Sensivity of Value Added Model Specifications: Measuring Socio-Economic Status

Mª Eugenia Ferrão
Universidade da Beira Interior. Departamento de Matemática. Covilhã, Portugal

Abstract

The paper describes the extent to which two different measures of Socio-Economic Status (SES) or the exclusion of that controlling variable from the Value Added Model (VAM) changes the estimates of school value added.

The statistical model used to estimate the school value added is a variance component model where pupils are level one unit and schools are level two units. Prior achievement is included as explanatory variable as well. The data used in this paper is derived from a school effectiveness research project (*Eficácia Escolar no Ensino da Matemática, 3EM* project) and was collected in Cova da Beira (NUT III), Portugal. It is a longitudinal data set which allows pupils to be followed through their schooling. However, for the purpose of this paper, we only used the data collected at the beginning and at the end of the academic year 2005-06 for the 1st, 3rd, 5th, 7th and 8th grades. SES variables considered are: (1) The student eligibility for Free School Meals and Books (FSM); (2) Parent's education (parent responsible for the pupil). Evidence shows that prior achievement shrinks ESE effects. Results also show that value-models are sensible to the way achievement is measured, even when there is only one subject.

Key Words: Value added models, socio-economical status, adjusting covariates, accountability, assessment systems.

SES in the VAM literature

Value Added Model is a family of statistical models that are employed to make inferences about the effectiveness of educational units, usually schools and/or teachers (Braun & Wainer, 2007). Educators, researchers and policymakers all generally agree that schooling is only one of many factors that affect student achievement and learning. One of the other factors that has long been recognised to contribute to a student's educational progress is his/her socioeconomic status, which is a strong predictor of student performance. Sociologists use the term socioeconomic status to refer to the relative position of a family or individual in a hierarchical social structure, based on their access to, or control over, wealth, prestige and power. The SES position in this hierarchy affects their educational opportunities and a measure of SES is usually used as a control variable in VAM or school effects model.

Thomas and Mortimore (1996) have compared five multilevel models of varying complexity in order to choose the best VAM and chose the one whose range of individual intake variables was students' prior achievement in verbal, quantitative, and non-verbal cognitive ability tests, their gender, age, ethnicity, mobility and entitlement for free school meals. The prior achievement measures were found to be the most important factors to control, and similar findings were also been reported by Gray et al. (1995). In line with this, Sammons et al. (1997, p. 43) demonstrated that the prior achievement is the most important factor required to control intake differences in measuring value added (using a unique sample of inner London), and they also showed that the inclusion of socio-economic factors in the analysis is highly relevant.

Rubin, Stuart and Zanutto (2004) argue in favour of considering SES as control variable in VAMs:

> Thus we see that many complications exist when thinking about an ideal randomized experiment, and even more complications arise when thinking about using observational data, of course, is the more realistic scenario. With observational data, one key goal is to find treated and control units that look as similar as possible on background covariates. If the groups look very different on background covariates, the results are likely to be based on untestable modelling assumptions and extrapolation. […] Because the values of ‘percent minority’ and ‘percent in poverty’ differ widely in different schools, as illustrated in Table 2 in Tekwe at al. (2004), it is likely that the estimates adjusting for such covariates using models rely heavily on extrapolation, even if students were randomly assigned to those schools after being subclassified into blocks (with dramatically different probabilities of treatment assignment between blocks but similar probabilities within blocks). This situation implies extreme sensitivity to these models' assumptions. If school A has no students who ‘look like’ students in the other schools, it is impossible to estimate the effect of school A relative to the comparison schools without making heroic assumptions.

There are some VAM experiences that do not include such variables (Ladd & Walsh, 2002) or conclude that SES and demographic variables at the student level had little effect on the value added assessment of teachers, since the longitudinal history of a student's performance serves as a substitute for those ‘omitting’ variables (Ballou et al., 2004).

Ladd & Walsh (2002) do not control VA estimates for SES. They propose an inclusion of more than one year of prior achievement as instrumental variable for adjusting for measurement error in a growth model. However, they repeatedly refer to the influence and importance of that construct in terms of getting reliable school value-added estimates.

