Educational inequalities between children of marriage migrants and those of local-born parents - quantile regression results from Taiwan

Tobias Haepp[a] and Mei Hsu[b1]

[a]National School of Development, Peking University, 5 Yiheyuan Road, Haidian District, 100871 Beijing, P.R. China

[b]Department of Business Administration, College of Management, National Taiwan Normal University, 162 He-ping East Road, Section 1, 10610 Taipei, Taiwan (R.O.C.)

ABSTRACT

This paper studies educational inequalities between children of marriage migrants and those of local-born parents using student test scores from Taiwan. We first find an average raw score gap of 9.2% for children of foreign parents. We then employ quantile regression methodology and find that, after introducing our extensive list of covariates, score gaps vanish for children with a mother from mainland China in all quantiles and for children with a Southeast Asian mother in the highest quantiles. In contrast, we identify large residual score gaps for children with a parent from a culturally more distant country in all quantiles. Different from previous studies which found the largest impact for language usage and parent education, we find that the physical home environment has the highest explanatory power in our analysis.

KEYWORDS

Educational inequality; Human capital; Quantile regressions; Second-generation marriage immigrants

1. Introduction

As a response to the ongoing demographic changes in East Asian societies, such as declining fertility rates, ageing populations and skewed sex ratios, cross-border marriages have become an important socioeconomic trend (Kim and Oh 2011). In Taiwan, the number of international marriages has averaged around 6.7% of total weddings during the past _ve years. As a result, the number of foreign-born spouses residing in Taiwan had risen to 510,250 by the end of 2015, with spouses from mainland China making up approximately two thirds and spouses from Southeast Asian countries constituting about 28.1%. Due to cultural factors and a relative dearth of domestic brides, the vast majority of the foreign spouses are females (92.2%). As a consequence, 6.2% of newborns in Taiwan were born to a mother with foreign nationality in 2015.[2]

1.Corresponding author: Mei Hsu, email: mhsu@ntnu.edu.tw. Hsu acknowledges funding from the Ministry of Science and Technology (R.O.C.) under MOST 104-2410-H-003-127.

2 The statistics can be accessed through the National Statistics website compiled by the Directorate General of Accounting, Budget and Statistics (Executive Yuan) at http://eng.stat.gov.tw.

An issue of concern is how the differences in culture and socioeconomic status between local- and foreign-born parents transmit from the parent generation to their children. Research from the US has pointed out that immigrants, in particular those from poorer countries, are at a significant disadvantage in terms of their earnings, and that the catching up process from one generation to the next is slow (Borjas 2006). Neal and Johnson (1996) have demonstrated that much of these later-life earnings differentials can be explained by test score results before labor market entry. Educational performance as a major determinant of later-life outcomes has therefore recently received significant attention in the literature on second-generation immigrants.

The case of Taiwan is of academic interest for studies in migration economics, because the composition of immigrants is such that a large number of them are native speakers of Mandarin-Chinese and already speak the language of class instruction to their children at home, hence allowing us to disentangle the language effect from socioeconomic and other relevant factors pertaining to foreign-born parents. A previous study on second-generation political immigrants in Taiwan found that these outperformed their local counterparts in terms of educational attainment after controlling for other relevant factors (Tsay 2006). More recently, the percentage share of students with a foreign parent in primary and middle schools has risen from a mere 1.63% in 2004 to 10.28% in 2014.[2] An analysis of the performance of these students relative to their native peers is therefore of importance for educational research, as well as for social policy in Taiwan and more broadly as part of a major trend in East Asian societies.

The specific dataset we employ in the present study is derived from the Taiwan Assessment of Student Achievement (TASA) database. The TASA data are a long-term project implemented by the National Academy for Educational Research (NAER) that focuses on the educational performance of students in primary school, middle school and high school in the subjects Chinese, English, mathematics, natural studies and social studies. The database contains data on student scores and also information from student questionnaires asking about factors such as birthplace and educational attainment of parents, the interaction of students with parents, classmates and teachers, the physical environment at home, languages used at home and questions regarding student preferences, such as towards study subjects. The current study therefore also makes a contribution by jointly analyzing a range of factors that determine student performance and weighing their relative impact on test scores through the lens of the native-immigrant performance gap.

The rest of this paper is structured as follows. In section 2, we review the literature on inequalities between second-generation immigrants and natives, with a particular focus on education. In section 3, we introduce our dataset in more detail and present descriptive statistics. Section 4 then introduces our quantile regression methodology and presents our research findings. In section 5 we discuss our findings and provide policy recommendations.

2. Literature review

While a large number of early studies on native-immigrant gaps has focused on labour market outcomes, more recent research has turned to the question how immigrants

[2]The data have been retrieved from the website of the Department of Statistics at the Ministry of Education under https://stats.moe.gov.tw/.

differ in the course of their acquisition of labor market relevant skills through schooling. Schnepf (2007) employs data from the Trends in International Maths and Science Study (TIMSS), the Programme of International Student Assessment (PISA) and the Programme of International Reading Literacy Study (PIRLS) and finds that immigrants perform little worse than natives in English speaking countries such as Australia, Canada, New Zealand the UK. In continental European countries, however, they fare significantly worse than natives. In addition to language ability, the author points to socioeconomic background and school segregation as factors behind these findings. Dustmann, Frattini, and Lanzara (2012) employ PISA data and confirm that test score gaps differ widely across OECD countries, and that parental background characteristics are an important determinant of them. They find that languages spoken at home are the single most important factor behind the score gaps. They add that higher school and peer quality in the destination countries raise the performance of immigrants relative to a comparison group in their country of origin for the case of their subsample of Turkish immigrants. In line with these cross-country study results, studies focusing on individual countries also confirm that language ability or languages spoken at home as well as parent characteristics are the key factors behind native-immigrant educational gaps.

Regarding the languages spoken at home, Dustmann, Machin, and Schönberg (2010) find that differences in mother tongues can explain between 37% and 64% of the educational achievement gap between white British and ethnic minority children in the UK and are the single most important factor behind these gaps. Bleakley and Chin (2008) demonstrate that parents' ability to speak the language of instruction significantly impacts children's language proficiency, performance relative to their age appropriate grade, and school dropout rates in their US census dataset.

Focusing on the effects of parental characteristics on interethnic score differentials in the US, Fryer and Levitt (2004) find that the black-white test score gap among kindergardeners disappears after controlling for factors such as parental education, parental occupational status, household income and the number of books at home. In contrast to their findings in their earlier study, the authors find that a gap emerges that cannot be explained by the same covariates during the subsequent four years in a later study (Fryer and Levitt 2006). Baert and Cockx (2013) analyze the differences in family endowments between natives and immigrant grandchildren in Belgium and decompose the gap into differences in these endowments and a residual "pure ethnic gap". They find that the overall gaps can be largely explained by family endowments until tertiary education, but pure ethnic gaps grow as students progress through the education system and are large when considering school-to-work-transitions. Using immigrant data measuring upper secondary school performance in Denmark, Colding, Husted, and Hummelgaard (2009) find that family characteristics only explain part of the gap and therefore points out that other constraints exists, in particular behavioral ones. Turning attention to these processes, Villiger, Wandeler, and Niggli (2014) is amongst the few studies that analyze the channels through which parental endowments are transmitted to the next generation. They find that immigrant students receive less emotional support from their parents which in turn negatively affects reading motivation and achievement in their dataset from Switzerland.

To summarize the literature on the determinants of native-immigrant educational performance gaps, various measures of family characteristics and languages spoken at home have been instrumental in explaining varying degrees of these gaps depending on the particular country. Our paper thus contributes to the literature on native-immigrant gaps by analyzing the factors from previous studies in the context of mar-

Table 1. Sample composition by parent origin

Class	Taiwan-Taiwan	Taiwan-China	Taiwan-SE Asia	Other country	Rest	Total
Primary-4	26034	511	684	1137	4734	33100
	78.65	1.54	2.07	3.44	14.30	100.00
Primary-6	73929	1025	1381	984	5200	82519
	89.59	1.24	1.67	1.19	6.30	100.00
Junior-2	74770	744	914	379	3058	79865
	93.62	0.93	1.14	0.47	3.83	100.00
Total	174733	2280	2979	2500	12992	195484
	89.38	1.17	1.52	1.28	6.65	100.00

The first row of each group shows absolute numbers, while the second row displays percentage shares of each group in the total student body of the respective grade or the total sample.

riage migration in Asia and by introducing several additional channels from research in the economics of education.

3. Data and descriptive statistics

Our dataset is derived from the Taiwan Assessment of Student Achievement Database (TASA) which is a long-term project implemented by the National Academy for Educational Research (NAER). The long-term assessment project implements exams and background questionnaires for pupils at various stages during their pre-university education. The years covered are primary school year 4 (primary-4), primary school year 6 (primary-6), junior high school year 2 (junior-2), and the second year of senior high school (senior-2) and vocational schools (vocational-2).[3] A drawback of the TASA database is that the questionnaires differ slightly over time and each age group has only been investigated in a selective number of years, rendering it impossible to trace a specific cohort over time. In the face of these constraints, we select primary-4 from 2006, primary-6 from 2007 and 2009, and junior-2 from 2007 and 2010 as the sample for our analysis.[4] Our final sample consists of data from 195,484 student exams in the subjects Chinese, English, mathematics, natural studies and social studies.[5]

The focal point of our analysis is the country of origin of each student's parents, which has been investigated as part of a questionnaire handed out to each exam taker. We select four parent combinations that are both meaningful from a policy perspective and also sizeable in terms of their sample shares. The first group are children with parents that were both born in Taiwan (henceforth: the TT group). The second and third group reflect the trends discussed in the introduction and are

[3]The Taiwanese education system begins with six years of primary school education, followed by three years of junior high school education and three years of senior high school education. The students generally enter primary school at age 7. The four age groups are therefore 10, 12, 14 and 17.

[4]In the senior high school sample, only 2.4% of the pupils have a foreign parent. We therefore exclude the senior-2 group due to insufficient sample size.

[5]Social studies has not been tested for the primary-4 subsample.

Table 2. Average scores and score gaps by parent origin

Taiwan-Taiwan	Taiwan-China	Taiwan-SE Asia	Other country	Foreign (avg.)
0.634	0.633	0.564	0.537	0.575
(0.208)	(0.199)	(0.203)	(0.211)	(0.208)
Gap (%)	0.114	11.071***	15.272***	9.205***
T-statistic	0.164	18.295	23.136	24.213

Standard deviations are displayed in parentheses. The T-statistic is a test on the equality of group means. The *** symbol denotes significance at the 1% level of this test result. The foreign group includes all pupils from the TC, TS and OT groups.

made up of children whose father was born in Taiwan with a spouse from mainland China (TC) and Southeast Asia (TS), respectively. To facilitate comparative evidence for other origins, we select our fourth group as children with at least one parent from another origin (OT). The latter group can be generally regarded as the most culturally distant group, which is likely to bring about additional inequalities that cannot be fully captured by the covariates employed in our analysis. Net of the observations with unknown country of parent origin, these four groups make up about 98.4% of our observations. Our final group consists of the remainder of children and has been included merely for completeness, but will not be discussed further, because the exact composition of parent origins in this group is uncertain.

