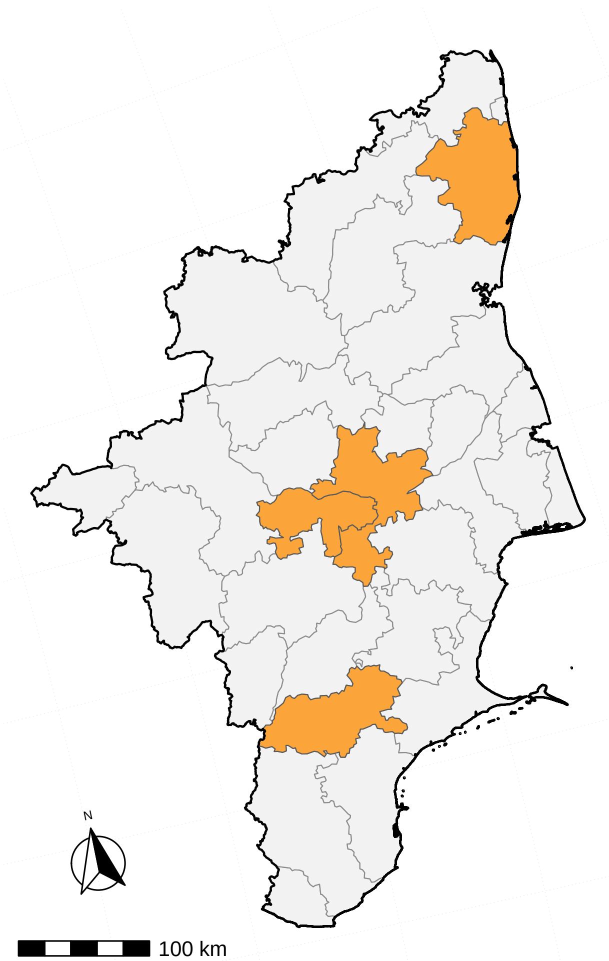
Online Appendix for "The productivity of public and private preschools (and schools): Evidence from India" by Berg, Romero, and Singh

September 4, 2025

A Additional tables and figures

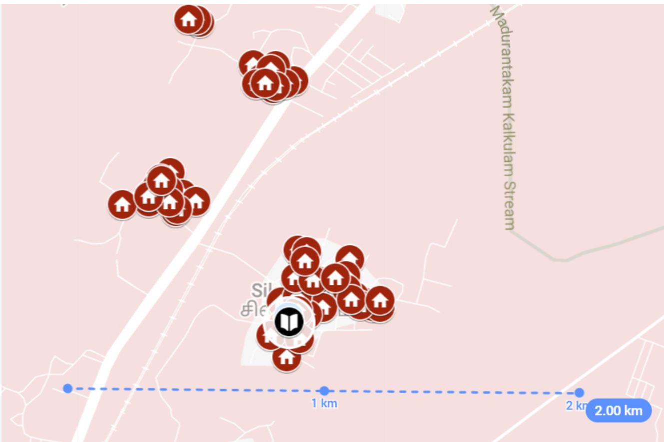
Figure A.1: Map of sample districts in Tamil Nadu



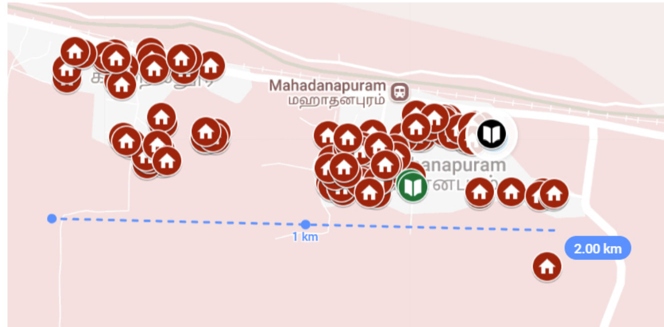
Note: This figure shows the four sample districts (Kancheepuram, Karur, Tiruchchirappalli, and Virudhunagar) included in the data collection.

Figure A.2: Villages and the households we sample (some examples)

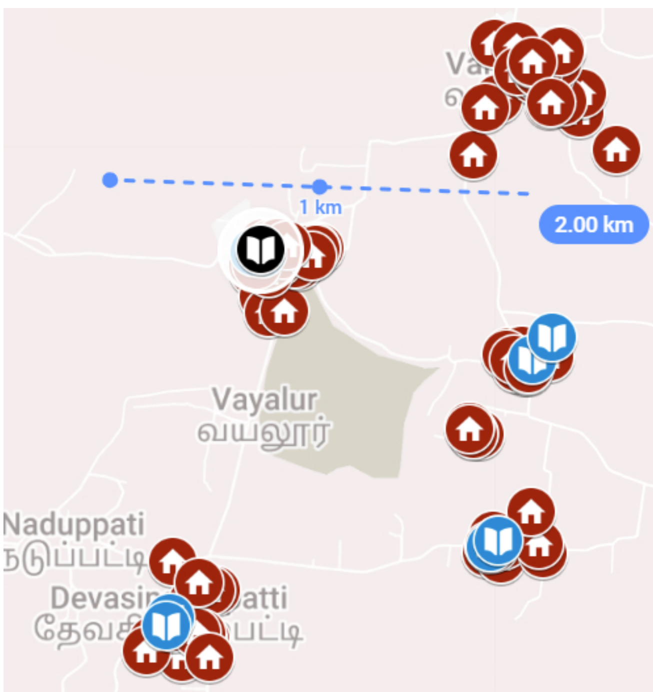
(a) Village 1



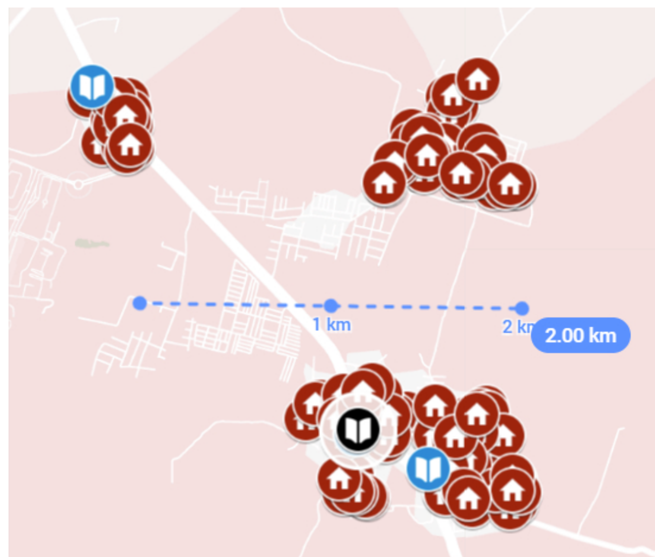
(b) Village 2



(c) Village 3

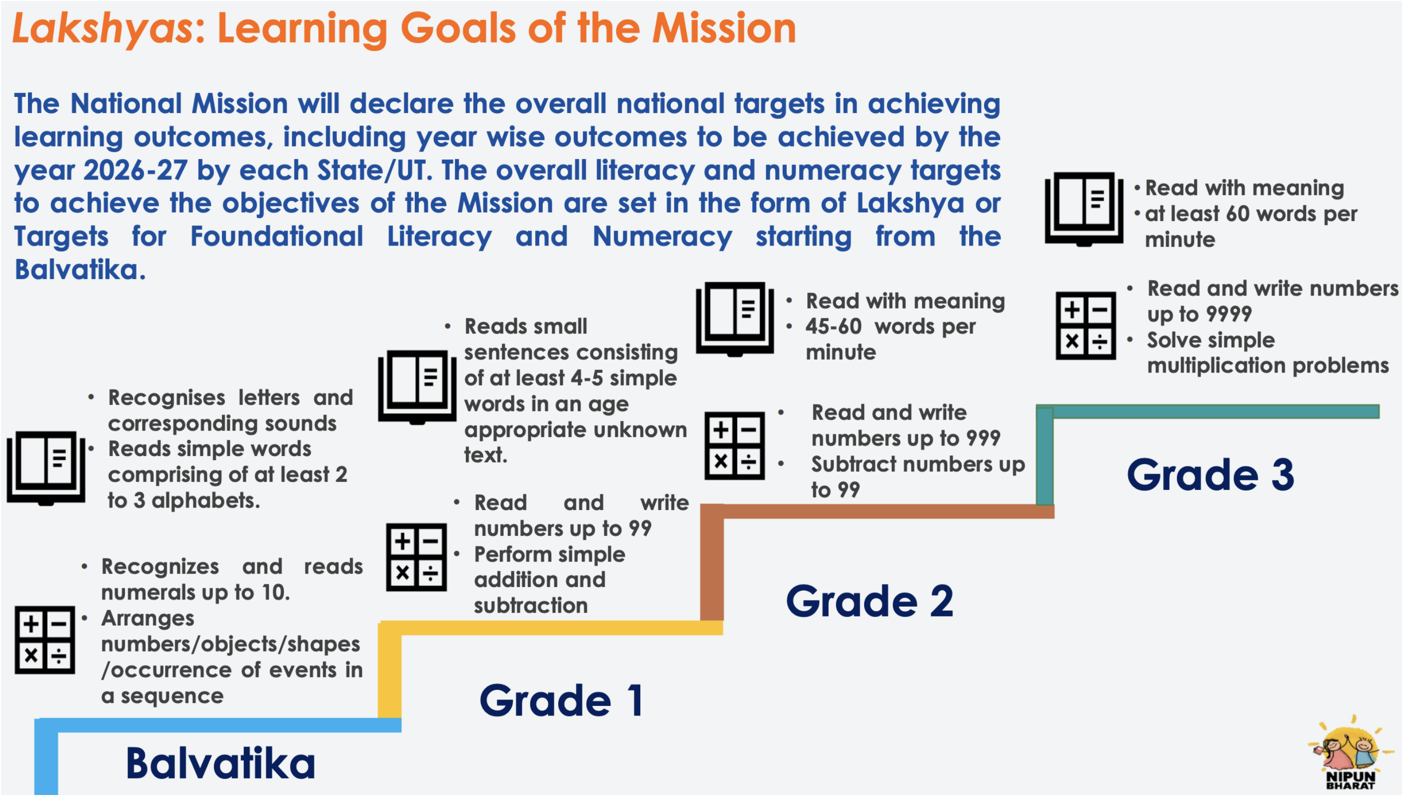


(d) Village 4



Note: This figure shows 4 villages in our sample. The black point with a "book" icon represents the "reference point" (i.e., the anganwadi). Blue "book" icons represent other public schools (across all levels), and green "book" icons represent other private schools (across all levels). The red "house" icons represent households in our sample (i.e., with children). The villages were selected based on their position at the 20th, 40th, 60th, and 80th percentiles of the maximum distance between the reference point and a household in our sample. As can be seen, in most cases, 'villages' encompass several clusters of households that are located near each other. The 2-kilometer rule was not strictly enforced, and in most cases, either all households in a cluster were included or none.

Figure A.3: Learning goals in foundational literacy and numeracy in India



Note: This figure shows official learning goals for early childhood education as presented in the National Initiative for Proficiency in reading with Understanding and Numeracy (NIPUN) by the Ministry of Education in India. Retrieved from https://static.pib.gov.in/WriteReadData/specificdocs/documents/2021/jul/doc20217531.pdf

Table A.1: Comparing baseline sample to NFHS

	NFHS-V (1)	Baseline (2)	Difference (3)
Panel A: Assets and household characteristics			
Internet	0.59	0.48	-0.11***
	(0.49)	(0.50)	p=0.00
Washing machine	0.14	0.09	-0.06***
	(0.35)	(0.28)	p=0.00
Fridge	0.56	0.46	-0.10***
	(0.50)	(0.50)	p=0.00
Computer	0.09	0.08	-0.01
	(0.28)	(0.27)	p=0.20
Television	0.94	0.93	-0.02***
	(0.23)	(0.26)	p=0.01
Fan	0.97	0.97	-0.00
	(0.16)	(0.17)	p=0.83
Electricity	0.99	0.94	-0.05***
	(0.08)	(0.23)	p=0.00
Car	0.05	0.05	-0.01
	(0.22)	(0.21)	p=0.34
Tractor	0.02	0.03	0.00
	(0.15)	(0.16)	p=0.35
Bike	0.77	0.75	-0.02
	(0.42)	(0.43)	p=0.14
Bicycle	0.46	0.36	-0.10***
	(0.50)	(0.48)	p=0.00
Number of children (3-10 yrs old)	1.62	1.55	-0.07***
	(0.68)	(0.62)	p=0.00
Scheduled caste	0.36	0.33	-0.03
	(0.48)	(0.47)	p=0.16
Owns land	0.31	0.25	-0.06***
	(0.46)	(0.43)	p=0.00
Observations	2,561	17,486	
Panel B: Maternal education			
Mother education: at least some primary	0.96	0.96	-0.00
	(0.20)	(0.21)	p=0.38
Mother education: at least some secondary	0.87	0.93	0.06***
	(0.33)	(0.25)	p=0.00
Observations	2,542	16,280	

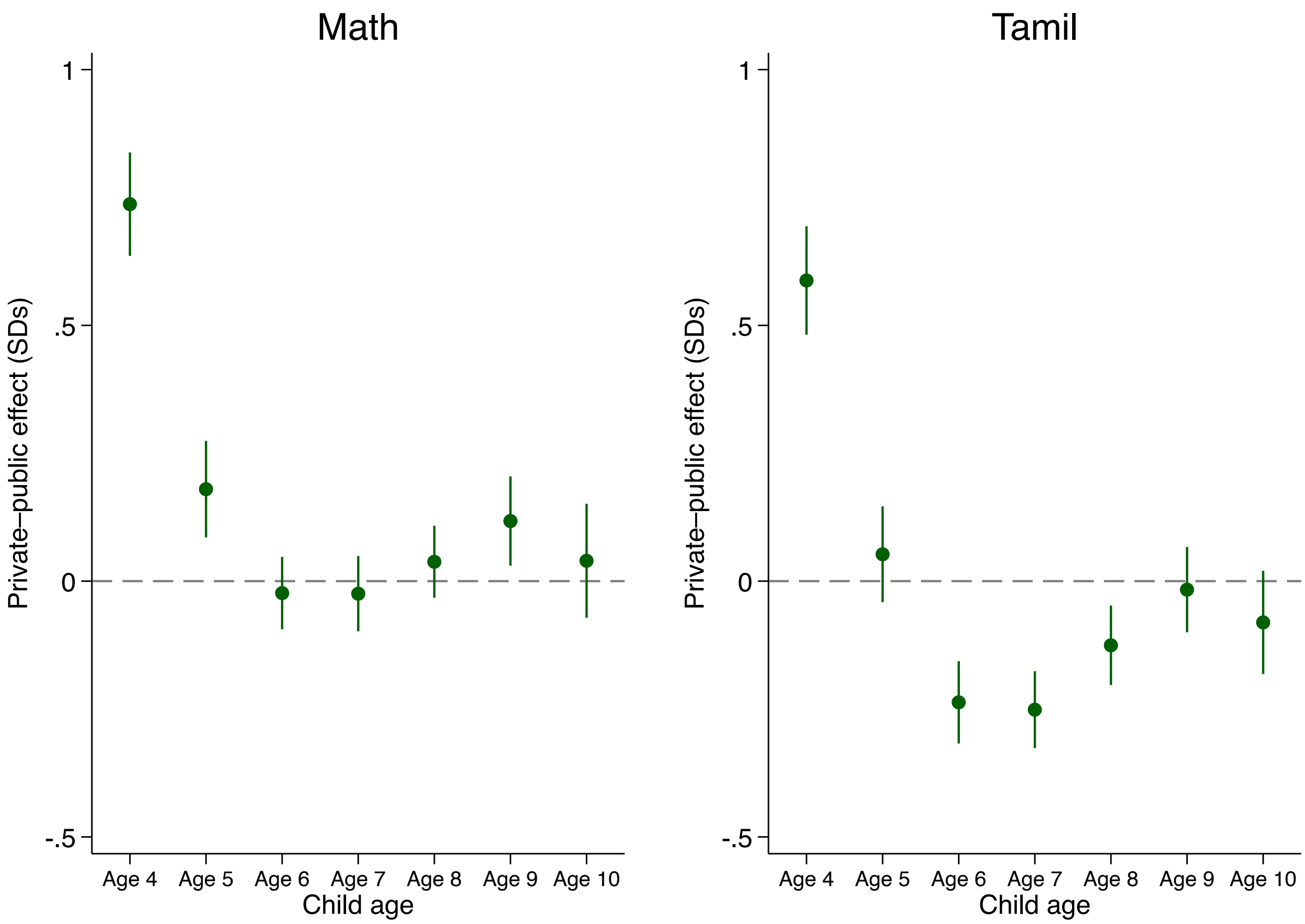
Notes: The table presents means and standard deviations for households in Tamil Nadu with children aged 3–10 in the NFHS-V survey (Column 1) and households in our baseline sample in 2022 (Column 2). Column 3 shows differences and statistical significance (clustering standard errors at the sampling cluster level for NFHS-V and the village level in our sample). $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Table A.2: Private school value-added in preschool and primary school for children aged 5

	Math (1)	Tamil (2)
Private school	0.715***	0.590***
	(0.0587)	(0.0562)
Primary school	1.051***	1.072***
	(0.0487)	(0.0475)
Private school $\times$ Primary school	-0.698***	-0.698***
	(0.0693)	(0.0672)
Math IRT score in 2022	0.171***	
	(0.0236)	
Tamil IRT score in 2022		0.148***
		(0.0216)
Constant	-0.146***	-0.164***
	(0.0400)	(0.0405)
Controls	All	All
Observations	2,841	2,841

Notes: Robust standard errors, clustered at the village level, in parentheses. The regressions include only children aged 5, around half of whom are already enrolled in primary school. The coefficient on the private school dummy captures the private premium in preschools. The sum of this coefficient and that of the interaction between private and primary school captures the private premium in primary school. Lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, as well as child gender are included in both regressions. Test scores refer to the IRT Expected A Posteriori (EAP) scores, standardized with respect to children aged 5 in the 2022 assessments. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Figure A.4: Private school value-added by age



Notes: These figures show the private premium by age (as opposed to estimating the premium for 6- to 10-year-olds together). The coefficients are analogous to those in Table 2, but they are estimated separately for children of different ages.

Table A.3: Private school value-added in math competencies

	Dependent variable: Proportion correct on math items							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Age 4	*Addition*	*Subtraction*	*Number identification*	*Quantitative comparison*				
Private school	0.251***	0.248***	0.326***	0.188***				
	(0.026)	(0.027)	(0.019)	(0.020)				
Math IRT score in 2022	0.072***	0.081***	0.059***	0.054***				
	(0.013)	(0.015)	(0.011)	(0.010)				
Constant	0.516***	0.414***	0.270***	0.651***				
	(0.019)	(0.021)	(0.016)	(0.015)				
Observations	1,839	1,839	1,839	1,839				
Number of items	2	1	7	5				
Public sector avg.	0.414	0.317	0.189	0.577				
Panel B: Ages 6–10	*Addition*	*Subtraction*	*Number identification*	*Quantitative comparison*	*Multiplication & division*	*problems Applied*	*Geometry*	*Measurement*
Private school	0.012*	0.010	-0.011	0.022*	0.014	0.051***	-0.023	0.005
	(0.005)	(0.006)	(0.008)	(0.011)	(0.013)	(0.009)	(0.032)	(0.011)
Math IRT score in 2022	0.047***	0.044***	0.057***	-0.043***	0.059***	0.058***	0.049**	0.076***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.006)	(0.004)	(0.015)	(0.005)
Constant	0.756***	0.661***	0.893***	0.699***	0.402***	0.709***	0.311***	0.630***
	(0.002)	(0.002)	(0.002)	(0.004)	(0.007)	(0.003)	(0.022)	(0.006)
Observations	14,344	14,344	5,201	14,344	9,142	14,344	805	9,142
Number of items	12	6	4	6	4	5	3	2
Public sector avg.	0.782	0.684	0.890	0.665	0.457	0.740	0.377	0.700

Notes: Robust standard errors, clustered at the village level, in parentheses. Lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender included. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Table A.4: Private school value-added in Tamil competencies

	Dependent variable: Proportion correct on Tamil items				
	(1)	(2)	(3)	(4)	(5)
Panel A: Age 4	*Letter & word recognition*	*Oral comprehension*			
Private school	0.221***	0.096***			
	(0.020)	(0.017)			
Tamil IRT score in 2022	0.059***	0.057***			
	(0.009)	(0.009)			
Constant	0.450***	0.583***			
	(0.014)	(0.013)			
Observations	1,839	1,839			
Number of items	7	5			
Public sector avg.	0.364	0.504			
Panel B: Ages 6–10	*Letter & word recognition*	*Oral comprehension*	*Word & sentence comprehension*	*Sentence & story comprehension*	*Spelling*
Private school	-0.058***	-0.024	-0.003	-0.024*	0.001
	(0.006)	(0.014)	(0.004)	(0.010)	(0.034)
Tamil IRT score in 2022	0.075***	0.042***	0.034***	0.085***	0.049**
	(0.003)	(0.007)	(0.002)	(0.005)	(0.019)
Constant	0.797***	0.818***	0.872***	0.412***	0.624***
	(0.002)	(0.005)	(0.002)	(0.006)	(0.033)
Observations	13,525	1,837	14,344	9,142	805
Number of items	14	5	12	16	2
Public sector avg.	0.846	0.800	0.894	0.504	0.713

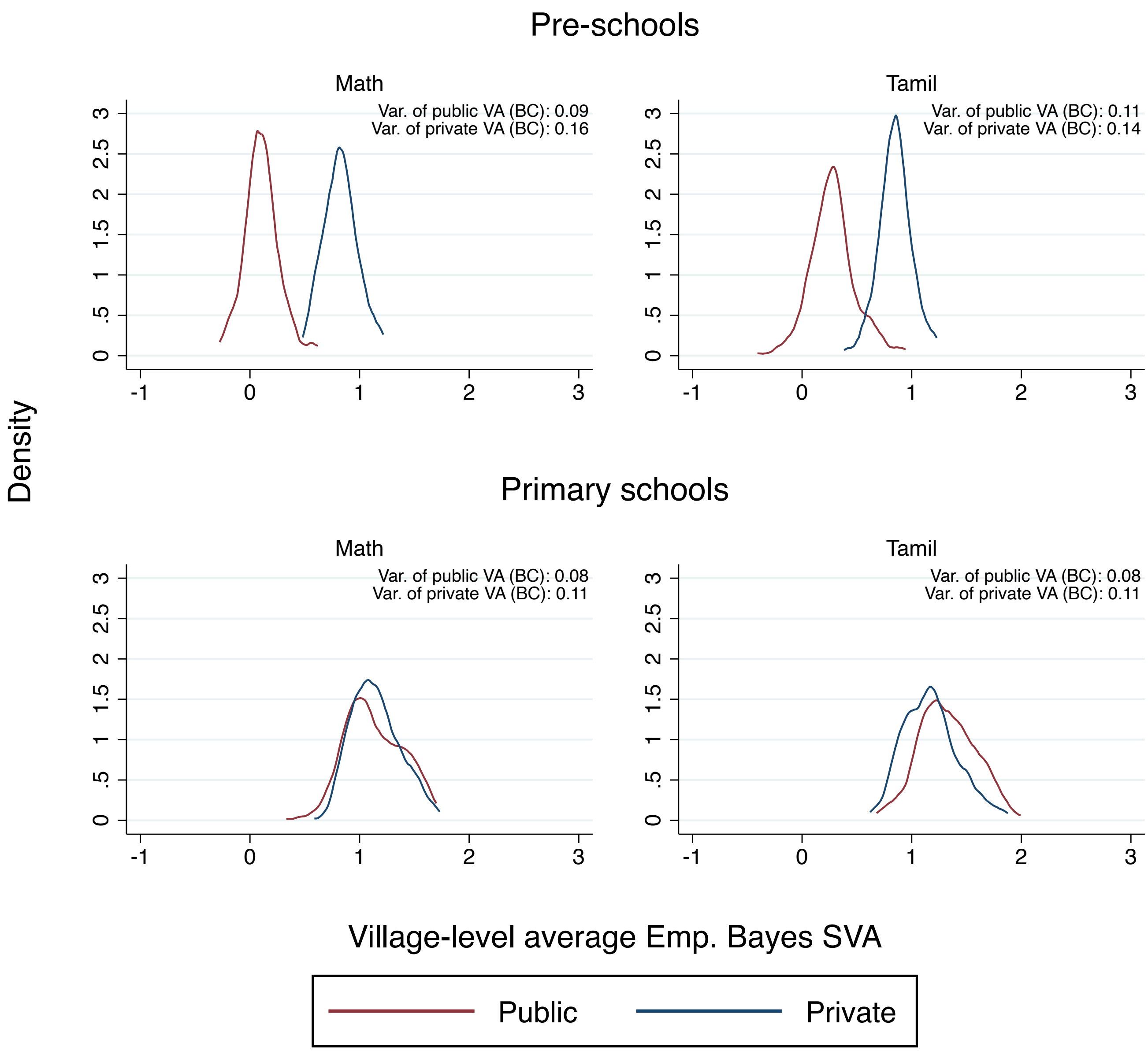
Notes: Robust standard errors, clustered at the village level, in parentheses. Lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender included. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Table A.5: Value-added of public/private preschools relative to no enrolment, age 4

	Math		Tamil	
	(1)	(2)	(3)	(4)
Public	0.229	0.244	0.302*	0.291*
	(0.126)	(0.125)	(0.147)	(0.146)
Private	1.039***	0.982***	0.946***	0.882***
	(0.126)	(0.127)	(0.146)	(0.146)
Math IRT score in 2022		0.176***		
		(0.0275)		
Tamil IRT score in 2022				0.198***
				(0.0266)
Constant	-0.801***	-0.582***	-0.819***	-0.541***
	(0.120)	(0.127)	(0.140)	(0.146)
Controls	Village FE	All	Village FE	All
Observations	1,906	1,906	1,906	1,906

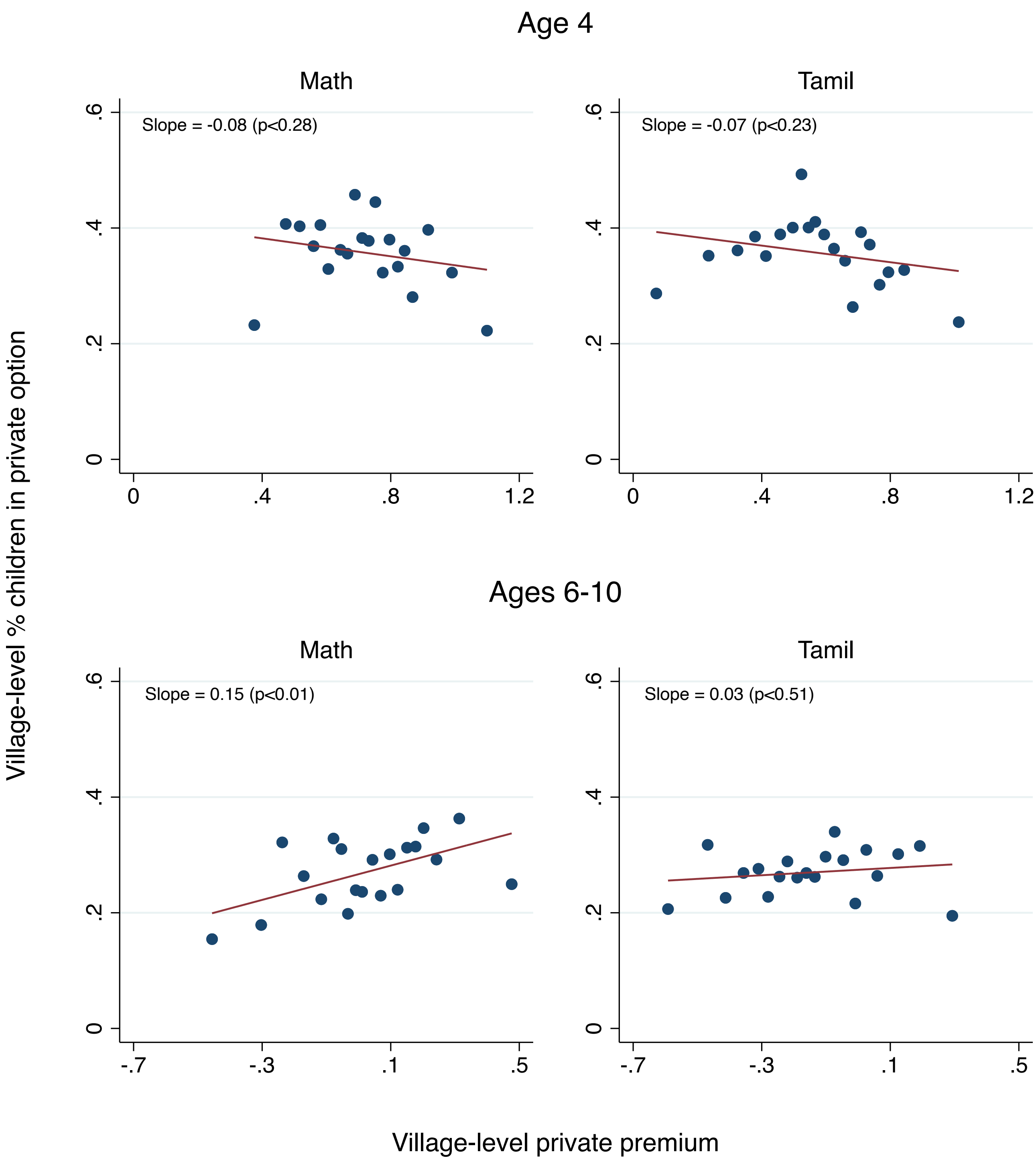
Notes: Robust standard errors, clustered at the village level, in parentheses. The omitted category is children not enrolled in any type of education (67 children). Columns 1 and 3 report raw test score differences by type of school attended, within villages. Columns 2 and 4 include village fixed effects and controls for lagged scores, deciles of the SES wealth index, paternal and maternal education, as well as child gender. Test scores refer to the IRT Expected A Posteriori (EAP) scores, standardized with respect to children aged 5 in the 2022 assessments. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Figure A.5: Distributions of village-level average value-added