> Had South Carolina … many of the schools serving low-performing students (which also tend to be those serving students with low socioeconomic status) would have been declared more effective than they appeared to be according to the state’s ranking and the reverse would have been true for schools serving high-performing students (p. 11). […] The combined result may well be that high quality teachers and administrators try to avoid schools serving low SES students in favour of schools serving high SES students. While anecdotal evidence from North Carolina is consistent with this view, we are not aware of any systematic study of the magnitude of this effect and believe it deserves further investigation. The larger this incentive effect, the more the accountability system would reduce the quality of education in the schools where achievement gains are most needed.

Despite being in favour of including SES as a student background variable, McCaffrey, Lockwood, Koretz and Hamilton (2003, p.69-70; 2004) conclude that controlling for student-level socioeconomic and demographic factors alone will not be sufficient enough to remove the effects of background characteristics in all school systems, especially those systems which serve heterogeneous students.

SES measures and the model

For the purpose of this paper two variables are used as *proxy* for SES: Parents' Education and student eligibility for Free School Meals and Books. These variables are often used in VA or school/teacher effects studies. We are not completely sure that these variables are *valid* for representing the attribute SES. Work is being done to develop a compound index for student socioeconomic and cultural status, which includes (1) parents' education, occupation; (2) cultural capital (how many times, in the past year, the student attended a concert, went to a museum, art gallery or theatre, etc.); (3) social capital ('obligations, expectations, and trustworthiness', Coleman, 1988). The great deal of data needed to obtain the SES index implies that, only due to the SES, the amount of missing data increases by 26%, which constitutes a limitation to its use.

We consider a two-level variance component model with pupils (indexed by *i*) at level 1 and class-schools (indexed by *j*) at level 2. Thus, value added is quantified by adjusted residuals ($\hat{u}_{os}$) of the equation of level 2; $\hat{u}_{os}$ represents the deviation of the class-school performance ($\hat{\beta}_{os}$) to the overall mean ($\hat{\gamma}_{00}$), adjusting (or controlling) for student prior-achievement ($x1_{js}$) and student and school SES ($x2_{js}$ and $x3_s$, respectively). The model we wish to estimate, based on true values, is written as

$$\begin{aligned} y1_{js} &= \beta_{0s} + \beta_1 x1_{js} + \beta_2 x2_{js} + \beta_3 x3_s + \varepsilon_{js} \\ \beta_{0s} &= \gamma_{00} + u_{0s} \\ \varepsilon_{js} &\sim N(0, \sigma_e^2) \\ u_{0s} &\sim N(0, \sigma_{u0}^2) \end{aligned} \tag{1}$$

The response variable is a normalised maths score (score_2) equated[1] with math prior achievement (score_1).

The data used in this paper is derived from a school effectiveness research project (*Eficácia Escolar no Ensino da Matemática*, 3EM project) and was collected in Cova da Beira (NUT III), Portugal. Students enrolled in compulsory education (primary –four years–, elementary –two years– and lower secondary –three years) define the target population. The random sample is representative of the county level and NUT III region (Vicente, 2006). The initial sample was oversampled in order to take account of parents non-agreement and dropout or attrition, which is a known problem in longitudinal studies. The largest dropout rate is 4.8% at the 8th grade. In primary education classes the rate is less than 1%. The actions of teachers and principals strongly contributed to keep the rate at a low level.

The dropout and missing responses, mainly due to the parent's education variable, reduce the number of cases by 6.1%, 5.7%, 10.0%, 8.1% and 10.3% at each grade, respectively. For the purpose of parameter estimation, missing responses are assumed 'missing at random' (Little & Rubin, 2002). The survey design is longitudinal, which allows pupils to be followed through their schooling, and consists of three waves –2005-06, 2006-07 and 2007-08, and data is collected at the beginning and at the end of each academic year. For the purpose of this paper, we only use the data collected in the 2005-06 academic year for the 1st, 3rd, 5th, 7th and 8th grades.

Results

Table I presents the number of statistical units involved in the analysis and in Table 2 some descriptive statistics of the SES variables are presented, such as the proportion of student eligibility for Free School Meals and Books (FSM), the standard deviation of the proportion across schools (a measure of SES heterogeneity between schools), the standard deviation of parents' education per school (a measure of SES heterogeneity between schools).

([1]) Equalization via common items.

