The impact of early childcare on children's socio-economical disadvantages

Candidate: Vittoria Dicandia

Università Commerciale Luigi Bocconi
MSc in Economic and Social Sciences

Supervisor: Prof. Paola Antonia Profeta

Discussant: Prof. Carlo Devillanova

Agli insegnanti, che svolgono un ruolo fondamentale

troppo spesso non riconosciuto,

e a mio padre, l'insegnante migliore che io conosca.

Desidero ringraziare la Fondazione "Roberto Franceschi", che scegliendo di finanziare il mio progetto di ricerca ha reso possibile la realizzazione di questo studio.
Grazie anche a tutte quelle persone ed istituzioni che hanno reso possibile la stesura di quest'elaborato, in particolare Dirigenti Scolatici, docenti e collaboratori degli Istituti Comprensivi di Mantova 1, Mantova 2 e Mantova 3, e il Kikolle Lab di Milano.

Un grazie sentito va alla Professoressa Paola Profeta, per la presenza costante e il supporto dimostrati durante lo svolgimento di questo lavoro, e non solo.
Vorrei ringraziare, infine, la Professoressa Pronzato per i preziosi consigli, e il Professor Devillanova per la sua disponibilità.

Contents

1 Introduction

The purpose of this study is to investigate the impact of early childcare on children achievements at the end of pre-school education and its role in the reduction of early age inequalities. This has been done distinguishing between outcomes regarding personal development, defined as degree of autonomy and ability to interact with others, and the ones related competences acquired, such as linguistic, logic and graphic abilities. The analysis will also explore whether the effects observed are only valid in absolute terms, or if they are stronger for children coming from disadvantaged backgrounds. The reasoning behind this research question comes from the main assumption, widely demonstrated in economic literature through the works of James J. Heckman and others, that early interventions do have an impact on achievements and technology of learning for future inputs. The absence of a systematic economic analysis in this field in Italy, mainly due to lack of data, represents itself compelling grounds for economic research in this field.

I have conducted a pilot study in the town of Mantova, evaluating children's outcomes and parents' and schooling inputs for children currently enrolled in the first year of primary school. I have collected new data in the form of comprehensive records regarding children's achievements at the end of the pre-school educational path, of information regarding the type of care received in their early years and their familiar background. The total number of 6-years-old children present in the selected municipality of Mantova amounts to 370, while the sample used for the analysis comprises 190 individuals.

The realization of this project has involved the creation of an original dataset, obtained thanks to the collaboration of primary schools, that provided records and had a "bridge" role in building a connection with the families of the children taken into consideration. For this kind of project to be practically feasible, it has been decided to focus on a single municipality and to choose a medium-sized town according to features related to population characteristics and early childcare diffusion. This choice allows to control for fixed effects related to territorial characteristic and to avoid issues with the sampling process that would have been much more relevant, had the analysed environment been larger and consequently required a random selection of the schools involved. The final aim,

hence, has been to develop a pilot project that can be reproduced in other municipalities and at a bigger scale, in order to provide a more complete and detailed picture of the impact of early childcare in the Italian context.
From a methodological standpoint, the project's main structure is composed by four distinct stages.

- The first step was to collect and analyse the evaluation records filled out by teachers of pre-primary education institutions at the end of the children's educational pre-schooling path. Given the high level of participation rate at this educational level, this represents the biggest pool from which it is possible to extract data on 5-years-old children's outcomes.

- The second phase consisted of the creation, delivery and collection of questionnaires addressed to children's families, with the aim of gathering information on the type of care received, type of family unit, socio-economical background (defined through parents' education level, type of occupation and income), geographical origin.

- After an appropriate codification of the data obtained through the completion of the first two phases, a comprehensive dataset was created.

