CefES

Center for European Studies

PAPER SERIES

———

Impact of Early Childcare on Immigrant Children's Educational Performance

Luca Corazzini, Elena Meschi
and Caterina Pavese

The Center for European Studies (CefES-DEMS) gathers scholars from different fields in Economics and Political Sciences with the objective of contributing to the empirical and theoretical debate on Europe.

DIPARTIMENTO DI ECONOMIA, METODI QUANTITATIVI E STRATEGIA DI IMPRESA

UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

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
[‡]Università degli Studi di Milano-Bicocca - Department of Economics, Management and Statistics
DEMS; Università degli Studi di Milano-Bicocca - Center for European Studies (CefES)

October 23, 2020

Abstract

This paper investigates the impact of attending early childcare on immigrant children's cognitive outcomes. Our analysis makes use of administrative data on the entire population of students in the 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 early childcare availability 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 attendance. Our results point out that the effect of early childcare attendance differs between native and immigrant children. Estimates show a positive and significant effect on the language test scores of immigrant children, with the effect being mostly driven by females, by children with low-educated mothers and by children who, at home, speak a language highly dissimilar to Italian. Unlike immigrants, native students are negatively impacted by early childcare attendance, as reflected in both language and math test scores. Effects are stronger on math test scores for females and for children with highly educated mothers.

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

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

---

[*]Corresponding author: email elena.meschi@unimib.it

1 Introduction

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

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

In this paper, we aim to estimate the impact of early childcare on second-generation immigrant children's cognitive outcomes. Our focus on early education traces back to the findings of Heckman and co-authors, who suggested 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 at this stage of development, investments made early in life show higher returns as well as dynamic complementarities with respect to investments made later in life (Cunha and Heckman [2007] and Heckman and Masterov [2007]). This is particularly true for children from disadvantaged backgrounds taking part in targeted interventions (Blau and Currie [2006]).

However, 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 administrative data on three cohorts of primary school students in Italy in order to investigate the impact of attending early childcare on later cognitive outcomes. In particular, we use the rich dataset on children's cognitive skills administered yearly to the entire population of students in the fifth grade by the Italian Institute for the Evaluation of the Educational System (henceforth INVALSI) for the school years 2014/2015 to 2016/2017, merged with unique administrative records on publicly available early childcare slots at the municipal level, collected by Antonelli and Grembi [2011]. In the INVALSI dataset, cognitive skills are measured by standardised and externally marked tests in language and math. Together with test scores, INVALSI collects data on schools', families' and students' characteristics, from which we take the information concerning whether each student attended early childcare or not.

We believe that our focus on Italy is particularly appropriate for two reasons: Italy only recently became a destination for international migration, and it is one of the OECD countries with the greatest educational disparity between native and immigrant students (OECD [2012]). Combined, these two features contribute to an invaluable setting for our analysis.

We contribute to the existing literature along several dimensions. First, while an extensive number of studies examine interventions geared toward children aged 3-5 years, only a few focus on children aged 0-2 years (i.e. early childcare), and the evidence they provide is mixed. Some studies find positive effects (Felfe and Lalive [2018] and Drange and Havnes [2019] 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 make use of two alternative measures of students' performance: INVALSI standardised test scores, and oral marks assigned by teachers. The two measures are likely to capture different dimensions of students' abilities, providing an additional understanding of the impact of early childcare attendance. We can also perform several tests of heterogeneity and thus identify some of the channels at work, such as the role of linguistic distance of immigrants' native language from Italian.

The identification of the causal effect of early childcare attendance is challenging: it involves the decision of parents to enrol children in the service, as well as the choice of early childcare providers to accept those children whose attendance could potentially benefit them the most. The variables underlying these decision-making processes are neither entirely observable nor measurable. Consequently, the selection of children into early childcare services is not random. To deal with the 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. For immigrant children, we find a positive and significant effect of early childcare attendance on language test scores. To investigate the channels driving our results, we explore whether the effect differs within the group of immigrant children, and we analyse the role played by students' gender, mother's educational level and the linguistic distance of immigrant children's mother tongue to the host-country language (i.e. Italian). As highlighted by our findings, children with low-educated mothers and who speak a language that differs significantly from Italian capture the largest benefits from attending early childcare, with the effect concentrated on female students. Overall, our results suggest that attending early childcare is an effective tool for fostering immigrant children's achievements, helping to close existing gaps with respect to native children.

For native students, we document the negative effects of early childcare attendance on both language and math test scores. In line with the effects uncovered in Fort et al. [2019], females and children with highly educated mothers are the populations most harmed in this context. These results might be driven by the relative lack of educational content of early childcare services in Italy, which is mostly care oriented, as well as the inadequacy of other institutional features such as the low adult-to-children ratio that might be detrimental to children with relatively better alternative modes of care.

In terms of policy implications, our analysis sheds light on the potential target(s) for policy interventions, uncovering (i) which group of immigrant students might benefit the most from early childcare attendance, and (ii) one of the crucial mechanisms behind the resulting effect. As a matter of fact, positive effects are primarily enjoyed by immigrant children coming from less-advantaged families, and linguistic distance plays a key role in explaining the positive impacts observed. This role should be taken into account when designing and implementing policies aimed at the successful integration of migrants into society. Also, our evidence invites consideration of how early childcare attendance negatively affects native students, arguably due to the relative lack of educational content and the low teacher-students ratio, which might be particularly detrimental to this group of students.

The structure of the paper is as follows. In Section 2, we provide a summary of the existing literature on early childcare and children's outcomes, together with an underlining of our paper's contributions. In Section 3, we describe the Italian system of public early childcare. Next, we

present our data and provide some relevant descriptive statistics in Section 4. In Section 5, we illustrate our identification strategy and discuss the threats to its validity. We then present and comment on our results in Sections 6, and we provide several robustness checks in Section 7. Finally, in Section 8, we offer some concluding remarks and discuss the possible policy implications of our findings.

2 Literature review

This paper draws on and contributes to two strands of literature. First, it relates to the large body of literature on the impact of early life experiences and early interventions on children's cognitive skills. Second, it adds to the literature on the educational gap between native and immigrant students (Cornelissen et al. [2018], Drange and Telle [2015]).

The existing literature on early interventions and cognitive outcomes has mainly focused on interventions geared toward children aged 3-5, and has typically found that childcare has positive effects on child development, especially for children from disadvantaged background (see Havnes and Mogstad [2015], Felfe et al. [2014], and Berlinski et al. [2009], among others). Research on the effects of childcare on younger children (0-2 years old) - which we call *early childcare* - is even more scarce, and the evidence is mixed. Desirable effects on cognitive and social skills from attending early childcare are stronger and more robust in countries with high quality and large availability of early childcare service, as in the northern European countries. For example, 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 [2019] provide evidence concerning the impact of early enrolment in childcare on children's educational attainments at age 7 in Norway. Their identification strategy exploits a randomised lottery employed in the city of Olso to allocate slots as an instrument for the age of the child, when the child first attends childcare. Entering the lottery implies attending the service about four months earlier. Their results point out that children enrolled early in day care facilities perform better on both the language and math tests when 7 years old, with the effect being driven by children from economically disadvantaged families.
However, the evidence of desirable effects is hardly equivocal. Evidence of negative effects emerged in various studies evaluating the universal early childcare expansion that occurred in Quebec during the 90s. Among these, the work of Baker et al. [2008] stands out: exploiting a difference-in-difference strategy to compare children's outcomes before and after the reform, the authors found childcare utilisation to be detrimental to children's behavioural and health outcomes following the policy change. These negative outcomes are confirmed in a subsequent study by the same authors, which also found little impact on cognitive test scores (Baker et al. [2019]). 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.

Our paper contributes to this literature by studying the impact of early childcare in Italy. Early childhood interventions in Italy received research attention only recently, thanks to the increasing availability of reliable data on children's educational outcomes in the past few years. The findings are also mixed in this case. Brilli et al. [2016] find a positive impact of early childcare availability on children's language cognitive outcome, with the effect being driven by areas with a higher level of rationing. The same conclusions are reached by Del Boca et al. [2016a], who, implementing an instrumental variable approach, provide evidence that early childcare attendance has a positive effect on children's later educational achievements. Carta and Rizzica [2018] exploiting a reform

that introduces the possibility of entering the fully-subsidised preschool service earlier, argue that entering preschool at a younger age has no significant effect on children's 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 the growing share of immigrant students in all OECD countries, very few studies investigated the impact of early childhood programmes on children with an immigrant background. In this group, Cornelissen et al. [2018] evaluate the heterogeneous impact of a universal childcare programme in Germany by using unique administrative data and exploiting the staggered implementation of a federal policy reform across municipalities that entitles every child to a slot when he or she turns 3 until their school entry at age 6. The authors detect a pattern of reverse selection on gains: children more (less) likely to attend these services benefit the least (most) from the attendance. In particular, they find that early childcare attendance works as an 'equaliser': children with immigrant ancestry are less likely to attend early childcare, but experience the largest gains in terms of overall school readiness in comparison to native children. Results are driven by students coming from families with less resources. In the same vein, the study of Drange and Telle [2015] takes advantage of a targeted intervention, one providing 4 hours of free childcare in some districts of Olso, 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 for girls and children coming from a disadvantaged background.

Our paper contributes to the scarce literature on the impact of early childhood programmes and focusing on immigrant children, a fast-growing and low-investigated group. Further, given the richness of our data, we are able to test whether the effect is heterogeneous along various dimensions, such as gender, parents' education and immigrant children's linguistic distance to the host-country language. To the best of our knowledge, no other studies investigated the role of linguistic distance in shaping immigrant children's early education experience. Additionally, our data contain information concerning individual early childcare attendance. Unlike many existing papers in the early education literature that focus on intention-to-treat effects (Carta and Rizzica [2018], Del Boca et al. [2016a] and Havnes and Mogstad [2011], among others), we are able to identify the causal effect of early childcare *attendance* rather than the effect of early childcare *eligibility*. Later, in our empirical analysis to instrument individual early childcare attendance, we use the supply of early childcare spots defined at the *municipal* level when the student was born. Compared to other existing studies (see for instance Del Boca et al. [2016a]), our instrument definition is more granular and coincides with the actual supply when the child was younger than 1 year old.

