

Corazzini, Luca; Meschi, Elena; Pavese, Caterina

Working Paper

Impact of Early Childcare on Immigrant Children's Educational Performance

GLO Discussion Paper, No. 394

Provided in Cooperation with:

Global Labor Organization (GLO)

Suggested Citation: Corazzini, Luca; Meschi, Elena; Pavese, Caterina (2019) : Impact of Early Childcare on Immigrant Children's Educational Performance, GLO Discussion Paper, No. 394, Global Labor Organization (GLO), Essen

This Version is available at:

<https://hdl.handle.net/10419/202642>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Impact of Early Childcare on Immigrant Children's Educational Performance*

Luca Corazzini[†], Elena Meschi[‡] and Caterina Pavese[†]

[†]Department of Economics, Ca' Foscari University of Venice, 30121 Venice, Italy; Correspondences:
luca.corazzini@unive.it

[‡]Department of Economics, Management and Statistics DEMS, University of Milano-Bicocca, 20126 Milan, Italy

Abstract

This paper investigates the impact of attending early childcare on second generation immigrant children's cognitive outcomes. Our analysis draws on administrative data on the entire population of students in fifth grade collected by the Italian Institute for the Evaluation of the Educational System (INVALSI) for school years 2014/2015 to 2016/2017 matched to unique administrative records on the early childcare public available slots at the municipal level. Our identification strategy exploits cross-sectional and time series variation in the provision of early childcare service across Italian municipalities as an instrument for individual early childcare attendance. Our results point out that the effect of early childcare attendance differs between native and immigrant children. Although we find no effects for Italian children, our estimates show a positive and significant effect on literacy test scores for immigrant children of low educated mothers, which suggests that early childcare may be particularly relevant for immigrant children from a disadvantaged background.

Keywords: Childcare, Cognitive skills, Immigrant children, IV.

JEL Classification Numbers: J13, J15, H75, I20, I28.

*We would like to thank Orazio Attanasio, Emanuele Bracco, Giorgio Brunello, Elena Fumagalli, Irene Mammi, Paolo Pinotti, Lorenzo Rocco, Marianne Simonsen for useful comments and discussions. Special thanks go to Patrizia Falzetti and Michele Cardone (INVALSI) for giving us access to INVALSI microdata and to Veronica Grembi for kindly providing us data on childcare coverage at the municipal level. We also thank the conference participants at Società Italiana Economia Pubblica, Padova (2018) and Invalsi seminar, Bari (2018).

1 Introduction

In recent years, immigrant inflows to OECD countries increased substantially, leading to rising number of immigrant children in European school systems: more than one-quarter of the school-aged population in Europe will have a migrant background by the early 2020s, according to estimates of the Migration Policy Institute Europe (see [Ahad and Benton, 2018]).

Children who are migrants or have immigrant parents face significant educational barriers, such as lack of language proficiency or limited knowledge of the school system in host countries, which may lead to significant educational disadvantage relative to native students, as documented by the recent economic literature (see for example [Alesina et al., 2018] , [Dustmann et al., 2012]). Performance gaps at school translate in persistent differences in literacy levels, dropout rates and eventually in labour market outcomes between immigrant students and their native peers, thus perpetuating inequalities and slowing the integration of migrants in host countries. Therefore, addressing the needs of diverse learners and fostering the school performance of immigrant students are key challenges for European school systems.

Our paper aims at estimating the impact of early childcare on second-generation immigrant children’s cognitive outcomes. The focus on early education traces back to the findings of Heckman and co-authors suggesting that ability gaps between individuals and socioeconomic groups open up at early ages for both cognitive and non-cognitive skills. Given the high degree of brain plasticity, investments made early in life show both higher returns as well as dynamic complementarities with respect to investments made later in life ([Cunha and Heckman, 2007] and [Heckman and Masterov, 2007]). The latter was shown to be particularly true for children from disadvantaged backgrounds taking part to targeted interventions ([Blau and Currie, 2006]).

However, the evidence on the effects of untargeted publicly provided early childcare programmes is less conclusive, relatively scarce and not specifically focused on children with an immigrant background.

This paper makes use of novel administrative data on the entire population of several cohorts of primary school students in Italy to investigate the impact of attending early childcare on later cognitive outcomes. We exploit a rich dataset on children’s cognitive skills administered yearly to the entire population of students in second, fifth, eighth and tenth grade by the Italian Institute for the Evaluation of the Educational System (INVALSI henceforth) for the school years 2014/2015 to 2016/2017, merged with unique administrative records on early child care public available slots at the municipal level, collected by [Antonelli et al., 2011]. In the INVALSI data-set, pupils’ cognitive skills are measured by standardised and externally marked tests in literacy and numeracy. Together with test scores, INVALSI collects information on students’, families’ and schools’ characteristics.

We believe that the focus on Italy is particularly appropriate for our analysis: Italy only recently became a destination for international migration together with being one of the OECD countries with the highest educational gap between native and non-native students ([OECD, 2012]). The combination of these two features offers an invaluable setting for our analysis.

We contribute to the existing literature along several dimensions. First, while there is an extensive literature looking at interventions toward children aged 3-5 (preschool programmes), only few studies focus on children 0-2 (early childcare) and the evidence they provide is mixed. Some studies find positive effects ([Felfe and Lalive, 2018] and [Drange and Havnes, 2018] [Cornelissen et al., 2018], among others) while others find negative or no effects ([Herbst, 2017]; [Carta and Rizzica, 2018] and [Fort et al., 2019]). Second, we focus on students of immigrant background

and in particular on second-generation immigrants, a growing and relatively under-investigated group. Third, thanks to the rich set of information available in our dataset, we can perform several tests of heterogeneity and thus identify the channels at work, such as for example the role of linguistic distance of immigrants' native language from Italian.

The identification of the causal effect of early childcare attendance on immigrant children's cognitive and non-cognitive skills is challenging. It involves both the decision of parents to enrol children to the service as well as the choice of nurseries to accept those children who can potentially benefit the most by attending the service. The variables underlying these decision processes - the one of parents and the one of nurseries - are neither entirely observable nor measurable. Consequently, selection of children into early childcare is not random.

To deal with non-random sorting of children into early public childcare, we exploit the cross-sectional and time series exogenous variation in early childcare supply across Italian municipalities as an instrument for individual attendance.

Our results point out that the effect of early childcare attendance differs between native and immigrant children. While we find no effects for Italian children, our estimates show a positive and significant effect of early childcare attendance on INVALSI literacy test scores for immigrant children of low educated parents, which highlights the potential of these programs to improve outcomes for young children, particularly for those from disadvantaged families and reduce inequality in child development.

To further investigate the channels driving our results, we explore whether the effect differs within the group of immigrant children. In particular, we analyse the role played by the linguistic proximity of immigrant children's own language to the host-country one. To the best of our knowledge, this is the first time that such investigation is applied in the context of early education. Our findings highlight that children speaking a language that differs significantly from Italian capture larger benefits from attending early childcare. Overall, our results suggest that attending early childcare is an effective tool to foster immigrant children achievement helping closing existing gaps with respect to native children.

The structure of the paper is as follows. In Section 2 we provide a summary of the existing literature on early childcare and children outcomes together with underlining the contributions of our paper. In Section 3 we describe the Italian system of public early childcare. We then present our data and provide a descriptive analysis (Section 4). In Section 5 we illustrate our identification strategy and the threats to its validity running some relevant robustness checks (Section 7). We discuss our results and offer some conclusions (Sections 6 and 8).

2 Literature review

This paper contributes to the literature on the impact of early life experiences and early interventions on children's cognitive skills along several dimensions.

The literature distinguishes between interventions toward children aged 0-2 and interventions toward children aged 3-5. Interventions targeting the second of the two groups have been extensively examined in the literature where most of the studies find a positive impact of the service on children from disadvantaged background.¹

¹See [Havnes and Mogstad, 2015], [Felfe et al., 2014], and [Berlinski et al., 2009] among others.

On the contrary, there is lack of consensus on the effects of programmes involving children aged 0-2. As matter of fact, the few studies on the issue provide mixed evidence.

Desirable effects on cognitive and social skills from attending early care are stronger and robust in countries with high quality and large availability of the early childcare service, as in the northern European countries. By using administrative data from a large West German state, [Felfe and Lalive, 2018] develop a marginal treatment effect framework to assess the effect of early childcare on children’s school readiness. They find that early childcare is particularly beneficial for children with low educated mothers or foreign-born parents. [Drange and Havnes, 2018] provide evidence on the impact of early childcare enrolment on children educational attainments at age 7 in Norway. Their identification strategy exploits a randomized lottery used in the city of Oslo to allocate slots as an instrument for the age of the child, when he first attends childcare. Getting a lottery implies attending the service about four months earlier. Their results point out that children entered early in day care facilities perform better on both the literacy and numeracy test when 7 years old, with the effect driven by children from economically disadvantaged families.

However, the evidence of desirable effects is hardly equivocal. Evidence of negative effects emerged in many seminal studies evaluating the universal early childcare expansion occurred in Quebec during the ’90. In this group stands out the work of [Baker et al., 2008] that, exploiting a difference-in-difference strategy to compare children’s outcomes before and after the reform, find that childcare utilization is detrimental for children’s behavioural and health outcomes after the policy change. These negative outcomes are confirmed in a subsequent study of the same authors that also find little impact on cognitive test score ([Baker et al., 2015]). Similar results have been found by other studies, such as [Herbst, 2017] [Havnes and Mogstad, 2015] and [Herbst, 2013] for children from more advantaged families.

With reference to Italy, only recently early childhood interventions received some research attention, probably because of the greater availability of reliable data on children’s educational outcomes.

Mixed evidence on the impact of early childcare on children’s development emerged also in the literature focusing on Italy. [Brilli et al., 2016] find a positive impact of early childcare availability on children’s language cognitive outcome with the effect driven by areas with a higher level of rationing. Same conclusions are reached by [Del Boca et al., 2016] who, implementing an instrumental variable approach, provide evidence that early childcare attendance has a positive effect on children later educational achievements. [Carta and Rizzica, 2018] exploiting a reform introducing the possibility to enter earlier the fully-subsidized preschool service, outline that the entering in preschool at younger age has no significant effect on children cognitive outcomes at age 7. On the contrary, implementing a regression discontinuity design, [Fort et al., 2019] find that early childcare has a detrimental effect on cognitive and non-cognitive outcomes of girls coming from more affluent families.

Our focus on immigrant children is the second contribution to the existing literature.

