2015/45

Higher education value added using multiple outcomes

Joniada Milla, Ernesto San Martín
and Sébastien Van Bellegem

CORE

DISCUSSION PAPER

Center for Operations Research
and Econometrics

Voie du Roman Pays, 34
B-1348 Louvain-la-Neuve
Belgium
http://www.uclouvain.be/core

CORE
Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve, Belgium.
Tel (32 10) 47 43 04
Fax (32 10) 47 43 01
E-mail: immaq-library@uclouvain.be
http://www.uclouvain.be/en-44508.html

CORE DISCUSSION PAPER
2015/45

Higher Education Value Added Using Multiple Outcomes*

Joniada Milla[a], Ernesto San Martín[a,b,c] and Sébastien Van Bellegem[a]

October 21, 2015

Abstract

We build a multidimensional value added model to analyze jointly the test scores on several outcomes. Using a unique Colombian data set on higher education within a seemingly unrelated regression equations (SURE) framework we estimate school-outcome specific value added indicators. These are used to measure the relative contribution of the school on a certain outcome, which may serve as an internal accountability measure. Apart from the evident estimation efficiency gains, a joint value added analysis is preferable to the unidimensional one. First, unless modeled in a multidimensional framework, the comparison of value added estimates for different outcomes within a school is not well defined; our model circumvents this issue. Second, even in the case of a separate major field of study analysis there still exists unobserved heterogeneity due to institutional diversity. This makes it more compelling to employ a rich set of outcomes in computing value added indicators. In the end, we aggregate the outcome-specific value added estimates to produce a composite value added index that reflects the combined value added contribution of all the subjects for each school.

Keywords: multidimensional value added, multiple outcomes, quality of higher education

JEL-Classification: I23, A22, C31, C51

---

* This research is the result of a scientific collaboration between the authors and the Instituto Colombiano para la Evaluación de la Educación (ICFES). Partial financial support of the ANILLO Project SOC 1107 and the FONDECYT Project N° 1141030 from the Chilean Government are gratefully acknowledged. Previous versions of this paper have been presented during the 2014 AERA conference in Philadelphia, Pennsylvania, U.S.A; the 2015 NCME conference in Chicago, U.S.A.; the 2015 AEFP conference in Washington D.C., U.S.A.; and the 2015 Workshop on *School Effectiveness and Value Added Models: Recent Progress on Modeling and Interpretation* in Santiago, Chile. Comments from the participants of the meetings are greatly appreciated. The authors thank Julian Mariño, Adriana Molina Mantilla, Richard Shavelson, Benjamin W. Domingue and Henry Braun for constructive feedback on earlier versions of this paper. Address correspondence to Joniada Milla, Voie du Roman Pays, 34, L1.03.01, B-1348 Louvain-la-Neuve, Belgium. E-mail: joniada.milla@uclouvain.be.
[a] Center for Operations Research and Econometrics (CORE), Univeristé catholique de Louvain, Belgium.
[b] Faculty of Mathematics and Faculty of Education, Pontificia Universidad Católica de Chile, Chile.
[c] Measurement Center MIDE UC, Chile

1 Introduction

Measuring the quality of higher education is an important but challenging task in policy analysis. The rationale of its increasing popularity relates to its ability of overcoming the asymmetric information in the typical principal-agent problem within the educational provision framework (Figlio & Loeb, 2011). Since governments rely increasingly more on student academic performance to evaluate school accountability for educational decision making, the precision in the estimation of accountability measures is of primary importance. In the literature it is common to analyze separately the test scores of different subjects or modules, where each is a dependent variable. This leads to a separate value added analysis for each individual test score. We propose to analyze jointly the test scores on several subjects. Our primary purpose in this article is to build a value added model (VAM) that analyzes these dependent variables together by modeling the inherent correlation among them.

There are several reasons why a joint analysis is preferable to a separate value added analysis. First, schools contribute to the human capital of the students by developing several aspects of their cognitive and social skills, which renders the dependent variables inherently and structurally linked. Usually the researcher is limited to measuring a single skill of the students by means of a test score, which lacks a complete representation of the school or teacher effects (Sammons, Nuttall, & Cuttance, 1993; Hill & Rowe, 1996).

Second, the data availability allows us to conduct the value added analysis separately by major field of study. Due to institutional diversity there is a lot of unobserved heterogeneity even within field of study. This is because different schools prepare students differently based on the weight they put e.g. on theoretical versus practical training. For instance, some schools within a certain field of study are stronger in developing critical thinking in students and others invest more effort in developing quantitative skills. To make a fair judgment in estimating value added indicators for each of these schools it is preferable to have a score that measures each of the two skills. This is another compelling rationale on why we need to consider multiple outcomes.

Third, the value added estimates are sensitive to the model used, but they are even more sensitive to the outcome considered. This issue is raised in Lockwood et al. (2007) and analyzed further in Papay (2011) for different items of the same test. This is because each outcome provides to some extent new information from the other and pulls the school-effect in a different direction. Hence, there is need to develop a value added estimate that uses all possible sources of information (test scores on multiple outcomes) about the student.

Fourth, when comparing value added indicators produced from different outcomes, the common practice prevailing in VAM literature ignores the fact that the academic outcomes have an inherent correlation among them. This renders comparison of value added within school for different outcomes not possible, unless estimated in a joint framework. Therefore, it is necessary to account for this correlation when compiling value added estimates of multiple outcomes, which is an important task that has informative value to the institutions for internal accountability purposes.

From the above discussion it follows that a joint analysis is desirable because it delivers more powerful statistical estimates and produces more powerful tests than in the unidimensional case. The higher the correlation among the outcomes, the higher is the additional power of the test (Snijders & Bosker, 2012, p.283). Finally, the multidimensional model enables us to compare the values of the parameter estimates. For instance, we can compare the marginal effect of a certain independent variable on the various dependent variables used in the model. The statistical comparison on the magnitude of the marginal effects can be provided by the multidimensional model.

In the literature, it is now well accepted that for a fair comparison of the effectiveness between schools, it is crucial to condition the estimation on a set of own student background, among which a prior attainment score that precedes the outcome of interest (Kupermintz, 2003; Ballou, Sanders, & Wright, 2004; Martineau, 2006; OECD, 2008; Lenkeit, 2012). Another point that we emphasize in this paper is the common practice of using the lagged dependent variable (or a similar test score taken earlier than the outcome) as covariates in a hierarchical linear model (HLM), given that both the current and the lagged test scores are measured during the time that the student was conducting his/her studies in the same school. Relevant references among many is the popular TVAAS introduced in Sanders, Saxton, and Horn (1997), McCaffrey, Lockwood, Koretz, Louis, and Hamilton (2004) and papers reviewed therein and the more recent Rothstein (2010) and Chetty, Friedman, and Rockoff (2014). In this paper we use a unique administrative, Colombian data set to estimate a university value added model with multiple outcomes. In our data, the lagged test score is measured *before* the students enter the institutions for which the value added indicators are being estimated. Our lagged test score is an entry exam (university admission score) that students take upon graduating high school, and the outcome is an exit exam score that student write close to completing the requirements to graduate from a university program. Hence, the school effect in our case is by construction independent of the lagged test score used as a covariate and as a measure for students ability.

The multidimensional value added model that we develop allows us to measure two important quantities for policy purposes: (1) the value added of a university on a single outcome that we call the *outcome-specific value added*, and (2) the across-outcome average value added for a given university that we call the *composite value added*. The outcome-specific value added is notably different from those estimated from a single-outcome (or unidimensional) value added model. The fundamental difference is that it is constructed after a marginalization of the multidimensional model, and therefore it takes into account all the available information also coming from the correlation with the other outcomes. It is important to emphasize that these outcome-specific value added indicators are comparable between them and, therefore, provide information that can be used for internal accountability purposes. The composite value added index that we build is a synthesis of all the information provided by the outcome-specific value added indicators.

A much discussed issue in VAMs is related to the question of whether the value added has a causal effect interpretation (Rubin, Stuart, & Zanutto, 2004; Schatz, Von Secker, & Alban, 2005; Kane & Staiger, 2008; Rothstein, 2010; Koedel & Betts, 2011; Kinsler, 2012; Chetty et al., 2014; Rothstein, 2014; Bacher-Hicks, Kane, & Staiger, 2014). A comprehensive literature review on causal inference of value added models and beyond can be found in Koedel, Mihaly, and Rockoff (2015). One factor that is believed to confound the value added estimates is the self-selection or sorting of the students in the post-secondary programs and universities. Chetty et al. (2014) and Lenkeit (2012) show that the prior attainment scores take care of most of the bias in value added indicators. However we can not rule out the possibility that there may exist unobserved characteristics of the students and schools in the sorting process (Rothstein, 2009). The administrative data we use provides information on the age, gender, number of semesters since entry into the university before the student took the exit test (measure for time to graduation), a socio-economic status variable known as INSE (*Indice de Nivel Socio-Económico*), and the type of the institution they are attending (whether a university or other institution that offers vocational, professional training). We incorporate these variables in our estimation in order to control for the observable heterogeneity of the students and approximate for its unobserved part.

To our best knowledge, the only paper that also estimates a value added assessment model using multiple outcomes is Broatch and Lohr (2012). Our model departs from this paper in several aspects. First, we focus on institution rather than teacher value added models. Hence the structure of the model is different. Second, in our case all outcomes of interest are continuous random variables rather than categorical or dummy variables. This makes the choice of the estimator and the estimation procedure fundamentally different. As a matter of fact, and despite its apparent simplicity, the hierarchical linear model used in our approach is not standard and we could not find any implemented procedure in common statistical softwares (as STATA or SAS) for its estimation. Our approach uses a two-nested-level model in which the micro-level model is multidimensional. The model can be seen as the combination of the common two-level hierarchical model together with the Seemingly Unrelated Regression (SUR) model (Zellner, 1962). We develop an estimation procedure based on the method of moments.

The case of Colombia is interesting regarding the availability of a national exam at the end of the undergraduate studies (called "pregrado" in Colombia), and for this reason it is relevant to discuss briefly the external validity of our value added model. In fact, the model can be applied to any set of continuous outcomes which are not necessarily related to the higher education. Through the case study of Colombia we show that there exist striking differences between the value added estimates produced by the conventional single-outcome models versus multiple-outcome models. This result encourages the application of exit examinations on several domains, if the aim is to evaluate or rank a multi-field institution such as the tertiary education colleges and universities.

The rest of the paper is organized as follows. In Section 2, we recall the notion of value added and construct the multidimensional model. We then derive its statistical properties and the identification of the parameters of interest. We define the value added estimator in Section 3. Section 4 overviews the identification and estimation procedure. We present a Monte Carlo simulation exercise in Section 4.2. In Section 5 we review briefly the Colombian educational system and we describe the data in Section 6. The results are presented in Section 7. Section 8 concludes.

2 The Heteroscedastic Multidimensional Hierarchical Linear Mixed Model (HMHLM)

2.1 School-outcome effects

In this section we provide the intuition for the structural model that we use to describe the data. In our data each student is submitted to an entry examination (at the end of the secondary school) and also writes a national exam upon graduation in his tertiary institution. Each of the exams contain several outcomes (e.g. quantitative reasoning, critical reading, language, etc.). In addition, we observe a set of characteristics at the student and school level. We denote by Y_{imj} the score in the exit exam in test m of student i who belongs to the school j, with $m \in \{1, \dots, M\}$, $i \in \{1, \dots, n_j\}$ and $j \in \{1, \dots, J\}$. So, in the exit exam we have M tests, n_j is the number of students in the university j that wrote the test, and J is the total number of tertiary institutions in our sample. Similarly, for every i, m, j we denote by Z_{imj} the column vector of size K that contains the scores of the entry exam and all other covariates, including the possible intercept.

It should be emphasized that, although three indices are considered in the notation for Y_{imj} and Z_{imj}, the data is not designed as a three-level nested system. In every university j, the student outcome is a vector of scores (one for each test of the exit exam). In our model we assume two hierarchical levels: the coarsest level is the institution, and the detailed level is the student.

Since a vector of M test scores is observed at the student level, it is possible to identify a latent source of variation at the school-outcome level. We assume that random variables, denoted by γ_{mj}, model the unobserved heterogeneity arising across schools and outcomes, that is, the γ_{mj}'s explain the heterogeneity that is present in the M test scores and that it is not explained by the observed characteristics of students and universities. Consequently, we assume that the test scores are independent between students conditional on γ_{mj} and the covariates, an assumption typically called *axiom of local independence* (Lazarsfeld, 1950). If $A^\top$ denotes the transpose of the matrix A, we construct the $n_j \times K$ design matrix of covariates $Z_{mj}^\top = (Z_{1mj}, \dots, Z_{n_j mj})$. The local independence assumption then writes

$$\mathop{\perp\!\!\!\perp}_{1\leq i\leq n_j} Y_{imj} \mid \gamma_{mj}, Z_{mj}$$

for every test m and institution j. In order to be more specific on the structure of the model, we assume that the expected final attainment score is individual-wise linearly dependent on the covariates and the school effect. More precisely, for every test m taken by a student i belonging to an institution j, we assume there exists a vector of K parameters, denoted by β_m such that

$$E(Y_{imj} \mid Z_{imj}, \gamma_{mj}) = Z_{imj}^{\top}\beta_m + \gamma_{mj},$$

which can equivalently be rewritten as

$$Y_{imj} = Z_{imj}^{\top}\beta_m + \gamma_{mj} + u_{imj},$$

where the idiosyncratic error u_{imj} is defined as $u_{imj} := Y_{imj} - E(Y_{imj} \mid Z_{imj}, \gamma_{mj})$. It is assumed that the idiosyncratic errors are independent across student, test and institution, and independent from any Z_{mj} and any γ_{mj}. The vector of K parameters, β_m, depends on m meaning that the marginal effect of the prior score or other covariates may differ for each outcome, but not across schools. The latent variable, γ_{mj}, is assumed to be independent across schools, but a non-vanishing correlation is allowed within school among two different outcomes. The next subsection summarizes the structural model and the dependency structure assumed among all variables in the model.

2.2 Structural model

Denote by $Y_{mj} = (Y_{1mj}, Y_{2mj}, \ldots, Y_{n_j mj})^{\top}$ the n_j-dimensional column vector of final scores in outcome m in school j. For every school, the multidimensional model can be written as

$$E(Y_{mj} \mid Z_{mj}, \gamma_{mj}) = Z_{mj}\beta_m + \gamma_{mj}\iota_{n_j} \tag{2.1}$$

where $\iota_{n_j} = (1, 1, \ldots, 1)^{\top}$ is a n_j-dimensional column vector of ones. The vector $u_{mj} = (u_{1mj}, \ldots, u_{n_j mj})^{\top}$ of idiosyncratic errors is accordingly defined as $u_{mj} := Y_{mj} - E(Y_{mj} \mid$

Z_{mj}, γ_{mj}).

Define Y_j a vector of size Mn_j which contains the stacked data vectors for all tested outcomes within a school j, i.e., $Y_j = (Y_{1j}^\top, \ldots, Y_{Mj}^\top)^\top$. We also define the vector $\gamma_j = (\gamma_{1j}, \ldots, \gamma_{Mj})^\top$ and the matrix of covariates at the school level to be the block diagonal matrix $Z_j = diag\{Z_{1j}, Z_{2j}, \ldots, Z_{Mj}\}$. Finally we stack all schools in the vector $Y = (Y_1^\top, \ldots, Y_J^\top)^\top$ of dimension MN with $N = \sum_j n_j$, the total number of students. Similarly, we define the $MN \times MK$ matrix of covariates $Z = (Z_1^\top, \ldots, Z_J^\top)^\top$ and the vector of parameters $\beta = (\beta_1^\top, \ldots, \beta_M^\top)^\top$ of size MK. The resulting model is

$$E(Y \mid Z, \gamma) = Z\beta + H^\top \gamma, \tag{2.2}$$

where H denotes the appropriate matrix of dimension $MJ \times MN$ with entries 0 and 1 (for details, see Technical Appendix section B.2). The idiosyncratic error is finally defined as $u := Y - E(Y \mid Z, \gamma)$. Equation (2.2) is complemented with the following structural assumptions:

(A.1) Exogeneity. The matrix of covariates Z is independent of the vector of random effects γ.

(A.2) Independence and Heteroscedasticity of the Random Effects. The school-outcome random effects are such that $(\gamma_j \mid Z_j) \sim \text{ID}\,(0, \Lambda_j)$, where Λ_j are $M \times M$ positive definite matrices; that is, the γ_j's are mutually independent and the distribution of each γ_j have mean 0 and variance-covariance matrix Λ_j.

(A.3) IID and Homoscedasticity of the idiosyncratic error. The error term is such that $(u_j \mid Z_j, \gamma_j) \sim \text{IID}\,(0, \sigma^2 I_{Mn_j})$ for some $\sigma^2 > 0$; that is, the u_j's are independent and identically distributed with a common distribution that have mean 0 and variance-covariance matrix $\sigma^2 I_{Mn_j}$.

(A.4) Local independence. $\underset{1 \le i \le n_j}{\perp\!\!\!\perp} Y_{imj} \mid \gamma_{mj}, Z_{mj}$

The exogeneity (A.1) is a common assumption in the context of HLM. In the context of value added for tertiary education in Colombia, this assumption is reasonable because the prior attainment score is measured before the students enter university. In other cases, this assumption is questionable and its violation leads to a more complex interpretation of the value added. This is discussed in detail in Manzi, San Martín, and Van Bellegem (2014) and Bates, Castellano, Rabe-Hesketh, and Skrondal (2014). Note that in Assumption (A.2) the variance-covariance matrix of the random effects are allowed to be school-specific, that is, they are different for every j. This means, in particular, that the intensity of the correlation between the outcome specific effect may vary from a school to another. Assumption (A.3) on homoscedasticity can be easily relaxed to take into account a varying variance according to group of schools. We assume constant variance in order to simplify the notation. In practice, the variation allowed in the matrices Λ_j between schools is already very rich to detect distinct variance structures across schools. It should be remarked that assumptions (A.2) and (A.3) rest on probability distributions that are fully known up to their mean and variance-covariance matrix. Finally the assumption (A.4) on local independence has been discussed above.