Table I. 1st wave sample composition			
Grade	**Number of:**		
	Students	**Classes**	**Schools**
1	309	35	35
3	327	37	37
5	306	19	9
7	287	18	11
8	248	16	11

Table II. SES Descriptive Statistics

Grade	FSM		Standard Deviation of Parents' Education (school average)
	Proportion P(FSM=Yes)	**SD (Average of FSM school proportion)**	
1	0.19	0.16	0.54
3	0.13	0.12	0.57
5	0.39	0.20	0.49
7	0.31	0.13	0.49
8	0.33	0.17	0.46

By comparing the probability distribution of FSM in primary education with that in elementary and lower secondary education we can observe extremely different values, which are unlikely to be accurate, considering it is the same underlying population in terms of SES distribution. While in primary education the responsibility and management of the student social support fund is attributed to local government (autarchy), in the elementary and higher levels of education that responsibility and management is attributed to each school. Criteria and resources are different in each subsystem of education. Thus the FSM appear to be a SES measure with error, which is usually known as misclassification. Further work and research is needed to adjust for misclassification. Ferrão and Goldstein (2008) show the impact of measurement error in VA estimates.

FIGURE I. Parents' education distribution

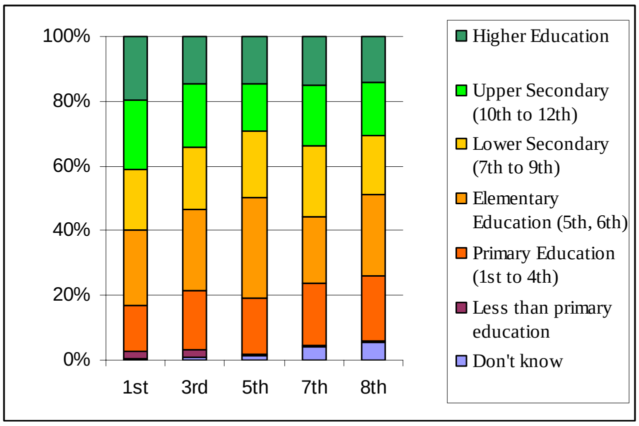


Figure I shows the distribution of students by parents' education. The raising of parents' education is markedly visible at the categories of high school and university. In the 5th year these categories represent about 30% while in the 1st year represent 41%.

Value Added Model: Parameter Estimates

Tables in the Annex A present the parameter estimates for VAM, model specification (1), with different set of controlling variables:

	X2, **student SES**[2]	*X3*, **school SES**
Model 0	---	---
Model 1	Parents' Education	---
Model 2	Parents' Education	Average of Parents' Education
Model 3	FSM	---
Model 4	FSM	Proportion of FSM

The Model 0 fixed parameters are all statistically significant (α=5%) and their estimates show the strong correlation between the response variable (score_2) and prior achievement (score_1).

The proportion of variance explained by model 0 (score_1) is 19%, 26%, 48%, 34% 19%, respectively for each grade. This confirms the relevance of prior achievement in the VAM. Particularly in some grades, prior achievement is moderately correlated with parents' education[3], for example in grade 3 it is -0.33, in grade 5 it is -0.38 and in grade 7 it is -0.34.
Concerning the effect of parents' education on Maths scores, model 1 results show a negative relationship (α=0.1 for 8th grade) with the exception of 7th grade. The coefficient of determination is 52% for 5th grade.

The results for model 2, which include contextual variable for parent's education, do not add any other relevant finding, unless the fixed parameter related to the contextual variable is not statistically significant.

Model 3 parameter estimates that FSM is only statistically significant in grade 5. Once we have already mentioned the unreliability of FSM as an SES variable, we should develop further work on this issue before commenting upon results. The same applies for model 4 in which the results suggest that the contextual variable based on FSM is only statistically significant at grade 1 (α=0.05) and 5 (α=0.1).