Despite of the growing share in all OECD countries, very few studies investigated the impact of early childhood programmes on children with an immigrant background. In this group, [Dustmann et al., 2013] evaluate the impact of a universal childcare programme in Germany using unique administrative data and exploiting the staggered implementation of a federal policy reform across municipalities that entitles every child to a slot when turned 3 until school entry at age 6. They find that early childcare has a positive impact on children with immigrant ancestry, reducing their language and motor skills problems, while no significant effect emerges for native children. In the same vein, the study of [Drange and Telle, 2015] takes advantage of

a targeted intervention providing 4 hours of free childcare in some district of Oslo to estimate the impact of early childcare on children’s enrolment. Their results suggest that the intervention succeeded in increasing childcare enrolment and improved immigrant children’s cognitive development, especially if coming from a disadvantaged background.

Our paper contributes extending the scarce literature on the impact of early childhood programmes and focusing on immigrant children, a fast-growing and little investigated group. Further, given the uniqueness of our data, we are able to analyse several channels through which the effect of the day-care service may operate such as mother’s education, child’s gender and immigrant children’s linguistic proximity to the host-country language. To the best of our knowledge, no other studies investigated the role of the linguistic distance in shaping immigrant children’s early education experience.

3 Background

3.1 Immigrant children in Italy

Over the past three decades, Italy became target of a massive inflow of immigrants. In 2017, the share of immigrants represent 10% of the total population and their number increased by more than 30% between 2009 and 2017, going from 4.5 million to 5.9 million. Most immigrants come from European countries and main countries of origin are Romania, Albania, Morocco, China and Ukraine.² They are typically characterized by low educational levels and are mostly concentrated at the bottom of the native wage distribution. Relative to Italians, they tend to have a lower socio-economic and cultural background.³

The surge in migration has deeply affected the Italian educational system. In the academic year 2016/2017 the number of immigrant children represents 10.8% of students in Italian primary school.⁴ Figure 1 shows the number of immigrant children in primary Italian school from the school year 2010/11 to 2016/2017 and it highlights the continual and rapid growth of non-native students.

²Istat, "Demografia in cifre" (www.demografia.istat.it).

³See [[Frattini and Vigezzi, 2018](#)].

⁴Miur, "Portale dei dati sulla scuola" (www.dati.istruzione.it).

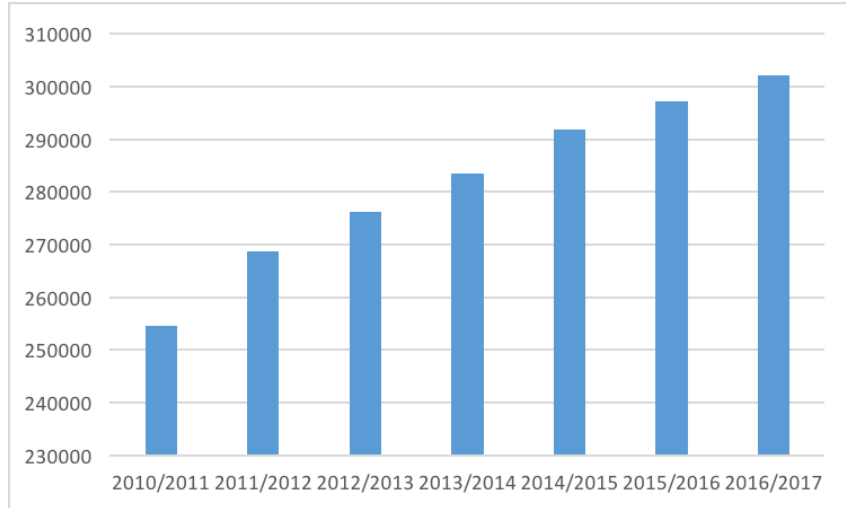


Figure 1: Number of immigrant children in Italian primary schools for the academic years 2010/2011 - 2016/2017.
Source: elaborazione su dati MIUR - Ufficio Statistica e studi

One of the most relevant aspects in the evolution of this phenomenon is the stable growth of second generation immigrant students. In the five academic years 2012/13-2016/2017 this group of students rose from 371.000 to 503.000 units with an increase of 35.4% across all the school levels and represent 61% of the total number of children without Italian citizenship.

In terms of the geographical distribution, Veneto, Umbria, Piemonte and Lombardia are the regions that host the higher number of second generation immigrant students out of the total number of students without Italian citizenship in elementary schools (between 57 and 61%).⁵ In Figure 2 we show second generation immigrant students geographical distribution at the municipal level for the 2016/17 academic year. Looking at the figure, it is apparent that most of second generation immigrant students are concentrated in Northern and Central regions.

⁵See [MIUR, 2018].

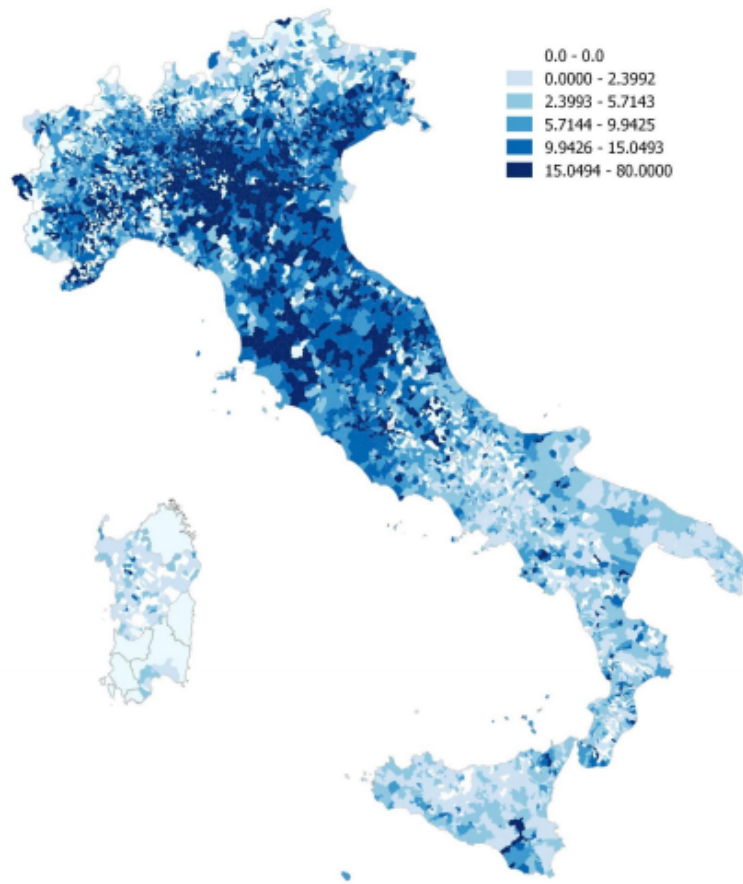


Figure 2: Geographic distribution of immigrant students in Italian primary schools for the academic year 2016/2017 at the municipal level (percentage values on the total number of students). Source: elaborazione su dati MIUR - Ufficio Statistica e studi.

3.2 Early Childcare in Italy

Formal education in Italy is compulsory from age six, when children start primary school, until age sixteen. For children younger than six, the provision of childcare is split into two stages: early childcare (nurseries) that is offered for children between three months and three years and preschool that is available for children aged three to six. There is no legal entitlement to childcare provision and while the access to the preschool service is almost universal and provided free of charge, early childcare is highly rationed and exhibits substantial cost variability across all Italian regions. The early childcare service is decentralized at the municipal level, provided by both public and private bodies and, differently from preschool and formal schooling, is not under the responsibility of the Italian Ministry of Education. Municipalities handle the direct provision of the service, and depending on the available budget and their (social and political) preferences, decide the number of slots to offer and the eligibility requirements used to allocate such slots.⁶

⁶There is high degree of variability in the criteria established by each municipality to assign the available slots. Besides the absolute priority given to children with disabilities, some municipalities benefit more families with disadvantaged socio-economic conditions, others parents' employment status (e.g. whether one or both parents are employed or not) and others family's composition (whether the child is orphan or has siblings).

The other tiers of government are responsible for more general issues. Regions establish the criteria for construction standards, management and appropriateness of personnel qualifications. The central government is only responsible for the definition of common standards and resources allocation among regions. On average, children spend 30.6 hours per week in early childcare and concerning the educational staff, the minimum qualification requirement is an upper secondary school diploma or a master degree in pedagogy.

Early childcare supply is very heterogeneous across regions and municipalities and accommodates 22.8 % of the potential demand (children under age 3) with relevant differences between northern and southern regions. While public and private childcare facilities in the North cover about 30% of 0–2 years old children, in the South the same figure drops to 10% (Figure 3).⁷

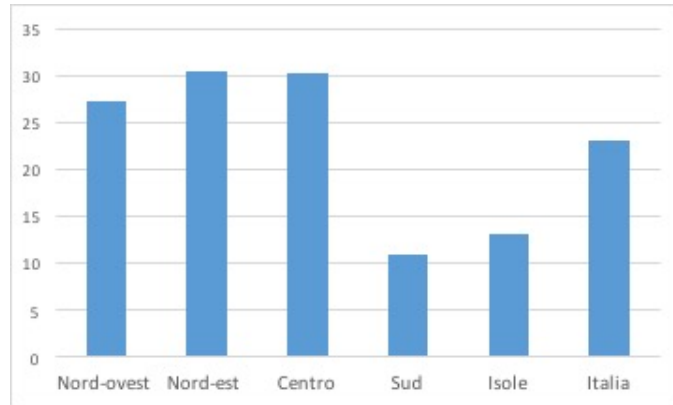


Figure 3: Early childcare coverage, year 2015. Notes: the coverage rate is provided by ISTAT as the number of early childcare authorized spots divided by the number of children 0-2

4 Data

4.1 INVALSI data

Our analysis draws on administrative data from INVALSI, the national agency that carries out a yearly testing of students' attainment in literacy and numeracy. Standardized tests are administered every year to the entire population of students in second, fifth, eighth and tenth grade. We focus our analysis on students in the fifth grade (last year of primary school) from the school years 2014/2015 to 2016/2017, the most recent ones.

The tests administered by INVALSI are standardized and are compulsory for all Italian schools and students attending the grades of interest. As all standardized tests, they have to satisfy two conditions. First, all students are required to answer the same questions during the same time interval. Second, correction and grading of the test follow standardized procedures that are set a priori and independently from who makes the actual correction of the tests. In other words, tests are anonymously and externally marked. These features of the INVALSI tests make students' results fully comparable across Italian schools and guarantee the tests to be objective.

All students take a test in Literacy and Numeracy in their own classroom on two different dates, typically in late May. The literacy test aims at assessing students' proficiency in reading comprehension that involves grammatical, lexical and pragmatic competences. The numeracy test

⁷ISTAT 2015, The municipal early child care supply and other early childhood services.

evaluates students' knowledge and mastering of mathematics specific contents such as geometry, probability and algebra. Both tests consist of multiple-choice and open-ended questions. Test scores are adjusted by a cheating factor directly provided by INVALSI taking into account the probability that student's class was involved in cheating behaviours during the test.⁸

INVALSI also provides marks assigned by teachers in math and reading at the end of the first term. Marks are distinguished in written and oral marks. We regard oral marks to be particularly interesting. Differently from INVALSI test scores, they are non-blind marks that reasonably take into account more than just cognitive skills of students since teachers' marks are relative and likely to be affected by the class and students' behaviour.