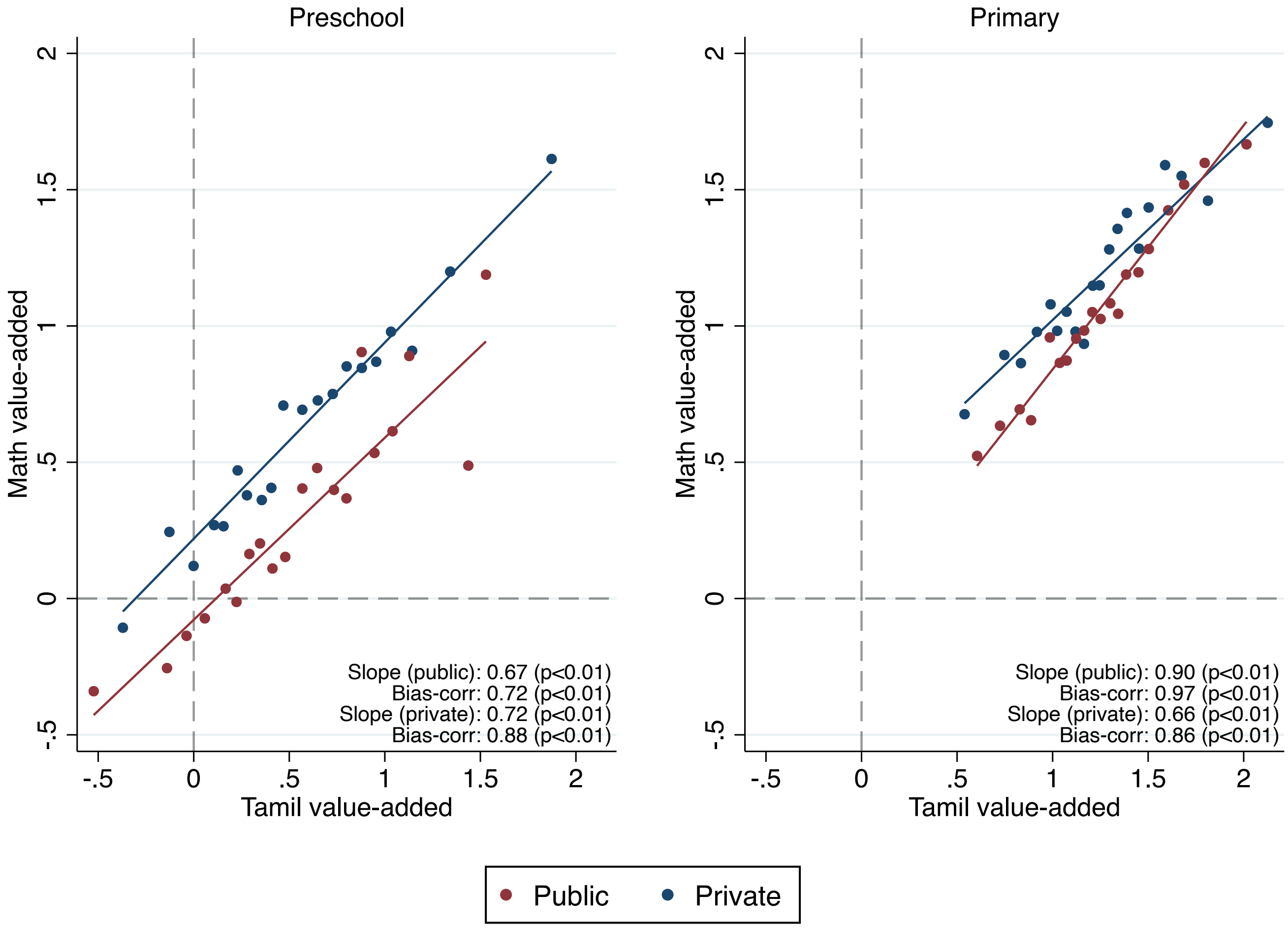
Notes: These figures show kernel density plots of village-level average school value-added by sector (public/private) and level (preschool/primary) using Empirical Bayes measures as described in Section E.1. These are generated in a regression that includes children aged 4–10, as described in Section 4.1. Bias-corrected variances of the underlying value-added parameters are computed as described in Appendix E.2.

Figure A.6: Village-level private premium and share of private enrolment



Notes: These figures depict the correlation between (1) the village-level share of children in a private school and (2) the difference between the average *private* and *public* school value-added in the same village (i.e., the private premium). These value-added measures are Empirical Bayes estimates, estimated as described in Appendix E.1.

Figure A.7: Cross-subject correlations of village-level value-added



Notes: These figures show local polynomial fits of village-level value-added in math vs. Tamil at the preschool/primary school level, respectively. Both raw and bias-corrected slopes are shown in the bottom right of the figures (see Appendix E.2 for details on bias-correction).

Table A.6: Regressions of private premia on market size and village average SES

	Preschool private premium		Primary school private premium	
	Math (1)	Tamil (2)	Math (3)	Tamil (4)
Number of children in market (std)	0.008	0.010	-0.002	-0.002
	(0.013)	(0.017)	(0.017)	(0.015)
Village average SES (std)	-0.021	-0.046*	0.016	-0.022
	(0.014)	(0.018)	(0.015)	(0.018)
Constant	0.720***	0.571***	0.020	-0.147***
	(0.012)	(0.015)	(0.015)	(0.015)
Observations	200	200	211	211

Notes: Robust standard errors in parentheses. This table shows village-level regressions of private premia, across sectors and levels, on the standardized number of children aged 6–10 and the standardized average SES percentile in the village. One standard deviation of number of children corresponds to 45 children; for village-level average SES, it corresponds to 9.7 percentiles of the SES distribution. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

B Attrition between 2022 and 2023 survey waves

This section investigates whether attrition between the 2022 and 2023 survey waves correlates with socioeconomic status, age, or baseline ability.

B.1 Age distribution in the main sample

In 2022, children from age 3 were surveyed and assessed. The 2023 survey was administered just over a year after the 2022 wave, which means that children aged slightly more than one year between the waves (1.2 years, on average, in our main sample). Hence, some children who were 3 years old at baseline had turned 5 by the endline. Likewise, children aged 2 at baseline — some of whom would be 4 by the endline — were not assessed in 2022 and are therefore excluded. This results in a slightly skewed age distribution in the main sample, such that children aged 4 at the endline are underrepresented. The main sample contains 1,837 children with endline age 4, 2,840 with endline age 5, and 14,344 with endline age 6–10.

B.2 Attrition between 2022 and 2023

Since the main sample focuses on children aged 4–10 at the time of the endline survey, most of these children were aged 3–9 at baseline. However, some children who were 9 years old at baseline will have aged out of the sample frame (e.g., turned 11) by the endline. This makes it non-trivial to specify exactly which children at baseline should be considered for possible attrition.

To address this issue, we look at attrition for *all* children aged 3–9 at baseline. While a subset of these children will not be included in the main sample (i.e., those who turned 11 by the endline), this gives a fair representation of attrition in the relevant age span. Table B.1 shows the results of this analysis.

Table B.1: Attrition of children aged 3–9 in the 2022 survey wave

	Re-surveyed (1)	Attrited (2)	Difference (village FE) (3)
Child age (years)	5.79	6.04	0.30***
	(1.82)	(2.08)	p=0.00
SES percentile	49.24	51.86	0.27
	(28.42)	(29.79)	p=0.62
Mother Edu: $<$ Gr.9	0.25	0.24	0.00
	(0.43)	(0.43)	p=0.64
Mother Edu: Gr. 9-11	0.41	0.40	-0.01
	(0.49)	(0.49)	p=0.13
Mother Edu: Gr. 12+	0.34	0.36	0.01
	(0.47)	(0.48)	p=0.26
Father Edu: $<$ Gr.9	0.38	0.38	0.01
	(0.49)	(0.48)	p=0.28
Father Edu: Gr. 9-11	0.42	0.41	-0.02**
	(0.49)	(0.49)	p=0.02
Father Edu: Gr. 12+	0.20	0.22	0.01
	(0.40)	(0.41)	p=0.11
Enrolled in private	0.21	0.24	0.01
	(0.41)	(0.43)	p=0.12
Math (2022)†	-0.01	0.03	0.02
	(0.90)	(0.91)	p=0.16
Tamil (2022)†	-0.00	0.01	0.01
	(0.94)	(0.96)	p=0.49
Observations	19,200	6,161	

Notes: This table presents means and standard deviations (in parentheses) for children aged 3–9 in the 2022 survey wave along a number of characteristics measured in 2022. The first column displays this information for children who were successfully re-surveyed in 2023, and the second for those who were not. The third column displays the differences between these groups, along with the p-value of the difference, controlling for village fixed effects. Standard errors of differences are clustered at the village level. † Math and Tamil (2022) baseline scores correspond to the residuals after regressing the original IRT scores on age brackets in years. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

B.3 Inverse probability weighting

Tables B.2 and B.3 reproduce the main results on the average private premium (Table 2) and the SES decomposition (Table 4) using inverse probability weighting. In particular, we predict the probability of being observed in 2023, conditional on being observed in 2022, using data on child age, gender, SES, parental education, baseline test scores, and village indicators in a probit regression. We then reproduce the main results, weighing each child inversely by the probability of being observed. The results change little under this alternative strategy.

Table B.2: Private school value-added in preschool and primary school: IPW

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Private school	0.805***	0.735***	0.341***	0.192***	0.133***	-0.002
	(0.046)	(0.050)	(0.046)	(0.049)	(0.023)	(0.019)
Math IRT score in 2022		0.185***		0.244***		0.310***
		(0.028)		(0.026)		(0.010)
Panel B: Tamil						
Private school	0.631***	0.576***	0.161***	0.049	-0.106***	-0.180***
	(0.051)	(0.053)	(0.046)	(0.048)	(0.024)	(0.024)
Tamil IRT score in 2022		0.198***		0.210***		0.339***
		(0.027)		(0.023)		(0.011)

Notes: Robust standard errors, clustered at the village level, in parentheses. Columns 1, 3 and 5 show raw means by private school attendance within villages. Columns 2, 4 and 6 include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender. Test scores refer to equated IRT scores, standardized with respect to children aged 5 in the 2022 assessments. All regressions are weighted inversely proportional to their estimated probability of being observed in 2023, conditional on being observed in 2022. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Table B.3: Decomposition of SES gap (top/bottom 25%) in 2023 test scores, preschool and primary level: IPW

	Age 4			Age 5			Ages 6–10		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel A: Math									
Top 25% SES	0.441***	0.420***	0.172*	0.324***	0.266***	0.152*	0.248***	0.174***	0.161***
	(0.069)	(0.069)	(0.078)	(0.057)	(0.055)	(0.067)	(0.032)	(0.028)	(0.026)
Private school			0.667***			0.263**			0.034
			(0.078)			(0.080)			(0.028)
Panel B: Tamil									
Top 25% SES	0.378***	0.351***	0.154*	0.148*	0.085	0.054	0.098**	0.052	0.110***
	(0.070)	(0.069)	(0.072)	(0.061)	(0.062)	(0.068)	(0.031)	(0.027)	(0.030)
Private school			0.527***			0.074			-0.155***
			(0.082)			(0.078)			(0.031)

Notes: Robust standard errors, clustered at the village level, in parentheses. Village fixed effects and controls for child gender are included in all regressions. Test scores refer to the IRT EAP scores, standardized with respect to children aged 5 in the 2022 assessments. The SES index is based on questions regarding the availability of household amenities and computed with PCA. The omitted category contains students in households with an SES index below the 25th percentile. Households with an SES index between the 25th and 75th percentiles are excluded from the regressions. All regressions are weighted inversely proportional to their estimated probability of being observed in 2023, conditional on being observed in 2022. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Village-level results robustness to attrition

In this section, we show that our village-level results are robust to attrition. Specifically, we present the most important village-level results separately for villages with above- and below-median attrition rates, and show that they are remarkably similar. In Table B.4, we show the average private premium at the pre- and primary school levels, both with and without Empirical Bayes shrinkage, as well as correlations in value-added across sectors and levels. We do this separately for all villages, both below- and above-median attrition. Average private premiums are very similar across low- and high-attrition villages, both with and without Empirical Bayes shrinkage. While the sector correlations are more imprecisely estimated, the qualitative pattern of stronger cross-sector correlation at the primary, relative to preschool level, still holds. The correlations in value-added across levels appear to differ between villages with high and low attrition rates. In villages with low attrition, the correlation between primary and preschool value-added appears to be stronger in the public sector than in the private sector.

Table B.4: Results for above- and below median attrition villages

	Math			Tamil		
	All (1)	Below (2)	Above (3)	All (4)	Below (5)	Above (6)
Panel A: Private premia						
Private premium (preschool, EB)	0.72	0.73	0.71	0.57	0.58	0.56
	(0.01)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)
Private premium (primary, EB)	0.02	-0.01	0.05	-0.15	-0.14	-0.16
	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Private premium (preschool)	0.73	0.73	0.72	0.58	0.62	0.53
	(0.04)	(0.05)	(0.05)	(0.04)	(0.06)	(0.05)
Private premium (primary)	0.02	-0.00	0.04	-0.13	-0.09	-0.17
	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)	(0.03)
Panel B: Sector correlations						
Private/public BC corr. (preschool)	0.80	0.84	0.69	0.61	0.61	0.51
	(0.13)	(0.17)	(0.20)	(0.11)	(0.14)	(0.18)
Private/public BC corr. (primary)	0.88	0.86	0.95	0.99	0.96	0.85
	(0.08)	(0.13)	(0.14)	(0.10)	(0.14)	(0.16)
Panel C: Level correlations						
Primary/preschool BC corr. (public)	0.54	0.56	0.31	0.52	0.48	0.47
	(0.09)	(0.10)	(0.10)	(0.08)	(0.09)	(0.11)
Primary/preschool BC corr. (private)	0.47	0.26	0.55	0.59	0.40	0.62
	(0.06)	(0.08)	(0.08)	(0.08)	(0.10)	(0.11)
Number of villages	215	107	108	215	107	108
Share of attrition	0.26	0.35	0.16	0.26	0.35	0.16

Notes: This table reproduces the main findings on village value-added separately for villages above and below the median in terms of share of attriting children from 2022 to 2023. Panel A shows the average private premia, both using Empirical Bayes and raw value-added estimates, at the pre- and primary school levels. Panel B shows the slope coefficient of private on public value-added, using the bias correction procedure described in Appendix E.2. Panel C shows the corresponding slope coefficients across levels (pre- and primary).

C Measuring socioeconomic status

We construct a household-level socioeconomic status (SES) index based on ownership of a set of assets in the 2022 survey round. Households were asked whether they own a washing machine, refrigerator, grinder, mixer, computer, TV, fan, electric lights, car, tractor, motorbike/scooter, bicycle, and a telephone. Furthermore, we recorded whether the household owns agricultural land and the house in which they reside, as well as whether they have access to running water. These responses are coded as binary variables, and combined into a single index using principal component analysis (PCA): the first eigenvector constitutes our SES index. Finally, this index is transformed into percentiles. Table C.1 reports descriptive statistics of household asset ownership and maternal education by quartiles of our constructed SES index, as well as for the full sample.

Table C.1: Household characteristics by quartiles of the constructed SES index

	SES quartiles				
	First (1)	Second (2)	Third (3)	Fourth (4)	All (5)
Panel A: Household assets					
Internet	0.20	0.21	0.60	0.87	0.46
	(0.40)	(0.41)	(0.49)	(0.34)	(0.50)
Washing machine	0.01	0.01	0.02	0.28	0.07
	(0.09)	(0.09)	(0.15)	(0.45)	(0.26)
Refrigerator	0.09	0.22	0.58	0.91	0.44
	(0.29)	(0.41)	(0.49)	(0.28)	(0.50)
Grinder	0.56	0.97	1.00	1.00	0.88
	(0.50)	(0.16)	(0.07)	(0.03)	(0.33)
Mixer	0.62	0.99	1.00	1.00	0.90
	(0.49)	(0.10)	(0.05)	(0.00)	(0.30)
Computer	0.02	0.03	0.05	0.21	0.07
	(0.14)	(0.18)	(0.22)	(0.41)	(0.26)
TV	0.75	0.97	1.00	1.00	0.93
	(0.43)	(0.16)	(0.07)	(0.03)	(0.26)
Fan	0.89	1.00	1.00	1.00	0.97
	(0.31)	(0.04)	(0.02)	(0.02)	(0.17)
Electric lights	0.90	0.96	0.97	0.98	0.95
	(0.30)	(0.20)	(0.17)	(0.14)	(0.22)
Car	0.01	0.01	0.03	0.14	0.04
	(0.08)	(0.11)	(0.16)	(0.35)	(0.20)
Tractor	0.01	0.01	0.01	0.08	0.03
	(0.10)	(0.11)	(0.12)	(0.27)	(0.16)
Motorbike	0.38	0.71	0.94	0.99	0.75
	(0.49)	(0.45)	(0.24)	(0.12)	(0.43)
Bicycle	0.27	0.39	0.30	0.48	0.36
	(0.44)	(0.49)	(0.46)	(0.50)	(0.48)
Telephone	0.92	0.99	1.00	1.00	0.98
	(0.27)	(0.10)	(0.01)	(0.02)	(0.15)
Owns land	0.16	0.23	0.24	0.43	0.26
	(0.36)	(0.42)	(0.43)	(0.50)	(0.44)
Owns house	0.83	0.86	0.93	0.96	0.89
	(0.38)	(0.35)	(0.25)	(0.19)	(0.31)
Running water	0.16	0.22	0.27	0.48	0.27
	(0.36)	(0.42)	(0.44)	(0.50)	(0.45)
Panel B: Maternal education					
Mother's education: $<$ Grade 9	0.35	0.28	0.20	0.16	0.25
	(0.48)	(0.45)	(0.40)	(0.37)	(0.43)
Mother's education: $\geq$ Grade 12	0.21	0.28	0.36	0.52	0.34
	(0.41)	(0.45)	(0.48)	(0.50)	(0.47)
Observations	5,013	4,781	5,216	4,304	19,314

Notes: This table shows means and standard deviations (in parentheses) of household asset ownership that forms the basis of our SES index (Panel A) and maternal education (Panel B). The sample is split along quartiles of the SES index in Columns 2–5, and Column 6 shows descriptives for the full sample.

D Robustness of the average private premium

D.1 Details on robustness values and bounds

For robustness values and bounds, we follow closely the approach suggested in Cinelli & Hazlett (2019), and implement the procedure using the provided STATA software `sensemakr`.

We provide an outline of this procedure below. Consider a notation where test scores are denoted as Y, private enrolment as D, the set of control variables included as $\mathbf{X}$, and the hypothetical confounder as Z. Under a scenario where Z is equally powerful in predicting Y and D, the robustness value is defined as:

$$R^2_{Y\sim Z|\mathbf{X},D} = R^2_{D\sim Z|\mathbf{X}} = RV_q, \tag{1}$$

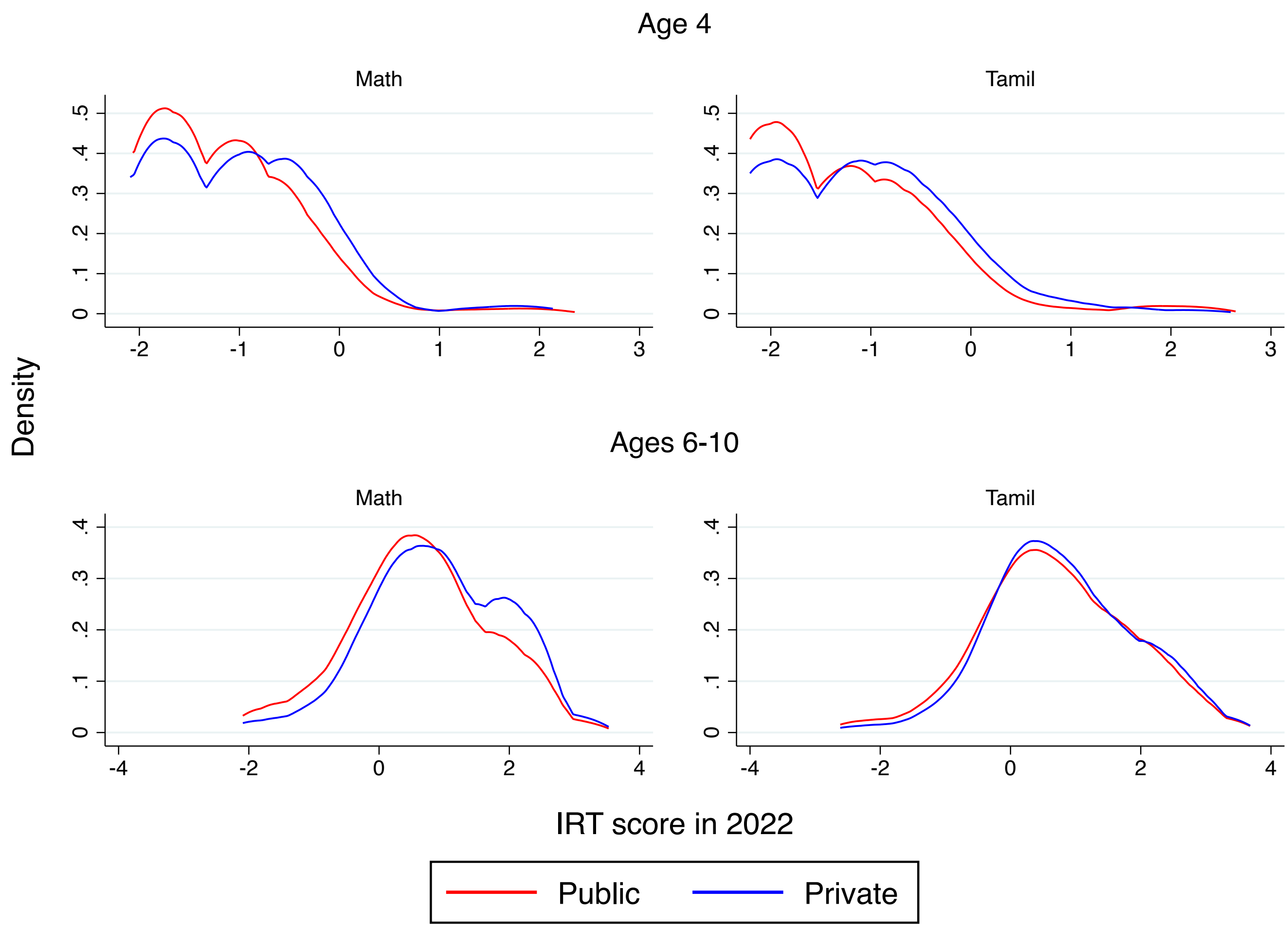
where RV_q is the value of these partial R^2's strong enough to reduce the estimated treatment effect by $100q\%$. In our analysis, we report two robustness values: RV, which sets $q = 1$ and hence asks how strong a predictor Z would have to be to completely nullify the treatment effect of D, and $RV_{\alpha=0.05}$ which asks how much is necessary to render the treatment effect statistically insignificant at the 5% level.

The strength of this procedure is that we can reason about bounds for $R^2_{Y\sim Z|\mathbf{X},D}$ and $R^2_{D\sim Z|\mathbf{X}}$ in a transparent way. We consider a scenario under which the confounder Z is as predictive of the residual variation in Y and D, as a set of benchmarking covariates $\mathbf{B}$. In our analysis, this includes the core covariates of the value-added analysis: baseline test score, deciles of our SES index, parental education, and child gender. Cinelli & Hazlett (2019) shows that under this assumption, we can recover bounds for the relevant partial R^2's and compare them to the robustness values to evaluate how problematic such a confounder Z would be.

We perform this exercise separately for children aged 4 and 6–10, as in our main analysis. Results are shown in Table 3. We find robustness values around 27–31% in math, and 20–24% in Tamil. This means that a problematic confounder would have to explain roughly a quarter of the residual variation in both test scores and private enrolment to nullify the estimated private premia. For the partial R^2's in our baseline specification, we find that (conditional on private enrolment) the benchmarking covariates $\mathbf{B}$ explain a nontrivial share of the variation in private enrolment ($\sim 16\%$), but fairly little of the variation in test scores ($\sim$ 7%). These values fall in the more extensive specifications in columns (2) through (4) — this is because the partial R^2 of the baseline test score and SES controls become less predictive once we control flexibly for baseline test score in the other subject and other proxies for SES, such as caste. Hence, the most informative values are found in column (1). Importantly, these values are well below the robustness values, suggesting that a confounder Z as strong as $\mathbf{B}$ would not nullify our private premia.

D.2 Common support of covariates

Figure D.1: Distributions of baseline test scores



Notes: These figures show the kernel densities of IRT scores in 2022 (baseline) for children enrolling in private and public institutions over the following year, separately by subject and age groups. While children in private institutions are generally positively selected on baseline scores, the distributions broadly overlap.

We estimate the probability of attending private school separately by age using a probit regression. The predictors used are the full set of value-added controls, lagged scores in both subjects, as well as village fixed effects, deciles of the SES index, paternal and maternal education, and child gender. In Tables D.1-D.3, showing the average impact of private enrolment, we restrict the sample to children with an estimated probability of attending private school within 5, 10, and 15 percentage points of the age-specific median. This excludes a significant share of the main analysis sample: one standard deviation of the estimated probability of private enrolment is similar across ages at around 28 percentage points. The results are similar to those in the main analysis.

