unit.

η_{ij} is a $K \times 1$ unobservable true variables vector.

X_{ij} is a $K \times 1$ explanatory variables vector of the i-th level one unit that belongs to the j-th level two unit.

θ_j is the effect of the j-th level two unit.

$Y_j = (Y_{1j}, Y_{2j}, ..., Y_{n_j j})^T$ is an $n_j \times 1$ vector.

$X_j = (X_{1j}, X_{2j}, ..., X_{n_j j})^T$ is an $n_j \times K$ matrix.

$\eta_j = (\eta_{1j}, \eta_{2j}, ..., \eta_{n_j j})^T$ is an $n_j \times K$ matrix.

$Y = (Y_1^T, Y_2^T, ..., Y_J^T)^T$ is an $N \times 1$ vector.

$X = (X_1^T, X_2^T, ..., X_J^T)^T$ is an $N \times K$ matrix.

$\theta = (\theta_1, \theta_2, \ldots \theta_J)^T$ is is a $J \times 1$ vector.

$\eta = (\eta_1^T, \eta_2^T, \ldots \eta_J^T)^T$ is an $N \times K$ matrix.

A.2 Variance of (Y_j)

$$
\begin{aligned}
V(Y_j) &= V(E(Y_j \mid \eta_j, \theta_j)) + E(V(Y_j \mid \eta_j, \theta_j)) \\
&= V(\alpha 1_{n_j} + \eta_j \beta + \theta_j 1_{n_j}) + E(\sigma^2 I_{n_j}) \\
&= V(\eta_j \beta) + V(\theta_j 1_{n_j}) + \sigma^2 I_{n_j} \\
&= V(\mathrm{vec}(\eta_j \beta)) + V(\theta_j 1_{n_j}) + \sigma^2 I_{n_j} \\
&= V(\mathrm{vec}(I_{n_j} \eta_j \beta)) + Var(\theta_j 1_{n_j}) + \sigma^2 I_{n_j} \\
&= V(\beta^T \otimes I_{n_j} \mathrm{vec}(\eta_j)) + V(\theta_j 1_{n_j}) + \sigma^2 I_{n_j} \\
&= (\beta^T \otimes I_{n_j})(\Psi_{k\times k} \otimes I_{n_j})(\beta^T \otimes I_{n_j})^T + \omega^2 J_{n_j} + \sigma^2 I_{n_j} \\
&= (\beta^T \Psi_{K\times K} \beta \otimes I_{n_j}) + \omega^2 J_{n_j} + \sigma^2 I_{n_j} \\
&= \beta^T \Psi_{K\times K} \beta I_{n_j} + \omega^2 J_{n_j} + \sigma^2 I_{n_j} \\
&= (\phi + \sigma^2) I_{n_j} + \omega^2 J_{n_j}
\end{aligned}
$$

A.3 Covariance of $(Y_j, \mathbf{Vec}(X_j))$

$$
\begin{aligned}
Cov(Y_j, \mathrm{Vec}(X_j)) &= Cov(E(Y_j \mid \eta_j, \theta_j), \mathrm{vec}(X_j)) \\
&= Cov(\alpha 1_{n_j} + \eta_j \beta + \theta_j 1_{n_j}, \mathrm{vec}(X_j)) \\
&= Cov(\eta_j \beta, E(\mathrm{vec}(X_j) \mid \mathrm{vec}(\eta_j))) \\
&= Cov(\eta_j \beta, \mathrm{vec}(\eta_j)) \\
&= Cov(I_{n_j} \eta_j \beta, \mathrm{vec}(\eta_j)) \\
&= Cov(\beta^T \otimes I_{n_j} \mathrm{vec}(\eta_j), \mathrm{vec}(\eta_j)) \\
&= (\beta^T \otimes I_{n_j}) Var(\mathrm{vec}(\eta_j)) \\
&= (\beta^T \otimes I_{n_j})(\Psi_{K\times K} \otimes I_{n_j}) \\
&= (\beta^T \Psi_{K\times K} \otimes I_{n_j})
\end{aligned}
$$

