

Grau en Estadística

Títol: Anàlisi dels factors determinants en el rendiment de matemàtiques de l'alumnat de 15 anys a Espanya (PISA 2018)

Autora: Patrícia Casado i Vila

Directora: Mireia Fernández-Ardèvol

Departament: Econometria, Estadística i Economia Aplicada (UB)

Convocatòria: Juny 2020



Resum

Aquest treball analitza els factors determinants en el rendiment de matemàtiques de l'alumnat de 15 anys a Espanya. S'han utilitzat les microdades proporcionades per l'estudi PISA (Programa Internacional per a l'avaluació d'estudiants) dut a terme per l'OCDE, recollides l'any 2018 en la seva darrera aplicació. Donada l'estructura jeràrquica de les dades, la metodologia emprada és una regressió multinivell que engloba la informació de l'estudiantat i de les escoles. Els resultats obtinguts mostren que el rendiment en matemàtiques de l'estudiantat de 15 anys no es pot explicar exclusivament en base als factors relatius a les escoles. És essencial tenir en compte la component personal, el context familiar i les actituds tant de l'alumnat com del professorat.

Paraules clau: rendiment educatiu, resultats matemàtiques, PISA 2018, Espanya, anàlisi multinivell, R.

Analysis of the determining factors in the performance of mathematics of 15-year-old students in Spain (PISA 2018)

Abstract

This paper analyses the determining factors in the performance of mathematics of 15-year-old students in Spain. The microdata used is provided by the PISA (International Program for Student Assessment) study conducted by the OECD, collected in 2018 in its latest application. Given the hierarchical structure of the data, the methodology used is a multilevel regression that encompasses information from students and schools. The results show that the mathematical performance of 15-year-olds cannot be explained solely based on school-related factors. It is essential to consider the personal component, the family context and the attitudes of both students and teachers.

Key words: education performance, mathematical results, PISA 2018, Spain, multilevel analysis, R.

Classificació AMS (*American Mathematical Society*)

62P25 - Applications to social sciences

Índex

1. INTRODUCCIÓ.....	- 2 -
2. METODOLOGIA.....	- 4 -
2.1. MODEL MULTINIVELL.....	- 4 -
2.1.1. Criteris d'informació.....	- 5 -
2.1.2. Software.....	- 6 -
2.2. ANÀLISI DE CORRESPONDÈNCIES MÚLTIPLES.....	- 6 -
3. SELECCIÓ DE LES VARIABLES.....	- 7 -
3.1. ESTUDIANTAT.....	- 8 -
3.2. ESCOLES.....	- 10 -
4. PREPARACIÓ DE LES DADES.....	- 14 -
4.1. DADES MANCANTS.....	- 14 -
4.2. RECODIFICACIÓ DE LES VARIABLES.....	- 14 -
4.3. CÀLCUL DE NOVES VARIABLES.....	- 15 -
5. RESULTATS.....	- 18 -
5.1. ESPECIFICACIÓ DEL MODEL JERÀRQUIC.....	- 18 -
5.2. MODEL DEFINITIU.....	- 22 -
6. CONCLUSIONS.....	- 26 -
BIBLIOGRAFIA.....	- 29 -
ANNEXES.....	- 31 -
ANNEX 1: ANÀLISI DESCRIPTIVA DE LES VARIABLES.....	- 31 -
ANNEX 2: ANÀLISI DE CORRESPONDÈNCIES MÚLTIPLES.....	- 39 -
ANNEX 3: MODELS.....	- 40 -
ANNEX 4: CODI R.....	- 47 -

1. Introducció

L'ensenyament de les matemàtiques ha sigut i segueix sent un desafiament per a totes les escoles, sobre tot, a l'educació secundària. És preocupant el rebuig i la desmotivació que provoca aquesta disciplina a una gran part de l'estudiantat i no existeix una explicació única. Aquest estudi sorgeix d'una vocació professional encarada a la docència i una inquietud personal per transformar l'assignatura de matemàtiques en un aprenentatge atractiu.

Quins factors relacionats amb les escoles poden influir en els resultats de matemàtiques de l'alumnat de secundària? L'objectiu principal d'aquest treball és identificar aquests factors, fent servir les dades proporcionades per l'estudi PISA (Programa Internacional per a l'avaluació d'estudiants), dissenyat i dut a terme per l'Organització per la Cooperació i el Desenvolupament Econòmic (OCDE). Es tracta de la principal font d'informació sobre el rendiment de l'estudiantat en àrees temàtiques clau i es realitza cada tres anys. En concret, l'àmbit geogràfic d'aquest treball és Espanya i s'analitzen les dades de la seva darrera aplicació, recollides l'any 2018.¹

S'han utilitzat les microdades proporcionades per l'OCDE. En concret, s'ha disposat de dues bases de dades, una amb variables referents a l'estudiantat i l'altra amb variables referents a les escoles. La mostra seleccionada consta de 35.943 alumnes i de 1.089 escoles. Les variables seleccionades són de diverses tipologies: les respostes extretes directament dels qüestionaris i els índexs elaborats per PISA.

Les dades relacionades amb l'educació acostumen a tenir una estructura jeràrquica on l'estudiantat forma part d'un nivell superior que, en aquest cas, són els centres educatius. Tenint en compte que les característiques de l'alumnat poden ser similars dins d'una mateixa escola però diferents entre elles, la utilització de tècniques d'anàlisi tradicionals pot conduir a resultats esbiaixats. Per representar aquesta variabilitat la metodologia més adient és la regressió multinivell que té en compte el caràcter niat de les dades calculant una equació diferent per a cada nivell.

Partint del model nul, la base de comparació de models més complexos, s'ha seguit una anàlisi sistemàtica per aconseguir implementar un model definitiu. Aquesta modelització ha permès extreure conclusions al voltant de les hipòtesis inicials plantejades. Els resultats obtinguts mostren que el rendiment en matemàtiques de l'estudiantat de 15 anys no es pot explicar exclusivament en base als factors relatius a les escoles. És essencial tenir en compte la component personal, el context familiar i les actituds tant de l'alumnat com del professorat.

¹ www.OECD.org/pisa/data

Agraïment especial a qui ha entès i ha fet que jo entengui la tutoria com una tasca transcendent, aportadora de valors i carregada de responsabilitat. Moltes gràcies Mireia Fernández-Ardèvol.

2. Metodologia

L'anàlisi empírica d'aquest estudi es basa en l'aplicació de tècniques de regressió multinivell ja que les microdades disponibles tenen una estructura jeràrquica (OCDE, 2009). Les dades relacionades amb l'educació acostumen a tenir un disseny niat, l'estudiantat forma part d'un nivell superior que en aquest cas son els centres educatius. Quan les característiques de l'alumnat son similars dins d'una mateixa escola però diferents entre elles, la utilització de tècniques d'anàlisi tradicionals pot conduir a resultats esbiaixats. La regressió multinivell, per tant, té en compte aquest caràcter niat de les dades a dins d'unitats més amplies calculant una equació diferent per a cada nivell. Aquesta tècnica, que no s'ha estudiat durant el grau, es desenvolupa a l'apartat 2.1.

A més, a l'hora d'analitzar el conjunt de preguntes relatives a l'ús de les tecnologies digitals als centres educatius, es disposa de la pregunta SC155 amb onze ítems (detallada a l'apartat Escolles3.2). Tractant-se de respostes tancades múltiples, s'ha realitzat una anàlisi per reduir-ne la dimensionalitat. Degut al caràcter qualitatiu de les variables, s'ha implementat una Anàlisi de Correspondències Múltiples (ACM) al conjunt d'ítems (apartat 2.2).

2.1. Model multinivell

El model multinivell (també anomenat models lineals jeràrquics) és un model estadístic paramètric que varia en més d'un nivell i es tracta d'una metodologia d'anàlisi per dades amb patrons de variabilitat complexos. Degut a la presència d'efectes fixes i aleatoris, els models multinivell també s'anomenen models d'efectes mixtes.

Aquests models tenen en compte les estructures jeràrquiques de les dades reconeixent les components residuals per a cada nivell de la jerarquia. En el nostre cas, un model a dos nivells que permet agrupar l'alumnat en escoles, inclourà residus tant per l'estudiantat com pels centres educatius (Rasbash, 2019). D'aquesta manera la variància residual està composta per la component entre escoles (variància dels residus pel nivell de les escoles) i dintre de les escoles (variància dels residus pel nivell de l'estudiantat). Els residus de les escoles, també anomenats "efectes", representen les característiques no observades d'aquest nivell que afecten al rendiment de l'alumnat.

Amb aquests models es pot conèixer no només el valor mig dels efectes de les variables explicatives sobre la variable endògena sinó que també s'especifica la variació d'aquests efectes en el mateix nivell. Per altra banda, amb aquesta modelització es poden expressar les relacions que es produeixen entre variables d'un mateix nivell i com influeixen les variables d'un nivell a l'altre.

Per realitzar l'anàlisi d'aquest estudi, s'utilitza la formulació a dos nivells proposada per Raudenbush i Bryk (2002), on el primer nivell correspon a les dades de l'estudiantat i el segon nivell representa la influència dels factors de les escoles. Seguint la metodologia de l'estudi de Guio-Jaimes (2014) i disposant de grups com a entitats úniques (Snijders, 2011), s'ha emprat un model amb l'intercepte aleatori i els pendents dels efectes fixes.

A les següents equacions (Guio-Jaimes, 2014), X_{kij} representa les covariables del nivell de l'estudiantat i Z_{lj} les covariables de les escoles, el segon nivell. β_{0j} i β_{1j} indiquen els coeficients de la regressió. u_{0j} representa els interceptes aleatoris definits per a cada escola j , incorrelacionats entre les escoles i amb les covariables. ε_{ij} son els residus del nivell de l'estudiantat, incorrelacionats entre l'alumnat i les escoles, incorrelacionats amb u_{0j} i amb les covariables.

Model del nivell 1:

$$Y_{ij} = \beta_{0j} + \sum_{k=1}^n \beta_{1j} X_{kij} + \varepsilon_{ij}$$

Model del nivell 2:

$$\beta_{0j} = \gamma_{00} + \sum_1 \gamma_{01} Z_{lj} + u_{0j}$$

$$\beta_{1j} = \gamma_{10}$$

Amb $u_{0j} \sim N(0; \tau_{00})$

Model complet:

$$Y_{ij} = \gamma_{00} + \gamma_{10} X_{kij} + \gamma_{01} Z_{lj} + u_{0j} X_{kij} + u_{0j} + \varepsilon_{ij}$$

2.1.1. Criteris d'informació

L'estimació d'un model jeràrquic exigeix una anàlisi sistemàtica partint del model més simple possible, el model nul. Aquest model no inclou predictors en cap dels nivells, només estima la mitjana global del rendiment i la variància de cada un dels nivells (Raudenbush, 2002). El model nul és la base de comparació de models més complexos que s'acceptaran o es rebutjaran si ajusten de manera significativa millor que el nul. Per dur a terme aquestes comparacions s'ha emprat la raó de verosimilitud de cada model, anomenat *Deviance*. El criteri d'informació de la desviància (DIC) és una generalització jeràrquica del criteri d'informació d'Akaike (AIC) i del criteri d'informació Bayesià (BIC) que també s'han tingut en consideració i la idea principal és que son preferibles els models amb menor DIC.

2.1.2. Software

Totes les anàlisis s'han dut a terme amb el paquet estadístic R (Team, 2018). L'anàlisi de les dades s'ha realitzat amb la funció *lmer* que ajusta un model lineal d'efectes mixtes (LMM) amb el criteri de màxima versemblança. S'ha de tenir en compte que els models amb efectes aleatoris no disposen de les distribucions asimptòtiques clàssiques i que hi ha un extens debat sobre quines eines estadístiques son apropiades per avaluar aquests models i per utilitzar com a inferència. Una de les dificultats trobades, per tant, ha sigut la visualització de la significació dels paràmetres del model. S'ha emprat la metodologia proposada per Kuznetsova (2017) que especifica instal·lar el paquet *lmerTest-package* a R (Christensen, 2017). Aquest paquet proporciona p-valors al test *anova* i a les taules resum (*summary*) amb el mètode aproximatiu de Satterthwaite (Luke, 2017) dels graus de llibertat. Per últim, la modelització emprada descarta les observacions que tenen dades mancants als predictors així que la mida de la mostra canvia segons el model analitzat.

2.2. Anàlisi de correspondències múltiples

L'anàlisi de correspondències múltiples (ACM) és una tècnica per detectar i representar patrons de relacions entre variables categòriques dependents. Aquest estudi es pot interpretar com una extensió de l'Anàlisi de Correspondències (AC) quan es disposa de més de dues variables categòriques o de l'Anàlisi de Components Principals (ACP) quan es tenen variables categòriques en canvi de quantitatives (Abdi, 2010). L'objectiu d'aquests mètodes és identificar grups d'individus amb perfils similars i establir possibles associacions entre els atributs establerts.

S'ha emprat el paquet *FactoMineR* (Sebastien Le, 2008) que proposa diversos mètodes exploratoris d'anàlisi multivariant (Husson F. L., 2010). Un cop implementada la tècnica d'Anàlisi de Correspondències Múltiples, s'ha realitzat un algorisme d'agrupament jeràrquic (en anglès *hierarchical clustering*) per dividir la població, en aquest cas els centres educatius, en diferents grups. Aquesta classificació es construeix de manera que les observacions d'un mateix grup tenen respostes similars als ítems en qüestió. Es tracta d'un mètode de mineria de dades d'aprenentatge no supervisat on el model s'ajusta a les observacions i no consta d'un coneixement previ. En aquest cas, el paquet necessari per realitzar el clustering és *flashClust* (Peter Langfelder, 2012).