2.3 Parameter Identification

The identification of the parameters is a necessary prerequisite for estimation and for an interpretable empirical analysis using the model. Parameters under interest are the MK coefficients in vector β, the variance covariance matrices of the outcome-effects within schools $\Lambda_1, \ldots \Lambda_J$, and the variance σ^2. To check the identification of those parameters, we need to show that they can all be expressed as a function of characteristics of the distribution of the observable variables, that is the statistical model bearing on the observable variables only. The statistical model is, consequently, derived after integrating out the unobserved multidimensional school effect γ_j. The following Lemma provides the joint distribution of $(Y_j^\top, \gamma_j^\top)^\top$, from which the statistical model is easily derived; for a proof, see Technical Appendix B.3.

Lemma 2.1 *The structural model specified by (2.2) and the structural assumptions (A.1), (A.2), (A.3) and (A.4) is such that, conditionally on the explanatory variables Z_j, the vector $(Y_j^\top, \gamma_j^\top)^\top$ is distributed according to a $Mn_j + M$-multidimensional distribution such that*

$$\left(\begin{array}{c} Y_j \\ \gamma_j \end{array} \Bigg| Z_j \right) \sim ID \left[\left(\begin{array}{c} Z_j\beta \\ 0 \end{array} \right); \left(\begin{array}{cc} \Lambda_j \otimes (\iota_{n_j}\iota_{n_j}^\top) + \sigma^2 I_{Mn_j} & \Lambda_j \otimes \iota_{n_j} \\ \Lambda_j^\top \otimes \iota_{n_j}^\top & \Lambda_j \end{array} \right) \right],$$

where $A \otimes B$ denotes the Kronecker product between two any matrices A and B, and ι_{n_j} is the unit vector, $(1, 1, \ldots, 1)^\top$, of length n_j.

From Lemma 2.1, we can easily deduce that, for each school, the distribution generating Y_j given the exogenous variable Z_j is a Mn_j-multivariate distribution of mean $Z_j\beta$ and variance-covariance matrix $\Lambda_j \otimes (\iota_{n_j}\iota_{n_j}^\top) + \sigma^2 I_{Mn_j}$. It follows that β is identified if the rank of Z_j is complete. Furthermore, the variance $V(Y_j \mid Z_j)$ has the following typical elements:

1. Diagonal elements, same school and outcome, and same student:

$$Var(Y_{imj} \mid Z_j) = Var(\gamma_{mj} + u_{imj} \mid Z_j) = \Lambda_{j;mm} + \sigma^2$$

2. Off-diagonal elements, same school and outcome, different students:

$$Cov(Y_{imj}, Y_{i'mj} \mid Z_j) = Cov(\gamma_{mj}, \gamma_{mj} \mid Z_j) = \Lambda_{j;mm}$$

3. Off-diagonal elements, same school and student, different outcome:

$$Cov(Y_{imj}, Y_{im'j} \mid Z_j) = Cov(\gamma_{mj}, \gamma_{m'j} \mid Z_j) = \Lambda_{j;mm'}$$

The two last set of moments show that Λ_j are identified from the intraschool covariances of the scores. This, together with the first set or moments, imply that the idiosyncratic variance σ^2 is identified as well.

3 Value Added Analysis

3.1 Outcome-specific value added index

The aim of value added indicators is to provide by institution a measure of student achievement growth acquired as a result of the school policy or practice. Its definition varies according to the nature of the school characteristic we want to highlight in the indicator (e.g. school practice, students group composition policy, student selection, etc.). It is a relative and data-driven indicator, meaning that the sample average of indicators is zero and each of them compares the level of the school within a fixed group of institutions under study (which can be sometimes very large).

The aim of the multidimensional model is to define an M-dimensional school value-added vector, where each coordinate represents the school value-added of a specific outcome. A model-free definition of value added is defined as a difference between conditional expectations in Manzi et al. (2014) using a uni-dimentional value added model. It extends straightforwardly to the multidimensional model, with the only difference that it yields in this case the following M-dimensional vector for each school j instead of a scalar:

$$VA_j = \frac{1}{n_j} \sum_{i=1}^{n_j} \left[E(Y_{ij} \mid Z_j, \gamma_j) - E(Y_{ij} \mid Z_j) \right].$$

Given the assumptions (A.1) to (A.4), the above expression reduces to the random vector γ_j. This random effect is predicted by using the empirical Bayes predictor in which the unknown parameters are replaced by their estimators. Using the following identity, which is valid under the linearity assumption of the conditional expectations $E(Y \mid Z, \gamma)$ and $E(\gamma)$ (for a proof, see Florens, Marimoutou, & Péguin-Feissolle, 2007),

$$E(\gamma_j \mid Y_j, Z_j) = E(\gamma_j \mid Z_j) + Cov(\gamma_j, Y_j \mid Z_j) \left[Var(Y_j \mid Z_j) \right]^{-1} (Y_j - E(Y_j \mid Z_j))$$

we derive an expression to predict the school-outcome specific value-added, as follows:

$$\tilde{\gamma}_j \equiv E(\gamma_j | Y_j, Z_j) = \Lambda_j^\top \otimes \iota_{n_j}^\top \left[\Lambda_j \otimes (\iota_{n_j} \iota_{n_j}^\top) + \sigma^2 I_{Mn_j} \right]^{-1} (Y_j - Z_j \beta) \tag{3.1}$$

where $\tilde{\gamma}_j$ is a vector of dimension M, that contains all outcome-specific value added estimates for school j.

As specified in the model, the school-outcome specific value added indicators are correlated between them. It is worth noting that the estimation of the outcome-specific value added indicator, γ_{jm}, not only depends on information related to the m-th outcome itself, but also to the other outcomes. This is due to the term $\Lambda_j \otimes (\iota_{n_j} \iota_{n_j}^\top)$ that contains all the elements of the covariance matrix Λ_j. To see this point more explicitly, consider the following simplified case that without loosing generality illustrates the main contribution of the multidimentional model in value added models. Suppose that students are tested on two different outcomes so that $M = 2$. In this case, equation (3.1) reduces to

$$\begin{pmatrix} \tilde{\gamma}_{1j} \\ \tilde{\gamma}_{2j} \end{pmatrix} = \begin{pmatrix} a(Y_{1j} - Z_{1j}\beta_1) + b(Y_{2j} - Z_{2j}\beta_2) \\ c(Y_{1j} - Z_{1j}\beta_1) + d(Y_{2j} - Z_{2j}\beta_2) \end{pmatrix},$$

where

$$
\begin{aligned}
a &= \tfrac{1}{(n_j\Lambda_{j;11}+\sigma^2)}\left[1+\tfrac{n_j^2(\Lambda_{j;12})^2}{\omega_{12}}\right]\Lambda_{j;11}-\tfrac{n_j(\Lambda_{j;12})^2}{\omega_{12}};\\
b &= -\tfrac{n_j\Lambda_{j;12}\Lambda_{j;11}}{\omega_{12}}+\tfrac{(n_j\Lambda_{j;22}+\sigma^2)\Lambda_{j;12}}{\omega_{12}};\\
c &= \tfrac{1}{(n_j\Lambda_{j;11}+\sigma^2)}\left[1+\tfrac{n_j^2(\Lambda_{j;12})^2}{\omega_{12}}\right]\Lambda_{j;12}-\tfrac{n_j\Lambda_{j;12}\Lambda_{j;22}}{\omega_{12}};\\
d &= -\tfrac{n_j(\Lambda_{j;12})^2}{\omega_{12}}+\tfrac{(n_j\Lambda_{j;22}+\sigma^2)\Lambda_{j;22}}{\omega_{12}},
\end{aligned}
$$

with $\omega_{12}=(n_j\Lambda_{j;11}+\sigma^2)(n_j\Lambda_{j;22}+\sigma^2)-n_j^2(\Lambda_{j;12})^2$. It is obvious from the above formulas that all the elements of the covariance matrix, Λ_j, are used in the calculation of each outcome-specific effect.

Now consider the case when the outcome-specific random effects are mutually independent, that is the covariace matrix Λ_j is a diagonal matrix with zero covariance elements. In that case,

$$
\tilde{\gamma}_j=\begin{pmatrix} \frac{\Lambda_{j;11}}{n_j\Lambda_{j;11}+\sigma^2}(Y_{1j}-Z_{1j}\beta_1)\\ \vdots \\ \frac{\Lambda_{j;MM}}{n_j\Lambda_{j;MM}+\sigma^2}(Y_{Mj}-Z_{Mj}\beta_M)\end{pmatrix}.
$$

and each component of $\tilde{\gamma}_j$ is equivalently obtained by performing a separate value added analysis for each outcome separately. This also shows how different the estimation of the outcome-specific value added is if the correlation between the outcomes is ignored by performing a separate and independent value added analysis.

3.2 A composite value added index

The outcome-specific value added provides a multidimensional index for each school. However, an all-information indicator of the school value added may sometimes be indispensable for policymakers. For this purpose we build an index that combines all the multiple value added indices into a single one. For each school, we first normalize each M-dimensional vector of VA indices by the inverted square root of the covariance matrix and then average them. By normalizing the indices we ensure that the averaged quantities are expressed in the same scale. It also avoids the over-representation of the individual VA indicators that are highly correlated. Recall that assumption (A.2) specifies that $Var(\gamma_j|Z_j)=\Lambda_j$ for every school j. Accordingly, we propose the following definition for the composite value added indicator,

$$\theta_j = \frac{1}{M}\gamma_j^\top \Lambda_j^{-1/2}\iota_M. \tag{3.2}$$

In the case when the outcome-specific random effects are independent and the covariance matrix Λ_j is a diagonal matrix, then we can write the composite value added index as follows,

$$\theta_j = \frac{1}{M}\sum_{m=1}^{M}\frac{\gamma_{mj}}{\sqrt{\Lambda_{j,mm}}}. \tag{3.3}$$

The independence assumption of the school random effects between the different outcomes is however a very strong assumption that in reality is difficult to guarantee in any educational system. Therefore, equation (3.2) represents a synthesis of the school value added information that best represents the multi-dimensional contribution of the school in the human capital formation of their students.

4 Estimation Procedure

4.1 Estimation

In order to compute the vector of school-outcome specific value added indicators, $\tilde{\gamma}_j$ in equation (3.1), we need to estimate all the unknown parameters using the data. These are the vector of regression coefficients β, covariance matrix elements of Λ_j for each school j and σ^2, the homoscedastic conditional variance of the idiosyncratic error term in equation (2.2). In this paper we consider a method of moments estimation procedure. In this section we summarize the estimation procedure using a sample of observations (Y, Z). From model (2.2) we can write: $Y = Z\beta + H^\top\gamma + u$.

An estimator of the σ^2 follows from the analysis of the residuals in a within regression, as we define below. We define a constant matrix of dimension n_j such that $\overline{J}_{n_j} = \iota_{n_j}\iota_{n_j}^\top/n_j$. Then, we define the within operator for school j as $W_j = I_{Mn_j} - I_M \otimes \overline{J}_{n_j}$. It follows that the within operator for the panel of schools is $W = diag(W_1, W_2, \ldots, W_J)$. Then the within regression is $WY = WZ\beta + Wu$ since by construction $WH^\top = 0$. The within regression estimator is $\hat{\beta}^w = (Z^\top WZ)^{-1}Z^\top WY$. The residuals, $\hat{e}^w = WY - WZ\hat{\beta}^w$, are such that $E\left[(\hat{e}^w)^\top(\hat{e}^w)\right] = \sigma^2 M(N - J - K^*)$, where $K^* \leq K$ is the number of non-zero covariates in the within regression. Therefore, $\hat{\sigma}^2 = (Y - Z\hat{\beta}^w)^\top W(Y - Z\hat{\beta}^w)/M(N - J - K^*)$.

The estimation of Λ_j for each school is more elaborate. We use the least squares residuals $\hat{e} = Y - Z\hat{\beta}$, where $\hat{\beta} = (Z^\top Z)^{-1}Z^\top Y$. First, we denote by $\hat{e}^j_m$ the sub-vector of $\hat{e}$ for school j in module m. Next, we denote by P^j_m a matrix such that $\hat{e}^j_m = P^j_m \hat{e}$. It can be shown (see the Technical Appendix for details) that

$$E\left[(\hat{e}^j_m)^\top(\hat{e}^j_m)\right] = \sigma^2 tr(P^j_{m'} Q {P^j_m}^\top) + \Lambda_{1;mm'}\alpha^{j,1}_{mm'} + \Lambda_{2;mm'}\alpha^{j,2}_{mm'} + \cdots + \Lambda_{J;mm'}\alpha^{j,J}_{mm'} \quad (4.1)$$

where $Q = I - Z(Z^\top Z)^{-1}Z^\top$, and $\alpha^{j,l}_{mm'}$ for $l = 1, \ldots, J$ are known quantities depending only on the covariates matrix Z. The estimation of Λ_j's are found as the solution of the linear system (4.1), after the expectation is replaced by the empirical mean and σ^2 is replaced by its estimate $\hat{\sigma}^2$ from above. Hence,

$$\frac{1}{n_j}(\hat{e}^j_m)^\top(\hat{e}^j_m) = \hat{\sigma}^2 tr(P^j_{m'} Q {P^j_m}^\top) + \Lambda_{1;mm'}\alpha^{j,1}_{mm'} + \Lambda_{2;mm'}\alpha^{j,2}_{mm'} + \cdots + \Lambda_{J;mm'}\alpha^{j,J}_{mm'}$$

for all $m \leq m'$, and for $j = 1, \ldots, J$. Finally, we estimate the vector β by the Generalized Least Squares estimator, i.e. $\hat{\beta} = (Z^\top \widehat{\Omega}^{-1} Z)^{-1} Z^\top \widehat{\Omega}^{-1} Y$, where $\widehat{\Omega}$ is the estimator of $\Omega = Var(H^\top \gamma + u \mid Z) = H^\top \Lambda H + \sigma^2 I_{MN}$ where Λ and σ^2 are replaced by their estimates. In the Technical Appendix we describe the structure of the Ω matrix and also explain how we invert it with reduced computational time.

An alternative method of estimation under the Normal assumption is the maximum likelihood estimation. Analyzing the efficiency of the estimators is beyond the scope of this paper and it is left for future research. Nevertheless, as a first check in the next section we provide the Monte Carlo simulation for given and known model.

4.2 Monte Carlo simulation study

In this section we discuss a Monte Carlo simulation that we conduct. First, we describe the data generating process that we use in this exercise in order to mimic as closely as possible the structure of the data and the context. This exercise simulates a synthetic data set with three outcomes indexed by $m = 1, 2, 3$, so $M = 3$; and for $J = 20$ schools. Each school is assigned a fixed, equal number of students. Note that this restriction is for simplicity purposes and does not affect the performance of the estimator. We do not impose the assumption of equal number of students per school in the VAM and the estimation procedure we introduce in this paper.

These are the parameters values that we fix:

1. We generate the school random effects for each outcome m such as γ_1 , γ_2 and γ_3 that are three random variables that have a joint normal distribution with a fixed correlation matrix of $\begin{pmatrix} 1 & 0.5 & 0.5 \\ 0.5 & 1 & 0.5 \\ 0.5 & 0.5 & 1 \end{pmatrix}$ and fixed equal variances of $\tau^2 = 0.25$. As a result, the covariance matrix of the school random effects is the following $\begin{pmatrix} 0.25 & 0.15 & 0.15 \\ 0.15 & 0.25 & 0.15 \\ 0.15 & 0.15 & 0.25 \end{pmatrix}$
2. We generate three covariates, that play the role of the pre-attainment scores in this case, as standard normal variables: X_1, X_2, X_3, such that each $X_m \sim \mathcal{N}(0,1)$.
3. For each $m = 1,2,3$, we set $\beta_m = \begin{pmatrix} \beta_{0m} \\ \beta_{1m} \\ \beta_{2m} \\ \beta_{3m} \end{pmatrix} = \begin{pmatrix} 5 \\ 0.1 \\ 0.2 \\ 0.3 \end{pmatrix}$
4. Finally, we generate the outcome variables for each $m = 1,2,3$ as a function of the covariates, school random effects and an error term such that $u_{mj} \sim \mathcal{N}(0, 0.04)$, so $\sigma^2 = 0.04$. Then each outcome variable is generated as follows:

$$Y_{mj} = \beta_{0m}\iota_{n_j} + \beta_{1m}X_1 + \beta_{2m}X_2 + \beta_{3m}X_3 + \gamma_{mj}\iota_{n_j} + u_{mj}.$$

The results from the simulation exercise are shown in Table 1, where we report the mean squared error (MSE) for each coefficient. We investigate how this statistic changes as we change the number of the students per school, and as we change the correlation coefficient between the three outcomes, Y_{mj}'s, which comes as a result of the correlation between the school random effects, γ_{mj}'s.

Between Case 1, 2 and 3, we increase the number of students per school firstly from 30 to 50, and than to 100, respectively. Note that moving between these three cases the MSE for the estimation of the σ^2 improves notably as the number of students per school increases. Between Case 1, 4 and 5, we increase the correlation coefficient between the school random effects firstly from 0.1 to 0.5, and then to 0.9. Comparing these three cases we notice that there no change in the estimation of the σ^2 (given that we keep the number of students per school equal), but there appears to be improvement in the estimation of the covariance matrix of the random effect (i.e. in the $\Lambda_{mm'}$'s) and the β's.

We also provide a display of the Monte Carlo simulation results through a boxplot distribution of the coefficient values for each of the cases discussed in Table 1. We present these results in the Appendix in Figures A.1, A.2, A.3, A.4, and A.5. In each figure the top panel displays the results from the separate unidimensional procedure, and the bottom panel show results from the joint analysis. This is in particular useful when comparing the MSE of the $\Lambda_{mm'}$'s between Cases 4 and 5, where the only change is the higher correlation coefficient between the school-outcome effects. The higher this correlation, the lower the MSE for the items in the covariance matrix of the school-outcome effects for the case of the joint multidimensional analysis.