The composition of our sample is summarized in Table 1. Reflecting the demographic trends elaborated in the introduction, 7.05% of our primary-4 students belong to the three second-generation immigrant groups selected, while the respective shares are lower in the student populations of primary-6 (4.10%) and junior-2 (2.54%). Overall, our sample contains 7759 pupils from the three second-generation marriage immigrant groups, hence constituting 4.97% of our whole data set.

A second key input of our analysis is our dependent variable, which consists of student grades in the five exam subjects. Because the number of valid exam questions differs across exams, we calculate the student scores as the percentage of correct answers in each exam. When linking this variable to our four origin groups, significant score gaps can be observed, and the highest test score gap is observed for children with mothers from Southeast Asia relative to children with both parents from Taiwan (see Table 2). The raw scores for students with mother from mainland China are insignificantly different from the TT group at about 0.1%. In contrast, the score gaps amounts to 11.1% for students with a Southeast Asian mother and 15.3% for students with a parent from neither of the other countries, and both of these score gaps are significant. The average foreign student score is significantly lower than the local student score and the gap amounts to 9.2%.

Table 3. Descriptive statistics by parent origin (parent variables)

	Taiwan-Taiwan	Taiwan-China	Taiwan-SE Asia	Other country
Gender (male=1)	0.512	0.498	0.519	0.513
	(0.500)	(0.500)	(0.500)	(0.500)
Siblings (#)	2.652	2.360	2.663	3.079
	(1.180)	(1.169)	(1.359)	(1.789)
Parents together	0.819	0.717	0.767	0.707
	(0.385)	(0.451)	(0.423)	(0.455)
Father high school	0.243	0.216	0.146	0.138
	(0.429)	(0.411)	(0.354)	(0.345)
Father university	0.246	0.198	0.095	0.220
	(0.431)	(0.399)	(0.293)	(0.414)
Mother high school	0.288	0.202	0.133	0.164
	(0.453)	(0.401)	(0.340)	(0.371)
Mother university	0.215	0.212	0.122	0.220
	(0.411)	(0.409)	(0.328)	(0.414)
Parents care about kid	0.729	0.701	0.665	0.670
	(0.445)	(0.458)	(0.472)	(0.470)
Talk about school	0.513	0.450	0.371	0.418
	(0.500)	(0.498)	(0.483)	(0.493)
Help with homework	0.591	0.534	0.490	0.499
	(0.492)	(0.499)	(0.500)	(0.500)
Parents supervise homework	0.226	0.201	0.140	0.251
	(0.419)	(0.401)	(0.347)	(0.434)
Bookstore/library	0.185	0.164	0.121	0.186
	(0.389)	(0.370)	(0.326)	(0.390)
Museum/exhibition	0.132	0.111	0.085	0.154
	(0.339)	(0.315)	(0.279)	(0.361)
Watch performance	0.113	0.102	0.087	0.146
	(0.316)	(0.302)	(0.281)	(0.353)
Dad speaks Mandarin to kid	0.756	0.818	0.589	0.652
	(0.430)	(0.386)	(0.492)	(0.476)
Dad speaks English to kid	0.002	0.004	0.001	0.026
	(0.043)	(0.066)	(0.026)	(0.159)
Mum speaks Mandarin to kid	0.807	0.877	0.620	0.703
	(0.395)	(0.329)	(0.485)	(0.457)
Mum speaks English to kid	0.002	0.004	0.009	0.018
	(0.044)	(0.059)	(0.093)	(0.132)

All variables are dummy variables, except for the number of siblings. The first row are mean values for each variable followed by standard deviations in parentheses.

A striking feature of the TASA database is the rich array of variables that can be calculated from other parts of the student questionnaires. We first discuss the parent-related variables which include variables measuring family structure, the educational attainment of each parent, a range of questions related to parent-child interaction, and the language spoken to the child by each parent. For brevity, we compare the mean values for the four groups and point out the groups with the highest and lowest values. The vast majority of parents from the TT group are staying together (81.9%), while almost a third of parents from the TC and OT groups are living separated (28.3% and 29.3%, respectively), and many of these are likely to stay outside of Taiwan. The TT group also exhibits the highest educational attainment of parents with 24.6% of fathers and 21.5% of mothers possessing a university degree. For both parents, the TS group is the one with the lowest parental educational attainment.

The second part of our parent-related variables focuses on parent behavior by asking the children about whether their parents undertake certain activities with them or talk about certain topics to them. The pupils are given four answer options: "never", "sometimes", "often" and "always". To facilitate the discussion of their marginal effects in section 4, we code these answers into dummy variables equal to one if something happens at least often ("often" or "always") and zero if it rarely happens ("sometimes" or "never"). The parents in the TT group generally care more about their child (72.9%) and are more likely to talk to their child about school (51.3%). They are also the most able parents in terms of helping their children with their homework (59.1%). TS parents score lowest on the first two categories, while TS and OT parents score almost equally low in terms of their ability to help with homework. The latter, however, are most likely to supervise their student homework (25.1%), and lay more emphasis on leisure activities, such as going to visit a museum or an exhibition (15.4%), or to go and watch a performance (14.6%). TT parents and OT parents are about equally likely to talk to their child about going to a bookstore or library (18.5% and 18.6%, respectively). From the behavioral variables, we can infer that TT parents are focused on school and homework, while OT parents pay relatively more attention to leisure activities that are not related to studying.

The languages spoken at home have been identified as another key determinant of school outcomes in previous research. The TC group is the one with the highest percentage of fathers (81.8%) and mothers (87.7%) speaking Mandarin-Chinese to their children, which is because parents from the TT group frequently speak local dialects at home. The percentages of Mandarin speakers are lowest in the TS group. The OT group is the highest in terms of the incidence of English spoken at home, but the percentages are still low at 2.8% for fathers and 1.8% for mothers.

Table 4 summarizes the remaining variables pertaining to the physical environment at home, student attitudes towards each subject, as well as to student interaction with classmates and teachers. In terms of the physical environment at home, children from the TT group have a better information technology environment at home in terms of higher likelihood to possess a computer (94.5%), internet access (89.0%) and study software (71.8%). They are also most likely to have their own bookshelf (86.8%) with study reference books (84.3%), study-related reading materials (89.1%), dictionary (98.6%) and encyclopedia (63.5%) that can be used in a quiet reading environment (82.6%). The TS group generally scores lowest in all of the above variables. TT pupils are also at an advantage in terms of subject- and classmate-related variables, whereas OT pupils score lowest on these variables.

Overall, children whose parents were both born in Taiwan exhibit a situation that is conducive to better educational performance relative to the other three groups in terms of parental characteristics and behavior, and also in terms of the study and IT environment at home. The differences between the different groups with foreign parents will serve as the backdrop for our analysis of the channels that contribute to the native-immigrant score performance gap analyzed in the following section 4.

Table 4. Descriptive statistics by parent origin (other variables)

	Taiwan-Taiwan	Taiwan-China	Taiwan-SE Asia	Other country
Computer at home	0.945	0.914	0.850	0.867
	(0.228)	(0.280)	(0.358)	(0.340)
Internet at home	0.890	0.849	0.736	0.811
	(0.313)	(0.358)	(0.441)	(0.391)
Study software at home	0.718	0.682	0.552	0.626
	(0.450)	(0.466)	(0.497)	(0.484)
Own bookshelf at home	0.868	0.838	0.725	0.811
	(0.338)	(0.368)	(0.447)	(0.392)
Study references at home	0.843	0.789	0.649	0.711
	(0.364)	(0.408)	(0.478)	(0.453)
Study readings at home	0.891	0.854	0.767	0.807
	(0.312)	(0.354)	(0.423)	(0.395)
Quiet reading area at home	0.789	0.740	0.641	0.703
	(0.408)	(0.439)	(0.480)	(0.457)
Dictionary at home	0.986	0.974	0.969	0.957
	(0.118)	(0.159)	(0.174)	(0.202)
Encyclopedia at home	0.634	0.499	0.364	0.568
	(0.482)	(0.500)	(0.481)	(0.495)
Likes subject	0.641	0.629	0.629	0.585
	(0.480)	(0.483)	(0.483)	(0.493)
Finds subject easy	0.554	0.549	0.554	0.530
	(0.497)	(0.498)	(0.497)	(0.499)
Classmates talk about homework	0.393	0.353	0.345	0.290
	(0.488)	(0.478)	(0.476)	(0.454)
Help from classmates	0.655	0.630	0.585	0.510
	(0.475)	(0.483)	(0.493)	(0.500)
Quarrels with classmates	0.060	0.079	0.083	0.134
	(0.237)	(0.270)	(0.276)	(0.341)
Teacher cares about kid	0.657	0.691	0.704	0.600
	(0.475)	(0.462)	(0.457)	(0.490)
Gets help from teacher	0.493	0.531	0.554	0.415
	(0.500)	(0.499)	(0.497)	(0.493)
Good interaction with teacher	0.598	0.634	0.614	0.532
	(0.490)	(0.482)	(0.487)	(0.499)
Likes teacher	0.584	0.625	0.627	0.543
	(0.493)	(0.484)	(0.484)	(0.498)

All variables are dummy variables. The first row are mean values for each variable followed by standard deviations in parentheses.