Comparison of VA Estimates

Both level 2 residuals (VA estimates) and ranks produced by all models were compared for each grade. Matrices 1 and 2 show the correlation between VA estimates at grade 1 and the correlation between ranks. Among the models that include the SES variable the correlation is smaller (even with fairly high value) in models 1, 2 and 4

MATRIX I. Correlation between VA estimates-grade 1

Mod5	Mod4	Mod3	Mod2	Mod1	Mod0	
Mod5	1.0000					
Mod4	0.9875	1.0000				
Mod3	0.9201	0.9311	1.0000			
Mod2	0.9028	**0.8881**	0.9818	1.0000		
Mod1	0.9032	**0.8875**	0.9795	0.9998	1.0000	
Mod0	0.8821	0.9106	0.9685	0.9843	0.9818	1.000

MATRIX II. Correlation between ranks-grade 1

Mod5	Mod4	Mod3	Mod2	Mod1	
Mod5	1.0000				
Mod4	0.9801	1.0000			
Mod3	0.9157	0.9273	1.0000		
Mod2	0.8940	**0.8856**	0.9742	1.0000	
Mod1	0.8918	**0.8806**	0.9710	0.9985	1.0000
Mod0	0.8879	0.8801	0.9966	0.9742	0.9710

([2]) Parent's Education –inverted scale and standardised.
([3]) Inverted scale.

Figures II, IIa and IIb illustrate the impact of a different model on school VA estimates. The dispersion of VA estimates resulting from model 1 and 4 show a general ‘trend agreement’ between estimates, with the exception of school marked in red. The larger difference between the school position in the rank given by model 1 and its position in the rank given by model 4 is 15 positions (Figure IIa and IIb).