A.4 Variance of (X_j)

$V(X_j) = V(\text{vec}(X_j)) = (T_{K\times K} + \Psi_{K\times K}) \otimes I_{n_j}$

A.5 Definition of Value Added - First Conditional Expectation

$$\begin{aligned}
E(Y_{ij}|X_{ij},\theta_j) &= E[E(Y_{ij}|\eta_{ij},\theta_j)|X_{ij},\theta_j] \\
&= E[(\alpha+\eta_{ij}^T\beta+\theta_j)|X_{ij},\theta_j] \\
&= \alpha + E(\eta_{ij}^T\beta|X_{ij},\theta_j) + E(\theta_j|X_{ij},\theta_j) \\
&= \alpha + E(\eta_{ij}^T\beta|X_{ij}) + \theta_j
\end{aligned}$$

A.6 Definition of Value Added - Second Conditional Expectation

$$\begin{aligned}
E(Y_{ij}|X_{ij}) &= E(E(Y_{ij}|X_{ij},\theta_j)|X_{ij}) \\
&= E[\alpha + E(\eta_{ij}^T\beta|X_{ij}) + \theta_j|X_{ij}] \\
&= \alpha + E(\eta_{ij}^T\beta|X_{ij}) + E(\theta_j|X_{ij}) \\
&= \alpha + E(\eta_{ij}^T\beta|X_{ij})
\end{aligned}$$

APPENDIX B

Appendix of Chapter 3

B.1 Conditional Expectation of $(Y_j|X_j, \theta_j)$

$$
\begin{aligned}
E(Y_j|X_j,\theta_j) &= E[E(Y_j|\eta_j,\theta_j)|X_j,\theta_j] \\
&= E[\alpha 1_{n_j} + \eta_j\beta + \theta_j 1_{n_j}|X_j,\theta_j] \\
&= \alpha 1_{n_j} + E(\eta_j\beta|X_j,\theta_j) + \theta_j 1_{n_j} \\
&= \alpha 1_{n_j} + (\beta^T\Psi \otimes I_{n_j})[(\Psi + T)\otimes I_{n_j}]^{-1}\mathrm{vec}(X_j) + \theta_j 1_{n_j} \qquad \text{(B.1)}
\end{aligned}
$$

B.2 Variance of $(Y_j|X_j)$

$$
\begin{aligned}
V(Y_j|X_j) = V(Y_j|\mathrm{vec}(X_j)) = \omega^2 J_{n_j} + \sigma^2 I_{n_j} + (\beta^T\Psi\beta \otimes I_{n_j}) \\
- (\beta^T\Psi \otimes I_{n_j})[(\Psi+T)\otimes I_{n_j}]^{-1}(\Psi^T\beta \otimes I_{n_j})
\end{aligned}
$$

B.3 Covariance of $(Y_j, \theta_j | X_j)$

$$
\begin{aligned}
Cov(Y_j, \theta_j | X_j) &= Cov(E(Y_j | X_j, \theta_j), \theta_j) | X_j \\
&= Cov(\alpha 1_{n_j} + (\beta^T \Psi \otimes I_{n_j})[(\Psi + T) \otimes I_{n_j}]^{-1} \mathrm{vec}(X_j) + \theta_j 1_{n_j}, \theta_j | X_j) \\
&= Cov(\theta_j 1_{n_j}, \theta_j | X_j) \\
&= Cov(\theta_j 1_{n_j}, \theta_j) \\
&= \omega^2 1_{n_j}
\end{aligned}
$$

Bibliography

[1] Douglas M Bates and José C Pinheiro, *Computational methods for multilevel modelling*, University of Wisconsin, Madison, WI (1998), 1–29.

[2] Michela Battauz, Ruggero Bellio, and Enrico Gori, *Covariate measurement error adjustment for multilevel models with application to educational data*, Journal of Educational and Behavioral Statistics **36** (2011), no. 3, 283–306.