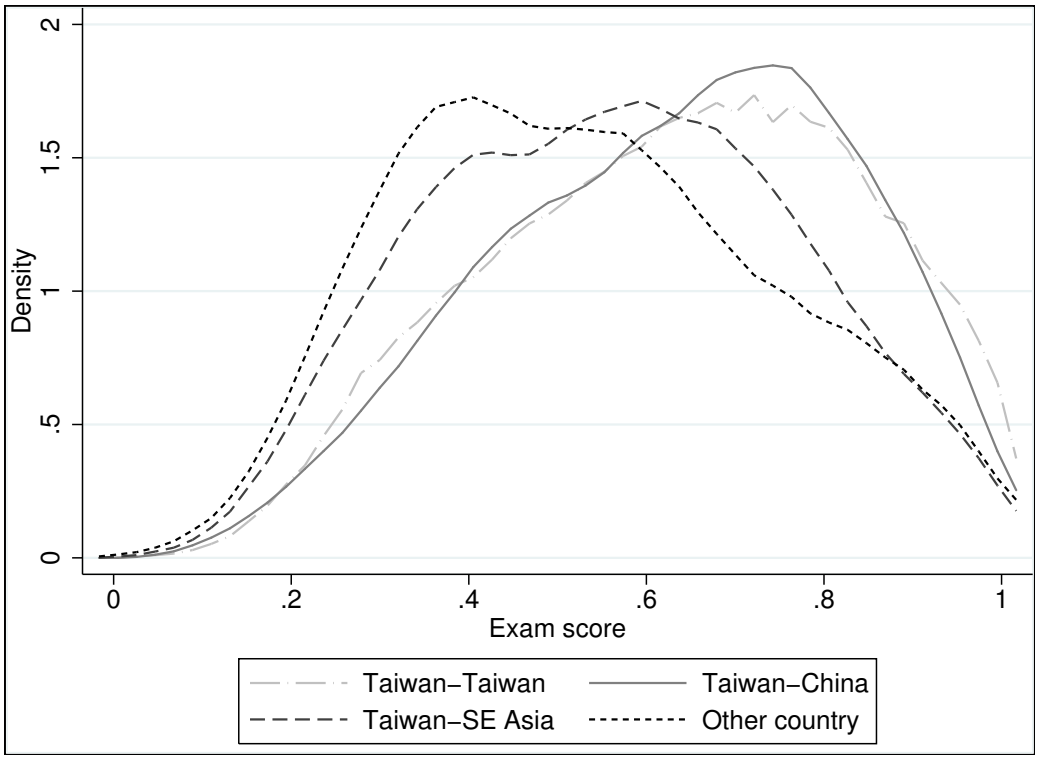


Figure 1. Density functions of score distributions by parent group

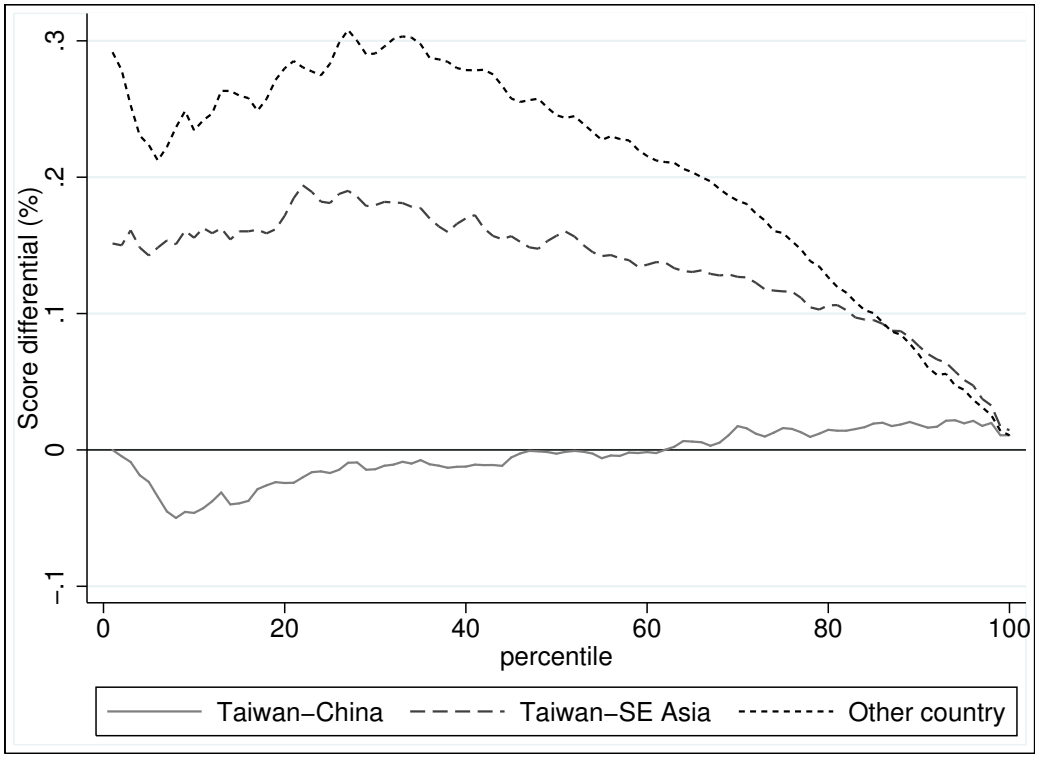


Figure 2. Score differentials at different percentiles by parent group

4. Analytical setting and results

4.1. *Basic quantile regression methodology*

We now turn to the question to what extent the above characteristics can be used to explain student test scores and the score gaps between natives and immigrants found in the previous section. In particular, we apply quantile regression methodologies to tackle this research question. The use of quantile regressions is motivated by the fact that the student test score distributions are not normally distributed and the test score gaps differ across quantiles. The relationship between the observable characteristics and our outcome variable is therefore likely to be nonlinear.

As depicted in Figure 1, the score distribution for the students from the OT group is slightly skewed to the left, while the other three distributions are skewed to the right with the highest skewness occurring for the strongest performing TT group. Figure 2 displays the resulting score differentials for each of the three foreign groups relative to students whose parents are both from Taiwan. We find that parent origin matters less for student performance in the upper tail of the score distributions, where second generation marriage immigrants are able to largely compensate for the adverse effects of their background. For students in the lower half of the distribution, we find large raw score gaps for the TS and OT groups. For the TC group, the raw score differentials in all percentiles are small and this group even slightly outperforms local students in the upper tail of the score distribution.

In order to obtain a preliminary assessment regarding the various factors behind student performance, we follow Koenker and Bassett (1978) and implement conditional quantile estimations with student scores as our dependent variable. Using y to denote the score of student i and X_i to denote a vector of covariates, the τth quantile of the conditional distribution of y_i as a linear function of the covariates can be written as:

$$q_\tau(y_i|X_i) = X_i\beta_\tau \tag{1}$$

Following the notation in Koenker and Hallock (2001), the quantile regression estimator for each β_τ solves the following minimization problem:

$$E(y_i|X_i) = \min_{\beta\in\mathbb{R}} \sum_{i=1}^{n} \rho_\tau(y_i - \xi(X_i, \beta)) \tag{2}$$

4.2. *Preliminary findings*

Table 5 presents our quantile regression results for the 25th, 50th and 75th percentiles. For our TC parent group, we find that the introduction of our parent variables, home environment variables and class variables renders our TC variable insignificant in the lower and middle tail of the score distribution. In the upper tail, the TC variable is insignificant after introducing any of our controls. Controlling for parent and home environment effects also significantly reduces the TS and OT differentials. These coefficients, however, remain significant even after introducing all of our control variables. Only our TS variable turns insignificant at the 5% level in the upper tail of the score distribution when we control for all factors listed in Tables 3 and 4. We also include interaction terms between our parent origin groups and student gender. For this part, we find no gender differences for the TC group, but significant differences in the higher

Table 5. Quantile regressions results with different control variable groups

	Q25 (1)	Q25 (2)	Q25 (3)	Q25 (4)	Q25 (5)	Q25 (6)	Q50 (1)	Q50 (2)	Q50 (3)	Q50 (4)	Q50 (5)	Q50 (6)	Q75 (1)	Q75 (2)	Q75 (3)	Q75 (4)	Q75 (5)	Q75 (6)
Dad (TW) Mum (CH)	−0.032**	−0.011	−0.037***	−0.014	−0.013	−0.008	−0.026**	−0.003	−0.033***	−0.009	−0.017*	0.001	−0.011	0.002	−0.013	0.003	−0.013	0.002
	(0.011)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.008)	(0.007)
Dad (TW) Mum (SA)	−0.090***	−0.059***	−0.072***	−0.051***	−0.082***	−0.040***	−0.078***	−0.042***	−0.059***	−0.035***	−0.072***	−0.021**	−0.061***	−0.027***	−0.045***	−0.028***	−0.057***	−0.014*
	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)
Dad (OT) Mum (OT)	−0.129***	−0.098***	−0.115***	−0.097***	−0.101***	−0.066***	−0.139***	−0.106***	−0.120***	−0.114***	−0.118***	−0.084***	−0.108***	−0.081***	−0.097***	−0.091***	−0.094***	−0.067***
	(0.011)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)	(0.008)	(0.007)
Dad (TW) Mum (CH) (male)	0.024	0.023	0.028	0.015	0.021	0.016	0.006	0.007	0.014	0.006	0.004	0.002	−0.010	−0.009	−0.013	−0.006	−0.011	−0.004
	(0.016)	(0.014)	(0.015)	(0.014)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.011)	(0.010)
Dad (TW) Mum (SA) (male)	0.006	0.006	0.012	0.018	0.010	0.013	−0.020	−0.023*	−0.028**	−0.021*	−0.019	−0.016	−0.036***	−0.038***	−0.032***	−0.021*	−0.036***	−0.027**
	(0.014)	(0.012)	(0.013)	(0.012)	(0.012)	(0.011)	(0.012)	(0.011)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.009)	(0.009)	(0.009)	(0.010)	(0.009)
Dad (OT) Mum (OT) (male)	−0.002	−0.006	−0.007	−0.009	−0.002	−0.006	−0.022	−0.028*	−0.024*	−0.012	−0.012	−0.004	−0.023*	−0.032**	−0.031**	−0.017	−0.022*	−0.020*
	(0.016)	(0.014)	(0.014)	(0.013)	(0.013)	(0.012)	(0.013)	(0.012)	(0.012)	(0.012)	(0.012)	(0.011)	(0.011)	(0.010)	(0.010)	(0.010)	(0.011)	(0.010)
Basic controls	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes	yes
Parent controls		yes				yes		yes				yes		yes				yes
Language controls			yes			yes			yes			yes			yes			yes
Home controls				yes		yes				yes		yes				yes		yes
Class controls					yes	yes					yes	yes					yes	yes
Observations	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121	152121
Pseudo R^2	0.070	0.121	0.087	0.113	0.123	0.178	0.069	0.118	0.087	0.112	0.120	0.176	0.074	0.114	0.088	0.107	0.113	0.161

T-statistics are displayed in parentheses. The significance symbols denote: $^{*}p < 0.1$, $^{**}p < 0.05$, $^{***}p < 0.01$. The basic regression conctrols are student gender, number of siblings, as well as dummy variables for study subject and grade. The parent controls are parent background characteristics and variables measuring parent behaviour, as displayed in table 3. Language controls are variables for each parent measuring whether the parent speaks each of the following languages to the child: Mandarin, Hokkien, Hakka, English or an aboriginal language. The home controls are the ten variables measuring the IT and study environment at home, as displayed in table 4. Class controls refer to the variables measuring student-teacher and student-classmate interaction displayed in the same table.

quantiles for the other two groups. The adverse effect of having a mother from Southeast Asia is particularly pronounced for female students in higher quantiles.

In summary, we find that our controls fully remove the score gap between students with a mother from mainland China and a father from Taiwan relative to students whose parents are both from Taiwan. They also almost fully remove the score gap for students with a mother from Southeast Asia in the upper part of the score distribution, whereas they are more useful in explaining parts of the score gap for students with parents from other countries in the lower part of the wage distribution. We briefly discuss how each individual variable affects the student scores in different parts of the distribution, before turning to a detailed composition of these score gaps at different quantiles.