3 Background

### 3.1 Immigrant children in Italy

Over the past three decades, Italy became subject to a massive inflow of immigrants. In 2017, the share of immigrants represented 10% of the total population, and their number increased by more than 30% between 2009 and 2017, rising from 4.5 million to 5.9 million. Most immigrants come from Eastern European countries, and the main countries of origin are Romania, Albania, Morocco, China and Ukraine. [1] They are typically characterised by low educational levels and

---

[1]Istat, "Demografia in cifre" (www.demo.istat.it).

are mostly concentrated at the bottom of the native wage distribution. Relative to Italians, they tend to have a lower socio-economic background Frattini and Vigezzi [2018].

The surge in migration has deeply affected the Italian educational system. In the school year 2016/2017, immigrant children represented 10.8% of students in Italian primary schools, the educational stage that absorbs the highest number of students without Italian citizenship.[2] Figure 1 shows the number of immigrant children in Italian primary schools by generational status, from school years 2012/13 to 2016/2017. Overall, the figure points out the continual and stable growth in the numbers of non-native students, particularly those of second-generation status. In the five school years considered, second-generation immigrant students (the focus of our analysis) rose from 182.315 to 221.643 units, with an increase of 27%, representing 76.6% of the total number of primary school students without Italian citizenship in the school year 2016/17.

Figure 1: Number of immigrant children in primary school by generational status (s.y. 2012/13 - 2016/17).
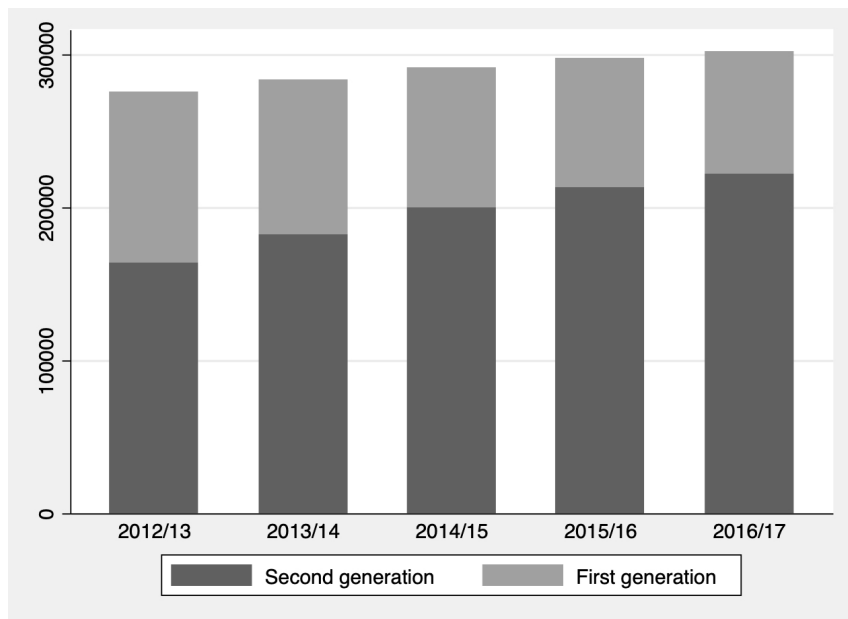
Source: Elaboration on MIUR data - Statistical office.

Figure 2 reports the geographical distribution of immigrant students across municipalities for the 2016/17 school year and reveals that while most of immigrants are concentrated in the Northern and Central regions, there is substantial variability across different municipalities within regions.

---

[2]Miur, "Portale dei dati sulla scuola" (www.dati.istruzione.it).

Figure 2: Percentage of immigrant students in primary schools across municipalities. School year 2016/2017.
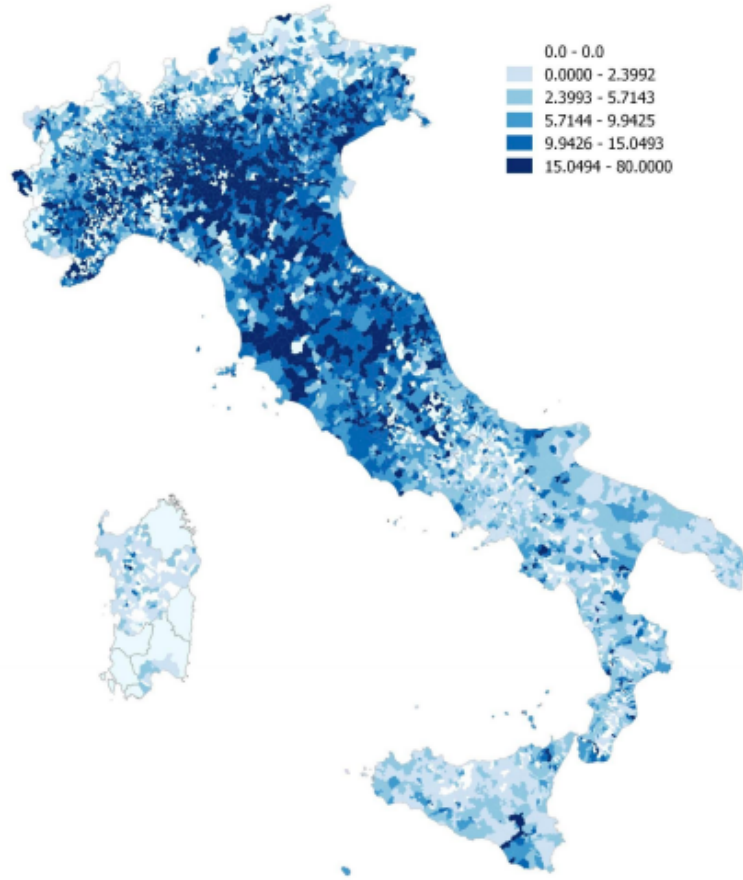
Source: MIUR - Ufficio Statistica e studi.

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 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 access to preschool services is almost universal and provided free of charge, early childcare is highly rationed and exhibits substantial cost variability across all Italian regions.

Further, early childcare service is decentralised at the municipal level, provided by both public and private bodies, and, in contrast to 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.[3] 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

---

[3] There 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 place a greater emphasis on families with disadvantaged socio-economic conditions; others, on parents' employment status (e.g. whether one or both parents are employed or not); and still others, on a family's composition (i.e. whether the child is orphan or has siblings). For further details, see Del Boca et al. [2016b].

central government is only responsible for the definition of common standards and resource allocation among regions. Both municipalities and families contribute to meeting the total cost of public early childcare, and fees are means-tested against the family income. On average, children spend 30.6 hours per week in early childcare, and the minimum qualification requirement for the educational staff is an upper secondary school diploma or a master degree in pedagogy. The supply of early childcare spots accommodates only 22.8 % of the potential demand of 0–2-year-old children, with relevant differences across municipalities. We exploit this heterogeneity for our identification strategy. Further details are provided in subsection 4.2.

4 Data

4.1 INVALSI data

Our analysis draws on administrative data from INVALSI, the national agency that carries out yearly testing of students' competences in language and math. Standardised tests are administered every year to the entire population of students in the second, fifth, eighth and tenth grades. We focus our analysis on students who attended the fifth grade (the last year of primary school) in school years 2014/2015 to 2016/2017, the most recent years for which data is available. The tests administered by INVALSI are standardised and are compulsory for all Italian schools and students attending the grades of interest. Like all standardised tests, they must satisfy two conditions: first, all students must answer the same questions during the same time interval; and second, correction and grading of the test must follow standardised procedures that are established a priori and independently of those making these corrections. In other words, tests are to be anonymously and externally marked. These features of the INVALSI tests make students' results fully comparable across all Italian schools and guarantee the tests to be objective.
All students take the test in language and math in their own classroom on two different dates, typically in late May. The language test aims at assessing students' proficiency in reading comprehension, which involves grammatical, lexical and pragmatic competences. The math test evaluates students' knowledge and mastering of mathematics-specific content 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 a student's class was involved in cheating behaviours during the test.[4] INVALSI also provides marks assigned by teachers in math and language at the end of the first term. Marks are distinguished as either written or oral, and we regard oral marks to be particularly interesting. In contrast to INVALSI test scores, they are non-blind marks that reasonably take into account more than just the cognitive skills of students, since teachers' marks are relative and likely to be affected by the class' and students' behaviour.

Together with these test scores, INVALSI provides detailed information on families', students' and schools' characteristics. This 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 language or math test and a number of general information queries concerning students' performances and characteristics, as provided by the school administrative staff. Among these questions, we have information as to whether each child attended early childcare and pre-school programs.

Additionally, in our analysis we use information on a rich set of other individual and family characteristics such as parents' education and working status; students' gender, month of birth and citizenship (native and second-generation immigrant children); and a synthetic index of

---

[4]For further details see Quintano et al. [2009].