[3] Dennis D Boos et al., *Introduction to the bootstrap world*, Statistical science **18** (2003), no. 2, 168–174.

[4] John Buonaccorsi, Eugene Demidenko, and Tor Tosteson, *Estimation in longitudinal random effects models with measurement error*, Statistica Sinica (2000), 885–903.

[5] James R Carpenter, Harvey Goldstein, and Jon Rasbash, *A novel bootstrap procedure for assessing the relationship between class size and achievement*, Journal of the Royal Statistical Society: Series C (Applied Statistics) **52** (2003), no. 4, 431–443.

[6] Raymond J Carroll, David Ruppert, Leonard A Stefanski, and Ciprian M Crainiceanu, *Measurement error in nonlinear models: a modern perspective*, CRC press, 2006.

[7] Andrew Chesher, *The effect of measurement error*, Biometrika **78** (1991), no. 3, 451–462.

[8] Hengjian Cui, Kai W Ng, and Lixing Zhu, *Estimation in mixed effects model with errors in variables*, Journal of multivariate analysis **91** (2004), no. 1, 53–73.

[9] Anthony Christopher Davison and David Victor Hinkley, *Bootstrap methods and their application*, vol. 1, Cambridge university press, 1997.

[10] Arthur P Dempster, Donald B Rubin, and Robert K Tsutakawa, *Estimation in covariance components models*, Journal of the American Statistical Association **76** (1981), no. 374, 341–353.

[11] Thomas J DiCiccio and Bradley Efron, *Bootstrap confidence intervals*, Statistical science (1996), 189–212.

[12] B Efron, *Bootstrap methods: another look at*, (1979).

[13] Bradley Efron and Robert J Tibshirani, *An introduction to the bootstrap*, CRC press, 1994.

[14] Maria Eugénia Ferrão and Harvey Goldstein, *Adjusting for measurement error in the value added model: evidence from portugal*, Quality & Quantity **43** (2009), no. 6, 951–963.

[15] Christopher A Field and Alan H Welsh, *Bootstrapping clustered data*, Journal of the Royal Statistical Society: Series B (Statistical Methodology) **69** (2007), no. 3, 369–390.

[16] Jean-Pierre Florens, Vêlayoudom Marimoutou, and Anne Péguin-Feissolle, *Econometric modeling and inference*, Cambridge University Press, 2007.

[17] Wayne A Fuller, *Measurement error models*, vol. 305, John Wiley & Sons, 2009.

[18] Andrew Gelman and Jennifer Hill, *Data analysis using regression and multilevel/hierarchical models*, Cambridge University Press, 2006.

[19] Harvey Goldstein, *Multilevel mixed linear model analysis using iterative generalized least squares*, Biometrika **73** (1986), no. 1, 43–56.

[20] ______, *Methods in school effectiveness research*, School effectiveness and school improvement **8** (1997), no. 4, 369–395.

[21] ______, *Multilevel statistical models*, vol. 922, John Wiley & Sons, 2011.

[22] Harvey Goldstein, Daphne Kounali, and Anthony Robinson, *Modelling measurement errors and category misclassifications in multilevel models*, Statistical Modelling **8** (2008), no. 3, 243–261.

[23] H Kim and Diane Lalancette, *Literature review on the value-added measurement in higher education*, Paris, France: OECD. Retrieved May **2** (2013), 2015.

[24] Milorad Kovacevic, Rong Huang, and Yong You, *Bootstrapping for variance estimation in multi-level models fitted to survey data*, ASA Proceedings of the Survey Research Methods Section (2006), 3260–3269.

[25] Ita GG Kreft, Jan De Leeuw, and Leona S Aiken, *The effect of different forms of centering in hierarchical linear models*, Multivariate behavioral research **30** (1995), no. 1, 1–21.

[26] Nan M Laird and James H Ware, *Random-effects models for longitudinal data*, Biometrics (1982), 963–974.