Per establir el nombre necessari de grups s'han computat les diferències (distàncies) entre les observacions del data set. Com no es poden aplicar les distàncies euclidianes ja que les dades no son numèriques, s'ha emprat la fórmula de Gower amb la funció d'R "*daisy*". Amb la confirmació de l'eina gràfica del dendrograma, s'han establert tres grups com a classificació de la variable.

3. Selecció de les variables

El Programa Internacional per a l'Avaluació d'Estudiants (PISA) es centra, a la seva darrera aplicació de 2018 (OCDE, 2020), en l'avaluació de la competència lectora en l'entorn digital. Tot i no ser el centre d'estudi, a l'informe també es recullen dades de les altres dues competències: matemàtiques i ciències. A l'hora de seleccionar les variables, per tant, no es tenen en compte les que fan referència a la comprensió lectora si no que es consideren les variables relatives a les característiques dels centres educatius i del respectiu alumnat.

A les bases de dades seleccionades hi consten diverses tipologies de variables. Per una banda es troben les respostes als qüestionaris que van realitzar tant la direcció de les escoles com l'estudiantat. En el qüestionari dels centres educatius es demana informació de caràcter general sobre el centre, la gestió escolar o el personal docent, entre d'altres. El nom d'aquestes variables es caracteritzen pel número de la pregunta amb el prefix "SC" (les inicials de la paraula escola en anglès) seguit dels possibles subapartats de la pregunta. Respecte a l'alumnat, les respostes dels qüestionaris proporcionen informació personal, de l'entorn familiar i de percepcions respecte al seu centre. En aquest cas, el prefix és "ST" (inicials d'estudiant/a en anglès).

Per una altra banda, PISA proporciona índexs que s'elaboren amb valors estandarditzats per descriure les característiques de l'alumnat i dels centres en base a les preguntes dels qüestionaris. S'utilitza la Teoria de Resposta a l'Ítem (IRT) per comprovar el comportament esperat d'aquests índexs i validar la comparabilitat entre països. Una descripció més detallada dels mètodes es troba a l'OCDE (forthcoming). Aquests índexs es construeixen transformant aritmèticament, recodificant o escalant múltiples ítems extrets de les preguntes dels qüestionaris. Els valors es transformen de manera que, en el conjunt de països de la OCDE, tenen mitjana zero i desviació estàndard igual a u. Per tant, la interpretació no es pot efectuar individualment, si no que s'ha de comparar amb la mitjana de l'OCDE.

Respecte a les dades de l'alumnat, s'ha de fer referència a altres dues tipologies de variables que es tenen en compte a l'estudi, els pesos i els valors plausibles. Tot i que l'alumnat inclòs a la mostra final de PISA es va escollir de manera aleatòria, les probabilitats de selecció de cada estudiant varien. Els pesos han de ser incorporats a l'anàlisi per assegurar que cada participant representa aproximadament el nombre correcte d'estudiants de la població. Els pesos, a més de reflectir les probabilitats desiguals de pertànyer a la mostra, tenen en compte possibles ajustaments deguts a una falta de resposta. La variable utilitzada per l'anàlisi multinivell, és la que s'especifica a l'annex A8 de l'OCDE (2007), "W_FSTUWT", i correspon al pes final per tot l'estudiantat que ha realitzat el test.

Per últim, al nivell de l'alumnat es disposa de deu valors plausibles que corresponen al rendiment de matemàtiques (PV1MATH, ..., PV10MATH). Aquests valors són la **variable endògena** del model. Amb la Teoria de Resposta a l'Ítem (TRI) s'obtenen estimacions de les puntuacions de l'alumnat en una mateixa escala que els fa comparables, independentment del subconjunt d'ítems que es proporioni a cada alumne/a. PISA computa les distribucions a posteriori de tal manera que les puntuacions obtingudes tenen una mitjana de 500 punts i, d'aquesta manera, dues tercers parts de tot l'alumnat avaluat té un resultat comprés entre els 400 i els 600 punts.

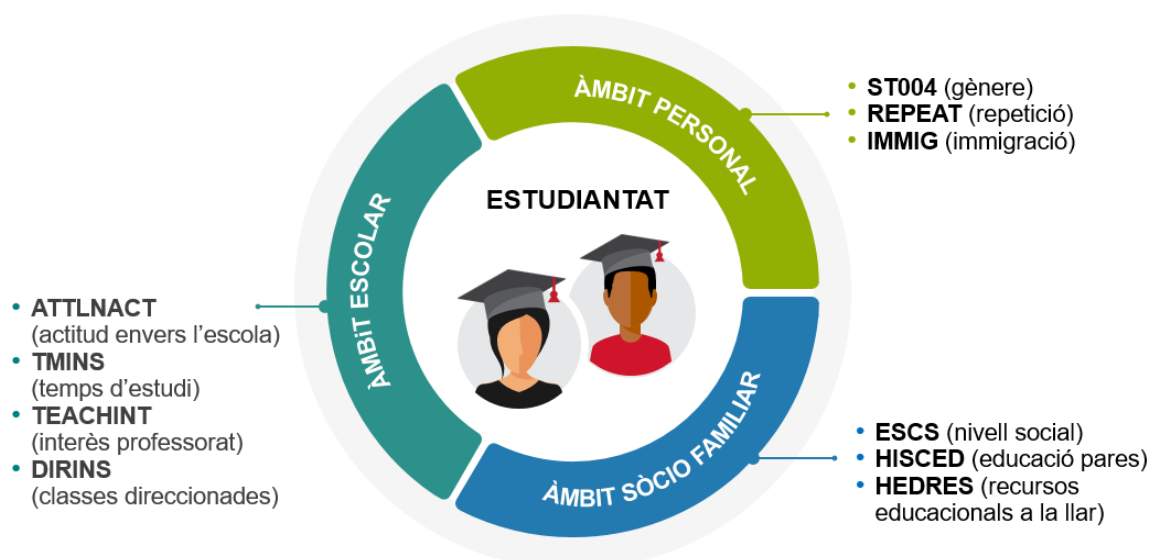
Els valors plausibles no es poden agregar al nivell de l'estudiantat, s'han de realitzar els estadístics amb els deu valors independentment i a posteriori tenir en compte el terme mitjà. Segons estableix l'OCDE (2009, pág. 44) si es tracta d'una base de dades de grans dimensions, no representa una gran diferència basar l'estudi en un o en els deu valors plausibles. Tot i així, es recomana utilitzar un valor en el cas de l'anàlisi exploratori i reportar els resultats tenint en compte els deu valors plausibles disponibles.

Els apartats 3.1 i 3.2 descriuen les **variables explicatives** del rendiment en matemàtiques de l'alumnat.

3.1. Estudiantat

Tenint com a objectiu d'estudi els factors relatius a les escoles, en el nivell que fa referència a l'estudiantat, s'han escollit variables genèriques de l'alumnat i de la seva relació amb l'escola. A la Il·lustració es mostren les tres categories en les quals s'han agrupat les variables d'aquest nivell.

Il·lustració 3.1 Agrupacions de les variables del nivell de l'estudiantat



a) Àmbit personal

En un primer grup es troben tres variables que descriuen la situació individual de l'alumnat. En primer lloc s'ha considerat el gènere, que s'ha extret de la pregunta ST004 on es proporcionen dues opcions de resposta: dona (0) i home (1).

L'edat no s'ha tingut en consideració perquè es tracta d'una variable molt homogènia (mitjana: 15.84; DS: 0.29). En canvi, com a variable que aproxima millor l'edat s'ha considerat el grau de repetició de l'estudiantat REPEAT que es computa amb la pregunta ST127 i pren dos valors: 0 quan l'alumnat no ha repetit cap curs, 1 en cas contrari (al menys ha repetit un curs).

I en tercer lloc, es disposa de l'índex de l'historial d'immigració de l'alumnat, IMMIG, que es calcula en base a la pregunta ST019 i pren tres possibles valors: (1) estudiantat nadiu (tenen al menys un genitor nascut al país); (2) estudiantat de segona generació (nascuda al país però els pares no); i (3) estudiantat de primera generació (ni ella ni els seus genitors son nascuts al país d'anàlisi).

b) Àmbit socio-familiar

Als índexs proporcionats per PISA 2018 hi ha certes variables que fan referència a les característiques socioeconòmiques de l'alumnat i de les seves famílies. L'estat econòmic, social i cultural (ESCS) de l'estudiantat es deriva de tres variables relacionades amb l'entorn familiar (pares i mares): el màxim nivell d'educació calculat com a índex dels anys d'escolarització (PARED), l'estat ocupacional calculat com un índex de l'estatus màxim ocupacional dels genitors (HISEI) i les possessions a casa (HOMEPOS).

A més, s'ha seleccionat l'índex que indica el màxim nivell educacional dels dos genitors HISCED en base a la classificació ISCED (1997) on hi ha sis categories. La variable ha sigut recodificada de tal manera que pren valor 0 si el nivell és inferior a l'educació terciària i 1 si és superior. Finalment s'ha seleccionat la variable HEDRES, que es basa en els ítems del qüestionari que mesuren l'existència de recursos educacionals a casa incloent un escriptori, un lloc tranquil per estudiar, llibres que ajudin a l'alumnat a estudiar o un ordinador per us escolar. És un índex de caràcter continu on els valors positius indiquen que els recursos educacionals disponibles a casa són superiors a la mitjana de l'OCDE.

c) Àmbit escolar

PISA proporciona algunes preguntes dedicades a valorar la motivació de l'alumnat i la seva actitud envers el centre escolar. L'índex ATTLNACT representa l'opinió de l'alumnat envers la idea d'escola en general. Es construeix en base a la pregunta ST036 on es demana si estan d'acord amb el fet que esforçar-se a l'escola és important, si els ajudarà a entrar a la

universitat o a obtenir una feina. Els valors positius en aquest índex, per tant, indiquen que l'estudiantat valora l'escola de manera més positiva que la mitjana dels països de l'OCDE.

En aquest apartat també es té en compte la variable **TMINS** que indica el temps d'aprenentatge proporcionat pel centre educatiu. Es computa multiplicant la duració de les classes pel nombre de classes setmanals.

A més, s'han seleccionat dues variables declaratives que s'han de valorar tenint en compte que son la percepció de l'estudiantat. Per una banda, **TEACHINT** especifica l'interès del professorat que es deriva de la pregunta **ST213**. Valors positius impliquen que l'alumnat percep el professorat més entusiasmada del que ho percep la resta de l'estudiantat de l'OCDE en mitjana. I per una altra banda, l'índex **DIRINS** estableix quant de direccionades son les classes, és a dir, si el mètode docent és més tradicional on el professorat és el centre de la classe. En aquest cas, valors positius estableixen que l'alumnat percep que el professorat realitza una classe dirigida amb més freqüència respecte a la mitjana de la OCDE.

3.2. Escoles

Els factors relatius als centres educatius són el focus de l'estudi i, per tant, s'han considerat tant variables directament extretes dels qüestionaris com índexs construïts. Les variables escollides s'han agrupat en quatre categories (com s'especifica a la Il·lustració) per tal de poder definir les escoles.

Il·lustració 3.2 Agrupacions de les variables del nivell de l'estudiantat



a) Tipologia d'escola

En un primer grup es troben les variables que permeten descriure l'escola des d'un punt de vista objectiu i amb les que es poden categoritzar les seves dimensions. La variable SCHLTYPE classifica les escoles en privades, concertades o públiques en base a qui té el poder de prendre les decisions i els valors corresponen a les respostes de la pregunta SC013 del qüestionari. L'índex consta de tres categories: (1) escoles privades independents del govern; (2) escoles privades dependents del govern (concertades); i (3) escoles públiques.

A més, s'ha seleccionat una variable que indica el nombre d'alumnat total del centre educatiu (SCHSIZE) per establir la grandària aproximada de l'escola. Per últim, a la pregunta SC001 del qüestionari es demana la mida del municipi on està situat el centre i es dona a escollir entre cinc opcions: (1) un poblet, llogaret o àrea rural (menys de 3.000 habitants); (2) un poble (entre 3.000 i 15.000 habitants aproximadament); (3) una ciutat petita (de 15.000 a 100.000 persones); (4) una ciutat mitjana (de 100.000 a 1.000.000 de persones); i (5) una gran ciutat (de més d'1.000.000 de persones).

b) Estructura de les classes

Un cop definida l'estructura del centre, es pretén comprovar la manera de gestionar les classes. Per fer una anàlisi categòrica, s'ha seleccionat la ràtio d'alumnes per classe (CLSIZE). Originalment, l'índex podia prendre nou possibles valors però s'ha reclassificat en tres categories: (1) classe petita (fins a 20 alumnes per classe); (2) classe mitjana (entre 21 i 30); i (3) classe gran (més de 30).