Economic, Social and Cultural Status (ESCS index).[5] In our robustness checks, we also add a set of school and class characteristics, such as their average size and their square, and the share of female, immigrant and retained students and poorly educated parents per class. [6] The granularity of our data allow us to distinguish between Italian and non-Italian students. In Italy, this distinction is made according to the *Ius sanguinis* principle, according to 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 parents 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 early childcare slots available at the municipal level, collected by Antonelli and Grembi [2014] and Antonelli and Grembi [2011]. Administrative records were collected using municipal final balances provided by the Ministry of the Interior. They measure the availability of all early childcare public spots directly and indirectly provided by Italian municipalities for the years 2005, 2006 and 2007, when the students in our sample were aged 0 and their parents could apply to have their children admitted to childcare facilities. In particular, the dataset contains information on the number of early childcare public spots provided by each municipality, normalised by 10,000 residents. Using municipality and cohort identifiers, we assign to each student observed in INVALSI the actual childcare provision available at the time of his birth. Information is available for about 600 Italian municipalities with populations of residents above the 10,000 threshold.[7] In our INVALSI dataset, about 70% of the observed students live in municipalities with a population greater than or equal to 10,000 residents; in those same municipalities, the largest share of the immigrant population residing in Italy is concentrated.[8]

As we will discuss in the next section, we use this detailed information on early childcare supply as an instrument for the individual early childcare attendance. The instrument is particularly suitable in the case of Italy: the country is characterised by its low availability of public childcare slots, with a national average coverage rate of around 15 %, albeit highly differentiated across regions. This data is well behind the 2002 European target of childcare slots for at least 33% of children under three years of age. Therefore, despite various policy initiatives implemented since the end of the 1990s that aim to increase childcare availability, the demand for public childcare is higher than the supply in all Italian regions. In fact, the data on waiting lists reported by Cittadinanza Attiva (2011) indicates that the shares of children who submitted an application

---

[5]The ESCS index is built by applying principal component analysis using the information provided by students and schools about families' cultural resources, such as internet connection and number of books, and parents' educational levels 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 on 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.

[6]The share of low-educated parents is defined as the share of parents with an educational level lower than or equal to a lower secondary school diploma per class.

[7]For municipalities under the 10,000 threshold, it was not possible to proceed with the integration of some relevant pieces of information.

[8]In Italy, there are about 8,000 municipalities; most are of small dimension, having populations below 10,000 residents.

but were not given a slot ranged from than 12% in Lombardia to more than 40% in Sicilia. This means that many Italian households are still confronted with availability constraints.

Figure 3 shows both the average and the standard deviation of early childcare supply in different Italian regions in 2005 and 2007. It indicates that there is substantial heterogeneity across regions, with the lowest supplies in Calabria and Campania, and the highest in Emilia-Romagna and Toscana. 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 reduced it. Molise is the most striking example, with a decrease in supply of about 30%. Crucially, Figure 3 highlights that the supply of childcare is heterogeneous not only across regions, but also within regions across different provinces and municipalities. Noticeably, the standard deviation relative to the mean shows that there is substantial variation within most of the regions in our analysis. There are many factors that might explain territorial differences in early childcare supply over time, ranging from variability of funding at the central and local levels as well as political variables such as policy makers' preferences. In sections 5.1 and 7, we address at length whether these elements might threaten our identification assumption.

Figure 3: Public early childcare (ECC) supply across regions - 2005 and 2007.
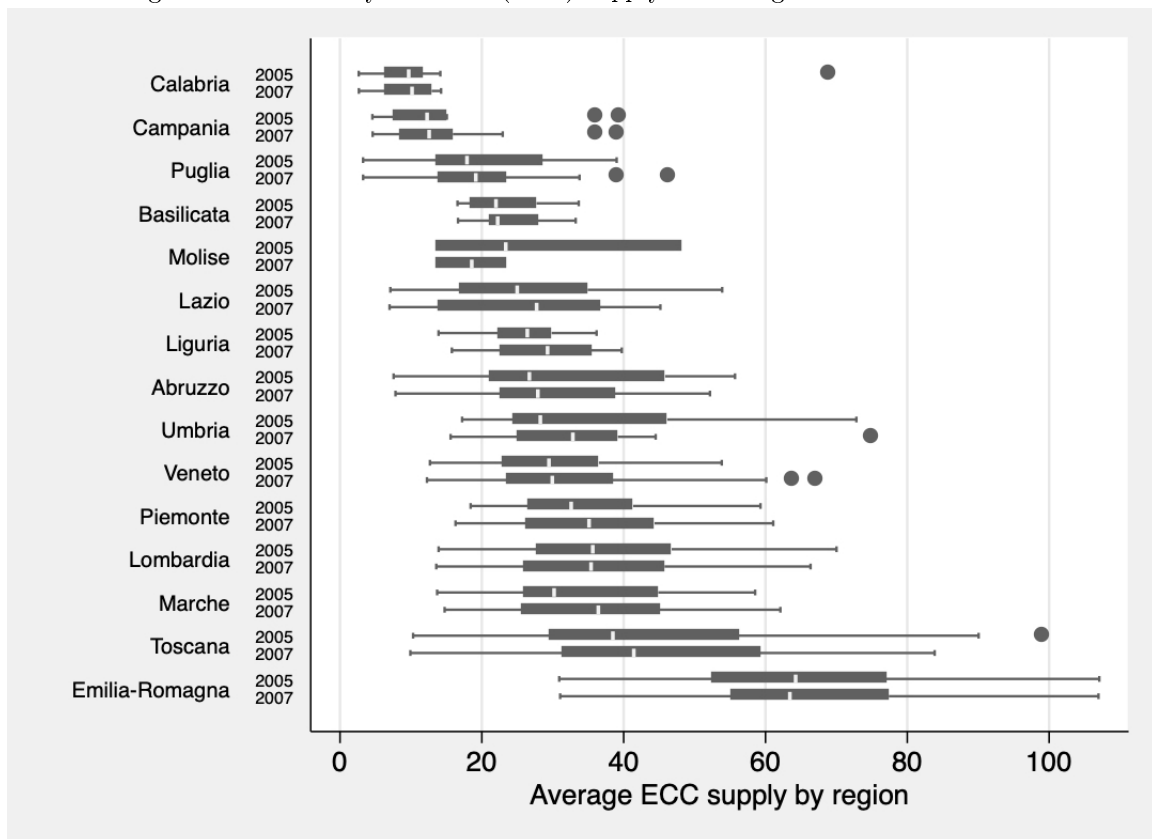
Notes: the variable *early childcare supply* is defined as the number of municipal early childcare public spots normalised by 10,000 residents. Source: Antonelli and Grembi [2014] and Antonelli and Grembi [2011].

Overall, our final dataset covers 362,816 students (333,241 natives and 29,575 immigrants) in about 600 municipalities and 2,763 schools located in 15 Italian regions.[9] For each school year

---

[9]Italy is composed of 20 regions, five of which have a special statute. We exclude these regions from our analysis due to the lack of data on municipal elections and the resulting political faction of the mayor, which we include as a control variable in our main specification. Thus, our sample comprises 15 regions and about 85 provinces (intermediate administrative bodies between regions and municipalities).

in the analysis, the share of the immigrant population residing in these municipalities is about 54% of the total immigrant population residing in Italy.[10]

Table 11 in the Appendix lists and defines all of the variables included in our empirical analysis.

4.3 Descriptive evidence

As a preliminary descriptive analysis, in Figures 4 and 5 we provide the kernel density estimation of the probability density function of the language and math test scores, respectively. Each figure reports the distribution of standardised test scores in language (Figure 4) and math (Figure 5) 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. As expected, these differences are more pronounced in the language test score distribution. For both native and immigrant students, there seems to be a positive difference between the performances of children who attended early childcare and those who did not, which is a first signal of positive selection into the service.

Figure 4: Kernel density language test score by early childcare attendance.
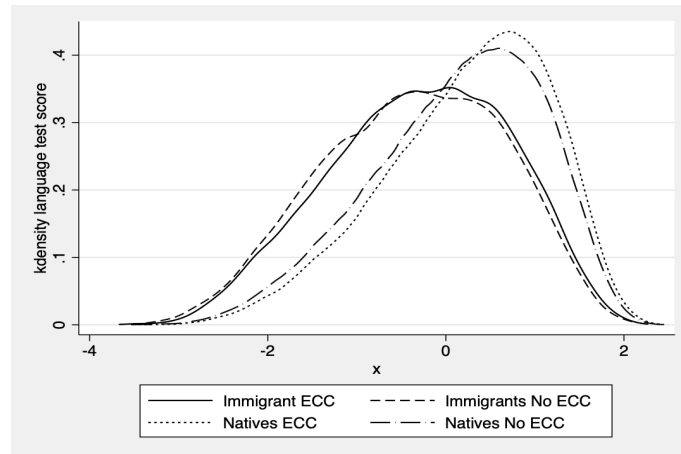
Source: INVALSI

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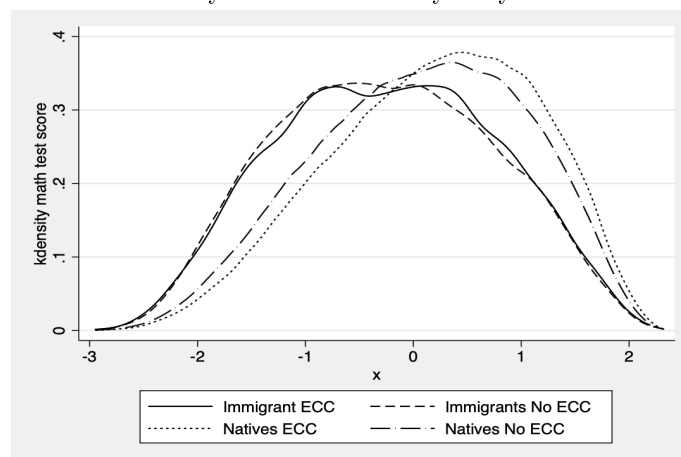
Source: INVALSI.

[10]This percentage is calculated by merging our data with yearly data on the number of immigrant residents per municipality, available at "http://dati.istat.it/".

11

Table 1 presents descriptive statistics 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 of the observable characteristics included in our analysis (Columns [3] and [6]), which further confirms that the process behind the service attendance is likely to be endogenous. On average, children enrolled in early childcare come from a more advantaged background than children not enrolled, independently from citizenship. They are more likely to have a working mother, parents with a high educational level and higher values for the ESCS index. This may be due to the specific access criteria applied by municipalities that tend to favour families with both parents working full time. In terms of observable individual and parental background characteristics, this descriptive evidence thus suggests children's positive selection into early childcare, which seems to be less pronounced for immigrant children than for native ones.