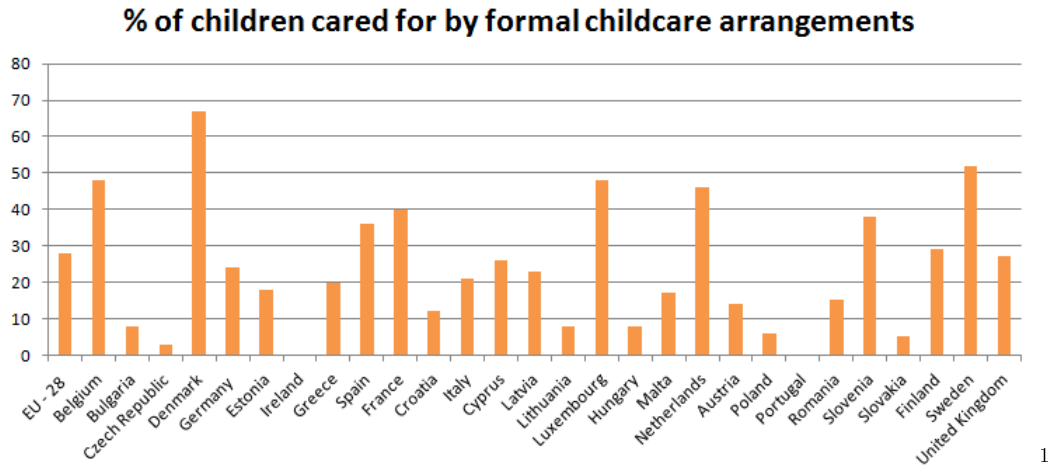
- The fourth and last step involved an empirical analysis on the basis of the original dataset. The findings of this study show that care provided by formal arrangements in the early years has a positive impact on all of the children's outcomes considered, and that the effect is stronger for children who come from disadvantaged backgrounds only in the outcome regarding linguistic competences acquired.

The plan of this work is as follows. Section 2 provides background information on the current Italian situation for the provision of childcare. Section 3 summarizes the findings of previous studies. Section 4 describes the process of data collection and the main features of the sample at hand. Section 5 discusses methodology and empirical results. Section 6 concludes.

2 Background information

In the past few years, the theme of early childcare has strongly made its way into the political debate, both at the national and at the European level. While this policy is particularly relevant in two different dimensions, child development and conciliation of work and family time, its "popularity" in the public opinion is mainly related to concerns about the participation of women into the labour market, and thus to the role that this kind of institution can play as a policy that favours work-life balance improvements. As a matter of fact, in 2002, the European Council held in Barcelona addressed this issue with the Employment Committee, establishing a goal for all of its members to create facilities that would be able to provide early childcare for at least 33% of children under 3 years of age by 2010.

Italy is among the countries that haven't been successful at reaching this target: in 2012 (latest available data), the percentage of children cared for by formal arrangements was only 21%, with a balanced share between centres operative for more or less than 30 hours per week, against a EU-28 average of 28%. Leading positions for this indicator are obviously held by Nordic countries, while the Eastern European republics are characterized by extremely low participation rates (even below 10%). Imbalance between demand and supply is a characteristic common to all of the countries, even though, clearly, for some the portion of unsatisfied demand is much more relevant.



[1]

[1]Source: personal elaboration on Eurostat – EU SILC

In most of the European countries, and Italy is no exception, there is a gap of more than 2 years between the end of childcare leave with an adequate compensation and the beginning of children's legal entitlement to pre-school childcare. This is mainly a consequence of the dominance in the Union of the split system as organizational form of pre-school childcare. The system's main feature is a different structure of the service provided, according to the child's age with the threshold being in most cases 3 years, and a tendency to support mainly the oldest group, for example through the establishment of legal entitlement to formal care. This is clearly an incisive element in understanding the existing differences in the offer of facilities and services provided to the two different age groups. Italy represents a perfect example in this sense, with 91% of children between 3 years and the minimum compulsory schooling age cared for by formal arrangements. [2]

Moreover, provisions for children belonging to the oldest group, are usually more formal in terms of educational objectives and higher level qualifications are expected for staff members. In Italy, for instance, no more than an upper secondary education level is required to be part of a staff dealing with the younger children, while the same cannot be said about the second and older group. This lower emphasis on worker's quality and preparation in the earliest childcare structures is also reflected in effectiveness measurements adopted by the various countries, which focus mainly on the care aspect of the provision, leaving educational objectives out of the picture. A fundamental factor determining the pervasiveness of childcare facilities is the amount of public expenditure committed to the creation, functioning and maintenance of such structures. While Eurostat does not provide specific data for expenses destined to early childcare, but only the aggregate value for pre-school education, the OECD delivers separate assessments for the 2 age groups. According to the latest entries available, referring to 2011, Italy spends less than 0.2% of its GDP on childcare for children under 3 years of age, while estimates for most of the western European countries examined are well above 0.4%.[3] These findings match the disparities underlined above in population coverage of the service supply. Some might

[2]Source: Eurydice (2014)
[3]Source: OECD (2014)

argue that this governmental deficiency might be overcome by the presence of a private sector participating in the market for childcare provision. However, in Italy private provisions, both self-financed and public financed, are already available, like in most of the countries in continental Europe, but this doesn't seem to make up for the shortage caused by a low level of public investment. The most successful countries, such as the Nordic ones, don't even have self-financed private settings, and are still able to provide a more extensive offer of childcare facilities. This is not to say that private institutions should not be available, but that the role of public expenditure, and consequently the role of public interest, is a predominant factor.