Table D.1: Private school value-added in preschool and primary school (probit restriction: 5 percentage points)

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Private school	0.967***	0.789***	0.291*	0.267*	-0.083	-0.060
	(0.167)	(0.185)	(0.131)	(0.131)	(0.051)	(0.048)
Math IRT score in 2022		0.449**		0.434		0.315***
		(0.147)		(0.234)		(0.023)
Panel B: Tamil						
Private school	1.131***	0.997***	0.005	-0.011	-0.209***	-0.186***
	(0.157)	(0.214)	(0.133)	(0.140)	(0.054)	(0.052)
Tamil IRT score in 2022		0.245		-0.264		0.288***
		(0.133)		(0.169)		(0.024)
Controls	Village FE	All	Village FE	All	Village FE	All
Observations	148	148	272	272	2,007	2,007

Notes: Robust standard errors, clustered at the village level, in parentheses. Columns 1, 3 and 5 show raw means by private school attendance within villages. Columns 2, 4 and 6 include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender. Test scores refer to equated IRT scores, standardized with respect to children aged 5 in the 2022 assessments. The sample is restricted to children within 5 percentage points of the age-specific median probability of private enrolment. These probabilities are estimated using age-specific probit regression of private enrolment on lagged test scores, SES, parental education, gender, and village fixed effects. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

Table D.2: Private school value-added in preschool and primary school (probit restriction: 10 percentage points)

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Private school	0.750***	0.695***	0.196	0.204*	-0.019	-0.018
	(0.111)	(0.107)	(0.101)	(0.097)	(0.033)	(0.032)
Math IRT score in 2022		0.270***		0.299***		0.325***
		(0.075)		(0.088)		(0.017)
Panel B: Tamil						
Private school	0.670***	0.641***	0.058	0.067	-0.193***	-0.183***
	(0.102)	(0.100)	(0.089)	(0.088)	(0.038)	(0.036)
Tamil IRT score in 2022		0.273***		0.050		0.316***
		(0.060)		(0.083)		(0.017)
Controls	Village FE	All	Village FE	All	Village FE	All
Observations	373	373	565	565	4,036	4,036

Notes: Robust standard errors, clustered at the village level, in parentheses. Columns 1, 3 and 5 show raw means by private school attendance within villages. Columns 2, 4 and 6 include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender. Test scores refer to equated IRT scores, standardized with respect to children aged 5 in the 2022 assessments. The sample is restricted to children within 10 percentage points of the age-specific median probability of private enrolment. These probabilities are estimated using age-specific probit regression of private enrolment on lagged test scores, SES, parental education, gender and village fixed effects. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

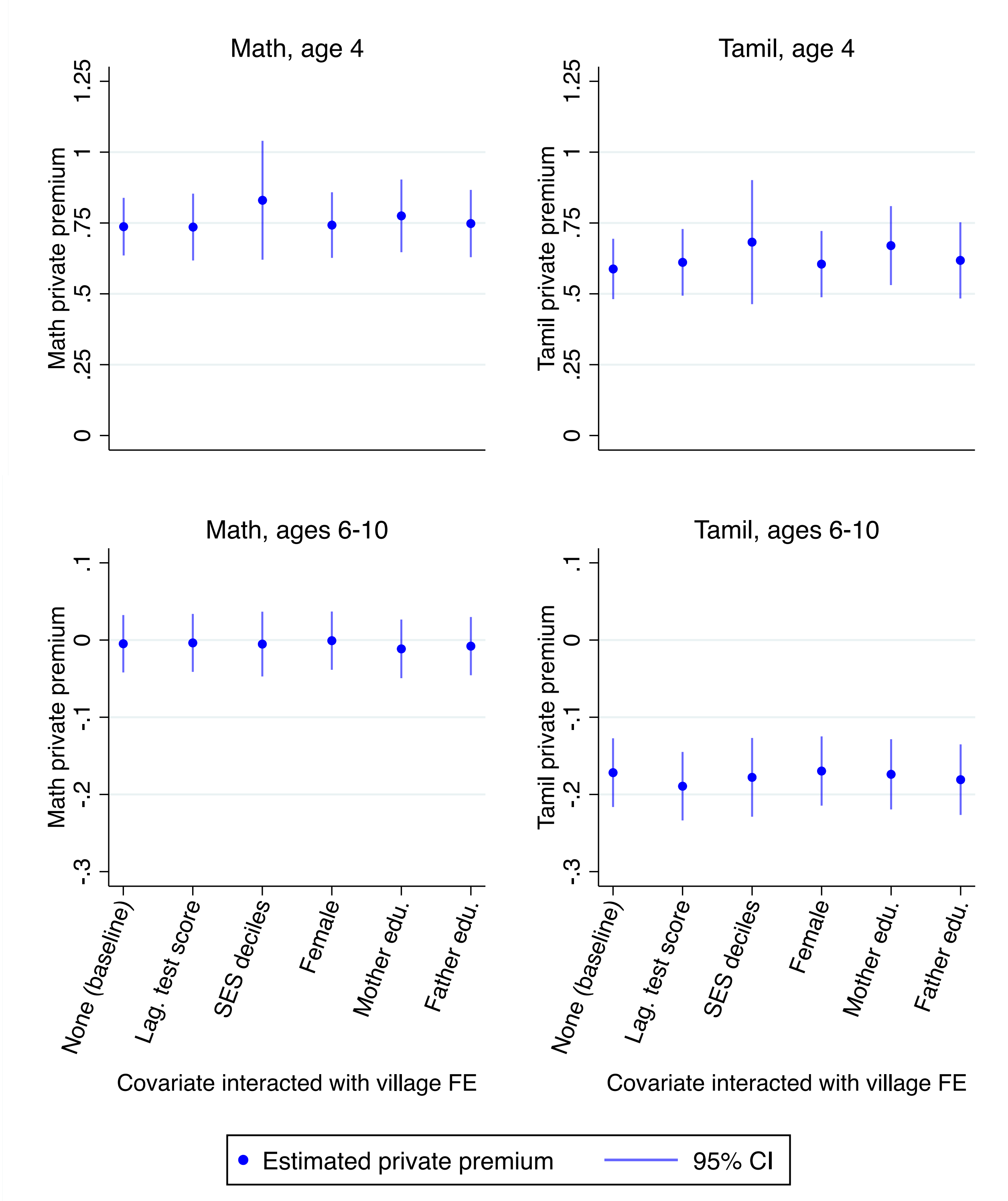
Table D.3: Private school value-added in preschool and primary school (probit restriction: 15 percentage points)

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Private school	0.740***	0.709***	0.117	0.123	0.018	0.000
	(0.077)	(0.077)	(0.076)	(0.075)	(0.029)	(0.028)
Math IRT score in 2022		0.197***		0.228***		0.310***
		(0.055)		(0.055)		(0.014)
Panel B: Tamil						
Private school	0.622***	0.610***	-0.034	-0.014	-0.179***	-0.179***
	(0.081)	(0.082)	(0.073)	(0.072)	(0.032)	(0.030)
Tamil IRT score in 2022		0.261***		0.126*		0.316***
		(0.053)		(0.056)		(0.015)
Controls	Village FE	All	Village FE	All	Village FE	All
Observations	607	607	926	926	6,243	6,243

Notes: Robust standard errors, clustered at the village level, in parentheses. Columns 1, 3 and 5 show raw means by private school attendance within villages. Columns 2, 4 and 6 include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender. Test scores refer to equated IRT scores, standardized with respect to children aged 5 in the 2022 assessments. The sample is restricted to children within 15 percentage points of the age-specific median probability of private enrolment. These probabilities are estimated using age-specific probit regression of private enrolment on lagged test scores, SES, parental education, gender and village fixed effects. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

D.3 Interacting covariates with village indicators

Figure D.2: Robustness of private premium to interacting controls with village FEs at different ages



Notes: These figures show the estimated average private premium under alternative specifications that allow each of the control variables specified along the horizontal axis to vary across villages. In practice, this amounts to interacting the covariate with village-level indicators.

E Empirical Bayes shrinkage and bias-corrected value-added covariances

E.1 Empirical Bayes shrinkage

This section details the construction of the Empirical Bayes estimates used throughout the paper. We follow a simplified version of the approach used by Andrabi et al. (2025). Let

$$y_{islv} = \theta_{slv} + \mathbf{\Gamma X}_{islv} + \epsilon_{islv} \tag{2}$$

where y_{islv} is the test score of child i in the private/public sector s and pre-school/primary level l, in village v. θ_{slv} is the village-level average value-added in a given sector, at a given level. $\mathbf{X}_{islv}$ is a vector of controls (lagged test scores, SES index deciles, maternal and paternal education, gender, and age), and ϵ_{islv} is an idiosyncratic error term. The variance of value-added, denoted σ^2_{sl}, is common across villages but allowed to differ between sectors. The variance of the error term is denoted by σ^2_{ϵ}. Both are assumed to be independent and homoskedastic. We denote the number of children in a given village-level-sector cell as N_{slv}.

Our estimate of θ_{slv} (i.e., the village-level-sector fixed effect) is

$$\hat{\theta}_{slv} = \theta_{slv} + \frac{1}{N_{slv}} \sum_{i \in slv} \epsilon_{islv} \tag{3}$$

The variance of this estimate is equal to

$$\begin{aligned} Var(\hat{\theta}_{slv}) &= E\Big[\Big(\theta_{slv} + \frac{1}{N_{slv}} \sum_{i \in slv} \epsilon_{islv}\Big)^2\Big] && (4) \\ &= E(\theta^2_{slv}) + E\Big(\frac{1}{N^2_{slv}} \sum_{i \in slv} \epsilon^2_{islv}\Big) && (5) \\ &= \sigma^2_{sl} + E\Big(\frac{1}{N_{slv}} \sigma^2_{\epsilon}\Big) && (6) \end{aligned}$$

The second equality follows from the assumption that ϵ_{islv} is independent and identically distributed at the child level. Rearranging terms, the variance of value-added purged of estimation error is equal to

$$\sigma_{sl}^2 = Var(\hat{\theta}_{slv}) - E\Big(\frac{1}{N_{slv}}\sigma_\epsilon^2\Big) \tag{7}$$

We can obtain an estimator of the left-hand side by plugging in moment estimators on the right-hand side. $Var(\hat{\theta}_{slv})$ is estimated as the sample variance of the fixed effects.[1] The variance of the error term, σ_ϵ^2, is estimated using residuals from Equation (2). An estimate of σ_{sl}^2 in Equation (7) is obtained by taking the average of the right-hand side.

Given a standard hierarchical model with normal priors, the Empirical Bayes scaling term is then given by

$$h_{slv} = \frac{\sigma_{sl}^2}{\sigma_{sl}^2 + \frac{1}{N_{slv}}\sigma_\epsilon^2} \tag{8}$$

We shrink each fixed effect toward its level-sector mean $\hat{\mu}_{sl} = \frac{1}{V}\sum_{v=1}^{V}\hat{\theta}_{slv}$, where V is the number of villages. The Empirical Bayes estimate of average value-added in a village-level-sector cell is therefore given by:

$$\hat{\theta}_{slv}^{EB} = h_{slv} \cdot \hat{\theta}_{slv} + (1 - h_{slv}) \cdot \hat{\mu}_{sl} \tag{9}$$

Intuitively, as the sample size of a given cell (N_{slv}) approaches infinity, h_{slv} tends to 1 such that the Empirical Bayes estimate is simply equal to the fixed effect. At the other extreme, the Empirical Bayes estimate shrinks the coefficient completely to the level-sector mean $\hat{\mu}_{sl}$.

[1] V denotes the number of villages: $Var(\hat{\theta}_{slv}) = \frac{1}{V}\sum_{v=1}^{V}(\hat{\theta}_{slv} - \hat{\mu}_{sl})^2$, where $\hat{\mu}_{sl}$ is equal to $\frac{1}{V}\sum_{v=1}^{V}\hat{\theta}_{slv}$.

E.2 Bias-corrected estimation of $\mathbf{\Sigma}_\theta$

Let θ_{sljv} be the estimate of value-added in sector s, level l, subject j and village v. To simplify notation, let k denote the combinations of levels, sectors and subjects, so that:

$$k = \begin{cases} 1 & \text{public, preschool, math} \\ 2 & \text{private, preschool, math} \\ 3 & \text{public, primary school, math} \\ 4 & \text{private, primary school, math} \\ ... \end{cases}$$

and so on for Tamil. Hence, θ_{1v} is the math test score gains in public preschools in village v. Denote the 8×8 covariance matrix of these value-added estimates across sectors, levels and subjects as $\mathbf{\Sigma}_\theta$, where element k, m is given by:

$$\sigma_{km} = Cov(\theta_{kv}, \theta_{mv}). \tag{10}$$

Sample moments of our estimated value-added parameters will not, in general, provide unbiased estimators of $\mathbf{\Sigma}_\theta$ due to excess estimation noise. Instead, we closely follow the approach of Angrist et al. (2025) in estimating this covariance matrix. We can write the covariance in Equation (10) as quadratic forms of the underlying (true) value-added parameters:

$$\sigma_{km} = \boldsymbol{\theta}' \mathbf{A}_{km} \boldsymbol{\theta}, \tag{11}$$

where $\boldsymbol{\theta} = (\theta_{11}, \theta_{12}, ..., \theta_{1V}, \theta_{21}, ..., \theta_{8V})'$ is a $8V \times 1$ vector of value-added parameters: eight for each of a total of V villages. $\mathbf{A}_{km}$ is a $8V \times 8V$ matrix of $V \times V$ blocks of zeros, except for the (k, m)th block which is a centering matrix: $(V-1)^{-1}(\mathbf{I}_V - V^{-1} l_V l_V')$. Here, $\mathbf{I}_V$ is the $V \times V$ identity matrix, and l_V is a column vector of 1's with length V.

Even if our estimates $\hat{\boldsymbol{\theta}}$ are unbiased (i.e., $E[\hat{\boldsymbol{\theta}}] = \boldsymbol{\theta}$), each of its elements contains some degree of noise. Let the sampling variance matrix of $\boldsymbol{\theta}$ be denoted by $8V \times 8V$ matrix $\mathbf{V} = E[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})]$.

As shown in Walters (2024), the bias in sample analogues of σ_{km} can be written as:

$$E[\hat{\boldsymbol{\theta}}' \mathbf{A}_{km} \boldsymbol{\theta}] = \boldsymbol{\theta}' \mathbf{A}_{km} \boldsymbol{\theta} + tr(\mathbf{A}_{km} \mathbf{V}). \tag{12}$$

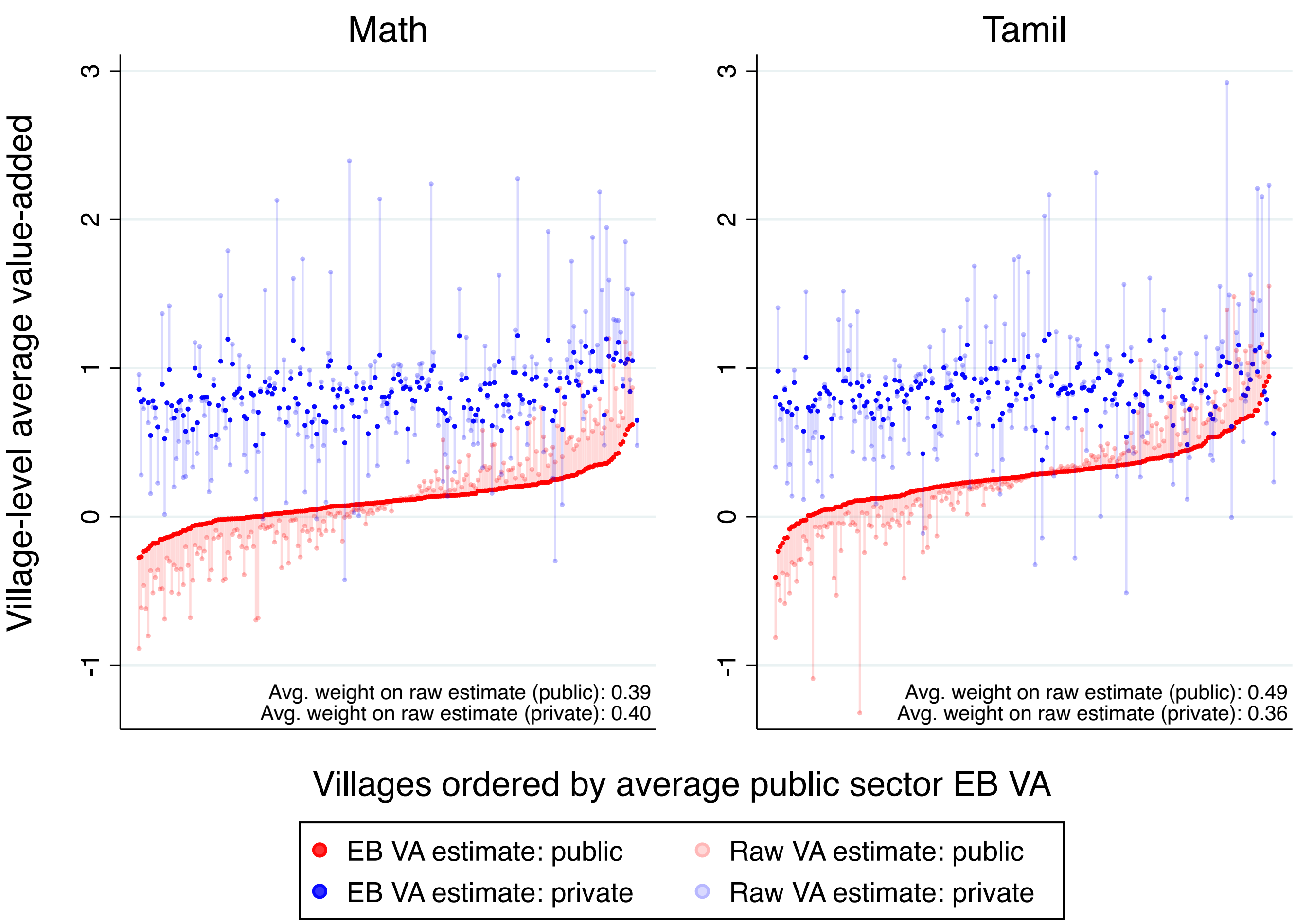
Hence, estimation error gives rise to an additional term on the left-hand side. The formula for bias-corrected estimates of σ_{km} subtracts this bias term directly:

$$\hat{\sigma}_{km} = \hat{\boldsymbol{\theta}}' \mathbf{A}_{km} \hat{\boldsymbol{\theta}} - tr(\mathbf{A}_{km} \hat{\mathbf{V}})$$

where $\hat{\mathbf{V}}$ is the heteroskedasticity-robust covariance matrix obtained from estimation of $\hat{\boldsymbol{\theta}}$ in seemingly unrelated regressions, stacking regressions on each of the outcomes on level-sector-village indicators and controls. A bias-corrected estimate of the slope of a regression of $\hat{\theta}_{kv}$ on $\hat{\theta}_{mv}$ is simply given by $\hat{\gamma}_{km} = \hat{\sigma}_{km} / \hat{\sigma}_{mm}$, i.e., the ratio of the bias-corrected covariance and the bias-corrected variance of the "left-hand side" variable. For instance, $\hat{\gamma}_{21}$ would denote the slope of a regression of private on public preschool math value-added across villages.

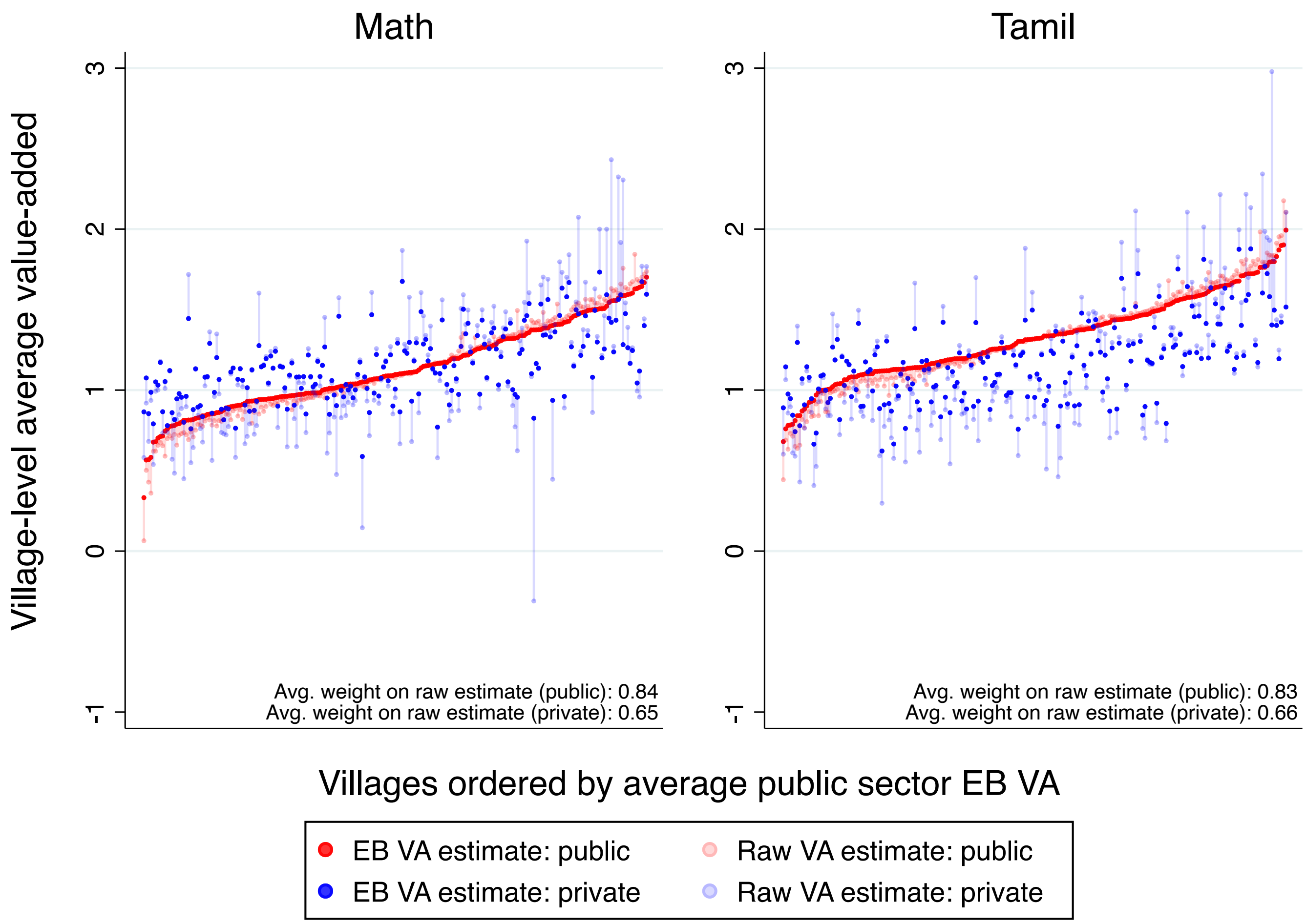
E.3 Impact of Empirical Bayes shrinkage

Figure E.1: Raw and shrunken village value-added estimates: preschool



Notes: These figures show Empirical Bayes estimates (solid colors) and raw value-added estimates (faded colors) separately for public and private preschools. Villages are ordered, along the horizontal axis, by their Empirical Bayes estimates of public sector productivity.

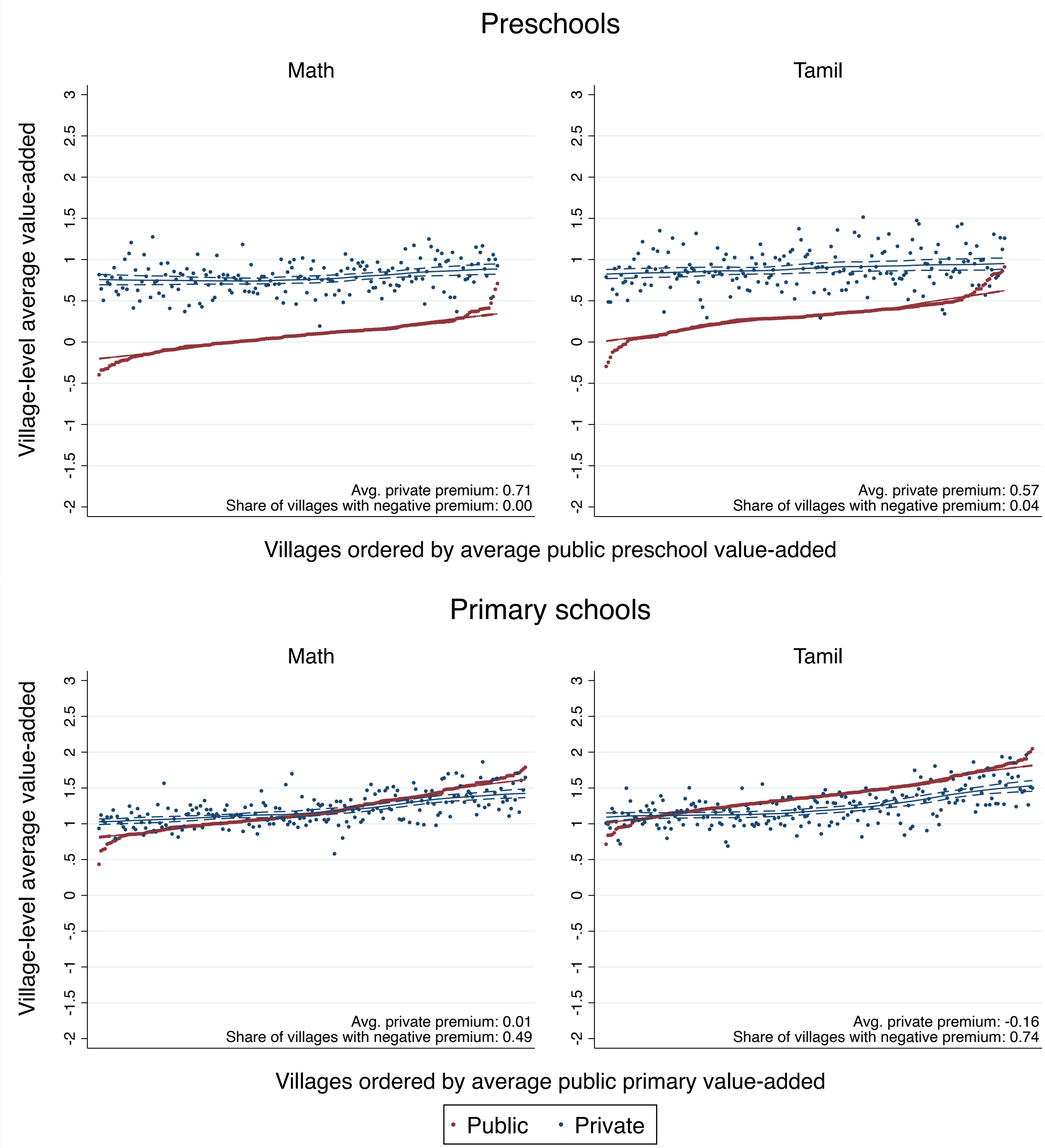
Figure E.2: Raw and shrunken village value-added estimates: primary school



Notes: These figures show Empirical Bayes estimates (solid colors) and raw value-added estimates (faded colors) separately for public and private primary schools. Villages are ordered, along the horizontal axis, by their Empirical Bayes estimates of public sector productivity.

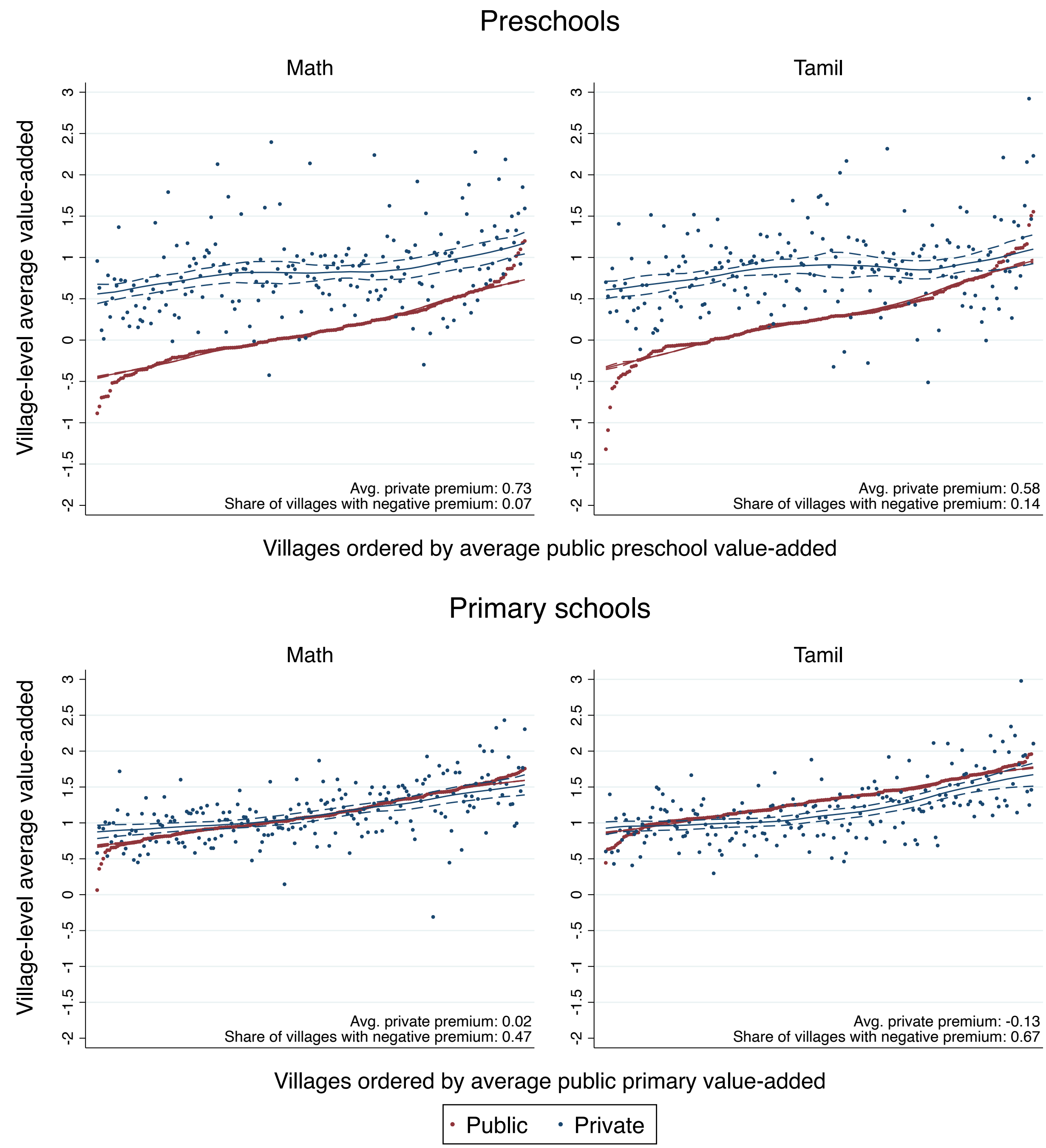
E.4 Village private premia without Empirical Bayes, 5-year-olds, and both

Figure E.3: Village-level average value-added: without 5-year-olds



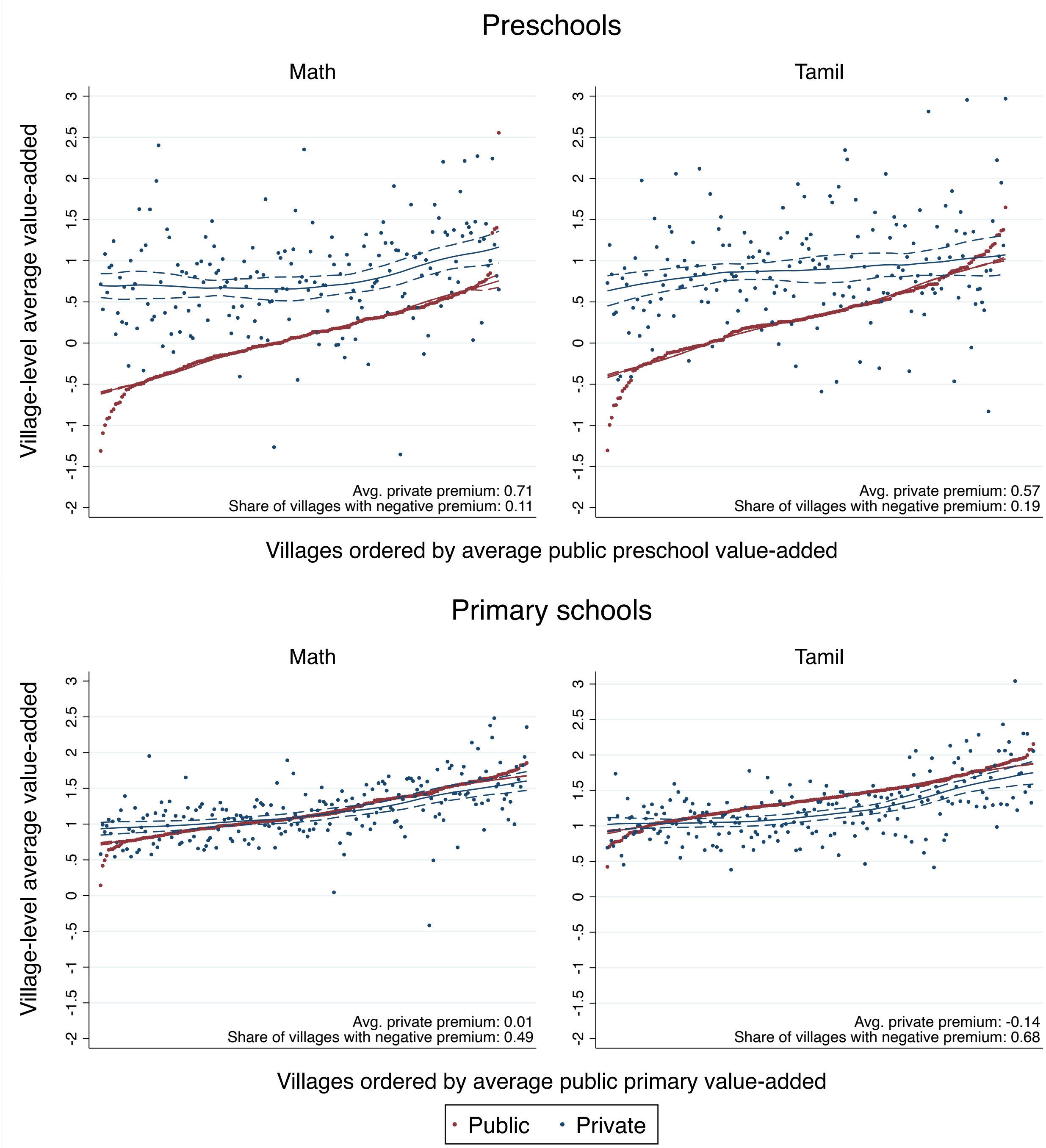
Notes: These figures show village-level average school value-added by sector (public/private) and level (preschool/primary) using Empirical Bayes estimates, excluding children of age 5. Villages are ordered along the x-axis by their average value-added in government schools. The regression specification generating these estimates is given by Equation 2.