Overall, what we take away from the results of this simulation study are aligned with our expectations. We observe that the higher the number of observations per school, the lower the MSE will be for the σ^2 and the β coefficients under both estimation procedures (unidimensional and joint/multidimensional). The MSE for the variances of the school-outcome effects (Λ_{mm}'s) is only moderately affected by a higher number of students per school. Using the multidimensional/joint estimation procedure, the MSE of $\Lambda_{mm'}$'s improves significantly once we increase the correlation coefficient of γ_{mj}'s.

Table 1: Simulation results

	Case 1		Case 2		Case 3		Case 4		Case 5	
	σ^2= 0.04		σ^2= 0.04		σ^2= 0.04		σ^2= 0.04		σ^2= 0.04	
	n_j= 30		n_j= 50		n_j= 100		n_j= 30		n_j= 30	
	ρ= 0.1		ρ= 0.1		ρ= 0.1		ρ= 0.5		ρ= 0.9	
	M=1	M=3	M=1	M=3	M=1	M=3	M=1	M=3	M=1	M=3
MSE(β_{01})	0.01282	2.290107	0.012648	0.198061	0.012637	12.67667	0.01282	0.099638	0.01282	1.037357
MSE(β_{11})	6.7e-05	0.000135	3.8e-05	4.7e-05	2.1e-05	2.2e-05	6.7e-05	0.00014	6.7e-05	7.4e-05
MSE(β_{21})	6.7e-05	8.1e-05	3.6e-05	4.2e-05	2e-05	2.1e-05	6.7e-05	0.000258	6.7e-05	9.1e-05
MSE(β_{31})	7.1e-05	1e-04	4.4e-05	4.5e-05	2.3e-05	2.1e-05	7.1e-05	8.5e-05	7.1e-05	0.000115
MSE(β_{02})	0.012842	0.16467	0.012667	0.263991	0.012734	32.06873	0.013474	0.399212	0.013509	0.621364
MSE(β_{12})	6.7e-05	0.000132	4.6e-05	4.8e-05	2e-05	0.000139	6.7e-05	0.000107	6.7e-05	9.1e-05
MSE(β_{22})	6.9e-05	8.5e-05	3.7e-05	0.000173	2e-05	5.3e-05	6.9e-05	0.000124	6.9e-05	0.00011
MSE(β_{32})	7.1e-05	0.000107	4e-05	4.3e-05	2.1e-05	0.000218	7.1e-05	0.00022	7.1e-05	0.000124
MSE(β_{03})	0.013592	0.056882	0.013609	0.088915	0.01364	56.36187	0.012974	0.076772	0.012486	1.280793
MSE(β_{13})	6.7e-05	7.7e-05	4.3e-05	7e-05	2e-05	0.000176	6.7e-05	7.8e-05	6.7e-05	7.5e-05
MSE(β_{23})	7e-05	8.5e-05	3.9e-05	7.6e-05	1.9e-05	7.7e-05	7e-05	0.000123	7e-05	0.000134
MSE(β_{33})	6.8e-05	0.000124	4.3e-05	4.4e-05	1.9e-05	4.9e-05	6.7e-05	8.7e-05	6.7e-05	8e-05
MSE(σ_1^2)	5e-06	2e-06	3e-06	1e-06	2e-06	1e-06	5e-06	2e-06	5e-06	2e-06
MSE(σ_2^2)	5e-06		3e-06		2e-06		5e-06		5e-06	
MSE(σ_3^2)	6e-06		3e-06		2e-06		6e-06		6e-06	
MSE(Λ_{11})	0.009372	0.008397	0.009339	0.008321	0.009628	0.008342	0.009372	0.008397	0.009372	0.008397
MSE(Λ_{22})	0.008086	0.054713	0.008121	0.054832	0.008058	0.054773	0.007805	0.02039	0.0087	0.00802
MSE(Λ_{33})	0.007164	0.055502	0.006499	0.055419	0.007423	0.05549	0.007074	0.020883	0.008541	0.008135
MSE(Λ_{12})		0.056521		0.056376		0.056334		0.021598		0.008358
MSE(Λ_{13})		0.003068		0.003077		0.003056		0.003774		0.007021
MSE(Λ_{23})		0.054653		0.054675		0.054495		0.020352		0.008213

5 Tertiary Education in Colombia

In the 1990s Colombia experienced a dramatic expansion of the tertiary education[1] following the introduction of the new Constitution and reforms in the education system that came with it. The expansion resulted mainly in the widening of the private sector and abundance of the non-university institutions offering professional and vocational training. This surge in the supply of tertiary education institutions has made instruction quality a major issue for Colombia. Consequently, the Colombian government introduced a national evaluation system whose objectives are described in the 2009 decree[2] of the Ministry of Education. The objectives consisted in assessing the level of development of competencies in students who are finishing their undergraduate education, producing value added indicators of the higher education programs, and providing information to compare not only higher education programs and institutions, but also teaching methodologies. In order to reach these objectives, the state collected test scores of the students at two specific points in their educational path, upon high school graduation and upon completion of their undergraduate studies through two national state exams, namely Saber 11 and Saber Pro .

The Colombian youth leaves compulsory upper secondary education at age 16 (equivalent to US high school). Those that aspire to continue post-secondary education have to take a national state exam at the last grade (grade 11). This test is formally known as Saber 11 and includes evaluations in core subjects, such as Spanish, mathematics, biology, chemistry, physics, philosophy, social sciences, and foreign languages (English). The score of this test has no effect in the graduation decision from the upper secondary education but it is the official national admission test into tertiary education since 1980, and thus compulsory for the students with tertiary education aspirations. Most of the institutions in Colombia (78%) use Saber 11 results for admission and each institution decides on the minimum acceptable Saber 11 score which may change each year depending on the demand. In order to be eligible for entry in tertiary education the Colombian students need both, the Saber 11 score and the upper secondary education diploma.

Upon completion of their undergraduate degree requirements, the students are submitted to another mandatory battery of tests, officially known as Saber Pro (former

[1]Tertiary education in this context will refer to all types of post-secondary education that will include university and non-university education, vocational and technical education that may not be granting degrees.

[2]Ministerio de Educación Nacional, Decreto 3963, 14-10-2009.

ECAES). Saber Pro is an exit examination from tertiary education which became compulsory for graduation as of 2009. It is designed to measure the academic capital that the students have built in the institution in various domains of competencies and proficiency. The exit exam, Saber Pro , is for each student a vector of five tests corresponding to the following modules or subjects: written communication, English, quantitative reasoning, critical reading and citizenship competencies.

There are four types of tertiary education institutions in Colombia. They are classified as follows: (1) Universities, which offer academic undergraduate programs and graduate programmes (master's or doctoral) with a focus on scientific research ; (2) University Institutions, which offer undergraduate degrees up to professional degree level and a type of graduate programme known as "specialisation" that is above a bachelor's degree but below a master's degree; (3) Technological institutions, which offer programmes up to a technologist level and is distinguishable from professional technical level by their scientific basis; (4) Professional technical institutions, which offer professional/technical level training for a particular job or career. Both technological institutions and professional technical institutions provide short term training with relatively lower tuition that respond to local labour market demand with flexibility to changes. Some of these institutions are private and some are public. The public ones are funded by the government and are autonomous in the way they allocate their funds. All tertiary institutions (except SENA (Servicio Nacional de Aprendizaje) centers) charge tuition to students, which vary a lot by institution. The private institutions rely on student tuition for most of their income. By law, all tertiary institutions are required to maintain a non-for-profit status. For more details, see OECD/International Bank for Reconstruction and Development/The World Bank (2012).

Given the diverse nature of the post-secondary education institutions described above, in this paper we focus only on Universities and University Institutions (i.e. type (1) and (2) above), and drop the other post-secondary education institutions from the sample.

6 Data

Our data are administrative records of the students that wrote the Saber 11 exam, who then enrolled into a tertiary education program and upon completion also wrote the

Saber Pro test. The data were generously provided by ICFES.[3] It is the two instances of these examinations that are in the center of our value added analysis. The prior attainment score considered in this study is a vector of six scores of the Saber 11 tests. Each student is tested on the following six subjects: Mathematics, Physics, Chemistry, English, Social Sciences, Spanish. The final attainment score is for each student a vector of five test scores from the Saber Pro test corresponding to the following competencies: Quantitative Reasoning, Critical Reading, English, Citizenship Competencies and Written Communication.

SABER 11 test is compulsory and the students' performance in this test is the primary admission criteria for entry in a higher education institution. This feature warranties that the students take the test seriously and put effort in producing their answers. We may say the same for the Saber Pro tests written since 2009, which is the year that it became compulsory in order to graduate, even though its outcome does not affect whether the student qualifies for graduation.

Initially, the Saber database contains 187,698 observations. We use the data belonging to the students who took the Saber Pro test in 2012 and 2013 (those who took Saber Pro in 2011 are excluded from the analysis). The sample is restricted further to the students who wrote Saber 11 four to six years prior to Saber Pro , i.e. the first group have written Saber 11 between 2006–2008 and the second group between 2007–2009. This is a reasonable sample modification since the analysis we perform requires coherence in the knowledge investment between the two tests. Hence, we will be working with the 2012 and 2013 tertiary education (potential) graduates. The test scores are equated, so we do not need to worry about test differences across years as a result of difficulty level. After discarding the above items, our working sample contains 139,205 observations. Apart from the above sample restrictions, we also have 1,946 missing observations in total in the test scores data. The test on Written Communication has the majority of the missing entries due to non-response. So, after excluding the missing observations the final database contains 137,278 observations.

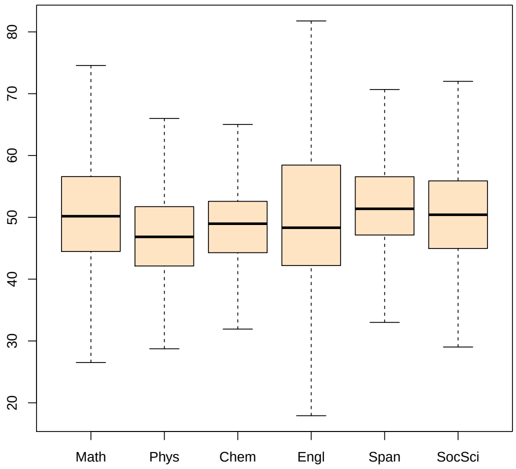
ICFES aggregated the detailed program classification into 18 broader groups, which we refer to as fields of study. Hence, each student who registered to a program is automatically classified into one field of study. The fields of study are listed in Table A.1, that also

[3] ICFES is the abreviation of "Instituto Colombiano para el Fomento de la Educación Superior" (Colombian Institute for the Promotion of Higher Education).

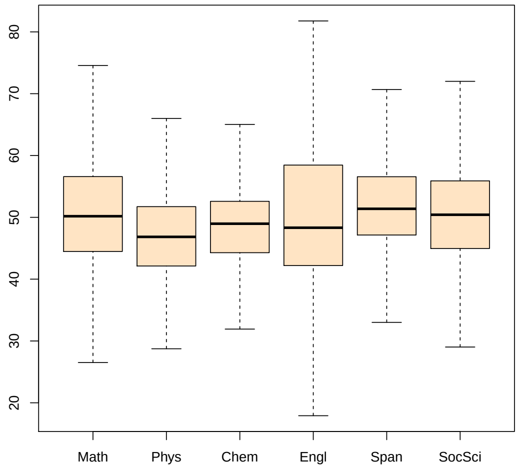
documents the number of students studying in each of them. The last column in Table A.1 shows that a big proportion of students (more than half) in the data graduated from only four fields of study: Engineering, Administration, Education and Law. Engineering attracted the highest proportion of the student body which is almost one-fourth of the student population in the data (23.4%). The fields that attract the least number of students (less than one percent) are Humanities and Military and Naval Sciences (0.7% and 0.3%, respectively). The value added model aims to compare institutions within a single field of study. In favor of space, in what follows we focus on only two major fields of study: Engineering and Law. In Figures 1 and 2 we show the boxplots of the tests scores using the original sample before cleaning the data. We also show adjacent to these plots how the distribution of these variables looks in the final sample that we use for the empirical analysis. Each box indicates the maximum value, third quartile, median, first quartile and the minimum value. It is obvious that the sample composition is left unaffected from the data cleaning.

In addition to the previously introduced variables, the database contains further information on each student, such as the gender, the residential area of the student at the time of the Saber 11 and Saber Pro exams were written, the field of study, the number of semesters in higher education that the student completed till he wrote the exit test, the INSE that is a continuous socioeconomic measure constructed by ICFES[4], the name of the school where the student completed his/her post-secondary education, the type of the institution, a categorical variable indicating the tuition paid, and the type of degree. In Table 2 we show summary statistics for all the variables we use in the empirical analysis, except for the school composition variables. We observe a similar distribution in terms of outcome variables, with no obvious differences in Saber Pro test scores between the Engineering and the Law students. However, note that the Engineering student perform about half a standard deviation better in Saber 11 math score. They also seem to do better in physics, chemistry and English as a secondary language. The INSE index is almost the same for both majors, but the proportion of females (0.623) in Law programs relative to males is much higher than in Engineering programs (0.408). Finally, about 80 percent choose academic track (rather than "Normal" or technical track). The proportion of students in the higher tuition-paying categories (expressed in million Colombian pesos)

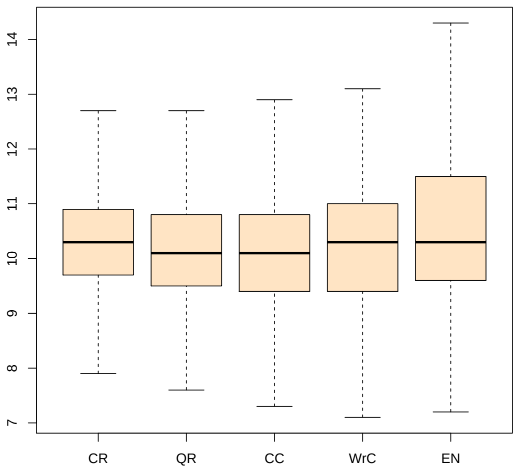
[4]See the internal report *Metodología de construcción del indice de nivel socioeconomico de los estudiantes - INSE - y de la clasificación socioeconómica - CSE - de los colegios*, ICFES, Junio 2010.



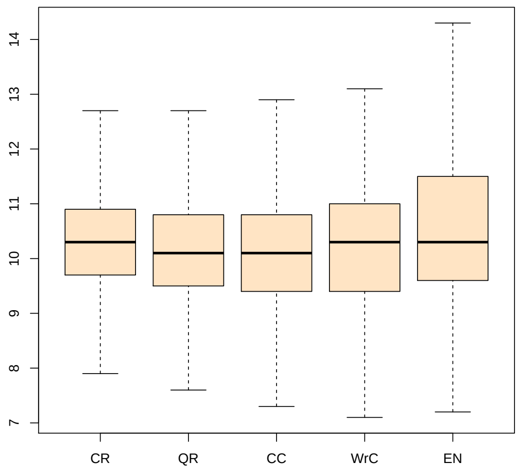
(a) Saber 11 , Raw data , N=32582



(b) Saber 11 , Final sample, N=30857

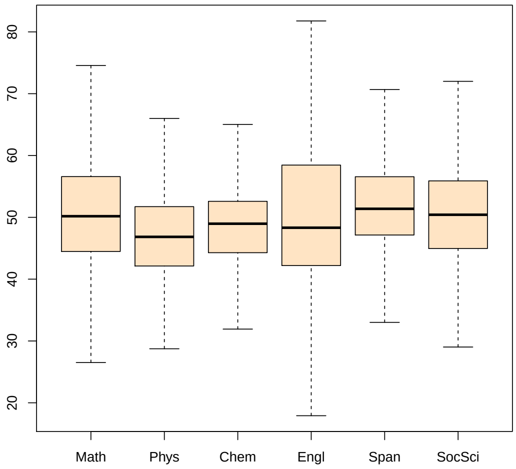


(c) Saber Pro , Raw data, N=32582

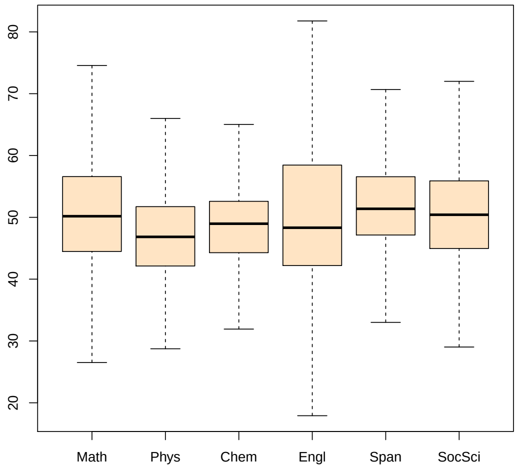


(d) Saber Pro , Final sample , N=30857

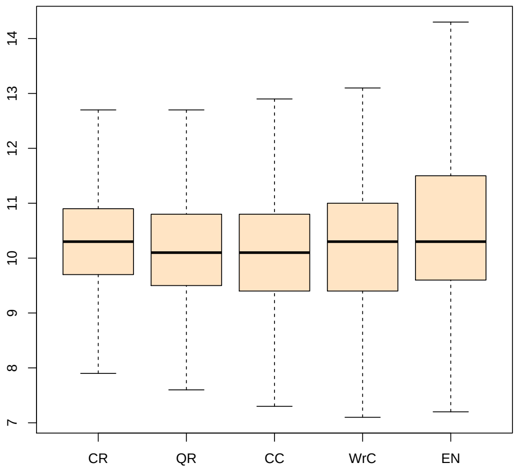
Figure 1: Boxplots of the test scores for the Engineering students



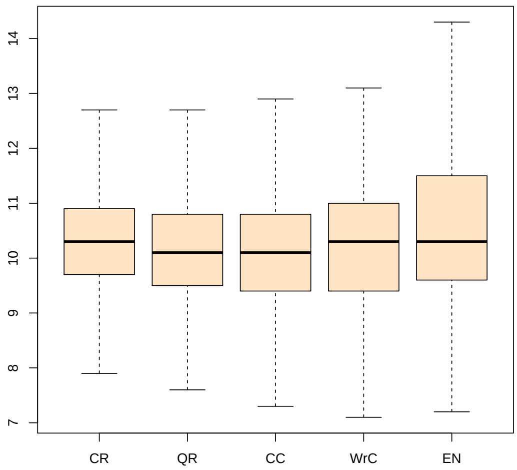
(a) Saber 11 , Raw data , N=11990



(b) Saber 11 , Final sample, N=11198



(c) Saber Pro , Raw data, N=11990



(d) Saber Pro , Final sample , N=11198

Figure 2: Boxplots of the test scores for the Law students

is higher for the Law students than for the Engineering students. Table A.2 provides the correlation coefficients between the pre-attainment scores and the outcome test scores.