It is important to notice also that in Italy, public funding for early childcare is delegated entirely to regional and local authorities. While this could represent an advantage in terms of ability to address local needs and contrast peculiar obstacles created by territorial and geographic characteristics and public opinion resistance or lack of interest in this policy area, the main effect brought about by this decentralized public funding structure is a great regional disparity in terms of number and spread of facilities. The national estimates provided by Istat for 2012, show that only 54.6% of the municipalities offer this service, with a territorial coverage of 81.1% and a children participation rate of 13,5%. However, they also show the large dissimilarities existing between the centre-north area, with North-East of Italy being the most virtuous macro-region (80.7% of municipalities offering the service, territorial coverage of 94.1% and children participation rate of 19.1%), and the South and islands area, with the macro-region of South at the very bottom (service available in only 32.0% of municipalities, 58.8% of territorial coverage and 4.0% of children cared for by some sort of formal arrangements).

Notwithstanding the country and regional differences in type and expansion of the service supply provided, a feature characterizing early childcare that is common throughout Europe is the focus on its relevance for employment and work-balance policies. Little emphasis has been given to its educational function and to the impact that it might have in terms of children's behaviour, competences and achievements.

Early childcare in Italy is mainly provided through the "asilo nido". This kind of facility

accounts for almost the entirety of children using a formal care arrangement, and can have a public or private setting. Regardless of its institutional feature, attendance of the asilo nido requires parental economic contribution, usually higher in private institutions, and characterized by needs-based criteria for fees reduction in public ones. The institution of the asilo nido was formally established in 1971, with a law that even today represents the legislative landmark for this policy area and that defined it as a "social service of public interest". Therefore, at least from a formal standpoint, there was no mention of educational elements, which consequently constrained its functions to a merely welfare dimension. There is an acknowledgment deficit of the role that this institution embodies by representing the first formal setting children encounter in their lives, with potential impacts on their development and growth, that results in the lack of emphasis on the quality of the service offered.

Children coming from socio-economical disadvantaged environments often begin their schooling life showing gaps with respect to their classmates, that can also generate negative effects in the long-term. As Heckman stated in an interview: "The accident of birth is a huge, huge imperfection". For instance, children born in families of immigrants might have higher linguistic difficulties, and attending a formal care arrangements might be useful in the acceleration of the learning process of the Italian language, resulting also in an improvement of the child's ability and potential to socialize. Thus, it is important to address the disadvantage dimension driven by the social and economic context of origin, regardless of the child's innate capability, in order to verify whether being cared for by a formal arrangement at an early age can guarantee same opportunities for all individuals at the beginning of their academic life, or at least contribute to a reduction in the existing gaps.

3 Literature review

Academic attention and consequently amount of research performed on childcare arrangements and child outcomes has grown sensibly in the last decade. Main contributions in this field are provided by seminal works of Todd and Wolpin (2003) and James J. Heckman and co-authors. They modelled children's outcomes according to cumulative longitudinal processes that take into account inputs supplied by schools attended and by parents. Their point is that skills acquired in a certain phase of an individual's life has an impact on initial conditions and technology of learning experienced in the following stages, and thus investment in early childcare affects outcomes in the long-term.

In particular, Heckman and Carneiro (2003) show that gaps in college attendance and delays are largely explained by factors related to early family. They also report that early interventions show a higher return with respect to the ones implemented later on, which have a more remedial or compensatory nature. Heckman et al. (2010) assess the High/Scope Perry Preschool Program, a targeted early intervention addressed towards disadvantaged African-Americans, in terms of its rate of return. Findings show that it is positive and significant for both males and females.

These results validate the need for investigating the impact of different forms of care received in early childhood on short and long-term individual's achievements.

While several psychological studies have provided evidence in support of a positive effect of formal childcare on children's outcomes[4], findings in economic research show some discrepancies regarding impacts on specific capabilities. A relevant portion of the studies developed in this area, focuses on medium-term outcomes, mainly cognitive and measured through academic scores available, and on long-term ones, such as educational attainment and labour market participation[5] and wages[6]. However, since this research project investigates impact on achievements of children about to start their primary school educational path, I am going to report findings obtained on short-term outcomes.

