ANNEX B – CONTRAST CODING USED IN CONDITIONING

Table B.1 PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Grade	ST01Q01	7-14 Ungraded Missing	value – mode 0 0 0 0 0 0 1 1
Study programme	ST02Q01	National categories	If there is at least one school with more than one study programme in a country, national study programmes are dummy coded with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Age of student	AGE	Value (decimal) Missing	value – median 0 0 1
Gender	ST04Q01	1 Female 2 Male Missing	Two dummies if missing data is present and one dummy if no missing data with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
ISCED 0	ST05Q01	1 No 2 Yes, one year or less 3 yes, more than one year Missing (or invalid)	Three dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Age when started ISCED 1	ST06Q01	Value Missing	value – median 0 0 1
Repeated grade at ISCED 1	ST07Q01	1 No 2 Yes, once 3 Yes, twice or more Missing (or invalid)	Three dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Repeated grade at ISCED 2	ST07Q02	1 No 2 Yes, once 3 Yes, twice or more Missing (or invalid)	Three dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Repeated grade at ISCED 3	ST07Q03	1 No 2 Yes, once 3 Yes, twice or more Missing (or invalid)	Three dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Arrived late for school	ST08Q01	1 None 2 One or two times 3 Three or four times 4 Five or more times Missing (or invalid)	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Skipped the whole school day	ST09Q01	1 None 2 One or two times 3 Three or four times 4 Five or more times Missing (or invalid)	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Skipped some classes	ST115Q01	1 None 2 One or two times 3 Three or four times 4 Five or more times Missing (or invalid)	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Mother	ST11Q01	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Father	ST11Q02	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Brother(s)	ST11Q03	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Sister(s)	ST11Q04	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Grandparent(s)	ST11Q05	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Lives at home with you – Other(s)	ST11Q06	1 Yes 2 No Missing (or invalid)	Two dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Occupational status of Mother (SEI)	BMMJ1	16-90 (decimal) Missing	value – median 0 0 1
Occupational status of Father (SEI)	BFMJ2	16-90 (decimal) Missing	value-median 0 0 1
Educational level of Mother (MISCED)	ST13Q01 ST14Q01 ST14Q02 ST14Q03 ST14Q04	5 None 4 ISCED 1 3 ISCED 2 2 ISCED 3B, C 1 ISCED 3A, Missing 1 Yes 2 No Missing	Item ST13Q01 was recoded as (5=0), (4=1), (3=2), (2=3), (3=4). Item ST14Q04 was recoded as (1=4), (2=0) Item ST14Q03 was recoded as (1=5), (2=0) Item ST14Q02 was recoded as (1=5), (2=0) Item ST14Q01 was recoded as (1=6), (2=0). New variable MISCED was created as maximum value of five items, thus having categories from 0 to 6. Plus one category for missing (when all five items are missing) Seven dummy variables were created based on the value of MISCED and with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)

[Part 2/6]

Table B.1 **PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables**

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Educational level of Father (FISCED)	ST17Q01	5 None 4 ISCED 1 3 ISCED 2 2 ISCED 3B, C 1 ISCED 3A, Missing	Item ST17Q01 was recoded as (5=0), (4=1), (3=2), (2=3), (3=4). Item ST18Q04 was recoded as (1=4), (2=0) Item ST18Q03 was recoded as (1=5), (2=0) Item ST18Q02 was recoded as (1=5), (2=0) Item ST18Q01 was recoded as (1=6), (2=0). New variable FISCED was created as maximum value of five items, thus having categories from 0 to 6. Plus one category for missing (when all five items are missing) Seven dummy variables were created based on the value of FISCED and with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
	ST18Q01 ST18Q02 ST18Q03 ST18Q04	1 Yes 2 No Missing	
What Mother is currently doing	ST15Q01	1 Working full-time 2 Working part-time 3 Not working, looking 4 Other Missing (or invalid)	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
What Father is currently doing	ST19Q01	1 Working full-time 2 Working part-time 3 Not working, looking 4 Other Missing (or invalid)	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Immigration status (IMMIG)	ST20int (CTSELF) (CTFATHER) (CTMOTHER)	1 Native 2 Second-Generation 3 First-Generation Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Country arrival age	ST21Q01	Value N/A (born in country) Missing (or >17)	(copy) 0 0 0 0 -1
Language at home	ST25int	1 Language of test 2 Other language Missing	-1 -1 01 00 00 01
Family wealth (WEALTH)	ST26Q02 ST26Q06 ST26Q13 ST26Q14 ST26Q15 ST26Q16 ST26Q17 ST27Q01 ST27Q02 ST27Q03 ST27Q04 ST27Q05	1 Yes 2 No Missing 1 None 2 One 3 Two 4 Three or more Missing	All items of Q26 were recoded as (Yes=1, No=0) and all items of Q27 were recoded as (1=0, 2=1, 3=2, 4=3). Total score was calculated as a ratio of a sum of all items over maximum score of valid responses (items with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Home educational resources (HEDRES)	ST26Q01 ST26Q03 ST26Q04 ST26Q05 ST26Q10 ST26Q11 ST26Q12	1 Yes 2 No Missing	All items were recoded as (Yes=1, No=0). Total score was calculated as a ratio of a sum of all items over maximum score of valid responses (items with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Cultural possessions at home (CULTPOS)	ST26Q07 ST26Q08 ST26Q09	1 Yes 2 No Missing	All items were recoded as (Yes=1, No=0). Total score was calculated as a ratio of a sum of all items over maximum score of valid responses (items with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
How many books at home	ST28Q01	1 0-10 books 2 11-25 books 3 26-100 books 4 101-200 books 5 201-500 books 6 More than 500 books Missing	Six dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Mathematics interest (INTMAT)	ST29Q01 ST29Q03 ST29Q04 ST29Q06	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (4=0),(3=1),(2=2),(1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Instrumental motivation for mathematics (INSTMOT)	ST29Q02 ST29Q05 ST29Q07 ST29Q08	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Subjective norms in mathematics (SUBNORM)	ST35Q01 ST35Q02 ST35Q03 ST35Q04 ST35Q05 ST35Q06	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Mathematics self-efficacy (MATHEFF)	ST37Q01 ST37Q02 ST37Q03 ST37Q04 ST37Q05 ST37Q06 ST37Q07 ST37Q08	1 Very confident 2 Confident 3 Not very confident 4 Not at all confident Missing	Items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 3/6]

PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables

Table B.1

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Mathematics anxiety (ANXMAT)	ST42Q01 ST42Q03 ST42Q05 ST42Q08 ST42Q10	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (4=0), (3=1), (2=2), (1=3).Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Mathematics self-concept (SCMAT)	ST42Q02 ST42Q04 ST42Q06 ST42Q07 ST42Q09	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items except item 02 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Item 02 was coded as (1=0), (2=1), (3=2), (4=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Perceived control of success in mathematics	ST43Q01 ST43Q02 ST43Q03 ST43Q04 ST43Q05 ST43Q06	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items 01, 02, 05 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Items 03, 04, 06 were coded as (1=0), (2=1), (3=2), (4=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Attributions to failure in mathematics (FAILMAT)	ST44Q01 ST44Q03 ST44Q04 ST44Q05 ST44Q07 ST44Q08	1 Very Likely 2 Likely 3 Slightly likely 4 Not at all likely Missing	Items were coded as (1=0), (2=1), (3=2), (4=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Mathematics work ethic (MATWKETH)	ST46Q01 ST46Q02 ST46Q03 ST46Q04 ST46Q05 ST46Q06 ST46Q07 ST46Q08 ST46Q09	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Mathematics intentions (MATINTFC)	ST48Q01 ST48Q02 ST48Q03 ST48Q04 ST48Q05	1 Courses after school -Maths 2 Courses after school -Test Language Missing 1 Major in college - Math 2 Major in college - Science Missing 1 Study harder - Math 2 Study harder - Test Language Missing 1 Maximum classes - Math 2 Maximum classes - Science Missing 1 Pursuing a career - Math 2 Pursuing a career -Science Missing	All items were reversely recoded as (2=0), (1=1). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Mathematics behaviour (MATBEH)	ST49Q01 ST49Q02 ST49Q03 ST49Q04 ST49Q05 ST49Q06 ST49Q07 ST49Q09	1 Always or almost always 2 Often 3 Sometimes 4 Never or rarely Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Learning strategies	ST53Q01 ST53Q02 ST53Q03 ST53Q04	Choice of three strategies 1 First strategy 2 second strategy 3 third strategy No response (Missing)	Items 01 and 02 were recoded as ('1'=1), ('2'=0), ('3'=0). Item 03 was recoded as ('1'=0), ('2'=1), ('3'=0). Item 04 was recoded as ('1'=0), ('2'=0), ('3'=1). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Out of school lessons in - test language (hours a week)	ST55Q01	1 'I do not attend' 2 'Less than 2 hours' 3 'Between 2 and 4 hours' 4 'Between 4 and 6 hours' 5 '6 or more hours' Missing	The item was recoded as ('1'=0), ('2'=1), ('3'=3), ('4'=5), ('5'=7), representing an approximate time in hours. Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Out of school lessons in mathematics (hours)	ST55Q02	Value Missing	The item was recoded as ('1'=0), ('2'=1), ('3'=3), ('4'=5), ('5'=7), representing an approximate time in hours. Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Out of school lessons in - science (hours)	ST55Q03	Value Missing	The item was recoded as ('1'=0), ('2'=1), ('3'=3), ('4'=5), ('5'=7), representing an approximate time in hours. Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Out of school lessons in - other subjects (hours)	ST55Q04	Value Missing	The item was recoded as ('1'=0), ('2'=1), ('3'=3), ('4'=5), ('5'=7), representing an approximate time in hours. Two dummy variable were created as follows: Value value – median 0 Missing 0 1

[Part 4/6]