Together with test scores, INVALSI provides detailed information on families', students' and schools' characteristics. These additional information is collected by means of a "Family Questionnaire" sent to each family before the test, a "Student Questionnaire" compiled by each student in the fifth and tenth grade on the same day of the literacy or numeracy test, and a number of general information concerning students' performances and characteristics provided by the school administrative staff.

In our analysis, we use information on a rich set of individual and family characteristics such as parents' education and working status and students' gender, month of birth, citizenship (native and second-generation immigrant children), attendance to pre-school and a synthetic index of Economic, Social and Cultural Status (ESCS index).⁹ We also add a set of school and class characteristics such as the average size and their squares, the school average ESCS index, the share of female per class, the fraction of immigrant and retained students per class and the proportion of poorly educated parents per class. The data allow us to distinguish between Italian and non-Italian students. In Italy, this distinction is made according to the *Ius sanguinis* principle by which citizenship is determined by having one or both parents who are Italian citizens and not by place of birth. Concerning immigrant students, we can distinguish between first generation immigrant children, foreign-born students whose parents were also foreign-born and, second-generation immigrant children, students born in Italy whose parent were foreign-born. This study focuses on second generation immigrant children, only. Available data do not allow us to properly take into account first generation immigrant children age at arrival and therefore, to assess whether or not, they had the chance to attend early childcare in Italy.

4.2 Data on early childcare provision

We match INVALSI data with unique administrative records on the early childcare public available slots at the municipal level, collected by [Antonelli et al., 2011] [Antonelli and Grembi, 2014]. Administrative records were collected using municipal final balances provided by the Ministry of Interior. They refer to the availability of all the early childcare public spots directly and indirectly provided by Italian municipalities for the years 2005, 2006, 2007. Information is available for more than 700 hundred Italian municipalities with the number of residents above

⁸For further details see [Quintano et al., 2009]

⁹The ESCS index is built applying principal component analysis using the information provided by students in the Questionnaire and by schools about families' cultural resources, such as internet connection and the number of books, and parents' educational level and working status. By construction, the index has null mean and unit standard deviation. It implies that a student with a strictly positive individual value of the ESCS index has a socio-economic-cultural background more favourable than the Italian average. The index is provided at the individual, class and school level.

the threshold of 10,000.¹⁰ This data set allows us to assign to each student observed in the three cohorts, the actual childcare provision available in his municipality when he was 1 year old.

As we will discuss in next section, we use this detailed information on early childcare supply as an instrument for the children's attendance rate to the early childcare service.

Figure 4 shows our instrument variability across Italian regions and over time and points out that there is substantial variability between regions, with the lowest supply in Calabria and Campania and the highest in Emilia-Romagna and Provincia Autonoma di Trento. Such territorial differences have different trends in time. In fact, while some regions between 2005 and 2007 increased their supply of early childcare spots others shorten it. Molise is the most striking example, with a decrease in the supply of about 30%. The supply of childcare is heterogeneous not only across regions but also within regions across different provinces and municipalities. Table 1 shows some descriptive statistics of the supply of early childcare spots by regions in 2007. Noticeably, apart from Valle d' Aosta¹¹, the standard deviation relative to the mean shows that there is substantial variation within most of the regions in our analysis.

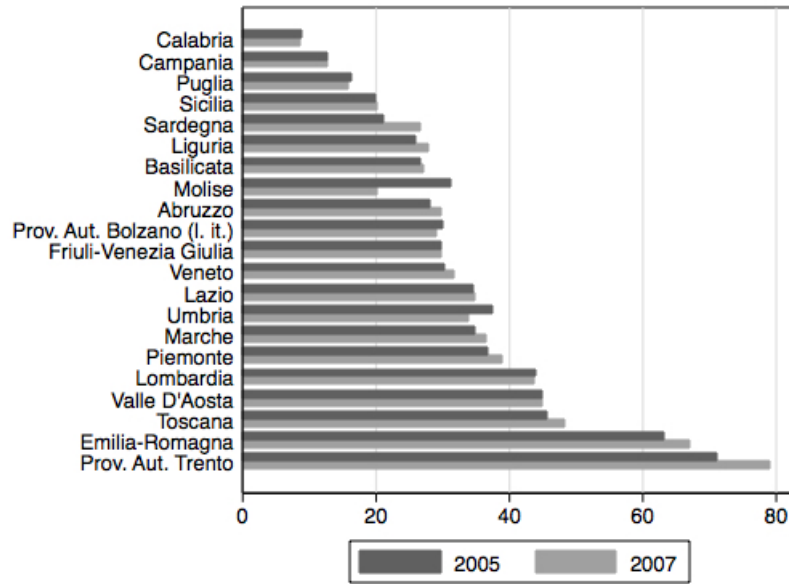


Figure 4: Variability of public childcare provision across regions - 2005 and 2007. The instrument is defined as the number of municipal early childcare public spots normalized by 10,000 residents.

¹⁰In the original data set, for municipalities under the 10,000 threshold was not possible to proceed at the integration of some relevant pieces of information.

¹¹Aosta is the only city in Valle d'Aosta having a number of residents higher than 10,000, thus the only one included in our analysis.

Table 1: Descriptive evidence of the supply of early child care spots by regions in 2007.

Regions	N	Mean	Sd	Min	Max
	(1)	(2)	(3)	(4)	(5)
Abruzzo	3,728	29.98	22.96	7.823	141.1
Basilicata	1,166	27.15	5.279	16.64	33.24
Calabria	3,278	8.650	4.130	2.659	14.26
Campania	8,035	12.85	5.885	4.6	39.28
Emilia-Romagna	17,185	67.13	20.42	15.13	119.1
Friuli-Venezia Giulia	2,869	29.98	6.384	19.67	38.32
Lazio	13,946	34.99	11.05	6.540	48.31
Liguria	4,207	27.98	8.246	15.77	46.60
Lombardia	25,447	43.83	15.83	13.57	74.12
Marche	3,819	36.64	14.55	14.75	62.10
Molise	498	20.35	4.467	13.78	23.38
Piemonte	12,555	39.00	10.14	16.29	61.09
Prov. Aut. Bolzano	266	28.95	5.561	24.41	36.52
Prov. Aut. Trento	450	79.06	1.653	64.35	91.91
Puglia	7,278	15.87	9.574	3.258	46.42
Sardegna	2,148	27.02	13.76	12.40	74.53
Sicilia	8,770	20.29	12.17	2.992	65.85
Toscana	9,406	48.54	18.26	9.923	94.04
Umbria	3,479	33.97	12.20	15.60	75.19
Valle d'Aosta	268	44.92	0	44.92	44.92
Veneto	11,801	31.88	13.07	11.71	67.25

Overall, our final data set covers 391,098 students (361,236 natives and 29,862 immigrants) 702 municipalities and 30865 schools located in the 20 Italian regions. For each academic year in the analysis, the share of immigrant population residing in these municipalities is about 60 % of the total immigrant population residing in Italy.¹²

Table 2 summarises all variables included in our empirical analysis.

¹²This percentage is calculated merging our data with yearly data on the number of immigrant residents per municipality <http://dati.istat.it/>

Table 2: Definition of variables and sources.

Variable	Description	Source	
Outcome variables			
Reading test score	Percentage of correct answers in Language test	INVALSI	
Math test score	Percentage of correct answers in Math test	INVALSI	
Reading oral grade	Grade got past academic year	INVALSI	
Math oral grade	Grade got past academic year	INVALSI	
Childcare variable			
Childcare coverage	Public childcare slots over pop 0-2 by municipality	[Antonelli et al., 2011]	
Endogenous variables			
Childcare attendance	dummy=1 if the child attended to early childcare		
Control variables (individual level)			
Female	dummy=1 if female		
Immigrant I generation	dummy=1 if child is Immigrant II generation		
Immigrant II generation	dummy=1 if child is Immigrant I generation		
Age	child year of birth		
II quarter month of birth	dummy=1 if the child is born in second quarter		
III quarter month of borth	dummy=1 if the child is born in third quarter		
Low level of paternal/maternal education	dummy=1 if mother/father has an educ. level lower or equal to middle school	INVALSI	
Medium level of paternal/maternal educ.	dummy=1 if mother/father has an educ. level higher than middle school or equal to high school		
High level of paternal/maternal educ.	dummy=1 if mother/father has an educ. level higher than high school		
Mother's working status	dummy=1 if the mother works		
Father's working status	dummy=1 if the father works		
Language mostly spoken at home	dummy=1 if the child mostly speaks a language different from italian at home		
Pre-school attendance	dummy=1 if the child attend the pre-school service		
Students average characteristics			
Share of female	Share of female in each fifth grade class		
Late students	Share of retained students in each class		
School and class characteristics			
Immigrant students	Share of immigrant students per class	INVALSI	
School size	Total number of students in fifth grade per school	INVALSI	
School size sq.	School size squared	INVALSI	
Average class size	Average number of students in each fifth grade cass	INVALSI	
Average class size sq.	Average class size squared	INVALSI	
Low educated parents share	Share of low educated parents per class	INVALSI	
ESCS School	Average ESCS per school	INVALSI	

4.3 Descriptive evidence

As a preliminary descriptive analysis, in Figure 5 and 6 we provide the kernel density estimation of the probability density function of the literacy and numeracy test score, respectively. Each figure reports the distribution of standardised test scores in literacy (Figure 5) and numeracy (Figure 6) by immigrant status and early childcare attendance. The figures highlight the substantial achievement gap between native and immigrant students. In fact, the distribution for Italian children is left-skewed and has a higher mode compared to the distribution of immigrants, which in turn appears rather symmetric and with a higher variance. These differences are more pronounced in the literacy test score distribution. For both natives and immigrant students, there seems to be a positive difference between the performances of children attending the early childcare and not attending it, which is a first signal of children positive selection into the service.