Table 1: Descriptive statistics and balancing test.

	Natives			Immigrants		
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	ECC	No ECC	p-value	ECC	No ECC	p-value
Background characteristics						
Female	0.484	0.502	0.000	0.496	0.505	0.125
Father compulsory education	0.233	0.334	0.000	0.337	0.428	0.000
Father high school	0.480	0.467	0.000	0.496	0.505	0.000
Father higher education	0.285	0.198	0.000	0.177	0.129	0.000
Mother compulsory education	0.143	0.248	0.000	0.339	0.438	0.000
Mother high school	0.487	0.511	0.000	0.452	0.406	0.000
Mother higher education	0.369	0.240	0.000	0.208	0.154	0.000
Father employed	0.967	0.955	0.000	0.882	0.889	0.070
Mother employed	0.824	0.640	0.000	0.528	0.358	0.000
Preschool	0.983	0.947	0.000	0.969	0.912	0.000
Late students	0.005	0.009	0.000	0.037	0.053	0.000
ESCS students	0.539	0.234	0.000	-0.293	-0.453	0.000
Outcome variables						
Language test score	0.237	0.120	0.000	-0.302	-0.375	0.000
Math test score	0.187	0.062	0.000	-0.238	-0.265	0.000
Language oral mark	0.264	0.161	0.000	-0.375	-0.430	0.000
Math oral mark	0.251	0.133	0.000	-0.270	-0.308	0.000
Observations	145,006	188,235		12,997	16,578	

Note: the table reports means of covariates by citizenship and by child's early childcare service (ECC) attendance. Column (3) and (6) report the p-value of the t-statistic for equality of means in the treated and control groups.

5 Identification strategy

We start our analysis by estimating the following linear model that links a child's cognitive outcomes (Y_{imt}^j) to his or her early childcare attendance ($Earlychildcare_{imt}$), while controlling for a set of individual, family and municipality characteristics as well as cohort fixed effects, province fixed effects and province-specific trends across cohorts:

$$Y_{imt}^{j} = \gamma_0 + \gamma_1 ECC_{imt} + \mathbf{X'}_{imt}\boldsymbol{\gamma_2} + \boldsymbol{\Phi'}_{mt}\boldsymbol{\gamma_3} + \eta_t + \lambda_{im} + \eta_t \times \lambda_{im} + \epsilon_{imt} \qquad (1)$$

where i indexes individuals, m indexes municipalities, t indexes cohorts, and j language/math test score/oral mark), ECC is the main explanatory variable indicating whether or not the child attended the childcare service and $\mathbf{X}$ is a vector of socio-demographic variables. Specifically, we control for preschool attendance, gender, quarter of birth, studies regularity and parents' education. The choice of individual control variables is motivated by the literature related to children's human capital production function (Cunha and Heckman [2007], Todd and Wolpin [2003]). Crucially, the vector $\boldsymbol{\Phi}$ controls for some key municipality characteristics such as population size, taxable income and political orientation. Importantly, we add a vast array of fixed effects. We include province fixed effect, λ_{im}, to control for province-specific, time-invariant unobservable characteristics that might systematically affect early childcare attendance as well as school attainment, and cohort fixed effects, η_t, in order to control for common shocks across cohorts. Eventually, we include province dummies interacted with a cohort time trend, $\eta_t \times \lambda_{im}$, to control province-specific supply or demand shocks affecting childcare attendance over time. Finally, ϵ_{imt} is the residual idiosyncratic error term.

We consider alternatively four dependent variables: language and math standardised INVALSI test scores, and language and math oral marks assigned by teachers. Test scores are defined as the percentage of correct answers and range from 0 to 100, while marks assigned by teachers range from 0 to 10, where 6 is the passing grade.[11] To increase comparability across cohorts, we standardised language and math test scores as well as oral marks to have means equal to 0 and standard deviations equal to 1 in each cohort and subject. All specifications are run separately for natives and second-generation immigrant students.[12]

The parameter of interest is γ_1 which 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 with the service based on their (un)observable characteristics and preferences. Therefore, comparing the outcomes of children who did or did not attend the service would result in biased estimates, even when conditioned by a rich set of control variables.

The literature addressed this endogeneity issue using different strategies. The most common methodologies rely on a difference in differences approach that exploits reforms in the provision of childcare (see for instance Baker et al. [2008], Havnes and Mogstad [2011] and Felfe et al. [2014]), a regression discontinuity design around the admission threshold Fort et al. [2019]) or instrumental variable methods that exploit some sources of exogenous variation in the probability of individual childcare attendance (Del Boca et al. [2016a], Felfe and Huber [2017], and Drange and Havnes [2019]). 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 determining children's probability of attending the care service. In a context where childcare provision is highly rationed, childcare availability clearly affects the probability of individual 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 early childcare attendance

---

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

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

in Italy is mostly driven by the supply - rather than the demand - for the service. However, our instrument does not affect parents who are not interested in enrolling their children in the early childcare service.

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

$$ECC_{imt} = \beta_0 + \beta_1 COVER_{mt} + \mathbf{X'}_{imt}\boldsymbol{\beta_2} + \boldsymbol{\Phi'}_{mt}\boldsymbol{\beta_3} + \eta_t + \lambda_{im} + \eta_t \times \lambda_{im} + \omega_{imt} \qquad (2)$$

where $COVER_{mt}$ is the public childcare coverage at the municipal level at time t and ω_{imt} is the error term.

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. [2016a], which instrumented early childcare with the regional supply of childcare).

In Figure 6, 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. The existence of a positive linear relationship between the two variables clearly emerges, which suggests that our instrument is valid. In our case, it implies that there are at least some parents whose enrolment decision is affected by the public early childcare supply of their 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 first stage F-statistics. Overall, our instrument proves to capture relevant variation in early childcare attendance, the endogenous variable.

Figure 6: First stage graphical representation.
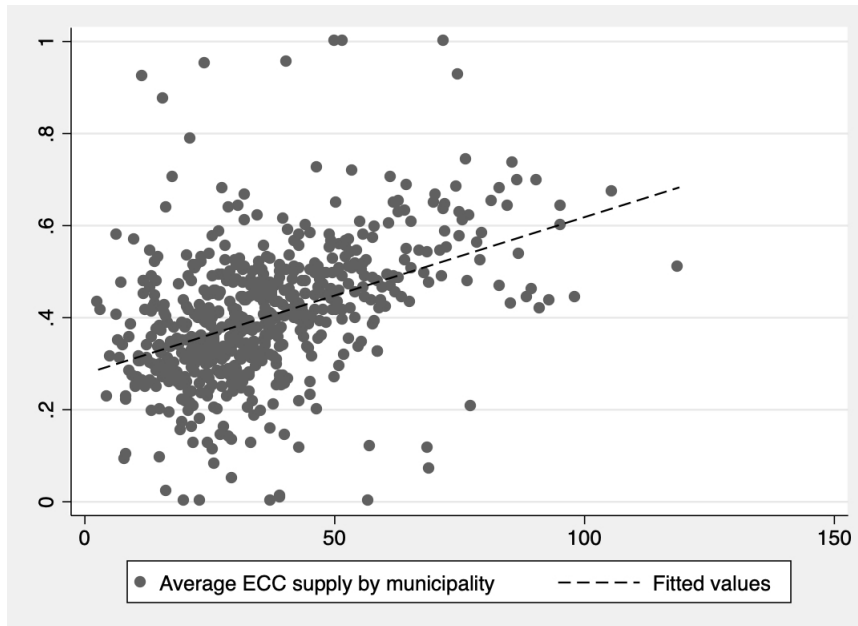
Notes: each dot in the figure shows the early childcare supply (ECC) by municipality.

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. The exclusion restriction implies that the variation in public childcare availability across municipalities and over time affects children's outcomes only through their childcare attendance. In other words, the public supply of early childcare spots must not have a direct effect on children's cognitive outcomes and should be uncorrelated with any factors determining children's early childcare attendance and/or cognitive outcomes. While it seems plausible to assume that the public supply of early childcare does not have a direct effect on cognitive outcomes, 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 why the exclusion restriction might not 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 each point in turn.

Regarding the first concern, we cannot include municipalities' fixed effect because, since we are using only three adjacent years, our instrument does not exhibit enough within-municipalities over-time variation. However, we believe this is not problematic in our case for the following reasons. First, we include a rich set of controls at the municipality level in order to capture several dimensions of heterogeneity between municipalities at the time when children in our sample were born. In particular, we include in our specification a time-varying measure of taxable income [13], an indicator of mayors' political faction, and population size. Second, we include in the model province fixed effect to capture all unobservable characteristics shared by municipalities within a fairly homogeneous area. Furthermore, we add province by cohort fixed effects to account for differential trends in childcare provision across provinces and for any other province-specific, time-varying shock. It is also important to notice that since we are only considering municipalities above 10,000 inhabitants, we can reasonably expect that within narrow geographical areas such as provinces, municipalities with similar size will have comparable social, economic and cultural characteristics. Therefore, we argue that, once we include the above-mentioned controls, the remaining variation in the instrument can be considered as good as random.

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 were in favour of more inclusive policies.[14] 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 into different school environments as well as in more (less) inclusive policies for immigrant children during early childhood. Controlling for the mayor's political faction in our specification is also intended to capture these potentially relevant differences across municipalities. In particular, we control for the municipalities' political orientation, looking at whether the incumbent mayor was supported by a left or a right party when elected. The incumbent mayor's political faction is defined as a dummy variable coded equal to 1 if the mayor

---

[13]Data at the municipal level for the years 2005-2007 are drawn from the Ministry of Economy and Finance "Analisi Statistiche - Open data dichiarazioni" www.finanze.gov.it

[14]For further details see Bracco et al. [2018]. Elaborated data in this paper are kindly provided by the authors.

was endorsed by left parties, or equal to 0 if they were supported by right parties.[15]

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 immigrants' location decision are employment opportunities in the area and the presence of co-ethnic immigrants, while we regard the availability of early childcare spots not to play a crucial role in their decision (see for example Damm [2009]).