Gormley and Gayer (2005), use the age cut-off for children to be enrolled in the Tulsa's

[4]Brilli et al. (2013)
[5]Havnes and Mogstad (2011)
[6]Goodman and Sianesi (2005)

Pre-K Program, a public preschool policy adopted in Oklahoma in the 90s, for their analysis. They observe a positive effect on different dimensions of children's outcomes, i.e. language, cognitive and motor skills. Gormley (2008) exploits the same setting only for Hispanic children and finds a positive impact on cognitive test scores.

Berlinski, Galiani and Gertler (2009) results on the evaluation of an Argentinian program for the provision of free and universal access to preschool, show a positive impact on behavioural outcomes and on Spanish and math test scores.

In Europe, interest on early age inputs and on their effects on children's outcomes had a more recent development. Some have used data available from the Millennium Cohort Study. Hansen and Hawkes (2009) assess effects of different categories of early childcare (formal group, formal non-group, partner care and other informal care) and demonstrate that formal group care has a positive impact on school readiness and no association with problem behaviour, while it affects negatively vocabulary test scores. Using the same dataset, Del Boca et al. (2014), focus on grandparents' care role and compare it to formal childcare. According to their findings, the latter has a positive effect with respect to the former on school readiness, pattern construction, number skills and problem solving. Moreover, for the first 3 capabilities, the impact seems to be stronger for children coming from more disadvantaged family settings. On the other hand, grandparents' care is shown to have a positive influence on child vocabulary.

Felfe and Lalive (2012) analysis, instead, evaluates impact of early childcare in West Germany and verifies that it has a positive effect on social skills and language achievements, and that such finding is stronger for children with younger and less educated mothers.

There is little evidence of childcare impacts based on Italian individuals, and this is mainly due to the lack of data for children outcomes. However, a few studies have been carried out in the most recent years. For instance, Del Boca et al. (2012) use data available from the ISFOL-PLUS dataset and show that childcare has a positive impact on children's academic outcomes, offsetting negative effects caused by working mothers. These findings are stronger for children coming from households characterized by lower incomes and educational levels.

4 Data and main features

A major element for the realization of this research project has undoubtedly been the collection of original data. While most studies realized in this area have used test scores and information on family characteristics by using national statistics, in this work a new approach for the acquisition of information has been adopted. Moreover, since the time and money restrictions didn't permit the realization of such study on a national level, and that this research is not aimed at evaluating the impact of a specific or experimental program, the choice of the universe to be analysed (identified with one municipality) has represented an important decision process.

4.1 Choosing the right municipality

Even though the geographic proximity of the municipality to the research base location (city of Milan) has been of undeniable relevance for obvious logistic reasons, other factors have influenced the final selection of the town of Mantova.

First of all, the great variety in availability and usage of early childcare facilities that characterises the Italian territory drove the focus on the country's northern regions. In order to verify the impact of a certain policy, it is important to identify the area where the likelihood of being able to benefit from it is higher. Moreover, the northern part of the country is generally characterized by a lower degree of cultural constraint against the usage of childcare facilities, meaning that prejudice against mothers choosing a formal care arrangement at the beginning of the children's life is not as pronounced as it is in the south (even though this is an element that persists in the country as a whole compared to other European nations), and so parents are able to choose more freely on the type of care that they believe to be more appropriate or convenient. Both the availability of a higher number of facilities and the higher level of openness to their use, lead to a higher level of consumers demand and this results in the need, especially for those facilities with public settings, of establishing selection criteria, which are usually in favour of low income families, for the children that will actually be the final users of the service provided. As a consequence, in municipalities, and especially in the medium and big-sized town, present

in the country northern regions it is more likely that access to the asilo nido will be granted to children coming from socio-economical disadvantaged backgrounds.

Data supplied by Istat on information related to childcare provision and usage are only available at provincial level, and show that Mantova belongs to the group of virtuous districts, with a territorial coverage of 96.7% (superior to both the national and macro-area average) and a rate of formal care arrangement used of 18% (once again higher than the territorially larger reference points). With respect to the city of Mantova, the municipality's website provides some insight related to the offer of early childcare facilities: there are 4 institutions with a public setting and 9, officially recognized, with a private one. Unfortunately there is no data available regarding the specific capacity in terms of number of children cared for in each infrastructure and in total.