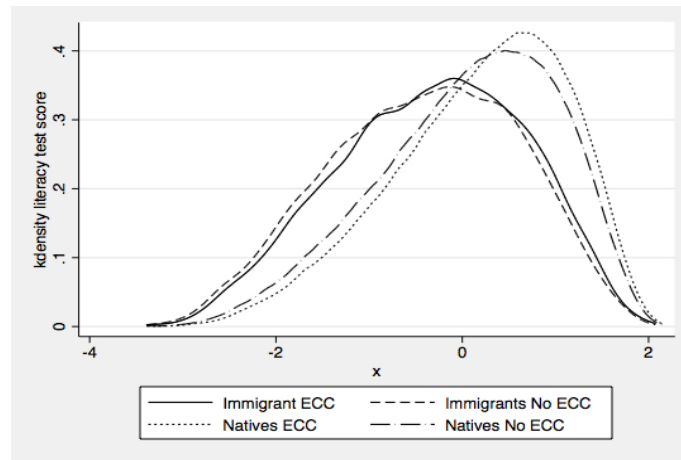


Figure 5: Kernel density literacy test score by early childcare attendance.
Source: INVALSI

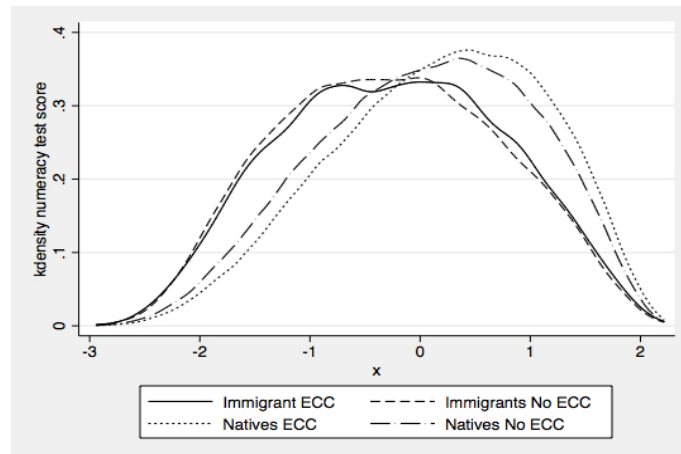


Figure 6: Kernel density numeracy test score by early childcare attendance.
Source: INVALSI.

Table 3 presents descriptive statistics of the individual and mean characteristics of children in our sample. Columns (2)-(5) refer to native children, while columns (6)-(9) refer to immigrant ones. Columns (2) and (6) report values for children who attended the early childcare service (ECC), the treated group, while columns (3) and (7) refer to children who did not attend the

service (No ECC), the control group.

For both native and immigrant children, we find significant differences in all the observable characteristics included in our analysis (Columns (3) and (6)), which is a preliminary signal of the endogeneity of the process behind the attendance to the daycare service.

On average, children enrolled in early childcare come from a more advantaged background than children not enrolled, independently from nationality. They are more likely to have a working mother, parents with a high educational level and higher values for the ESCS index. In terms of observable individual and parental background characteristics, descriptive evidence suggests children's positive selection into early childcare. This positive selection seems to be less pronounced for immigrant children than for native ones.

Table 3: Descriptive statistics and balancing test.

Variable	Natives			Immigrants		
	(1)	(2)	(3)	(4)	(5)	(6)
	ECC	No ECC	p-value	ECC	No ECC	p-value
Background characteristics						
Female	0.483	0.502	0.000	0.495	0.505	0.066
Father compulsory education	0.235	0.341	0.000	0.336	0.429	0.000
Father high school	0.478	0.462	0.000	0.485	0.443	0.000
Father higher education	0.285	0.195	0.000	0.177	0.127	0.000
Mother compulsory education	0.146	0.258	0.000	0.34	0.441	0.000
Mother high school	0.486	0.506	0.000	0.451	0.408	0.000
Mother higher education	0.366	0.235	0.000	0.208	0.15	0.000
Father employed	0.958	0.944	0.000	0.876	0.884	0.027
Mother employed	0.818	0.632	0.000	0.528	0.361	0.000
Preschool	0.98	0.945	0.000	0.976	0.914	0.000
Late students	0.005	0.01	0.000	0.038	0.056	0.000
ESCS students	0.548	0.232	0.000	-0.271	-0.44	0.000
Outcome variables						
Literacy test score	0.193	0.072	0.000	-0.33	-0.411	0.000
Numeracy test score	0.17	0.043	0.000	-0.24	-0.28	0.000
Language oral mark	0.22	0.11	0.000	-0.423	-0.485	0.000
Math oral mark	0.212	0.086	0.000	-0.309	-0.356	0.000
Observations	203,291	157,945		12,947	16,915	

Notes: the table reports means of covariates by nationality and by whether the child attended the early child care service (ECC). Column (3) and (6) report the p-value of the t-statistic for equality of means in the treated and untreated group.

5 Identification strategy

We start our analysis by estimating the following linear model that links child's cognitive outcomes (Y_{imt}^j) to her early childcare attendance ($Earlychildcare_{imt}$) while controlling for characteristics of the child, the family, the class, the school as well as provincial fixed effects.

$$Y_{imt}^j = \gamma_0 + \gamma_1 Earlychildcare_{imt} + \gamma_2 YEAR_t + \gamma_3 PROV_{im} + \mathbf{X}'_{imt}\gamma + \epsilon_{imt} \quad (1)$$

where i indexes individuals, m indexes municipalities, t indexes cohorts, and j literacy/numeracy test score/oral mark), *Earlychildcare* is the variable indicating whether or not the child attended the childcare service, *YEAR* and *PROV* are year and province dummies, \mathbf{X} is a vector of baseline controls and ϵ_{imt} is the error term. In particular, we control for preschool attendance, gender, quarter of birth, studies regularity, mother's and father's education and employment condition, share of female and late students per class, school size and its square, average class size and its square, share of migrants per class, low educated parents' share and ESCS per school. The choice of control variables is motivated by the literature related to children's human capital production function. ([Cunha and Heckman, 2007], [Todd and Wolpin, 2003]).

We consider four dependent variables: literacy and numeracy standardized test scores, language and math oral marks assigned by teachers. Test scores are defined as the fraction of correct answers and they range from 0 to 100, while marks assigned by teachers range from 0 to 10, where 6 is the passing grade.¹³ To increase comparability across cohorts we standardized literacy and numeracy test scores as well as oral marks to have mean equal to 0 and standard deviation equal to 1 in each cohort and subject. All the specifications are run separately for Italian and second-generation immigrant students.¹⁴ The parameter of interest is γ_1 that captures the impact of early childcare attendance.

Clearly, the problem with estimating this linear regression by OLS is the non-random sorting of children into early childcare. In fact, parents voluntarily decide to enrol children to the service based on their observable and unobservable characteristics and preferences. Therefore, comparing outcomes of children that attended and did not attend the service would result in biased estimates, even conditionally on a rich set of control variables.

The literature addressed this endogeneity issue using different strategies. The most common methodologies rely on difference in differences approach exploiting reforms in the provision of childcare (see for instance [Baker et al., 2008], [Havnes and Mogstad, 2011], and [Felfe et al., 2014]), regression discontinuity design around the admission threshold [Fort et al., 2019] or an instrumental variable method ([Del Boca et al., 2016], [Felfe and Huber, 2017], and [Drange and Havnes, 2018]). Given the absence of a major reform affecting the early childcare service in Italy, we implement an instrumental variable strategy. In particular, we exploit cross-sectional and longitudinal variation in the early childcare supply across Italian municipalities as an instrument for children's probability to attend the nursery service. In a context where childcare provision is highly rationed, childcare availability clearly affects the probability of individual childcare attendance, and is arguably exogenous to children's characteristics and to unobservables in the model (we will discuss this point in more detail in Section 5.1). This approach is consistent with the evidence of [Del Boca and Vuri, 2007] showing that attendance to early childcare in Italy is mostly driven by the supply - rather than the demand- for the service. However, our instrument does not affect parents that are not interested in enrolling their children to the early childcare service.

Our instrumental variable is defined as the number of early childcare publicly provided spots normalized for 10,000 residents. Equation (2) describes the first stage of our IV-model:

$$Earlychildcare_{imt} = \beta_0 + \beta_1 COVER_{mt} + \beta_2 YEAR_t + \beta_3 PROV_{im} + \mathbf{X}_{imt}\beta + \omega_{imt} \quad (2)$$

¹³Due to the high number of missing observations, we are not able to use teachers' written marks.

¹⁴In order to test the hypothesis that regressors slopes are different across native and immigrant students, we run a Chow test that strongly rejected the data poolability assumption suggesting that in our framework separate regressions provide a better fit.

where $COVER_{mt}$ is the public childcare coverage at the municipal level at time t and ω_{imt} is the error term. Standard errors of both the first and second stage are clustered at the municipality level.

We contribute to the existing IV literature in this field, by adopting a tighter identification strategy that exploits variation in the instrument not only across local areas (the main variation used in previous studies) but also across cohorts, thus enabling us to control for time constant unobserved area characteristics. Moreover, our instrument is defined on a much finer geographical aggregation compared to existing studies in the literature (see for instance [Del Boca et al., 2016] that instrumented early childcare with the regional supply of childcare).

In Figure 7 we report a graphical representation of the first stage that illustrates the relationship between the supply of early childcare spots and the average early childcare attendance by municipality. It emerges clearly the existence of a positive linear relationship between the two variables which suggests that our instrument is valid. In our case, it implies that there are at least some parents that are affected in their enrolment decision by the public early childcare supply of the municipality. We will further check the validity of our IV, by examining the first stage of the instrument to be significantly different from zero. In section 6, tables of results provide the first stage instrument coefficient estimates, as well as the Kleibergen-Paap F-statistics. Overall, our instrument shows to capture relevant variation in the early childcare attendance, the endogenous variable.

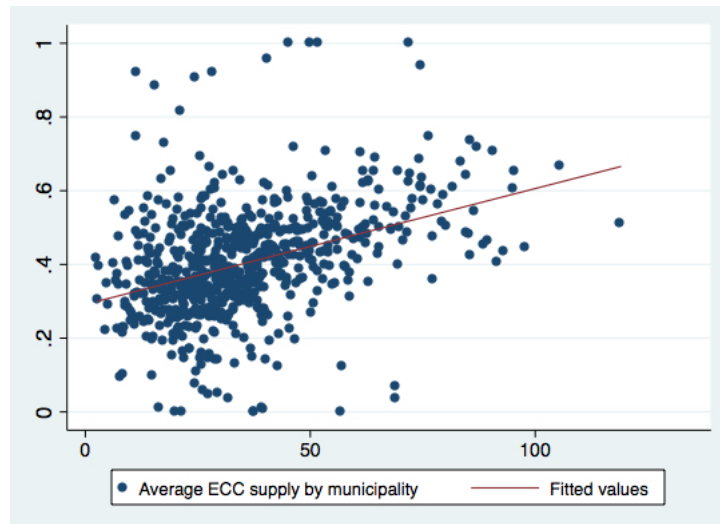


Figure 7: First stage graphical representation. Each dot shows the early childcare supply (ECC) by municipality.