FIGURE II. Dispersion of VA estimates in Grade 1

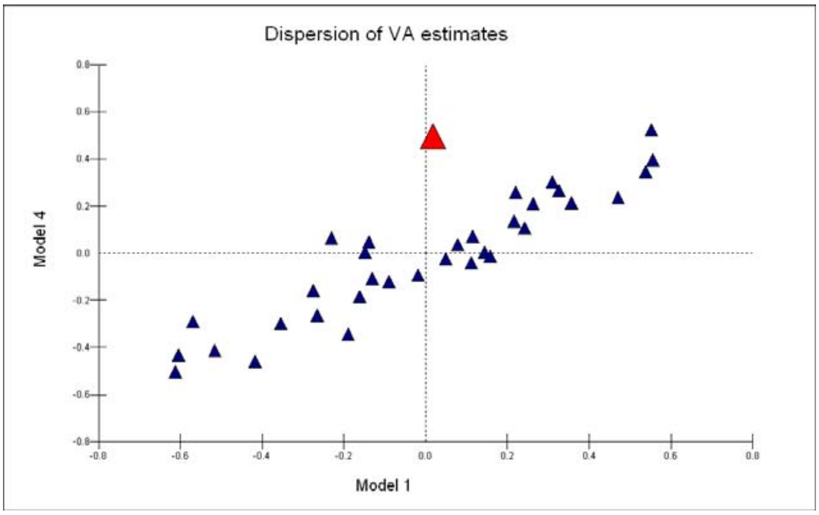


FIGURE IIa and IIb. Confidence Interval (95%) for VA

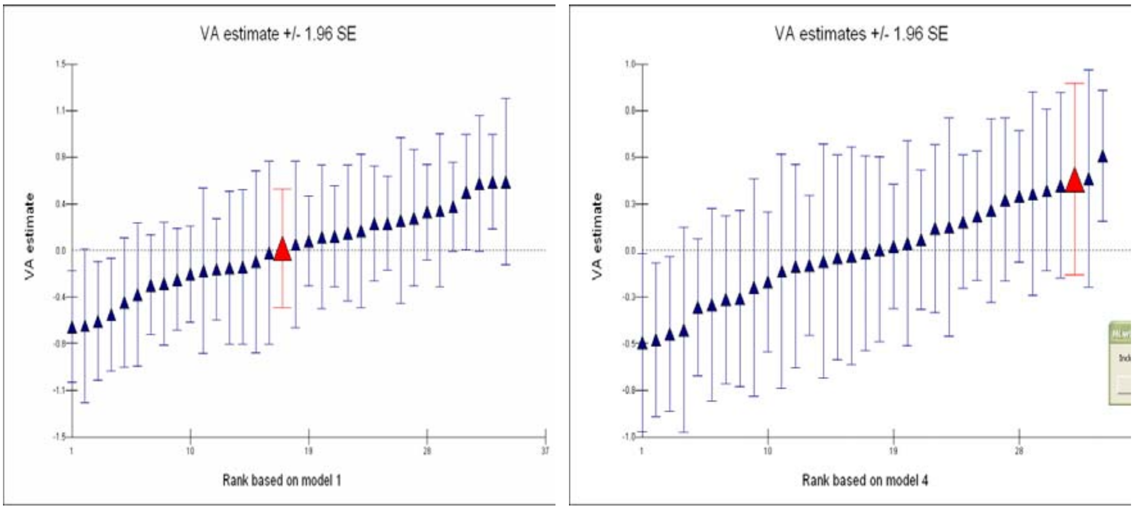
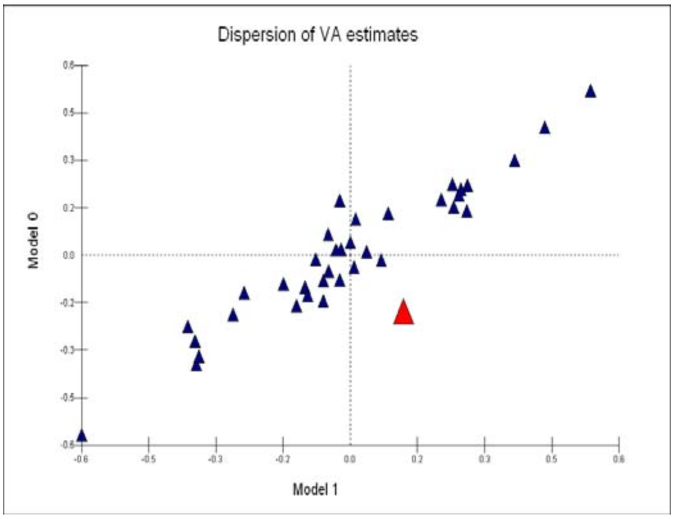


FIGURE III. Dispersion of VA estimates in Grade 3



In grade 3, the correlation between VA estimates produced by different models is larger than 0.93. The smaller value of correlation between ranks is 0.88, which corresponds to a larger difference of 20 positions in the ranks produced by the models.

MATRIX I. Correlation between VA estimates-grade 5

	Mod 3	Mod 2	Mod1	Mod 0
Mod 3	1.0000			
Mod 2	0.6091	1.0000		
Mod 1	0.9980	0.6035	1.0000	
Mod 0	0.9155	0.8032	0.9199	1.0000

In the 5th grade the correlation between VA estimates generated by model 2 and those by model 3 is 0.61 (see Figures IV and V). The correlation in terms of rank position is 0.96.

FIGURE IV. Dispersion of VA estimates in Grade 5

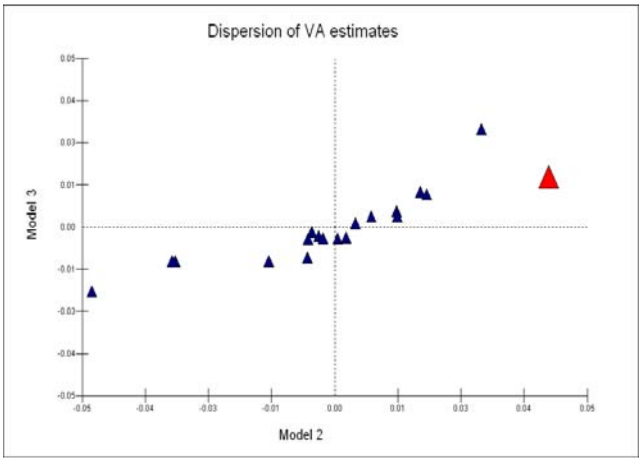
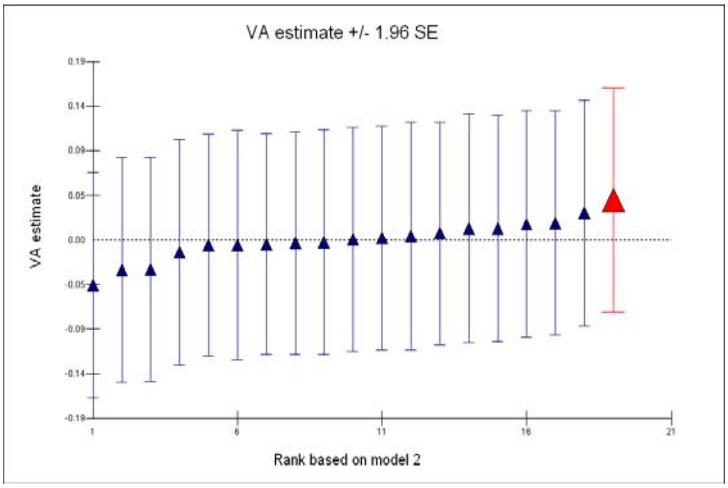