7 Results

7.1 Unidimensional versus multidimensional value added analysis

The conventional approach to the estimation of the VA indicators is to estimate equation (2.1) separately for each module m. In terms of the regression estimation details, the vector Y contains all the modules tested in the Saber Pro test. The matrix Z contains the six domains of the Saber 11 test, namely Mathematics, Spanish, Chemistry, Social Sciences, English language and Physics. It also includes a unit vector to capture the intercept term. Other covariates include the INSE and a set of school composition variables that are computed as the average of the Saber 11 test scores of the current peers in the same university. Tables A.3 and A.4 show the regression results for the two fields of study, Engineering and Law, respectively. In these two tables each column is a separate regression.

As also noted earlier in paper, the outcome variables are inherently correlated. We can see that in Table 3, that shows the correlation coefficients for the raw test scores. They vary approximately between 0.205 to 0.631 for both Engineering and Law, and the majority are higher than 0.5.

Moreover, in order to support the claims made elsewhere in the literature about the sensitivity of value added indicators to the outcome variable used, Table 4 displays the rank correlation among the unidimensional value added estimates between the five outcomes. The correlation coefficients are as high as 0.713 (between CR and CC for Law)[5], and as low as 0.270 (between QR and EN for both Engineering and Law). Obviously, the varying correlation is a reflection of sensitivity.

From a policy point of view, the information provided from the unidimensional value added model is to some extent different for each of the five outcome variables. So, one faces the decision of choosing one. The multidimensional model accounts for the correlations shown in Table 3 and aggregates the information provided by the separate unidimensional

[5] For a list of acronyms please see Appendix.

Table 2: Summary statistics by field of study

Variable	Mean	St.Dev.	Min	Max
		Engineering		
Spro WrC	10.185	1.043	6.80	13.10
Spro EN	11.046	1.507	6.00	15.00
Spro CR	10.414	0.954	5.80	14.70
Spro QR	10.891	1.156	5.80	16.30
Spro CC	10.191	1.009	5.70	14.00
S11 Span	52.875	6.832	18.33	99.69
S11 Math	55.512	11.535	14.92	121.49
S11 SocSci	52.033	7.983	12.87	97.35
S11 Chem	51.520	7.070	17.70	93.33
S11 Phys	50.098	8.129	13.60	97.06
S11 Engl	54.754	14.654	12.90	111.94
INSE	54.121	9.482	20.059	73.135
Academic track	0.767		0	1
Female dummy	0.408		0	1
Tuition 1–3 million peso	0.241		0	1
Tuition 3–5 million peso	0.217		0	1
Tuition >5 million peso	0.233		0	1
		Law		
Spro WrC	10.563	1.099	6.80	13.1
Spro EN	10.551	1.417	6.00	15
Spro CR	10.519	0.979	6.20	14.6
Spro QR	9.984	0.918	6.70	16
Spro CC	10.617	0.993	5.70	14
S11 Span	52.351	7.112	18.21	87.44
S11 Math	49.529	9.744	11.30	110.73
S11 SocSci	51.355	8.035	6.74	81.73
S11 Chem	48.161	6.337	21.83	82.04
S11 Phys	46.646	7.322	11.26	81.64
S11 Engl	51.810	14.403	12.90	111.94
INSE	55.346	9.223	20.024	73.135
Academic track	0.803		0	1
Female dummy	0.623		0	1
Tuition 1–3 million peso	0.333		0	1
Tuition 3–5 million peso	0.275		0	1
Tuition >5 million peso	0.276		0	1

Table 3: Correlation coefficients among the raw test scores

Field of Study		QR	CR	EN	CC	WrC
Engineering	QR	1	0.268	0.273	0.205	0.276
	CR		1	0.497	0.546	0.468
	EN			1	0.557	0.624
	CC				1	0.535
	WrC					1
Law	QR	1	0.303	0.334	0.257	0.335
	CR		1	0.48	0.457	0.47
	EN			1	0.495	0.631
	CC				1	0.482
	WrC					1

Table 4: Rank correlation coefficients among the unidimensional VA estimates

			Unidimensional Value Added				
	Field of Study		QR	CR	EN	CC	WrC
Unidimensional Value Added	Engineering	QR	1	0.665	0.27	0.537	0.376
		CR		1	0.363	0.663	0.454
		EN			1	0.368	0.293
		CC				1	0.509
		WrC					1
	Law	QR	1	0.547	0.27	0.64	0.469
		CR		1	0.291	0.713	0.408
		EN			1	0.246	0.205
		CC				1	0.495
		WrC					1

indicators. Tables A.5 and A.6 show the estimation results of equation (2.2), when the outcomes are analyzed jointly in a SURE for the field of study of Engineering and Law, respectively. It can be easily seen that there is a gain in estimation efficiency in the case of SURE regressions relative to the separate estimation procedure. Moreover, we are able to compare, say the effect of the mathematics score on Critical Reading versus Written Communication for the Students of Engineering in Table A.5. As we can see, the coefficient estimate in the former is almost twice as high as in the latter. Instead, by referring to Tables A.3 and A.4 we are not able to make a similar comparison because the inherent correlation between the outcome variables (Saber Pro test scores) is not taken into account in those regression.

Table 5: Rank correlation coefficients among the unidimensional and multidimensional VA estimates

	Field of Study		Multidimensional Value Added QR	CR	EN	CC	WrC	$ave(\hat{\theta}_j)$	$w(\hat{\theta}_j)$
Unidimensional Value Added	Engineering	QR	0.135					0.338	0.299
		CR		0.14				0.422	0.209
		EN			0.092			0.314	0.042
		CC				0.186		0.35	0.151
		WrC					0.183	0.328	0.172
	Law	QR	0.251					0.470	0.313
		CR		0.119				0.474	0.236
		EN			0.167			0.346	0.179
		CC				0.361		0.563	0.384
		WrC					0.269	0.463	0.125

In Table 5 we show the rank correlation coefficients among the value added indicators produced through the unidimensional and the multidimensional models. The low bilateral correlation coefficients indicate that these two sets provide very different ranking of the tertiary institutions. Differently from the separate analysis, the multidimensional model allows us to combine the information into a single ranking index by averaging (equally weighted or weighted by the inverse of the covariance matrix) the individual subject VA estimates. In the last two columns of Table 5 we can see the correlation coefficients of the combined value added indicators and the unidimensional value added indicators for each subject. Again, the correlation coefficients are quite low.

7.2 Within-institution effectiveness comparison

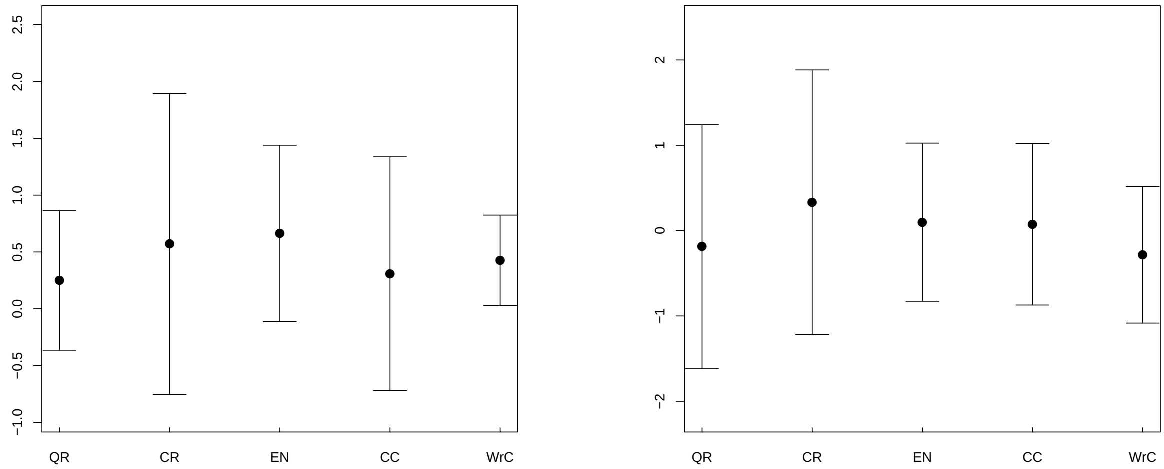
Another advantage of the multidimensional analysis is related to its use for internal accountability purposes for each university. In Figure 3 we plot the VA estimates for each module, together with their confidence bands based on the following formula for the lower and upper band: $\left[\hat{\gamma}_{mj} - (\sqrt{\Lambda^{j}_{mm}})t_{\alpha/2}, \quad \hat{\gamma}_{mj} + (\sqrt{\Lambda^{j}_{mm}})t_{\alpha/2}\right]$. In Table 6 we report the plotted numerical values. Using this information, this particular university, for instance, can infer that on average for the engineering programs, their students performance in Citizenship Competencies (CC) and Written Communications (WrC) is much better than in the other domains of the Saber Pro exam. This is relative to the other institutions in the sample that also offer programs in the field of Engineering. We should keep in mind that the universities may not be the same in each major grouping, so the interpretation is always relative to the universities for each field of study. For the Law majors, this particular university that we have reported has a much lower VA estimate on its students for Quantitative Reasoning (QR) and Written Communication (WrC) than for the other tests.

Table 6: Multidimensional value added estimates and 95% confidence bounds for an unnamed university

Field of Study		QR	CR	EN	CC	WrC
Engineering	Lower bound of γ_{mj}	-0.365	-0.753	-0.113	-0.720	-0.027
	γ_{mj}	0.249	0.570	0.663	0.309	0.426
	Upper bound of γ_{mj}	0.863	1.893	1.439	1.337	0.824
Law	Lower bound of γ_{mj}	-1.613	-1.218	-0.827	-0.872	-1.084
	γ_{mj}	-0.186	0.333	0.099	0.073	-0.284
	Upper bound of γ_{mj}	1.241	1.883	1.026	1.019	0.515

7.3 Between-institutions effectiveness comparison

Finally, the multidimensional VAM allows us to also make a comparative analysis between universities for a single test outcome. For example, consider the VA estimates shown in Table 7 for three anonymous universities for each field of study. We can see that for the test on Quantitative Reasoning university A is performing better than C, which is better than B.



(a) Engineering (b) Law

Figure 3: Multidimensional VA estimates and their 95% confidence bands for a single university

Table 7: Multidimensional value added estimates for three unnamed universities

Field of Study		QR	CR	EN	CC	WrC
Engineering	University A	0.419	1.244	0.261	0.273	0.335
	University B	-0.390	0.383	0.954	-0.197	0.557
	University C	0.217	-0.350	0.148	-0.407	0.171
Law	University A	0.246	-0.058	-0.020	0.219	-0.414
	University B	-0.319	0.999	-0.149	0.260	0.077
	University C	0.017	-0.321	0.125	-0.334	-0.247

8 Conclusion

Teachers and schools contribute to their students' learning process by developing several aspects of their cognitive and social skills. This and other related reasons render the test scores on different subjects inherently and structurally linked. In the value added literature it is common to analyze individually the test scores of different subjects. However, it has been recognized that the unidimensional value added estimates are sensitive to the outcome used and yield different rankings, yet with some degree of dependency that is difficult to model (see Lockwood et al., 2007; Papay, 2011). Our contribution to the literature consists in modeling this dependency and accounting for it in the estimation of the value added indicators. We do so by developing a multidimensional value added model within the framework of SURE models. The key idea is that the test scores on different subjects (or, multiple outcomes) depend on both a set of covariates and a multidimensional school effect, the components of which are correlated among them. Conditional on the covariates, it is this correlation that induces the conditional dependency between the different outcomes. We point out in the paper that the value added indicators produced through the unidimensional and multidimensional value added models are quite different. We show empirically how different the ranking induced by those two value added analysis is, the rank correlation coefficients of which are not higher than 0.37 (see Table 5).

In practice, the multidimensional value added model that we propose in this paper can be exploited in conducting two very informative effectiveness analysis.

1. First, the model allows for a within-school effectiveness comparison. The multidimensional value added model produces value added indicators for each outcomes, which are comparable between them for the same school. A school or institution can exploit this information to evaluate its relative performance in different outcomes and identify areas for improvement.

2. Second, the model also allows for a between-schools effectiveness comparison. This comparison can be made outcome-by-outcome fashion, or in an all-information level. For the later we produce the composite value added indicator by weighting each outcome-specific indicator by the inverted square root of the covariance matrix (see equation (3.2)). In this way, the composite index takes into account the school specific inherent dependency of the multiple outcomes and a more informative between-school comparison can be performed.

The case study we analyze in this paper deals with the quality of higher education. The interest on quality assurance in higher education is expanding worldwide, generating new requirements and international standards for the accreditation of institutions and programs (Hou, 2011). The accreditation process typically focuses its attention on organizational aspects on programs and institutions. However, some authors argue that accreditation should be complemented with internal improvement processes at the students level (see Schwarz & Westerheijden, 2004; Dano & Stensaker, 2007). The value added analysis of the tertiary educational system of Colombia can be considered as an input of an accreditation process which focuses its attention on students' progress. Both the within- and between-institution comparisons provide information relevant to internal and external accountability processes.

References

Bacher-Hicks, A., Kane, T. J., & Staiger, D. O. (2014). Validating teacher effect estimates using changes in teacher assignments in los angeles. *National Bureau of Economic Research, Working Paper No. 20657*.

Ballou, D., Sanders, W., & Wright, P. (2004). Controlling for student background in value-added assessment of teachers. *Journal of Educational and Behavioral Statistics*, *29*, 37–65.

Bates, M. D., Castellano, C. E., Rabe-Hesketh, S., & Skrondal, A. (2014). Handling correlations between covariates and random slopes in multilevel models. *Journal of Educational and Behavioral Statistics*, *39*, 524-549.

Broatch, J., & Lohr, S. (2012). Multidimentional assessment of value added by teachers to real-world outcomes. *Journal of Educational and Behavioral Statistics*, *37*(2), 256-277.

Chetty, R., Friedman, J., & Rockoff, J. (2014). Measuring the impact of teachers i: Evaluating bias in teacher value-added estimates. *American Economic Review*, *104(9)*, 2593-2632.

Dano, T., & Stensaker, B. (2007). Still Balancing Improvement and Accountability? Developments in External Quality Assurance in the Nordic Countries 1996-

2006. *Quality in Higher Education*, *13*, 81–93.

Engle, R. E., Hendry, D. F., & Richard, J. F. (1983). Exogeneity. *Econometrica*, *51*, 277-304.

Figlio, D., & Loeb, S. (2011). School accountability. In E. A. Hanushek, S. Machin, & L. Woessmann (Eds.), (Vol. 3, p. 383-420). Elsevier B.V.

Florens, J. P., Marimoutou, V., & Péguin-Feissolle, A. (2007). *Econometric Modeling and Inference*. Cambridge University Press.

Florens, J.-P., Mouchart, M., & Rolin, J.-M. (1990). *Elements of Bayesian Statistics*. New York: Marcel Dekker.

Hill, P. W., & Rowe, K. J. (1996). Multilevel modelling in school effectiveness research. *School Effectiveness and School Improvement*, *7*, 1–34.

Hou, A. Y. (2011). Quality assurance at a discrete: International accreditation in Taiwan higher education. *Higher Education*, *61*, 179-191.

Kane, T. C., & Staiger, D. O. (2008). Estimating teacher impacts on student achievement: An experimental evaluation. *National Bureau of Economic Research, Working Paper No.14607*.

Kinsler, J. (2012). Assessing Rothstein's critique of teacher value-added models. *Quantitative Economics*, *3*, 333–362.

Koedel, C., & Betts, J. R. (2011). Does student sorting invalidate value-added models of teacher effectiveness? An extended analysis of the Rothstein critique. *Education Finance and Policy*, *6*(1), 18–42.

Koedel, C., Mihaly, K., & Rockoff, J. E. (2015). Value-added modeling: A review. *Economics of Education Review*.

Kupermintz, H. (2003). Teacher effects and teacher effectiveness: A validity investigation of the Tennessee Value Added Assessment System. *Journal of Educational and Behavioral Statistics*, *25*(3), 287–298.

Lazarsfeld, P. F. (1950). The Logical and Mathematical Foundation of Latent Structural Analysis. In S. S. Stout (Ed.), (pp. 362–412). Science Editions.

Lenkeit, J. (2012). Effectiveness measures for cross-sectional studies: a comparison of value-added models and contextualised attainment models. *School Effec-*

tiveness and School Improvement, 1–25.

Lockwood, J. R., McCaffrey, D. F., Hamilton, L. S., Stecher, B., Le, V.-N., & Martinez, J. F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, *44*(1), 47-67.

Manzi, J., San Martín, E., & Van Bellegem, S. (2014). School Effectiveness by Value Added Analysis under Endogeneity. *Psychometrika*, *79*, 130–153.

Martineau, J. A. (2006). Distoring value added: The use of longitudinal, vertically scaled student achievement data for growth-based, value-added accountability. *Journal of Educational and Behavioral Statistics*, *31*, 35–62.

McCaffrey, D. F., Lockwood, J. R., Koretz, D., Louis, T. A., & Hamilton, L. S. (2004). Models for value-added modeling of teacher effects. *Journal of Educational and Behavioral Statistics*, *29*, 67-101.

OECD. (2008). Measuring improvements in learning outcomes. Best practices to assess the value-added of schools.

OECD/International Bank for Reconstruction and Development/The World Bank. (2012). Reviews of National Policies for Education: Tertiary Education in Colombia 2012.

Papay, J. P. (2011). Different tests, different answers: The stability of teacher value-added estimates across outcome measures. *American Educational Research Journal*, *48*(1), 163-193.

Rothstein, J. (2009). Student sorting and bias in value-added estimation: Selection on observables and unobservables. *Education Finance and Policy*, *4*(4), 537–571.