4.3. *Channels affecting the student scores*

In line with previous research findings by Schnepf (2007) and Dustmann, Frattini, and Lanzara (2012), parental characteristics are a significant determinant of student academic performance. The educational background of both father and mother exerts a positive impact on student grades throughout all quantiles. Interestingly, these effects are generally larger in lower quantiles, and the highest positive effect is found for fathers with university education in these quantiles. Other parent variables that are found significant are whether parents are staying together with the child, talking about library or bookstore, talking about what happened in school and the ability to help with a child's homework. Children whose parents frequently talk about going to watch performances, however, score lower across all quantiles.

Variables measuring language usage by father and mother are shown in Figure 4. Unsurprisingly, the highest coefficients are found for Mandarin usage, which is also the language of instruction in schools, hence corroborating the finding in Bleakley and Chin (2008). Speaking Hakka or an Aboriginal language at home, however, correlates negatively with student performance.

The quantile regression coefficient results for home environment variables in Figure 5 confirm our expectation that this is a key channel affecting student performance. In particular, possessing a computer and books, such as dictionary, study-reference books and study-related reading material all positively affect student performance. Another striking feature is that the book-related variables exhibit a U-shaped pattern. These variables are therefore most relevant around the median of the score distribution, whereas they affect the best and worst performing students to a lesser extent. The effects of owning a computer and internet access are significant and positive at a similar magnitude throughout the score distribution. We find positive coefficients for subject-related variables, and getting help from classmates or talking to classmates about homework can help poorly performing students, whereas quarrelling with classmates harms these students most (see Figure 6). Coefficients of the teacher-related variables are generally small or insignificant.

Overall, we find significant effects for parent background characteristics, language usage at home, home environment variables, and also for attitude towards the exam subjects and student-classmate interaction. Bearing in mind the differences in endowments and background characteristics between different student groups as depicted in Tables 3 and 4 and the effect of each of these variables on student performance shown in the figures in this subsection, it is straightforward to conjecture that the various variables play different roles in determining the test score gap for different student