Table B.1 PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Homework set by teacher (hours)	ST57Q01	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
How many hour of homework with somebody overlooking and helping	ST57Q02	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Work with a personal <tutor>	ST57Q03	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Classes through a commercial company	ST57Q04	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Study with parents or family	ST57Q05	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Train school lessons on a computer	ST57Q06	Value Missing	Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Experience with applied mathematics tasks at school (EXAPPLM)	ST61Q01 ST61Q02 ST61Q03 ST61Q04 ST61Q06 ST61Q08	1 Frequently 2 Sometimes 3 Rarely 4 Never Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Experience with pure mathematics tasks at school (EXPUREM)	ST61Q05 ST61Q07 ST61Q09	1 Frequently 2 Sometimes 3 Rarely 4 Never Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Experience with applied mathematics tasks at school (FAMCON)	ST62Q01 ST62Q02 ST62Q03 ST62Q06 ST62Q07 ST62Q08 ST62Q09 ST62Q10 ST62Q12 ST62Q15 ST62Q16 ST62Q17 ST62Q19	1 Never heard of it 2 Heard of it once or twice 3 Heard of it a few times 4 Heard of it often 5 Know it well Missing	All items were coded as (1=0), (2=1), (3=2), (4=3), (5=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Experience with applied mathematics tasks at school (FAMCONC)	ST62Q04 ST62Q11 ST62Q12	1 Never heard of it 2 Heard of it once or twice 3 Heard of it a few times 4 Heard of it often 5 Know it well Missing	All items were coded as (1=0), (2=1), (3=2), (4=3), (5=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Average time per week on <language> (LMINS)	ST69Q01 ST70Q01	Value Missing	The value is the product of ST69Q01*ST70Q01. Two dummy variable were created as follows: Value value – mean 0 Missing 0 1
Average time per week on mathematics (MMINS)	ST69Q02 ST70Q02	Value Missing	The value is the product of ST69Q02*ST70Q02. Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Average time per week on science (SMINS)	ST69Q03 ST70Q03	Value Missing	The value is the product of ST69Q03*ST70Q03. Two dummy variable were created as follows: Value value – median 0 Missing 0 1
Total number of classes per week	ST71Q01	Value Missing	value – median 0 0 1
Number of students attending <language> class	ST72Q01	Value Missing	value – median 0 0 1
Experience with these types of problems at school	ST73Q01 ST73Q02 ST74Q01 ST74Q02 ST75Q01 ST75Q02 ST76Q01 ST76Q02	1 Frequently 2 Sometimes 3 Rarely 4 Never Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Teacher support (TEACHSUP)	ST77Q01 ST77Q02 ST77Q04 ST77Q05 ST77Q06	1 Every lesson 2 Most lessons 3 Some lessons 4 Never or hardly ever Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Teacher behaviour: teacher-directed instruction (TCHBEHTD)	ST79Q01 ST79Q02 ST79Q06 ST79Q08 ST79Q15	1 Every lesson 2 Most lessons 3 Some lessons 4 Never or hardly ever Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Teacher behaviour: student orientation (TCHBEHSO)	ST79Q03 ST79Q04 ST79Q07 ST79Q10	1 Every lesson 2 Most lessons 3 Some lessons 4 Never or hardly ever Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 5/6]

PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables

Table B.1

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Teacher behaviour: formative assessment (TCHBEHFA)	ST79Q05 ST79Q11 ST79Q12 ST79Q17	1 Every lesson 2 Most lessons 3 Some lessons 4 Never or hardly ever Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Cognitive activation in mathematics lessons (COGACT)	ST80Q01 ST80Q04 ST80Q05 ST80Q06 ST80Q07 ST80Q08 ST80Q09 ST80Q10 ST80Q11	1 Always or almost always 2 Often 3 Sometimes 4 Never or rarely Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Disciplinary climate (DISCLIMA)	ST81Q01 ST81Q02 ST81Q03 ST81Q04 ST81Q05	1 Every lesson 2 Most lessons 3 Some lessons 4 Never or hardly ever Missing	All items were coded as (1=0), (2=1), (3=2), (4=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Mathematics teacher's support anchoring vignettes (ANCMTSUP)	ST82Q01 ST82Q02 ST82Q03	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Mathematics teacher's support (MTSUP)	ST83Q01 ST83Q02 ST83Q03	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Mathematics teacher's classroom management anchoring vignettes (ANCCLSMAN)	ST84Q01 ST84Q02 ST84Q03	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Mathematics teacher's classroom management (CLSMAN)	ST85Q01 ST85Q02 ST85Q03 ST85Q04	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items 01, 02, 03 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Item 04 was coded as (4=3), (3=2), (2=1), (1=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Teacher-student relations (STUDREL)	ST86Q01 ST86Q02 ST86Q03 ST86Q04 ST86Q05	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Sense of belonging to school (BELONG)	ST87Q01 ST87Q02 ST87Q03 ST87Q04 ST87Q05 ST87Q06 ST87Q07 ST87Q08 ST87Q09	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items 01, 04 and 06 were coded as (4=3), (3=2), (2=1), (1=0). Items 02, 03, 05, 07, 08 and 09 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Attitude towards school: learning outcomes (ATSCHL)	ST88Q01 ST88Q02 ST88Q03 ST88Q04	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items 03 and 04 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Items 01 and 02 were coded as (4=3), (3=2), (2=1), (1=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Attitude towards school: learning activities (ATTLNACT)	ST89Q02 ST89Q03 ST89Q04 ST89Q05	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Perceived control of success at school	ST91Q01 ST91Q02 ST91Q03 ST91Q04 ST91Q05 ST91Q06	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items 01, 02 and 05 were reversely recoded as (4=0), (3=1), (2=2), (1=3). Items 03, 04 and 06 were coded as (4=3), (3=2), (2=1), (1=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Students' perseverance (PERSEV)	ST93Q01 ST93Q03 ST93Q04 ST93Q06 ST93Q07	1 Very much like me 2 Mostly like me 3 Somewhat like me 4 Not much like me 5 Not at all like me Missing	Items 04, 06 and 07 were reversely recoded as (5=0), (4=1), (3=2), (2=3), (1=4). Items 01 and 03 were coded as (5=4), (4=3), (3=2), (2=1), (1=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1
Openness for problem solving (OPENPS)	ST94Q05 ST94Q06 ST94Q09 ST94Q10 ST94Q14	1 Very much like me 2 Mostly like me 3 Somewhat like me 4 Not much like me 5 Not at all like me Missing	All items were reversely recoded as (5=0), (4=1), (3=2), (2=3), (1=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value — value – mean — 0; Missing — 0 — 1