Figure E.4: Village-level average value-added: without Empirical Bayes



Notes: These figures show village-level average school value-added by sector (public/private) and level (preschool/primary) using raw value-added estimates (rather than Empirical Bayes estimates). Villages are ordered along the x-axis by their average value-added in government schools. The regression specification generating these estimates is given by Equation 2.

Figure E.5: Village-level average value-added: without Empirical Bayes and 5-year-olds



Notes: These figures show village-level average school value-added by sector (public/private) and level (preschool/primary) using raw value-added estimates (rather than Empirical Bayes estimates), excluding children of age 5. Villages are ordered along the x-axis by their average value-added in government schools. The regression specification generating these estimates is given by Equation 2.

E.5 Precision dependence and Empirical Bayes shrinkage

This procedure may be problematic in settings where effect sizes (i.e., value-added) are correlated with precision. In our context, such correlation could arise if villages with higher private-sector value-added also have greater private school enrolment, reflecting parental demand for test score gains. In that case, lower value-added estimates would tend to come from villages with fewer children enrolled in private options, resulting in less precise estimates that are disproportionately shrunk toward the sectoral mean. This would be problematic for our investigation of the relationship between private premia and village-level market shares of the private sector.

We address this by adopting an alternative Empirical Bayes shrinkage approach that allows for precision dependence (Chen 2024; Walters 2024). This approach models value-added in a particular sector and pre-/primary school level — for simplicity denoted as θ_v — as a non-linear function of its standard error s_v, plus a constant:

$$\theta_v = \psi_0 + \psi_1 \log s_v + s_v^{\psi_2} r_v,$$

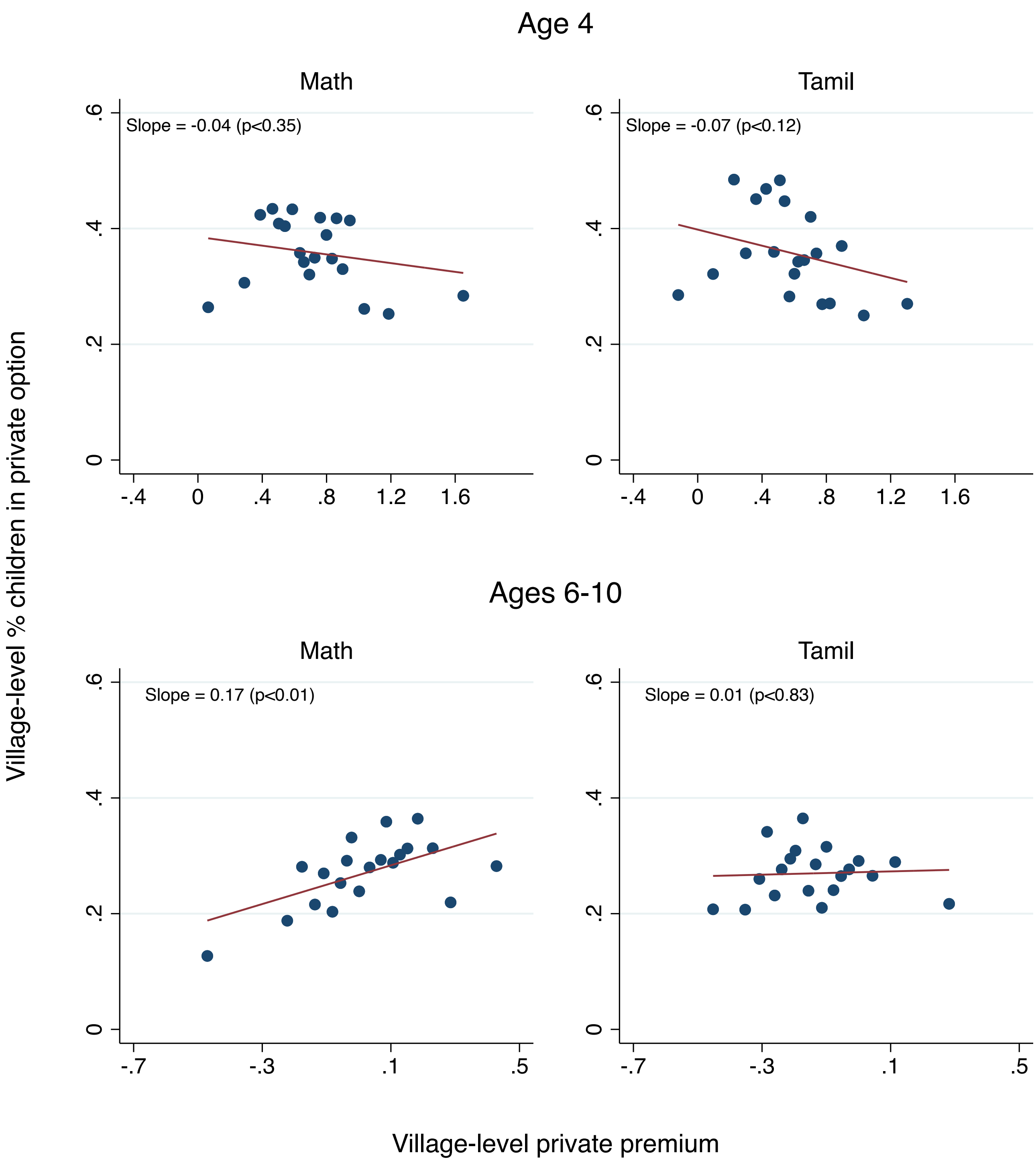
where $r_v|s_v$ is drawn from a normal distribution with mean zero and variance σ_r^2. It can be shown that the Empirical Bayes posterior prediction of θ_v can be written as:

$$\theta_v^{EB} = \hat{\psi}_0 + \hat{\psi}_1 \log s_v + s_v^{\hat{\psi}_2} \hat{r}_v^*,$$

where $\hat{r}_v^* = [\hat{\sigma}_r^2 / (\hat{\sigma}_r^2 + s_v^{2(1-\hat{\psi}_2)})] \cdot [s_v^{\hat{\psi}_2} / (\hat{\theta}_v - \hat{\psi}_0 - \hat{\psi}_1 \log s_v)]$. The estimates $\hat{\psi}_0$ and $\hat{\psi}_1$ are coefficients from a regression of value-added $\hat{\theta}_v$ on $\log s_v$ plus a constant. Estimates $\hat{\psi}_2$ and $\hat{\sigma}_r^2$ come from a nonlinear least squares regression of $[(\hat{\theta}_v - \hat{\psi}_0 - \hat{\psi}_1 \log s_v)^2 - s_v^2]$ on $s_v^{2\psi_2} \sigma_r^2$.

Figure E.6 shows the correlation between private sector market shares and the private premium across villages. In practice, the alternative Empirical Bayes approach gives similar results as in the main analysis.

Figure E.6: Village-level private premium and share of private enrolment under precision-dependent Empirical Bayes shrinkage

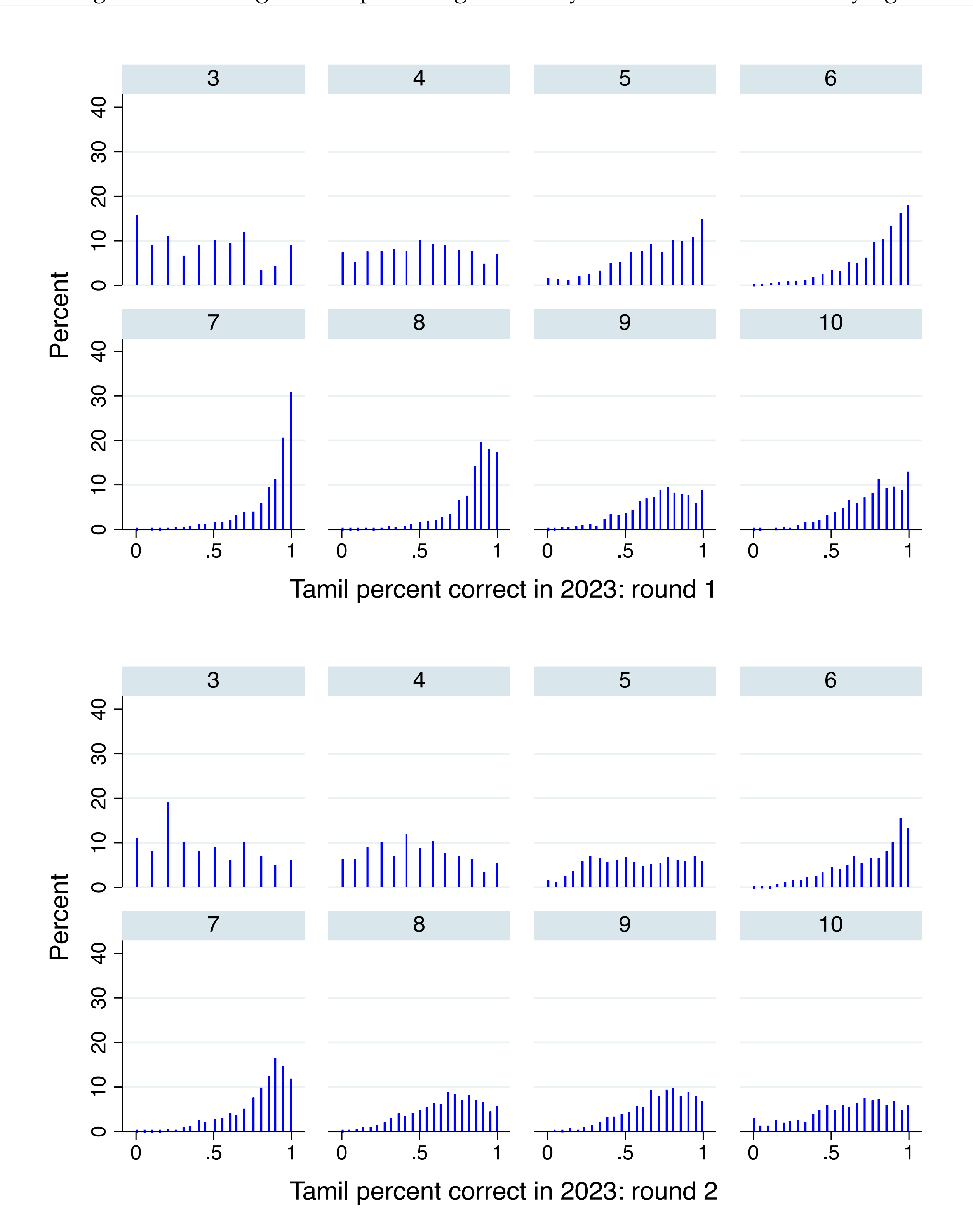


Notes: These figures depict the correlation between (1) the village-level share of children in a private school and (2) the difference between the average *private* and *public* school value-added in the same village (i.e., the private premium). These value-added measures are Empirical Bayes estimates allowing for value-added estimates to be correlated with the size of their standard errors, estimated as described in Appendix E.5.

F Main value-added results using round 2 data in 2023

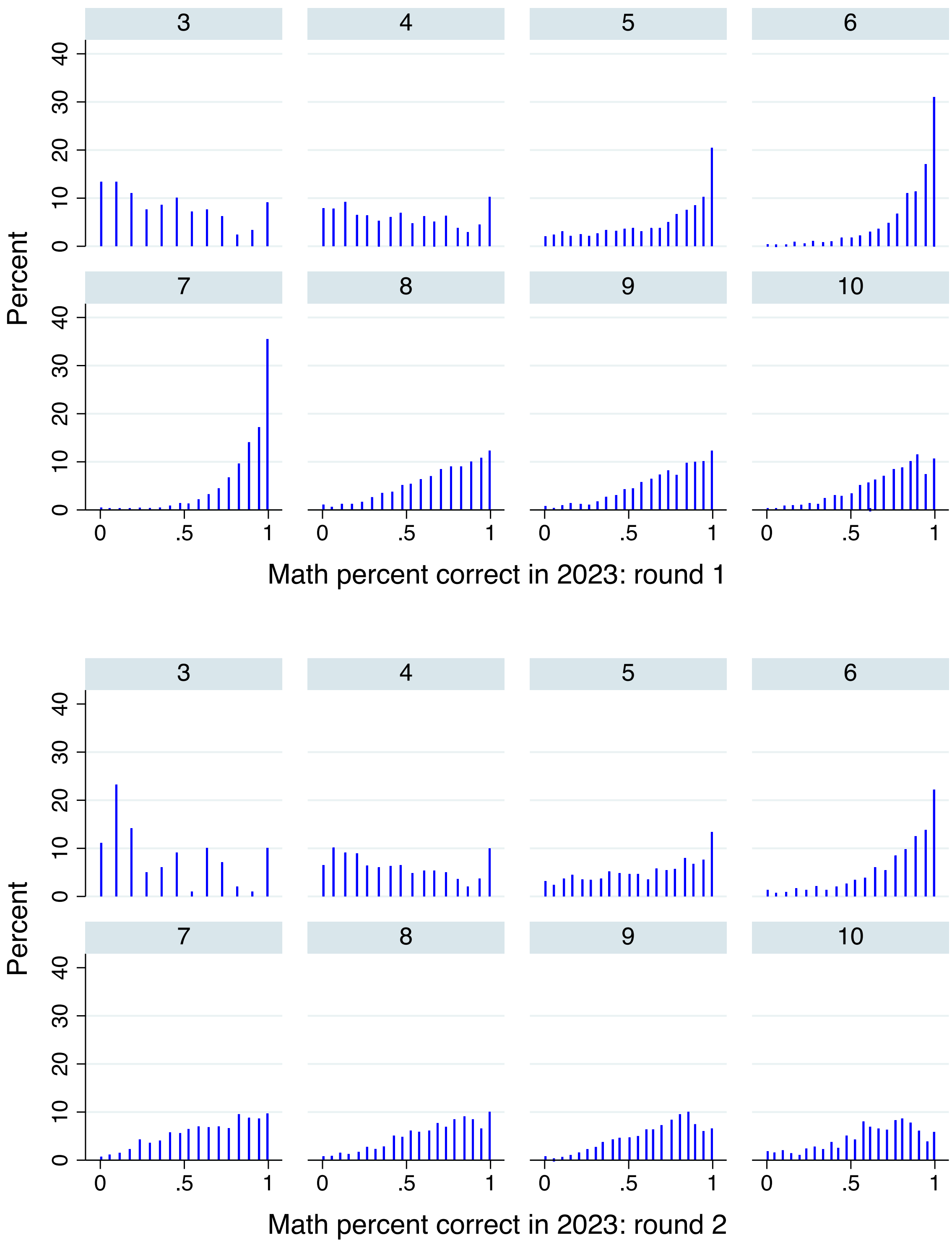
F.1 Test score distributions in 2023 by round

Figure F.1: Histograms of percentage correctly answered Tamil items by age



Notes: These figures show histograms of percent correctly answered Tamil items, separately by age, in the two assessment rounds of 2023. Each bin captures the density by steps of 1 percentage point.

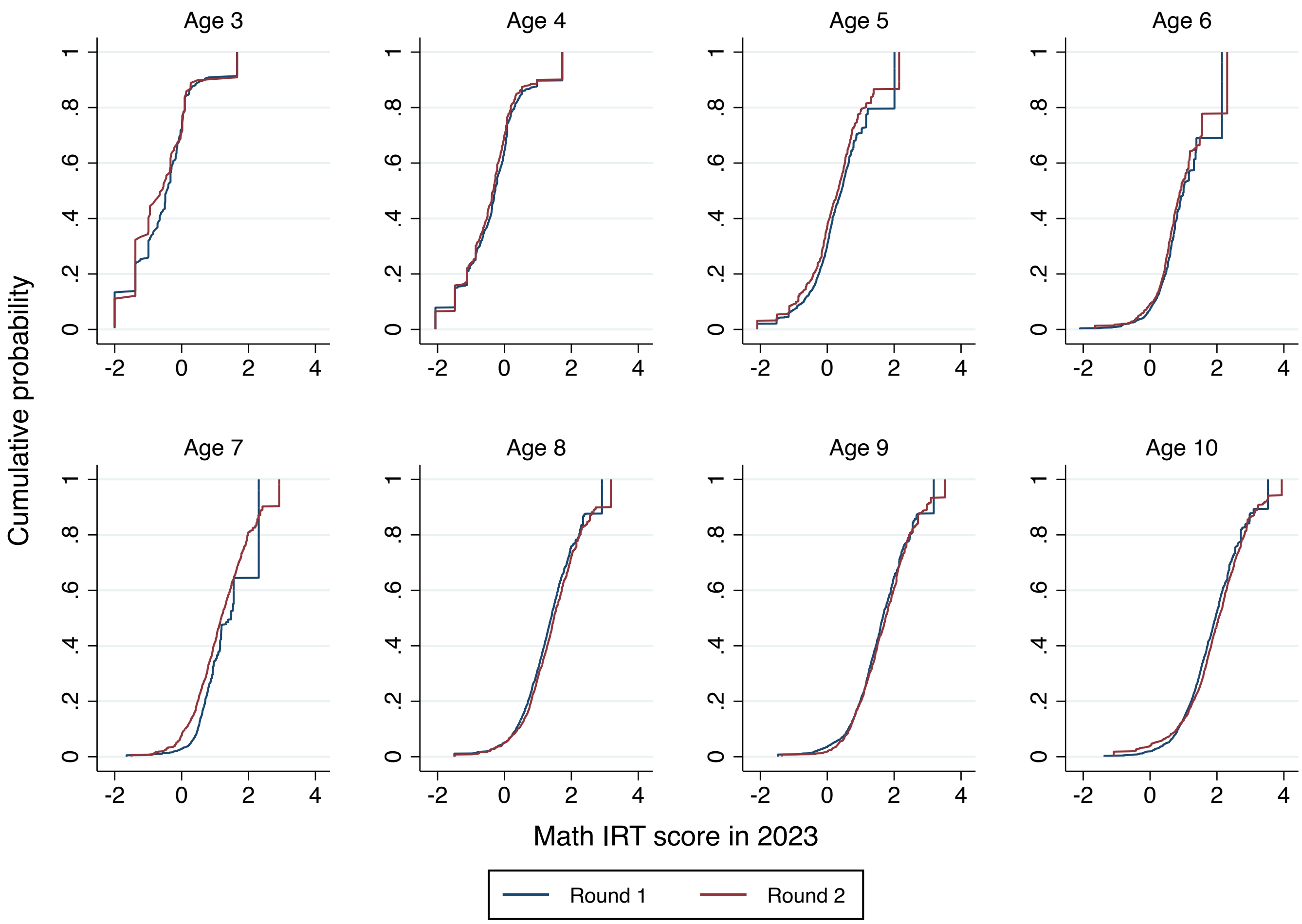
Figure F.2: Histograms of percentage correctly answered math items by age



Notes: These figures show histograms of percent correctly answered math items, separately by age, in the two assessment rounds of 2023. Each bin captures the density by steps of 1 percentage point.

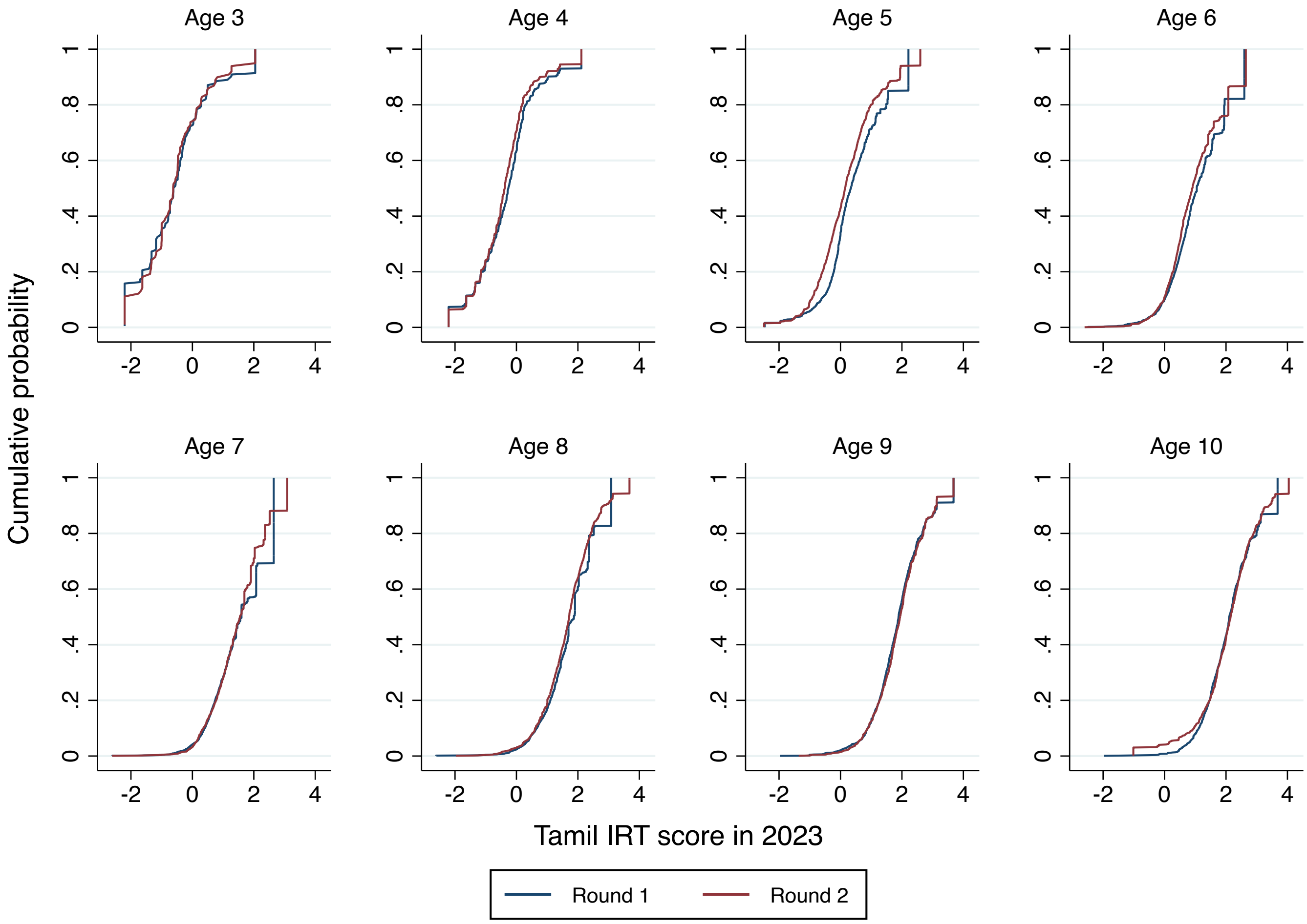
F.2 IRT score distributions in 2023 by age and round

Figure F.3: CDFs of math IRT scores in 2023 by age and round



Notes: These figures show cumulative distribution functions (CDFs) of IRT EAP scores in math in 2023, by child age and assessment round.

Figure F.4: CDFs of Tamil IRT scores in 2023 by age and round



Notes: These figures show cumulative distribution functions (CDFs) of IRT EAP scores in Tamil in 2023, by child age and assessment round.

F.3 The private premium using only round 2 assessments in 2023

Table F.1: Private school value-added in preschool and primary school, round 2 assessments in 2023

	Age 4		Age 5		Ages 6–10	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Math						
Private school	0.873***	0.817***	0.395***	0.254**	0.094**	0.003
	(0.082)	(0.087)	(0.071)	(0.076)	(0.034)	(0.031)
Math IRT score in 2022		0.142***		0.213***		0.313***
		(0.040)		(0.038)		(0.015)
Constant	-0.619***	-0.443***	0.218***	0.414***	1.440***	1.194***
	(0.028)	(0.052)	(0.027)	(0.043)	(0.009)	(0.014)
Panel B: Tamil						
Private school	0.742***	0.690***	0.230***	0.096	-0.098**	-0.145***
	(0.089)	(0.088)	(0.066)	(0.076)	(0.034)	(0.030)
Tamil IRT score in 2022		0.163***		0.198***		0.321***
		(0.035)		(0.037)		(0.015)
Constant	-0.608***	-0.398***	0.189***	0.382***	1.659***	1.401***
	(0.030)	(0.051)	(0.025)	(0.044)	(0.009)	(0.014)
Controls	Village FE	All	Village FE	All	Village FE	All
Observations	748	748	1,114	1,114	5,528	5,528

Notes: Robust standard errors, clustered at the village level, in parentheses. Columns 1, 3, and 5 show raw means by private school attendance within villages. Columns 2, 4, and 6 include lagged scores, village fixed effects, and controls for deciles of the SES wealth index, paternal and maternal education, and child gender. Test scores refer to equated IRT scores, standardized with respect to children aged 5 in the 2022 assessments. The sample is restricted to children assessed in the second round of 2023. $p < 0.01$ = ***, $p < 0.05$ = **, $p < 0.1$ = *.