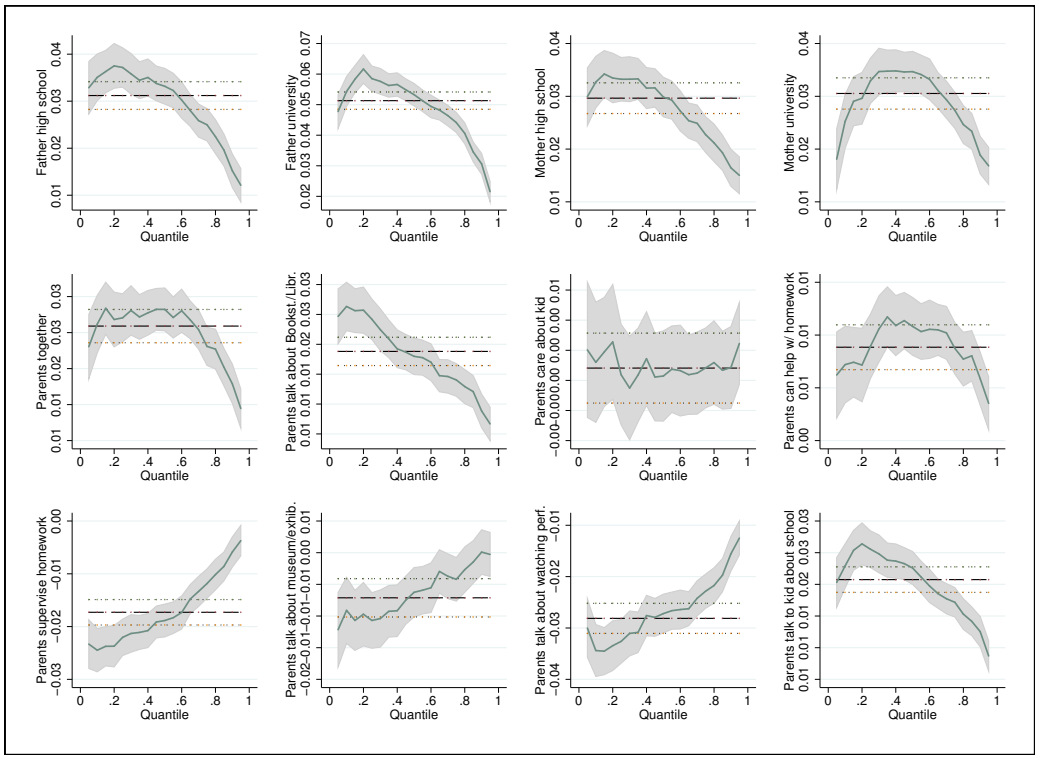


Figure 3. Coefficients of parent variables across quantiles

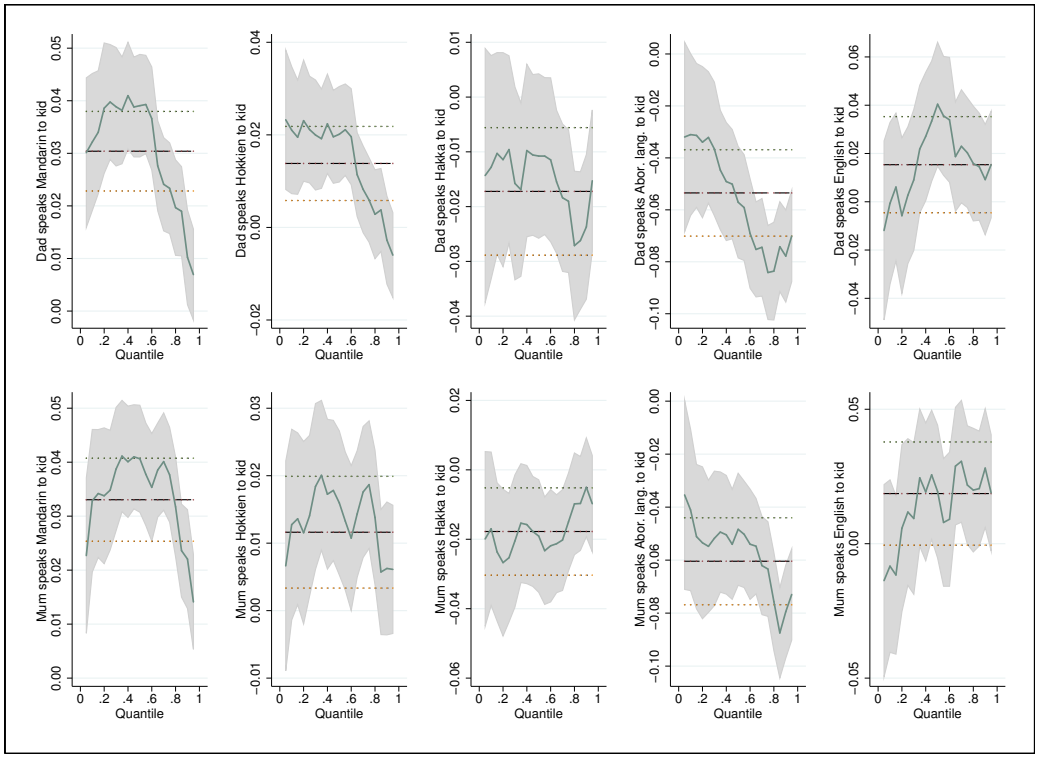


Figure 4. Coefficients of home language variables across quantiles

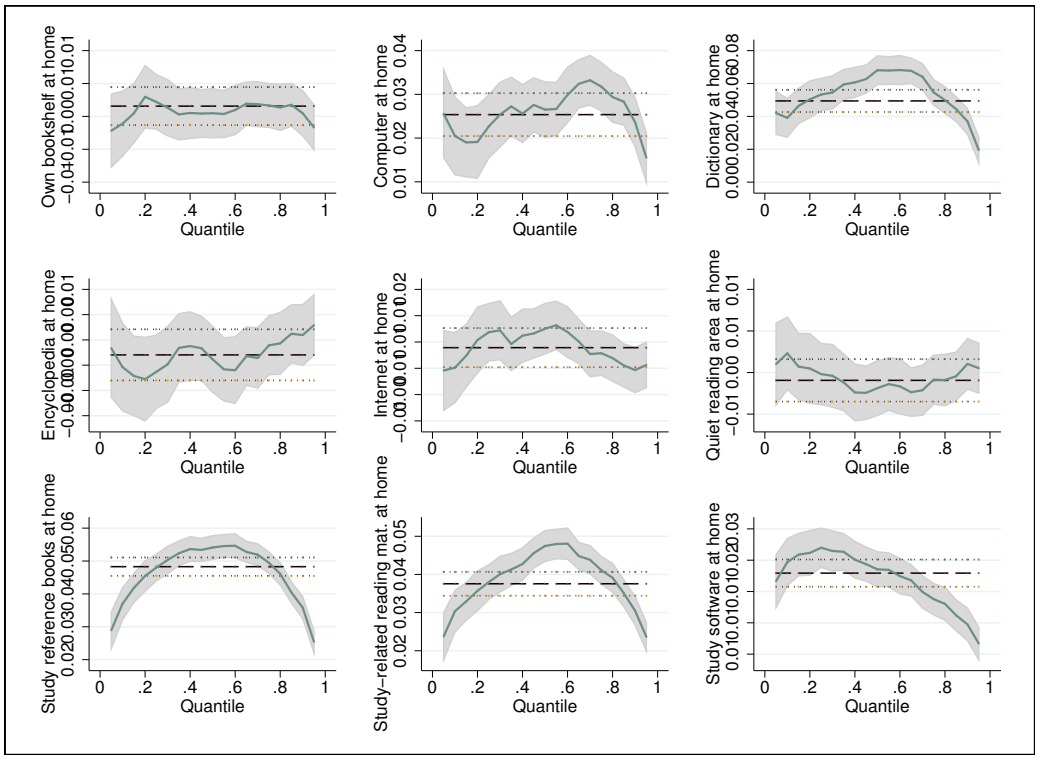


Figure 5. Coefficients of home environment variables across quantiles

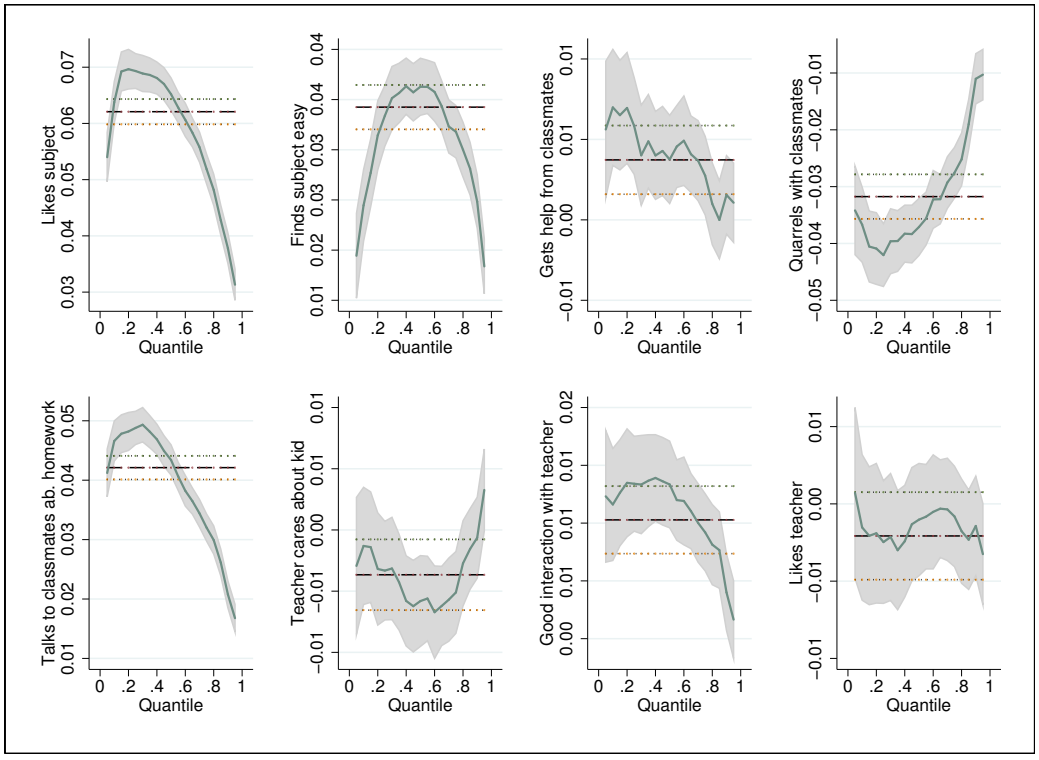


Figure 6. Coefficients of class environment variables across quantiles

groups.

4.4. *Quantile decompositions of score gaps by variable group*

We now move on to the decomposition of the student test score gaps between the three second generation immigrant groups and the local students. The starting point of our test score gap decomposition are the decomposition techniques at mean values introduced into the labor economics literature by Oaxaca (1973) and Blinder (1973) (henceforth: OB). The basic idea of this approach is that differences in an outcome variable Y can be decomposed into the differences between the endowments of covariates X in each group and differences in the returns to these endowments. It is then straightforward to decompose the explained part into the contribution of each individual variable or variable group, because the total explained part is the sum of each individual contribution (Jann 2008).[6]

Based on our preceding analysis, we expect these effects to differ between different quantiles in the score distribution. The approach we adopt for the decomposition at different quantiles is the unconditional quantile approach developed by Firpo, Fortin, and Lemieux (2009). This method involves first estimating a recentered influence function (RIF) and then replacing the outcome variable with the values obtained from this estimation. The OB approach can then be implemented as a local approximation for the effect of changes in each covariate on the unconditional distribution.[7]

Using q_τ to denote the population τ-quantile of the unconditional distribution of our student test scores y and $\mathrm{IF}(y, q_\tau)$ to denote the corresponding influence function, the RIF can be written as:

$$\mathrm{RIF}(y, q_\tau) = q_\tau + \mathrm{IF}(y, q_\tau) = c_{1,\tau} \cdot \Pi\{y > q_\tau\} + c_{2,\tau} \tag{3}$$

Where $c_{1,\tau} = 1/f_y(Q_\tau)$ and $c_{2,\tau} = Q_\tau - c_{1,\tau} \cdot (1 - \tau)$, meaning that the RIF for each quantile is essentially an indicator variable telling whether the variable is less than or equal to the quantile Q_τ, multiplied by a factor and plus a constant. Using g to denote each of our student groups, i.e. g $\in$ {TT, TC, TS, OT}, or simply g $\in$ {TT, IM}, where IM denotes either of the three second-generation immigrant groups, the coefficient estimates of the unconditional quantile regression at each quantile τ can be written as:

$$\hat{\gamma}_{g,\tau} = \left(\sum_{i \in G} X_i \cdot X_i^\tau \right)^{-1} \cdot \sum_{i \in G} \widehat{\mathrm{RIF}}(Y_{gi}; Q_{g,\tau}) \cdot X_i \tag{4}$$

[6] We do not decompose the unexplained part because of the problems with attributing the unexplained part to covariates, in particular dummy variables, discussed in the literature (see, e.g., Jones (1983) and Fortin, Lemieux, and Firpo (2011)).

[7] The major advantage of the method proposed by Firpo, Fortin, and Lemieux (2009) is that their decomposition is path independent and a detailed decomposition for different variable groups can therefore be implemented. Other approaches, such as the ones proposed by Machado and Mata (2005) and Chernozhukov, Fernández-Val, and Melly (2013) are path dependent and the coefficient estimates for variable groups differ depending on the sequence in which these variables are introduced into the model, thus rendering a detailed decomposition cumbersome.

An OB decomposition at any quantile can then be performed as follows:

$$\underbrace{\hat{\Delta}^{\tau}}_{\text{total gap}} = \underbrace{\bar{X}_B\left(\hat{\gamma}_{TT,\tau} - \hat{\gamma}_{IM,\tau}\right)}_{\text{unexplained}} + \underbrace{\left(\bar{X}_{TT} - \bar{X}_{IM}\right)\hat{\gamma}_{IM,\tau}}_{\text{explained}} \tag{5}$$

An important feature is that the explained part is then again the sum of the contributions of each of our seven variable groups k, as in the initial OB decomposition:

$$\left(\bar{X}_{TT} - \bar{X}_{IM}\right)\hat{\gamma}_{IM,\tau} = \sum_{k=1}^{K}\left(\bar{X}_{TTk} - \bar{X}_{IMk}\right)\hat{\gamma}_{IMk,\tau} \tag{6}$$

We first present the OB decomposition results at mean values for our complete sample as displayed in Table 6. Our key findings are as follows. For the TC group, the insignificance of the performance gap is due to several offsetting factors. The home environment, parental behavior and parental education levels are the key factors contributing to the advantage of the TT group. On the other hand, the higher incidence of Mandarin spoken in TC households as well as an unexplained residual remove the gap produced by the above variables.[8] For the TS group, our covariates are able to explain about two thirds of the initial gap and reduce it from an initial 11.0% to about 3.9%. The home environment and parent educational attainment emerge as the key factors behind the educational performance inequalities between the TS group and the TT group. In particular, for the TS group, the home environment variables alone contribute more than half of the explanatory power of all variables combined. For the OT group, our covariates are significant, but can only explain a small portion of the performance gap of this group. The largest factor is again the home environment, which explains about 16.5% of the initial gap. After introducing all factors, the initial gap of 15.3% is reduced to an unexplained residual gap of 12.0%. Overall, we identify the home environment as the key factor explaining the immigrant-local student test score gap, and the two variables measuring the availability of reference books and study-related reading materials contribute 57, 65 and 63 per cent of the overall home environment effect for the TC, TS and OT groups, respectively.

Figure 7 moves beyond the decomposition at mean values and shows how the relative impact of the different explanatory variables varies across nine quantiles along the score distribution.[9] We find that parent behavior is the major force working against TC students in higher quantiles, whereas the home environment is the main factor in lower quantiles. Speaking Mandarin at home reduces the gap vis-à-vis TT students across all quantiles, while parent education levels increase the gap in all quantiles. For the TS group, the home environment is the key factor across all quantiles and this group alone contributes more than half of the explanatory power in lower quantiles. Parental education again exerts an important impact in all quantiles. For the OT group, a range of factors matter for the test score gap. While the effects of parent education and student-classmate interaction are almost constant across quantiles, parent behaviour matters more in higher quantiles, while the physical home environment effect is again largest in lower quantiles. From a policy perspective, our quantile results underline

[8]The insignificant gap and the positive language effect also corroborate the finding on mainland Chinese political immigrants in Taiwan reported in Tsay (2006). As conjectured by the author, these outperformed native Taiwanese due to linguistic advantages and a psychological effect caused by the refugee experience, which induced their parents to increase investments into their education.

[9]In these graphs, we only retain the variables that are significant at least at the 10% level.

Table 6. Oaxaca decomposition for complete sample

	Taiwan-China	Taiwan-SE Asia	Other country
Prediction (TT)	0.634	0.634	0.634
	(0.000)	(0.000)	(0.000)
Prediction (IM)	0.633	0.564	0.537
	(0.004)	(0.004)	(0.004)
Difference	0.001	0.070***	0.097***
	(0.004)	(0.004)	(0.004)
Basic	−0.011***	−0.009***	−0.027***
	(0.001)	(0.001)	(0.001)
Education	0.003***	0.013***	0.003***
	(0.001)	(0.000)	(0.001)
Behavior	0.005***	0.005***	0.008***
	(0.001)	(0.000)	(0.000)
Home	0.007***	0.023***	0.016***
	(0.001)	(0.001)	(0.001)
Subject	0.001	0.001	0.004***
	(0.001)	(0.001)	(0.001)
Language	−0.003***	0.010***	0.008***
	(0.000)	(0.001)	(0.001)
Classmates	0.002***	0.003***	0.007***
	(0.001)	(0.000)	(0.001)
Teacher	−0.001***	−0.000	0.001***
	(0.000)	(0.000)	(0.000)
Total	0.005**	0.045***	0.021***
	(0.002)	(0.002)	(0.003)
Unexplained	−0.004	0.025***	0.076***
	(0.004)	(0.003)	(0.004)
Observations	177013	177712	177233

TT refers to the Taiwan-Taiwan group. IM refers to the second-generation immigrant group in the respective column. The notation for statistical significance is the same as in Table 5.