Rothstein, J. (2010). Teacher quality in educational production: Tracking, decay and student achievement. *The Quarterly Journal of Economics*, *125 (1)*, 175-214.

Rothstein, J. (2014). *Revisiting the impacts of teachers.* Retrieved from `http://eml.berkeley.edu/~jrothst/workingpapers/rothstein_cfr.pdf`

Rubin, D. B., Stuart, E. A., & Zanutto, E. L. (2004). A potential outcome view of value added assessment in education. *Journal of Educational and Behavioral Statistics*, *29*(1), 103–116.

Sammons, P., Nuttall, D., & Cuttance, P. (1993). Differential school effectiveness: Results from a reanalysis of the inner london education authority?s junior school project data. *British Educational Research Journal*, *19*, 381–405.

Sanders, W., Saxton, A., & Horn, S. P. (1997). The Tennesse Value-Added Assessment System: A quantitative outcome-based approach to educational assessment. In J. Millman (Ed.), (p. 137-160). Thousand Oaks, CA: Corwin Press, Inc.

Schatz, C., Von Secker, C., & Alban, T. (2005). Balancing accountability and improvement: Introducing value-added models to a large school system. In R. W. Lissitz (Ed.), (pp. 1–18). Maple Grove, MN: JAM Press.

Schwarz, S., & Westerheijden, D. F. (2004). *Accreditation and Evaluation in the European Higher Education Area*. Springer.

Snijders, T. A. B., & Bosker, R. J. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). SAGE Publications.

Zellner, A. (1962). An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. *Journal of the American Statistical Association*, *57*, 348–368.

A Appendix

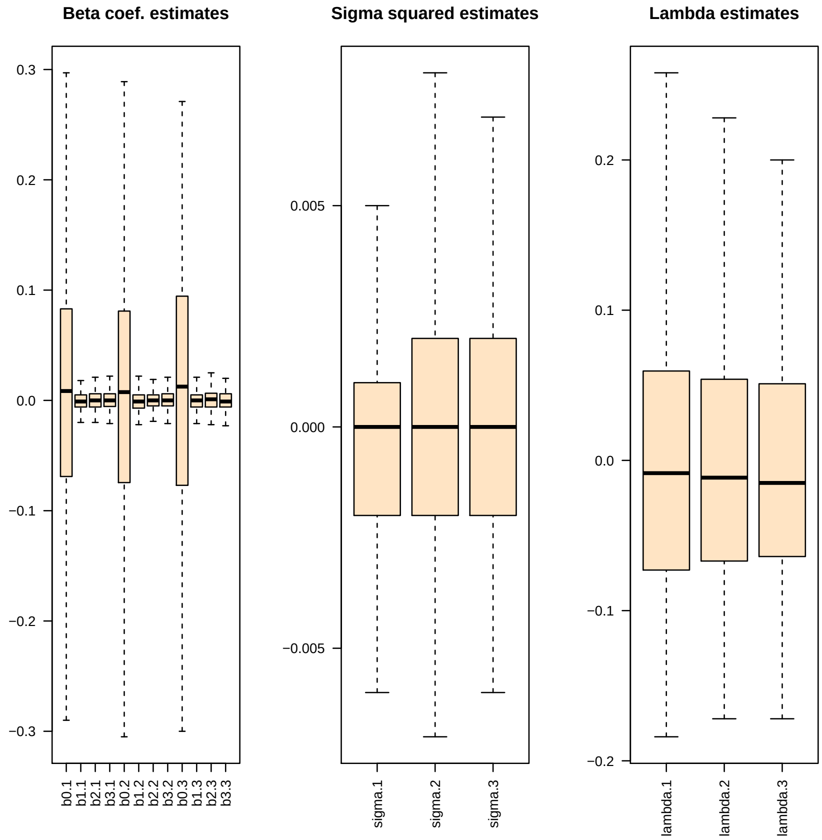
A.1 Acronyms

Variable names

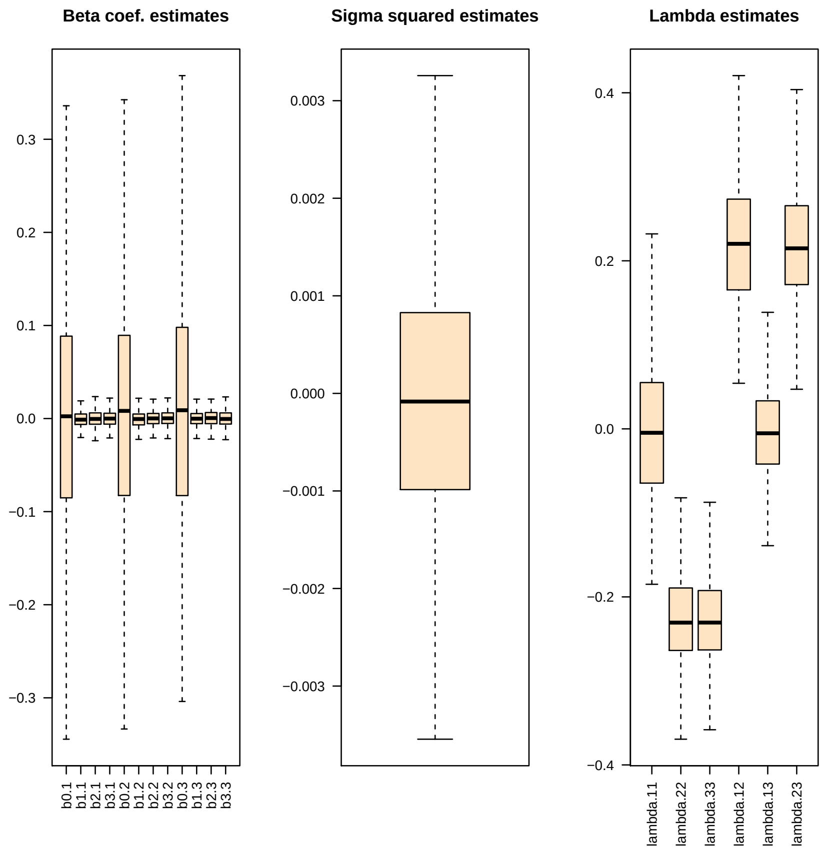
- QR (or Spro QR): SABER PRO test, Quantitative Reasoning score
- CR (or Spro CR): SABER PRO test, Critical Reading score
- EN (or Spro EN): SABER PRO test, English as a foreign language score
- CC (or Spro CC): SABER PRO test, Citizenship Competencies score
- WrC (or Spro WrC): SABER PRO test, Written Communication score
- Engl (or S11 Engl): SABER 11 test, English score
- Math (or S11 Math): SABER 11 test, Mathematics score
- Phys (or S11 Phys): SABER 11 test, Physics score
- Chem (or S11 Chem): SABER 11 test, Chemistry score
- SocSci (or S11 SocSci): SABER 11 test, Social Sciences score
- Span (or S11 Span): SABER 11 test, Spanish score

Table elements

- VA: Value added
- Qu.: Quartile
- NA: Not Applicable
- # : Number of Observations
- stdev: Standard Deviation
- *ave(.)*: equally-weighted average
- *w(.)*: weighted average, where the weight is the inverse of the variance