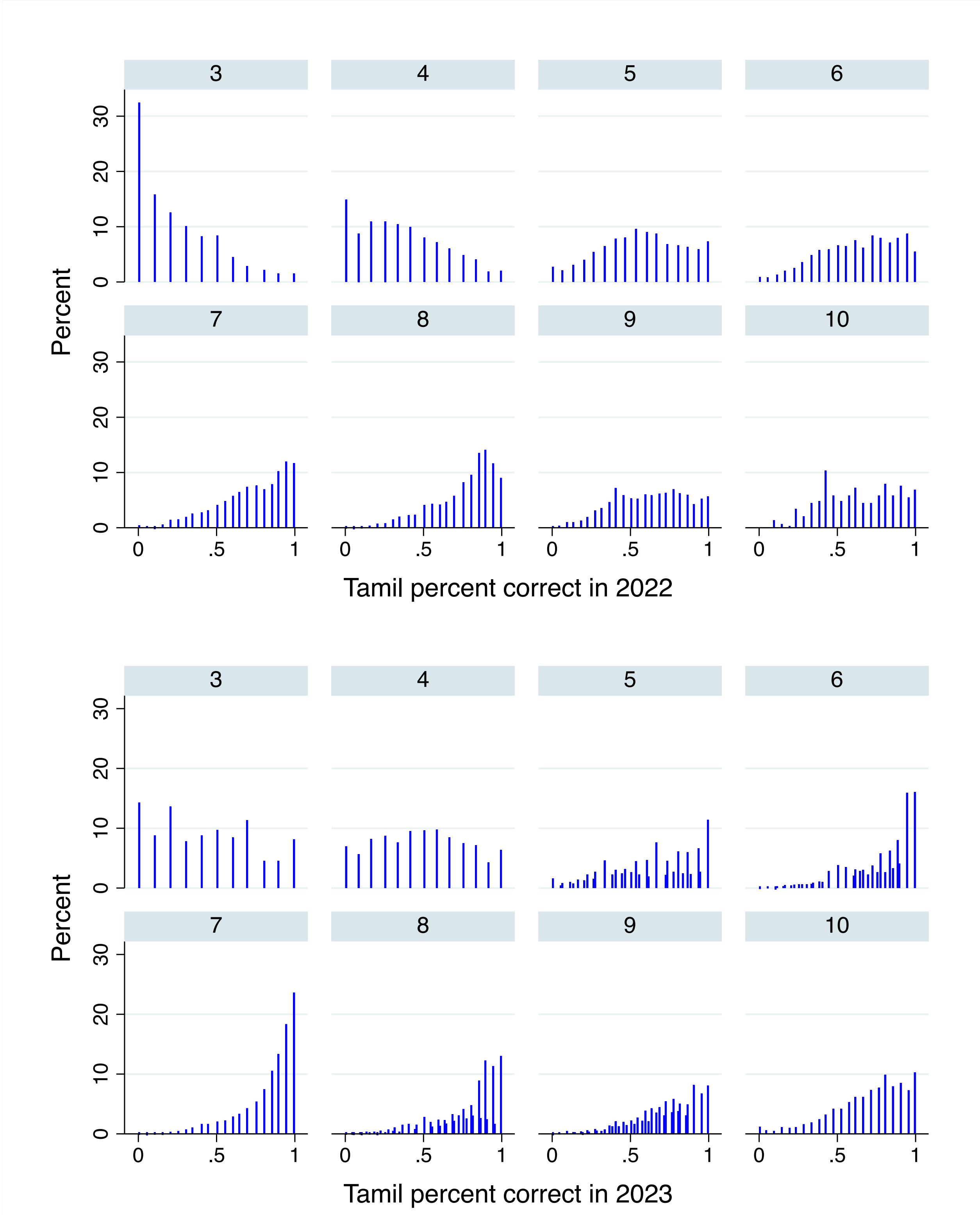
G Details on assessment

Since our data are similar to those used by Singh et al. (2024), details on test content and vertical linking across rounds and ages can be found in the Online Appendix of that paper. A key difference between the data sets used in Singh et al. (2024) and in this paper is that we have an additional round of data collection. Moreover, unlike Singh et al. (2024), we do not use any pre-pandemic data, focusing exclusively on the period after schools reopened.

We provide an overview of the basic details of the test here. Tests were individually administered in Tamil by enumerators during household visits. All tests in both rounds demonstrated strong internal consistency, with Cronbach's alpha exceeding 0.85 in most cases, except for Math and Tamil for 3-year-olds in 2022, which had Cronbach's alpha values of 0.8263 and 0.8070, respectively.

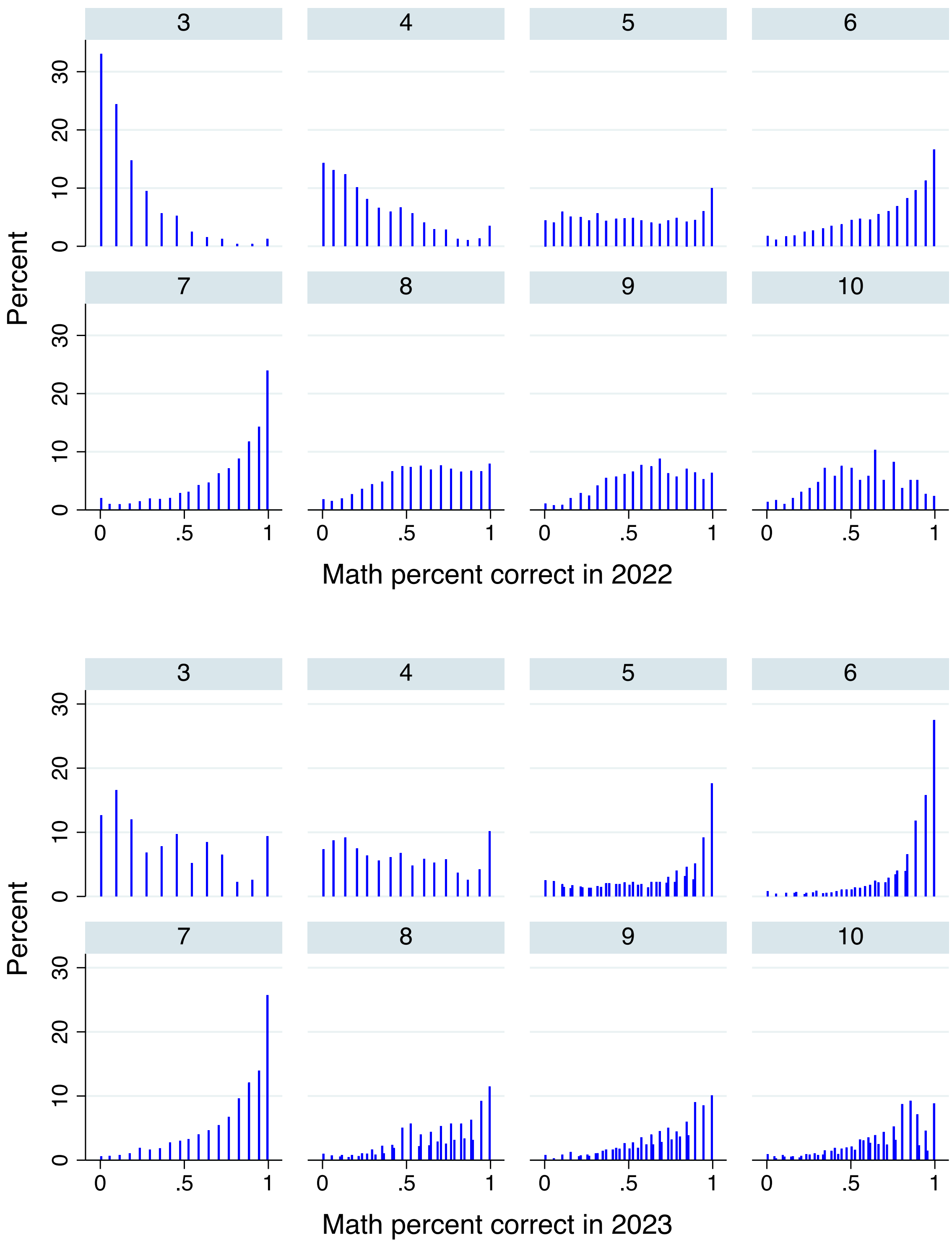
Figures G.1-G.2 show the distribution of correct items by age and year for each subject. While there are some ceiling and floor effects, as mentioned in the main text, the order in which students were tested each year was randomized. In 2023, during the second round of testing, we adapted the tests to prevent ceiling and floor effects (see Appendix F).

Figure G.1: Histograms of percentage correctly answered Tamil items by age



Notes: These figures show histograms of percent correctly answered Tamil items, separately by age, in the assessment waves of 2023 (top) and 2022 (bottom). Each bin captures the density by steps of 1 percentage point.

Figure G.2: Histograms of percentage correctly answered math items by age



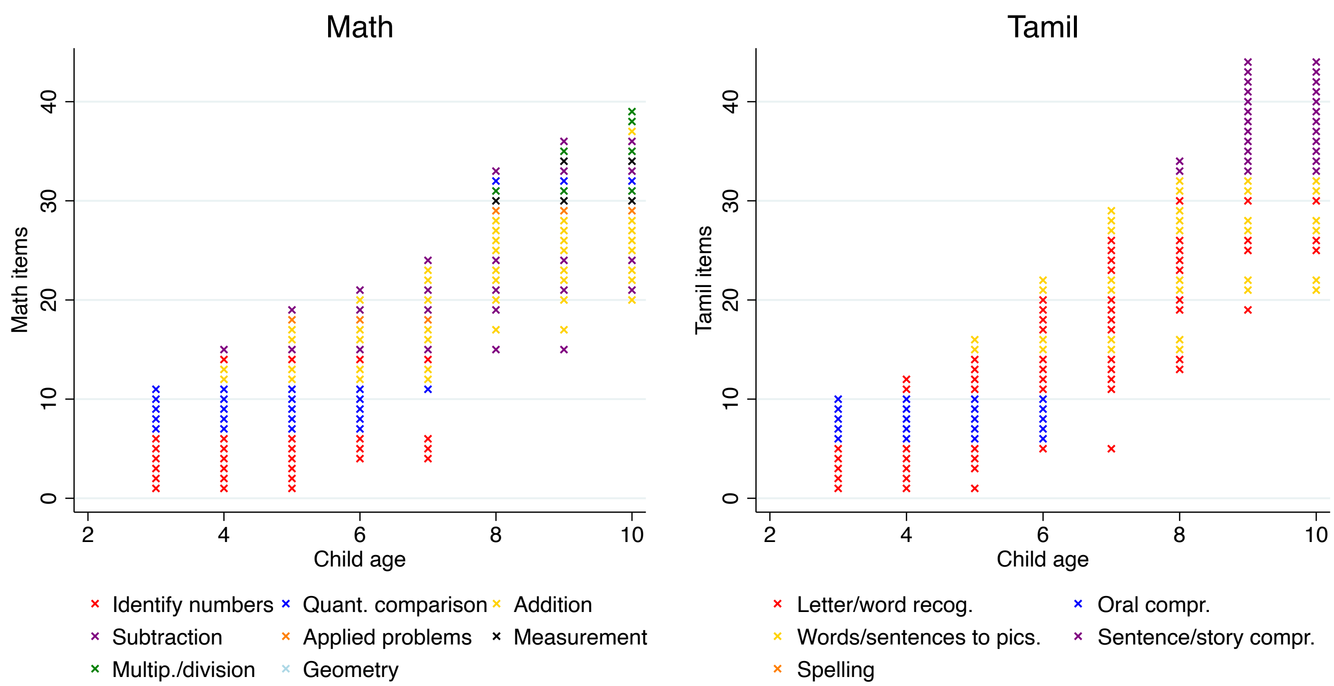
Notes: These figures show histograms of percent correctly answered math items, separately by age, in the assessment waves of 2022 (top) and 2023 (bottom). Each bin captures the density by steps of 1 percentage point.

As a summary measure of the competencies tested and the overlap across ages and rounds, Figure G.3 displays the items asked for each age across years and the competency to which they belong.

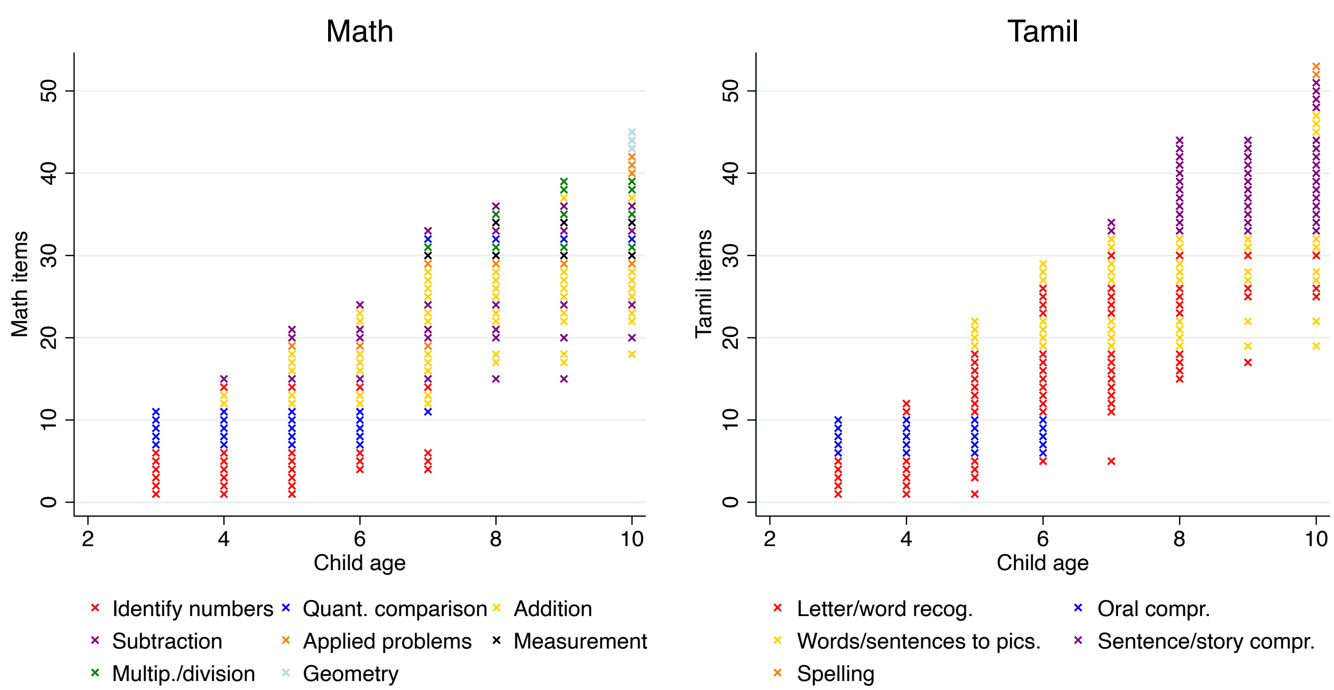
We order the items by the younger age group to which they are administered. There is a high degree of overlap across years and ages to ensure the vertical linking is robust. Figure G.4 shows the items by age and year of inclusion in the assessments. Most items were included in both years; however, some were added in 2023 to mitigate floor and ceiling effects, as discussed above.

Figure G.3: Assessment item maps

(a) 2022

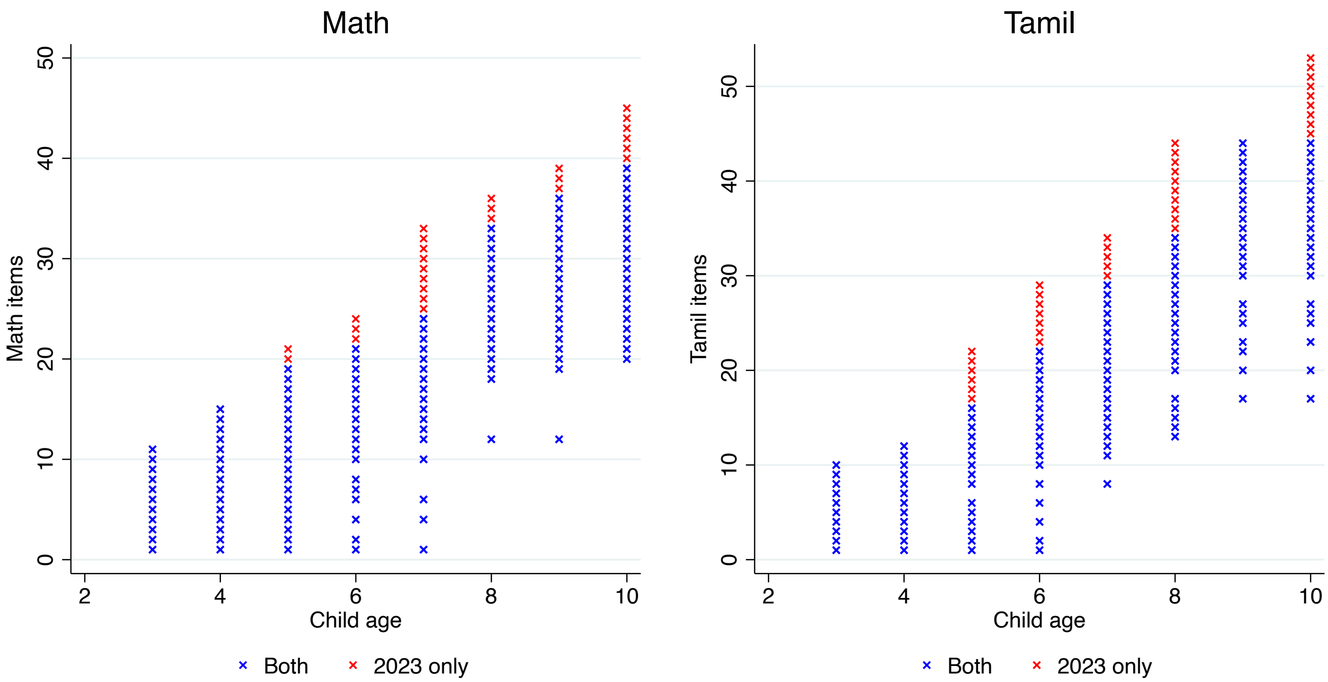


(b) 2023



Notes: These figures show the items administered to children across ages, by competency of the item. The first panel displays the assessment from 2022, while the second panel shows the same assessment from 2023.

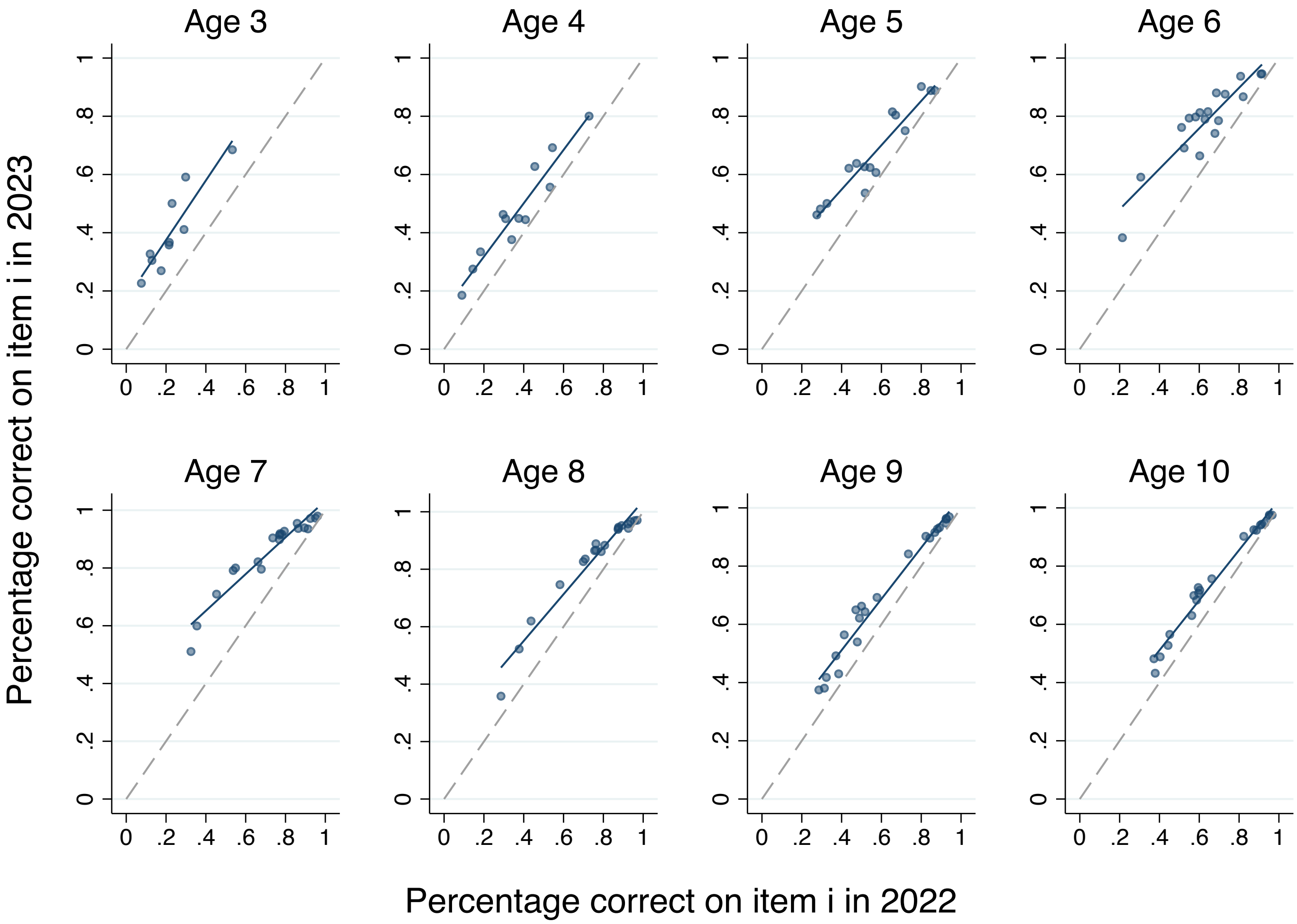
Figure G.4: Assessment items by year



Notes: These figures show the items administered to children across ages, by year of inclusion in the assessments.

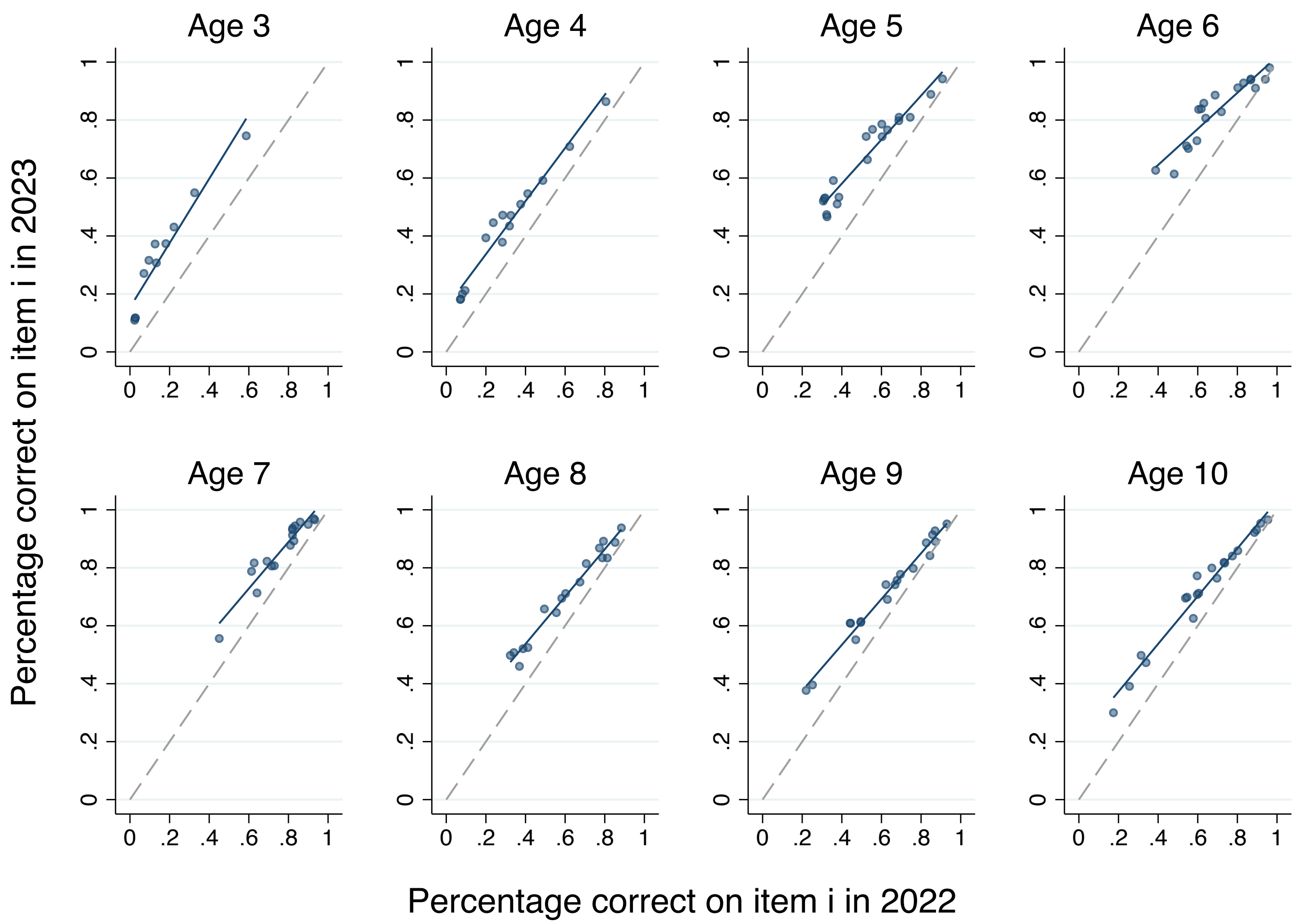
Reassuringly, we find a strong correlation in the proportion of students answering common items correctly across rounds within each age group, indicating stable item-level response patterns (see Figures G.5-G.6). However, performance in 2023 is consistently higher, shifting results vertically above the 45-degree line. This upward shift likely reflects either general improvements in educational outcomes over time.

Figure G.5: Percentage correct on Tamil items across years by age



Notes: These figures show scatter plots of percentage correct on specific Tamil test items, separately by age, in 2022 and 2023. Each dot in the figure corresponds to one item, administered in both rounds.

Figure G.6: Percentage correct on math items across years by age

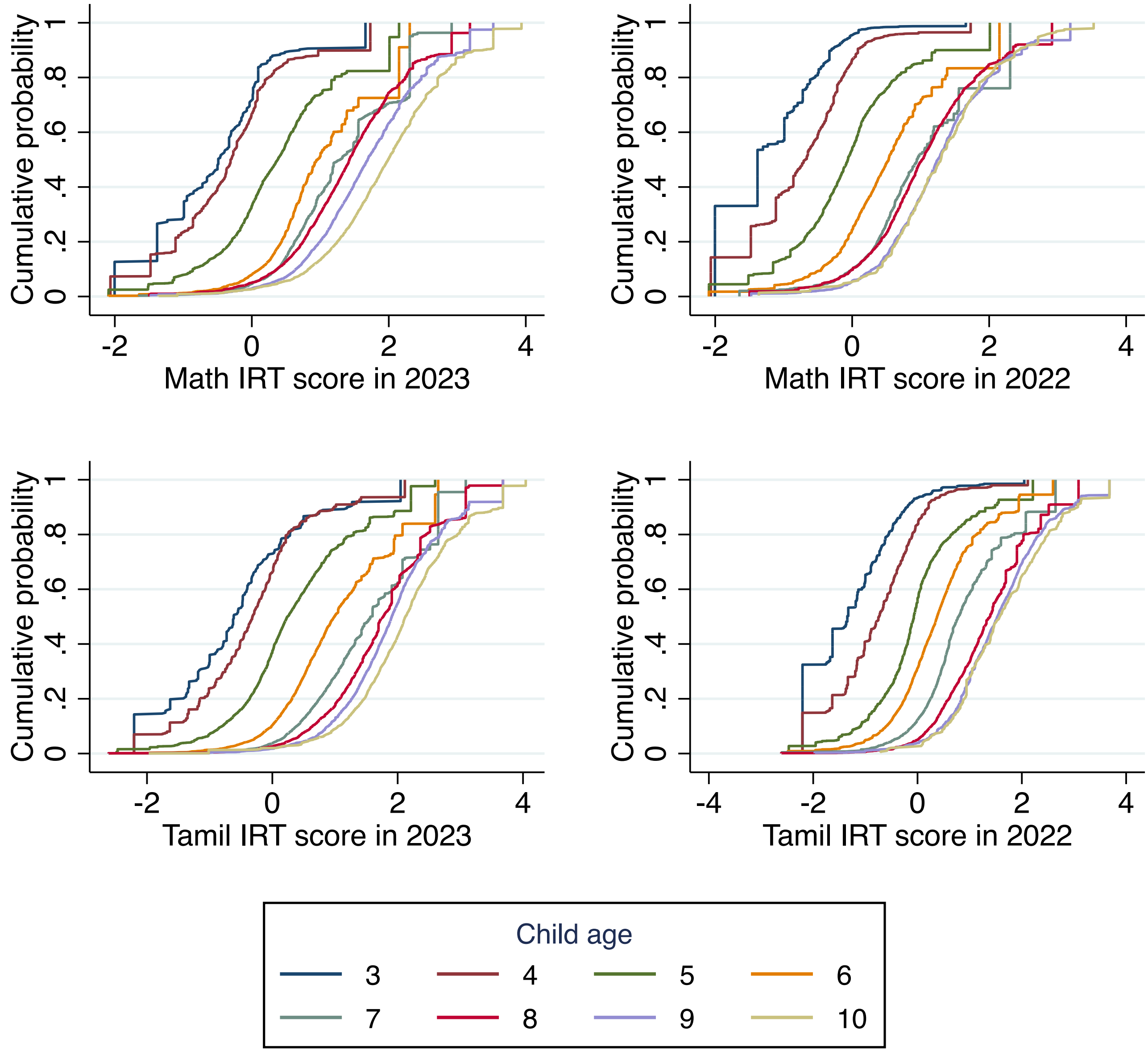


Notes: These figures show scatter plots of percentage correct on specific math test items, separately by age, in 2022 and 2023. Each dot in the figure corresponds to one item, administered in both rounds.