As previously mentioned, use of early childcare in Italy is partially covered by parental contributions, even for public facilities, and the town of Mantova is no exception. Records provided by the municipality's official documentation show that early childcare fees are adjusted for the family's ISEE (indicator for the equivalent economic situation) level, going from a minimum monthly rate between €. 20.00 and €. 40.00 for those belonging to the lowest bracket, to a maximum of €. 491.13 for those with an ISEE of €. 22724.01 and above. These reductions do not apply to non-resident families, who are charged €. 800.00 per month, independently from their economic situation. Data on the costs of private early childcare facilities is not available, but they are usually higher with respect to the public ones. However, for children signed up in public care facilities waiting lists, it is possible to obtain a voucher that covers the 50% of a private accredited institution rate. Clearly, choosing a formal care arrangement represents a consistent expense, but this is generally the case across the country.

Another important aspect considered for the selection of the town of Mantova has been the dimension of the reference population. As a matter of fact, the universe to be chosen had to be not too large, in order not to compromise the feasibility of information collection, and not too small, so that the findings validity would not be nullified by an analysis based on few observation. This sort of in-between/balance number of children currently

attending the first year of primary school has been estimated, after a consultation with an expert in this kind of research activity, to be around 400, which basically meant an overall population of about 50,000 individuals. Indeed, Mantova has 47,223 inhabitants and the number of children enrolled in the educational year of interest is 370.

Given the study's interest to investigate the impact of early childcare for children coming from disadvantaged backgrounds, with a particular concern for immigrant families, it was important to choose a municipality with a significant presence of foreign individuals. Indicatively, the minimum share of immigrant population that would account for a noteworthy foreign presence had been identified around 10%. The town of Mantova did indeed fit this criterion as well, with a 14.20% of non-Italian population. The most consistent groups (more than 400 individuals) hailing from Morocco, Romania, Brazil, Albania, Ukraine, China and Bangladesh.

Last but not least, an influential factor was represented by the organizational structure of primary schools. The role of this institution has been of fundamental importance for the process of data collection. On one hand, these schools were the final recipients of the end of pre-primary schooling evaluation records, since they are mainly used for the formation of balanced classes. On the other, primary schools represented the means through which it was possible to gain access to children's parents, and therefore hand them the questionnaires necessary for the completion of the second phase of the study. Procedures regarding the achievement of agreements on participation to a research project and of the authorization to dispose of sensitive data for privacy reasons can be time costly. Thus, the fact that the 9 primary schools in Mantova are under the domain of only 3 "Comprehensive Institutes" allowed for a relatively smoother and faster preliminary course of action in obtaining the green light to proceed from the school head.

4.2 Data collection

The first contacts with the primary school heads have taken place in the spring of 2014: upon presentation of documents regarding the aim and structure of the research project, a declaration guaranteeing data protection and a copy of the questionnaire meant for the

children's parents, school councils approved the decision to collaborate for the realization of the study. The institutions involved, all of which comprise 3 primary schools, are Mantova 1 ("S. Allende", "P. Pomponazzo", "Don E. Tazzoli"), Mantova 2 ("R. Ardigò", "Don P. Mazzolari", "I. Nievo") and Mantova 3 ("Martiri di Belfiore", "Don Leoni", "Don Minzoni").

The evaluation records coming from pre-primary education facilities were filled in at the end of the academic year and handed in to the children's primary school of choice. Then, they were gathered by the 3 coordinating Institutes and made available for consultation. The process of obtaining the documents necessary for the extraction of this first information to be included in the new dataset took approximately between 3 and 4 months.

Following the completion of the first stage, the list of students enrolled in each class of every primary school was collected. It was necessary not only to obtain the official number of children attending the first year of primary school but also to distribute the questionnaire with codes that allowed for the matching with evaluation records. In agreement with the schools involved, it was indeed decided to proceed with the process of making the individuals anonymous prior to the diffusion of questionnaires, in order to strengthen the families' sense of data protection and thus increase the chances of their collaboration. Therefore, a code was assigned to each student according to the official lists provided, and the evaluation records were classified correspondingly.

Lastly, the questionnaire was distributed to parents of all of the children enrolled in their first year of primary school. During this phase of the realization of the project, there were direct contacts with all of the schools and teachers that participated to the study. They were the ones who handed the questionnaires to the children's families and collected them. Some of the educators involved also provided assistance to individuals for whom the Italian language was a barrier in understanding and filling the form. This final process of data collection was completed in about 3 months.