A la pregunta SC042, composta per dos ítems, es pregunta a l'equip directiu de quina manera s'agrupa a l'estudiantat, en cas de fer-ho. S'han seleccionat les dues variables corresponents als ítems que estableixen si l'alumnat s'agrupa en funció de les seves capacitats en diferents classes (SC042Q01) o dins de la mateixa classe (SC042Q02). Ambdues s'han recodificat i prenen dos valors: 1 quan l'estudiantat s'agrupa de certa manera a almenys una matèria i 0 en cas contrari (a cap matèria).

c) Valors i compromís

A continuació, s'ha seleccionat l'índex d'activitats extraescolars CREATIV que ve derivat de la pregunta SC053 del qüestionari. En aquesta es proposen tres grups d'activitats: 1) banda, orquestra o cor; 2) obra de teatre o musical escolar; i 3) club d'art o activitats artístiques. El centre indica quins grups proporciona a l'alumnat i la variable està creada a partir de la suma del nombre de grups d'activitats seleccionades. Per tant, l'índex pren valors enters entre 0 i 3.

A més, hi ha dues variables relacionades amb l'actitud de l'estudiantat (STUBEHA) i del professorat (TEACHBEHA). La pregunta del qüestionari per crear els dos índexs és la SC061 que interpel·la sobre quins factors dificulten l'aprenentatge de l'alumnat. La primera variable s'ha creat en base als ítems relatius a les actituds de l'alumnat i la segona en base al comportament del professorat. Ambdues estan tipificades i el valor mitjà és zero. Valors alts en aquests índexs indiquen que la percepció de l'equip directiu del centre respecte a les respectives actituds dificulta l'educació de l'alumnat. Per tant, les actituds dels centres amb valors negatius no perjudiquen l'aprenentatge de l'estudiantat.

Cal recordar que els índexs construïts s'han d'analitzar en comparació a la mitjana de l'OCDE.

d) Recursos humans

Els dos últims grups fan referència a l'equip professional de l'escola i, en aquest apartat, s'han seleccionat índexs ja creats. Amb la variable STAFFSHORT s'estudia una possible manca de personal docent al centre i s'ha creat a partir de la pregunta SC017 del qüestionari. Valors positius en aquest índex indiquen que la quantitat i/o la qualitat del personal és un obstacle per l'educació de l'alumnat, en comparació amb la mitjana de la OCDE.

La ràtio estudiant-professor STRATIO s'ha obtingut dividint el nombre d'alumnes (SCHSIZE) entre el nombre de professors/es (TOTAT). A l'hora de computar l'índex el professorat a temps parcial es pondera amb 0,5.

Per últim, a la base de dades proporcionada per l'OCDE, es troben quatre variables que indiquen la proporció de professorat certificat en base a la classificació estandarditzada de 1997 dels nivells educatius (ISCED). Per a l'estudi s'ha tingut en compte la variable amb el nivell més alt (PROAT6, proporció del professorat amb un nivell ISCED 6) que equival a haver cursat un doctorat. Amb aquesta informació es pretén definir una tipologia d'escola amb un percentatge de professorat altament format.

e) Recursos materials

De la mateixa manera que està construït l'índex STAFFSHORT, per aquest apartat es disposa de la variable EDUSHORT, definida com manca de material educatiu. Els valors positius, per tant, indiquen que la quantitat i/o la qualitat del material educatiu disponible a l'escola és un obstacle per a l'educació de l'alumnat, en comparació amb la mitjana de la OCDE.

En referència als recursos informàtics i tecnològics s'ha extret la informació de dues preguntes del qüestionari, la SC155 amb onze ítems de resposta múltiple i la SC56 amb vuit ítems de resposta dicotòmica.

La primera pregunta que s'ha tingut en consideració, SC155, planteja al/la director/a avaluar la capacitat del centre de millorar l'aprenentatge i l'ensenyament mitjançant l'ús de dispositius digitals. Les onze variables estan detallades a l'Annex 1: Anàlisi descriptiva de les variables però, per explicar un exemple, el sisè ítem pretén avaluar si el professorat té les habilitats tècniques i pedagògiques necessàries per integrar dispositius digitals en l'ensenyament. Les quatre categories de resposta, en escala de Likert, per a les onze variables són: (1) "Totalment en desacord"; (2) "En desacord"; (3) "D'acord"; i (4) "Totalment d'acord". Per incloure aquesta informació a l'estudi s'ha aplicat una Anàlisi de Correspondències Múltiples (veure l'apartat 4.3) per reduir-ne la dimensionalitat.

La segona pregunta que fa referència a les tecnologies digitals és la SC156 i està composta per vuit ítems. En aquest cas es demana si el centre disposa de certes característiques i, per tant, les dues categories de resposta són "Sí" (1) i "No" (2). Algunes de les característiques qüestionades son per exemple: "[tenir] el seu propi programa sobre l'ús de dispositius digitals" o "[tenir] un programa específic que prepari l'alumnat per a un comportament responsable a Internet". La representació més detallada dels ítems es troba a l'Annex 1: Anàlisi descriptiva de les variables. Tractant-se d'un conjunt de variables binàries s'ha unificat la informació recodificant els valors com s'exposa a l'apartat 4.2 de recodificació de les variables.

4. Preparació de les dades

Un cop seleccionades les bases de dades i les variables d'interès, és necessari realitzar un procés de preparació de les dades. Amb l'objectiu de validar que els valors que prenen les variables son correctes es realitza una anàlisi descriptiva (detallada a l'Annex 1: Anàlisi descriptiva de les variables) i es compara amb la informació proporcionada al *diccionari de variables*². Havent realitzat totes les modificacions prèvies a la modelització, la mostra resultant consta de 33.514 alumnes i 1011 escoles.

4.1. Dades mancants

El primer pas per realitzar un processament de les dades és el tractament dels valors mancants. Al nivell de les escoles, havent triat diverses variables extretes directament del qüestionari, es disposa d'alguns centres educatius que no han contestat les preguntes del qüestionari. S'ha optat per eliminar de la mostra les 25 escoles (2,30%) i els alumnes corresponents als centres (Dong, 2019).

A més, PISA estableix una nomenclatura especial per a les dades mancants en base al motiu. A les bases de dades tractades en aquest estudi només es disposa de valors especificats amb una "M" que representen les respostes que no han sigut contestades per elecció (OCDE, 2009). Aquests valors han estat codificats com valors perduts (*missing values*).

Segons Husson (2020), al programa R, per realitzar una anàlisi de correspondències múltiples, els valors mancants es tracten com una categoria addicional. A l'apartat 4.3 s'aplica aquesta tipologia d'anàlisi per als ítems de la variable SC155 i, per tant, s'han eliminat les observacions que tenen dades mancants en aquestes variables.

4.2. Recodificació de les variables

Amb les bases de dades "netes" s'ha realitzat un procés de recodificació per facilitar la interpretació dels resultats de l'estudi. En alguns casos només s'han canviat els valors de les etiquetes i en altres s'han modificat les agrupacions de valors. Al nivell dels centres educatius s'han modificat els valors de la pregunta SC042 (agrupaments per nivells), la variable CLSIZE (nombre d'alumnes per classe) i l'índex HISCED (el nivell educatiu més alt dels genitors). De l'estudiantat només s'han canviat els valors de les etiquetes de la variable gènere (ST004). Els valors obtinguts un cop realitzades les recodificacions estan especificats a la selecció de variables (apartat 3).

² www.OCDE.org/pisa/data

Per últim, s'ha creat una única variable derivada de la pregunta SC156 del qüestionari de les escoles en relació a les tecnologies digitals. Aquesta pregunta està composta per vuit ítems amb una resposta dicotòmica (sí o no). La variable s'ha creat de la mateixa manera que l'índex que proporciona PISA, CREATIV, a partir de la suma de les opcions seleccionades (OCDE, 2018). Per tant, aquesta variable pren valors de 0 a 8 en base al nombre d'ítems als que s'ha contestat afirmativament.

4.3. Càlcul de noves variables

Transformacions

A l'hora d'estimar els models de l'estudi, no és convenient disposar de variables amb magnituds d'ordre molt diferents. Realitzant l'anàlisi descriptiva univariant dels valors que prenen les variables s'ha considerat necessari modificar els predictors TMINS (temps d'estudi de l'alumnat en minuts) i SCHSIZE (nombre d'alumnes a l'escola). En el primer cas, el valor màxim que prenia la variable eren 3000 minuts i, per tant, s'han calculat els valors en hores.

Respecte a SCHSIZE, la variable prenia valors entre 19 i 2698 alumnes i en aquest cas s'ha optat per utilitzar un mètode de tipificació. S'ha emprat la funció "scale" del paquet "base" (2018) del programa R. En primer lloc es centren les dades restant la mitjana de la columna i a continuació s'escalen dividint per la desviació estàndard. D'aquesta manera els valors varien de -1,56 a 4,48 amb mitjana 0,099. Valors més alts de la variable indiquen escoles de més alumnat.

Aplicació de l'Anàlisi de Correspondències Múltiples

Per últim, respecte al conjunt de preguntes relatives a l'ús de les tecnologies digitals als centres educatius, es disposa de la pregunta SC155 amb onze ítems (detallats a l'Annex 1: Anàlisi descriptiva de les variables taula A1.26). Tractant-se de respostes tancades múltiples, s'ha estudiat la possibilitat de reduir la dimensionalitat per facilitar una posterior interpretació. Degut al caràcter qualitatiu de les variables, s'ha implementat una Anàlisi de Correspondències Múltiples (ACM) al conjunt d'ítems.

Cal recordar que les variables incloses a l'anàlisi estableixen la capacitat del centre per millorar l'aprenentatge i l'ensenyament mitjançant l'ús de dispositius digitals. Es mesuren en una escala de Likert on 1 equival a "totalment en desacord" i 4 a "totalment d'acord". Amb l'anàlisi implementada, els centres s'agrupen en tres classes i, per tant, la variable consta de les tres categories corresponents. Al primer grup es troben la majoria d'escoles (76% de les observacions) i els ítems d'estudi prenen valors similars a la mitjana de la mostra. Per tant, els directors/es d'aquestes escoles han contestat que tenen suficients dispositius digitals, disposen d'un bon programari i d'una bona plataforma i que el professorat té les habilitats

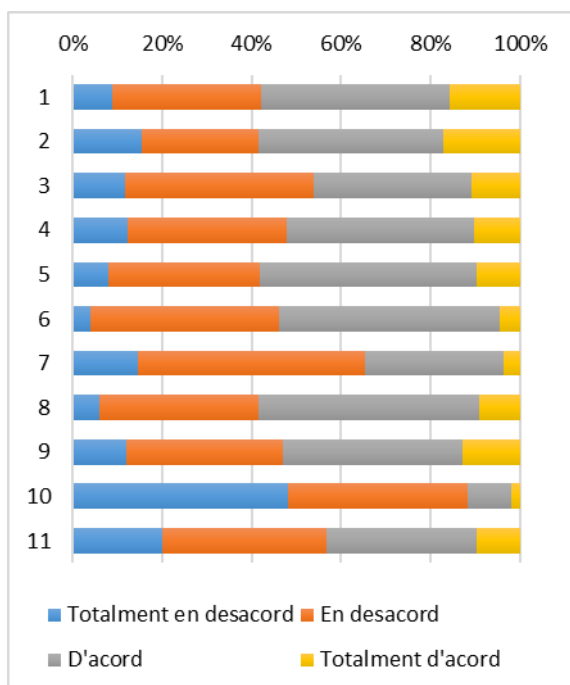
tècniques i pedagògiques necessàries per integrar dispositius digitals en l'ensenyament. La segona classe està composta per poques escoles (9%) i son les que han contestat amb valors més alts a les preguntes. Es tracta de centres que tenen majors capacitats per millorar l'aprenentatge i l'ensenyament mitjançant l'ús de dispositius digitals. Per últim, la tercera classe agrupa el 15% de les escoles i prenen valors per sota de la mitjana, indicant que les seves capacitats son inferiors. La darrera classe s'ha recodificat amb valor zero per establir-la com a categoria base pel model ja que és la que pren els valors inferiors.

Taula 4.1: Resultats pregunta SC155

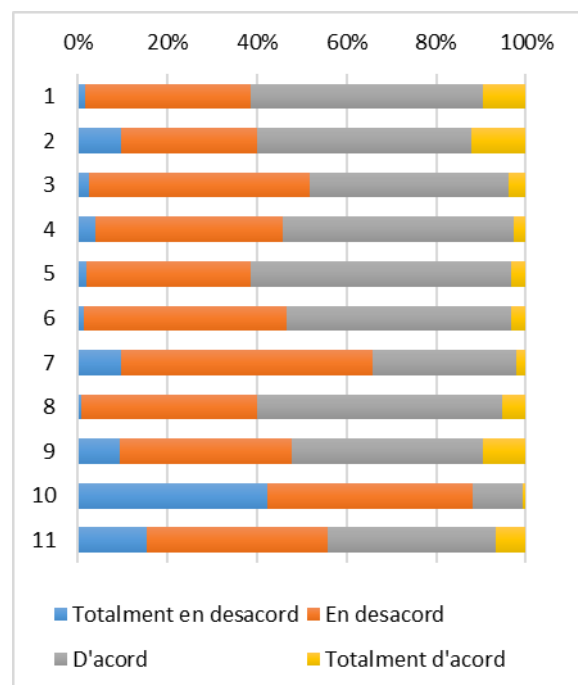
	Clúster 1	Clúster 2	Clúster 3
Codificació	1	2	0
n	149	772	90
%	15%	76%	9%

A continuació es mostren els gràfics de la pregunta SC155 amb els percentatges de resposta per a tota la mostra d'escoles i per a les escoles de cada grup del *clustering*. Els onze ítems de la pregunta es troben detallats a la taula A1.29 de l'Annex 1: Anàlisi descriptiva de les variables.