[Part 6/6]

PISA 2012 Main Survey contrast coding used in conditioning for the student questionnaire variables

Table B.1

Variable	Variable name	Variable coding	Contrast coding
STUDENT QUESTIONNAIRE			
Problem-sloving strategy: systematic strategies	ST96Q02 ST101Q01 ST101Q02 ST104Q01	1 I would definitely do this 2 I would probably do this 3 I would probably not do this 4 I would definitely not do this Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Problem-sloving strategy: unsystematic strategies	ST96Q01 ST101Q03 ST101Q05 ST104Q04	1 I would definitely do this 2 I would probably do this 3 I would probably not do this 4 I would definitely not do this Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Problem-sloving strategy: seeking help	ST96Q03 ST96Q05 ST104Q05 ST104Q06	1 I would definitely do this 2 I would probably do this 3 I would probably not do this 4 I would definitely not do this Missing	All items were reversely recoded as (4=0), (3=1), (2=2), (1=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 1/2]

PISA 2012 Main Survey contrast coding used in conditioning for the ICT questionnaire variables

Table B.2

Variable	Variable name	Variable coding	Contrast coding
ICT QUESTIONNAIRE			
ICT availability at home (ICTHOME)	IC01Q01 IC01Q02 IC01Q03 IC01Q04 IC01Q05 IC01Q06 IC01Q07 IC01Q08 IC01Q09 IC01Q10 IC01Q11	1 Yes, and I use it 2 Yes, but I don't use it 3 No Missing	Items were reversely recoded as (3=0), (2=1), (1=2). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
ICT availability at school (ICTSCH)	IC02Q01 IC02Q02 IC02Q03 IC02Q04 IC02Q05 IC02Q06 IC02Q07	1 Yes, and I use it 2 Yes, but I don't use it 3 No Missing	Items were reversely recoded as (3=0), (2=1), (1=2). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
How old were you when you first used a computer	IC03Q01	1 6 y/o or younger 2 Between 7 and 9 y/o 3 Between 10 and 12 y/o 4 13 y/o or older 5 Never used Missing (or invalid)	Five dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
How old were you when you first accessed the internet	IC04Q01	1 6 y/o or younger 2 Between 7 and 9 y/o 3 Between 10 and 12 y/o 4 13 y/o or older 5 Never used Missing (or invalid)	Five dummies with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Internet use at school on a typical weekday	IC05Q01	01 No time 02 1-30 minutes 03 31-60 minutes 04 Between 1 and 2 hours 05 Between 2 and 4 hours 06 Between 4 and 6 hours 07 More than 6 hours Missing (or invalid)	Items were recoded to represent numerical value in half-hours as follows ('01'=0), ('02'=1), ('03'=2), ('04'=4), ('05'=8), ('06'=12), ('07'=18). Two dummy variables were created as follows: Value value – median 0 Missing 0 1
Internet use outside of school on a typical weekday	IC06Q01	01 No time 02 1-30 minutes 03 31-60 minutes 04 Between 1 and 2 hours 05 Between 2 and 4 hours 06 Between 4 and 6 hours 07 More than 6 hours Missing (or invalid)	Items were recoded to represent numerical value in half-hours as follows ('01'=0), ('02'=1), ('03'=2), ('04'=4), ('05'=8), ('06'=12), ('07'=18). Two dummy variables were created as follows: Value value – median 0 Missing 0 1
Internet use on a typical weekend day	IC07Q01	01 No time 02 1-30 minutes 03 31-60 minutes 04 Between 1 and 2 hours 05 Between 2 and 4 hours 06 Between 4 and 6 hours 07 More than 6 hours Missing (or invalid)	Items were recoded to represent numerical value in half-hours as follows ('01'=0), ('02'=1), ('03'=2), ('04'=4), ('05'=8), ('06'=12), ('07'=18). Two dummy variables were created as follows: Value value – median 0 Missing 0 1
ICT entertainment use (ENTUSE)	IC08Q01 IC08Q02 IC08Q03 IC08Q04 IC08Q05 IC08Q06 IC08Q07 IC08Q08 IC08Q09 IC08Q11	1 Never or hardly ever 2 Once or twice a month 3 Once or twice a week 4 Almost every day 5 Every day Missing	Items were coded as (1=0), (2=1), (3=2), (4=3), (5=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 2/2]