5.1 Threats to identification

Our estimates identify the causal effect of early childcare attendance on children’s cognitive outcomes if the supply of early childcare spots at the municipal level is uncorrelated with the unobserved child’s characteristics conditional on a rich set of observable covariates ($\mathbf{X}_{imt}\gamma$), provincial ($\gamma_3 PROV_{im}$) and cohort fixed effects ($\gamma_2 YEAR_t$). In other words, public supply of early childcare spots must not have a direct effect on cognitive children’s outcome and should be uncorrelated with any characteristics determining children’s early childcare attendance and/or cognitive outcomes. While it seems plausible to assume the public supply of early childcare spots not to have a direct effect on cognitive children’s outcome, a major concern is the existence of

an association between the instrument and some potential determinants of early childcare attendance. In particular, there are three key reasons that might lead the exclusion restriction not to hold. First, municipalities might differ along several dimensions other than the supply of early childcare spots. Second, availability of early childcare and the quality of its provision to immigrant children might be negatively affected by an active policy of local politicians to hinder immigrant families to reside in their municipality. Third, immigrant families might choose to settle down in a municipality where the supply of early childcare spots is higher. We address in turn each point.

Regarding the first concern, we use provincial fixed effect in order to remove systematic differences between provinces in time. It is noteworthy to underline that we look at municipalities with a number of residents greater than 10,000 in the same province.¹⁵ We expect municipalities with a number of residents greater than 10,000 in the same province to be characterized by a high level of internal homogeneity both in terms of social and economic characteristics. As showed in Figures 1 and 4, our instrument captures between municipalities and over time variation but the latter is not enough to allow us to use municipality fixed effects. Indeed, the variation we exploit is not generated by a sharp policy change or a reform able to produce sizeable changes over time. Thus, a threat to our identification may still derive from the existence of characteristics that correlate with both the outcome variable and the instrument that are not captured by the province fixed effects. In particular, one may argue that more resourceful municipalities supply a higher number of early childcare spots together with providing more opportunities for children's cognitive development. We deal with this potential threat by including in our specification an additional variable that aim to capture possible remaining heterogeneity between municipalities within the same provinces, which may affect the availability of childcare spots and may also be correlated with unobservable factors in the model that may influence children cognitive outcomes. In particular, as a robustness check we include a time varying measure of taxable income per capita at the municipal level for the years 2005-2007.¹⁶

Regarding the second concern of local anti-immigrant active policies, during the last two decades right parties in Italy, especially in the North, strongly supported anti-migration positions while left parties where in favour of more inclusive policies.¹⁷ One may argue that different attitudes toward immigrants could have had spill-over effects on many social and economic aspects of immigrant families residing in the municipality. In our context different orientations toward immigrants might translate in different school environments as well as in more (less) inclusive policies for immigrant children during early childhood. Controlling in our specification for the mayor political faction, is intended to capture these potentially relevant differences across municipalities.

Regarding the third concern, one may inquire whether immigrant families decided to settle down in a specific municipality because of the greater availability of early childcare spots. The main drivers in immigrant families' residence decision are employment opportunities while we regard the availability of early childcare spots not to play a crucial role in their decision. In order to substantiate these points, in Section 7 we run a battery of robustness checks. Results are presented in Table 12 and 13.

Under our identifying assumptions, our approach identifies the **causal** effect of early childcare attendance on the child's cognitive outcomes.

¹⁵Provinces are intermediate administrative divisions between municipalities and regions. The number of Provinces is equal to 107.

¹⁶Data are drawn from the Ministry of Economy and Finance "Analisi Statistiche - Open data dichiarazioni" www.finanze.gov.it

¹⁷For further details see [Bracco et al., 2018]

6 Results

This section reports our main results on the effect of early childcare attendance on children cognitive outcomes. All specifications are run separately for native and immigrant students.

6.1 The effect of Early childcare attendance on standardized test scores

We start with the analysis of INVALSI standardized test scores. Table 4 reports OLS and IV estimates for literacy and numeracy test scores. In all specifications we cluster standard errors at the municipal level and we control for province and cohort fixed effects as well as for all control variables discussed in section 5.

As illustrated in previous sections, OLS estimates might be biased. Early childcare enrolment is the result of a set of decisions made by parents and nurseries that are based on unobserved determinants. In other words, children attending early childcare is likely to be a selected group. Our IV estimates takes into account the endogeneity of early childcare attendance and exploits only the variation in the aggregate childcare attendance rate within municipalities and across cohorts. As a consequence, the positive sign of the OLS estimates in Panel A of Table 4 does not necessarily imply early childcare attendance having a positive impact. In particular, the difference between IV and OLS estimates outlines a positive selection into childcare, translating in OLS estimates to be upward biased. In fact, we expect more skilled parents to be more likely to be employed in high-paying jobs and, therefore, to have higher incentives to enrol their children into early childcare in order to keep the household's income constant. Further, children of more skilled parents are more likely to grow up in stimulating and favourable environments that lead them to perform better than children coming from more disadvantaged families, independently from attending or not the early childcare service. Evidence in Table 4 are consistent with these considerations: OLS estimates are positive and significant, while when turning to IV estimates they are no longer significant. The only exception is represented by the OLS estimates for immigrants' numeracy test score that are negative and significant. The direction of the selection in this case is less straightforward to interpret and could be caused by unobservable characteristics that we are not able to control for. On average, we find that, neither for native children, nor for those with an immigrant background, early childcare attendance has a significant impact on literacy and numeracy test scores when the child is 10.

For the sake of brevity, we did not include in our tables of results control variables' coefficients. Reassuringly, their sign is consistent with the evidence from previous studies underpinning the goodness of our model. In particular, having working and/or highly educated parents has a positive and significant impact on children's outcomes as well as attending preschool. Interestingly, also being born earlier (we control for children's quarter of birth) has a positive and significant effect on children's results.

In Panel B, we report first stage estimates, which show that public municipal supply of early childcare strongly determines children's attendance to the service, both for native and non-native students. First stage F-statistics indicate that the instrument is not weakly defined, implying that it is able to predict relevant variation of the endogenous variable.

Table 4: Effects of early childcare attendance on children’s test score at grade 5.

LITERACY					NUMERACY			
Panel A: Second Stage								
	Natives		Immigrants		Natives		Immigrants	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
Early childcare	0.011*** (0.004)	-0.081 (0.119)	-0.006 (0.012)	0.220 (0.205)	0.036*** (0.005)	-0.255 (0.160)	-0.047*** (0.011)	0.053 (0.196)
Observations	361,236	361,236	29,862	29,862	361,236	361,236	29,862	29,862
Panel B: First Stage								
	(2)		(4)		(6)		(8)	
	WS		WS		WS		WS	
Early childcare public supply	0.002*** (0.000)		0.003*** (0.000)		0.002*** (0.000)		0.003*** (0.000)	
First stage F-stat	41.8***		38.56***		42.19***		38.48***	
Observations	361,236		29,862		361,236		29,862	
Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Entries in the Tabel are the estimated coefficients of early childcare attendance on standardized literacy (columns 1-4) and numeracy (columns 5-8) scores. Columns 1,3,5,7 reports OLS estimates, while columns 2,4,6 and 8 IV estimates. Columns 1-2 and 5-6 report estimates for native students, while columns 3-4 and 7-8 for immigrant students. All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students’ average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors (adjusted for clustering at the municipal level) in parenthesis, ***p<0.01, **<0.05, *p<0.1.

6.2 Heterogeneous effects

Results reported in Table 4 are averaged on several individual characteristics. Thus, they could mask substantial heterogeneity in the effect of early childcare across different sub-groups of the population. We proceed by exploring whether attendance to early childcare is experienced differently across children from various backgrounds. In particular, we investigate whether the impact differs by children’s social and cultural background and by gender.

In Table 5 and 6, we report IV estimates of the impact of early childcare on INVALSI literacy and numeracy test scores for children with high or low educated parents ¹⁸, respectively for native and immigrant children. Parents’ educational level proxy children’ social and cultural background. Table 7 suggests that parental background does not play a role in shaping the effect of childcare programs on literacy scores of native students, but interestingly highlights that for immigrant students the impact of early childcare is indeed not homogenous across parental backgrounds. In fact, when we run separate regressions by parental education, we find that the impact of early childcare is positive and significant only for imigrant children from low educated families (see column 6) ¹⁹. These results are in line with previous studies underlying that childcare attendance is particularly beneficial for children from disadvantaged backgrounds, because they would probably experience lower-quality care in the absence of formal childcare, due to worse

¹⁸For each child we consider the highest level of education among the two parents and then we define a parent to be highly educated if she/he has an educational level equal or higher than high school diploma.

¹⁹Our results are consistent also when using father or mother educational level as a proxy for the child’s social and cultural background.

and less stimulating home environment (see [Havnes and Mogstad, 2011], [Drange and Havnes, 2018], [Felfe and Lalive, 2018], [Cornelissen et al., 2018]).

In terms of magnitude, early childcare utilization increases literacy test scores of immigrant children from low socioeconomic background by 72% of a standard deviation. To appreciate the size of the effect, consider that immigrants' literacy test scores are 0.25 standard deviations lower than natives'. Therefore, early access to preschool tend to reduce the achievement gap between immigrant and natives by about 30%.

Table 5: Heterogeneous effects by parents' educational level, literacy

LITERACY						
Panel A: Second Stage						
	(1) Whole Sample	Natives (2) High Edu	(3) Low Edu	(4) Whole Sample	Immigrants (5) High Edu	(6) Low Edu
Early childcare	-0.081 (0.119)	-0.011 (0.121)	-0.271 (0.203)	0.162 (0.201)	-0.060 (0.195)	0.727** (0.358)
Observations	361,236	286,690	74,546	30,896	18,036	11,826
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
First stage F-stat	42.12***	42.13***	38.63***	38.56***	33.20***	32.57***
Observations	361,236	286,690	74,546	30,896	18,036	11,826
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

In Table 6 we show estimates of the impact of early childcare attendance on numeracy test scores for children with low educated parents. Regardless of the immigrant status and parents' educational level we find no effect on numeracy test scores. Interestingly, the fact that results that we observe for literacy test scores are not mirrored for numeracy test scores, suggest that early childcare is likely to function as a device to facilitate children's learning of the host country language.

Table 6: Heterogeneous effects by parents' educational level, numeracy.

NUMERACY						
Panel A: Second Stage						
	(1)	Natives		(4)	Immigrants	
	Whole Sample	(2) High Edu	(3) Low Edu	Whole Sample	(5) High Edu	(6) Low Edu
Early child care	-0.255 (0.160)	-0.230 (0.160)	-0.394 (0.243)	0.052 (0.196)	-0.024 (0.207)	0.223 (0.319)
Observations	361,236	286,690	74,546	29,862	18,036	11,826
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early child care public supply	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
First stage F-stat	42.19***	42.13***	38.63***	38.48***	33.32***	32.60***
Observations	361,236	286,690	74,546	29,862	18,036	11,826
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the numeracy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1, 5 and 10 percent level.