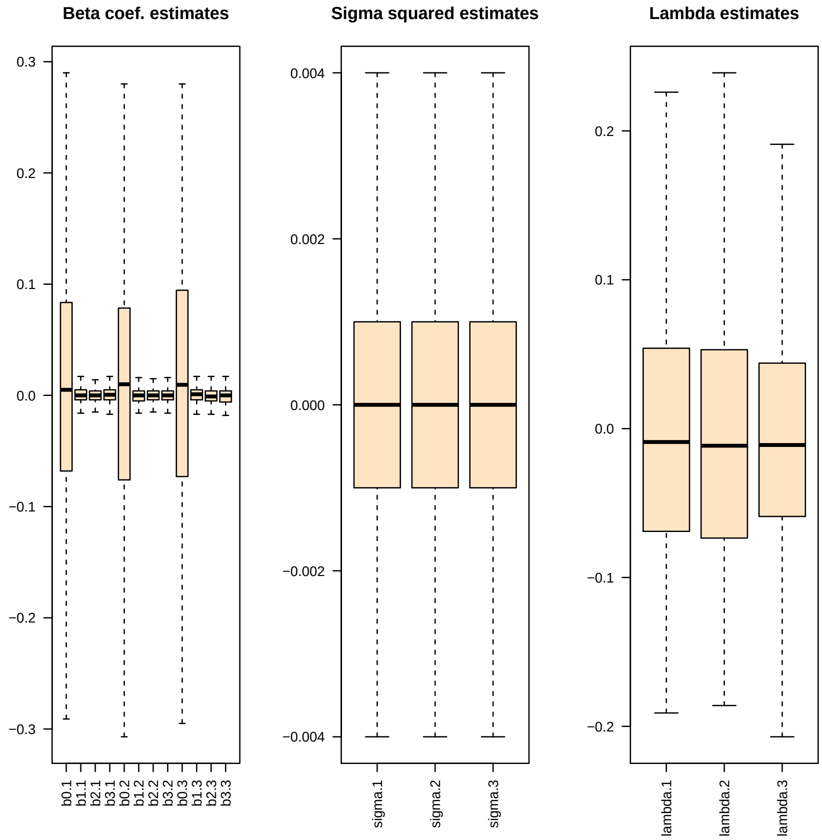


(a) Unidimensional/Separate value added analysis

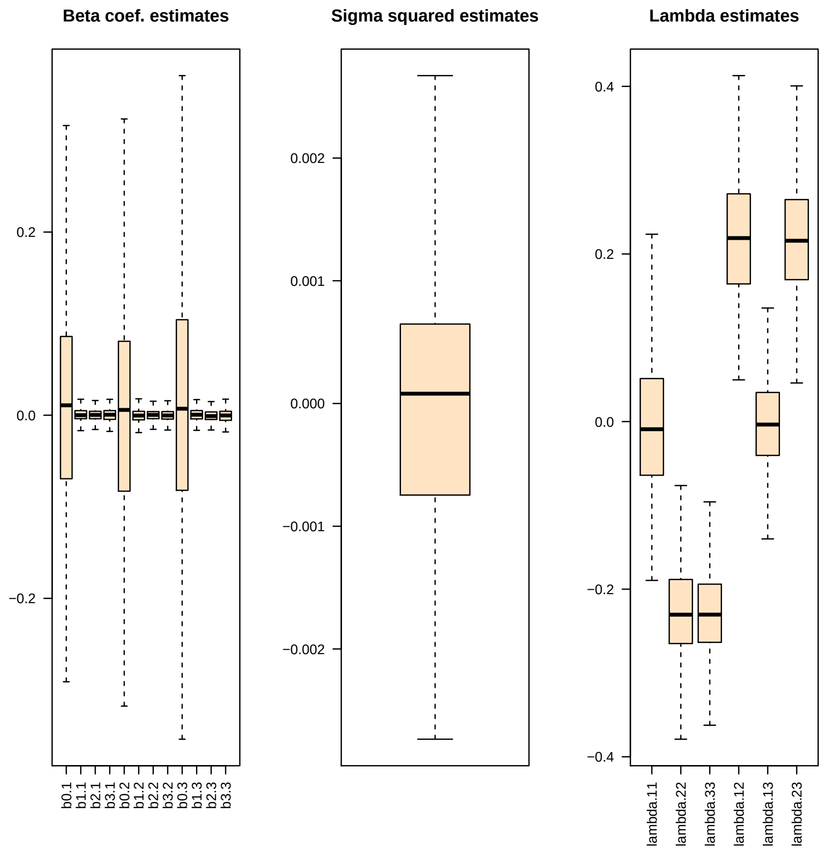


(b) Multidimensional/Joint value added analysis

Figure A.1: Distribution of the coefficient estimates from the simulation Case 1

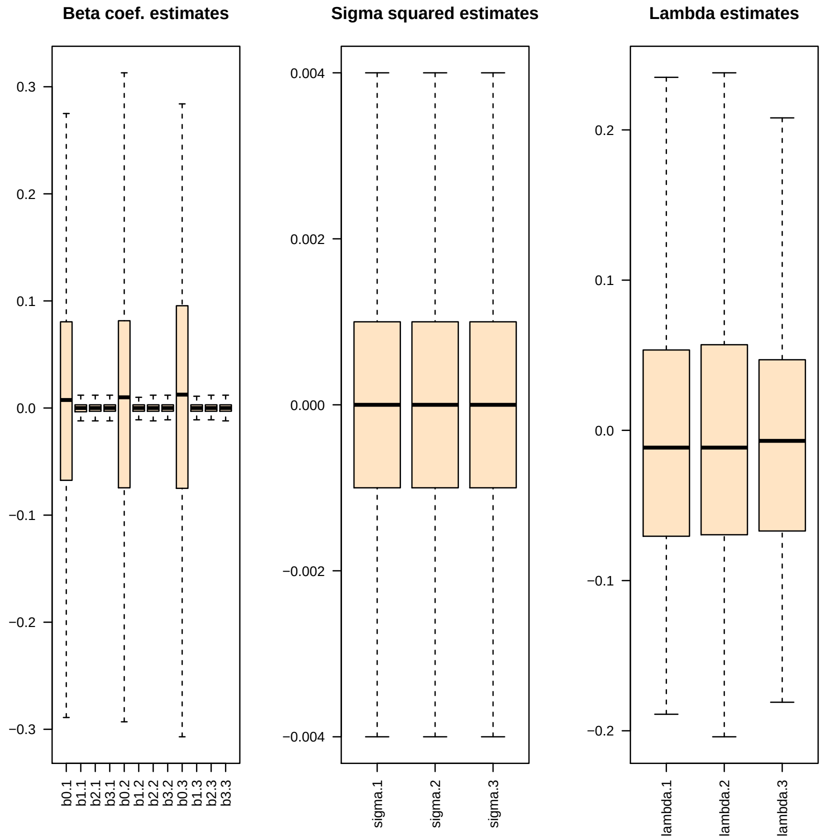


(a) Unidimensional/Separate value added analysis

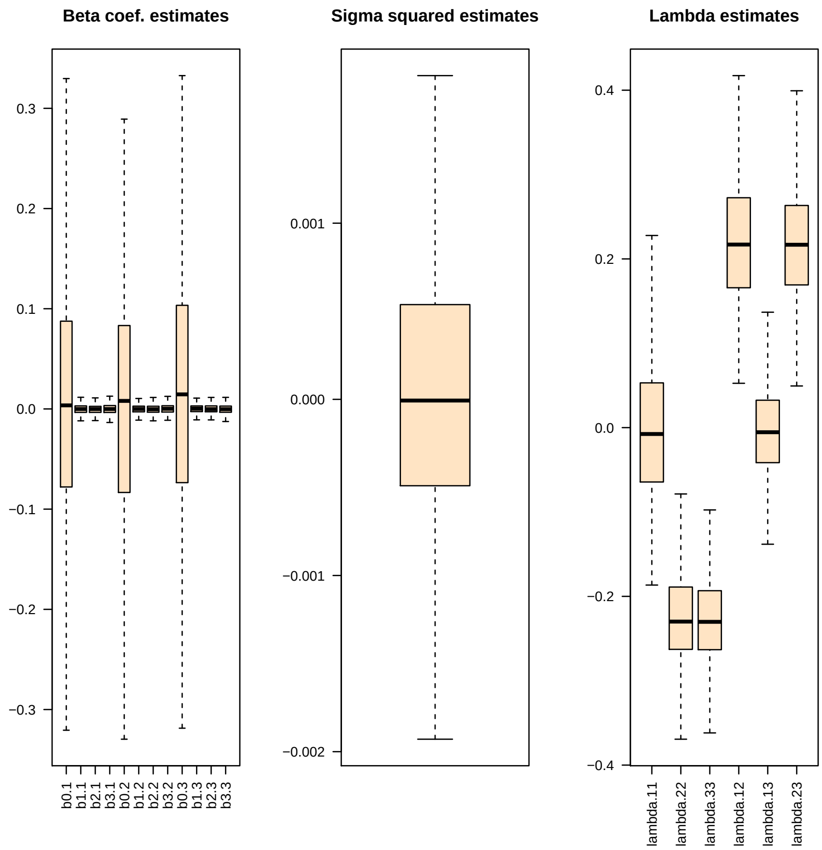


(b) Multidimensional/Joint value added analysis

Figure A.2: Distribution of the coefficient estimates from the simulation Case 2

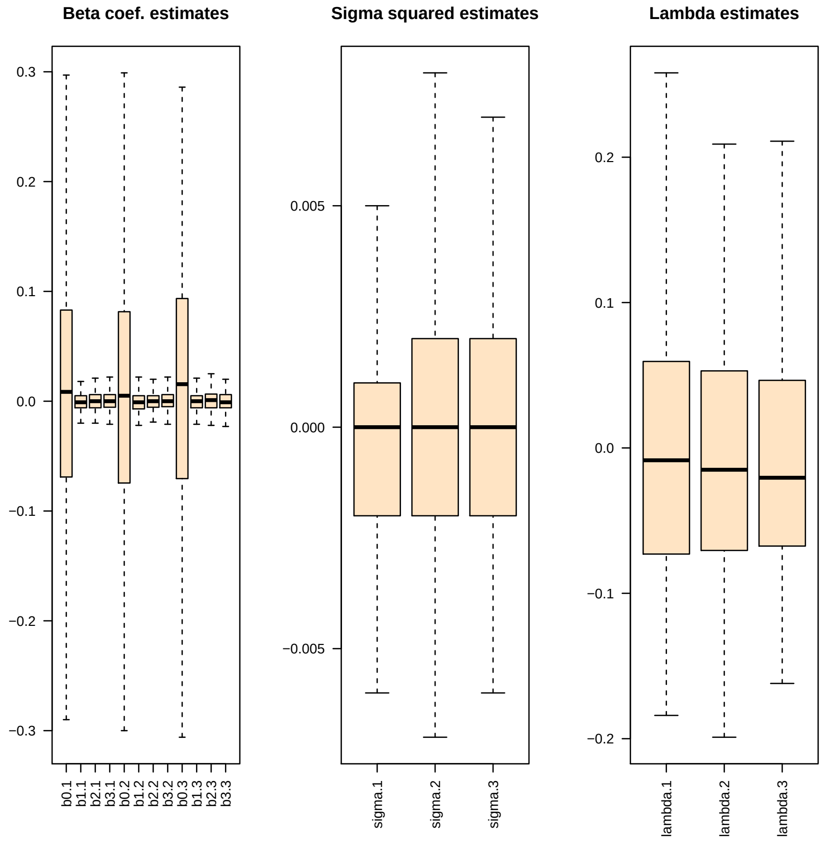


(a) Unidimensional/Separate value added analysis

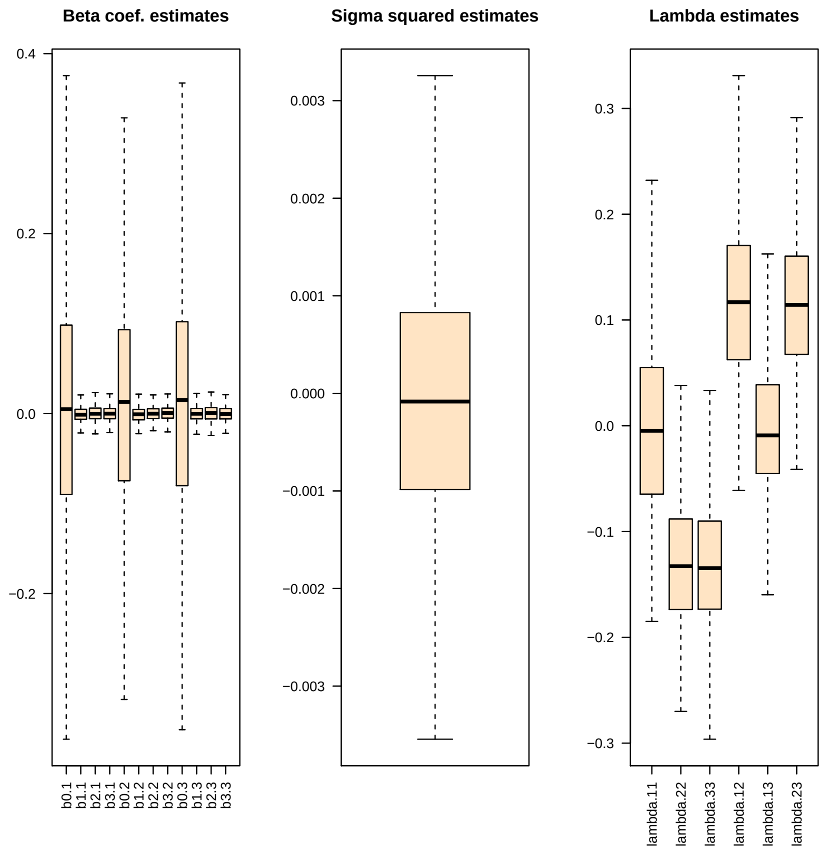


(b) Multidimensional/Joint value added analysis

Figure A.3: Distribution of the coefficient estimates from the simulation Case 3

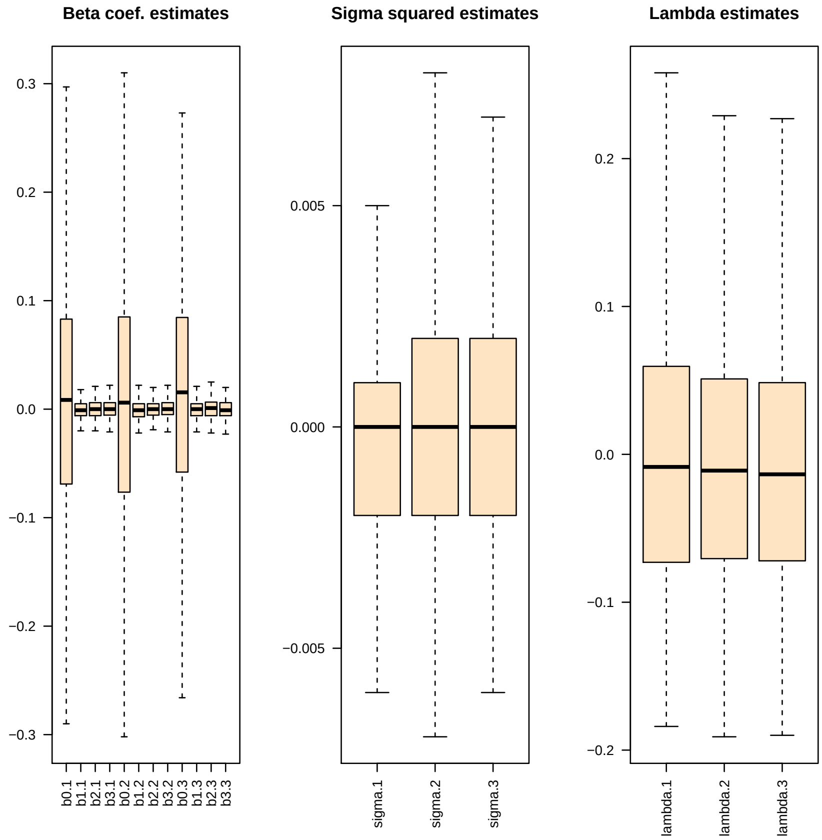


(a) Unidimensional/Separate value added analysis

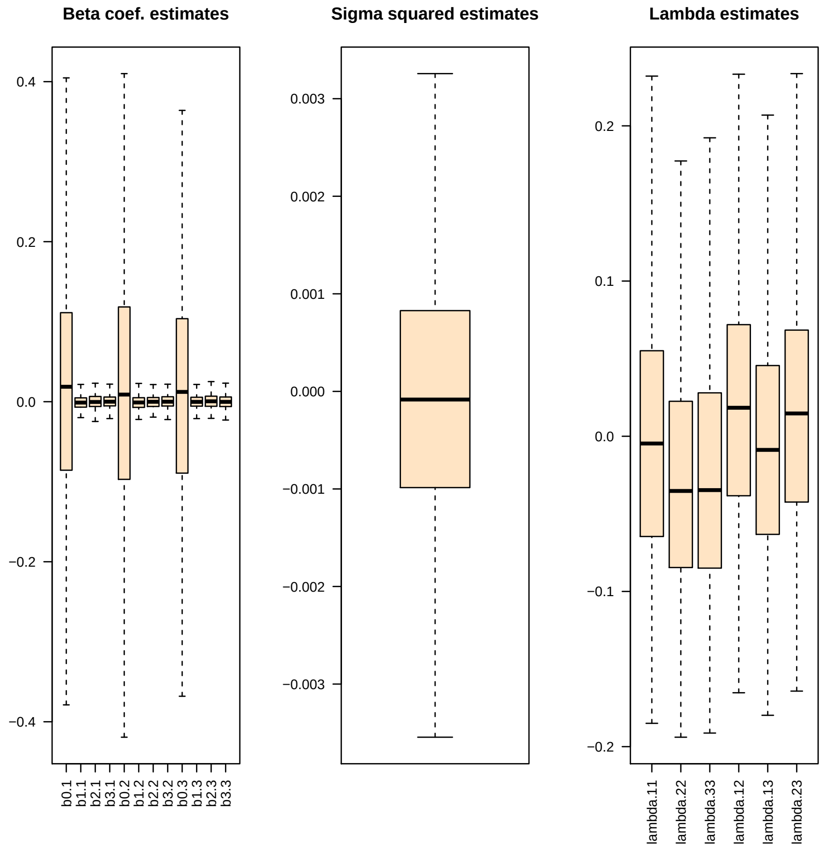


(b) Multidimensional/Joint value added analysis

Figure A.4: Distribution of the coefficient estimates from the simulation Case 4



(a) Unidimensional/Separate value added analysis



(b) Multidimensional/Joint value added analysis

Figure A.5: Distribution of the coefficient estimates from the simulation Case 5

Table A.1: Distribution of students in each reference study program

Field of study	Number of students	Proportion	Cumulative
Engineering	32174	0.234	0.234
Administration	24310	0.177	0.411
Education	11877	0.087	0.498
Law	11772	0.086	0.584
Accounting	9025	0.066	0.649
Health	7463	0.054	0.704
Communication Journalism and Publicity	6491	0.047	0.751
Arts and Design	5322	0.039	0.79
Psychology	5092	0.037	0.827
Social Sciences	4412	0.032	0.859
Nursing	3799	0.028	0.887
Economics	3343	0.024	0.911
Medicine	3071	0.022	0.934
Architecture and City Planning	2841	0.021	0.954
Natural and Exact Sciences	2533	0.018	0.973
Agricultural Sciences	2362	0.017	0.99
Humanities	916	0.007	0.997
Military and Naval Sciences	475	0.003	1

Table A.2: Correlation matrix between the Saber 11 and Saber Pro test scores

Engineering						
	S11 Span	S11 Math	S11 SocSci	S11 Chem	S11 Phys	S11 Engl
Spro WrC	0.23	0.2	0.25	0.20	0.15	0.25
Spro EN	0.47	0.52	0.46	0.49	0.42	0.77
Spro CR	0.46	0.43	0.5	0.44	0.34	0.45
Spro QR	0.44	0.61	0.48	0.52	0.46	0.50
Spro CC	0.44	0.43	0.51	0.43	0.35	0.42
Law						
	S11 Span	S11 Math	S11 SocSci	S11 Chem	S11 Phys	S11 Engl
Spro WrC	0.27	0.21	0.3	0.24	0.16	0.27
Spro EN	0.46	0.45	0.45	0.42	0.33	0.77
Spro CR	0.48	0.36	0.51	0.39	0.26	0.44
Spro QR	0.39	0.47	0.42	0.43	0.34	0.42
Spro CC	0.47	0.38	0.51	0.39	0.27	0.42

Table A.3: GLS unidimensional regression estimates for Engineering

	QR	CR	EN	CC	WrC
Intercept	4.302	4.465	3.766	4.765	5.634
	(0.508)	(0.425)	(0.631)	(0.439)	(0.591)
Mathematics	0.013	0.023	0.014	0.021	0.01
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Spanish	0.026	0.006	0.006	0.007	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Chemistry	0.016	0.027	0.009	0.033	0.013
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Social Sciences	0.023	0.015	0.012	0.016	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
English	0.009	0.003	0.005	0.005	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Physics	0.007	0.009	0.052	0.006	0.007
	(0)	(0)	(0.001)	(0)	(0.001)
INSE	0.002	0.001	0.013	0.001	0
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Academic track	-0.311	0.081	-0.095	0.053	0.263
	(0.01)	(0.009)	(0.011)	(0.01)	(0.012)
Female dummy	0.001	-0.051	0.039	-0.043	-0.02
	(0.018)	(0.017)	(0.02)	(0.018)	(0.022)
Tuition 1–3million	0.093	-0.044	0.106	-0.04	-0.023
	(0.026)	(0.024)	(0.029)	(0.025)	(0.032)
Tuition 3–5million	0.054	-0.031	0.081	-0.053	0.013
	(0.024)	(0.022)	(0.026)	(0.023)	(0.028)
Tuition >5million	-0.087	-0.056	-0.018	-0.075	-0.051
	(0.012)	(0.011)	(0.013)	(0.011)	(0.014)
Composition Spanish	-0.025	0.012	0.031	-0.016	0.03
	(0.015)	(0.013)	(0.019)	(0.013)	(0.018)
Composition Mathematics	0.027	0.006	0.013	0.01	-0.006
	(0.009)	(0.007)	(0.011)	(0.007)	(0.01)
Composition Social Sciences	0.089	0.056	-0.007	0.085	0.064
	(0.013)	(0.011)	(0.017)	(0.011)	(0.016)
Composition Chemistry	-0.044	-0.029	-0.012	-0.039	0.013
	(0.015)	(0.012)	(0.018)	(0.013)	(0.017)
Composition Physics	-0.009	-0.011	-0.033	-0.023	-0.045
	(0.014)	(0.012)	(0.017)	(0.012)	(0.016)
Composition English	-0.008	-0.01	0.028	-0.005	-0.007
	(0.006)	(0.005)	(0.007)	(0.005)	(0.007)
Composition INSE	-0.001	0.003	0.002	0.002	-0.001
	(0.004)	(0.003)	(0.005)	(0.003)	(0.005)
$\hat{\sigma}^2$	0.649	0.559	0.781	0.639	0.949
$\hat{\Lambda}_m$	0.010	0.006	0.018	0.006	0.013
mean $\hat{\gamma_m}$	0.000	0.000	0.000	0.000	0.000

Note: Standard errors in parenthesis. Dependent variable indicated on top of each column. Sample size for one module is 30857 for 126 schools.

Table A.4: GLS unidimensional regression estimates for Law

	QR	CR	EN	CC	WrC
Intercept	5.163	2.903	3.558	4.106	6.237
	(0.733)	(0.681)	(0.775)	(0.786)	(0.937)
Mathematics	0.01	0.027	0.01	0.025	0.012
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Spanish	0.019	0.004	0.005	0.007	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Chemistry	0.014	0.029	0.008	0.03	0.019
	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Social Sciences	0.023	0.015	0.008	0.015	0.011
	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)
English	0.009	0.001	0.004	0.003	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Physics	0.006	0.009	0.052	0.005	0.006
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
INSE	-0.001	-0.002	0.01	-0.003	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Academic track	-0.267	0.005	-0.053	-0.029	0.187
	(0.014)	(0.015)	(0.016)	(0.016)	(0.02)
Female dummy	0.079	0.038	0.116	0.072	0.042
	(0.047)	(0.048)	(0.053)	(0.05)	(0.064)
Tuition 1–3million	0.063	0.013	0.07	0.003	0.002
	(0.043)	(0.044)	(0.049)	(0.047)	(0.059)
Tuition 3–5million	0.031	0.054	0.088	-0.006	0.016
	(0.04)	(0.041)	(0.045)	(0.043)	(0.055)
Tuition >5million	-0.064	-0.089	0.003	-0.088	-0.064
	(0.018)	(0.019)	(0.021)	(0.019)	(0.025)
Composition Spanish	-0.018	0.035	-0.001	0.035	0.032
	(0.018)	(0.016)	(0.019)	(0.019)	(0.023)
Composition Mathematics	0.017	-0.021	0.03	-0.021	-0.002
	(0.014)	(0.013)	(0.015)	(0.015)	(0.018)
Composition Social Sciences	0.052	0.069	-0.015	0.047	0.057
	(0.014)	(0.013)	(0.015)	(0.015)	(0.018)
Composition Chemistry	-0.049	-0.051	0.004	-0.041	-0.027
	(0.024)	(0.022)	(0.025)	(0.026)	(0.031)
Composition Physics	0.024	0.041	-0.013	0.025	-0.024
	(0.026)	(0.024)	(0.027)	(0.027)	(0.033)
Composition English	-0.006	-0.02	0.019	-0.004	-0.006
	(0.007)	(0.006)	(0.007)	(0.007)	(0.009)
Composition INSE	-0.001	0.012	0.012	0.004	-0.002
	(0.006)	(0.006)	(0.006)	(0.006)	(0.008)
$\hat{\sigma}^2$	0.514	0.582	0.68	0.604	1.012
$\hat{\Lambda}_m$	0.014	0.011	0.015	0.016	0.021
mean $\hat{\theta}$	0.000	0.000	0.000	0.000	0.000

Note: Standard errors in parenthesis. Dependent variable indicated on top of each column. Sample size for one module is 11198 for 73 schools.

Table A.5: SURE GLS multidimensional regression estimates for Engineering programs

	QR	CR	EN	CC	WrC
Intercept	10.394	4.931	7.103	-3.383	1.107
	(0.273)	(0.273)	(0.273)	(0.273)	(0.273)
Mathematics	0.01	0.014	0.023	0.013	0.021
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Spanish	0.002	0.006	0.006	0.026	0.007
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Chemistry	0.013	0.009	0.028	0.016	0.033
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Social Sciences	0.003	0.012	0.015	0.023	0.016
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
English	0.001	0.005	0.003	0.009	0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Physics	0.007	0.052	0.009	0.007	0.006
	(0)	(0)	(0)	(0)	(0)
INSE	0	0.013	0.001	0.002	0
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Academic track	0.262	-0.095	0.078	-0.313	0.05
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Female dummy	-0.017	0.047	-0.025	0.019	-0.023
	(0.017)	(0.017)	(0.017)	(0.017)	(0.017)
Tuition 1-3million	-0.021	0.148	-0.006	0.134	-0.028
	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
Tuition 3-5million	0.021	0.114	0.006	0.093	-0.031
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Tuition >5million	-0.051	-0.017	-0.055	-0.085	-0.074
	(0.012)	(0.012)	(0.012)	(0.012)	(0.012)
Composition Spanish	-0.068	0.099	-0.147	0.263	-0.117
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)
Composition Mathematics	0.101	-0.132	-0.073	0.034	-0.013
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Composition Social Sciences	0.035	0.085	0.038	-0.129	0.253
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Composition Chemistry	0.043	0.29	0.102	0.037	0.086
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Composition Physics	-0.15	-0.338	0.058	0.043	-0.048
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Composition English	-0.037	0.103	0.025	-0.094	-0.093
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Composition INSE	0.025	-0.116	-0.016	0.025	0.019
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$\hat{\sigma}^2$	0.716				
mean $\hat{\gamma}_{qr}$	-0.144				
mean $\hat{\gamma}_{cr}$	0.287				
mean $\hat{\gamma}_{en}$	0.05				
mean $\hat{\gamma}_{cc}$	-0.029				
mean $\hat{\gamma}_{wrc}$	0.175				

Note: Dependent variable indicated on the first column of each panel. Sample size for one module is 30857 for 126 schools.