PISA 2012 Main Survey contrast coding used in conditioning for the ICT
Table B.2 questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
ICT QUESTIONNAIRE			
ICT use at home for school related tasks (HOMSCH)	IC09Q01 IC09Q02 IC09Q03 IC09Q04 IC09Q05 IC09Q06 IC09Q07	1 Never or hardly ever 2 Once or twice a month 3 Once or twice a week 4 Almost every day 5 Every day Missing	Items were coded as (1=0), (2=1), (3=2), (4=3), (5=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Use of ICT for school (USESCH)	IC10Q01 IC10Q02 IC10Q03 IC10Q04 IC10Q05 IC10Q06 IC10Q07 IC10Q08 IC10Q09	1 Never or hardly ever 2 Once or twice a month 3 Once or twice a week 4 Almost every day 5 Every day Missing	Items were coded as (1=0), (2=1), (3=2), (4=3), (5=4). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Use of computer in mathematics lessons by students (USEMATH1)	IC11Q01 IC11Q02 IC11Q03 IC11Q04 IC11Q05 IC11Q06 IC11Q07	1 Yes, students did this 2 Yes, but only the teacher demonstrated this 3 No Missing	Items were recoded as (1=1), (2=0), (3=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Use of computer in mathematics lessons by teacher demonstrating only (USEMATH2)	IC11Q01 IC11Q02 IC11Q03 IC11Q04 IC11Q05 IC11Q06 IC11Q07	1 Yes, students did this 2 Yes, but only the teacher demonstrated this 3 No Missing	Items were recoded as (1=0), (2=1), (3=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Attitudes towards computers: computer as a tool for school learning (ICTATTPOS)	IC22Q01 IC22Q02 IC22Q04	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (1=3), (2=2), (3=1), (4=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Attitudes towards computers: limitations of the computer as a tool for school learning (ICTATTNEG)	IC22Q06 IC22Q07 IC22Q08	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were coded as (1=0), (2=1), (3=2), (4=3). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 1/3]

PISA 2012 Main Survey contrast coding used in conditioning for the educational career
Table B.3 questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
EDUCATIONAL CAREER QUESTIONNAIRE			
Did you ever miss two or more consecutive months of ISCED 1	EC01Q01	1 No, never 2 Yes, once 3 Yes, twice or more Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Did you ever miss two or more consecutive months of ISCED 2	EC02Q01	1 No, never 2 Yes, once 3 Yes, twice or more Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Information about careers (INFOCAR)	EC03Q01 EC03Q02 EC03Q03 EC03Q04 EC03Q05 EC03Q06 EC03Q07 EC03Q08 EC03Q09 EC03Q10	1 Yes 2 No, never Missing	Items were reversely recoded as (1=1), (2=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Skills acquired at school (INFOJOB1) – How to find info on jobs – How to search for a job – How to write a resume – How to prepare for a job interview – How to find info on <ISCED 3-5> – How to find info on student grants	EC04Q01A EC04Q02A EC04Q03A EC04Q04A EC04Q05A EC04Q06A	1 Tick 2 No tick	Items were reversely recoded as (1=1), (2=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Skills acquired out of school (INFOJOB2) – How to find info on jobs – How to search for a job – How to write a resume – How to prepare for a job interview – How to find info on <ISCED 3-5> – How to find info on student grants	EC04Q01B EC04Q02B EC04Q03B EC04Q04B EC04Q05B EC04Q06B	1 Tick 2 No tick	Items were reversely recoded as (1=1), (2=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 2/3]

PISA 2012 Main Survey contrast coding used in conditioning for the educational career questionnaire variables