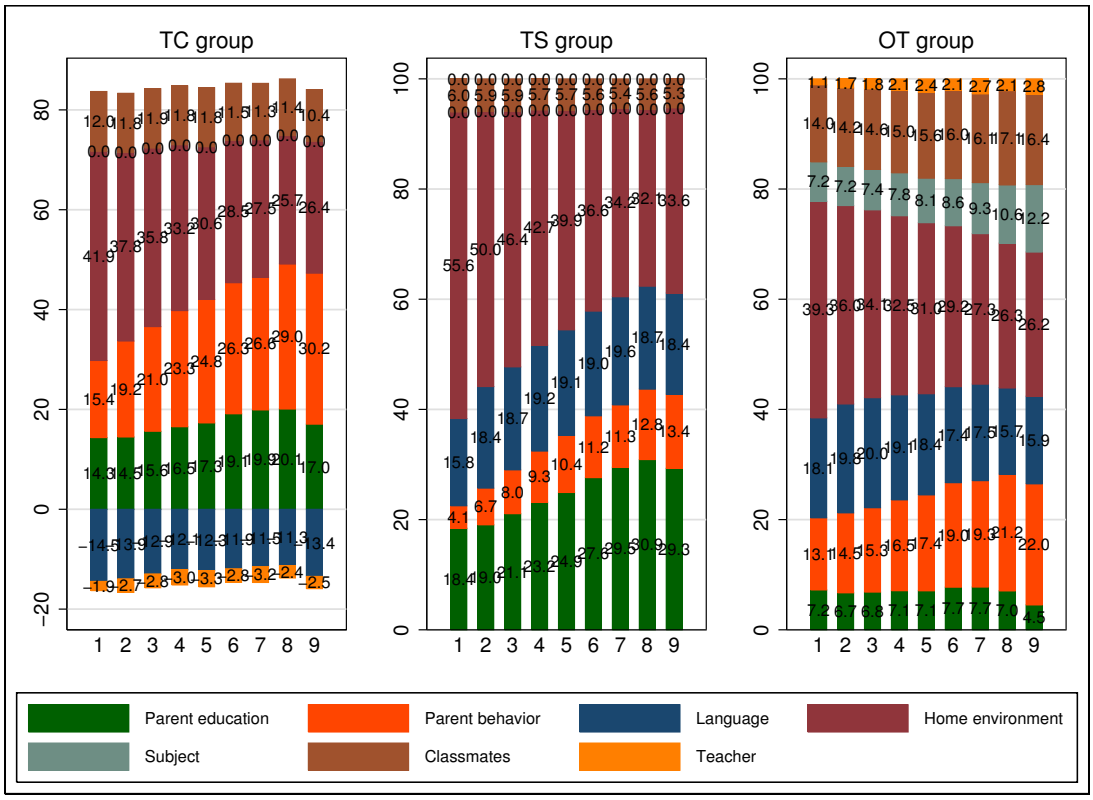


Figure 7. Factors explaining score gaps by parent group (complete sample)

the fact that potential remedies aimed at raising the performance of second-generation immigrants should not only be tailored according to parent origin country, but should also differ for students at different quantiles along the score distribution.

4.5. *Findings by study subject and age group*

4.5.1. Findings for five study subjects

In this subsection, we follow the same sequency as in the analysis of our complete sample above by first presenting raw score differentials (Figure 8), then conducting an OB decomposition (Table 7) and finally presenting a more detailed decomposition at different quantiles (Figure 9). Students from the TC group outperform TT students in Chinese and social studies, i.e. in the two subjects that rely most heavily on the ability to express oneself in the language of instruction. TC students perform similar to TT students in mathematics and natural studies, while TT students outperform all other students in English. As displayed in Table 7 below, the score gaps between TT and TC students in social studies, natural studies and mathematics are insignificant at mean values. Analogous to the results for our complete sample, OT students perform worst in all subjects except English. TS students are outperformed by TC and TT students in all subjects, and also by OT students in the subject English.

Moving on to the factors behind these gaps, the OB decomposition results in Table 7 reveal that the home environment is the key factor behind the score gaps across almost all subjects, with the TC differential in social studies that is mostly explained by parent education being the only exception. Notably, our covariates render the score

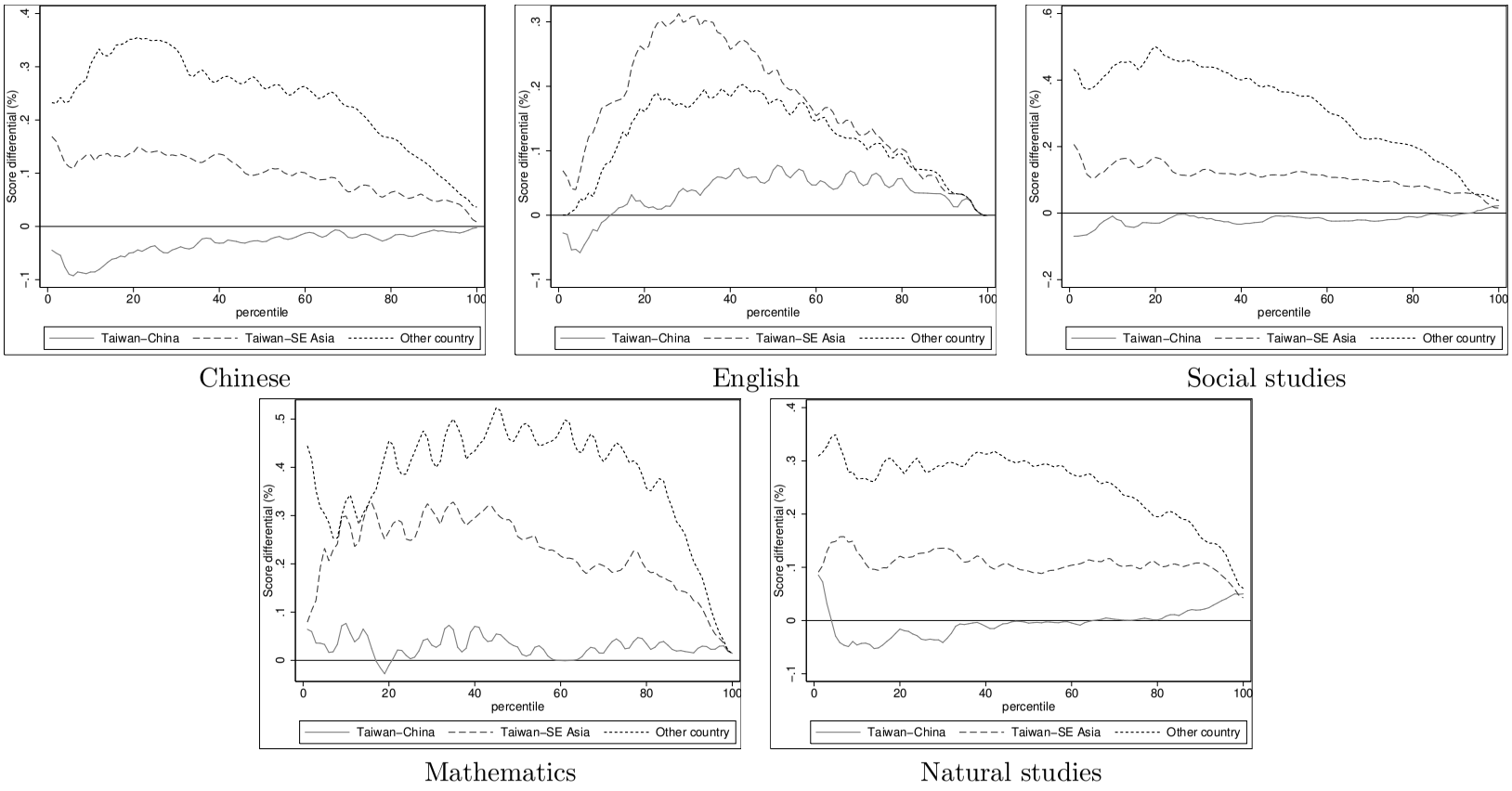


Figure 8. Score gaps for different subjects by parent group

Table 7. Oaxaca decomposition by study subject

	Chinese			English			Social studies			Natural studies			Mathematics		
	TC	TS	OT	TC	TS	OT	TC	TS	OT	TC	TS	OT	TC	TS	OT
Differential															
Prediction (TT)	0.666	0.666	0.666	0.715	0.715	0.715	0.595	0.595	0.595	0.581	0.581	0.581	0.603	0.603	0.603
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prediction (IM)	0.685	0.611	0.547	0.690	0.622	0.642	0.605	0.541	0.468	0.582	0.527	0.470	0.588	0.501	0.445
	(0.008)	(0.007)	(0.008)	(0.009)	(0.009)	(0.007)	(0.009)	(0.008)	(0.012)	(0.008)	(0.007)	(0.008)	(0.011)	(0.009)	(0.009)
Difference	−0.018**	0.055***	0.119***	0.026***	0.093***	0.074***	−0.010	0.054***	0.127***	−0.001	0.054***	0.111***	0.015	0.102***	0.158***
	(0.008)	(0.007)	(0.008)	(0.010)	(0.009)	(0.007)	(0.009)	(0.008)	(0.012)	(0.008)	(0.007)	(0.008)	(0.011)	(0.009)	(0.009)
Basic	−0.012***	−0.008***	−0.018***	−0.020***	−0.021***	−0.046***	−0.004***	−0.002**	0.003	−0.005***	−0.001	0.002	−0.002	0.002	0.023***
	(0.002)	(0.001)	(0.002)	(0.004)	(0.003)	(0.003)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)
Education	0.003***	0.011***	0.001	0.001	0.011***	0.007***	0.006***	0.016***	−0.005**	0.003***	0.012***	0.003**	0.005***	0.015***	0.003*
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Behavior	0.004***	0.007***	0.008***	0.007***	0.005***	0.012***	0.003**	0.006***	0.007***	0.004***	0.003***	0.006***	0.006***	0.004***	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Home	0.006***	0.022***	0.016***	0.011***	0.030***	0.021***	0.003*	0.018***	0.011***	0.005***	0.015***	0.011***	0.008***	0.024***	0.015***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Subject	−0.000	−0.001***	−0.000	0.010***	0.010***	0.013***	0.000	−0.002*	0.004**	−0.001	−0.002	−0.003**	0.002	0.009***	0.005
	(0.000)	(0.000)	(0.000)	(0.003)	(0.003)	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)
Language	−0.003***	0.011***	0.008***	−0.003***	0.007***	0.006***	−0.003***	0.010***	0.009***	−0.002***	0.008***	0.008***	−0.002	0.012***	0.011***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Classmates	0.002**	0.004***	0.007***	0.001	0.004***	0.006***	0.002	0.002*	0.006***	0.003***	0.002**	0.008***	0.003***	0.002*	0.008***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Teacher	−0.002**	−0.000	0.004***	0.001	0.001**	−0.000	−0.001	−0.001	0.002**	−0.001	−0.001	0.002**	−0.001	−0.001	0.002***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Total	−0.001	0.045***	0.025***	0.007	0.047***	0.019***	0.005	0.047***	0.035***	0.006*	0.038***	0.036***	0.020***	0.068***	0.076***
	(0.004)	(0.004)	(0.005)	(0.007)	(0.006)	(0.005)	(0.004)	(0.004)	(0.007)	(0.004)	(0.003)	(0.004)	(0.006)	(0.005)	(0.007)
Unexplained	−0.017**	0.011	0.094***	0.019**	0.046***	0.054***	−0.015*	0.007	0.092***	−0.007	0.017***	0.076***	−0.005	0.034***	0.082***
	(0.007)	(0.007)	(0.008)	(0.008)	(0.007)	(0.006)	(0.008)	(0.008)	(0.011)	(0.007)	(0.006)	(0.007)	(0.009)	(0.008)	(0.009)
Observations	37829	37937	37781	36397	36595	36750	29837	29944	29705	36534	36681	36547	36416	36555	36450

TT refers to the Taiwan-Taiwan group. IM refers to the immigrant group in the respective column. The notation for statistical significance is the same as in Table 5.