FIGURE V. Confidence Interval $(VA)_{95\%}$. Model 2



Model 1 suggests that, in grade 5, after controlling for prior achievement and parent's education, VA is not statistically different from zero.

In grade 7, the correlation between VA estimates produced by models 0 to 4 is larger than 0.96, and the correlation between the respective ranks is larger than 0.94. Prior achievement is the strongest predictor. In general, this evidence holds true for grade 8, and can be seen in figures 7 and 8, which illustrates the comparison between VA estimates based on model 0 (prior achievement as controlling variable) and model 1 (Prior achievement and parent's education as controlling variables). It is important to take into account that the variance partition coefficient (VPC) is quite low in elementary and lower secondary education (see Table III).

FIGURE VI. Confidence Interval $(VA)_{95\%}$. Model 1

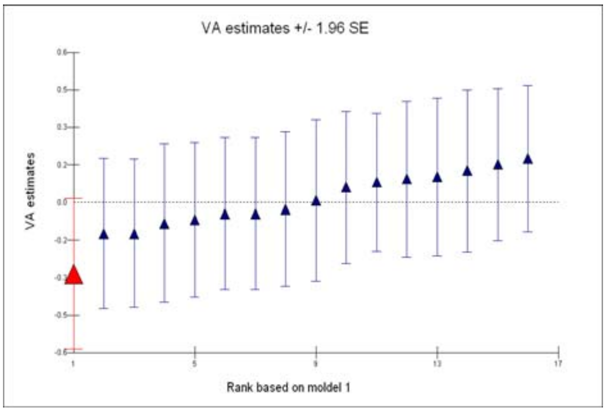


FIGURE VII. Dispersion of VA estimates in Grade 8

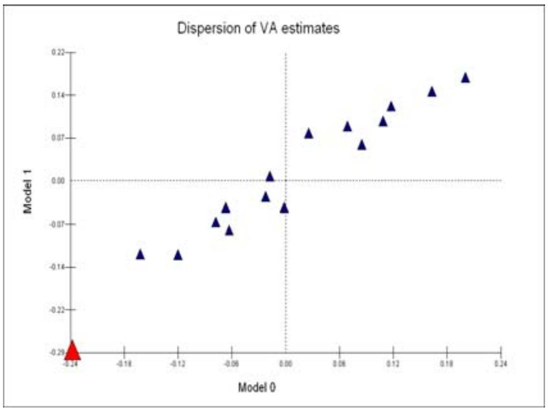


TABLE III. Null model estimates

	Grade				
	1st	3rd	5th	7th	8th
Intercept	-0.019 (0.092)	-0.122 (0.088)	-0.005 (0.086)	-0.038 (0.091)	-0.003 (0.100)
Random parameters					
σ_u^2	0.155 (0.067)	0.164 (0.066)	0.077 (0.045)	0.080 (0.048)	0.096 (0.057)
σ_e^2	0.848 (0.073)	0.857 (0.072)	0.918 (0.079)	0.911 (0.080)	0.897 (0.086)
VPC	0.15	0.16	0.08	0.08	0.10

Discussion

There are two main findings in this paper. The first is that the SES fixed parameter is statistically significant at all grades with the exception of grade 7. The second is that the impact of model choice (read as different set of controlling variables) on VA estimates is particularly important in primary education. Mainly in elementary and lower secondary, the results of the models seem to confirm Ballou et al. (2004) about the relationship between SES and prior achievement. The evidence presented above seems to suggest that along the school trajectory prior achievement encapsulates the effect of SES. More work needs to be done on this, but if it is true, this constitutes another challenge for further research about equity (to look for schools that actually ‘compensate’ for the SES disadvantage). The statistical model, which we specified in the background paper for estimating the school effect on equity, may no longer be adequate for investigating plausible hypothesis revealed in the data.