Table B.3

Variable	Variable name	Variable coding	Contrast coding
EDUCATIONAL CAREER QUESTIONNAIRE			
The first language learned at home	EC05Q01	1 Test language or other official dialect 2 Test language or other official dialect and another language 3 Language other than the test or other official language Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Age when started learning test language	EC06Q01	1 Between 0 and 3 y/o 2 Between 4 and 6 y/o 3 Between 7 and 9 y/o 4 Between 10 and 12 y/o 5 13 y/o or older Missing (or invalid)	Items were recoded to represent numerical value in years as follows ('1'=1), ('2'=5), ('3'=8), ('4'=11), ('5'=14). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Language spoken with my mother	EC07Q01	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language spoken with my father	EC07Q02	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language spoken with my brother(s)/sister(s)	EC07Q03	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language spoken with my best friend	EC07Q04	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language spoken with my schoolmates	EC07Q05	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language usually used for reading books, magazines, newspapers	EC08Q01	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language usually used for watching TV or movies	EC08Q02	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language usually used for surfing the internet	EC08Q03	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Language usually used for writing e-mails or letters	EC08Q04	1 Mostly my heritage language 2 About equally heritage and test languages 3 Mostly test language 4 Not applicable Missing	Four dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Attended <remedial lessons> in test language	EC09Q03	1 Yes 2 No, never Missing	Two dummy variables with default value of '00' and - national mode = '-1' in both dummies - corresponding category = '01' (including missing)
Attended lessons in my heritage language	EC11Q02	1 Yes 2 No, never Missing	Two dummy variables with default value of '00' and - national mode = '-1' in both dummies - corresponding category = '01' (including missing)
Attended instruction in school subjects through my heritage language	EC11Q03	1 Yes 2 No, never Missing	Two dummy variables with default value of '00' and - national mode = '-1' in both dummies - corresponding category = '01' (including missing)
Hours per week receiving systematic support for improving test language	EC10Q01	1 None 2 Less than 2 3 2 or more but less than 4 4 4 or more but less than 6 5 6 or more Missing (or invalid)	Items were coded to represent numerical value in hours per week as follows ('1'=0), ('2'=1), ('3'=2), ('4'=3), ('5'=5). Two dummy variables were created as follows: Value value – median 0 Missing 0 1

[Part 3/3]

Table B.3 PISA 2012 Main Survey contrast coding used in conditioning for the educational career questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
EDUCATIONAL CAREER QUESTIONNAIRE			
Hours per attending either lessons in heritage language or instructions through heritage language	EC12Q01	1 None 2 Less than 2 3 2 or more but less than 4 4 4 or more but less than 6 5 6 or more Missing (or invalid)	Items were coded to represent numerical value in hours per week as follows ('1'=0), ('2'=1), ('3'=2), ('4'=3), ('5'=5). Two dummy variables were created as follows: Value value – median 0 Missing 0 1
Mother born in country of test	ST22Q01	1 No 2 Yes Missing (or invalid)	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Acculturation: host culture oriented strategies (HOSTCUL)	ST23Q01 ST23Q03 ST23Q05 ST23Q07	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (1=3), (2=2), (3=1), (4=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Acculturation: heritage culture oriented strategies (HERITCUL)	ST23Q02 ST23Q04 ST23Q06 ST23Q08	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (1=3), (2=2), (3=1), (4=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1
Cultural distance between host and heritage culture (CULTDIST)	ST24Q01 ST24Q02 ST24Q03	1 Strongly agree 2 Agree 3 Disagree 4 Strongly disagree Missing	Items were reversely recoded as (1=3), (2=2), (3=1), (4=0). Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows: Value value – mean 0 Missing 0 1

[Part 1/3]

Table B.4 PISA 2012 Main Survey contrast coding used in conditioning for the parent questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
PARENT QUESTIONNAIRE			
Who will complete this questionnaire – Mother or female guardian	PA01Q01	1 Tick 2 No tick Missing (or invalid)	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Who will complete this questionnaire – Father or male guardian	PA01Q02	1 Tick 2 No tick Missing (or invalid)	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Who will complete this questionnaire – other	PA01Q03	1 Tick 2 No tick Missing (or invalid)	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Age of Father	PA02Q01	1 Younger than 36 y/o 2 Between 36 and 40 y/o 3 Between 41 and 45 y/o 4 Between 46 and 50 y/o 5 51 y/o or older Missing (or invalid)	Items were coded as follows ('1'=0), ('2'=1), ('3'=2), ('4'=3), ('5'=5). Two dummy variables were created as follows: Value value – median 0 Missing 0 1
Does the child's Father have any of the following qualifications - ISCED 5A, 6 - ISCED 5B - ISCED 4 - ISCED 3A (PQFISCED)	PA03Q01 PA03Q02 PA03Q03 PA03Q04	1 Yes 2 No Missing	Item PA03Q04 was recoded as (1=1), (2=0) Item PA03Q03 was recoded as (1=1), (2=0) Item PA03Q02 was recoded as (1=2), (2=0) Item PA03Q01 was recoded as (1=3), (2=0). New variable PQFISCED was created as maximum value of four items, thus having categories from 0 to 3. Plus one category for missing (when all four items are missing) Four dummy variables were created based on the value of PQFISCED and with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Occupational status Father, parents answer (SEI)	BFMJ4 (based on PA04Q01)	16-90(decimal) Missing	value – median 0 0 1
Does the child's Mother have any of the following qualifications - ISCED 5A, 6 - ISCED 5B - ISCED 4 - ISCED 3A (PQMISCED)	PA05Q01 PA05Q02 PA05Q03 PA05Q04	1 Yes 2 No Missing	Item PA05Q04 was recoded as (1=1), (2=0) Item PA05Q03 was recoded as (1=1), (2=0) Item PA05Q02 was recoded as (1=2), (2=0) Item PA05Q01 was recoded as (1=3), (2=0). New variable PQMISCED was created as maximum value of four items, thus having categories from 0 to 3. Plus one category for missing (when all four items are missing) Four dummy variables were created based on the value of PQMISCED and with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Occupational status Mother, parents answer (SEI)	BMMJ3 (based on PA06Q01)	16-90(decimal) Missing	value-median 0 0 1
Annual household income	PA07Q01	1 Less than <$A> 2 <$A> or more and less <$B> 3 <$B> or more and less <$C> 4 <$C> or more and less <$D> 5 <$D> or more and less <$E> 6 <$E> or more Missing	Items were coded as (6=5), (5=4), (4=3), (3=2), (2=1), (1=0). Two dummy variable were created as follows: Value value – median 0 Missing 0 1