4.3 Questionnaire design

The creation of the questionnaire was realized on the basis of the kind of information that was believed to be influential both on children's outcomes and on the families' choices of care arrangement in the early years of a child's life. With few exceptions, the majority of questions were presented in a multiple choice format, in order to ease the codification process and to obtain standardised replies. Moreover, this kind of format makes the questionnaire more user friendly, by reducing the amount of time required to fill it and by providing an extensive range of options, thus increasing the likelihood of parents' willingness to respond. Another factor that was taken into account when drafting the questionnaire was its length: since the response rate was merely based on parents' voluntary submission of the filled out document, it was important to keep it short and to make it easy to handle, therefore an A4 folder was chosen as the most appropriate format.

1. The first category of relevant data is related to the feature of the family nucleus. The respondents were asked to respond about the parents' civil status. 4 options were given (married/living together, divorced/separated, single parent, widowed), defined in such a way to give an rough indication of whether and to what extent both parents are present in the child's life. It is also asked to list the number, affiliation, sex and age of all the components of the family nucleus, and whether there are grandparents living nearby. The former is requested with the main purpose of getting hold of information on siblings, while the latter is to check if the most common care arrangement, besides care given by parents, is an available alternative to formal early childcare. Furthermore, respondents are asked to indicate their country of origin and the language(s) spoken with the children within the household.

2. The second type of data is connected to care arrangements, and in particular to the type of care actually received, the type of care preferred by parents in the best case scenario (without considering constraints such as costs and distance in the case of formal care, or lack of relatives availability in the case of informal type of assistance), and whether parents applied for a formal childcare facility. This last

variable is meant to be used in case there is interest in considering differently those children who were not cared for by a formal arrangement on grounds independent from parents' will.

3. The third and last group of data is related to parents' characteristics. Respondents are required to provide information on:

    - highest education qualification, on a scale from 1 to 6, with 1 representing a level below completion of primary school and 6 the attainment of a postgraduate degree;
    - occupation type, using the categorization of different kinds of jobs elaborated and used by the INVALSI (National Institute for the evaluation of the education system);
    - parents' participation rate to school meetings and activities (taken separately for mothers and fathers considering 3 possible levels – less than 50%, between 50% and 75%, above 75%;
    - income level (defined in accordance with the 5 brackets used by the national fiscal system for income taxation).

    Additional questions were asked on who brings the child to school and who picks him up, in order to obviate to the possibility of non response on the participation component by providing supplementary indicators on parents' presence in the child's school life. Questions about income level were, predictably, the ones that raised the biggest issues already during the preliminary steps related to school council and school heads agreement to collaborate for the realization of this study. Notwithstanding the fact that respondents were permitted to leave in blank fields they did not feel comfortable with, the Comprehensive Institute of Mantova 3 has not prohibited the inclusion of questions directly referring to earnings in the family, and the same has happened for the primary school "Don E. Tazzoli", belonging to the Comprehensive Institute of Mantova 1. While in the second case this provision concerned only one class of students (composed by 21 students, out of which only

4 returned the completed questionnaire), in the first there were 3 primary schools involved, for a total of 118 children. Given that the amount of missing values for this variables would be extremely consistent, another set of questions was added to the questionnaires, with the intention of finding an alternative method for the calculation of an indicator that could give an estimate of the families' income levels. The strategy implemented was to use the price paid per meal in the school canteen. As a matter of fact, the value of this fee changes according to the family's declared ISEE, adjusted for the number and age of other children belonging to the same family nucleus benefiting from the use of this service and for the family residence (in case this was in any municipalities other than Mantova, cost per meal would be full price, regardless of the nucleus income level). In this way, the number of missing values was reduced and it was influenced by both missing responses on the questionnaire and the number of children attending a primary school in Mantova but resident in another municipality. Moreover, even though this method does not allow for the distinction, and consequently the analysis of a potential difference in impact, of the income level of the mother and the father, it is helpful in reducing the potential bias given by distortion of information, which is a quite common issue in income data collection based on self-reporting. While information regarding the existence and personal data of siblings were originally part of a first draft of the document, other fields were added; more specifically, a question on the family officially registered municipality of residence and one on whether there were siblings using the canteen service. The ISEE classifies families incomes in 5 brackets:

from €. 0 to €. 5164.00

from €. 5164.01 to €. 10329.00

from €. 10329.01 to €. 15493.00

from €. 15493.01 to €. 22724.00

from €. 22724.01

Fees reductions of 50% on the cost per meal are applied to the oldest child of a resident family with 2 children using the service, and on all of the children but the oldest belonging

to a resident family composed by 3 children. The availability of this information and the kind of data collected through the questionnaire ultimately permitted the calculation of families ISEE levels.
A copy of the questionnaire is provided in the Appendix section.