One of the most important limitations of the work presented above is the validity and reliability of SES variables, particularly FSM. Both longer time periods of observation and population data are also highly relevant characteristics in order to know the *true* school value added. Research work is being done in order to complement the analysis presented with the longitudinal data collected in 2006-07 and 2007-08.

There are other concerns related to the issue that constitute topics for further research. For example, Lockwood et al. (2007) use longitudinal data from a cohort of middle school students from a school district and compare several VAMs (teacher effects). They found that the variation within teachers across achievement measures is larger than the variation across teachers. These results suggest that VAMs are sensitive to the ways in which student achievement is measured, even when only a particular subject. In the special issue of JEBS on

Value-Added Models, Reckase (2004) also points out the importance of assessment: 'The sophisticated statistical procedures described in these articles may be giving a glossy finish to misleading assessment results. Before putting a lot of confidence in the results of these analyses, the functioning of the assessments needs to be investigated in great detail'.

References

BALLOU, D., SANDERS, W. & WRIGHT, P. (2004). Controlling for Student Background in Value-Added Assessment of Teachers. *Journal of Educational and Behavioral Statistics*, 29, 37-66.

BRAUN, H. (2005). *Using student progress to evaluate teachers: a primer on value-added models.* Policy Information Perspective, Princeton, NJ: ETS.

BRAUN, H. & WAINER, H. (2007). Value-Added Modeling. In C.R. RAO & S. SINHARAY, S. (Eds.) *Handbook of statistics 26, Psychometrics* (475-501). Amsterdam: Elsevier.

COLEMAN, J. (1988). Social capital in the generation of human capital. *American Journal of Sociology*, *94*, 95-120.

FERRÃO, M. E. (2007). *Value-Added Models in Portugal*. Background paper prepared for the 1st meeting VAM project.

FERRÃO, M. E. & GOLDSTEIN, H. (in press). Adjusting for Measurement Error in the Value Added Model: evidence from Portugal. *Quality & Quantity*.

GRAY, J., JESSON, D., GOLDSTEIN, H. HEDGER, K. & RASBASH J. (1995). A multilevel analysis of school improvement: changes in schools' performances over time. *School Effectiveness and School Improvement*, 6, 97-114.

LADD, H. F. & WALSH, R. P. (2002). Implementing value-added measures of school effectiveness: Getting the incentives right. *Economics of Education Review*, 21, 1-17.

LITTLE, R. J. A. & RUBIN, D. B. (2002). *Statistical analysis with missing data*, (2nd edition). Wiley: New York.

LOCKWOOD, J. R., MCCAFFREY D. F., HAMILTON, L. S., STECHER, B. M., LE, V. & MARTINEZ, F. (2007). The sensitivity of value-added teacher effect estimates to different mathematics achievement measures. *Journal of Educational Measurement*, 44, 47-67.

MCCAFFREY, D., LOCKWOOD, J. R., KORETZ, D. M. & HAMILTON, L. S. (2003). *Evaluating Value-Added Models for Teacher Accountability*. Santa Mónica, CA: RAND Corporation.

RUBIN, D. B., STUART, E. A. & ZANUTTO, E. L. (2004). A potential outcomes view of value-added assessment in Education. *Journal of Educational and Behavioral Statistics*, 29, 103-116.

SAMMONS, P., THOMAS, S. & MORTIMORE, P. (1997). *Forging links: effective schools and effective departments*. London: Chapman.

THOMAS, S. & MORTIMORE, P. (1996). Comparison of value-added models for secondary school effectiveness. *Research Papers in Education*, *11*, 5-33.

VICENTE, P. (2006). *O plano amostral do projecto 3EM*. En M. E. FERRÃO, C. NUNES Y C. BRAUMANN (eds.), Proceedings of the XIV Annual Conference of the Portuguese Statistical Society. Lisboa: SPE.