[Part 2/3]

PISA 2012 Main Survey contrast coding used in conditioning for the parent
Table B.4   questionnaire variables

Variable	Variable name	Variable coding	Contrast coding
PARENT QUESTIONNAIRE			

Paid to educational providers in the last year	PA08Q01	1 Nothing
		2 More than $0 and less <$W>
		3 <$W> or more and less <$X>
		4 <$X> or more and less <$Y>
		5 <$Y> or more and less <$Z>
		6 <$Z> or more
		Missing
			Items were coded as (6=5), (5=4), (4=3), (3=2), (2=1), (1=0).
			Two dummy variable were created as follows:
			  Value value – median 0
			  Missing 0 1

Parents perception of school quality (PQSCHOOL)	PA09Q01
	PA09Q02
	PA09Q03
	PA09Q04
	PA09Q05
	PA09Q06
	PA09Q07
		1 Strongly agree
		2 Agree
		3 Disagree
		4 Strongly disagree
		Missing
			All items were coded as (1=3), (2=2), (3=1), (4=0).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Parental involvement in their child's school (PARINVOL)	PA10Q01
	PA10Q02
	PA10Q03
	PA10Q04
	PA10Q05
	PA10Q06
	PA10Q07
	PA10Q08
	PA10Q09
	PA10Q10
	PA10Q11
		1 Yes
		2 No
		Missing
			Items were recoded as (1=1), (2=0).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Which of the following best describes the schooling available to students in your location	PA11Q01	1 Two or more other schools
		2 One other school
		3 No other schools
		Missing
			All items were coded as (1=2), (2=1), (3=0).
			Two dummy variable were created as follows:
			  Value value – median 0
			  Missing 0 1

Parent school selection	PA12Q01
	PA12Q02
	PA12Q03
	PA12Q04
	PA12Q05
	PA12Q06
	PA12Q07
	PA12Q08
	PA12Q09
	PA12Q10
	PA12Q11
		1 Not important
		2 Somewhat important
		3 Important
		4 Very Important
		Missing
			All items were coded as (1=3), (2=2), (3=1), (4=0).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Parents current support of a child (PARSUPP)	PA13Q01
	PA13Q02
	PA13Q03
	PA13Q04
	PA13Q05
	PA13Q06
	PA13Q07
		1 Never or hardly ever
		2 Once or twice a year
		3 Once or twice a month
		4 Once or twice a week
		5 Every day or almost
		Missing
			Items were coded as (1=0), (2=1), (3=2), (4=3), (5=4).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Mathematics in child's career and job market	PA14Q01
	PA14Q02
	PA14Q03
	PA14Q04
		1 Strongly agree
		2 Agree
		3 Disagree
		4 Strongly disagree
		Missing
			All items were coded as (1=3), (2=2), (3=1), (4=0).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Academic and professional expectations in mathematics (PQMCAR)	PA15Q01
	PA15Q02
	PA15Q03
	PA15Q04
	PA15Q05
		1 Yes
		2 No
		Missing
			Items were recoded as (1=1), (2=0).
			Total score was calculated as a ratio of a sum of all questions over maximum score of valid responses (questions with missing value did not contribute to max score). Two dummy variables were created as follows:
			  Value value – mean 0
			  Missing 0 1

Child repeated a grade at <ISCED 1>	PA18Q01	1 No, never
		2 Yes, once
		3 Yes, twice or more
		Missing (or invalid)
			Two dummy variables with default value of '00' and
			  - national mode = '-1' in all dummies
			  - corresponding category = '01' (including missing)

Child repeated a grade at <ISCED 2>	PA18Q02	1 No, never
		2 Yes, once
		3 Yes, twice or more
		Missing (or invalid)
			Two dummy variables with default value of '00' and
			  - national mode = '-1' in all dummies
			  - corresponding category = '01' (including missing)