Turning on the heterogeneity by gender, we investigate whether male and female students experience early childcare attendance differently. While on average we do not find significant gender differences both for literacy and numeracy test scores, gender plays an important role in explaining our positive and significant result for immigrant children with low educated parents in the literacy test score. In particular, we find that immigrant girls with low educated parents are those gaining the largest benefit from the childcare attendance. Results are reported in Table 7. This finding is consistent with those in [Felfe et al., 2014] and [Havnes and Mogstad, 2011] who report that early childcare attendance is particularly beneficial for girls. This finding is also in line with the psychological literature emphasizing girls greater capability of reaping the benefits from an early interaction with adult native speakers (see [Fenson et al., 1994] [Bornstein et al., 2004]).

Table 7: Heterogeneous effects by parents' educational level and child's gender.

LITERACY								
Panel A: Second Stage								
	Natives				Immigrants			
	(1) High Edu Male	(2) Low Edu Male	(3) High Edu Female	(4) Low Edu Female	(5) High Edu Male	(6) Low Edu Male	(7) High Edu Female	(8) Low Edu Female
Early childcare	0.092 (0.133)	-0.362 (0.309)	0.024 (0.144)	0.0151 (0.232)	-0.086 (0.276)	0.627 (0.587)	-0.040 (0.246)	1.059** (0.4448)
Observations	157,954	24,732	153,487	25,063	10,576	4,324	10,592	4,370
Panel B: First stage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)	0.003*** (0.001)
First stage F-stat	40.02***	25.91***	41.22***	36.73***	25.04***	13.40***	25.77***	33.32***
Observations	157,954	24,732	153,487	25,063	10,576	4,324	10,592	4,370
Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(4) refer to native students, while columns (5)-(8) refer to immigrant ones. Both native and immigrant children are distinguished by gender. Column (2),(4),(6) and (8) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal than middle school diploma), while column (1),(3),(5), and (7) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

On the whole, our results outline that the impact of the early childcare service varies by sub-groups with the largest benefits observed for immigrant children from more disadvantaged backgrounds. In particular, females coming from less resourceful families are those gaining the largest benefits. Further, in all the specifications analysed, we find that early childcare attendance has no impact on native children neither in literacy nor in numeracy test scores. This finding stands in contrast with part of the recent literature reporting detrimental effect for native children (see [Fort et al., 2019]).

6.3 The effect of Early Childcare attendance on oral marks assigned by teachers

We now focus on the impact of early childcare on oral marks assigned by teachers in language and math at the end of the first term of the same academic year in which the INVALSI tests take place. We repeat the same analysis performed for INVALSI test scores but considering oral marks assigned by teachers as dependent variables. We regard oral marks to be particularly interesting because they capture something different than just cognitive skills: they are non-blind scores and are likely to be affected by teachers' evaluations regarding students behaviours and thus should grasp a mixture between cognitive and non-cognitive skills such as students' motivation, resilience and effort.

In Table 8, we report IV estimates for language and math oral marks. Results show that the attendance to early childcare has a positive and significant effect both on immigrant children's

reading and math oral marks assigned by teachers.

Table 8: Effects of early childcare attendance on children oral marks at grade 5.

	READING ORAL MARK		MATH ORAL MARK	
Panel A: Second Stage				
	Natives (1)	Immigrants (2)	Natives (3)	Immigrants (4)
Early childcare	-0.126 (0.169)	0.492* (0.272)	-0.117 (0.166)	0.438* (0.245)
Observations	346,955	28,573	346,358	28,545
Panel B: First Stage				
Ecc public supply	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.000)	0.003*** (0.001)
First stage F-stat	42.39***	37.43***	42.71***	37.38***
Observations	346,955	28,573	346,358	28,545
Province FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the reading oral mark for columns (1)-(2) and the math oral mark for columns (3)-(4). Columns (1)-(3) refer to native students, while columns (2)-(4) refer to immigrant ones. All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

In the following tables we investigate the interplay between the effect of early childcare attendance on reading (math) oral marks and the educational level of the students' parents ²⁰. In Tables 9 and 10 we report results from subsamples divided by immigrant status (columns 1-3 native children and columns 4-6 immigrant children) and by high and low parents' education (columns 2 and 5 high educated parents and columns 3 and 6 low educated parents) respectively for reading and math oral marks. In both Tables columns 1 and 4 show baseline estimates for ease of comparison. Our results show that early childcare has a positive and significant impact on language and math oral marks of immigrant children with low educated parents while we find no effect on native children.

All in all, oral marks results are consistent with those in previous section analysing the impact of early childcare on INVALSI test scores, however, differently from before our results are significant for math oral marks as well. The latter it is likely not be related with children's higher ability in math but to their oral exposure competences that strongly depend on their language fluency and comprehension.

²⁰As in previous subsection, for each child we consider the highest level of education among the two parents then, we define a parent to be highly educated if she/he has an educational level equal or higher than high school diploma.

Table 9: Heterogeneous effects by parents' educational level, reading oral marks.

READING ORAL MARK						
Panel A: Second stage						
	(1) Whole Sample	Natives (2) High Edu	(3) Low Edu	Immigrants (4) Whole Sample	(5) High Edu	(6) Low Edu
Early childcare	-0.126 (0.169)	0.030 (0.165)	0.024 (0.223)	0.492* (0.272)	0.420 (0.293)	0.641* (0.345)
Observations	346,955	299,087	47,868	28,573	20,276	8,297
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.002*** (0.000)
First stage F-stat	42.39***	42.97***	35.67***	37.43***	30.93***	29.54***
Observations	346,955	299,087	47,868	28,573	20,276	8,297
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the reading oral mark assigned by teachers. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal than middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

Table 10: Heterogeneous effects by parents' educational level, math oral marks.

MATH ORAL MARK						
Panel A: Second stage						
	(1) Whole Sample	Natives (2) High Edu	(3) Low Edu	Immigrants (4) Whole Sample	(5) High Edu	(6) Low Edu
Early childcare	-0.117 (0.166)	0.016 (0.163)	0.066 (0.237)	0.438* (0.242)	0.224 (0.245)	0.958** (0.379)
Observations	346,358	298,600	47,758	28,545	20,255	8,290
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
First stage F-stat	42.71***	43.30***	35.97***	37.38***	31.07***	29.00***
Observations	346,358	298,600	47,758	28,545	20,255	8,290
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the reading oral mark assigned by teachers. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated mother (educational level of the parent with the highest level of education lower or equal than middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

As in previous sub-section, we investigate whether the positive effect found for immigrant children with low educated parents shows significant gender differences. For both males and females with low educated parents, we detect a positive and significant effect (at the 10 percent level) of early childcare attendance on children language oral mark, with the effect being not mediated by gender ²¹.

6.4 The role of language proximity

To explore a possible channel at work in determining the effect of early childcare on second generation immigrant children, we expand our investigation to another dimension that might help us explain our results on cognitive skills, namely the linguistic proximity of immigrant children's language to Italian [Frattini and Meschi, 2019]. Indeed, one may inquire about the impact of early childcare attendance being experienced equally by diverse groups of immigrant children, identified according to their linguistic proximity to Italian. We expect children whose native language differs significantly from Italian to benefit the most from the attendance to early childcare being in there highly exposed to the Italian language. To the best of our knowledge, no other studies on early child development focused on this relevant dimension.

In order to carry out this analysis we make use of a linguistic proximity index, INDEX, based on information from Ethnologue (see [Adsera and Pytlikova, 2015]). INVALSI data contain unique

²¹Our results are consistent also when using father or mother educational level as a proxy for the child's social and cultural background.

information about the language children mostly spoken at home ²². We restrict the sample to second generation immigrant children and we match the individual information about the language children mostly speak at home with Ethnologue data. In our data, the index ranges between 0 and 1. It is set equal to 0 if the two languages do not share a common language family and equal to 1 if they do share one. Thus, the higher is the number of linguistic family tree's branches that two language share the higher is the index. In our sample the index ranges between 0 and 0.7 and has four unique values.

While on average we find no effect on children speaking a language more (less) distant from Italian, we report significant heterogeneity for children with low educated parents speaking a language distant from Italian. Result are reported in Table 11. We run separate regressions for second generation immigrant students whose language is more (less) distant from Italian on the basis of INDEX being above (below) the median and with low (high) educated parents. As in previous section we focus on our main outcome variables: INVALSI literacy and numeracy test scores.

As expected, our results confirm no effect for children speaking a language with low dissimilarity with respect to Italian both when they have high or low educated parents. However, consistently with our results in previous sections, we find that early childcare attendance has a positive and significant impact on literacy test scores for children speaking a language distant from Italian. We find no effect on numeracy test scores.

Table 11: Heterogeneous effects by child's linguistic proximity and parents' educational level.

	Low linguistic proximity (LLP)		High linguistic proximity (HLP)	
Panel A: Second Stage	(1)	(2)	(3)	(4)
Early childcare	0.330 (0.406)	2.410* (1.463)	-0.200 (0.274)	-0.472 (0.571)
Observations	5,556	3,779	3,305	657
Panel B: First Stage	(1)	(2)	(3)	(4)
Early childcare supply	0.003*** (0.001)	0.002** (0.001)	0.005*** (0.001)	0.005*** (0.002)
First stage F-stat	15.00***	5.00**	27.17***	8.35**
Observations	5,556	3,779	3,305	657
Province FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
Class controls	✓	✓	✓	✓

Notes: Columns (1)-(2) refer to immigrant students speaking a language with a low level of proximity with respect to Italian (LLP). Columns (3)-(4) refer to immigrant student speaking a language with high level of linguistic proximity with respect to Italian (HLP). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

In Figure 8 we plot the marginal effect of attending early childcare on literacy test scores by linguistic proximity at the 90% significant level. It emerges clearly that the effect is positive and significant with low levels of linguistic proximity and decreases, turning not significant, as the two languages become more similar.

²²INVALSI data do not provide immigrant children's country of origin.