Contact: Mª Eugenia Ferrão. Universidade da Beira Interior. Departamento de Matemática. Rua Marqués d'Avila e Bolama. 6200-001 Covilhã, Portugal. E-mail: mferrao@ubi.pt

Annex A
Parameter Estimates

Model 0	Grade 1		Grade 3		Grade 5		Grade 7		Grade 8	
Fixed	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
cons	-0.015	0.090	-0.075	0.071	-0.022	0.048	-0.022	0.074	0.000	0.075
score_1	0.476	0.050	0.502	0.049	0.683	0.043	0.587	0.050	0.410	0.060
Random										
Level 2 var	0.172	0.065	0.089	0.041	0.011	0.014	0.052	0.032	0.035	0.032
Level 1 var	0.637	0.055	0.654	0.055	0.513	0.044	0.606	0.054	0.773	0.074
-2**log likelihood*	745.828		788.950		636.784		650.076		606.510	

Model 1	Grade 1		Grade 3		Grade 5		Grade 7		Grade 8	
Fixed	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
cons	-0.014	0.093	-0.046	0.076	-0.039	0.046	-0.042	0.077	-0.008	0.078
score_1	0.468	0.051	0.443	0.052	0.656	0.047	0.599	0.056	0.374	0.063
par_edu_stand	-0.113	0.054	-0.209	0.054	-0.117	0.047	-0.037	0.057	-0.113	0.065
Random										
Level 2 var	0.179	0.068	0.112	0.048	0.005	0.013	0.054	0.034	0.037	0.035
Level 1 var	0.636	0.057	0.632	0.055	0.479	0.043	0.642	0.059	0.770	0.079
-2**log likelihood*	701.896		735.763		553.206		614.581		543.728	

Model 2	Grade 1		Grade 3		Grade 5		Grade 7		Grade 8	
Fixed	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
cons	-0.017	0.096	-0.054	0.077	-0.039	0.046	-0.046	0.077	-0.006	0.078
score_1	0.468	0.052	0.447	0.053	0.653	0.047	0.600	0.056	0.370	0.064
par_edu	-0.116	0.057	-0.225	0.058	-0.106	0.050	-0.046	0.060	-0.104	0.070
par_edu_sch	0.023	0.167	0.101	0.140	-0.060	0.108	0.079	0.172	-0.069	0.189
Random										
Level 2 var	0.179	0.068	0.112	0.048	0.004	0.013	0.053	0.034	0.036	0.034
Level 1 var	0.636	0.057	0.632	0.055	0.479	0.043	0.642	0.059	0.771	0.079
-2**log likelihood*	701.876		735.254		552.896		641.371		543.594	

Model 3	Grade 1		Grade 3		Grade 5		Grade 7		Grade 8	
Fixed	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**	**Estimate**	**SE**
cons	0.014	0.091	-0.070	0.074	0.071	0.058	-0.009	0.083	0.011	0.086
score_1	0.466	0.051	0.498	0.050	0.666	0.045	0.585	0.051	0.423	0.061
FSM	-0.172	0.131	-0.058	0.144	-0.235	0.092	-0.043	0.109	-0.027	0.131
Random										
Level 2 var	0.159	0.062	0.096	0.043	0.002	0.013	0.056	0.033	0.032	0.032
Level 1 var	0.637	0.055	0.657	0.055	0.513	0.046	0.608	0.054	0.788	0.077
-2**log likelihood*	744.142		788.816		587.235		639.959		589.532	

Model 4	Grade 1		Grade 3		Grade 5		Grade 7		Grade 8	
Fixed	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE	Estimate	SE
cons	0.206	0.121	-0.076	0.101	0.206	0.095	0.180	0.204	0.069	0.159
score_1	0.467	0.050	0.498	0.050	0.653	0.045	0.586	0.051	0.420	0.062
FSM	-0.101	0.135	-0.061	0.149	-0.172	0.099	-0.024	0.110	-0.009	0.138
Prop_FSM	-1.117	0.503	0.039	0.561	-0.413	0.236	-0.616	0.613	-0.193	0.444
Random										
Level 2 var	0.130	0.054	0.100	0.044	0.000	0.000	0.052	0.032	0.031	0.027
Level 1 var	0.635	0.055	0.658	0.055	0.510	0.044	0.608	0.054	0.789	0.077
-2**log likelihood*	739.456		788.856		584.232		638.969		589.344	