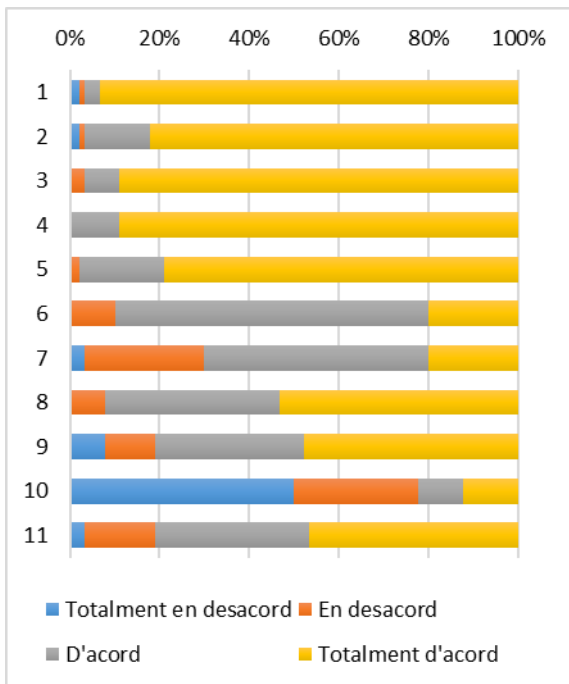
Il·lustració 4.2: Pregunta SC155, mostra escoles



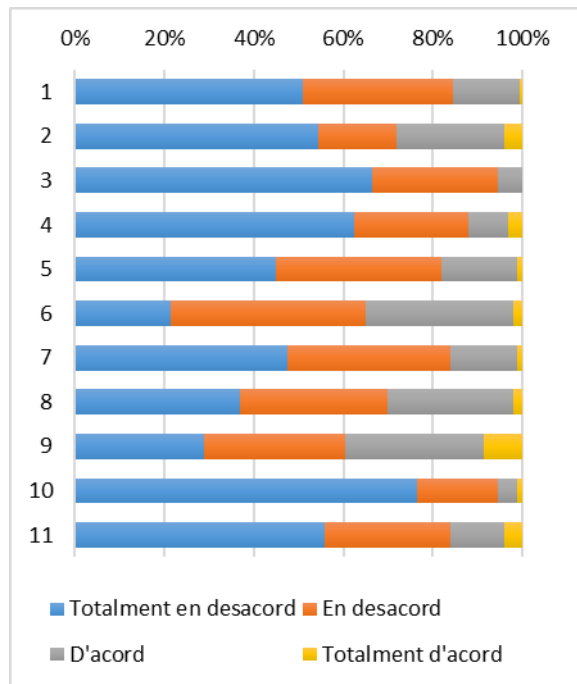
Il·lustració 3: Pregunta SC155, escoles clúster 1



Il·lustració 4.4: Pregunta SC155, escoles clúster 2



Il·lustració 4.5: Pregunta SC155, escoles clúster 3



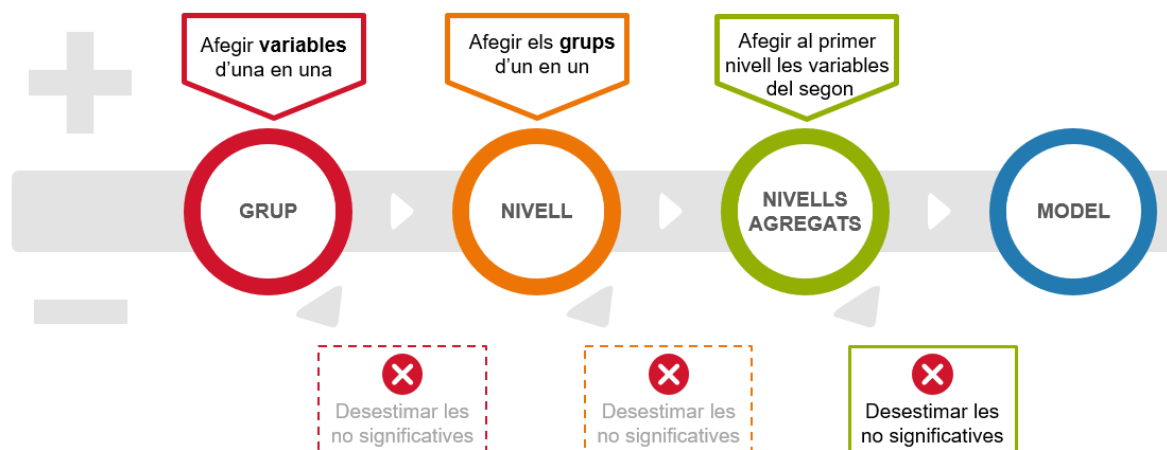
5. Resultats

5.1. Especificació del model jeràrquic

Com s'esmenta a l'apartat 2.1.1, s'ha realitzat una anàlisi sistemàtica per determinar l'especificació del model jeràrquic definitiu partint de la creació del model nul. Aquest model no inclou predictors en cap dels nivells, estima la mitjana global del rendiment de l'alumnat i ens permetrà analitzar la variància de cada un dels nivells (Raudenbush, 2002). Seguint la recomanació de l'OCDE, l'anàlisi exploratòria dels models s'ha realitzat amb el primer valor plausible com a variable endògena i, per tant, no s'analitzen els coeficients de les variables exògenes. Un cop establert el model definitiu s'ha emprat el valor mitjà dels estadístics realitzant els models pels deu valors plausibles proporcionats (OCDE, 2009).

A l'hora de realitzar l'anàlisi amb l'objectiu d'identificar un model multinivell que ajusti les dades, s'han seguit diversos procediments resumits a la Il·lustració .

Il·lustració 5.1 Esquema de l'especificació del model



A cada pas, les variables es descarten si el fet d'afegir-les al model no millora l'ajust de les dades. En un principi s'han tingut en compte els nivells i els grups en els quals s'han classificat les variables a l'apartat 3. Un cop s'han ajuntat els grups de variables per a cada nivell, es comprova la significació dels paràmetres. Desestimar els efectes amb un nivell de significació inferior al 5% és una condició molt restrictiva respecte als resultats obtinguts. Tenint en compte aquest factor, s'han descartat les variables els quals efectes no resulten significatius tant en el model complet com en el model amb les variables del respectiu grup. Per últim, tenint un model complet per a cada nivell, s'han inclòs els predictors resultants dels dos nivells i s'ha realitzat un model multinivell definitiu.

```

Linear mixed model fit by maximum likelihood . t-tests use
Satterthwaite's method [lmerModLmerTest]
Formula: PVI_MATH ~ (1 | CNTSCHID)
Data: ddm1
Weights: W_FSTUWT

          AIC          BIC      logLik  deviance  df.resid
406012.1  406037.4 -203003.1  406006.1    33511

Scaled residuals:
  Min       1Q   Median       3Q      Max
-7.5348 -0.4841  0.0631  0.5640  6.8312

Random effects:
 Groups      Name      Variance Std.Dev.
CNTSCHID (Intercept)  971.4    31.17
Residual              79043.8   281.15
Number of obs: 33514, groups: CNTSCHID, 1011

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept)  485.025      1.156  979.191  419.6  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

A la Il·lustració es mostra el resum del model nul i, a la part superior, es troben les especificacions amb les que s’ha aplicat un model lineal mixt ajustat amb el criteri de màxima versemblança (més detall a l’apartat 2.1.2). S’ha emprat la formulació referent a un model amb l’intercepte aleatori i que, per tant, realitza una regressió per a cada escola. En aquest cas, tractant-se d’un model sense predictors, s’utilitzen totes les observacions, els 33.514 alumnes i les 1.011 escoles. El valor que es té en compte a l’hora de realitzar les comparacions entre models és el que fa referència a la desviància, anomenat *deviance*, i en el model nul pren valor de 406.006,1.

La mitjana del rendiment de l’estudiantat total en el model nul s’estima en 485,025. La mitjana per a l’escola j , en canvi, és estimada com $485,025 + \hat{u}_{0j}$ on \hat{u}_{0j} son els residus dels centres educatius. Un centre amb $\hat{u}_{0j} \geq 0$ té la mitjana del rendiment superior al valor mitjà estimat i un centre amb $\hat{u}_{0j} \leq 0$ estarà per sota.

L’anàlisi de les variàncies explicades pel model es basa en la informació proporcionada pels efectes aleatoris (Szaragad, 2011). La variància entre les escoles (nivell 2) s’estima com $\hat{\sigma}_{u_0}^2 = 971,4$ i la variància entre l’estudiantat dintre d’un mateixa escola com $\hat{\sigma}_{\epsilon}^2 = 79.403,8$. La variància total, per tant, és $\hat{\sigma}_{u_0}^2 + \hat{\sigma}_{\epsilon}^2 = 971,4 + 79.403,8 = 80.015,2$.

Un cop obtinguts els resultats del model bàsic, s’han implementat diversos models afegint les variables predictores d’una en una i en grups per a cada nivell per separat. Els resultats del model on s’han afegit les variables de l’estudiantat (primer nivell) es mostren a la Il·lustració 5..

Il·lustració 5.5: Taula resum del model amb predictors de l'estudiantat

```

Linear mixed model fit by maximum likelihood . t-tests use
Satterthwaite's method [lmerModLmerTest]
Formula:
PVMATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + ATTLNACT +
  TMINS + TEACHINT + DIRINS + (1 | CNTSCHID)
Data: ddm1
Weights: W_FSTUWT

      AIC      BIC    logLik  deviance  df.resid
229256.8 229359.3 -114615.4 229230.8   19610

Scaled residuals:
  Min       1Q   Median       3Q      Max
-7.8234 -0.5041  0.0365  0.5530  5.5802

Random effects:
 Groups      Name      Variance Std.Dev.
CNTSCHID (Intercept)  352.8    18.78
Residual          48047.4   219.20
Number of obs: 19623, groups: CNTSCHID, 999

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  5.196e+02  2.544e+00  1.689e+04 204.272 < 2e-16 ***
ST0041       1.905e+01  9.910e-01  1.939e+04  19.226 < 2e-16 ***
REPEAT1     -8.208e+01  1.294e+00  1.961e+04 -63.432 < 2e-16 ***
IMMIG2      -1.495e+00  2.341e+00  1.961e+04 -0.638 0.523205
IMMIG3      -9.771e+00  2.144e+00  1.962e+04 -4.557 5.21e-06 ***
ESCS        1.955e+01  7.617e-01  1.918e+04 25.670 < 2e-16 ***
HISCED1     -1.917e+01  1.524e+00  1.948e+04 -12.584 < 2e-16 ***
ATTLNACT    1.746e+00  5.208e-01  1.957e+04  3.353 0.000801 ***
TMINS       8.850e-02  7.125e-02  1.960e+04  1.242 0.214235
TEACHINT    6.820e+00  5.525e-01  1.959e+04 12.345 < 2e-16 ***
DIRINS     -5.455e+00  6.063e-01  1.959e+04 -8.998 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Es pot comprovar que s'han afegit nou variables a la modelització: ST004 (gènere), REPEAT (el grau de repetició), IMMIG (historial immigrant), ESCS (estat econòmic, social i cultural), HISCED (màxim nivell educatiu dels genitors), ATTLNACT (actitud envers les activitats escolars), TMINS (temps d'estudi), TEACHINT (interès del professorat) i DIRINS (classes direccionades). La desviància, estadístic que indica l'ajust del model, és 229.230,8 i ha disminuït considerablement respecte al model nul, aproximadament a la meitat. El fet d'afegir les variables del nivell de l'estudiantat millora l'ajust del model. En aquest cas, havent afegit predictors amb valors mancants la mostra considerada és de 19.623 estudiants/es i 999 escoles. La variància total explicada per aquest model és de $\hat{\sigma}_{u_0}^2 + \hat{\sigma}_{\varepsilon}^2 = 352,8 + 48047,4 = 48400,2$.

Per una altra banda, el model amb les variables del segon nivell, referents a les escoles, es mostra a la Il·lustració 5..