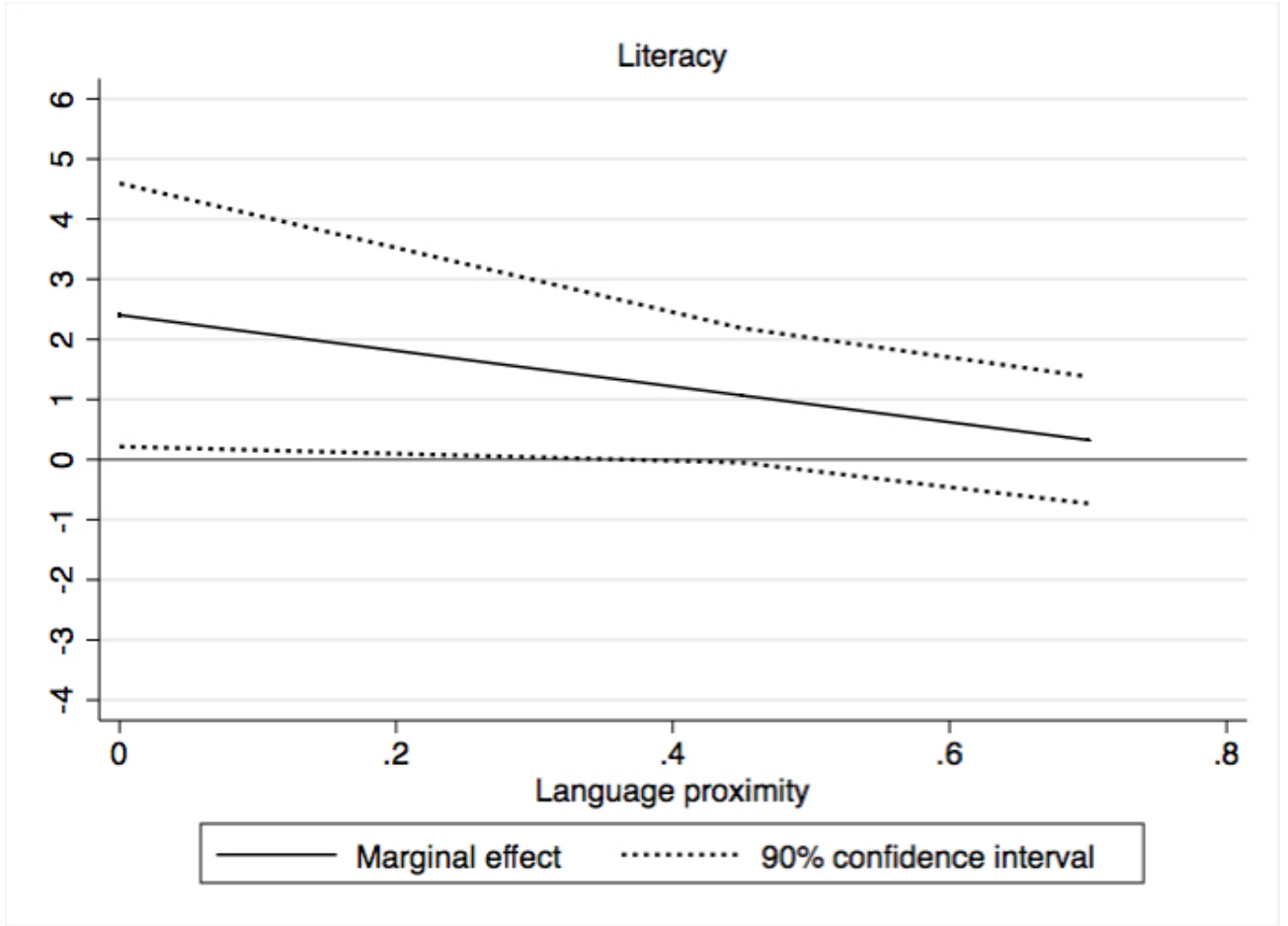


Figure 8: Marginal effect of early childcare attendance on literacy test score by language proximity and parents' low educational level.

Overall, results are consistent with those in previous sections, namely early childcare exerts positive effects on literacy outcomes of immigrant children with low educated parents. As a consequence, linguistic distance might well be considered a valid channel through which the effect operates.

7 Robustness checks

In this section we perform several robustness checks in order to support the validity of our identification strategy together with showing the robustness of our estimates to the potential threat associated with missing values in the variable `Early childcare`.

7.1 Robustness checks on IV assumptions

A key assumption for the validity of our IV strategy is the exclusion restriction, which implies the effect of the supply of early childcare spots to affect children's cognitive outcomes only through the children's attendance to the early childcare service. In section 5.1 we discussed in detail some potential threats to our identification strategy and in what follows we provide empirical evidence of the validity and robustness of our assumptions addressing most relevant issues.

Dealing with the fact that municipalities might differ on other dimensions rather than supply of early childcare spots, in Table 12 we estimate the same specification as in Table 5 but controlling also for taxable income per capita at the municipal level. We expect taxable income to proxy several types of resources provided by the municipality as well as the local labour market conditions. As we can see in Table 12 results are mostly unchanged and point in the direction of an even stronger positive impact of early childcare on immigrant children cognitive outcomes once we control for municipality taxable income.

Table 12: Heterogeneous effects by parents' educational level controlling for taxable income per capita.

LITERACY						
Panel A: Second Stage						
	(1) Whole Sample	Natives (2) High Edu	(3) Low Edu	(4) Whole Sample	Immigrants (5) High Edu	(6) Low Edu
Early childcare	-0.081 (0.119)	-0.030 (0.174)	0.28 (0.296)	0.162 (0.201)	0.117 (0.490)	1.614** (0.799)
Observations	361,236	310,412	49,476	30,896	21,142	8,684
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.001)	0.002*** (0.001)
First stage F-stat	42.12***	31.33***	33.56***	38.56***	9.44**	10.17***
Observations	361,236	310,412	49,476	30,896	21,142	8,684
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics, taxable income per capita and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1, 5 and 10 percent level.

As an additional robustness check, we control for the municipalities' political orientation looking at whether the incumbent mayor was supported by a left or a right party when elected²³. We merge our data set with elections data at the municipality level collected by [Bracco et al., 2018] from the Italian Ministry of the Interior²⁴. We use same specification as in Table 5 controlling also for the incumbent mayor political faction: a dummy variable coded equal to 1 if the mayor was endorsed by left parties or equal to 0 in case it was supported by right parties. Results are shown in Table 13. It emerges clearly that our main findings remain unaltered compared to our main specification both in terms of magnitude and significance.

²³In smaller municipalities it is common that mayors are endorsed by ad hoc unitary lists not directly related to national parties or that are grouped so that it is not possible to state neatly the political orientation. In our data mayors supported by unitary lists are coded as 0s.

²⁴Data are presented in [Bracco et al., 2018]. Elaborated data in this paper are kindly provided by the authors.

Table 13: Heterogeneous effects by parents' educational level controlling for municipality political orientation.

LITERACY						
Panel A: Second Stage						
	(1) Whole Sample	(2) High Edu	(3) Low Edu	(4) Whole Sample	(5) High Edu	(6) Low Edu
Early childcare	-0.081 (0.119)	0.068 (0.124)	-0.001 (0.233)	0.162 (0.201)	-0.113 (0.201)	0.894** (0.375)
Observations	361,236	283,721	42,475	30,8896	19,975	8,190
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.003*** (0.000)	0.003*** (0.001)	0.003*** (0.001)
First stage F-stat	42.12***	36.92***	31.82***	38.56***	28.98***	33.27***
Observations	361,236	283,721	42,475	30,896	19,975	8,190
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics, municipalities' political orientation and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

In sub-section 5.1 we argued that municipalities in the same province with a number of residents greater than 10,000 are likely to be characterized by a high level of internal homogeneity. However, one may consider the provincial capital of the municipality to be economically, culturally and socially different with respect to the other municipalities in the same province. Provincial capitals are usually the largest municipality in the province with a prominent economic and cultural role. In our framework, a provincial capital may provide better resources to parents and children together with a more vibrant environment potentially benefitting children's test outcomes. In order to address this potential threat to identification, we run same specification as in Table 5 but controlling also for a dummy equal to 1 if the municipality is a provincial capital and 0 otherwise. Results are again consistent with those in Table 5, pointing to an even larger effect of early childcare attendance on immigrant children's coming from a disadvantage background.

Table 14: Heterogeneous effects by parents' educational level controlling for provincial capital.

LITERACY						
Panel A: Second Stage						
	(1) Whole Sample	(2) High Edu	(3) Low Edu	(4) Whole Sample	(5) High Edu	(6) Low Edu
Early childcare	-0.081 (0.119)	0.076 (0.146)	0.027 (0.250)	0.162 (0.201)	-0.019 (0.404)	1.522* (0.810)
Observations	361,236	311,441	49,795	30,896	21,168	8,694
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.001)	0.002*** (0.001)
First stage F-stat	42.12***	37.54***	35.06***	38.56***	12.60***	9.30***
Observations	361,236	311,441	49,795	30,896	21,168	8,694
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics, provincial capital and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1, 5 and 10 percent level.

As a final robustness check, we control for taxable income per capita, mayor political faction and provincial capital dummy in the same specification at the same time. Results reported in Table 16 are again consistent to those in the main analysis and point to an even stronger effect of the early childcare attendance on immigrant children coming from disadvantaged social and cultural backgrounds.

Table 15: Heterogeneous effects by parents' educational level controlling for provincial capital.

LITERACY						
Panel A: Second Stage						
	(1) Whole Sample	(2) High Edu	(3) Low Edu	(4) Whole Sample	(5) High Edu	(6) Low Edu
Early childcare	-0.081 (0.119)	0.014 (0.189)	0.136 (0.342)	0.162 (0.201)	-0.159 (0.591)	1.927** (0.969)
Observations	361,236	283,050	42,345	30,896	19,912	8,163
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002*** (0.000)	0.002*** (0.000)	0.002*** (0.000)	0.003*** (0.000)	0.002*** (0.001)	0.002*** (0.001)
First stage F-stat	42.12***	25.54***	28.25***	38.56***	6.70**	8.27**
Observations	361,236	283,050	42,345	30,896	19,912	8,163
Province FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the literacy test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant ones. Column (2) and (5) refer to students with high educated parents (educational level of the parent with the highest level of education equal or higher than high school diploma), while column (3) and (6) refer to students with low educated parents (educational level of the parent with the highest level of education lower or equal to middle school diploma). All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics, municipalities' political orientation, taxable income per capita, provincial capital dummy and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F-stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1, 5 and 10 percent level.

Taken together, these findings reassure us that our instrument is not capturing other dimensions that are potentially correlated to children's cognitive outcomes.

7.2 Missing values

A potential concern with the use of the variable *early childcare* is the high incidence of missing values in this variable (31.05, 30.85 and 31.26 percent in the academic years 2017, 2016, 2015 respectively). In order to assess the relevance of this missing piece of information we include in our baseline IV specification (see Table 4) a dummy variable called **Missing** that takes values 1 when the information about student's early childcare attendance is not available and 0 otherwise.