Table A.6: SURE GLS multidimensional regression estimates for Law programs

	QR	CR	EN	CC	WrC
Intercept	5.32	-2.627	-0.487	1.336	-1.636
	(0.432)	(0.432)	(0.432)	(0.432)	(0.432)
Mathematics	0.012	0.01	0.027	0.01	0.025
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Spanish	0.003	0.005	0.004	0.019	0.007
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Chemistry	0.019	0.008	0.029	0.014	0.03
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Social Sciences	0.011	0.008	0.015	0.023	0.015
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
English	0.003	0.004	0.001	0.009	0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Physics	0.007	0.052	0.009	0.006	0.005
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
INSE	-0.002	0.01	-0.002	-0.001	-0.003
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Academic track	0.187	-0.053	0.006	-0.266	-0.028
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Female dummy	-0.009	0.083	0.073	0.091	0.068
	(0.042)	(0.042)	(0.042)	(0.042)	(0.042)
Tuition 1-3million	-0.048	0.052	0.035	0.062	-0.013
	(0.039)	(0.039)	(0.039)	(0.039)	(0.039)
Tuition 3-5million	-0.024	0.056	0.076	0.023	-0.023
	(0.035)	(0.035)	(0.035)	(0.035)	(0.035)
Tuition >5million	-0.058	0.006	-0.087	-0.06	-0.085
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Composition Spanish	0.102	-0.023	0.037	0.131	0.124
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Composition Mathematics	0.022	-0.186	0.072	-0.074	0.018
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Composition Social Sciences	0.004	-0.078	-0.004	-0.04	-0.005
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Composition Chemistry	0.105	0.107	-0.055	-0.026	-0.085
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Composition Physics	-0.157	0.288	0.071	0.116	0.139
	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
Composition English	-0.028	0.058	-0.049	0.019	-0.042
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
Composition INSE	-0.008	0.018	0.064	-0.025	0.013
	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$\hat{\sigma}^2$	0.679				
mean $\hat{\gamma}_{qr}$	0.085				
mean $\hat{\gamma}_{cr}$	-0.067				
mean $\hat{\gamma}_{en}$	-0.06				
mean $\hat{\gamma}_{cc}$	-0.017				
mean $\hat{\gamma}_{wrc}$	0.01				

Note: Dependent variable indicated on the first column of each panel. Sample size for one module is 11198 for 73 schools.

B Technical Appendix

In this Appendix, we derive the estimation procedure used in the paper. Proofs and derivations of other specific results established in the main text are also gathered in this Appendix.

B.1 Notation

In this section we set the notation we use in the Technical Appendix.

- Y_{imj} is the score (scalar) for person i, in school j, in module m, for $i = 1, \ldots, n_j$, $j = 1, \ldots, J$ and $m = 1, \ldots, M$. The following three vectors are accordingly defined:
  $Y_{mj} = (Y_{1mj}, Y_{2mj}, \ldots, Y_{n_j mj})^\top$, a vector of dimension $n_j \times 1$;
  $Y_j = (Y_{1j}^\top, Y_{2j}^\top, \ldots, Y_{Mj}^\top)^\top$, a vector of dimension $Mn_j \times 1$;
  $Y = (Y_1^\top, Y_2^\top, \ldots, Y_J^\top)^\top$, a vector of dimension $MN \times 1$ where $N = \sum_j^J n_j$.
- Z_{imj} is a vector of dimension $K \times 1$ of K-explanatory variables for person i, in school j and for module m, including the intercept.
  $Z_{mj}^\top = (Z_{1mj}, \ldots, Z_{n_j mj})$ a matrix of dimension $n_j \times K$.
  $Z_j = diag(Z_{1j}, \ldots, Z_{Mj}$ a matrix of dimension $Mn_j \times MK$.
  $Z = (Z_1^\top, \ldots, Z_J^\top)^\top$ a matrix of dimension $MN \times MK$.
- $\beta_m = (\beta_{0m}, \beta_{1m}, \ldots, \beta_{K-1,m})^\top$ a vector of dimension $K \times 1$, where β_{0m} corresponds to the coefficient regression of the intercept.
  $\beta = (\beta_1^\top, \beta_2^\top, \ldots, \beta_M^\top)^\top$ a vector of dimension $MK \times 1$.
- $\gamma_j = (\gamma_{1j}, \gamma_{2j}, \ldots, \gamma_{Mj})^\top$ a vector of dimension $M \times 1$.
  $\gamma = (\gamma_1, \ldots, \gamma_J)^\top$ a vector of dimension $MJ \times 1$.
- $\iota_{n_j} = (1, 1, \ldots, 1)^\top$ a vector of dimension $n_j \times 1$ for $j = 1, 2, \ldots, J$.
- $H_j = I_M \otimes \iota_{n_j}^\top$ a $M \times Mn_j$ matrix.
- $H = diag\{H_1, \ldots, H_J\}$ a $MJ \times MN$ matrix.
- $H^\top = diag\{H_1^\top, \ldots, H_J^\top\}$ a $MN \times MJ$ matrix.
- $\overline{H}_j = I_M \otimes \iota_{n_j}^\top / n_j$ a $M \times Mn_j$ matrix.
- $\overline{H} = diag\{\overline{H}_1, \ldots, \overline{H}_J\}$ a $MJ \times MN$ matrix. This is also known as the "between" transform.
- $J_{n_j} = \iota_{n_j} \iota_{n_j}^\top$, a matrix of dimension $n_j \times n_j$ of 1's.
  $\overline{J}_{n_j} = \iota_{n_j} \iota_{n_j}^\top / n_j$, a matrix of dimension $n_j \times n_j$ of $\dfrac{1}{n_j}$'s.
- $W = I_{MN} - H^\top \overline{H}$ a $MN \times MN$ matrix. This is also known as the "within module" transform. $W_j = I_{Mn_j} - I_M \otimes \overline{J}_{n_j}$ for each school j, and $W = diag\,(W_1, W_2, \ldots, W_J)$.
- Note that $H^\top \overline{H} = diag\{H_1^\top \overline{H}_1, \ldots, H_J^\top \overline{H}_J\}$, where

  $$H_j^\top \overline{H}_j = (I_M \otimes \iota_{n_j})(I_M \otimes \iota_{n_j}^\top / n_j) = I_M \otimes \iota_{n_j} \iota_{n_j}^\top / n_j = I_M \otimes \overline{J}_{n_j}.$$

  So $H^\top \overline{H} = diag\{I_M \otimes \overline{J}_{n_1}, \ldots, I_M \otimes \overline{J}_{n_J}\}$, a matrix of dimension $MN \times MN$.
- Below we also need to go from a stacked vector to school-level or module-in-school-level vectors. Denote by P^j the $(Mn_j) \times (MN)$ matrix such that $P^j Y = Y_j$. Similarly, define by P_m^j the $n_j \times (MN)$ that is such that $P_m^j Y = Y_{mj}$. The matrix P_m^j, for instance, is such that

$$P_m^j = \begin{pmatrix} \mathbf{0}_{n_j \times (Mn_1)} & \cdots & \mathbf{0}_{n_j \times (Mn_{j-1})} & \mathbf{0}_{n_j \times ((m-1)n_j)} & \mathbf{I}_{n_j} & \mathbf{0}_{n_j \times ((M-m)n_j)} & \mathbf{0}_{n_j \times (Mn_{j+1})} & \cdots & \mathbf{0}_{n_j \times (Mn_J)} \end{pmatrix}$$

where, for instance, $\mathbf{0}_{n_j \times (Mn_1)}$ is an $n_j \times (Mn_1)$ matrix with only zero entries.

B.2 Structural model

In order to specify the joint distribution generating (Y, Z, γ), we perform a marginal-conditional decomposition. Taking into account that we are dealing with J different schools, it is reasonable to assume that $\{(Y_j, Z_j, \gamma_j) : j = 1, \ldots, J\}$ are mutually independent. As an application of Theorem 7.6.9 in Florens, Mouchart, and Rolin (1990), this condition is equivalent to the following three conditions:

1. $\perp\!\!\!\perp_{1 \leq j \leq J} Y_j \mid (Z, \gamma)$.
2. For each school j, $Y_j \perp\!\!\!\perp (Z, \gamma) \mid (Z_j, \gamma_j)$.
3. $\perp\!\!\!\perp_{1 \leq j \leq J} (Z_j, \gamma_j)$.

Note that the third condition above implies that the school effects γ_j's are mutually independent. By doing so, the model specification is completed by decomposing the joint distribution generating (Y_j, Z_j, γ_j). The order of the decomposition is accordingly the following:

1. The exit test scores vector Y_j is stochastically determined by Z_j (which contains the entry test scores and other covariates) and the school effect γ_j. Empirical analysis reveals a correlation between Y_j and Z_j. Furthermore, it is expected that, after conditioning on Z_j, the school has an impact on the exit test scores. These relationships are consequently represented by the conditional distribution $p(Y_j \mid Z_j, \gamma_j)$.
2. The conditional distribution of Z_j given γ_j does not depend on γ_j because the entry test scores as well as other covariates were measured *before* the school's intervention. This means that

$$Z_j \perp\!\!\!\perp \gamma_j, \tag{B.1}$$

   or in other word, Z_j are exogenous covariates with respect to γ_j.
3. $\gamma_j \sim (0, \Lambda_j)$; that is, the distribution of γ_j is known up to the variance-covariance matrix Λ_j, which is specific to school j. Furthermore, it is allowed that the school-outcome specific effects are correlated among them.
4. The distribution of Z_j is left unspecified as it is typically done with exogenous variables; see Engle, Hendry, and Richard (1983).

The conditional model $p(Y_j \mid Z_j, \gamma_j)$ is specified as follows for each module $m = 1, \ldots, M$,

1. $Y_{mj} \perp\!\!\!\perp (Z_j, \gamma_j) \mid Z_{mj}, \gamma_{mj}$; that is, a module-wise relationship between module-exit scores and covariates.
2. $\perp\!\!\!\perp_{1 \leq i \leq n_j} Y_{imj} \mid Z_{mj}, \gamma_{mj}$; that is, the axiom of local independence.
3. For each $i = 1, \ldots, n_j$, $Y_{imj} \perp\!\!\!\perp Z_{mj} \mid Z_{imj}, \gamma_{mj}$; that is, individual-wise dependency of the exit score with respect to the covariates.
4. $(Y_{imj} \mid Z_{imj}, \gamma_{mj}) \sim (Z_{imj}^{\top} \beta_m + \gamma_{mj}, \sigma^2)$; that is, the conditional distribution of $(Y_{imj} \mid Z_{imj}, \gamma_{mj})$ is known up to its means and variance. Furthermore, $E(Y_{imj} \mid Z_{imj}, \gamma_{mj}) = Z_{imj}^{\top} \beta_m + \gamma_{mj}$ and $Var(Z_{imj}^{\top} \beta_m + \gamma_{mj}) = \sigma^2$.

After staking all outcomes, we obtain

$$E(Y_j \mid Z_j, \gamma_j) = Z_j \beta + \gamma_j \otimes \imath_{n_j}, \quad Var(Y_j \mid Z_j, \gamma_j) = \sigma^2 I_{Mn_j};$$

and after staking schools, we obtain

$$\begin{array}{rl} \text{(i)} & E(Y \mid Z, \gamma) = Z\beta + H^{\top}\gamma, \\ \text{(ii)} & V(Y \mid Z, \gamma) = \sigma^2 I_{MN}, \\ \text{(iii)} & Var(\gamma) = diag(\Lambda_1, \ldots, \Lambda_J). \end{array} \tag{B.2}$$

B.3 Proof of Lemma 2.1

We compute each element of the distribution separately. Based on the assumptions that frame our structural model and the Law of Iterated Expectations, it follows that

$$E(Y_j \mid Z_j) = E[E(Y_j \mid Z_j, \gamma_j) \mid Z_j] = Z_j\beta + E(\gamma_j \otimes \imath_{n_j} \mid Z_j) = Z_j\beta$$

because Z_j is exogenous with respect to γ_j and $E(\gamma_j \mid Z_j) = 0$ because Z_j is exogenous with respect to γ_j. For the variance, we compute $V(\gamma_j \mid Z_j) = \Lambda_j$ by definition. Then,

$$\begin{aligned}
V(Y_j \mid Z_j) &= V(E(Y_j \mid Z_j, \gamma_j) \mid Z_j) + E(V(Y_j \mid Z_j, \gamma_j) \mid Z_j) \\
&= V(\gamma_j \otimes \iota_{n_j}) + \sigma^2 I_{Mn_j} \\
&= \begin{bmatrix} \Lambda^j_{11}\iota_{n_j}\iota_{n_j}^\top & \Lambda^j_{12}\iota_{n_j}\iota_{n_j}^\top & \dots & \Lambda^j_{1M}\iota_{n_j}\iota_{n_j}^\top \\ \Lambda^j_{21}\iota_{n_j}\iota_{n_j}^\top & \Lambda^j_{22}\iota_{n_j}\iota_{n_j}^\top & \dots & \Lambda^j_{2M}\iota_{n_j}\iota_{n_j}^\top \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda^j_{M1}\iota_{n_j}\iota_{n_j}^\top & \Lambda^j_{M2}\iota_{n_j}\iota_{n_j}^\top & \dots & \Lambda^j_{MM}\iota_{n_j}\iota_{n_j}^\top \end{bmatrix} + \sigma^2 I_{Mn_j} \\
&= \Lambda_j \otimes (\iota_{n_j}\iota_{n_j}^\top) + \sigma^2 I_{Mn_j}.
\end{aligned}$$

and

$$\begin{aligned}
cov(Y_j, \gamma_j \mid Z_j) &= cov(E(Y_j \mid Z_j, \gamma_j) \mid Z_j) \\
&= cov(\gamma_j \otimes \iota_{n_j}, \gamma_j \mid Z_j) \\
&= cov(\gamma_j \otimes \iota_{n_j}, \gamma_j) \quad \text{by (B.1)} \\
&= \Lambda_j \otimes \iota_{n_j}.
\end{aligned}$$

The moments of γ_j are directly obtained from the model specification.

B.4 Definition of the Value Added

Following Manzi et al. (2014), the definition of value added for each school j and module m is given by

$$VA_{mj} = \frac{1}{n_j}\sum_{i=1}^{n_j} E\left(Y_{imj} \mid Z_{imj}, \gamma_{mj}\right) - \frac{1}{n_j}\sum_{i=1}^{n_j} E\left(Y_{imj} \mid Z_{imj}\right)$$

which, upon derivations using the structural model results to be equal to γ_{mj}. More specifically,

$$\begin{aligned}
E\left(Y_j \mid Z_j, \gamma_j\right) - E\left(Y_j \mid Z_j\right) &= (Z_j\beta + \gamma_j \otimes \iota_{n_j}) - E\left(E\left(Y_j \mid Z_j\gamma_j\right) \mid Z_j\right) \\
&= (Z_j\beta + \gamma_j \otimes \iota_{n_j}) - E(Z_j\beta + \gamma_j \otimes \iota_{n_j} \mid Z_j) \\
&= (Z_j\beta + \gamma_j \otimes \iota_{n_j}) - Z_j\beta - E(\gamma_j \otimes \iota_{n_j} \mid Z_j) \\
&= \gamma_j \otimes \iota_{n_j}
\end{aligned}$$

because $E(\gamma_j \otimes \iota_{n_j} \mid Z_j) = 0$ by (B.1).

B.5 Estimation by the method of moments

B.5.1 Estimation of σ^2

Let

$$Y = Z\beta + H^\top \gamma + u, \tag{B.3}$$

where $u = Y - E(Y \mid Z, \gamma)$. Applying the transform W-operator to this equation, we obtain

$$WY = WZ\beta + Wu,$$

where

$$Var(Wu) = \sigma^2 WW^\top = \sigma^2 W.$$

Then, using the fact that W is a projection matrix,

$$\widehat{\beta}^w = (Z^\top WZ)^{-1} Z^\top WY.$$

Let define the following residual:

$$\begin{aligned} \widehat{e}^w &= WY - WZ\widehat{\beta}^w \\ &= [I_{MN} - WZ(Z^\top WZ)^{-1}Z^\top W]WY, \text{ because } W \text{ is a projection matrix.} \\ &:= MWY. \end{aligned}$$

It can be verified that M is idempotent and symmetric (i.e., a projection matrix). Taking into account that $M(WY) = M(WZ\beta + Wu) = M(Wu)$, it follows that

$$\begin{aligned} E\left[(\widehat{e}^w)^\top (\widehat{e}^w)\right] &= E\left[(WY)^\top M (WY)\right] \\ &= \operatorname{tr}\left[M \operatorname{Var}(Wu)\right] \\ &= \sigma^2 \operatorname{tr}(MW) \\ &= \sigma^2 \left[\operatorname{tr}(W) - \operatorname{tr}(WZ(Z^\top WZ)^{-1}W)\right] \\ &= \sigma^2 \left[\operatorname{tr}(W) - tr(I_{MK})\right] \\ &= \sigma^2 \left[MN - MJ - MK^*\right] \\ &= \sigma^2 M(N - J - K^*), \end{aligned}$$

where $K^* \leq K$ is the number of non-zero covariates in the Within regression. Therefore, the estimation of σ^2 is given by

$$\widehat{\sigma}^2 = \frac{(Y - Z\widehat{\beta}^w)^\top W(Y - Z\widehat{\beta}^w)}{M(N - J - K^*}.$$

B.5.2 Estimation of Λ_j

In equation (B.3), define $v := H^\top \gamma + u$ as the unobserved component. From the model, the OLS estimator of β, $\hat{\beta} = (Z^\top Z)^{-1} Z^\top Y$, is a consistent estimator. The corresponding residuals are given by $\hat{e} = Y - Z\hat{\beta}$. Let us introduce a partition on the MN-vector of residuals $\hat{e}$ in the following way: Denote by $\hat{e}^j_m = P^j_m \hat{e}$ the subvector of $\hat{e}$ for school j in module m (the projection matrix P^j_m has been defined above in Section B.1). Observe that $(\hat{e}^j_m)^\top \hat{e}^j_{m'}$ is a biased estimator of $\Lambda^j_{mm'}$. We need to analyze the bias in order to find our final estimator.

Let $Q = I - Z(Z^\top Z)^{-1} Z^\top$ be the projection onto the vectorial space that is orthogonal to the columns of Z. We also denote by $\Pi^j_{mm'} = (P^j_m)^\top P^j_{m'}$ an $MN \times MN$ matrix. Note that $(\hat{e}^j_m)^\top \hat{e}^j_{m'} = \hat{e}^\top \Pi^j_{mm'} \hat{e}$.