4.4 Dataset

The first step in the creation of the dataset, was to accurately codify children's outcomes described in the pre-primary schools evaluation records.
In Mantova there is a total of 16 pre-primary institutions: 6 of them ("T. Ferrari", Da Feltre", "Strozzi e Valenti", "Montessori", "Visentini", "Calvi") are council schools, 2 are private equipollent institutions ("Casa dei bambini" and "Redentore – Ferrini") and 10 are public schools, i.e. fall under the organizational umbrella of the Comprehensive Institutes previously mentioned ("Berni" and "Pacchioni" – Mantova 1, "Rodari" and "Frank" – Mantova 2, "Campogalliani", "Collodi", "Ricordo ai Caduti" and "Sawyer" – Mantova 3). Moreover, there is a handful of children coming from facilities situated in municipalities close by. This means that children enrolled in their first year of primary school come from a certain variety of pre-school institutions. Unfortunately, there is not homogeneity among the evaluation records provided by such schools, not even at the town level. This is mainly due to the fact that there is no coordination nor guidelines for the provision of students' information at this stage of their educational path, and this is true throughout the country. Most receiving schools have created their own standard form, which have mainly been used by the public pre-primary institutions belonging to the same Comprehensive Institute. Certain council and private schools have also adapted by using these forms according to the child's primary school destination. Other options, adopted mainly by council schools, for the communication of children achievements, comprise the use of their own evaluation records form, information-passing through preliminary meetings with first-year primary school teachers or the delivery of a so-called "personal booklet", filled with the some of the child's works. Lastly, there were cases in which no information was provided, or was available in any official record. Thus, in total, there were 8 types of records forms

available, even though, at least for the most structured ones, there were common grounds besides general distinction of the outcomes, such as types of entries included or "grading" method.

The variety of inputs, considering both the differences between structured forms and other less "orthodox" sources of information, and the dissimilarities among the different types of structured forms, is what made the standardization and codification process particularly arduous. Moreover, it needs to be taken into account that for children whose information was collected through teachers conversations and personal booklets it was not possible to derive a complete indication of their capabilities, and so outcome levels were available only for certain fields.

First of all, it was important to identify precisely the skills areas of interest, which have been classified into 4 types of child outcomes. A scrutiny of the literature available on this kind of studies makes it clear that there is in general a main distinction between cognitive and non-cognitive skills (Cunha and Heckman, 2008, among others), which are seen as factors that affect children's future lives in different ways and on diverse grounds. Evidence from the collected evaluation records forms shows that this non-written rule to report separately such types of accomplishment has been followed, which ultimately led to the decision of analysing these two aspects separately. In the realm of non-cognitive skills, it was possible to discern two factors that define a child's capabilities. On one hand there was the level of autonomy accomplished, meaning self-care, eating, dressing and care of school material abilities, among others. On the other, social relationships were evaluated, including ability of relating to other children, to adults, and of respecting behavioural and school rules. These two sets of capabilities are identified respectively with the first and the second outcome considered for the analysis of early childcare effects.

When it comes to cognitive skills, the main component is clearly defined by linguistic competence and simple preliminary mathematical knowledge (such as ability to count). This set of capabilities is what constitutes the third type of outcome considered and is the one that is more strongly related to future learning achievements. In particular, language fluency represents a very important element for immigrant children, who might encounter

a higher degree of difficulty in their first schooling years when facing communication and comprehension barriers. Furthermore, by examining the evaluation records, it emerged that graphical competences make up for another discriminating factor, concerning issues of fine motor skills and ability to be aware of the human figure and common objects and to reproduce them. This last element constitutes the fourth and last outcome held to be relevant, and therefore included in the dataset, in analysing children development.

Once the 4 outcomes were identified, it was decided to use a scale going from 1 to 4 for the evaluation of children's achievements, following the method used by the most systematic pre-primary school forms. The value of 3 is what corresponds to a general good valuation, and the other values are consequently assigned in a comparative fashion. With respect to the outcomes measuring graphical competences, professional consultancy was involved in the evaluation of documentation provided by personal booklets.