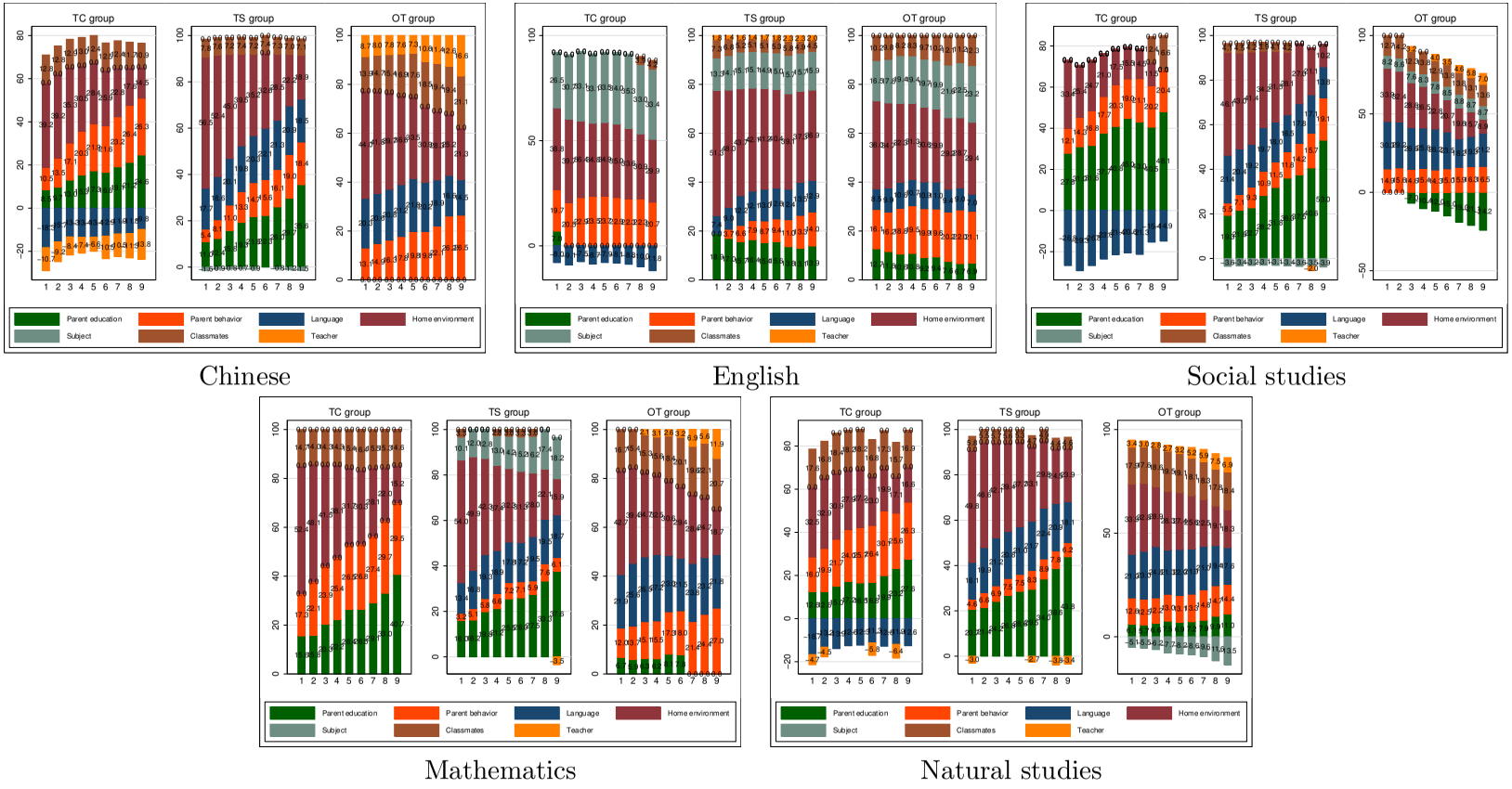


Figure 9. Factors explaining score gaps for different subjects by parent group

gaps for the TS group insignificant for the subject Chinese and for social studies, and the home environment, parent education and home languages contribute more than two thirds of the reduction in the gap. Yet, large gaps again remain even after introducing our covariates for the OT group, and the largest residual gap is found for Chinese, where about 79% of the gap remains unexplained.

Figure 9 depicts how these factors vary across quantiles in the score distribution. The home environment is again identified as a variable that matters across the whole distribution in most subjects and for all student groups, and the only part where it does not take an effect is the TC group in the subject social studies. Parent education matters for the TS group in all subject, and for the TC group it is the only significant factor raising the score gap in social studies. The home language is conducive to better performance for low-performing students in the TC group for the subjects Chinese, social studies and natural studies, but it raises the gap in all subjects for the TS and OT groups, in particular in lower quantiles.

4.5.2. Findings across three age groups

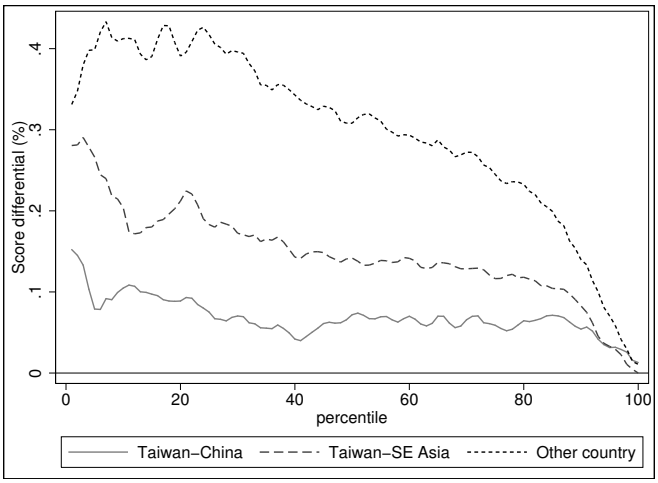
We now turn our attention to the three age groups in our analysis — primary-4, primary-6 and junior-2, which approximately correspond to student ages 10, 12 and 14, respectively.[10] As displayed in Figure 10, the relative overall performance of the three second-generation immigrant student groups does not change over time. The only major change is that the TC group performs worse than the TT group in primary-4, but is then able to close this gap as students proceed to primary-6 and junior-2.

Regarding the role of different factors in explaining the score gaps at different student ages, we find that several factors differ over time (Table 8). In particular, the role of the home environment decreases with student age, while other factors such as interaction with teacher and classmates and attitude towards study subjects gain importance over time. The attitude towards the subject generally matters more at higher student ages, and it matters more in higher quantiles for the case of the OT group (see Figure 11), hence exacerbating the score gap in this part of the distribution. For the OT group, classmate interaction is gaining importance as student age increases, rising to about one fifth of total explanatory power in our junior-2 group. We find the largest residual score gaps for our primary-4 students and the lowest for our junior-2 students, reflecting the fact that our covariates gain importance as students progress through the education system.

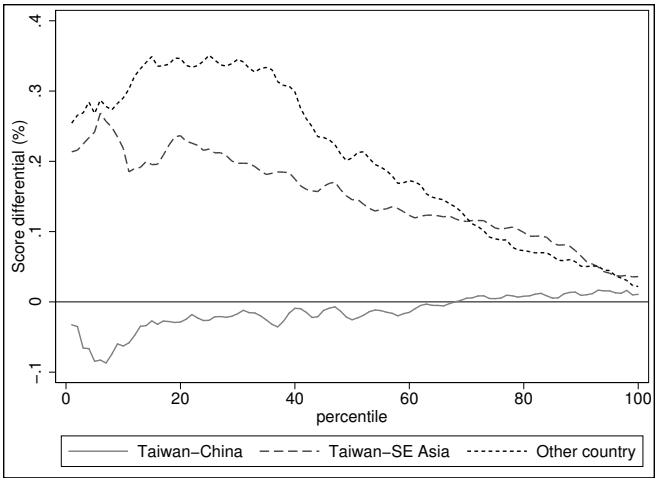
5. Conclusions

International marriages have become a pervasive demographic trend globally and in particular in East Asian societies. In this paper, we focus on an important consequence of this trend by analysing the incidence and determinants of educational inequalities between children of local parents and children of these cross-border marriages in Taiwan. While the average test score gaps of second-generation immigrants amount to approximately 9.2%, these differ significantly across different parent groups, with no significant gap for students with a father born in Taiwan and a mother born in mainland China. Marriage migration from a culturally highly similar location therefore does not cause educational inequalities for the second generation in our study. We find

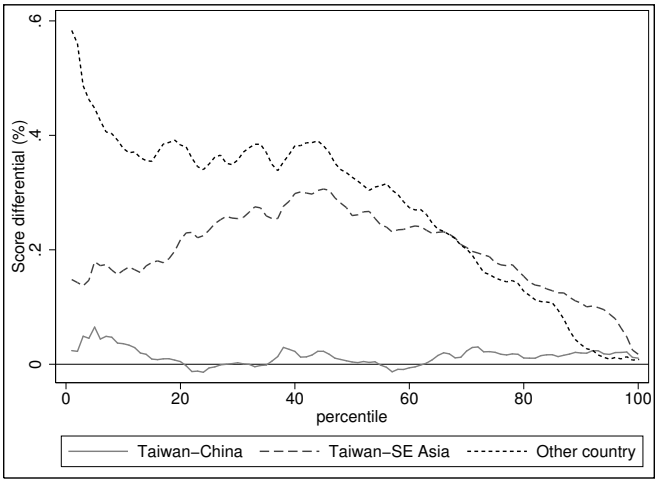
[10]To facilitate the comparison across age groups, we exclude the social studies test scores from this part of our analysis, because the subject has not been tested in primary-4.