We create comparable, vertically-linked test scores across ages and testing rounds (2022 and 2023) by pooling all responses and estimating Item Response Theory (IRT) scores. Each item was scored dichotomously (correct or incorrect). We employ a two-parameter logistic (2PL) item response theory (IRT) model, which is appropriate for our predominantly open-ended questions, using the `OpenIRT` package in Stata (see https://github.com/tristanz/OpenIRT).

We conduct some basic tests to ensure the vertical linking is adequate. For example, Figure G.7 shows that the distribution of IRT scores increases with age, as expected, across both rounds.

Figure G.7: Cumulative distributions of IRT scores by age

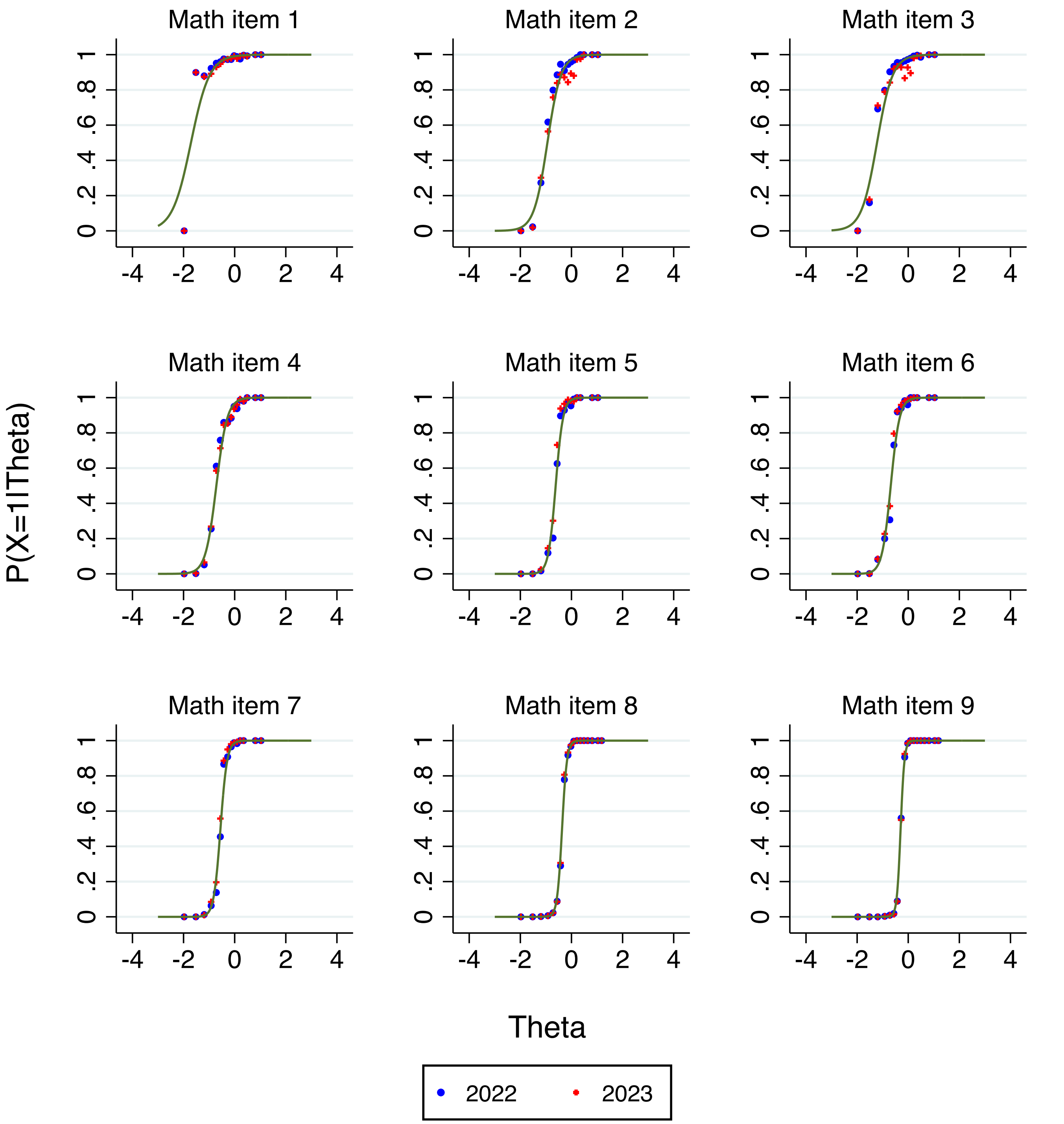


Notes: These figures show the cumulative distribution functions of IRT scores, estimated jointly using both assessment waves, by age and year.

We also show empirical fit to the estimated 2PL Item Characteristic Curve (ICC) for each round in Figures G.8 to G.18, and for children above/below 6 in Figures G.19 to G.29. Overall, questions are able to discriminate between students with different achievement levels (i.e., the ICC monotonically increases, meaning higher ability students are more likely to answer the question correctly), and there is no differential item functioning across rounds or age groups (i.e., students do not have an advantage in answering the question given by the timing of the survey round, and thus the likelihood of answering the question correctly depends on the ability and not the timing of the survey nor age of the child.).

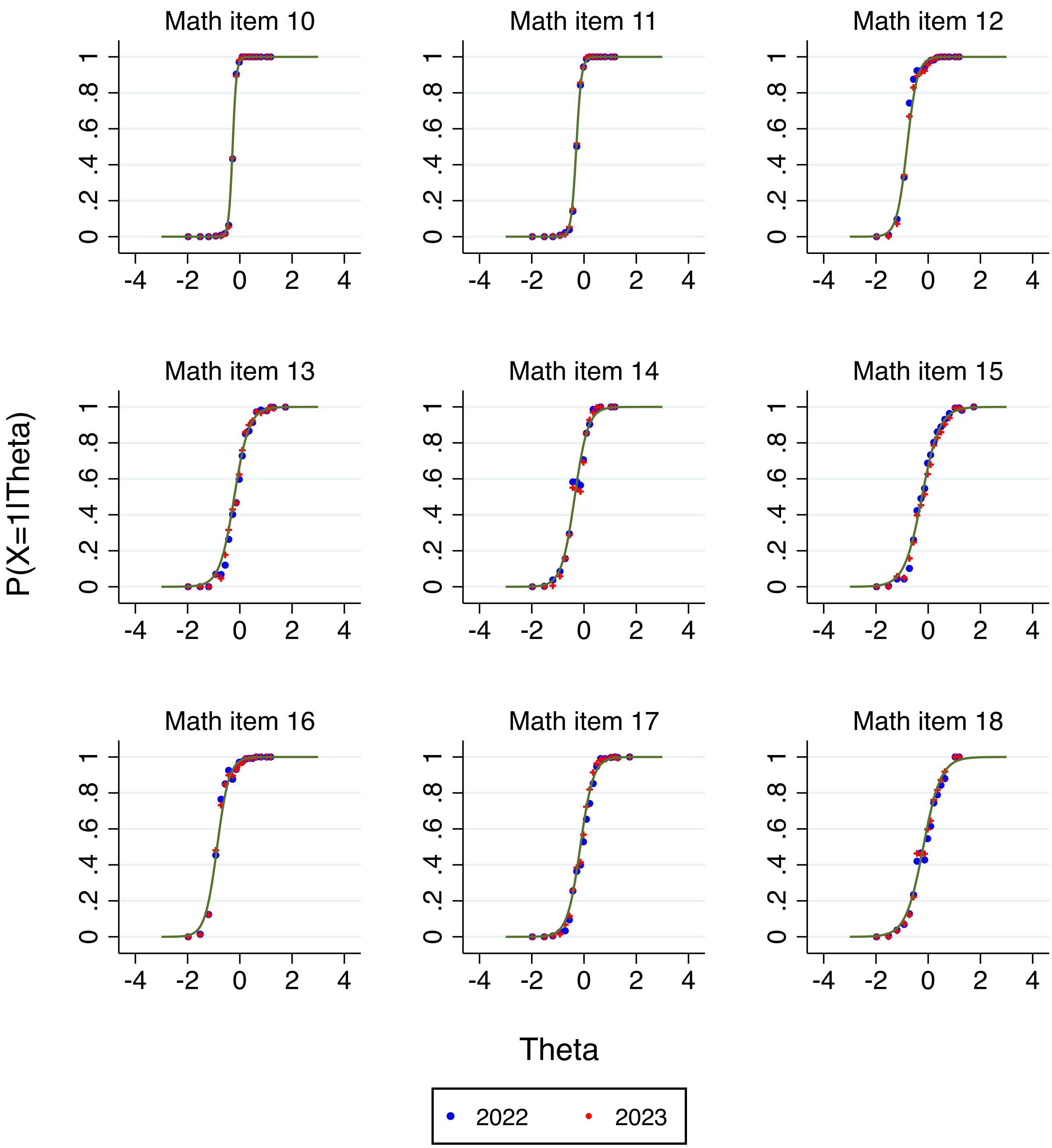
ICC curves and empirical fit by round

Figure G.8: Empirical fit to the estimated item characteristic curve (ICC) for math items by round



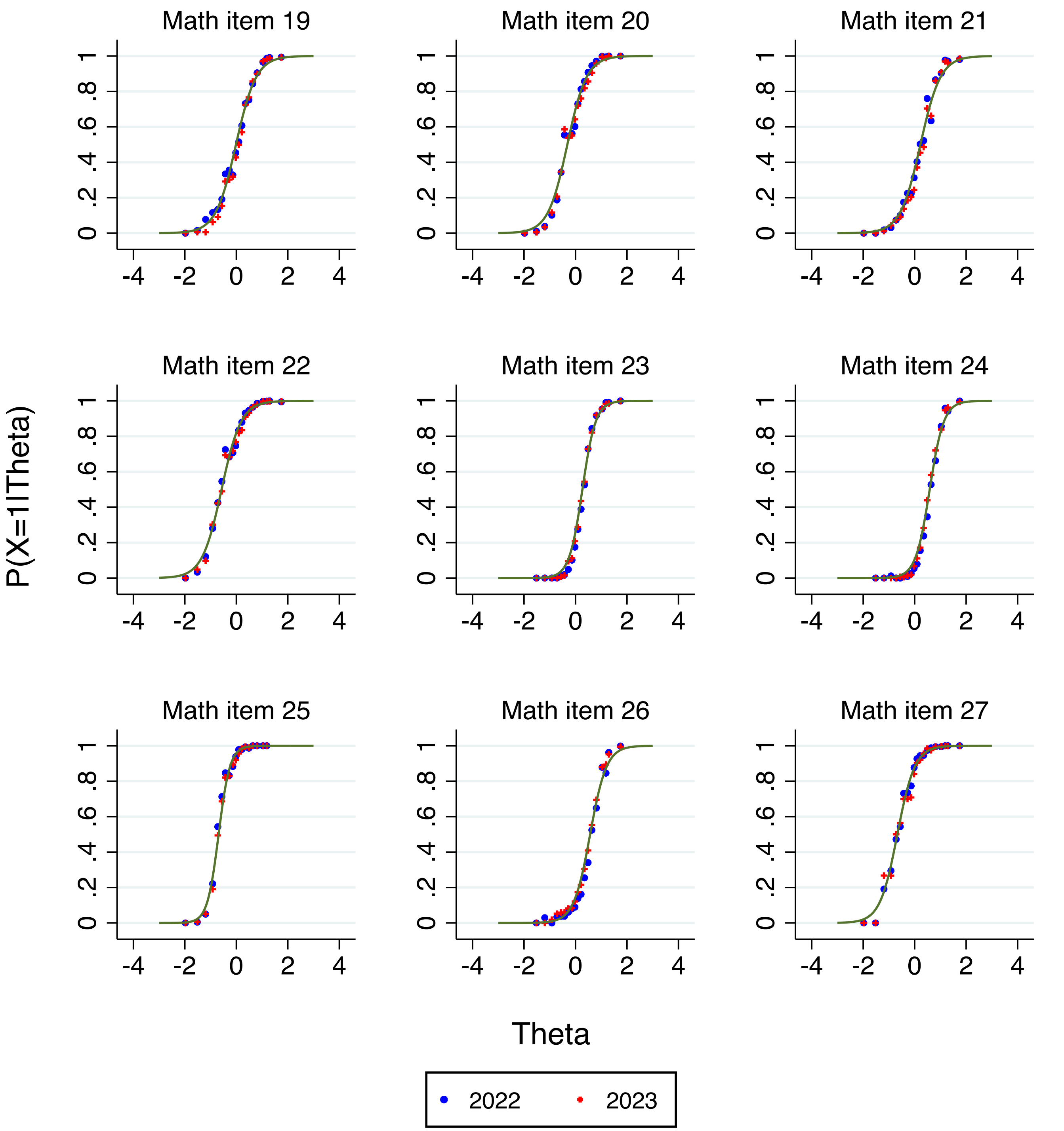
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.9: Empirical fit to the estimated item characteristic curve (ICC) for math items by round



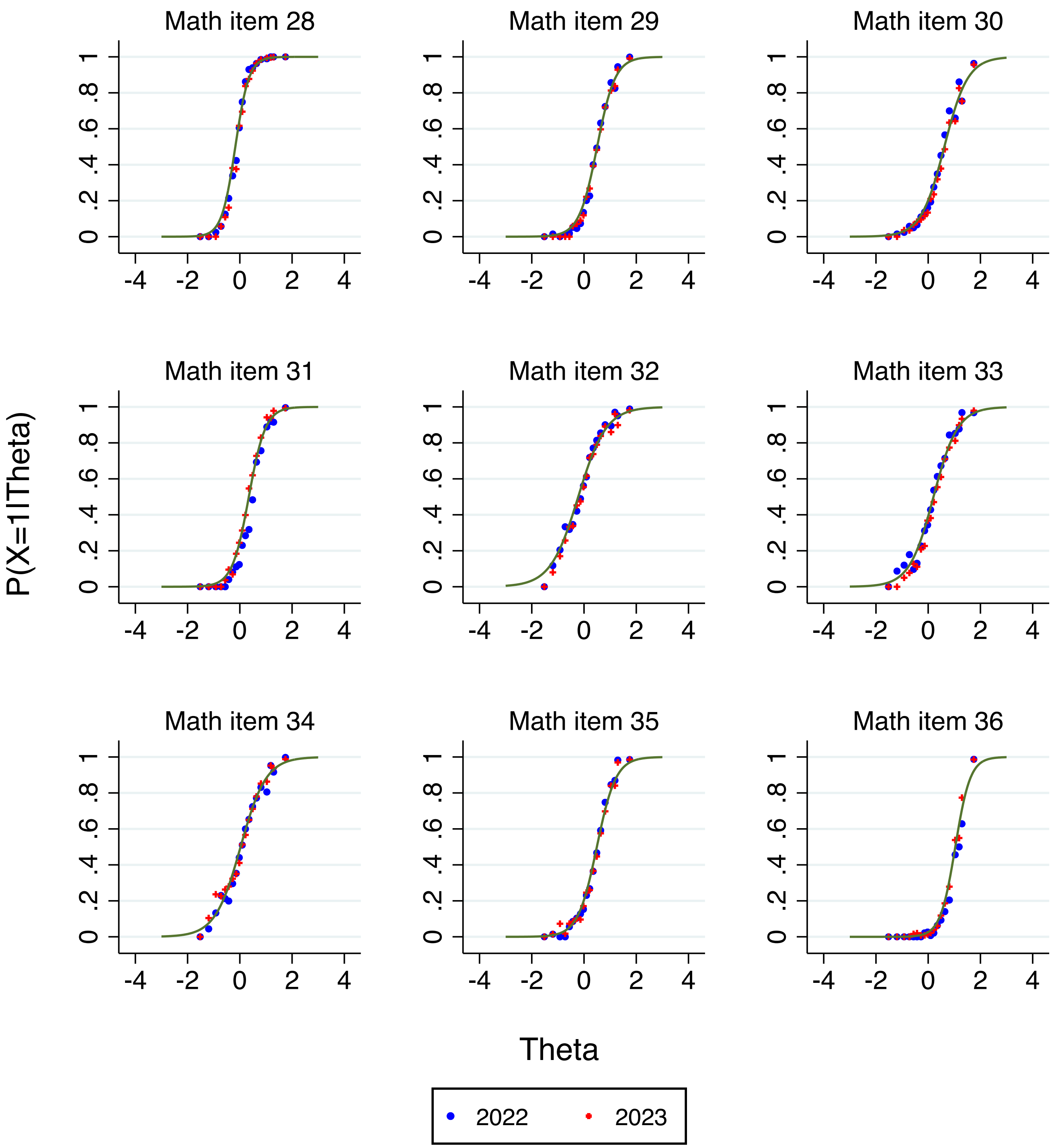
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.10: Empirical fit to the estimated item characteristic curve (ICC) for math items by round



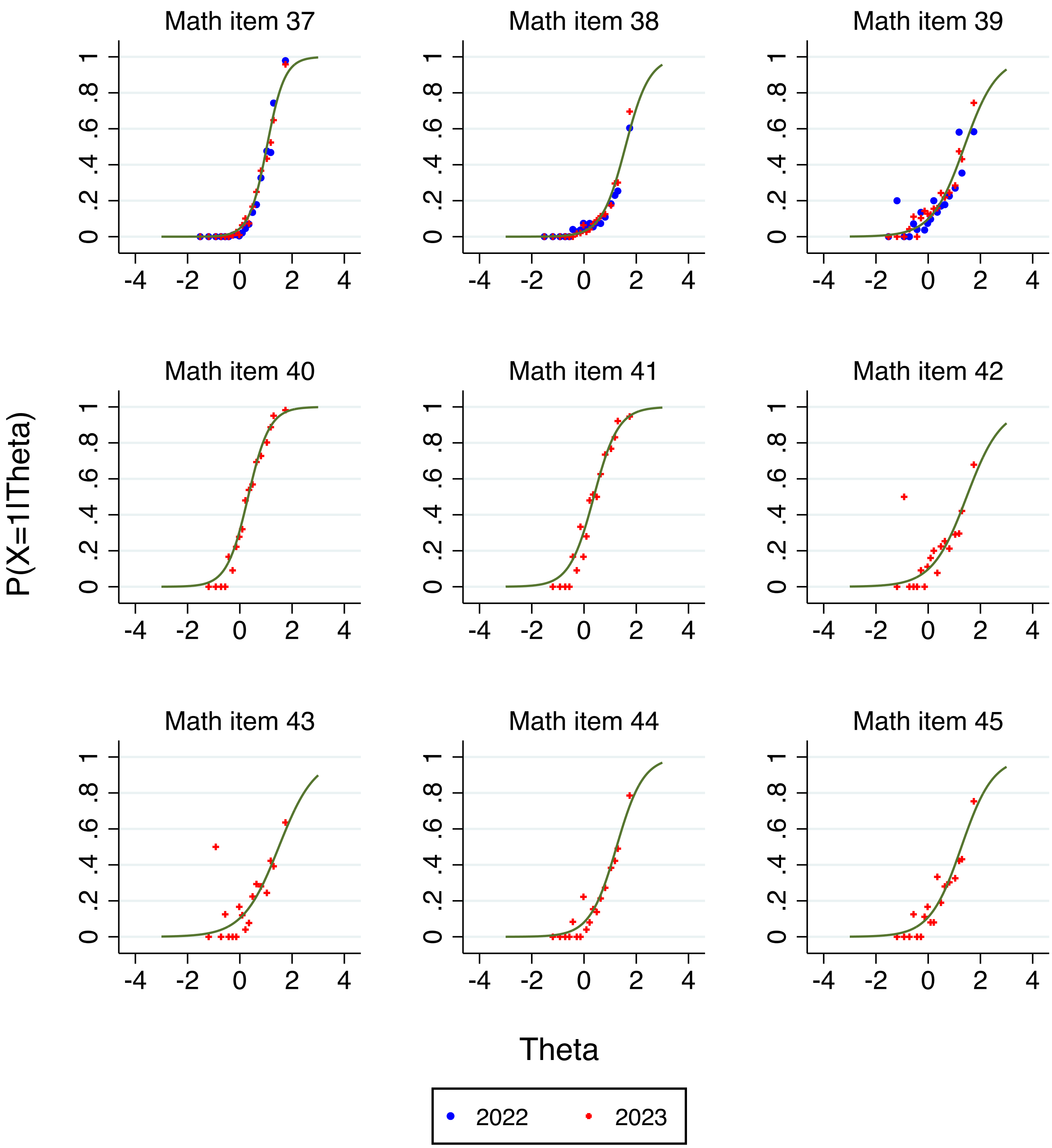
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.11: Empirical fit to the estimated item characteristic curve (ICC) for math items by round



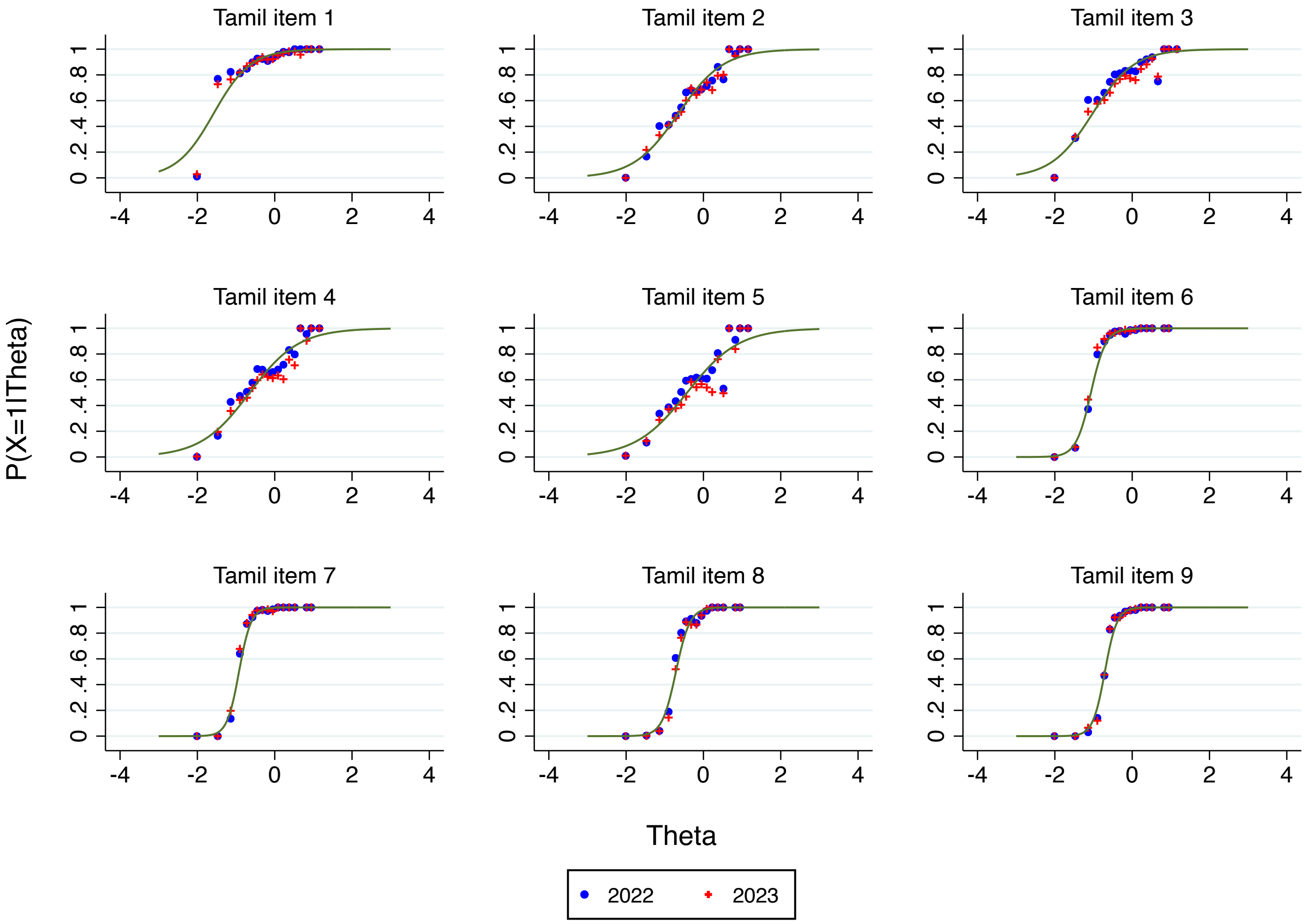
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.12: Empirical fit to the estimated item characteristic curve (ICC) for math items by round



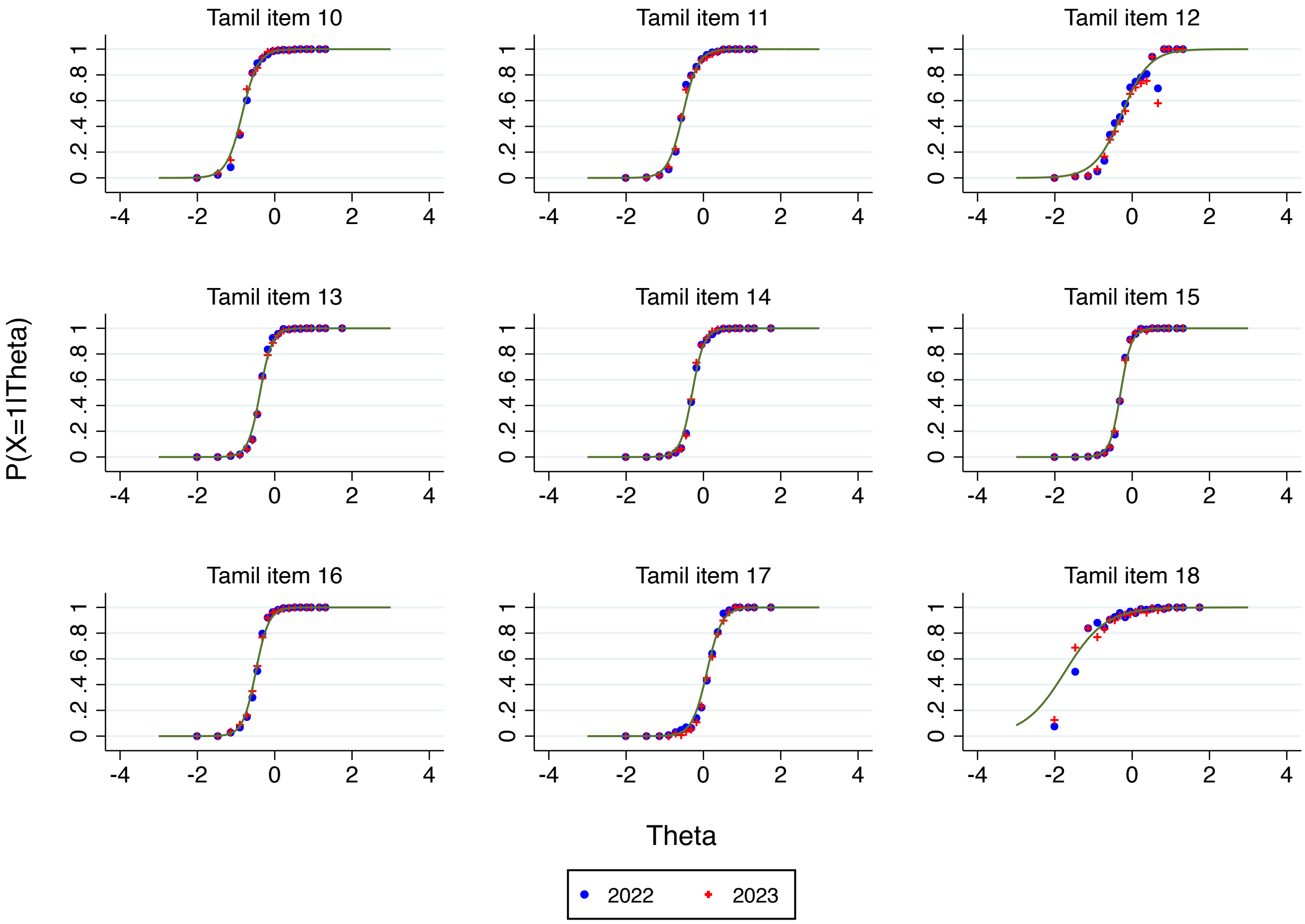
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.13: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