Using this notation, we can now compute the expectation of the estimator. Using that γ and u are independent, we obtain the following decomposition:

$$E\left[(\hat{e}_m^j)^\top \hat{e}_{m'}^j\right] = E\left(\gamma^\top H Q^\top \Pi_{mm'}^j Q H^\top \gamma\right) + E\left(u^\top Q^\top \Pi_{mm'}^j Q u\right). \tag{B.4}$$

Each term is the expectation of a quadratic form, that we derive each below. First,

$$\begin{aligned}
b_{mm'}^j &:= E\left(u^\top Q^\top \Pi_{mm'}^j Q u\right) \\
&= \sigma^2 tr(Q^\top \Pi_{mm'}^j Q) \\
&= \sigma^2 tr(Q^\top P_m^{j\top} P_{m'}^j Q) \\
&= \sigma^2 tr(P_{m'}^j Q P_m^{j\top})
\end{aligned}$$

which is known because we have already estimated σ^2 in the previous section.

Second,

$$\begin{aligned}
E\left(\gamma^\top H Q^\top \Pi_{mm'}^j Q H^\top \gamma\right) &= tr(\Lambda^{1/2} H Q^\top \Pi_{mm'}^j Q H^\top \Lambda^{1/2}) && \text{(B.5)} \\
&= tr(\Lambda H Q^\top \Pi_{mm'}^j Q H^\top) && \text{(B.6)}
\end{aligned}$$

and contains the unknown matrix Λ in a tricky expression. Denote by $G_{mm'}^j$ the known matrix, $HQ^\top \Pi_{mm'}^j Q H^\top$, that depends only on the covariates. Observe that

$$tr(\Lambda G_{mm'}^j) = \Lambda_{mm'}^1 \alpha_{mm'}^{j,1} + \Lambda_{mm'}^2 \alpha_{mm'}^{j,2} + \cdots + \Lambda_{mm'}^J \alpha_{mm'}^{j,J} \tag{B.7}$$

for some $\alpha_{mm'}^{j,l}$ entries of $G_{mm'}^j$, for every $m, m' = 1, \ldots, M$ and every $j, l = 1, \ldots, J$. By restricting to $m \leqslant m'$, (B.4) is, therefore, equivalently rewritten as

$$E\left[(\hat{e}_m^j)^\top \hat{e}_{m'}^j\right] = b_{mm'}^j + (\Lambda_{mm'}^1 \alpha_{mm'}^{j,1} + \cdots + \Lambda_{mm'}^J \alpha_{mm'}^{j,J}). \tag{B.8}$$

We now summarize the estimation procedure for Λ^j.

1. Compute the OLS estimator using *all* the data : $\hat{\beta} = (Z^\top Z)^{-1} Z^\top Y$.
2. Compute the residuals $\hat{e} = Y - Z\hat{\beta}^w$ and define the vector v^j of size $M(M+1)/2$ that is such that

$$v^j = \left(\hat{e}_1^{j\top} \hat{e}_1^j, \hat{e}_1^{j\top} \hat{e}_2^j, \ldots, \hat{e}_1^{j\top} \hat{e}_M^j, \hat{e}_2^{j\top} \hat{e}_2^j, \ldots, \hat{e}_2^{j\top} \hat{e}_M^j, \ldots, \hat{e}_M^{j\top} \hat{e}_M^j\right)^\top \tag{B.9}$$

   with $\hat{e}_m^j = P_m^j \hat{e}$ and consider the vector of size $JM(M+1)/2$

$$v = \left((v^1)^\top, \ldots, (v^J)^\top\right)^\top. \tag{B.10}$$

3. Let b^j be a vector of size $M(M+1)/2$ such that

$$b^j = (b_{11}^j, b_{12}^j, \ldots, b_{1M}^j, b_{22}^j, b_{23}^j, \ldots, b_{2M}^j, \ldots, b_{MM}^j)^\top \tag{B.11}$$

   with $b_{mm'}^j = \hat{\sigma}^2 tr(P_{m'}^{j\top} Q Q^\top P_m^j)$.
4. Consider the vectors of size $JM(M+1)/2$ such that

$$b = \left((b^1)^\top, \ldots, (b^J)^\top\right)^\top. \tag{B.12}$$

5. For every school j, denote by $diag(\alpha^{j,l})$ the diagonal matrix of size $M(M+1)/2$ having elements

$$(\alpha_{11}^{j,l}, \alpha_{12}^{j,l}, \ldots, \alpha_{1M}^{j,l}, \alpha_{22}^{j,l}, \alpha_{23}^{j,l}, \ldots, \alpha_{2M}^{j,l}, \ldots, \alpha_{MM}^{j,l})^\top$$

over the diagonal. Denote by Γ^j the matrix of size $M(M+1)/2 \times JM(M+1)/2$ concatenating the last matrices:

$$\Gamma^j = \left[diag(\alpha^{j,1}), \ldots, diag(\alpha^{j,J})\right]$$

which yields a matrix of dimensions $M(M+1)/2 \times JM(M+1)/2$ since $j = 1, \ldots, J$.

6. Stack the Γ^j for each school j to obtain a $JM(M+1)/2$ square matrix

$$\Gamma = \begin{bmatrix} \Gamma^1 \\ \vdots \\ \Gamma^J \end{bmatrix}.$$

7. Observe that: $v = b + \Gamma\tilde{\Lambda}^j$ where $\tilde{\Lambda}^j$ is defined in the next item.

8. Denote by $\tilde{\Lambda}^j = (\Lambda_{11}^j, \ldots, \Lambda_{1M}^j, \Lambda_{22}^j, \Lambda_{23}^j, \ldots, \Lambda_{2M}^j, \ldots, \Lambda_{MM}^j)^\top$ the vector given by

$$\Gamma^{-1}(v - b). \tag{B.13}$$

9. The estimated matrix Λ^j is an arrangement of the estimated vector $\tilde{\Lambda}^j$ ($\hat{\Lambda}_{mm'}^j = \tilde{\Lambda}_{mm'}^j$).

B.5.3 Estimation of β

The efficient estimation is given by the GLS estimator $\hat{\beta}^g = (Z^\top\Omega^{-1}Z)^{-1}Z^\top\Omega^{-1}Y$ where Ω is a consistent estimator of $Var(H^\top\gamma + u \mid Z)$. A consistent estimator is provided by

$$\hat{\Omega} = H^\top\hat{\Lambda}H + \hat{\sigma}^2 I_{MN}.$$

The practical inversion of that matrix is a difficulty, given its huge size. Two results are useful to simplify that step. First, observe that $\hat{\Omega}$ is a block diagonal matrix. Since there is no correlation between two schools, the inversion of $\hat{\Omega}$ simplifies to the inversion of each by-school block

$$\begin{aligned} \hat{\Omega}_j &= H_j\hat{\Lambda}H_j^t + \hat{\sigma}^2 I_{Mn_j} \\ &= \Lambda^j \otimes J_{n_j} + \hat{\sigma}^2 I_{Mn_j}. \end{aligned}$$

For each school, this empirical covariance matrix is of size $Mn_j \times Mn_j$. The second result is a recursive method to inverse the matrix in a given school, which is provided in the next section.

B.6 Inversion of Ω_j

It is useful to notice that $\hat{\Omega}_j$ is decomposed into blocks of size $n_j \times n_j$:

$$\hat{\Omega}_j = \begin{pmatrix} A_{11} & A_{12} & \ldots & A_{1M} \\ A_{21} & A_{22} & \ldots & A_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M1} & A_{M2} & \ldots & A_{MM} \end{pmatrix}$$

where

$$A_{mm} = \alpha_m \overline{J}_{n_j} + \beta_m E_{n_j} \qquad \text{with } \alpha_m = n_j(\hat{\Lambda}^j_{mm}) + \hat{\sigma}^2, \text{ and } \beta_m = \hat{\sigma}^2$$

and, for $m \neq p$,

$$A_{mp} = \zeta_{mp} \overline{J}_{n_j} \qquad \text{with } \zeta_{mp} = n_j \hat{\Lambda}^j_{mp}$$

and $E_{n_j} = I_{n_j} - \overline{J}_{n_j}$. Note that elements of A_{mp} depend on j (including $n = n_j$). We skipped the j to simplify the notation.

B.6.1 Inversion when $M = 2$

In the following, we use recursively the inversion formula for partitioned matrices: If $G = (D - CA^{-1}B)^{-1}$ it is easy to check that

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1}(I_n + BGCA^{-1}) & -A^{-1}BG \\ -GCA^{-1} & G \end{pmatrix}. \tag{B.14}$$

Using the spectral decomposition[6] of A_{11}:

$$\begin{aligned} A_{11}^{-1} &= \alpha_1^{-1}\overline{J}_n = \beta_1^{-1}E_n \\ G_2 := A_{22} - A_{21}A_{11}^{-1}A_{12} &= \left(\alpha_2 - \frac{\zeta_{12}\zeta_{21}}{\alpha_1}\right)\overline{J}_n + \beta_2 E_n \end{aligned}$$

and therefore

$$\Omega_j = \begin{pmatrix} \frac{1}{\alpha_1}\left(1 + \frac{\zeta_{12}\zeta_{21}}{\alpha_1\alpha_2 - \zeta_{12}\zeta_{21}}\right)\overline{J}_n + \frac{1}{\beta_1}E_n) & -\frac{\zeta_{12}}{\alpha_1\alpha_2 - \zeta_{21}\zeta_{12}}\overline{J}_n \\ -\frac{\zeta_{21}}{\alpha_1\alpha_2 - \zeta_{21}\zeta_{12}}\overline{J}_n & \frac{\alpha_2}{\alpha_1\alpha_2 - \zeta_{21}\zeta_{12}}\overline{J}_n + \frac{1}{\beta}_2 E_n \end{pmatrix}. \tag{B.15}$$

when $M = 2$

B.6.2 Recursive inversion: from $M - 1$ to M

Denote

$$\hat{\Omega}_j = \left(\begin{array}{ccc|c} & & & A_{1M} \\ & A & & A_{2M} \\ & & & \vdots \\ & & & A_{2,M-1} \\ \hline A_{M1} & A_{M2} & \dots\ A_{M,M-1} & A_{MM} \end{array}\right)$$

with

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1,M-1} \\ A_{21} & A_{22} & \dots & A_{2,M-1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{M-1,1} & A_{M-1,2} & \dots & A_{M-1,M-1} \end{pmatrix}.$$

Matrix A has just been inverted in step $M - 1$. For some x_m, y_m and z_m the entries of A^{-1} are

$$\begin{aligned} A_{mm}^{-1} &= x_{mm}\overline{J}_n + y_m E_n \\ A_{mp}^{-1} &= x_{mp}\overline{J}_n \qquad (m \neq p) \end{aligned}$$

[6] The columns of $\overline{J}_n$ and E_n are mutually orthonormal and are the eigenfunctions of A_{11}

and construct the matrix $X_{M-1} = (x_{mp})_{m,p=1}^{M-1}$ and the vector $\zeta_{M-1} = (\zeta_{M1}, \dots, \zeta_{M,M-1})^\top$. Then using the symmetry of Ω_j,

$$\begin{aligned} G_M &:= A_{MM} - (A_{M,1}, \dots, A_{M,M-1}) A^{-1} (A_{1,M}, \dots, A_{M-1,M})^\top \\ &= (\alpha_M + \zeta_{M-1}^\top X_{M-1} \zeta_{M-1}) \overline{J}_n \end{aligned}$$

Now, after some derivations and using the symmetry of X_{M-1}, we obtain

$$\hat{\Omega}_j^{-1} = \begin{pmatrix} A^{-1} + \frac{X_{M-1}^\top \zeta_{M-1} \zeta_{M-1}^\top X_{M-1}}{\alpha_M + \zeta_{M-1}^\top X_{M-1} \zeta_{M-1}} \overline{J}_n & -\frac{X_{M-1}^\top \zeta_{M-1}}{\alpha_M + \zeta_{M-1}^\top X_{M-1} \zeta_{M-1}} \overline{J}_M \\ -\frac{\zeta_{M-1}^\top X_{M-1}}{\alpha_M + \zeta_{M-1}^\top X_{M-1} \zeta_{M-1}} \overline{J}_M & \frac{1}{\alpha_M + \zeta_{M-1}^\top X_{M-1} \zeta_{M-1}} \overline{J}_n + \frac{1}{\beta_M} E_n \end{pmatrix}.$$

Recent titles

CORE Discussion Papers

2015/12 Mathieu PARENTI, Philip USHCHEV, Jacques-François THISSE. Toward a theory of monopolistic competition.
2015/13 A.B. TAYLOR, J.M. HENDRICKX and F. GLINEUR. Smooth strongly convex interpolation and exact worst-case performance of first-order methods.
2015/14 Christian HAFNER, Hans MANNER and Léopold SIMAR. The "wrong skewness" problem in stochastic frontier models: A new approach.
2015/15 Paul BELLEFLAMME, Nesssrine OMRANI Martin PEITZ. The Economics of Crowdfunding Platforms.
2015/16 Samuel FEREY and Pierre DEHEZ. Multiple Causation, Apportionment and the Shapley Value.
2015/17 Mehdi MADANI and Mathieu VAN VYVE. A MIP framework for Non-convex uniform price day-ahead electricity auctions.
2015/18 Yukihiro NISHIMURA and Pierre PESTIEAU. Efficient taxation with differential risks of dependence and mortality.
2015/19 Louis N. CHRISTOFIDES, Michael HOY, Joniada MILLA and Thanasis STENGOS. Nature or nurture in higher education? Inter-generational implications of the Vietnam-Era lottery.
2015/20 Philippe J. DESCHAMPS, Alternative lormulations of the leverage Effect in a stochatic volatility model with asymmetric leavy-tailed errors.
2015/21 Jean HINDRIKS and Yukihiro NISHIMURA. Equilibrium leadership in tax competition models with capital ownership: a rejoinder.
2015/22 Frédéric VRINS and Monique JEANBLANC. The Φ-Martingale.
2015/23 Wing Man Wynne LAM. Attack-dettering and damage dontrol investments in cybersecurity.
2015/24 Wing Man Wynne LAM. Switching costs in two-sided markets.
2015/25 Philippe DE DONDER, Marie-Louise LEROUX. The political choice of social long term care transfers when family gives time and money.
2015/26 Pierre PESTIEAU and Gregory PONTHIERE. Long-term care and births timing.
2015/27 Pierre PESTIEAU and Gregory PONTHIERE. Longevitiy variations and the welfare State.
2015/28 Mattéo GODIN and Jean HINDRIKS. A review of critical issues on tax design and tax administration in a global economy and developing countries
2015/29 Michel MOUCHART, Guillaume WUNSCH and Federica RUSSO. The issue of control in multivariate systems, A contribution of structural modelling.
2015/30 Jean J. GABSZEWICZ, Marco A. MARINI and Ornella TAROLA. Alliance formation in a vertically differentiated market.
2015/31 Jens Leth HOUGAARD, Juan D. MORENO-TERNERO, Mich TVEDE and Lars Peter ØSTERDAL. Sharing the proceeds from a hierarchical venture.
2015/32 Arnaud DUFAYS and Jeroen V.K. ROMBOUTS. Spare change-point time series models.
2015/33 Wing Man Wynne LAM. Status in organizations.
2015/34 Wing Man Wynne LAM. Competiton in the market for flexible resources : an application to cloud computing. 2015/35 Yurii NESTEROV and Vladimir SHIKHMAN. Computation of Fisher-Gale equilibrium by auction.
2015/36 Maurice QUEYRANNE and Laurence A. WOLSEY. Thight MIP formulations for bounded up/down tiems and interval-dependent start-ups.

Recent titles
CORE Discussion Papers – continued

2015/37 Paul BELLEFLAMME and Dimitri PAOLINI. Strategic promotion and release decisions for cultural goods.
2015/38 Nguyen Thang DAO and Julio DAVILA. Gender inequality, technologial progress, and the demographic transition.
2015/39 Thomas DEMUYNCK, Bram DE ROCK and Victor GINSBURGH. The transfer paradox in welfare space.
2015/40 Pierre DEHEZ. On Harsanyi dividends and asymmetric values.
2015/41 Laurence A. WOLSEY. Uncapacitated lot-sizing with stock upper bounds, stock fixed costs, stock overloads and backlogging: A tight formulation.
2015/42 Paul BELLEFLAMME. Monopoly price discrimination and privacy: the hidden cost of hiding.
2015/43 Pierre PESTIEAU and Gregory PONTHIERE. Optimal fertility under age-dependent labor productivity.
2015/44 Jacques DREZE. Subjective expected utility with state-dependent but action/observation-independent preferences
2015/45 Joniada MILLA, Ernesto SAN MARTÍN and Sébastien VAN BELLEGEM. Higher education value added using multiple outcomes.

Books

W. GAERTNER and E. SCHOKKAERT (2012), *Empirical Social Choice*. Cambridge University Press.
L. BAUWENS, Ch. HAFNER and S. LAURENT (2012), *Handbook of Volatility Models and their Applications*. Wiley.
J-C. PRAGER and J. THISSE (2012), *Economic Geography and the Unequal Development of Regions*. Routledge.
M. FLEURBAEY and F. MANIQUET (2012), *Equality of Opportunity: The Economics of Responsibility*. World Scientific.
J. HINDRIKS (2012), *Gestion publique*. De Boeck.
M. FUJITA and J.F. THISSE (2013), *Economics of Agglomeration: Cities, Industrial Location, and Globalization.* (2nd edition). Cambridge University Press.
J. HINDRIKS and G.D. MYLES (2013). *Intermediate Public Economics.* (2nd edition). MIT Press.
J. HINDRIKS, G.D. MYLES and N. HASHIMZADE (2013). *Solutions Manual to Accompany Intermediate Public Economics.* (2nd edition). MIT Press.
J. HINDRIKS (2015). *Quel avenir pour nos pensions ? Les grands défis de la réforme des pensions.* De Boeck.
P. BELLEFLAMME and M. PEITZ (2015). *Industrial Organization: Markets and Strategies* (2nd edition). Cambridge University Press.

CORE Lecture Series

R. AMIR (2002), Supermodularity and Complementarity in Economics.
R. WEISMANTEL (2006), Lectures on Mixed Nonlinear Programming.
A. SHAPIRO (2010), Stochastic Programming: Modeling and Theory.