Primary-4



Primary-6



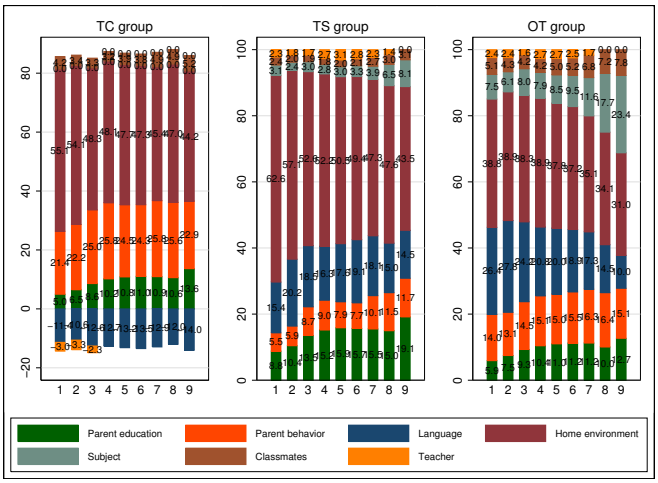
Junior high-2

Figure 10. Score gaps at different student age groups by parent group

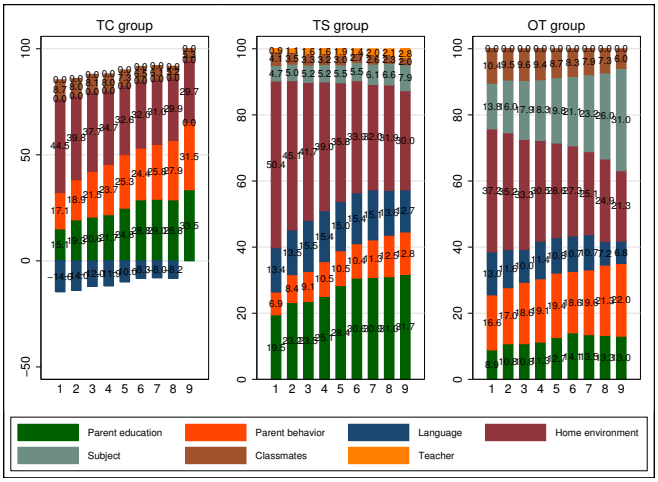
Table 8. Oaxaca decomposition by age group

	Primary-4			Primary-6			Junior high-2		
	TC	TS	OT	TC	TS	OT	TC	TS	OT
Prediction (TT)	0.680	0.680	0.680	0.648	0.648	0.648	0.603	0.603	0.603
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Prediction (IM)	0.639	0.602	0.538	0.652	0.578	0.551	0.603	0.514	0.497
	(0.009)	(0.008)	(0.006)	(0.006)	(0.005)	(0.007)	(0.008)	(0.007)	(0.012)
Difference	0.041***	0.078***	0.142***	−0.004	0.071***	0.097***	0.001	0.090***	0.106***
	(0.009)	(0.008)	(0.006)	(0.006)	(0.005)	(0.007)	(0.008)	(0.007)	(0.012)
Basic	−0.001	−0.006	0.003	−0.006***	−0.003	−0.028***	−0.006***	−0.000	0.003
	(0.005)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
Education	0.002***	0.007***	0.004***	0.004***	0.015***	0.005***	0.003**	0.015***	−0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Behavior	0.006***	0.004***	0.006***	0.004***	0.005***	0.010***	0.005***	0.005***	0.009***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Home	0.013***	0.026***	0.015***	0.006***	0.020***	0.016***	0.001	0.021***	0.011***
	(0.002)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)
Subject	0.000	0.002*	0.004***	0.002	0.001	0.009***	0.006***	0.005***	0.015***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)
Language	−0.003***	0.009***	0.008***	−0.002***	0.008***	0.006***	−0.002***	0.011***	0.011***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Classmates	0.001**	0.001***	0.002***	0.001**	0.002***	0.005***	0.003**	0.006***	0.014***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.002)
Teacher	−0.001*	0.001***	0.001***	0.000	0.001***	−0.000	−0.004***	−0.005***	0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)
Total	0.018***	0.044***	0.043***	0.009***	0.049***	0.022***	0.007	0.058***	0.056***
	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)	(0.008)
Unexplained	0.023***	0.035***	0.099***	−0.013**	0.022***	0.076***	−0.007	0.032***	0.050***
	(0.007)	(0.006)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.006)	(0.010)
Observations	26545	26718	27171	74954	75310	74913	75514	75684	75149

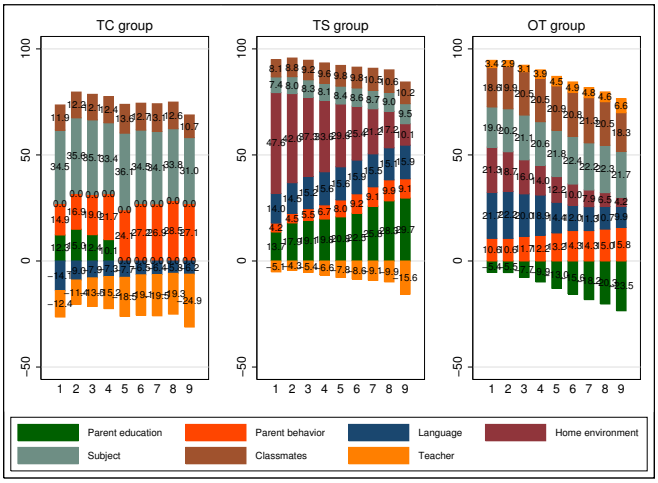
TT refers to the Taiwan-Taiwan group. IM refers to the respective immigrant group. The student age groups are: primary school grade 4 (PR-4), primary school grade 6 (PR-6) and lower secondary school grade 2 (JH-2). The notation for statistical significance is the same as in Table 5.



Primary-4



Primary-6



Junior high-2

Figure 11. Factors explaining score gaps at different student age groups by parent group

the largest gap of about 15.3% for children with at least one parent that is neither from Taiwan nor mainland China, nor from Southeast Asia. Children with a father who was born in Taiwan and a Southeast Asian mother constitute the intermediate case at raw score gaps of about 11.1%. We investigate what elements contribute to these score differentials and our findings reveal that these factors differ significantly by parent origin country and also across different quantiles of the score distribution.

In contrast to previous studies from other countries that were unable to simultaneously include parent characteristics, home language variables and variables pertaining to the physical home environment, we find that when jointly analyzing these and other factors, the physical home environment emerges as the most important determinant of the local-immigrant student test score gap. In particular, amongst the physical home environment variables, the effect of having non-electronic learning materials and study reference books contributes between half and two thirds of the home environment effect and the relative effect of these variables is largest in lower quantiles. The most affected student group are those with a mother from Southeast Asia. Being disadvantaged in terms of their home environment contributes more than half of the explanatory power of all variables combined for this student group. A policy implication of our study is therefore that subsidizing these materials can be instrumental in removing the local-immigrant test score gap, in particular for children with a parent from Southeast Asia or other more distant birthplaces.

Even though their role is not as important as the physical home environment, the educational attainment of parents and the languages spoken at home also contribute to the test score gaps in our analysis. In particular, the impact of parent education generally increases in higher quantiles, while the effect of speaking the language of class instruction decreases. Providing additional opportunities for low performing students with a mother from Southeast Asia or a parent from more distant locations to speak the language of class instruction is therefore also a necessary policy response that is likely to have the biggest effect on the test score gaps in the subjects Chinese and social studies.

As immigrant children grow older, student-teacher and student-classmate interaction, as well as attitudes towards study subjects emerge as additional important determinants of test score gaps. Starting from junior high school age, increased attention should be paid to these factors to reduce the local-immigrant test score gap. Another interesting finding that emerges from our comparison across age groups is that observable factors are generally most able to explain the test score gaps in our oldest student group and this is also where we find the smallest residual score gaps. This is hence also the age group where policy-interventions may be more effective, because unobserved or unmeasurable factors have a smaller impact compared to younger student age groups.

Overall, the nuanced findings in our study call for tailor-made policy responses in order to alleviate the educational performance inequalities that exist for children of foreign-born marriage migrants. Without such responses, the inequalities are likely to persist until senior high school age and labor market entry, and further long-term consequences for the affected groups at later stages throughout their careers are likely to occur.

References

Baert, Stijn, and Bart Cockx. 2013. "Pure ethnic gaps in educational attainment and school to work transitions: When do they arise?" *Economics of Education Review* 36: 276–294.

Bleakley, Hoyt, and Aimee Chin. 2008. "What Holds Back the Second Generation?" *Journal of Human Resources* 43 (2): 267–298.

Blinder, Alan. 1973. "Wage Discrimination: Reduced Form and Structural Estimates." *Journal of Human Resources* 8 (4): 436–455.

Borjas, George J. 2006. "Making It in America: Social Mobility in the Immigrant Population." *The Future of Children* 16 (2): 55–71.

Chernozhukov, Victor, Iván Fernández-Val, and Blaise Melly. 2013. "Inference on Counterfactual Distributions." *Econometrica* 81 (6): 2205–2268.

Colding, Björg, Leif Husted, and Hans Hummelgaard. 2009. "Educational progression of second-generation immigrants and immigrant children." *Economics of Education Review* 28 (4): 434–443.

Dustmann, Christian, Tommaso Frattini, and Gianandrea Lanzara. 2012. "Educational Achievement of Second Generation Immigrants: An International Comparison." *Economic Policy* 27 (69): 143–185.

Dustmann, Christian, Stephen Machin, and Uta Schönberg. 2010. "Ethnicity and Educational Achievement in Compulsory Schooling." *The Economic Journal* 120 (546): F272–F297.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux. 2009. "Unconditional Quantile Regressions." *Econometrica* 77 (3): 953–973.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. 2011. "Decomposition Methods in Economics." In *Handbook of Labor Economics*, edited by Orley Ashenfelter and David Card, Vol. 4, part A, 1–102. Elsevier.

Fryer, Roland G., and Steven D. Levitt. 2004. "Understanding the Black-White Test Score Gap in the First Two Years of School." *The Review of Economics and Statistics* 86 (2): 447–464.

Fryer, Roland G., and Steven D. Levitt. 2006. "The Black-White Test Score Gap Through Third Grade." *American Law and Economics Review* 8 (2): 249–281.

Jann, Ben. 2008. "The Blinder-Oaxaca decomposition for linear regression models." *Stata Journal* 8 (4): 453–479.

Jones, F. L. 1983. "On Decomposing the Wage Gap: A Critical Comment on Blinder's Method." *The Journal of Human Resources* 18 (1): 126–130.

Kim, Hyuk-Rae, and Ingyu Oh. 2011. "Migration and Multicultural Contention in East Asia." *Journal of Ethnic and Migration Studies* 37 (10): 1563–1581.

Koenker, Roger, and Kevin F. Hallock. 2001. "Quantile Regression." *Journal of Economic Perspectives* 15 (4): 143–156.

Koenker, Roger W, and Gilbert Bassett. 1978. "Regression Quantiles." *Econometrica* 46 (1): 33–50.

Machado, Jos A. F., and Jos Mata. 2005. "Counterfactual decomposition of changes in wage distributions using quantile regression." *Journal of Applied Econometrics* 20 (4): 445–465.

Neal, Derek A., and William R. Johnson. 1996. "The Role of Premarket Factors in Black-White Wage Differences." *Journal of Political Economy* 104 (5): 869–895.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review* 14 (3): 693–709.

Schnepf, Sylke Viola. 2007. "Immigrants' educational disadvantage: an examination across ten countries and three surveys." *Journal of Population Economics* 20 (3): 527–545.

Tsay, Wen-Jen. 2006. "The Educational Attainment of Second-Generation Mainland Chinese Immigrants in Taiwan." *Journal of Population Economics* 19 (4): 749–767.

Villiger, Caroline, Christian Wandeler, and Alois Niggli. 2014. "Explaining differences in reading motivation between immigrant and native students: The role of parental involvement." *International Journal of Educational Research* 64: 12–25.