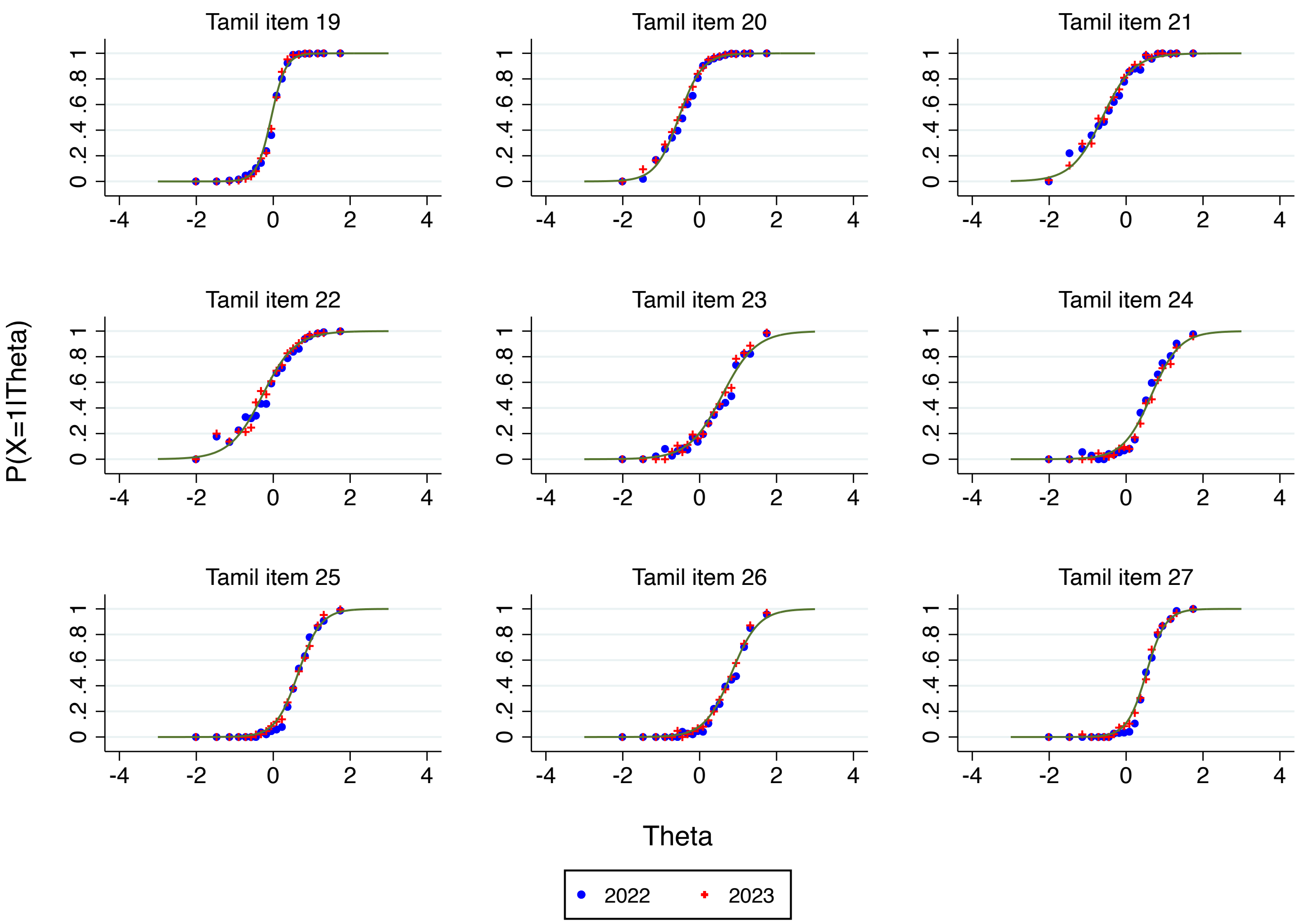
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.14: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



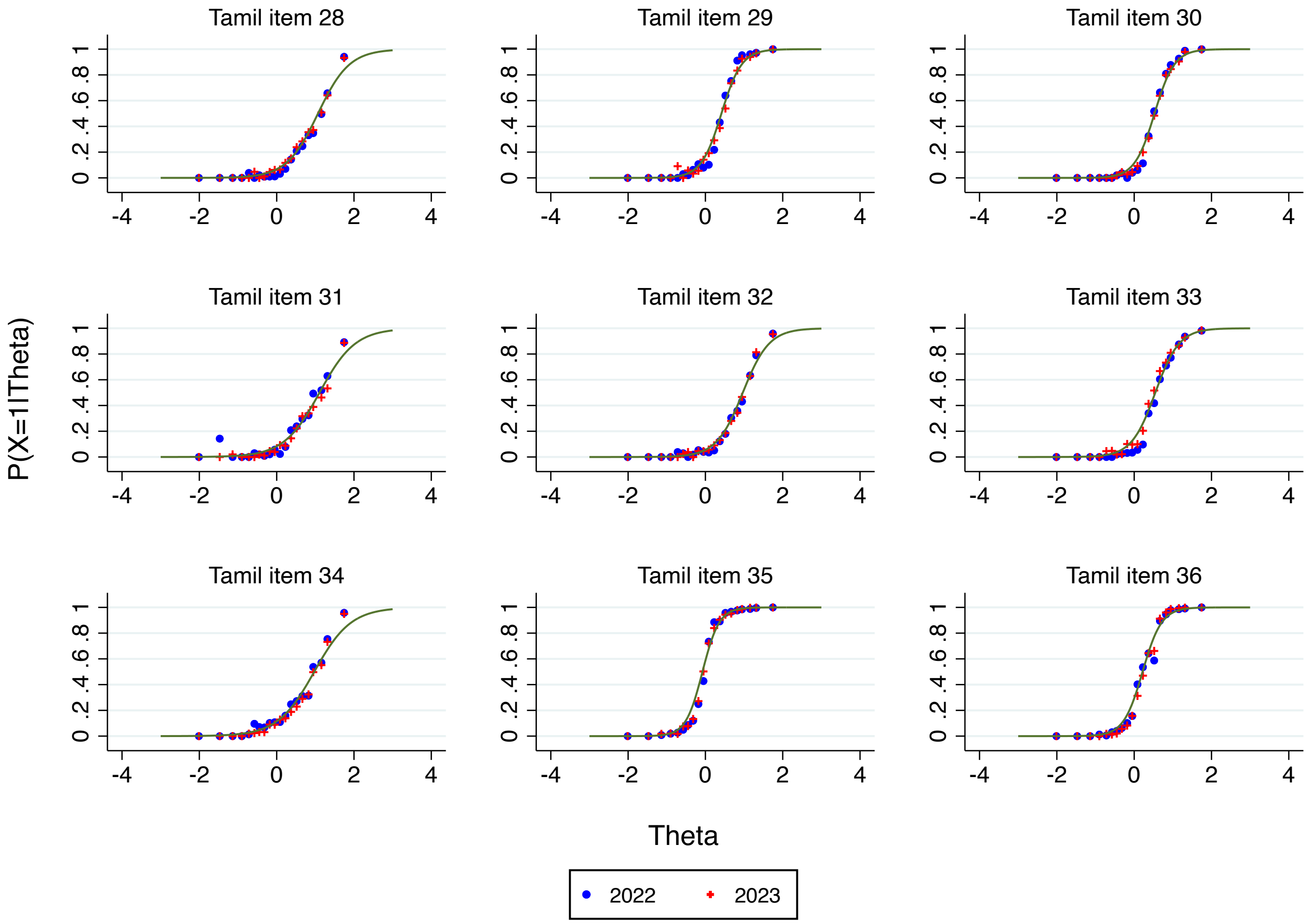
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.15: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



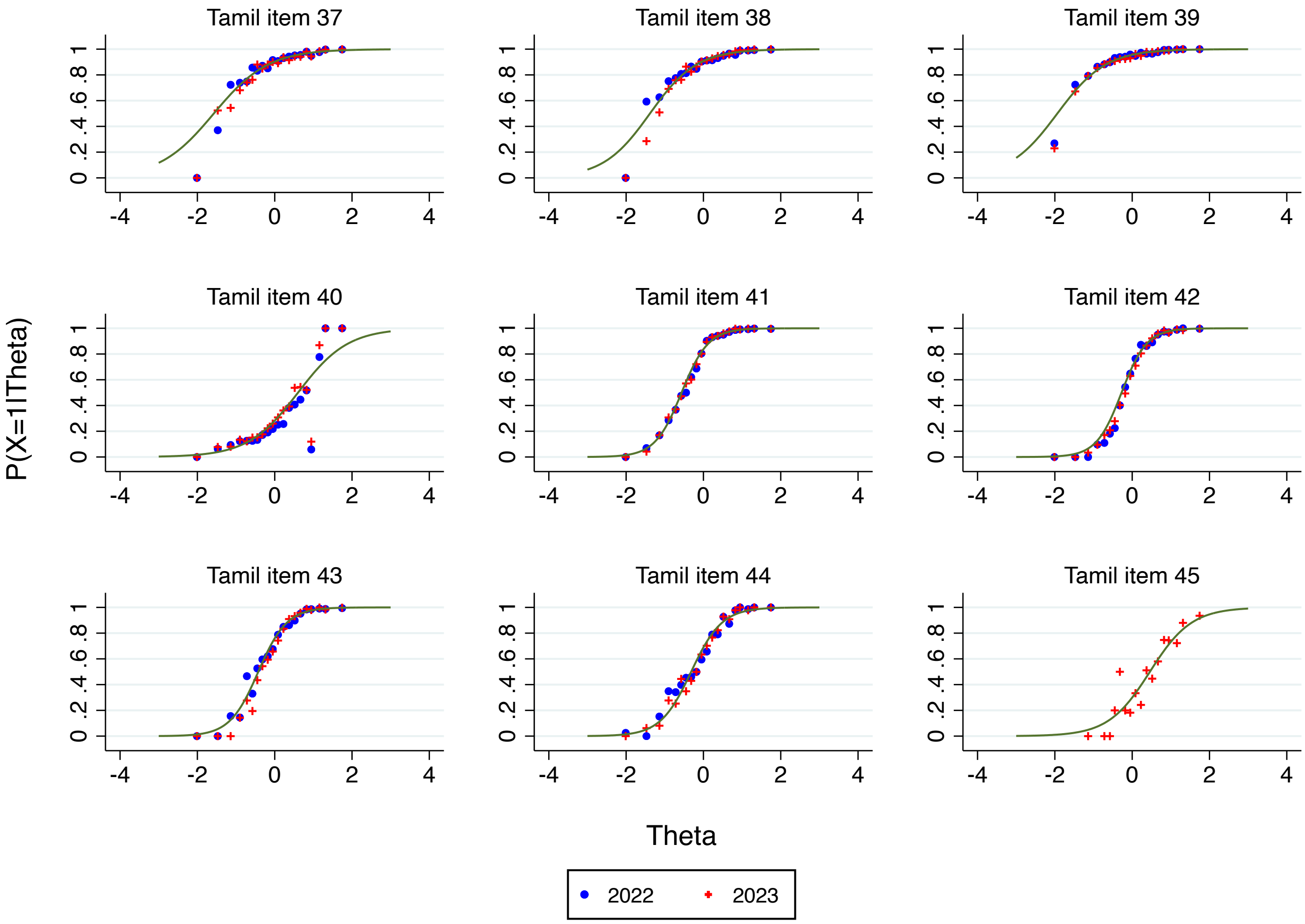
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.16: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



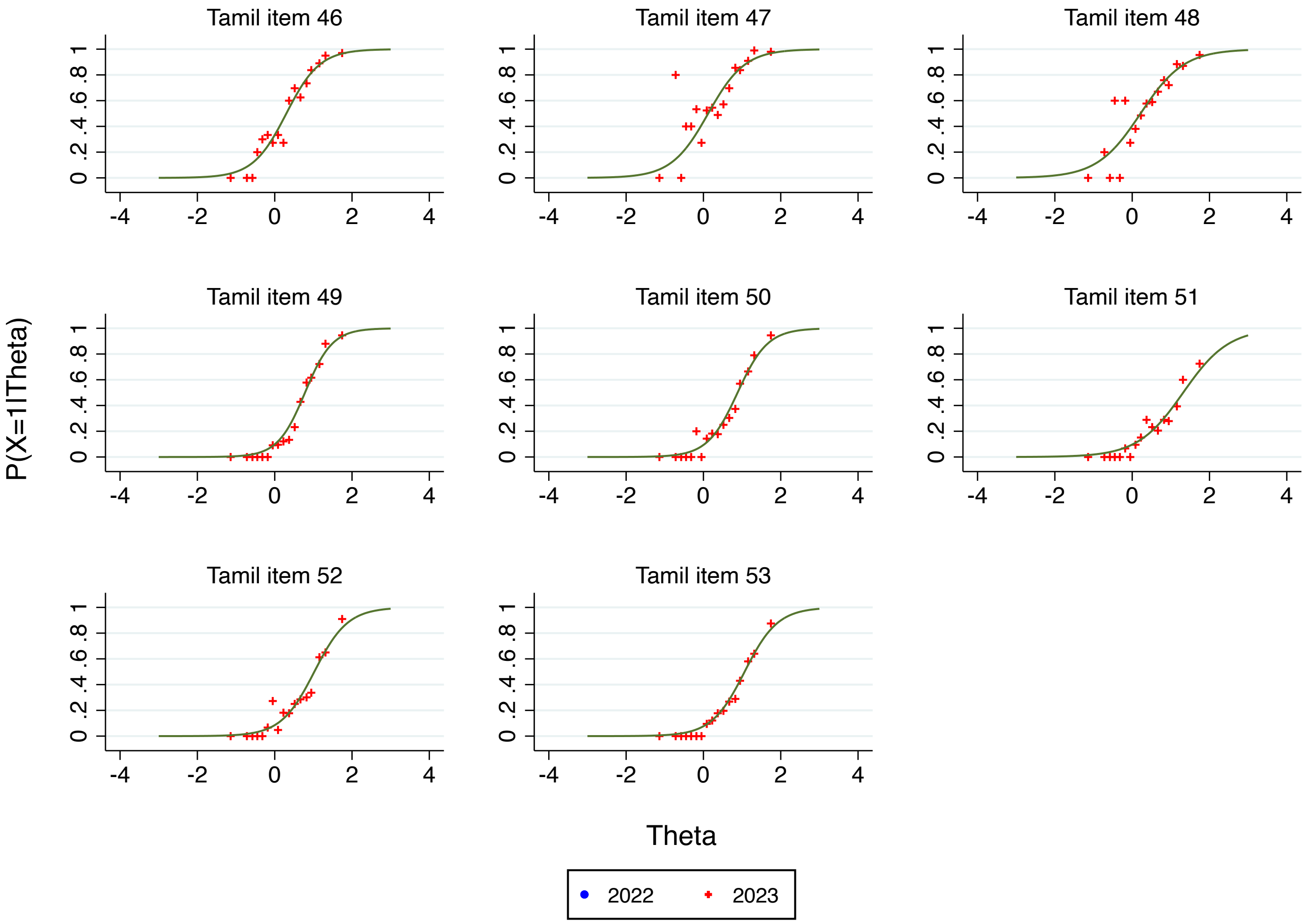
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

Figure G.17: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

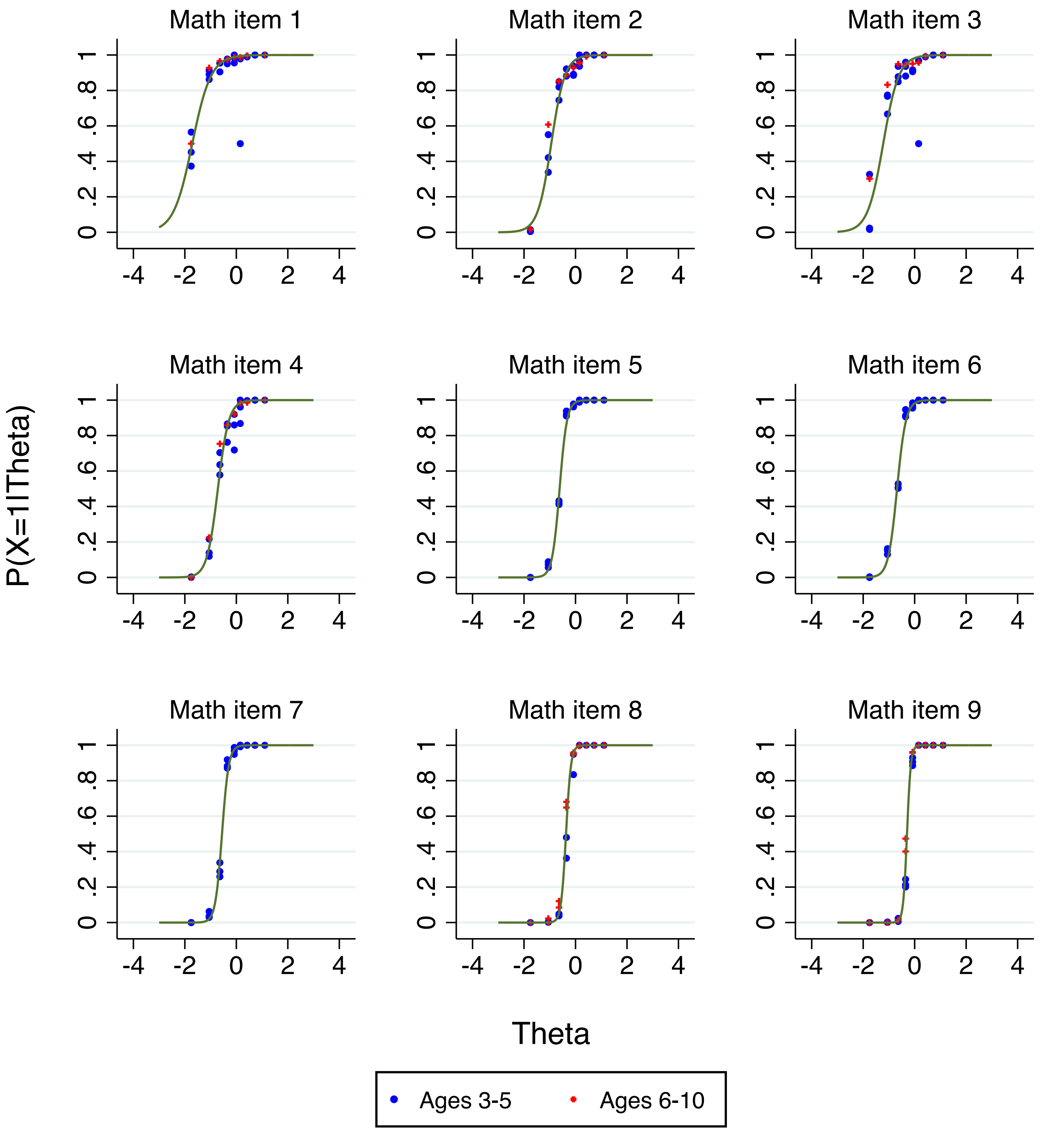
Figure G.18: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by round



Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately by assessment round.

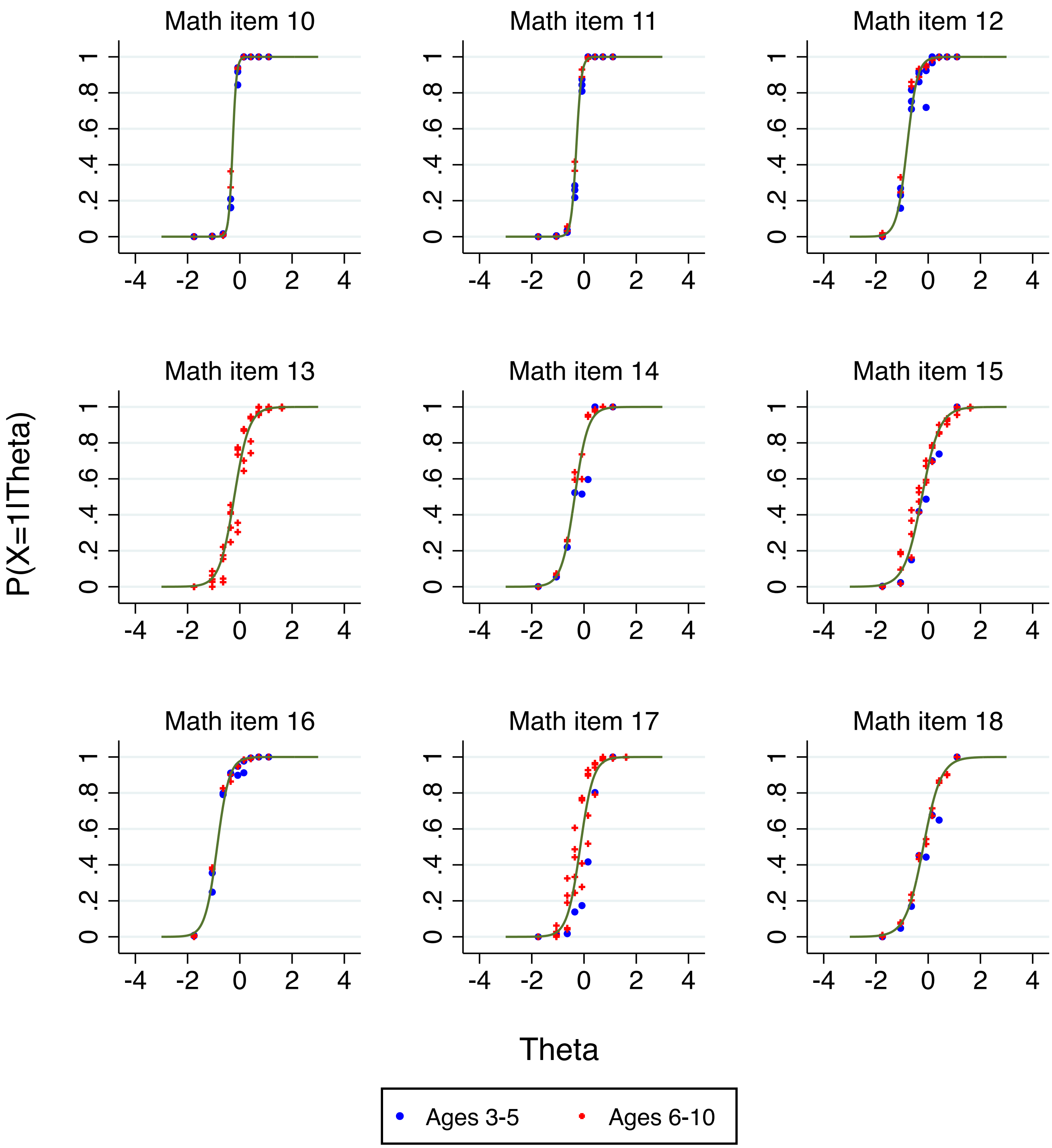
ICC curves and empirical fit by age

Figure G.19: Empirical fit to the estimated item characteristic curve (ICC) for math items by age



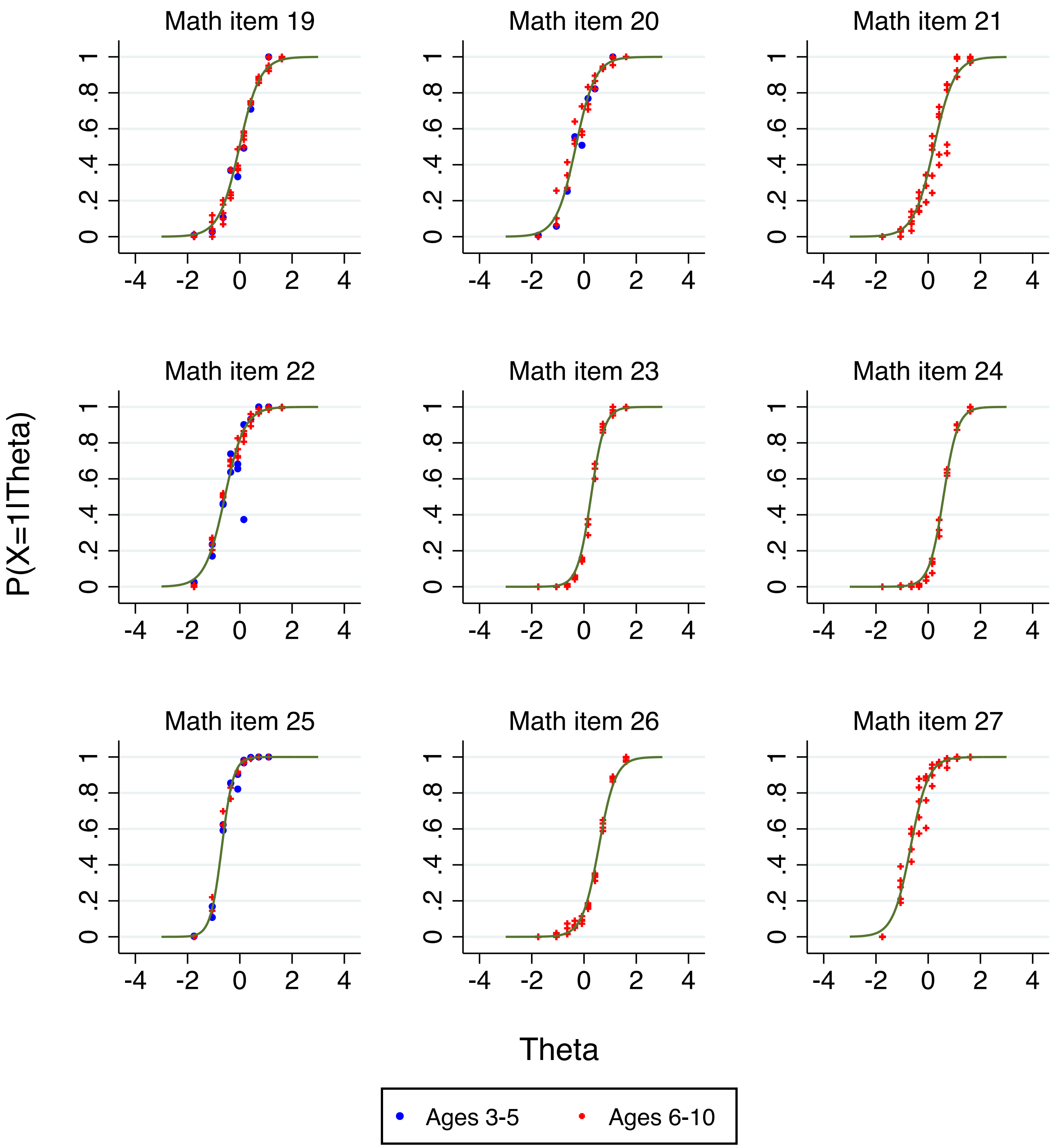
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.20: Empirical fit to the estimated item characteristic curve (ICC) for math items by age



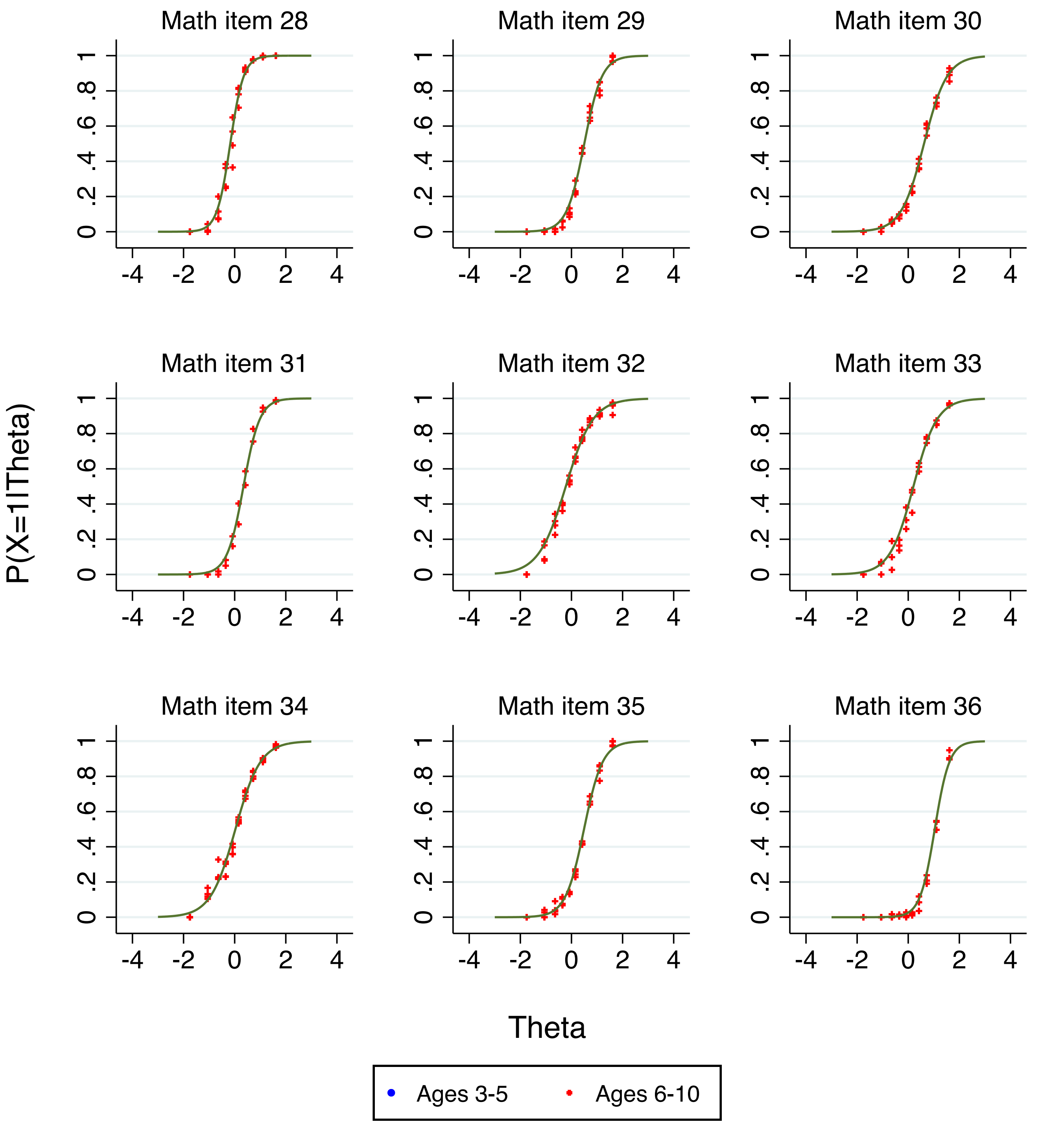
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.21: Empirical fit to the estimated item characteristic curve (ICC) for math items by age