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

6 Results

6.1 The effect of early childcare attendance on standardised test scores

The results of our main specification are reported in Table 2, which shows the impact of early childcare attendance on INVALSI standardised test scores in language (columns 1-4) and math (columns 5-9) for both native and immigrant students. In all specifications, we control for province and cohort fixed effects, for province linear time trends as well as for all control variables discussed in section 5. For the sake of brevity, we do not include control variables' coefficients in our tables of results. 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 does 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.

Focusing on our variable of interest, OLS estimates suggest that early childcare is associated with a higher performance of natives in both language and math (columns 1 and 4). For immigrant students, instead, early childcare seems to have no significant impact on language test scores (column 3) and a negative effect on math performance (column 7). Given the endogeneity of childcare attendance, these OLS estimates cannot be interpreted as causal. The positive coefficients of childcare attendance on test scores of natives simply reflects the positive selection into childcare services already observed in Table 1. Parents with higher education are in fact more likely to be employed in high-paying jobs; therefore, they tend to have higher incentives to enrol their children in early childcare and are more likely to obtain a slot if employment status constitutes a preferential criterion for access to public childcare. At the same time, children of more educated 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 the early childcare service. This non-random sorting of children into childcare lead the OLS estimates to be upward biased. For immigrant students, the negative sign of the childcare coefficient on math test scores is less straightforward to interpret, and suggests a negative selection into early childcare, probably caused by unobservable characteristics that we are not able to control.

Once we take the endogeneity of childcare into account by using an IV approach that exploits the exogenous variation in childcare supply across municipalities and over time, we find that early childcare attendance has a negative and significant impact on both language and math test scores of native students. This finding is consistent with part of the literature on early interventions that

---

[15]In 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.

document negative impacts on native children taking part in early childhood programmes (Fort et al. [2019], Herbst [2017], Havnes and Mogstad [2015], Herbst [2013], Cascio and Schanzenbach [2013]). In the next section, we better characterise this evidence by investigating which group of native students is driving the effect and how our results relate to the existing literature. Turning to immigrant students, interestingly, our IV estimates suggest a positive and significant impact of early childcare on language test scores, while no effect is detected on math test scores (see columns [4] and [8]). We believe that this finding is particularly thought-provoking and relevant from a policy perspective, as it suggests that early childcare may be a useful instrument for improving linguistic skills and thus for facilitating the integration of immigrant children. In terms of magnitude, our estimates indicate that early childcare utilisation increases language test scores of immigrant children by about 70% of a standard deviation. To appreciate the size of this effect, consider that one standard deviation in immigrant test scores corresponds to 17 percentage points in the original variable that gives the percentage of correct answers (mean value=61). Therefore, an increase by about 0.7 standard deviation (an increase of about 12 percentage points) is similar to moving language test scores from the median (score=63) to the 75th percentile (score=75). Twelve percentage points is also almost the difference between the median values of language test scores for second-generation immigrant (53.8) and native (64.7) students, which suggests that the effect is sizable and almost capable of closing the linguistic gap between second-generation immigrants and natives.

In Panel B, we report first stage estimates, which show that the public municipal supply of early childcare strongly determines children's service attendance, 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.

17

Table 2: Effects of early childcare attendance on children's test scores at grade 5.

	LANGUAGE				MATH			
	Natives		Immigrants		Natives		Immigrants	
Panel A: Second Stage								
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Early childcare	0.018***	-0.261***	0.001	0.671**	0.042***	-0.570***	-0.039***	0.155
	(0.003)	(0.101)	(0.011)	(0.266)	(0.003)	(0.108)	(0.011)	(0.253)
Observations	333,241	333,241	29,575	29,575	333,241	333,241	29,575	29,575
Panel B: First Stage								
		(2) WS		(4) WS		(6) WS		(8) WS
Early childcare public supply		0.002***		0.002***		0.002***		0.002***
		(0.000)		(0.000)		(0.000)		(0.000)
First stage F-stat		356.04***		65.34***		356.04***		65.34***
Observations		333,241		29,575		333,241		29,575
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Municipality controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Province by cohort FE	✓	✓	✓	✓	✓	✓	✓	✓

Notes: Entries in the Table are the estimated coefficients of early childcare attendance on INVALSI standard-ised language (columns [1]-[4]) and math (columns [5]-[8]) test scores. Columns (1),(3),(5) and (7) report OLS estimates, while columns (2),(4),(6) and (8) report IV estimates. Columns (1)-(2) and (5)-(6) report estimates for native students, and columns (3)-(4) and (7)-(8), for immigrant students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

6.2 Heterogeneous effects

Results reported in Table 2 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 children from various backgrounds experience early childcare attendance differently. In particular, we investigate whether the impact differs by children's social and economic background and by gender.

In Table 3 and 4, we report IV estimates of the impact of early childcare on INVALSI language and math test scores for children with high- and low-educated mothers, respectively, for native and immigrant children. We define a mother as high educated if she has an educational level equal to or higher than a high school diploma. The mother's educational level serves as a proxy for children's social and economic background as well as the 'quality' of the most likely care alternative to early childcare attendance: the mother. Table 3 suggests that parental background does not play a role in shaping the effect of childcare programmes on language test scores of native students, but interestingly highlights that, for immigrant students, the impact of early childcare is not homogeneous 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 immigrant children with low-educated mothers, and the size of the effect is higher than that for the whole sample of migrants (see Column [6]). [16] These results are in line with previous studies

---

[16] Our results are consistent also when using the father's educational level or the highest between mothers and fathers as a proxy for the child's social and cultural background.

emphasising 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, this due to worse and less stimulating home environments (see Havnes and Mogstad [2011], Drange and Havnes [2019], Felfe and Lalive [2018], Cornelissen et al. [2018]).

Table 3: Heterogeneous effects by mother' educational level, language.

	LANGUAGE					
	Natives			Immigrants		
Panel A: Second Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample	High Edu	Low Edu	Whole Sample	High Edu	Low Edu
Early childcare	-0.261***	-0.166	-0.158	0.671**	0.311	1.096**
	(0.101)	(0.106)	(0.224)	(0.261)	(0.287)	(0.508)
Observations	333,241	265,804	67,437	29,575	17,894	11,681
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002***	0.002***	0.002***	0.002***	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
First stage F-stat	356.04***	303.20***	89.25***	65.34***	49.28***	21.98***
Observations	333,241	265,804	67,437	29,575	17,894	11,681
Individual controls	✓	✓	✓	✓	✓	✓
Municipality controls	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓
Province by cohort FE	✓	✓	✓	✓	✓	✓

Notes: the table shows IV estimates in which the dependent variable is the INVALSI standardised language test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant students. Columns (2) and (5) refer to students with a high-educated mother (i.e. education level higher than a lower secondary school diploma), while columns (3) and (6) refer to students with a low-educated mother (i.e. educational level lower than or equal to a lower secondary school diploma). All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and father's educational level. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

In Table 4, we report the estimates of the impact of early childcare attendance on math test scores for children with a high- (low-) educated mother. Focusing on immigrant students (columns [4]-[6]), we find no effect on math test scores, regardless of the mother's educational level. Interestingly, the fact that the results we observe for language are not mirrored for math test scores suggests that for immigrant children, early childcare is likely to serve as a device for facilitating the learning of the host country's language. We investigate this hypothesis at length in subsection 6.4.

We now turn our attention to native students. Our results show that the average negative effect of childcare attendance on math test scores discussed above is due entirely to the larger negative impact on students from a highly educated background, while no effect is found for children with low-educated mothers (see columns [2] and [3]). This finding is consistent with evidence from previous studies that early childcare is more likely to adversely affect advantaged children from relatively affluent families able to provide a more stimulating context or to have access to high quality substitutes of formal childcare. In particular, in the Italian context, Fort et al. [2019] find similar results, adopting a quasi-experimental approach that exploits discontinuities in the admission thresholds in a regression discontinuity design, and thus showing that early childcare has a negative effect on cognitive outcomes of children from wealthier backgrounds.

Table 4: Heterogeneous effects by mother' educational level, math.

MATH

	Natives			Immigrants		
Panel A: Second Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
	Whole Sample	High Edu	Low Edu	Whole Sample	High Edu	Low Edu
Early childcare	-0.570***	-0.538***	-0.219	0.155	0.254	0.000
	(0.108)	(0.114)	(0.227)	(0.253)	(0.287)	(0.446)
Observations	333,241	265,804	67,437	29,575	17,894	11,681
Panel B: First stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002***	0.002***	0.002***	0.002***	0.003***	0.002***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.000)
First stage F-stat	356.04***	303.20***	89.25***	65.34***	49.28***	21.98***
Observations	333,241	265,804	67,437	29,575	17,894	11,681
Individual controls	√	√	√	√	√	√
Municipality controls	√	√	√	√	√	√
Cohort FE	√	√	√	√	√	√
Province FE	√	√	√	√	√	√
Province by cohort FE	√	√	√	√	√	√

Notes: the table shows IV estimates in which the dependent variable is the INVALSI standardised math test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant students. Columns (2) and (5) refer to students with a high-educated mother (i.e. educational level of the mother is higher than a lower secondary school diploma), while columns (3) and (6) refer to students with a low-educated mother (i.e. education level of the mother lower than or equal to a lower secondary school diploma). All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and father's educational level. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

We have so far neglected heterogeneity by gender, which we investigate now by running separate regressions for males and females, for both natives and immigrants. The impact of early childcare could differ by gender because of developmental differences of boys and girls in early childhood, due to both biological and social processes. In fact, the psychological literature suggests that, at the same age, girls tend to outperform boys in terms of language, social development and behaviour (see Magnuson et al., (2016) for a review). This advantage may allow girls to reap greater benefits in learning cognitive and behavioural skills from formal early childcare, but, at the same time, girls may also reap larger benefits from other childcare settings, such as stimulating parental care (see Del Boca et al. [2019] for an extensive discussion on the heterogeneous impact of childcare by gender).