Child repeated a grade at <ISCED 3>	PA18Q03	1 No, never
		2 Yes, once
		3 Yes, twice or more
		Missing (or invalid)
			Two dummy variables with default value of '00' and
			  - national mode = '-1' in all dummies
			  - corresponding category = '01' (including missing)

What level of education do you expect your child to complete
- ISCED 2
- ISCED 3B or C
- ISCED 3A
- ISCED 4
- ISCED 5B
- ISCED 5A, 6
(PQOCCASP)
	PA19Q01
	PA19Q02
	PA19Q03
	PA19Q04
	PA19Q05
	PA19Q06
		1 Tick
		Missing
			Item PA19Q01 was recoded as (1=1)
			Item PA19Q02 was recoded as (1=2)
			Item PA19Q03 was recoded as (1=3)
			Item PA19Q04 was recoded as (1=4)
			Item PA19Q05 was recoded as (1=5)
			Item PA19Q06 was recoded as (1=6).
			New variable PQOCCASP was created as maximum value of six items, thus having categories from 1 to 6. Plus one category for missing (when all six items are missing)
			Six dummy variables were created based on the value of PQOCCASP and with default value of '00' and
			  - national mode = '-1' in all dummies
			  - corresponding category = '01' (including missing)

Expected by parents occupation of the child (SEI)	BSMJ5 (based on PA20Q01)	16-90 (decimal) Missing	value-median 0 0 1

[Part 3/3]

PISA 2012 Main Survey contrast coding used in conditioning for the parent
Table B.4 **questionnaire variables**

Variable	Variable name	Variable coding	Contrast coding
PARENT QUESTIONNAIRE			
Immigration status of mother (PQIMMIGM) by analogy with ST20int	PA21int (PQCTmother) (PQCTMGmot) (PQCTMGfat)	1 Native 2 Second-Generation 3 First-Generation Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Immigration status of father (PQIMMIGF) by analogy with ST20int	PA21int (PQCTfather) (PQCTFGmot) (PQCTFGfat)	1 Native 2 Second-Generation 3 First-Generation Missing	Three dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01' (including missing)
Age of father when arrived to the country	PA22Q01	Value (decimal) Missing	value – median 0 0 1
Age of mother when arrived to the country	PA23Q01	Value (decimal) Missing	value – median 0 0 1
Citizenships of mother (PQCITIZM)	Based on PA24N0101 PA24N0102 PA24N0103 PA24N0104 PA24N0105 PA24N0106 PA24N0107 PA24N0108 PA24N0109 PA24N0110 PA24N0111 PA24N0112 PA24N0113	1 Country of test only 2 Country of test and other 3 Not country of test (inc. Missing)	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01'
Citizenships of father (PQCITIZF)	Based on PA24N0201 PA24N0202 PA24N0203 PA24N0204 PA24N0205 PA24N0206 PA24N0207 PA24N0208 PA24N0209 PA24N0210 PA24N0211 PA24N0212 PA24N0213	1 Country of test only 2 Country of test and other 3 Not country of test Missing	Two dummy variables with default value of '00' and - national mode = '-1' in all dummies - corresponding category = '01'
Language spoken at home – Father (PQLANGNF)	Based on PA25 in F	1 Language of test 2 Other language Missing	Two dummy variables with default value of '00' and - language of test = '-1' in all dummies - other language = '01' '00' - missing = '00' '01'
Language spoken at home – Mother (PQLANGNM)	Based on PA25 in M	1 Language of test 2 Other language Missing	Two dummy variables with default value of '00' and - language of test = '-1' in all dummies - other language = '01' '00' - missing = '00' '01'

Table B.5 **PISA 2012 Main Survey contrast coding used in conditioning for other variables**

Variable	Variable name	Variable coding	Contrast coding
OTHER VARIABLES			
School identification number	SCHOOLID	Unique 7-digit school ID	IDs for small schools (less than 8 students) were recoded into '9999999' for schools which did not administer UH booklet to students, '9999998' for schools which administered UH booklet to all students and '8888888' for schools which administered both UH and normal booklet to all students Total number of schools minus one dummies were created for school membership with default value of '00' and - largest school in the country = '-1' in all dummies - corresponding SCHOOLID= '01'
Booklet number	BOOKID	1 or 21 . . . . 2 or 22 . . . . 3 or 23 . . . . 4 or 24 . . . . 5 or 25 . . . . 6 or 26 . . . . 7 or 27 . . . . 8 . . . . 9 . . . . 10 . . . . 11 . . . . 12 . . . . 13 . . . . 20 (UH) . . . . 71 . . . . 72 . . . . 73 . . . . 74 . . . . 70 (UH)	01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 00 00 00 00 00 00 00 00 00 00 00 00 01 00 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 00 00 00 00 00 00 00 00 00 00 00 00 00 -1 -1 -1 11 11 11 11 11 11 11 11 11 11 01 00 00 11 11 11 11 11 11 11 11 11 11 00 01 00 11 11 11 11 11 11 11 11 11 11 00 00 01 11 11 11 11 11 11 11 11 11 11 00 00 00 11 11 11 11 11 11 11 11 11 11