[27] Nicholas T Longford, *On estimating standard errors in multilevel analysis*, Journal of the Royal Statistical Society: Series D (The Statistician) **49** (2000), no. 3, 389–398.

[28] Jorge Manzi, Ernesto San Martín, and Sébastien Van Bellegem, *School system evaluation by value added analysis under endogeneity*, Psychometrika **79** (2014), no. 1, 130–153.

[29] Ernesto San Martín and Fernando Quintana, *Consistency and identifiability revisited*, Brazilian Journal of Probability and Statistics (2002), 99–106.

[30] ABDERRAHIM OULHAJ-MICHEL MOUCHART, *The role of the exogenous randomness in the identification of conditional models*, Metron **64** (2006), no. 2, 253–271.

[31] Michel Mouchart and Renzo Orsi, *Building a structural model: parameterization and structurality*, Econometrics **4** (2016), no. 2, 23.

[32] OECD, *2012, reviews of national policies for education: Tertiary education in colombia*, OECD, Publishing, 2012.

[33] Laura E. Rumbley Philip G. Altbach, Liz Reisberg, *Trends in global higher education:tracking an academic revolution*, UNESCO, 2009.

[34] José C Pinheiro and Douglas M Bates, *Mixed-effects models in s and s-plus*, Statistics and computing (1978).

[35] Stephen W Raudenbush, *What are value-added models estimating and what does this imply for statistical practice?*, Journal of Educational and Behavioral Statistics **29** (2004), no. 1, 121–129.

[36] Stephen W Raudenbush and Anthony S Bryk, *Hierarchical linear models: Applications and data analysis methods*, vol. 1, Sage, 2002.

[37] Stephen W Raudenbush and JDouglas Willms, *The estimation of school effects*, Journal of educational and behavioral statistics **20** (1995), no. 4, 307–335.

[38] Charles S Roehrig, *Conditions for identification in nonparametric and parametric models*, Econometrica: Journal of the Econometric Society (1988), 433–447.

[39] Ernesto San Martín, Jorge González, and Francis Tuerlinckx, *Identified parameters, parameters of interest and their relationships*, (2009).

[40] Carl P Simon and Lawrence Blume, *Mathematics for economists*, vol. 7, Norton New York, 1994.

[41] Carmen D Tekwe, Randy L Carter, Chang-Xing Ma, James Algina, Maurice E Lucas, Jeffrey Roth, Mario Ariet, Thomas Fisher, and Michael B Resnick, *An empirical comparison of statistical models for value-added assessment of school performance*, Journal of Educational and Behavioral Statistics **29** (2004), no. 1, 11–36.

[42] Hoai-Thu Thai, France Mentré, Nicholas HG Holford, Christine Veyrat-Follet, and Emmanuelle Comets, *A comparison of bootstrap approaches for estimating uncertainty of parameters in linear mixed-effects models*, Pharmaceutical statistics **12** (2013), no. 3, 129–140.

[43] AB Tom, Tom AB Snijders Roel J Bosker, and Roel J Bosker, *Multilevel analysis: an introduction to basic and advanced multilevel modeling*, Sage, 1999.

[44] UNESCO, *Education 2030,towards inclusive and equitable quality education and lifelong learning for all*, UNESCO, 2015.

[45] Rien Van der Leeden, FMTA Busing, and Erik Meijer, *Applications of bootstrap methods for two-level models*, Multilevel Conference, Amsterdam, 1997.

[46] G Woodhouse, *Errors in variables in multilevel models*, London: Institute of Education (1996).

[47] Geoffrey Woodhouse, Min Yang, Harvey Goldstein, and Jon Rasbash, *Adjusting for measurement error in multilevel analysis*, Journal of the Royal Statistical Society. Series A (Statistics in Society) (1996), 201–212.

[48] Xu-Ping Zhong, Wing-Kam Fung, and Bo-Cheng Wei, *Estimation in linear models with random effects and errors-in-variables*, Annals of the Institute of Statistical Mathematics **54** (2002), no. 3, 595–606.