Our results are reported in Table 5 for language test scores and in Table 6 for math test scores, and are totally in line with the predictions of the theories above. In fact, we find significant differences across gender in the impact of childcare on both language and math test scores. In particular, our results suggest that females tend to be more affected by childcare attendance. In fact, females are driving all of the results discussed above: both the positive effect on immigrants' test scores in language and the negative impact on natives' language and math test scores stem entirely from females, while for males, the effects are not statistically significant.

Table 5: Heterogeneous effects by gender, language.

LANGUAGE

	Natives			Immigrants		
Panel A: Second Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
	Whole sample	Male	Female	Whole sample	Male	Female
Early childcare	-0.261***	-0.248	-0.266**	0.671**	0.598	0.660**
	(0.101)	(0.154)	(0.132)	(0.266)	(0.442)	(0.325)
Observations	333,241	168,382	164,859	29,575	14,742	14,833
Panel B: First Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002***	0.001***	0.002***	0.002***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
First stage F-stat	356.04***	154.56***	205.08***	65.34***	22.99***	44.33***
Observations	333,241	168,382	164,859	29,575	14,742	14,833
Individual controls	√	√	√	√	√	√
Municipality controls	√	√	√	√	√	√
Cohort FE	√	√	√	√	√	√
Province FE	√	√	√	√	√	√
Province by cohort FE	√	√	√	√	√	√

Notes: the table shows IV estimates in which the dependent variable is the INVALSI standardised language test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant students. Columns (2) and (5) refer to male students, while columns (3) and (6) refer to female students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

Table 6: Heterogeneous effects by gender, math.

MATH

	Natives			Immigrants		
Panel A: Second Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
	Whole sample	Male	Female	Whole sample	Male	Female
Early childcare	-0.570***	-0.500***	-0.616***	0.155	-0.007	0.191
	(0.108)	(0.163)	(0.143)	(0.253)	(0.431)	(0.303)
Observations	333,241	168,382	164,859	29,575	14,742	14,833
Panel B: First Stage						
	(1)	(2)	(3)	(4)	(5)	(6)
Early childcare public supply	0.002***	0.001***	0.002***	0.002***	0.002***	0.003***
	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
First stage F-stat	356.04***	154.56***	205.08***	65.34***	22.99***	43.33***
Observations	333,241	168,382	164,859	29,575	14,742	14,833
Individual controls	√	√	√	√	√	√
Municipality controls	√	√	√	√	√	√
Cohort FE	√	√	√	√	√	√
Province FE	√	√	√	√	√	√
Province by cohort FE	√	√	√	√	√	√

Notes: the table shows IV estimates in which the dependent variable is the INVALSI standardised math test score. Columns (1)-(3) refer to native students, while columns (4)-(6) refer to immigrant students. Columns (2) and (5) refer to male students, while columns (3) and (6) refer to female students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, regularity of studies and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

Our findings highlight how the effect of early childcare attendance is not uniform across subgroups of the population. The positive impact is concentrated among immigrant children and, in particular, on their test scores in language. This effect is mostly driven by females and by students coming from a disadvantaged background. For this group of students, early childcare attendance is successful at narrowing the educational gap.

Crucially, our results emphasise that the impact of early childcare depends on the predetermined characteristics of children enrolled in the service: whether they are native or immigrant, whether they come from a disadvantaged background and whether they are male or female. The existence of countervailing effects on different parts of the population of students is in line with the existing literature. On the one hand, a number of studies investigating the impact on disadvantaged children find that early intervention has a positive impact on students' cognitive outcomes, with the effect estimates being more pronounced for girls (Drange and Telle [2015], Felfe et al. [2014], Havnes and Mogstad [2011] Anderson [2008], among others). On the other hand, the findings of several studies analysing untargeted early interventions or the expansion of early childcare access report negative or no impact on native students, particularly if they come from an advantaged background (Baker et al. [2008], Herbst [2017], Drange et al. [2016]). Relative to the recent literature on Italy, our results are again in line with Fort et al. [2019], showing detrimental effects on native children's cognitive skills, with the effect being stronger for females.

Why is early childcare more beneficial (detrimental) to immigrant (native) children, and why does it have a stronger impact on female students? As we argue at further length in following sections, the positive effect of early childcare service attendance on immigrant students' language

22

test scores is likely to mirror the positive impact of early exposure to the host country's language. Unlike immigrants, the interpretation of our results for native students traces back to the alternative mode of care that children coming from an advantaged background might gain access to. In other words, results for native children speak to the quality of the early childcare service. In this context, some important quality indicators are the qualifications required of the staff, the adult-to-children ratio and the educational content provided by the service as a whole. In Italy, the minimum qualification for teachers/educational staff working with children aged 0-2 is set at the upper secondary level, while in countries with a unitary setting, they are usually required to be qualified at the tertiary level [17]. In these countries, the adult-to-children ratio is not set by a national standard and is relatively low (on average, about 1:6, varying across regions) compared to northern European countries, where it runs between 1:3–1:5, and the service is mostly care-oriented and without a clear and uniform educational content (Eurydice [2014]). These institutional features might explain our results for native students who might experience losses due to the lower quality of the service, that is, relative to the alternative modes of care they might enjoy at home.

Concerning the underlined heterogeneity by gender, the psychological literature has emphasised girls' greater capability of reaping the benefits from cognitive stimuli and early interaction with adults at an early stage of life (see Fenson et al. [1994], Bornstein et al. [2004]). Females' greater sensitivity to early interaction might arguably explain both the larger benefit we observe for immigrant female students exposed earlier to the destination country's language and the losses for native females who are provided less one-on-one interaction and a less stimulating environment compared to the one they would be exposed to at home.

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 school year in which the INVALSI tests take place. We repeat the same analysis performed for INVALSI test scores but consider 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 the following tables, we investigate the impact of early childcare attendance on language and math oral marks. In Table 7, we report results by subject (columns [1]-[4], language; columns [5]-[8], math) and immigrant status (columns [1]-[2] and [5]-[6], native children, and columns [3]-[4] and [7]-[8], immigrant children). For both language (columns [1] and [3]) and math (columns [5] and [7]), we report baseline estimates for INVALSI standardised test scores for ease of comparison. Our results show that, on average, early childcare has a positive and significant impact on language and math oral marks of immigrant children. The positive impacts of early childcare attendance do not vary by mother's educational level and child's gender, neither for language nor for math oral marks assigned by teachers. [18] Interestingly, no effect is detected on native children's oral marks both on average and by child's gender and mother's education.

---

[17]Unitary settings are those where pre-primary education is organised in a single phase in one setting, i.e. Nordic countries, Slovenia, Estonia and Lithuania

[18]For the sake of brevity, we do not report tables for the heterogeneity by gender and mother's educational level.

For immigrant students, evidence on oral marks strongly suggests that greater fluency in the host country's language is the channel driving the observed positive effects. In fact, results are consistent with those in the previous section analysing the impact of early childcare on INVALSI standardised test scores. However, differing from before, our coefficients turn highly significant for math oral marks as well. The latter is likely not related to children's higher ability in math but to their oral exposure competences that strongly depend on their language fluency and comprehension.

Table 7: Effects of early childcare attendance on oral marks assigned by teachers.

	LANGUAGE ORAL MARKS				MATH ORAL MARKS			
	Natives		Immigrants		Natives		Immigrants	
Panel A: Second Stage								
	(1) Invalsi std score	(2) Oral marks	(3) Invalsi std score	(4) Oral marks	(5) Invalsi std score	(6) Oral marks	(7) Invalsi std score	(8) Oral marks
Early childcare	-0.261***	-0.020	0.671**	1.030***	-0.570***	-0.035	0.155	0.878***
	(0.101)	(0.094)	(0.266)	(0.273)	(0.108)	(0.096)	(0.253)	(0.279)
Observations	333,241	333,241	29,575	29,575	333,241	333,241	29,575	29,575
Panel B: First Stage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare public supply	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
First stage F-stat	356.04***	356.04***	65.34***	65.34***	356.04***	356.04***	65.34***	65.34***
Observations	333,241	333,241	29,575	29,575	333,241	333,241	29,575	29,575
Individual controls	√	√	√	√	√	√	√	√
Municipality controls	√	√	√	√	√	√	√	√
Cohort FE	√	√	√	√	√	√	√	√
Province FE	√	√	√	√	√	√	√	√
Province by cohort FE	√	√	√	√	√	√	√	√

Notes: entries in the Table are the estimated coefficients of early childcare attendance on standardised language (columns [1]-[4]) and math (columns [5]-[8]) oral marks assigned by teachers. Columns (1), (3), (5) and (7) report INVALSI standardised test score estimates, while columns (2), (4), (6) and (8) report oral mark estimates. Columns (1)-(2) and (5)-(6) report estimates for native students, while columns (3)-(4) and (7)-(8) report estimates for immigrant students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

6.4 The role of language proximity

Up to now, we found that childcare attendance improves the test scores of children with immigrant backgrounds. This result may be explained by how childcare provides second-generation immigrants with early exposure to their host country's language, which is often different from the language spoken at home. In order to understand whether this is indeed the channel at work, we investigate whether the linguistic distance of the immigrant children's mother tongue to the Italian language plays a role in shaping the effect of childcare. If our hypothesis that early childcare improves immigrants' school achievement by improving their language and communication skills is correct, we should expect that children whose parents speak a language significantly different from Italian will benefit the most from early childcare attendance, as the context will grant them greater exposure to the new language.