Il·lustració 5.6 Taula resum del model amb predictors de les *escoles*

```

Linear mixed model fit by maximum likelihood . t-tests use
Satterthwaite's method [lmerModLmerTest]
Formula:
PV1MATH ~ SCHLTYPE + SCHSIZE + STUBEHA + TEACHBEHA + SC042Q01 +
CSC156 + STAFFSHORT + STRATIO + PROAT6 + (1 | CNTSCHID)
Data: ddm1
Weights: W_FSTUWT

          AIC          BIC      logLik deviance df.resid
300845.9  300951.5 -150410.0  300819.9   24821

Scaled residuals:
   Min       1Q   Median       3Q      Max
-7.6474 -0.4799  0.0608  0.5592  6.9148

Random effects:
 Groups Name Variance Std.Dev.
CNTSCHID (Intercept) 643.8 25.37
Residual 77312.8 278.05
Number of obs: 24834, groups: CNTSCHID, 747

Fixed effects:
              Estimate Std. Error    df t value Pr(>|t|)
(Intercept) 515.3089      6.8799 756.9265  74.901 < 2e-16 ***
SCHLTYPE2   -18.7126      5.0169 773.2144  -3.730 0.000206 ***
SCHLTYPE3   -33.3076      5.2105 770.8711  -6.392 2.83e-10 ***
SCHSIZE      8.0719      1.3801 723.2157   5.849 7.50e-09 ***
STUBEHA     -9.8028      1.4658 751.2073  -6.688 4.42e-11 ***
TEACHBEHA    2.5199      1.5580 721.4562   1.617 0.106239
SC042Q011   -4.9656      2.4231 733.5196  -2.049 0.040794 *
CSC156       0.3498      0.5680 791.4327   0.616 0.538138
STAFFSHORT  -2.3528      1.5019 735.3463  -1.567 0.117652
STRATIO     -0.3646      0.3232 713.2683  -1.128 0.259698
PROAT6      78.1312     22.1007 805.0189   3.535 0.000431 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Les nou variables del nivell de les escoles que s’han considerat pel model son: SCHLTYPE (titularitat de l’escola), SCHSIZE (número d’alumnes), STUBEHA (comportament de l’alumnat), TEACHBEHA (comportament del professorat), SC042Q01 (divisió en classes en base al nivell), CSC156 (predisposició envers els dispositius digitals), STAFFSHORT (manca de professorat), STRATIO (rati alumnat-professorat) i PROAT6 (proporció de professorat altament certificat). En aquest cas l’estadístic de la desviància és 300.819,9 i ha disminuït aproximadament una tercera part respecte al model nul. Afegir les variables del nivell de l’escola, per tant, millora l’ajust del model que té en consideració 24.834 alumnes i 747 escoles. La variància total s’estima com $\hat{\sigma}_{u_0}^2 + \hat{\sigma}_{\varepsilon}^2 = 643,8 + 278,05 = 921,05$.

Havent verificat que els models amb les variables predictores dels dos nivells per separat milloren l’ajust de les dades, s’han agregat els dos nivells. En aquesta fase comparativa s’han realitzat els càlculs amb el primer valor plausible per representar el rendiment de l’estudiantat.

Taula 5.1: Comparació de les desviàncies dels models

	N estudiantat	N escoles	Paràmetres	Desviància
Model nul	33514	1011	1	406006,1
Model amb estudiantat	19623	999	11	229230,8
Model amb escoles	24834	747	11	300819,9
Model complet	14820	737	22	173073,8

Taula 5.2: Comparació de les variàncies dels models

	Variància estudiantat ($\hat{\sigma}_{\varepsilon}^2$)	Variància escoles ($\hat{\sigma}_{u_0}^2$)	Variància total ($\hat{\sigma}_{u_0}^2 + \hat{\sigma}_{\varepsilon}^2$)
Model nul	79043,84	971,43	80015,27
Model amb estudiantat	48047,44	352,81	48400,25
Model amb escoles	77312,82	643,76	77956,58
Model complet	46575,27	316,77	46892,04

La diferència entre la desviància del model definitiu amb el model especificat sense cap variable predictora és de 232.932,3. Per tant, es comprova que es compleix l'especificació d'obtenir un estadístic de la desviància al model complet considerablement inferior al model nul (Taula 5.1: Comparació de les desviàncies dels models). A més, la modelització ajuntant les variables dels dos nivells explica el 41,4% de la variància total del model nul (Taula).

5.2. Model definitiu

A la Taula 5.3 s'exposen els principals resultats de l'estudi empíric que es comenten a continuació. Els valors obtinguts corresponen al valor mitjà dels deu models emprant cada valor plausible. La primera columna presenta els coeficients de la regressió logística a dos nivells, que informen sobre el signe de la relació entre la variable i el rendiment de l'alumnat. A més, es mostren els errors estàndard a la segona columna i el nivell de significació dels paràmetres.

Taula 5.3: Taula resum dels coeficients del *model definitiu*

Variables	Definició	Estimació	Std. Error	
(Intercepte)	Constant	520,532	6,252	***
ST0041	Gènere (0 = dona)	17,733	1,146	***
REPEAT1	Grau de repetició	-83,331	1,508	***
ESCS	Estat econòmic, social i cultural	17,147	0,965	***
IMMIG2	Immigrant de segona generació	-1,727	2,740	
IMMIG3	Immigrant de primera generació	-13,614	2,510	***
HISCED1	Màxim nivell educatiu dels genitors	-14,175	1,814	***
HEDRES	Recursos educatius a casa	-0,045	0,706	
TEACHINT	Interès del professorat	5,731	0,644	***
DIRINS	Classes direccionades	-5,568	0,708	***
TMINs	Temps d'estudi	0,217	0,084	**
ATTLNACT	Actitud amb activitats escolars	1,663	0,604	
SCHLTYPE2	Escola concertada	-3,398	4,039	
SCHLTYPE3	Escola pública	-2,485	4,224	
SCHSIZE	Número d'alumnes	4,257	1,137	***
STUBEHA	Actitud de l'alumnat	-4,016	1,208	***
TEACHBEHA	Actitud del professorat	0,975	1,283	
SC042Q011	Dividir l'alumnat per nivell	-2,922	2,011	
CSC156	Predisposició dispositius digitals	-0,118	0,470	
STAFFSHORT	Mancaça de personal	-0,428	1,243	
STRATIO	Ràtio alumnat professorat	-0,230	0,263	
PROAT6	Proporció professorat doctorat	29,362	18,082	**

Nota 1: *** p<0.01, ** p<0.05, * p<0.1

Nota 2: Les categories de referència son les següents:

Nivell estudiantat: ser dona (ST004); no haver repetit cap curs (REPEAT); estat econòmic, social i cultural mitjà (ESCS); ser nadiu (IMMIG); genitors amb nivell educatiu mínim (HISCED); nivell mitjà de recursos educatius a casa (HEDRES); percepció mitjana de l'interès del professorat (TEACHINT); percepció del mètode docent mitjà (DIRINS); temps d'estudi mitjà (TMINs); i actitud mitjana envers les activitats escolars (ATTLNACT).

Nivell escoles: ser privada (SCHLTYPE); nombre d'alumnes igual a la mitja de la mostra SCHSIZE); comportament mitjà de l'estudiantat i del professorat (STUBEHA i TEACHBEHA); no dividir els alumnes en base al seu nivell a cap matèria (SC042Q01); no tenir predisposició envers els dispositius digitals (SC156); tenir una mancaça de personal mitjà (STAFFSHORT); disposar d'una ràtio alumnat-professorat igual a la mitjana de la mostra (STRATIO); no tenir professorat altament qualificat (PROAT6).

Predictores del nivell de l'estudiantat

La primera variable amb una significació de l'1% és el gènere de l'alumnat, les noies obtenen, de mitjana, 17,7 punts menys que els nois. Segons l'estudi Calero (2010) amb les dades de PISA de 2006, el fet de ser dona incrementa en un 40% la probabilitat que un alumne obtingui una puntuació negativa a la prova, tot i que la taxa de fracàs escolar dels homes és major. Altres variables tenen paràmetres de major magnitud. Així, l'alumnat que no ha repetit cap curs té una mitjana de 83,3 punts més respecte a l'estudiantat que ha repetit un o més cursos.

La variable predictora de l'estat econòmic, social i cultural (ESCS) és significativa amb un nivell de confiança del 99%. Aquesta influeix de manera positiva, la mitjana del rendiment de matemàtiques de l'alumnat augmenta en un 17,1 per cada unitat de la variable. Els resultats de l'estudiantat immigrant de segona generació (persona nascuda al país i els genitors a l'estranger) no son significativament diferents respecte a l'estudiantat nadiu. En canvi, el fet d'haver nascut a l'estranger es reflecteix de manera significativa als resultats, en mitjana obtenen 13,6 punts menys que la gent nascuda al país. A l'estudi realitzat per Martín (2009), s'utilitza una única variable ("condició immigrant") que resulta significativa amb un nivell inferior a 0,01 i amb el signe del coeficient negatiu.

Fent referència a l'entorn familiar de l'estudiantat, tot i no tenir un efecte estadísticament significatiu, el fet d'afegir la variable HEDRES (recursos educatius a casa) millora el model i el signe es correspon amb l'esperat. Respecte al signe del coeficient de la variable HISCED (nivell educatiu dels genitors) resulta contradictori, quant més alt el nivell pitjor són els resultats. Una possible explicació de l'estrany funcionament de la predictora és que la variable ESCS està calculada a partir de tres índexs i una de les dimensions fa referència a l'educació dels genitors. Per altra banda, tant a Escardíbul (2008) com a Calero (2010) es mostren incongruències entre els signes del paràmetre en qüestió.

L'actitud del professorat és de gran rellevància a l'estudi i es disposa de dues variables sobre la percepció de l'estudiantat respecte al professorat. En primer lloc, el fet que l'alumnat tingui la percepció de que el professorat s'interessa a l'hora de donar les classes (TEACHINT) és significatiu amb un nivell de confiança del 99%. Augmentar una unitat de l'índex implica un augment en la mitjana del rendiment de l'alumnat de 5,7 punts. En segon lloc, el fet que l'estudiantat indiqui que les classes són molt direccionades (DIRINS) resulta en un rendiment, en mitjana, de 5,6 punts menys a matemàtiques. L'actitud de l'alumnat envers les activitats referents a l'escola no resulta significativa. Per últim, el temps d'estudi dedicat, amb una significació del 5%, també influeix positivament en els resultats. Cada hora setmanal de més dedicada a l'estudi augmenta la mitjana dels resultats de matemàtiques de 0,2 punts.

Predictores del nivell de les escoles

Respecte a les variables relacionades amb les escoles, la tipologia (SCHLTYPE) no té efectes estadísticament significatius per cap de les categories. El fet de no tenir efectes significatius està en línia amb l'estudi realitzat per Escardíbul (2008) i l'anàlisi de Guio-Jaimes (2014). El nombre d'alumnes és significativament important (p-valor inferior a 0,01) a l'hora d'explicar el rendiment de l'alumnat i influeix de manera positiva. La variable està tipificada i, per tant, cada unitat per sobre de la mitjana de la mostra augmenta la mitjana dels resultats en matemàtiques de 4,3 punts.

Valors elevats a la variable STUBEHA indiquen que la percepció de l'equip directiu del centre respecte a les actituds de l'estudiantat dificulta l'educació de l'alumnat. L'efecte d'aquesta variable és significatiu a l'1% i per cada unitat de la variable es disminueix la mitjana del rendiment de l'alumnat en 4 punts. Una escola amb una proporció de professorat altament qualificat, és a dir doctorat (variable PROAT6), obté un rendiment de l'estudiantat considerablement millor.

Per últim, cal destacar quatre variables predictores que no son significatives però el fet d'afegir-les al model millora l'ajust de les dades. La primera variable és CSC156, creada a partir de les respostes del qüestionari sobre la predisposició envers els dispositius digitals, que no té el signe esperat ja que planificar la política digital de l'escola s'esperaria que tingués signe positiu. En segon lloc, el fet de dividir l'alumnat en diverses classes en base al nivell de l'alumnat (SC042) tindria un efecte positiu. Per últim, STAFFSHORT i STRATIO tenen els coeficients amb signe negatiu i, per tant, corresponent al que caldria esperar ja que tenir una mancança de professorat elevada i una ràtio alta d'alumnes per professor influiria de manera negativa.

6. Conclusions

El present treball es centra en identificar els factors relacionats amb els centres educatius que influeixen en els resultats de matemàtiques de l'estudiantat de 15 anys. Es fa servir la principal font de dades disponible, l'estudi PISA (Programa Internacional per a l'avaluació d'estudiants), que examina el rendiment de l'alumnat en àrees temàtiques clau. En concret, l'àmbit geogràfic d'aquest treball és Espanya i s'analitzen les dades de la seva darrera aplicació, recollides l'any 2018. L'anàlisi empírica d'aquest estudi es basa en l'aplicació de tècniques de regressió multinivell ja que les microdades disponibles tenen una estructura jeràrquica (OCDE, 2009).

Les principals conclusions extretes de l'anàlisi multinivell de les dades indiquen que diversos factors relacionats amb les característiques de l'alumnat juguen un paper rellevant en el rendiment acadèmic en l'àmbit de les matemàtiques. En primer lloc, les variables de l'àmbit personal, del context familiar i del centre educatiu de l'alumnat incideixen de manera significativa en el rendiment en matemàtiques de l'estudiantat de 15 anys. Dins de l'àmbit personal, les noies presenten resultats inferiors respecte als nois i el fet de no haver repetit cap curs influeix de manera positiva en el rendiment. Respecte al context familiar, el nivell econòmic, social i cultural, definit per la variable ESCS, també és de gran rellevància per explicar els resultats de l'alumnat i té una relació positiva amb la variable endògena.

Els resultats obtinguts amb la condició d'immigració proporcionen una distinció del tipus d'alumnat immigrant (IMMIG). Tot i no ser el centre de l'estudi del projecte, aquesta informació mereix una menció especial. L'estudiantat immigrant de primera generació (tant la persona com els genitors nascuts a l'estranger) obté, de mitjana, una puntuació molt més baixa que l'alumnat nadiu. Aquesta diferència podria ser explicada per dificultats relacionades amb l'idioma o per un procés deficient d'adaptació al país. En canvi, amb la categoria referent a la migració de segona generació (persona nascuda al país i els genitors a l'estranger) es demostra que no hi ha diferències significatives respecte a l'alumnat nadiu.

El qüestionari respost per l'estudiantat proporciona informació sobre la seva percepció envers el professorat i el propi centre educatiu. Aquesta informació resulta molt rellevant a l'hora d'extreure conclusions en relació a l'objectiu de l'estudi. El fet que l'alumnat tingui la percepció que el professorat mostra interès a l'hora de donar la classe contribueix a incrementar el rendiment en matemàtiques. A més, les formes de docència dirigistes (DIRINS) influeixen negativament en el rendiment de l'estudiantat. Es tracta d'un mètode d'ensenyament més tradicional on el professorat imparteix la lliçó sense la intervenció de l'alumnat. Per últim, s'han tractat dues variables d'àmbit més personal, l'actitud de l'estudiantat envers les activitats escolars i el temps d'estudi dedicat. Ambdues tenen una

relació positiva amb el rendiment, millor actitud i més hores dedicades a l'estudi millors resultats en matemàtiques.

Respecte a les variables relacionades amb les escoles, que són el principal interès d'aquesta recerca, la titularitat del centre educatiu (SCHLTYPE) no té efectes estadísticament significatius. Per tant, possibles diferències en resultats el rendiment de l'estudiantat no es determinen en base a la tipologia del centre. La dimensió de l'escola, mesurada pel nombre d'alumnes (SCHSIZE), i la qualificació elevada del professorat, mesurada per la proporció de docents amb títol de doctor/a (PROAT6), influeixen de manera positiva en el rendiment de l'estudiantat. Finalment, el fet que el/la director/a de l'escola indiqui que el comportament de l'alumnat perjudica l'aprenentatge (STUBEHA) afecta negativament els resultats obtinguts a les proves de matemàtiques.