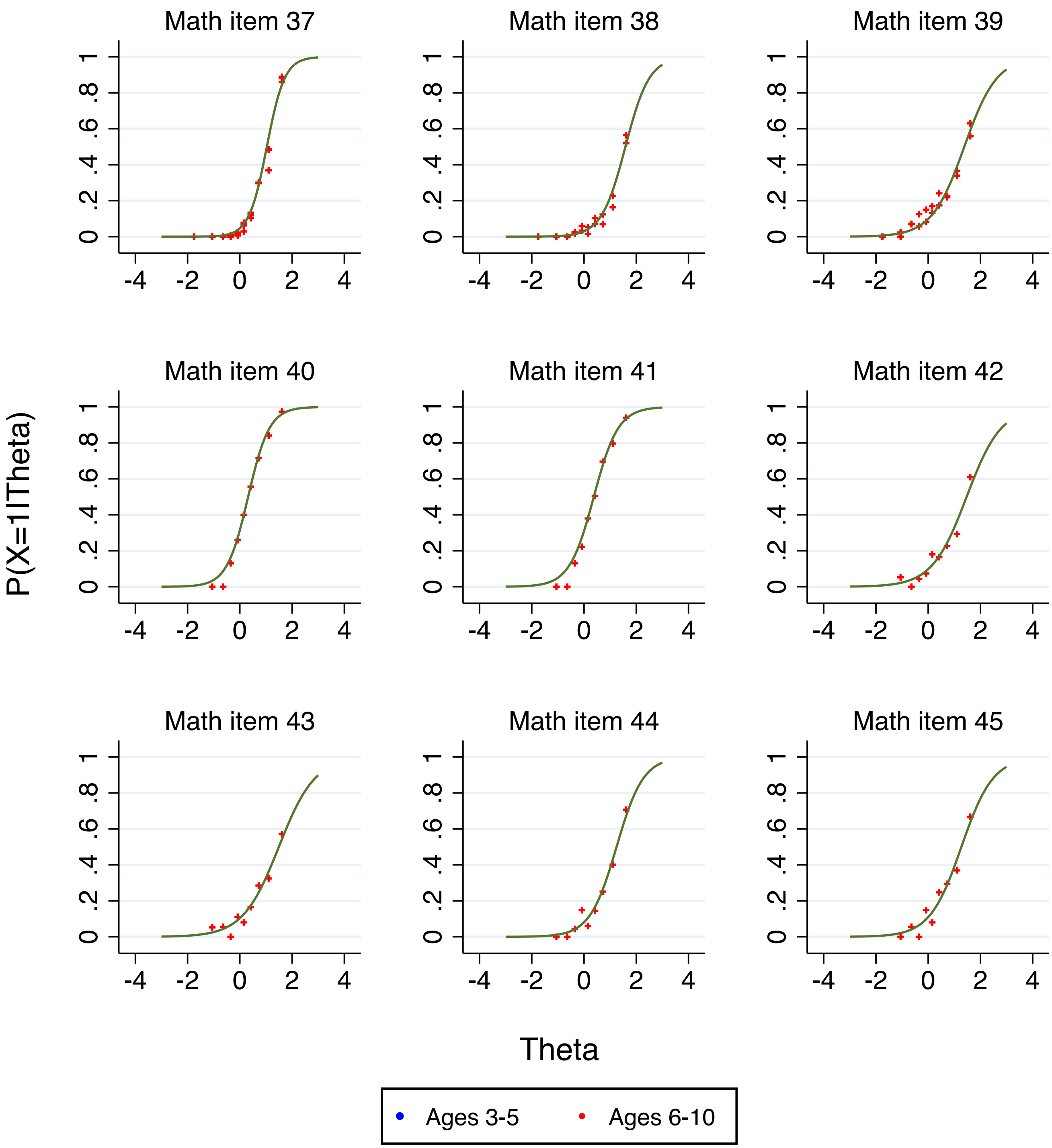
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.22: Empirical fit to the estimated item characteristic curve (ICC) for math items by age



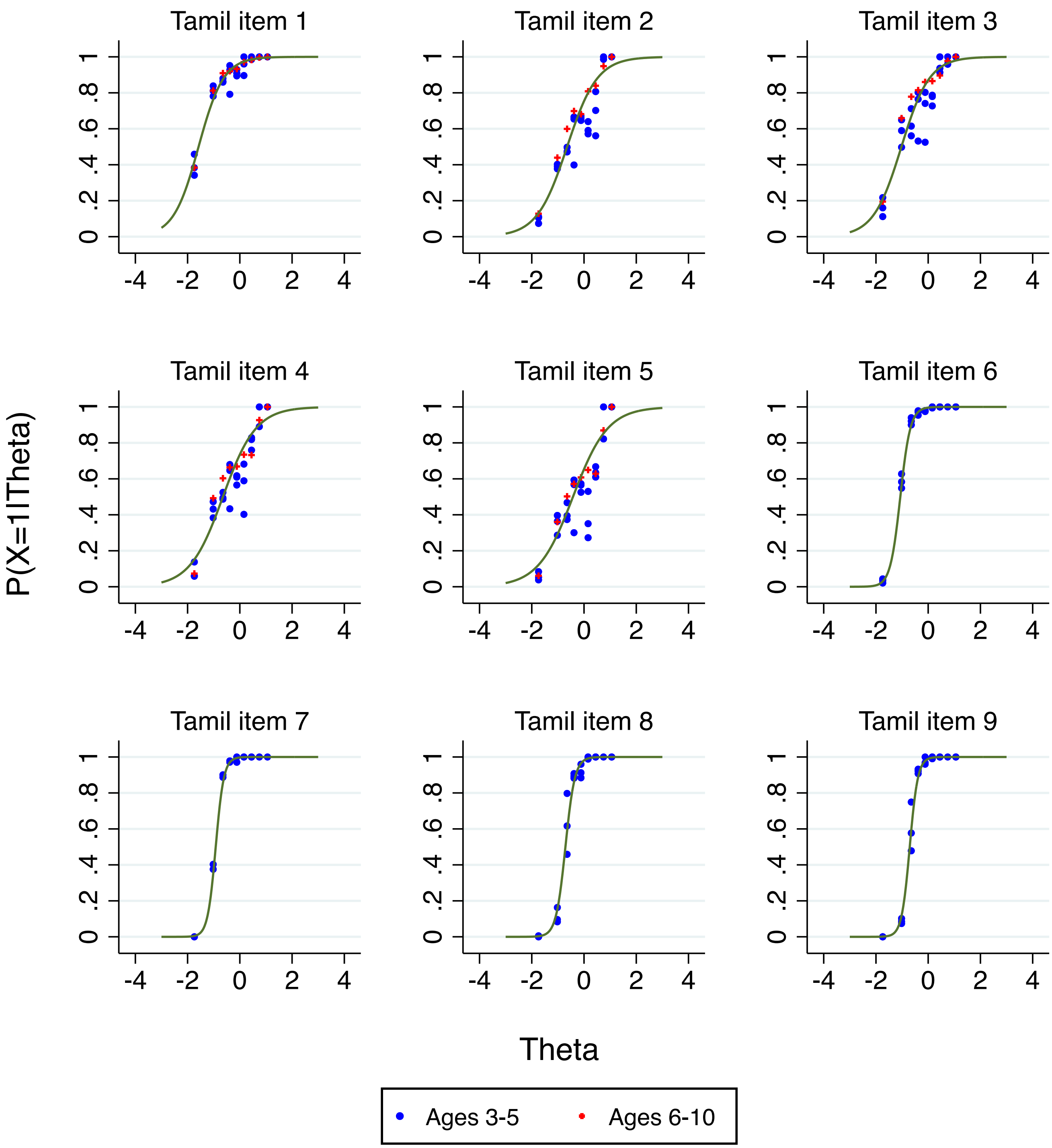
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.23: Empirical fit to the estimated item characteristic curve (ICC) for math items by age



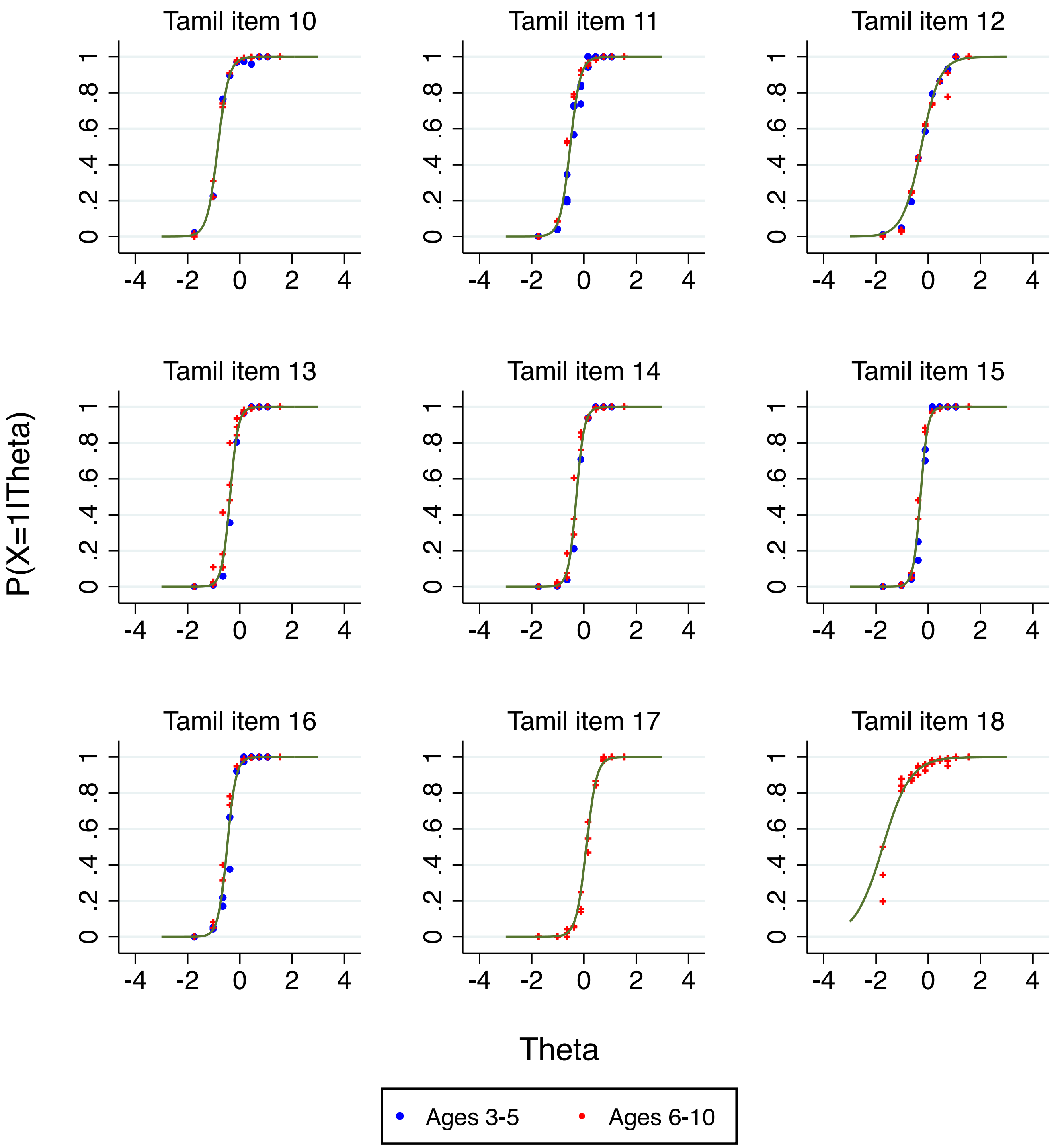
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.24: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



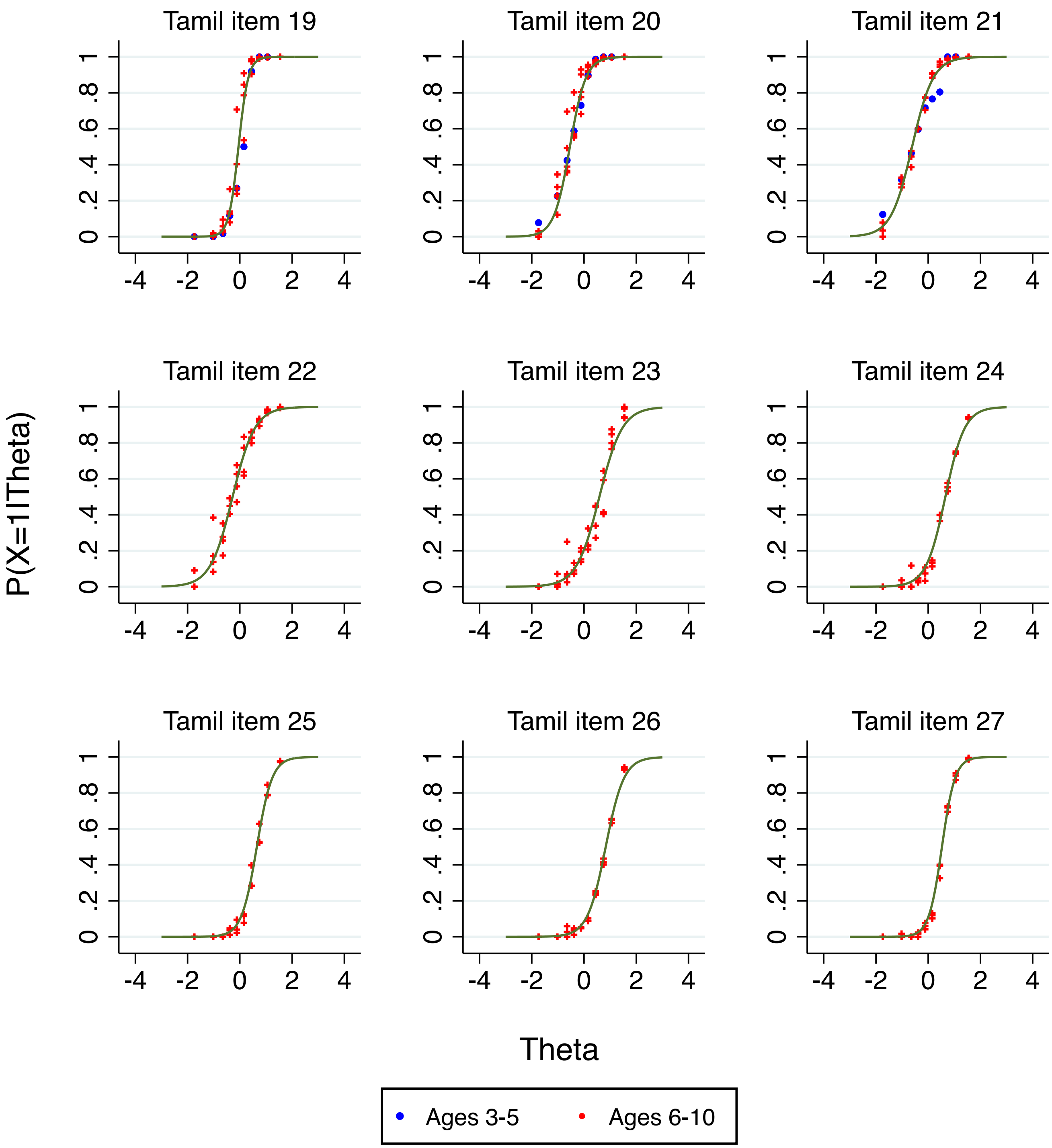
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.25: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



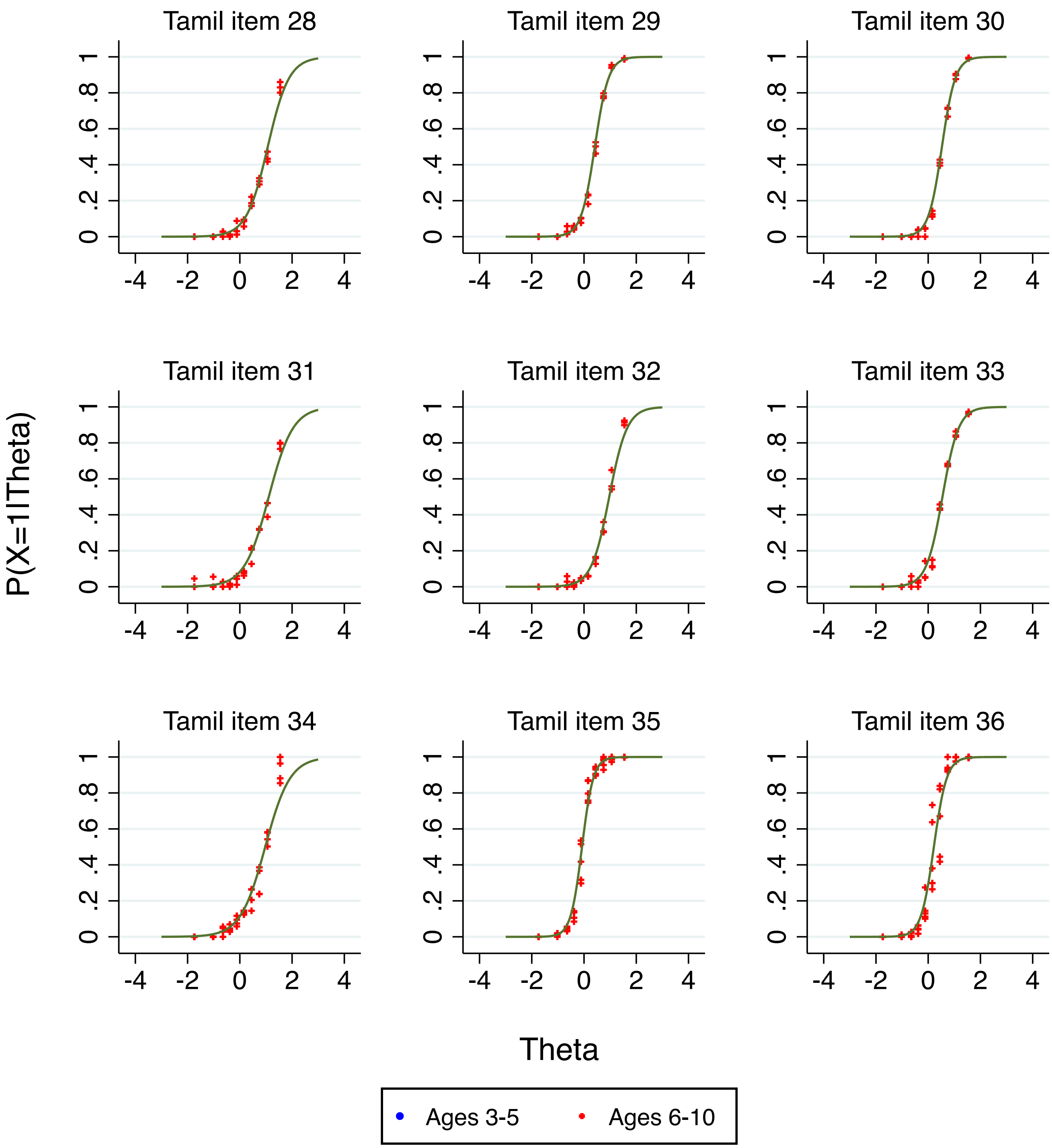
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.26: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



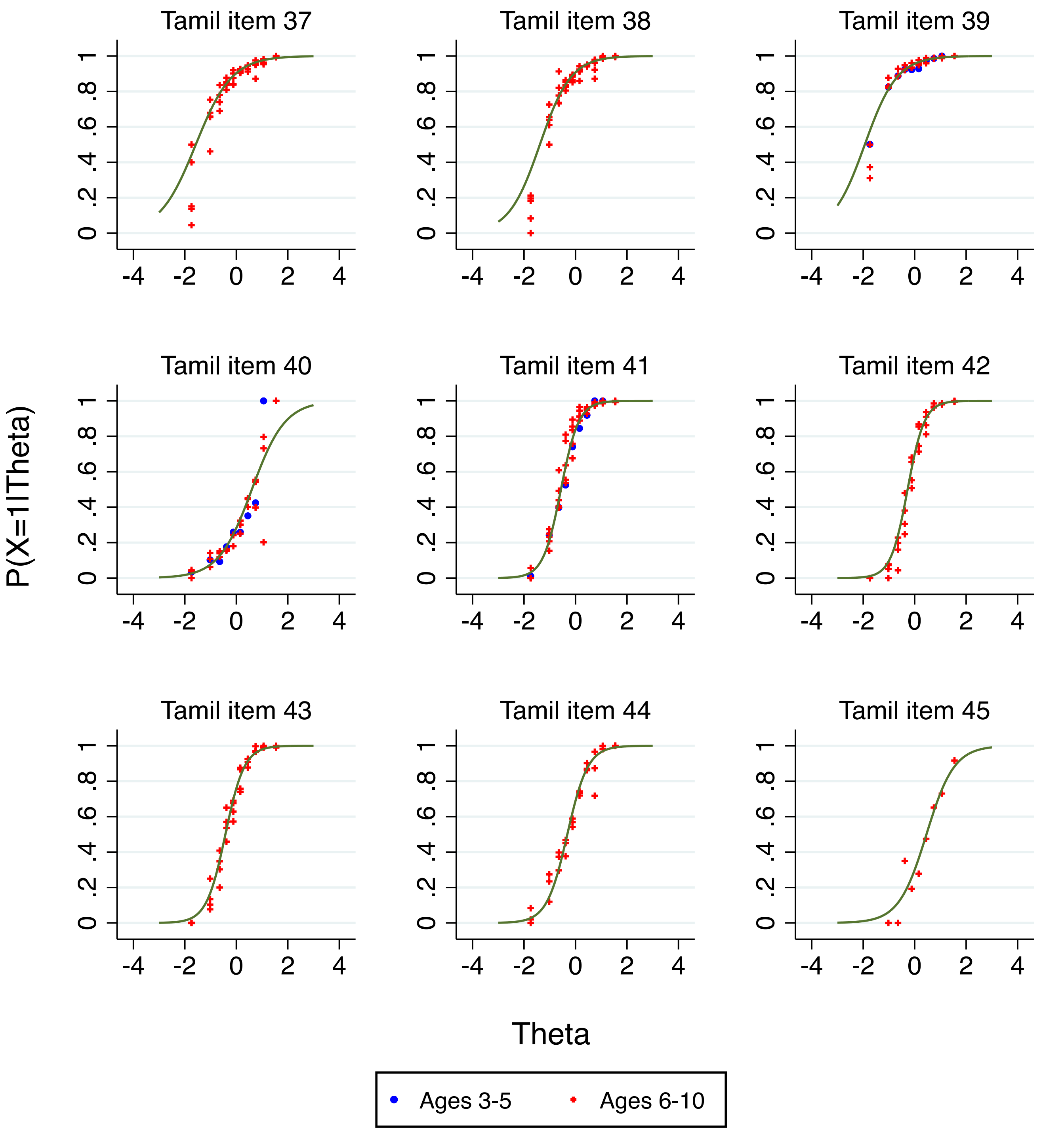
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.27: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



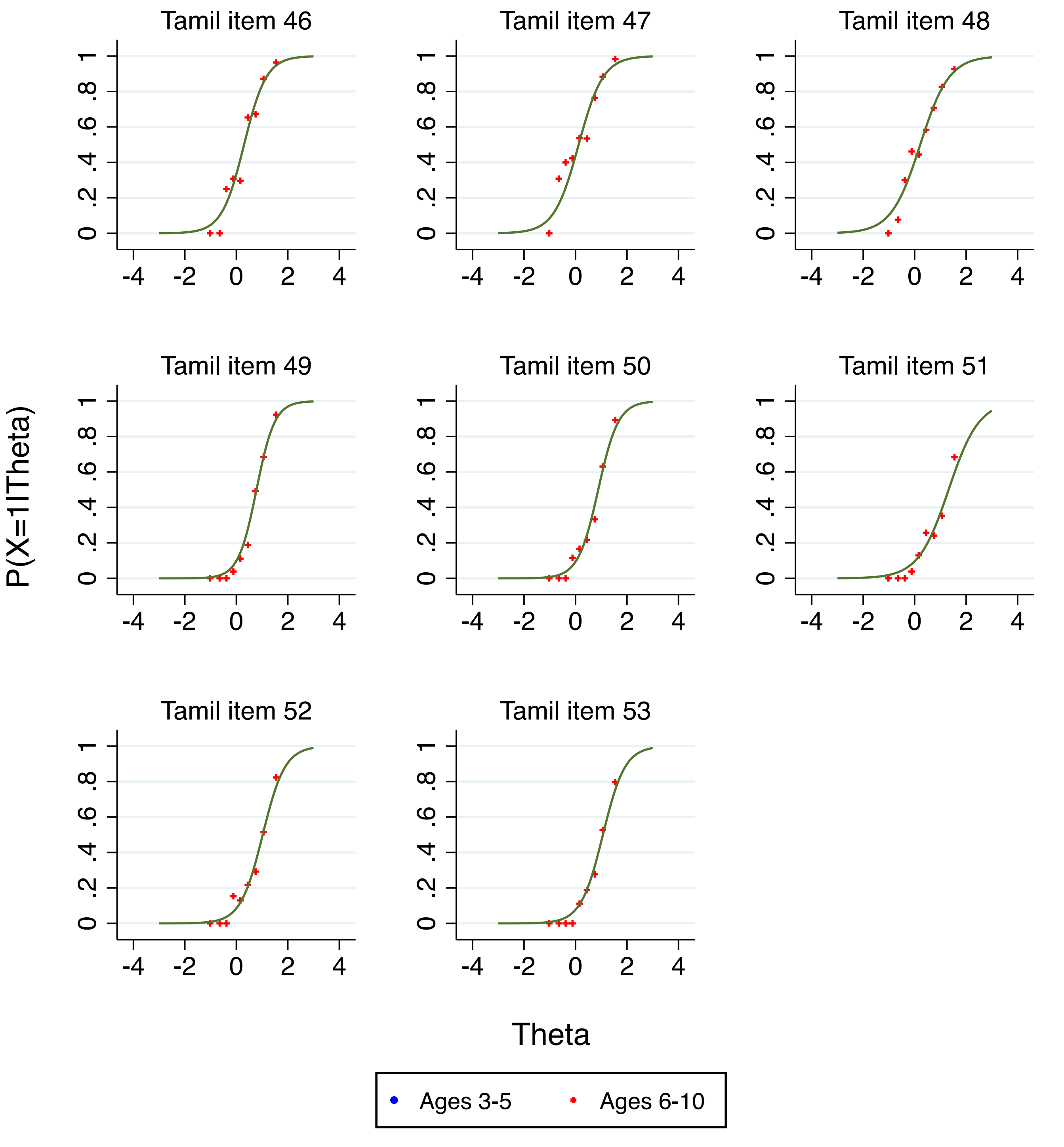
Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.28: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

Figure G.29: Empirical fit to the estimated item characteristic curve (ICC) for Tamil items by age



Notes: This figure presents the likelihood that students with different IRT scores answer different questions correctly, as well as the item characteristic curve for each question, separately for children aged 3-5 and 6-10.

References

Andrabi, T., Bau, N., Das, J., & Khwaja, A. I. (2025, January). Heterogeneity in school value added and the private premium. *American Economic Review*, *115*(1), 147–82. Retrieved from https://www.aeaweb.org/articles?id=10.1257/aer.20221422 doi: 10.1257/aer.20221422

Angrist, J., Hull, P., Legate-Yang, R., Pathak, P. A., & Walters, C. R. (2025, March). *Putting school surveys to the test* (Working Paper No. 33622). National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w33622 doi: 10.3386/w33622

Chen, J. (2024). *Empirical bayes when estimation precision predicts parameters.* Retrieved from https://arxiv.org/abs/2212.14444

Cinelli, C., & Hazlett, C. (2019, 12). Making sense of sensitivity: Extending omitted variable bias. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *82*(1), 39-67. Retrieved from https://doi.org/10.1111/rssb.12348 doi: 10.1111/rssb.12348

Singh, A., Romero, M., & Muralidharan, K. (2024). Covid-19 learning loss and recovery: Panel data evidence from India. *Journal of Human Resources*, *Forthcoming*.

Walters, C. (2024). Chapter 3 - Empirical Bayes methods in labor economics. In C. Dustmann & T. Lemieux (Eds.), (Vol. 5, p. 183-260). Elsevier. Retrieved from https://www.sciencedirect.com/science/article/pii/S1573446324000014 doi: https://doi.org/10.1016/bs.heslab.2024.11.001