In order to compute linguistic distance from Italian, we employ the continuous linguistic distance measure developed by the Max Planck Institute for Evolutionary Anthropology. This measure evaluates the phonetic dissimilarities between words in two languages, called the Levenshtein linguistic distance. The intuition behind the construction of this measure is to compare pairs of words on the basis of their pronunciation. In particular, a key set of 40 words, the so-called *Swadesh word list* (Swadesh [1952]) that contain terms of everyday life common to most of the world's languages, is transcribed using a phonetic alphabet, the Automatic Similarity Judgement Program. Then, for each word, the number of sounds that have to be changed (in order to express one language's word in another language) is calculated (Isphording and Otten [2013]).[19]

The Levenshtein linguistic distance has been previously exploited to investigate the impact of linguistic dissimilarity on immigrant destination country language acquisition (Isphording and Otten [2013], Isphording and Otten [2014]), migration determinants Adsera and Pytlikova [2015] and the heterogeneous impact of immigrant peers on native students' achievements in the classroom Frattini and Meschi [2019]. However, to the best of our knowledge, no other studies on early childhood development have focused on this relevant dimension.

INVALSI data contain unique information about the language children mostly speak at home. We merge this information to the Levenshtein linguistic distance measure.[20] More specifically, we restrict the sample to second-generation immigrant students, and we match individual information about the language mostly spoken at home with the Levenshtein measure. In our sample, the index ranges between 0 and 100, and has fourteen unique values.

Using the same specification as in section 6, we run separate regressions for second-generation immigrant students whose language is more or less distant from Italian based on the Levenshtein linguistic distance being above (below) the median in our sample.

Results are reported in Table 8. Columns (1)-(2) ([3]-[4]) show results for children speaking a language with a low (high) level of dissimilarity with respect to Italian, respectively, for INVALSI language and math standardised test scores. The estimates confirm our expectations and reveal that the positive effect of childcare on language test scores hold only for children speaking a language highly distant from Italian, whereas no effect is detected on math test scores and on children speaking a language relatively close to Italian. When we repeat the analysis by gender[21], we see that the latter result is again mainly driven by the effect on females. Although not statistically significant (presumably because of a very low sample size), results by linguistic distance and mother's educational level (not reported here for the sake of brevity) are consistent with the one in the main analysis pointing to a positive effect of early childcare attendance on children from less favourable backgrounds and speaking a language distant from Italian.

These findings reveal that the mechanism through which the effect might operate consists in a reduction in the cost of language acquisition due to students' early exposure to the destination country's language. As previously discussed, immigrant females reap the largest benefit from early childcare service attendance. This latter piece of evidence adds to our previous results, showing that among females, those speaking a language strongly dissimilar from Italian benefit the most from early childcare service attendance. Consistent with our argument, no effect is detected for both male and female students who speak a language closer to Italian.

Taken together, these results are consistent and add to those in previous sections. Namely, early childcare attendance exerts positive effects on language outcomes of second-generation immigrant

---

[19]For further details on the measure construction, see Bakker et al. [2009].

[20]INVALSI data do not provide immigrant children's country of origin.

[21]The results' table is not reported here but is available upon request.

students, with the effect concentrated on children speaking a language strongly dissimilar to Italian and on females.

Table 8: Effects of early childcare attendance on children's test scores by child's linguistic distance.

	LOW LINGUISTIC DISTANCE		HIGH LINGUISTIC DISTANCE	
	Language	Math	Language	Math
Panel A: Second Stage				
	(1)	(2)	(3)	(4)
Early childcare	0.046	-0.015	0.842*	0.109
	(0.309)	(0.313)	(0.488)	(0.450)
Observations	13,596	13,596	11,822	11,822
Panel B: First Stage				
	(1)	(2)	(3)	(4))
Early childcare public supply	0.003***	0.003***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)
First stage F-stat	40.78***	21.23***	40.78***	21.23***
Observations	13,596	13,596	11,822	11,822
Individual controls	√	√	√	√
Municipality controls	√	√	√	√
Cohort FE	√	√	√	√
Province FE	√	√	√	√
Province by cohort FE	√	√	√	√

Notes: the table shows IV estimates in which the dependent variable is the INVALSI language test score. Columns (1)-(3) refer to immigrant students speaking a language with a low level of linguistic distance vis-à-vis Italian (LLD). Columns (4)-(6) refer to immigrant students speaking a language with a high level of linguistic distance vis-à-vis Italian (HLD). All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, studies regularity and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

7 Robustness checks

This section presents a set of robustness checks of our main findings. First, we augment our specification (see section 5) with a number of control variables at the class and school levels. As discussed above, our specification allows for a qualified number of controls that allow us to identify the causal effect of early childcare attendance on standardised test scores. However, one might consider including a more extensive set of controls, taking into account school and class compositional differences in order to investigate whether their inclusion has a substantial impact on the estimated coefficient. In particular, we control for school and class size and their square, and the share of female, immigrant and retained students and low-educated parents[22] per class.

Second, 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 characterised by a high level of internal homogeneity. However, the provincial capital might differ economically, culturally and socially 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, which would potentially benefit children's test outcomes. In order to

---

[22]The share of poorly educated parents is defined as the share of parents with an educational level lower than or equal to a lower secondary school diploma per class.

address this potential threat to identification, we control in our specification also for a dummy equal to 1 if the municipality is a provincial capital, and 0 otherwise.

Estimates are reported in Table 9 for language test scores and in Table 10 for math test scores. For ease of comparison, in columns (1) and (4) of both tables, we report baseline estimates from Table 2 for native and immigrant students, respectively. In each table, columns (2) and (5) control for school and class characteristics, columns (3) and (6) control for the municipality being the provincial capital and columns (4) and (8) include all previous controls in the same specification. Reassuringly, augmenting our specification with additional controls, results are remarkably similar to those in Table 2. It is worth noting that estimates in columns (7) and (8) point in the direction of an even stronger positive impact of early childcare on immigrant children's cognitive outcomes.

Later we investigate whether our main findings concerning language test score outcomes for immigrant children by mother's educational level are sensitive to the usage of alternative proxies of students' socio-economic backgrounds. Specifically, we investigate whether using the father's educational level has an impact on the estimated coefficients. Again, results are almost identical in terms of both significance and magnitude.

In sum, robustness checks yield similar results to the one in the main analysis, confirming the overall pattern of our results.

Table 9: Effects of early childcare attendance on children's test scores at grade 5, including additional controls.

	LANGUAGE							
	Natives				Immigrants			
Panel A: Second Stage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare	-0.261***	-0.297***	-0.244**	-0.250**	0.671**	0.645**	1.067**	0.924*
	(0.101)	(0.107)	(0.124)	(0.120)	(0.266)	(0.292)	(0.494)	(0.484)
Observations	333,241	333,241	333,241	333,241	29,575	29,575	29,575	29,575
Panel B: First Stage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare public supply	0.002***	0.001***	0.001***	0.001***	0.002***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
First stage F-stat	356.04***	321.23***	235.66***	252.19***	65.34***	53.75***	21.79***	21.47***
Observations	333,241	333,241	333,241	333,241	29,575	29,575	29,575	29,575
Individual controls	✓	✓	✓	✓	✓	✓	✓	✓
Municipality controls	✓	✓	✓	✓	✓	✓	✓	✓
Cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓	✓	✓
Province by cohort FE	✓	✓	✓	✓	✓	✓	✓	✓
Class and school controls		✓		✓		✓		✓
Provincial Capital			✓	✓			✓	✓

Notes: entries in the Table are the estimated coefficients of early childcare attendance on INVALSI standard-ised language test scores. Columns (1)-(4) report estimates for native students, and columns (5)-(8), for immigrant students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, preschool attendance, regularity of studies and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. Class and school controls include: school and class size and their square, and the share of female, immigrant and retained students and low-education parents per class. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

Table 10: Effects of early childcare attendance on children's test scores at grade 5, including additional controls.

	MATH							
	Natives				Immigrants			
Panel A: Second Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare	-0.570***	-0.615***	-0.523***	-0.538***	0.155	-0.019	0.011	-0.205
	(0.108)	(0.114)	(0.132)	(0.128)	(0.253)	(0.278)	(0.437)	(0.441)
Observations	333,241	333,241	333,241	333,241	29,575	29,575	29,575	29,575
Panel B: First Stage	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Early childcare public supply	0.002***	0.001***	0.001***	0.001***	0.002***	0.002***	0.001***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
First stage F-stat	356.04***	321.23***	235.66***	252.19***	65.34***	53.75***	21.79***	21.47***
Observations	333,241	333,241	333,241	333,241	29,575	29,575	29,575	29,575
Individual controls	√	√	√	√	√	√	√	√
Municipality controls	√	√	√	√	√	√	√	√
Cohort FE	√	√	√	√	√	√	√	√
Province FE	√	√	√	√	√	√	√	√
Province by cohort FE	√	√	√	√	√	√	√	√
Class and school controls		√	√	√		√	√	√
Provincial Capital			√	√			√	√

Notes: entries in the Table are the estimated coefficients of early childcare attendance on INVALSI standardised math test scores. Columns (1)-(4) report estimates for native students, and columns (5)-(8), for immigrant students. All specifications include province, cohort and province dummies interacted with a time trend. Individual-level control variables include: gender, month of birth, attendance to preschool, studies regularity and parents' educational levels. Municipality-level control variables include: taxable income, resident population and the incumbent mayor's political orientation. Class and school controls include: school and class size and their square, and the share of female, immigrant and retained students and low-educated parents per class. See Table 11 in Section 9 for variable definitions and sources. Robust standard errors are given in parentheses; ***p<0.01, **<0.05, *p<0.1.