És necessari destacar la manca de significativitat de moltes variables relacionades amb les escoles. Com destaca Escardíbul (2008) una possible explicació és l'alta similitud entre els centres educatius en relació a la dotació de determinats recursos i a l'organització de l'ensenyament. L'estudi d'Escardíbul mostra que la variació dels resultats entre escoles és significativament menor a Espanya que al conjunt de països de l'OCDE i, per tant, l'alumnat mostra més diferències en una mateixa escola que en relació amb altres centres.

Cal senyalar que les predictores comportament del professorat (TEACHBEHA) i predisposició envers els dispositius digitals (CSC156) no han produït diferències significatives sobre el rendiment i els coeficients no tenen el signe esperat. Respecte a la primera variable, seria d'esperar que un comportament del professorat que dificultés l'aprenentatge de l'estudiantat afectés de manera negativa. La variable CSC156 s'ha creat sumant el nombre d'ítems seleccionats a la pregunta SC156 (veure apartat 4.2) i podria ser que aquesta aproximació no representés la informació de manera completa. Finalment, tot i no ser estadísticament significatives, les variables relacionades amb el fet de dividir l'alumnat en classes en base al nivell (SC042), tenir una manca de personal (STAFFSHORT) i una ràtio alumnat-professorat (STRATIO) per sobre de la mitjana de l'OCDE tenen un coeficient amb signe negatiu.

Actualment les Tecnologies de la Informació i de la Comunicació (TIC) són una de les prioritats educatives ja que contribueixen al procés d'ensenyament i d'aprenentatge de l'estudiantat. En el present estudi, tot i haver considerat informació extreta directament de les respostes de l'equip directiu de les escoles sobre la disposició de dispositius digitals, els resultats obtinguts no han estat concloents. Tractant-se d'una temàtica tant àmplia és recomanable realitzar una anàlisi apart més focalitzada en el tema. De totes maneres, com destaca Alderete (2017), la introducció de les TIC a les escoles representa un nou impediment per a l'estudiantat que es troba en desavantatge social i econòmic.

Per concloure, els resultats obtinguts porten cap a un camí comú: els factors relatius a les escoles no expliquen de manera completa els resultats de matemàtiques de l'alumnat de 15 anys. Tot i que la majoria de les variables d'aquest nivell no resulten estadísticament significatives, sí que influeix una component relacionada amb els valors dels centres educatius i les actituds del professorat. Respecte al nivell de l'alumnat, és essencial tenir en compte la component personal, el context familiar i les actituds de l'alumnat. S'ha demostrat que una actitud motivadora a l'hora d'impartir la classe és fonamental per augmentar els resultats en matemàtiques. El fet que la majoria de les variables predictorres del segon nivell, referents a les escoles, no resultin significatives pot ser explicat per una alta similitud en els mètodes d'ensenyament dels diversos centres educatius.

Bibliografía

- Abdi, H. &. (2010). *Principal component analysis*. *Wiley interdisciplinary reviews: computational statistics*, 2(4), 433-459.
- Alderete, M. V. (2017). *Acceso a las TIC y rendimiento educativo: ¿una relación potenciada por su uso? Un análisis para España*.
- Austin, P. C. (2001). *An introduction to multilevel regression models*. *Canadian journal of public health*, 92(2), 150-154.
- Bolker, B. (2018). *Glmm faq*.
- Buxton, R. (2008). *Statistics: Multilevel modeling*. Publicado por Mathematics Learning, Support Centre.
- Calero, J. C. (2010). *Determinantes del riesgo de fracaso escolar en España: una aproximación a través de un análisis logístico multinivel aplicado a PISA-2006*.
- Christensen, A. K. (2017). *ImerTest Package: Tests in Linear Mixed Effects Models*. doi:10.18637/jss.v082.i13
- Dong, X. &. (2019). *An exploration of impact factors influencing students' reading literacy in Singapore with machine learning approaches*. *Recollit de International Journal of English Linguistics*, 9(5).
- Escardíbul, J. O. (2008). *Los determinantes del rendimiento educativo en España. Un análisis a partir de la evaluación de PISA-2006*. *Investigaciones de Economía de la Educación*, 3, 153-162.
- Guio-Jaimes, J. M.-d.-M. (2014). *The evolution of school failure risk during the 2000 decade in Spain: Analysis of PISA results with a two-level logistic model*.
- Husson, F. J. (2020). *Package 'FactoMineR'*. *Package FactorMineR*.
- Husson, F. L. (2010). *Exploratory multivariate analysis by example using R, volume 20105550 of Chapman & Hall/CRC Computer Science & Data Analysis*. CRC Press, 30, 101-102.
- Kuznetsova, A. B. (2017). *ImerTest package: tests in linear mixed effects models*. *Journal of statistical software*, 82(13).
- Luke, S. G. (2017). *Evaluating significance in linear mixed-effects models in R*. *Behavior research methods*, 49(4), 1494-1502.
- Martín, E. L. (2009). *Estudio de variables determinantes de eficiencia a través de los modelos jerárquicos lineales en la evaluación PISA 2006: el caso de España*.
- Müllner, D. &. (2015). *Package 'fastcluster'*.

- OCDE. (2007). *PISA 2006: Science Competencies for Tomorrow's World. Vol. 1–Analysis*.
- OCDE. (2009). *PISA Data Analysis Manual: SAS, Second Edition*.
- OCDE. (2009). *PISA Technical Report*.
- OCDE. (2018). *Technical notes on analyses in this report", in Equity in Education: Breaking Down Barriers to Social Mobility*.
- OCDE. (2019). *PISA 2018 Results (Volume II): Where All Students Can Succeed*.
- OCDE. (2020). *PISA 2018 Results (Volume I): What Students Know and Can Do*.
- OCDE. (forthcoming). *PISA 2018 Technical Report*.
- Peter Langfelder, S. H. (2012). *Fast R Functions for Robust Correlations and Hierarchical Clustering*.
- Rasbash, J. (2019). *What are multilevel models and why should I use them*.
- Raudenbush, S. W. (2002). *Hierarchical linear models: Applications and data analysis methods (Vol. 1)*. sage.
- Sebastien Le, J. J. (2008). *FactoMineR: An R Package for Multivariate Analysis*.
doi:10.18637/jss.v025.i01
- Snijders, T. A. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage.
- Szmaragd, C. &. (2011). *Module 5: Introduction to Multilevel Modelling R Practical*.
- Team, R. C. (2018). *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing)
- Unesco. (1997). *International Standard Classification of Education-ISCED 1997: November 1997*.

Annexes

Annex 1: Anàlisi descriptiva de les variables

1.1 Estudiantat

a) Àmbit personal

Taula A1.1

ST004: gènere			
Dona (0)	Home (1)	No resposta	N
50,12%	49,88%	0%	33514

Taula A1.2

REPEAT: grau de repetició			
0	1	No resposta	N
73,65%	24,97%	1,38%	33514

Taula A1.3

IMMIG: historial immigrant				
Nadiu (1)	Primera generació (2)	Segona generació (3)	No resposta	N
85,42%	4,51%	7,03%	3,03%	33514

b) Àmbit socio-familiar

Taula A1.4

ESCS: estat econòmic, social i cultural							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-6,07	-0,74	0,03	-0,05	0,77	3,72	1,81%	33514

Taula A1.5

HEDRES: recursos educatius a casa							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-4,49	-0,69	-0,14	-0,10	1,18	1,21	1,36%	33514

Taula A1.6

HISCED: nivell educació genitors (original)								
0	1	2	3	4	5	6	No resposta	N
0,81%	3,37%	11,78%	3,46%	9,75%	20,27%	47,71%	2,85%	33514

Taula A1.7

HISCED: nivell educació genitors (recodificada)			
0	1	No resposta	N
29,17%	67,98%	2,85%	33514

c) Àmbit escolar

Taula A1.8

ATLANCT: actitud activitats escolars							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-2,54	-0,66	0,44	0,18	1,08	1,08	5,13%	33514

Taula A1.9

TMINS: temps d'estudi (original; unitats minuts)							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
120	1500	1650	1710,24	1840	3000	38,88%	33514

Taula A1.10

TMINS: temps d'estudi (transformada; unitats hores)							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
2	25	27,5	28,5	30,67	50	38,88%	33514

Taula A1.11

TEACHINT: interès del professorat							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-2,22	-0,64	0,17	0,07	0,87	1,82	0,21%	33514

Taula A1.12

DIRINS: classes direccionades							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-2,94	-0,76	-0,21	-0,18	0,37	1,82	2,08%	33514

1.2 Escoles

a) Tipologia d'escola

Taula A1.13

SCHLTYPE: tipologia d'escola				
Privada (1)	Concertada (2)	Pública (3)	No resposta	N
6,63%	28,68%	62,51%	2,18%	1011

Taula A1.14

SC001: mida municipi						
Un poblet, llogaret o àrea rural (1)	Un poble (2)	Una ciutat petita (3)	Una ciutat mitjana (4)	Una gran ciutat (5)	No resposta	N
5,93%	24,43%	34,12%	27,79%	7,42%	0,30%	1011

Taula A1.15

SCHSIZE: Nombre alumnes (original)							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
19	384,50	645,50	710,87	824,25	2698	6,23%	1011

Taula A1.16

SCHSIZE: Nombre alumnes (tipificada)							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
1,56	0,73	0,15	0,00	0,48	4,47	6,23%	1011

b) Estructura de les classes

Taula A1.17

CLSIZE:				
Classe petita (1)	Classe normal (2)	Classe gran (3)	No resposta	N
16%	63%	20%	0,79%	1011

Taula A1.18

SC042 Q1: Els estudiants s'agrupen en funció de les seves capacitats en diferents classes			
En cap matèria	En alguna matèria	No resposta	N
63%	33%	4,55%	1011

Taula A1.19

SC042 Q2: Els estudiants s'agrupen en funció de les seves capacitats dins la mateixa classe			
En cap matèria	En alguna matèria	No resposta	N
61%	36%	2,47%	1011

c) Valors i compromís

Taula A1.20

CREACTIV: Activitats extra curriculars					
0	1	2	3	No resposta	N
29,77%	35,41%	23,84%	9,40%	1,58%	1011

Taula A1.21

STUBEHA: Comportament de l'alumnat							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-3,38	-0,75	-0,09	-0,15	0,55	3,44	0,20%	1011

Taula A1.22

TEACHBEHA: Comportament del professorat							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-2,09	-0,56	-0,12	-0,10	0,61	3,79	0,20%	1011

d) Recursos humans

Taula A1.23

STAFFSHORT: Mancança de personal							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-1,46	-0,18	0,32	0,29	0,84	4,04	1,09%	1011

Taula A1.24

STRATIO: Ràtio estudiant professor							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
1	8,72	11,22	11,72	14,60	51,58	7,22%	1011

Taula A1.25

PROAT6: Professorat amb doctorat							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
0	0	0,02	0,03	0,04	1	17,31%	1011

e) Recursos materials

Taula A1.25

EDUSHORT: Mancança de material							
Mín	q1	Mediana	Mitjana	q3	Màx	Missing	N
-1,42	-0,68	0,10	0,13	0,80	2,96	0,59%	1011

Taula A1.26

Codi	Ítem de la pregunta SC155
SC155Q01	El nombre de dispositius digitals connectats a Internet és suficient.
SC155Q02	L'amplada de banda o la velocitat d'Internet al centre és suficient.
SC155Q03	El nombre de dispositius digitals per a l'ensenyament és suficient.
SC155Q04	Els dispositius digitals del centre tenen prou potència informàtica.
SC155Q05	La disponibilitat de programari adequat és suficient.
SC155Q06	El professorat té les habilitats tècniques i pedagògiques necessàries per integrar dispositius digitals en l'ensenyament.
SC155Q07	El professorat disposa de prou temps per preparar classes que integrin dispositius digitals.
SC155Q08	El professorat té recursos professionals efectius al seu abast per aprendre a utilitzar els dispositius digitals.
SC155Q09	Es disposa d'una bona plataforma de suport a l'aprenentatge en línia.
SC155Q10	El professorat rep incentius per integrar dispositius digitals en la seva docència.
SC155Q11	El centre disposa de suficient personal qualificat d'assistència tècnica.

Taula A1.27

SC155 (inicial)						
Pregunta	Totalment en desacord	En desacord	D'acord	Totalment d'acord	No resposta	N
1	8,45%	32,42%	40,77%	15,24%	3,12%	1089
2	15,06%	25,07%	39,76%	16,62%	3,49%	1089
3	11,11%	40,59%	34,34%	10,56%	3,40%	1089
4	11,94%	34,16%	40,31%	10,28%	3,31%	1089
5	7,62%	32,69%	46,74%	9,64%	3,31%	1089
6	3,95%	39,85%	47,93%	4,68%	3,58%	1089
7	13,96%	48,85%	30,21%	3,67%	3,31%	1089
8	5,97%	34,16%	47,84%	8,91%	3,12%	1089

9	11,29%	33,98%	39,12%	12,03%	3,58%	1089
10	45,73%	39,21%	9,73%	1,74%	3,58%	1089
11	19,28%	35,35%	32,32%	9,73%	3,31%	1089

Taula A1.28

SC155 (recalculada)				
Clúster 1 (1)	Clúster 2 (2)	Clúster 3 (0)	No resposta	N
76,36%	8,90%	14,74%	0%	1011

Taula A1.29

Codi	Ítem de la pregunta SC156
SC156Q01	El seu propi programa sobre l'ús de dispositius digitals.
SC156Q02	El seu propi programa sobre l'ús específic de dispositius digitals amb fins pedagògics.
SC156Q03	Un programa per utilitzar dispositius digitals per a l'ensenyament i l'aprenentatge d'assignatures concretes.
SC156Q04	Reunions regulars amb el personal docent sobre l'ús de dispositius digitals amb fins pedagògics.
SC156Q05	Un programa específic que prepari l'alumnat per a un comportament responsable a Internet.
SC156Q06	Una política específica sobre l'ús de les xarxes socials (Facebook, etc.) en l'ensenyament i l'aprenentatge.
SC156Q07	Un programa específic per promoure la col·laboració en l'ús de dispositius digitals entre el professorat.
SC156Q08	Un temps establert en què el professorat es reuneix per compartir, avaluar o elaborar materials i procediments didàctics que utilitzin els dispositius digitals.

Taula A1.30

SC156 (inicial)				
Pregunta	Sí	No	No resposta	N
1	56,11%	40,40%	3,49%	1089
2	33,24%	63,36%	3,40%	1089
3	35,45%	60,70%	3,86%	1089
4	33,06%	63,18%	3,76%	1089
5	51,52%	44,90%	3,58%	1089
6	47,57%	48,94%	3,49%	1089
7	29,66%	66,48%	3,86%	1089
8	21,40%	75,11%	3,49%	1090

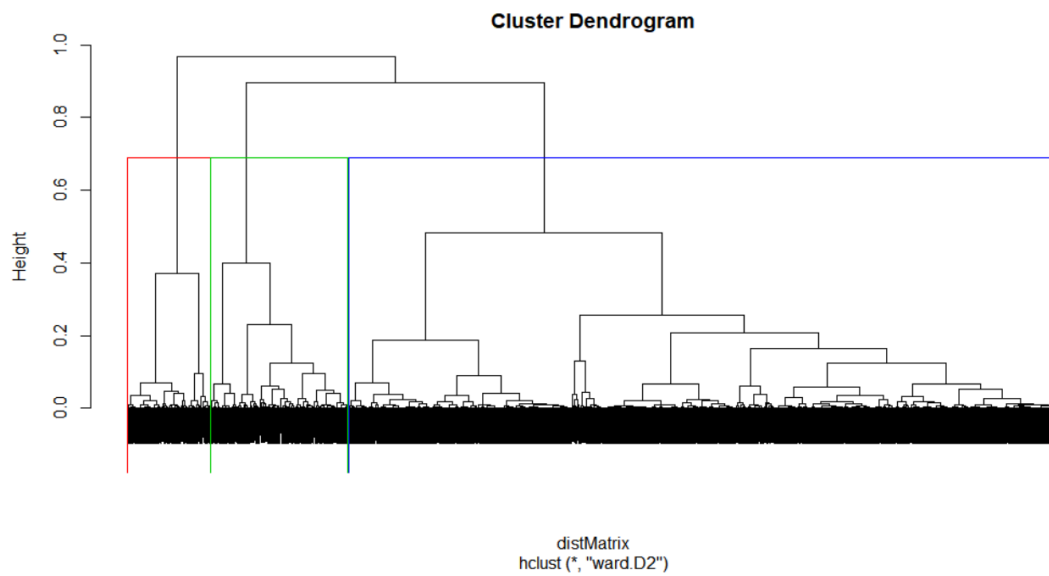
Taula A1.31

SC156 (recalculada)										
0	1	2	3	4	5	6	7	8	No resposta	N
12,0%	12,8%	18,7%	16,2%	12,5%	10,6%	9,2%	4,1%	4,1%	0%	1011

Annex 2: Anàlisi de correspondències múltiples

A la il·lustració A2.1 es mostra el dendrograma resultant de l'Anàlisi de Correspondències Múltiples de la pregunta SC155. La primera classe està representada pel rectangle blau amb la majoria d'observacions (76%), la segona està composta pel 9% de les escoles i correspon al rectangle vermell i la tercera classe, amb el 15% de les escoles, és el rectangle verd.

Il·lustració A2.1 Dendrograma, ACM pregunta SC155



Annex 3: Models

3.1 Grups estudiantat

a) Àmbit personal

Taula A3.1: Comparació models del primer grup de l'estudiantat

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + ST004	33514	1011	2	405910,2
m2: m1 + REPEAT	32052	1011	3	390350,1
m3: m2 + IMMIG	32480	1011	5	383057,4

Il·lustració A3.1: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq   Mean Sq NumDF DenDF   F value   Pr(>F)
ST004  28046908  28046908     1  32012  486.471 < 2.2e-16 ***
REPEAT 626407945 626407945     1  32479 10864.994 < 2.2e-16 ***
IMMIG   7786265   3893133     2  32427   67.526 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

b) Àmbit socio-familiar

Taula A3.2: Comparació models del segon grup de l'estudiantat

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + ESCS	32906	1011	2	395899,1
m2: m1 + HISCED	32539	1011	3	390999
m3: m2 + HEDRES	32523	1011	4	390756,4

Il·lustració A3.2: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq   Mean Sq NumDF DenDF   F value   Pr(>F)
ESCS   164552860 164552860     1  31440 2267.8057 < 2e-16 ***
HISCED 23195933 23195933     1  32294 319.6776 < 2e-16 ***
HEDRES   311587   311587     1  32403   4.2942 0.03825 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c) Àmbit escolar

Taula A3.3: Comparació models del tercer grup de l'estudiantat

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + ATTLNACT	31794	1010	2	384463,5
m2: m1 + TMINS	20250	1000	3	242971,6
m3: m2 + TEACHINT	20167	1000	4	241780,7
m4: m3 + DIRINS	20115	1000	5	240852,1

Il·lustració A3.3: ANOVA model 4