Table 16: Missing values

	LITERACY		NUMERACY	
Panel A: Second Stage				
	(1) Native	(2) Immigrant	(3) Native	(4) Immigrant
Early childcare	-0.098 (0.131)	0.173 (0.216)	-0.234 (0.170)	0.098 (0.197)
Missing	-0.048 (0.051)	0.047 (0.097)	-0.085 (0.064)	0.034 (0.086)
Observations	424,862	34,169	410,867	33,023
Panel B: First Stage				
	(1)	(2)	(3)	(4)
Early childcare public supply	0.002*** (0.000)	0.003*** (0.001)	0.002*** (0.000)	0.003*** (0.001)
Observations	424,862	34,169	410,867	33,023
Province FE	✓	✓	✓	✓
Year FE	✓	✓	✓	✓
Individual controls	✓	✓	✓	✓
School and class controls	✓	✓	✓	✓

Table 17: Notes: Columns (1) to (4) report IV estimates for native and immigrant students where the dependent variables are literacy and numeracy test score. All the specifications include province and cohort fixed effects. Control variables include individual level characteristics, students' average characteristics and school and class characteristics according to Table 2 in Section 4. Robust standard errors are clustered at the municipal level. First stage F stat. refers to the Kleibergen-Paap F-statistics. The symbols ***, **, * indicate that the coefficients are statistically significant, respectively at the 1,5 and 10 percent level.

As shown in Table 16 the variable **Missing** is not statistically significant and coefficients estimates are mostly unchanged, thus confirming reliability of our results.

8 Conclusions

The successful integration of immigrant children into the educational system is one of the most important challenges for many European countries. Children who are migrants or have immigrant parents face significant educational barriers which may lead to substantial educational disadvantages relative to native students. Performance gaps at school translate in persistent differences in literacy levels, drop-out rates and eventually in labour market outcomes between immigrant students and their native peers, thus perpetuating inequalities and slowing the integration of migrants in the host country.

In this paper, we analysed the impact of early childcare attendance on second-generation immigrant children's cognitive outcomes. In particular, we investigated whether childcare provided at very young age is an effective policy to close the existing performance gap between immigrant and native children. To carry out our analysis we drew on novel Italian administrative data and we make use of a neat IV identification strategy in order to deal with children's non-random sorting into early childcare.

While we find no significant effect for native children, our results point out to a positive and significant impact of early childcare attendance on immigrant children coming from a disadvantaged background, with the effects being stronger for female students. Further, we find that the impact of childcare is stronger for immigrant children whose mother tongue is further from Italian, which suggests that access to childcare may be a relevant policy tool to improve language

proficiency of immigrant children, thus fostering their integration and assimilation in the host country.

Our findings provide empirical support to the renewed interest of social scientists and policy makers in the attendance to early childcare and pre-primary school programs as an effective tool to facilitate immigrant children's integration and close their educational gap.²⁵ Unfortunately, despite of the encouraging empirical evidence and the rich policy debate, the access of children with immigrant background to early childcare facilities (and other pre-primary school programs) is still very limited in OECD countries (including Italy, the country where the data used in this study come from).²⁶ Institutional and economic factors play a crucial role in determining the limited access of immigrant children to these precious social services. In Italy, for instance, municipal authorities are responsible for administering early childcare programs and defining the rules to allocate the limited available slots over families. In some cases, this institutional situation has translated into inefficient and discretionary allocation rules, such as those restricting the access to early childcare programs to families satisfying time requirements with their legal residence in the municipality.²⁷ Our results strongly question the desirability and appropriateness of such legal requirements, as they are very likely to reduce the social benefit of early childcare programs and limit the effective integration of immigrant children in the education system of the host country.

²⁵See, for instance, the recent [OECD, 2018].

²⁶For Europe, see [CoE, 2017]. According to [OECD, 2015], first-generation immigrant students are almost half as likely as non-immigrant students to have attended pre-primary education, with significant exceptions such as Belgium Austria, Slovenia, Canada, and Norway where immigrant students are more likely than non-immigrant to have attended pre-primary school education.

²⁷For instance, in 2018 Veneto introduced the requirement of 15 years of legal residence in the municipality to get access to early childcare programs. This requirement was declared unconstitutional after few months.

References

- [Adsera and Pytlikova, 2015] Adsera, A. and Pytlikova, M. (2015). The role of language in shaping international migration. *The Economic Journal*, 125.586:F49–F81.
- [Ahad and Benton, 2018] Ahad, A. and Benton, M. (2018). Mainstreaming 2.0. Technical report, Brussels: Migration Policy Institute Europe.
- [Alesina et al., 2018] Alesina, A., Carlana, M., Ferrara, E. L., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. Technical report, National Bureau of Economic Research.
- [Antonelli et al., 2011] Antonelli, M., Grembi, V., et al. (2011). *Target centrali e finanza locale: Il caso degli asili nido in Italia*, volume 706. Carocci.
- [Antonelli and Grembi, 2014] Antonelli, M. A. and Grembi, V. (2014). Central targets and local agendas: Missing lisbon 2010. Public finance research papers 6, Istituto di Economia e Finanza, DIGEF, Sapienza University of Rome.
- [Baker et al., 2008] Baker, M., Gruber, J., and Milligan, K. (2008). Universal child care, maternal labor supply, and family well-being. *Journal of Political Economy*, 116(4):709–745.
- [Baker et al., 2015] Baker, M., Gruber, J., and Milligan, K. (2015). Non-cognitive deficits and young adult outcomes: The long-run impacts of a universal child care program. Technical report, National Bureau of Economic Research.
- [Berlinski et al., 2009] Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of public Economics*, 93(1-2):219–234.
- [Blau and Currie, 2006] Blau, D. and Currie, J. (2006). Pre-school, day care, and after-school care: who’s minding the kids? *Handbook of the Economics of Education*, 2:1163–1278.
- [Bornstein et al., 2004] Bornstein, M. H., Hahn, C.-S., and Haynes, O. M. (2004). Specific and general language performance across early childhood: Stability and gender considerations. *First language*, 24(3):267–304.
- [Bracco et al., 2018] Bracco, E., De Paola, M., Green, C. P., and Scoppa, V. (2018). The effect of far right parties on the location choice of immigrants: Evidence from lega nord mayors. *Journal of Public Economics*, 166:12–26.
- [Brilli et al., 2016] Brilli, Y., Del Boca, D., and Pronzato, C. D. (2016). Does child care availability play a role in maternal employment and children’s development? evidence from italy. *Review of Economics of the Household*, 14(1):27–51.
- [Carta and Rizzica, 2018] Carta, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children’s outcomes: Evidence from italy. *Journal of Public Economics*, 158:79–102.
- [CoE, 2017] CoE (2017). Fighting school segregation in europe through inclusive education: a position paper. Technical report, Concil of Europe, Commissioner for Human Rights.
- [Cornelissen et al., 2018] Cornelissen, T., Dustmann, C., Raute, A., and Schönberg, U. (2018). Who benefits from universal child care? estimating marginal returns to early child care attendance. *Journal of Political Economy*, 126(6):2356–2409.

- [Cunha and Heckman, 2007] Cunha, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.
- [Del Boca et al., 2016] Del Boca, D., Pasqua, S., and Suardi, S. (2016). Child care, maternal employment, and children’s school outcomes. an analysis of italian data. *European Journal of Population*, 32(2):211–229.
- [Del Boca and Vuri, 2007] Del Boca, D. and Vuri, D. (2007). The mismatch between employment and child care in italy: the impact of rationing. *Journal of Population Economics*, 20(4):805–832.
- [Drange and Havnes, 2018] Drange, N. and Havnes, T. (2018). Early child care and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics (forthcomings)*.
- [Drange and Telle, 2015] Drange, N. and Telle, K. (2015). Promoting integration of immigrants: Effects of free child care on child enrollment and parental employment. *Labour Economics*, 34:26–38.
- [Dustmann et al., 2012] Dustmann, C., Frattini, T., and Lanzara, G. (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, 27(69):143–185.
- [Dustmann et al., 2013] Dustmann, C., Raute, A., and Schönberg, U. (2013). Does universal child care matter? evidence from a large expansion in pre-school education.
- [Felfe and Huber, 2017] Felfe, C. and Huber, M. (2017). Does preschool boost the development of minority children?: the case of roma children. *Journal of the Royal Statistical Society (Series A)*.
- [Felfe and Lalive, 2018] Felfe, C. and Lalive, R. (2018). Does early child care affect children’s development? *Journal of Public Economics*, 159:33–53.
- [Felfe et al., 2014] Felfe, C., Nollenberger, N., and Rodríguez-Planas, N. (2014). Can’t buy mommy’s love? universal childcare and children’s long-term cognitive development. *Journal of population economics*, 28(2):393–422.
- [Fenson et al., 1994] Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., and Stiles, J. (1994). Variability in early communicative development. *Monographs of the society for research in child development*, pages i–185.
- [Fort et al., 2019] Fort, M., Ichino, A., and Zanella, G. (2019). Cognitive and non-cognitive costs of daycare 0–2 for children in advantaged families. *Jornal of Political Economy (forthcoming)*.
- [Frattini and Meschi, 2019] Frattini, T. and Meschi, E. (2019). The effect of immigrant peers in vocational schools. *European Economic Review*, 113:1–22.
- [Frattini and Vigezzi, 2018] Frattini, T. and Vigezzi, N. (2018). 2nd migration observatory report: “immigrant integration in europe and in italy”. Technical report, Centro Studi Luca D’Agliano and Collegio Carlo Alberto.
- [Havnes and Mogstad, 2011] Havnes, T. and Mogstad, M. (2011). No child left behind: Subsidized child care and children’s long-run outcomes. *American Economic Journal: Economic Policy*, 3(2):97–129.
- [Havnes and Mogstad, 2015] Havnes, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127:100–114.

- [Heckman and Masterov, 2007] Heckman, J. J. and Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29(3):446–493.
- [Herbst, 2013] Herbst, C. M. (2013). The impact of non-parental child care on child development: Evidence from the summer participation “dip”. *Journal of Public Economics*, 105:86–105.
- [Herbst, 2017] Herbst, C. M. (2017). Universal child care, maternal employment, and children’s long-run outcomes: Evidence from the us lanham act of 1940. *Journal of Labor Economics*, 35(2):519–564.
- [MIUR, 2018] MIUR (2018). Gli alunni con cittadinanza non italiana a.s. 2016/2017. Technical report, Italian Ministry of Education Statistical Office.
- [OECD, 2012] OECD (2012). Untapped skills. realizing the potential of immigrant students. Technical report.
- [OECD, 2015] OECD (2015). Help immigrant students to succeed at school and beyond. Technical report.
- [OECD, 2018] OECD (2018). Working together for local integration of migrants and refugees. Technical report.
- [Quintano et al., 2009] Quintano, C., Castellano, R., and Longobardi, S. (2009). *A fuzzy clustering approach to improve the accuracy of italian student data: An experimental procedure to correct the impact of outliers on assessment test scores*. Vita e pensiero.
- [Todd and Wolpin, 2003] Todd, P. E. and Wolpin, K. I. (2003). On the specification and estimation of the production function for cognitive achievement. *The Economic Journal*, 113.485:F3–F33.