8 Conclusions

The successful integration of immigrant children into the educational system is one of the most important challenges for the majority of European countries. Children who are migrants or who have immigrant parents face significant educational barriers, which may lead to substantial educational disadvantages relative to native students. Performance gaps at school then translate into persistent differences between immigrant students and their native peers in terms of literacy levels, drop-out rates and, eventually, in labour market outcomes, thus perpetuating inequalities and slowing the integration of migrants into 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 a very young age is an effective policy to close or at least reduce the existing performance gap between immigrant and native children. To carry out our analysis, we drew on novel Italian administrative data and made use of a neat IV identification strategy in order to deal with children's non-random sorting into early childcare.

Our results point to a positive and significant impact of early childcare attendance on immigrant children's cognitive outcomes, with the effect driven by females and by students with low-educated mothers. Further, we find that the impact of childcare is stronger for immigrant

children whose mother tongue is further from Italian, which strongly suggests that greater fluency in the host country language might well be the mechanism through which the effect operates. Then, access to early childcare may be a relevant policy tool to improve the language proficiency of immigrant children, thus fostering their integration and assimilation into the host country.

On the other hand, our estimates report a negative and significant effect on native students with regard to both language and math test scores. The detrimental effects are mostly concentrated among females and among students coming from more advantaged backgrounds. These findings might be interpreted in light of the relatively better alternative modes of care that native students might gain access to, and call policy makers' attention to the educational content and organisation of early childcare services.

Our findings provide empirical support to the renewed interest of social scientists and policy makers in early childcare and pre-primary school program attendance as an effective tool for facilitating immigrant children's integration and for closing the gap in their educational performance.[23] Unfortunately, despite the encouraging empirical evidence and the rich policy debate, the access of children with immigrant backgrounds to early childcare facilities (and other pre-primary school programs) is still very limited in OECD countries (including Italy, the country of origin for the data used in this study). [24] 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 across families. In some cases, this institutional situation has translated into inefficient and discretionary allocation rules, such as those restricting access to early childcare programs to families who satisfy time requirements with their legal residence in the municipality.[25] 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 into the education system of the host country.

---

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

[24]For Europe, see CoE [2017]. According to OECD [2015], first-generation immigrant students are almost half as likely as non-immigrant students to attend pre-primary education; there are, however, 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.

[25]For instance, in 2018, Veneto introduced the requirement of 15 years of legal residence in the municipality in order to obtain access to early childcare programs. However, this requirement was declared unconstitutional a few months after its introduction.

References

Adsera, A. and Pytlikova, M. (2015). The role of language in shaping international migration. *The Economic Journal*, 125(586):F49–F81.

Ahad, A. and Benton, M. (2018). Mainstreaming 2.0. Technical report, Brussels: Migration Policy Institute Europe.

Alesina, A., Carlana, M., La Ferrara, E., and Pinotti, P. (2018). Revealing stereotypes: Evidence from immigrants in schools. Working Paper 25333, National Bureau of Economic Research.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Antonelli, M. A. and Grembi, V. (2011). *Target centrali e finanza locale: Il caso degli asili nido in Italia*. Carocci editore.

Antonelli, M. A. and Grembi, V. (2014). Central targets and local agendas: Missing lisbon 2010. Public finance research papers 6, Available at SSRN 2161985.

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, M., Gruber, J., and Milligan, K. (2019). The long-run impacts of a universal child care program. *American Economic Journal: Economic Policy*, 11(3):1–26.

Bakker, D., Müller, A., Velupillai, V., Wichmann, S., Brown, C. H., Brown, P., Egorov, D., Mailhammer, R., Grant, A., and Holman, E. W. (2009). Adding typology to lexicostatistics: A combined approach to language classification. *Linguistic Typology*, 13(1):169–181.

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, D. and Currie, J. (2006). Pre-school, day care, and after-school care: who's minding the kids? 2:1163–1278.

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, 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(C):12–26.

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, F. and Rizzica, L. (2018). Early kindergarten, maternal labor supply and children's outcomes: Evidence from italy. *Journal of Public Economics*, 158(C):79–102.

Cascio, E. and Schanzenbach, D. W. (2013). The Impacts of Expanding Access to High-Quality Preschool Education. *Brookings Papers on Economic Activity*, 44(2):127–192.

CoE (2017). Fighting school segregation in europe through inclusive education: a position paper. Technical report, Concil of Europe Commissioner for Human Rights.

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, F. and Heckman, J. (2007). The technology of skill formation. *American Economic Review*, 97(2):31–47.

Damm, A. P. (2009). Determinants of recent immigrants' location choices: Quasi-experimental evidence. *Journal of Population Economics*, 22(1):145–174.

Del Boca, D., Martino, E. M., Meroni, E. C., and Piazzalunga, D. (2019). Early Education and Gender Differences. *Economia Italiana*, 3:11–36.

Del Boca, D., Pasqua, S., and Suardi, S. (2016a). Child care, maternal employment, and children's school outcomes. an analysis of italian data. *European Journal of Population*, 32(2):211–229.

Del Boca, D., Pronzato, C., and Sorrenti, G. (2016b). When Rationing Plays a Role: Selection Criteria in the Italian Early Childcare System. *CESifo Economic Studies*, 62(4):752–775.

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, N. and Havnes, T. (2019). Early childcare and cognitive development: Evidence from an assignment lottery. *Journal of Labor Economics*, 37(2):581–620.

Drange, N., Havnes, T., and Sandsør, A. M. (2016). Kindergarten for all: Long-run effects of a universal intervention. *Economics of Education Review*, 53(C):164–181.

Drange, N. and Telle, K. (2015). Promoting integration of immigrants: Effects of free child care on child enrollment and parental employment. *Labour Economics*, 34(C):26–38.

Dustmann, C., Frattini, T., and Lanzara, G. (2012). Educational achievement of second-generation immigrants: an international comparison. *Economic Policy*, 27(69):143–185.

Eurydice (2014). Key data on early childhood education and care in europe. 2014 edition. eurydice and eurostat report. Technical report, Luxembourg: Publications Office of the European Union.

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)*, 180(2):475–502.

Felfe, C. and Lalive, R. (2018). Does early child care affect children's development? *Journal of Public Economics*, 159(C):33–53.

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, 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*, 59:1–173.

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*, 128(1):158–205.

Frattini, T. and Meschi, E. (2019). The effect of immigrant peers in vocational schools. *European Economic Review*, 113(C):1–22.

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, 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, T. and Mogstad, M. (2015). Is universal child care leveling the playing field? *Journal of Public Economics*, 127(C):100–114.

Heckman, J. and Masterov, D. V. (2007). The productivity argument for investing in young children. *Applied Economic Perspectives and Policy*, 29(3):446–493.

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(C):86–105.

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.

Isphording, I. E. and Otten, S. (2013). The costs of babylon—linguistic distance in applied economics. *Review of International Economics*, 21(2):354–369.

Isphording, I. E. and Otten, S. (2014). Linguistic barriers in the destination language acquisition of immigrants. *Journal of economic Behavior & organization*, 105(C):30–50.

OECD (2012). Untapped skills. realizing the potential of immigrant students. Technical report.

OECD (2015). Help immigrant students to succeed at school and beyond. Technical report.

OECD (2018). Working together for local integration of migrants and refugees. Technical report.

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.

Swadesh, M. (1952). Lexico-statistic dating of prehistoric ethnic contacts: With special reference to north american indians and eskimos. *Proceedings of the American Philosophical Society*, 96(4):452–463.

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.

Declaration of interest

None.

Acknowledgments

Funding

9 Appendix

Table 11: Definition of variables and sources.

Variable	Description	Source
Outcome variables		
Language test score	Percentage of correct answers in Language test	INVALSI
Math test score	Percentage of correct answers in Math test	INVALSI
Language oral grade	Oral grade got past school year	INVALSI
Math oral grade	Oral grade got past school year	INVALSI
Childcare variable		
Childcare coverage	Public childcare slots over pop 0-2 by municipality	Antonelli and Grembi [2011]
Endogenous variables		
Childcare attendance	dummy=1 if the child attended to early childcare	INVALSI
Individual level control variables		
Female	dummy=1 if female	INVALSI
Immigrant I generation	dummy=1 if child is Immigrant II generation	INVALSI
Immigrant II generation	dummy=1 if child is Immigrant I generation	INVALSI
Age	child year of birth	INVALSI
II quarter month of birth	dummy=1 if the child is born in second quarter	INVALSI
III quarter month of borth	dummy=1 if the child is born in third quarter	INVALSI
Low level of paternal/maternal education	dummy=1 if mother/father has an educational level lower or equal to lower secondary school	INVALSI
Medium level of paternal/maternal educ.	dummy=1 if mother/father has an educational level higher than lower secondary school or equal to high school	INVALSI
High level of paternal/maternal educ.	dummy=1 if mother/father has an educ. level higher than high school	INVALSI
Mother's working status	dummy=1 if the mother works	INVALSI
Father's working status	dummy=1 if the father works	INVALSI
Language mostly spoken at home	dummy=1 if the child mostly speaks a language different from Italian at home	INVALSI
Pre-school attendance	dummy=1 if the child attended pre-school	INVALSI
Class and school level control variables		
Share of female	Share of female per class	INVALSI
Share of late students	Share of retained students per class	INVALSI
Share of migrants	Share of immigrants per class	INVALSI
Share of low-educated parents	Share of parents with an educational an educational level equal to or lower than lower secondary school diploma per class	INVALSI
Average class size	Average number of students per class	INVALSI
Average class size squared		INVALSI
School size	Total number of students per school	INVALSI
School size squared		INVALSI
Municipality level control variables		
Taxable income		Ministry of Economy and Finance
Resident population	Resident population in 2001	INVALSI
Incumbent mayor's political faction		Ministry of the Interior (for elaborated data see Bracco et al. [2018])