```
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
ATTLNACT  4480978  4480978     1 19926  70.582 < 2.2e-16 ***
TMINS      1299262  1299262     1 19984  20.465 6.107e-06 ***
TEACHINT  17544076 17544076     1 20104 276.344 < 2.2e-16 ***
DIRINS     16177029 16177029     1 19969 254.811 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.2 Grups escoles

a) Tipologia d'escola

Taula A3.4: Comparació models del primer grup de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + SCHLTYPE	32804	989	3	397340,5
m2: m1 + SCHSIZE	30751	928	4	372747,8
m3: m2 + SC001	30714	927	8	372284,7

Il·lustració A3.4: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
SCHLTYPE  12790607 6395303     2  928.18 81.4287 < 2.2e-16 ***
SCHSIZE    2917502  2917502     1  915.88 37.1473 1.612e-09 ***
SC001Q01TA  420300  105075     4  918.21  1.3379    0.254
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

b) Estructura de les escoles

Taula A3.5: Comparació models del segon grup de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + CLSIZE	33232	1003	2	402632,4
m2: m1 + SC042Q01	31855	958	3	385954
m3: m2 + SC042Q02	31029	935	4	375922,4

Il·lustració A3.5: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
          Sum Sq Mean Sq NumDF  DenDF  F value  Pr(>F)
CLSIZE    851020   851020     1  971.97  10.6766  0.001123 **
SC042Q01   373549   373549     1  915.64   4.6864  0.030660 *
SC042Q02TA  64197    64197     1  912.58   0.8054  0.369721
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

c) Valors i compromís

Taula A3.6: Comparació models del tercer grup de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + CREATIV	32988	995	2	399736,5
m2: m1 + STUBEHA	32988	995	3	399552
m3: m2 + TEACHBEHA	32988	995	4	399547,7

Il·lustració A3.6: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
          Sum Sq Mean Sq NumDF  DenDF  F value  Pr(>F)
CREACTIV    44122    44122     1  971.36   0.5546  0.45663
STUBEHA  13321767  13321767     1  989.47  167.4465 < 2e-16 ***
TEACHBEHA   343879   343879     1  985.48   4.3224  0.03787 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

d) Recursos humans

Taula A3.7: Comparació models del quart grup de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + STAFFSHORT	33133	1000	2	401272,1
m2: m1 + STRATIO	30696	928	3	372037,7
m3: m2 + PROAT6	26429	798	4	320209,6

Il·lustració A3.7: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
          Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
STAFFSHORT 3281851 3281851     1 754.52 42.3783 1.372e-10 ***
STRATIO    3408009 3408009     1 777.31 44.0074 6.127e-11 ***
PROAT6      184258  184258     1 795.45  2.3793  0.1233
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

e) Recursos materials

Taula A3.8: Comparació models del cinquè grup de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + EDUSHORT	33133	1000	2	401272,1
m2: m1 + CSC155	30696	928	3	372037,7
m3: m2 + CSC156	26429	798	4	320209,6

Il·lustració A3.8: ANOVA model 3

```
Type III Analysis of Variance Table with Satterthwaite's method
          Sum Sq Mean Sq NumDF   DenDF F value    Pr(>F)
EDUSHORT 2615981 2615981     1  967.08 33.1354 1.154e-08 ***
CSC155    81174   40587     2  958.16  0.5141 0.5982058
CSC156   1071646 1071646     1 1018.99 13.5740 0.0002414 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

3.3 Nivell estudiantat

Taula A3.9: Comparació models del nivell de l'estudiantat

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + grup 1	32480	1011	5	383057,4
m2: m1 + grup 2	31962	1011	8	375759,4
m3: m2 + grup 3	19620	999	12	229190,5

Il·lustració A3.9: ANOVA model 3 amb totes les variables

```

Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq   Mean Sq  NumDF  DenDF   F value    Pr(>F)
ST004    17745767  17745767     1  19391  369.4558 < 2.2e-16 ***
REPEAT   192692247 192692247     1  19605 4011.7329 < 2.2e-16 ***
IMMIG      970170    485085     2  19564  10.0992 4.133e-05 ***
ESCS     27942587   27942587     1  19121 581.7473 < 2.2e-16 ***
HISCED   7394456   7394456     1  19486 153.9480 < 2.2e-16 ***
HEDRES      378      378     1  19580  0.0079 0.9293150
ATTLNACT  531784     531784     1  19572 11.0714 0.0008783 ***
TMINS     81828     81828     1  19598  1.7036 0.1918328
TEACHINT  7268847   7268847     1  19586 151.3329 < 2.2e-16 ***
DIRINS   3857685   3857685     1  19589  80.3146 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Il·lustració A3.10: ANOVA model 3 amb variables descartades

```

Type III Analysis of Variance Table with Satterthwaite's method
      Sum Sq   Mean Sq  NumDF  DenDF   F value    Pr(>F)
ST004    17759311  17759311     1  19395  369.6204 < 2.2e-16 ***
REPEAT   193324064 193324064     1  19607 4023.6082 < 2.2e-16 ***
IMMIG      998803    499401     2  19568  10.3939 3.079e-05 ***
ESCS     31660195   31660195     1  19176 658.9362 < 2.2e-16 ***
HISCED   7608075   7608075     1  19481 158.3451 < 2.2e-16 ***
ATTLNACT  540175     540175     1  19572 11.2425 0.000801 ***
TMINS     74121     74121     1  19601  1.5427 0.214235
TEACHINT  7322095   7322095     1  19590 152.3930 < 2.2e-16 ***
DIRINS   3890221   3890221     1  19591  80.9663 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

3.4 Nivell escola

Taula A3.10: Comparació models del nivell de les escoles

	N estudiantat	N escoles	Paràmetres	Desviància
m0: model nul	33514	1011	1	406006,1
m1: m0 + grup 1	30714	927	8	372284,7
m2: m1 + grup 2	28598	862	11	346659,1
m3: m2 + grup 3	28358	854	14	343707,3
m4: m3 + grup 4	23835	719	17	288715,2
m5: m4 + grup 5	23835	719	21	288713,3

Il·lustració A3.11: ANOVA model 3 amb totes les variables

```

Type III Analysis of Variance Table with Satterthwaite's method
Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
SCHLTYPE 3206790 1603395 2 743.45 20.5063 2.144e-09 ***
SCHSIZE 2001651 2001651 1 679.81 25.5997 5.412e-07 ***
SC001Q01TA 205342 51336 4 728.69 0.6565 0.622392
CREACTIV 75460 75460 1 696.17 0.9651 0.326252
STUBEHA 3045418 3045418 1 721.48 38.9487 7.418e-10 ***
TEACHBEHA 109055 109055 1 702.22 1.3947 0.238006
CLSIZE 37192 37192 1 739.86 0.4757 0.490610
SC042Q01 210331 210331 1 720.79 2.6900 0.101417
SC042Q02TA 49687 49687 1 708.24 0.6355 0.425626
EDUSHORT 42659 42659 1 694.36 0.5456 0.460380
CSC155 44776 22388 2 711.65 0.2863 0.751106
CSC156 78343 78343 1 772.78 1.0020 0.317152
STAFFSHORT 100859 100859 1 716.08 1.2899 0.256444
STRATIO 68194 68194 1 694.48 0.8721 0.350686
PROAT6 718631 718631 1 791.27 9.1908 0.002512 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Il·lustració A3.12: ANOVA model 3 amb variables descartades

```

Type III Analysis of Variance Table with Satterthwaite's method
Sum Sq Mean Sq NumDF DenDF F value Pr(>F)
SCHLTYPE 3499790 1749895 2 770.01 22.6340 2.807e-10 ***
SCHSIZE 2644770 2644770 1 723.22 34.2087 7.498e-09 ***
STUBEHA 3457884 3457884 1 751.21 44.7259 4.425e-11 ***
TEACHBEHA 202238 202238 1 721.46 2.6158 0.1062390
SC042Q01 324667 324667 1 733.52 4.1994 0.0407935 *
CSC156 29327 29327 1 791.43 0.3793 0.5381379
STAFFSHORT 189729 189729 1 735.35 2.4540 0.1176524
STRATIO 98372 98372 1 713.27 1.2724 0.2596977
PROAT6 966245 966245 1 805.02 12.4979 0.0004307 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

3.5 Definitiu

Al cos del treball (apartat 5.2) es mostren els resultats obtinguts un cop haver especificat els deu models pels corresponents deu valors plausibles com a variable endògena. A la il·lustració A3.13 es mostra el model definitiu executat amb el primer valor plausible.

Il·lustració A3.13: Taula resum del model definitiu (amb PV1MATH)

```

Linear mixed model fit by maximum likelihood . t-tests use
Satterthwaite's method [lmerModLmerTest]
Formula:
PV1MATH ~ ST004D01T + REPEAT + ESCS + IMMIG + HISCED + HEDRES +
  TEACHINT + DIRINS + TMINS + ATTLNACT + SCHLTYPE + SCHSIZE +
  STUBEHA + TEACHBEHA + SC042Q01TA + CSC156 + STAFFSHORT +
  STRATIO + PROAT6 + (1 | CNTSCHID)
Data: ddm1
Weights: W_FSTUWT

      AIC      BIC   logLik deviance df.resid
173121.8 173304.2 -86536.9 173073.8    14796

Scaled residuals:
   Min       1Q   Median       3Q      Max
-6.6458 -0.5061  0.0354  0.5451  5.6717

Random effects:
 Groups      Name      Variance Std.Dev.
CNTSCHID (Intercept)  316.8    17.8
Residual            46575.3   215.8
Number of obs: 14820, groups: CNTSCHID, 737

Fixed effects:
              Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  5.245e+02  6.167e+00  1.093e+03  85.042 < 2e-16 ***
ST004D01T1   1.831e+01  1.133e+00  1.467e+04  16.155 < 2e-16 ***
REPEAT1     -8.304e+01  1.491e+00  1.480e+04 -55.695 < 2e-16 ***
ESCS        1.851e+01  9.546e-01  1.479e+04  19.386 < 2e-16 ***
IMMIG2      -2.066e+00  2.709e+00  1.481e+04  -0.763 0.445717
IMMIG3      -9.540e+00  2.482e+00  1.482e+04  -3.843 0.000122 ***
HISCED1     -1.718e+01  1.793e+00  1.473e+04  -9.580 < 2e-16 ***
HEDRES      1.250e-01  6.982e-01  1.478e+04  0.179 0.857896
TEACHINT    5.998e+00  6.371e-01  1.480e+04  9.414 < 2e-16 ***
DIRINS      -5.236e+00  6.998e-01  1.481e+04  -7.482 7.73e-14 ***
TMINS       1.848e-01  8.294e-02  1.481e+04  2.228 0.025873 *
ATTLNACT    2.236e+00  5.973e-01  1.477e+04  3.743 0.000183 ***
SCHLTYPE2   -7.198e+00  3.982e+00  7.331e+02  -1.808 0.071070 .
SCHLTYPE3   -5.496e+00  4.164e+00  7.489e+02  -1.320 0.187347
SCHSIZE     4.416e+00  1.121e+00  6.994e+02  3.940 8.97e-05 ***
STUBEHA     -2.598e+00  1.191e+00  7.298e+02  -2.181 0.029487 *
TEACHBEHA   -8.779e-04  1.265e+00  6.914e+02  -0.001 0.999446
SC042Q01TA1 -3.711e+00  1.982e+00  7.189e+02  -1.872 0.061603 .
CSC156      -1.350e-01  4.636e-01  7.669e+02  -0.291 0.770997
STAFFSHORT  -4.196e-01  1.226e+00  7.098e+02  -0.342 0.732195
STRATIO     -1.647e-01  2.590e-01  6.660e+02  -0.636 0.525101
PROAT6      3.842e+01  1.783e+01  8.323e+02  2.155 0.031455 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```


Annex 4: Codi R

```
### CARREGAR LES DADES ###

# Nivell escoles
read.csv("sch_qqq_sp.csv", header = TRUE, sep = ",")
# Nivell estudiantat
read.csv("stu_qqq_sp.csv", header = TRUE, sep = ",")

### PROCESSAMENT DE LES DADES ###

## Valors mancants
# Descartar les 25 escoles amb no resposta
ddsc[-c(which(ddsc$"SC001Q01TA"=="")),]
# Descartar les observacions (escoles) amb "M" a SC155
dd155orig <- ddsc[, c(colnames(ddsc)[which(startsWith(colnames(ddsc), "SC155
")]))]
M <- c()
for(i in 1:ncol(dd155orig)){
  M <- c(M, which(dd155orig[,i]=="M"))
}
ddsc <- ddsc[-M,]

## Càlcul de noves variables
# Transformacions
ddst$TMINS <- (ddst$TMINS/60)
ddsc$SCHSIZE <- scale(ddsc$SCHSIZE)

# Anàlisi de correspondències múltiples
dd155 <- ddsc[, c(colnames(ddsc)[which(startsWith(colnames(ddsc), "SC155"))
])]

acm155 <- MCA(dd155)
dades_clustering_acm <- acm155$ind$coord
distMatrix <- (daisy(dades_clustering_acm, metric = "gower", stand=TRUE))^2
h1 <- hclust(distMatrix,method="ward.D2")
```

```

suggested.level<-function(hc,min=3,max=10){
  if(min<2) stop("Min should be equal or higher than 2")
  intra <- rev(cumsum(hc$height))
  quot <- intra[min:(max)]/intra[(min - 1):(max - 1)]
  nb.clust = which.min(quot) + min - 1
  return(nb.clust)
}
suggested.level(h1)

plot(h1, labels = FALSE)
rect.hclust(h1, k = suggested.level(h1), border = 2:5) # add rectangle

class155 <- cutree(h1,suggested.level(h1))
ddsc[, (ncol(ddsc)+1)] <- class155
colnames(ddsc)[ncol(ddsc)] <- "CSC155"

### MODELITZACIÓ ###

## Especificació del model jeràrquic

# Unir les bases de dades
PISA2018 <- left_join(ddst, ddsc, by="CNTSCHID")

# Model nul
lmer(PV1MATH ~ (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 1 estudiantat
lmer(PV1MATH ~ ST004 + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)
lmer(PV1MATH ~ ST004 + REPEAT + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)
lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 2 estudiantat
lmer(PV1MATH ~ ESCS + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

```

```

lmer(PV1MATH ~ ESCS + HISCED + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ESCS + HISCED + HEDRES + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 3 estudiantat

lmer(PV1MATH ~ ATTLNACT + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ATTLNACT + TMINS + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ATTLNACT + TMINS + TEACHINT + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ATTLNACT + TMINS + TEACHINT + DIRINS + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 1 escoles

lmer(PV1MATH ~ SCHLTYPE + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 2 escoles

lmer(PV1MATH ~ CLSIZE + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ CLSIZE + SC042Q01 + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ CLSIZE + SC042Q01 + SC042Q02TA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 3 escoles

lmer(PV1MATH ~ CREATIV + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ CREATIV + STUBEHA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ CREATIV + STUBEHA + TEACHBEHA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

# Grup 4 escoles

lmer(PV1MATH ~ STAFFSHORT + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

```

```

lmer(PV1MATH ~ STAFFSHORT + STRATIO + (1|CNTSCHID), data=PISA2018, weights=
W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ STAFFSHORT + STRATIO + PROAT6 + (1|CNTSCHID), data=PISA2018,
weights=W_FSTUWT, REML=FALSE)

# Grup 5 escoles

lmer(PV1MATH ~ EDUSHORT + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, RE
ML=FALSE)

lmer(PV1MATH ~ EDUSHORT + CSC155 + (1|CNTSCHID), data=PISA2018, weights=W_F
STUWT, REML=FALSE)

lmer(PV1MATH ~ EDUSHORT + CSC155 + CSC156 + (1|CNTSCHID), data=PISA2018, we
ights=W_FSTUWT, REML=FALSE)

# Nivell estudiantat (ajuntant grups)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + (1|CNTSCHID), data=PISA2018, weight
s=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + (1|CNTSCHI
D), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT,
REML=FALSE)

# Nivell escoles (ajuntant grups)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + (1|CNTSCHID), data=PISA201
8, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC042Q01 + SC042Q
02TA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC042Q01 + SC042Q
02TA + CREATIV + STUBEHA + TEACHBEHA + (1|CNTSCHID), data=PISA2018, weight
s=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC042Q01 + SC042Q
02TA + CREATIV + STUBEHA + TEACHBEHA + STAFFSHORT + STRATIO + PROAT6 + (1|
CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC042Q01 + SC042Q
02TA + CREATIV + STUBEHA + TEACHBEHA + STAFFSHORT + STRATIO + PROAT6 + EDU
SHORT + CSC155 + CSC156 + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, RE
ML=FALSE)

# Agregació de nivells (a l'estudiantat, afegir grups de les escoles)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT,
REML=FALSE)

```

```

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + SCHLTYPE + SCHSIZE + SC001Q01TA + (1|CNTSCHID),
data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC04
2Q01 + SC042Q02TA + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC04
2Q01 + SC042Q02TA + CREATIV + STUBEHA + TEACHBEHA + (1|CNTSCHID), data=PISA
A2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC04
2Q01 + SC042Q02TA + CREATIV + STUBEHA + TEACHBEHA + STAFFSHORT + STRATIO +
PROAT6 + (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED + HEDRES + ATTLNACT +
TMINS + TEACHINT + DIRINS + SCHLTYPE + SCHSIZE + SC001Q01TA + CLSIZE + SC04
2Q01 + SC042Q02TA + CREATIV + STUBEHA + TEACHBEHA + STAFFSHORT + STRATIO +
PROAT6 + EDUSHORT + CSC155 + CSC156 + (1|CNTSCHID), data=PISA2018, weights=
W_FSTUWT, REML=FALSE)

## Model definitiu

lmer(PV1MATH ~ ST004 + REPEAT + ESCS + IMMIG + HISCED + HEDRES + TEACHINT +
DIRINS + TMINS + ATTLNACT + SCHLTYPE + SCHSIZE + STUBEHA + TEACHBEHA + SC04
2Q01 + CSC156 + STAFFSHORT + STRATIO + PROAT6 + (1|CNTSCHID), data=PISA2018
, weights=W_FSTUWT, REML=FALSE)

## Models analitzats al cos del treball

model.nul <- lmer(PV1MATH ~ (1|CNTSCHID), data=PISA2018, weights=W_FSTUWT,
REML=FALSE)

model.estudiantat <- lmer(PV1MATH ~ ST004 + REPEAT + IMMIG + ESCS + HISCED
+ ATTLNACT + TMINS + TEACHINT + DIRINS + (1|CNTSCHID), data=PISA2018, weigh
ts=W_FSTUWT, REML=FALSE)

model.escoles <- lmer(PV1MATH ~ SCHLTYPE + SCHSIZE + STUBEHA + TEACHBEHA +
SC042Q01 + CSC156 + STAFFSHORT + STRATIO + PROAT6 + (1|CNTSCHID), data=PISA
2018, weights=W_FSTUWT, REML=FALSE)

model.complet <- lmer(PV1MATH ~ ST004 + REPEAT + ESCS + IMMIG + HISCED + HE
DRES + TEACHINT + DIRINS + TMINS + ATTLNACT + SCHLTYPE + SCHSIZE + STUBEHA
+ TEACHBEHA + SC042Q01 + CSC156 + STAFFSHORT + STRATIO + PROAT6 + (1|CNTSCH
ID), data=PISA2018, weights=W_FSTUWT, REML=FALSE)

```