In total, the records collected amount to 256 (which correspond to 69.19% of children currently at the beginning of their schooling career). However, due to the lack of a well-structured form used by all the pre-school institutions, the number of children evaluated vary for each income. In particular, there are 223 observations available for the first outcome, 230 for the second, 232 for the third and 249 for the fourth. With the exception of the last type of outcome considered, the mode is represented by the value 3 (in outcome 4, the modal value is 4).

Furthermore, variables classifying the type of pre-primary school attended and the evaluation form type were included. Of particular relevance is the former, because it allows to control for fixed effects generated by the educational facilities. Clearly, since these institutions were attended by the child in the bridge period between the end of early childcare and the evaluation time, they might have had an influence on his capability outcomes.

The second step for the creation of the dataset was to insert the information collected through the questionnaire, which comprises variables regarding family nucleus composition, care provisions for the child and parents' features. For some of the variables, elaborations were needed in order to pull together different kinds of data or to format them in a more practical and functional way. The latter is the case, for instance, for

personal data provided for siblings, for a better classification of occupation type (by considering high-skill and low-skill jobs) and education level (by distinguishing between those holding a degree and those who don't). The former, instead, was the case for the creation of the variable indicating families ISEE level (built on other 3 available variables, as previously explained), parents participation rate to academic activities (built on self-reported levels and confronted with open-ended answers on who brings and picks up the child from school) and for the creation of a variable called "origin", that mixed data on parents country of origin with information of language spoken at home with the child. The aim of this variable is to provide a sort of immigration rank, by giving the lowest value to children coming from families where both parents are foreign and Italian is not one of the idioms used within the household, and the highest to the ones with both parents coming from Italy and speaking Italian.

The total number of questionnaires collected is 272, which accounts for the 73.51% of children enrolled in their first year of primary school. Looking at the disaggregated data, it is possible to notice that the response rate was not heterogeneous among different schools; instead, some observations went missing in every class/facility, which ultimately led to the loss of little less than 100 questionnaires.

Ability to perform the desired analysis requires the existence of both elements of the dataset (outcome variables and information obtained through the questionnaire). This necessary condition caused a cutback in the number of observations worthy of consideration to 190 with respect to the amount of data collected.

In order to verify whether the sample used for this study can be representative in terms of the population characteristics of families with a child enrolled in the first year of primary school in the municipality of Mantova, a two-test Wilcoxon-Mann-Whitney test is performed on all of the variables obtained through the questionnaire. Through this test it is indeed possible to check for potential disparities between the group of questionnaires included in the sample, and the ones that were left out because there were no matching evaluation records. If the p-value provided by the test is significant, it means that there exists a statistically significant difference between the underlying distributions. Results

show that p-values are non significant for all the variables considered, therefore we can conclude that, considering the set of all questionnaires received, the sample of observations used in the analysis is not biased and thus can be considered to be representative.
The two-sided Wilcoxon-Mann-Whitney test was preferred to the more widely used t-test because the distribution normality assumption does not apply. Instead, it is required that variables are ordered. For this reason, it is believed to be better suited for the nature of the available data.

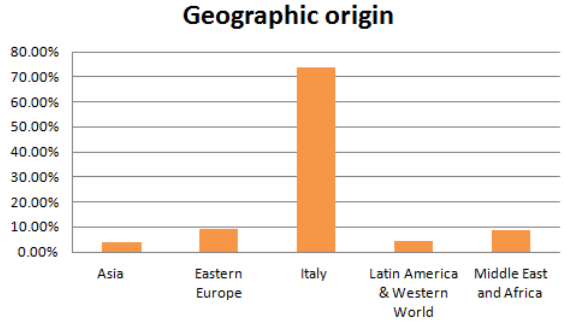
4.5 Descriptive statistics

Hereby I will introduce some relevant summary statistics that will give an insight on the main features of the sample at hand. The rate of children observed in the sample that have been cared for by formal arrangements, in either public or private facilities, is 64.74%. One possible explanation for this finding may derive from the fact that families were aware of the study main purpose, through official communications received by the primary school of reference, and hence those that benefited from this service were likely to be more inclined to take part to the research project. When the ISEE level is taken into account, by looking separately at observation that belong to the highest bracket versus all of the others, difference in share of treated children is minimal and statistically non significant. Lastly, if we separate the sample in order to detect the rate of children that received formal care in their early years among Italian families and families in which at least one of the parents is foreign, we see that in the former group it is nearly 70%, while in the latter it is close to half of the individuals considered.

	Sample	Native	Foreign	Highest ISEE	Other ISEE
% treated	64.74	69.23	51.06	65.96	63.54

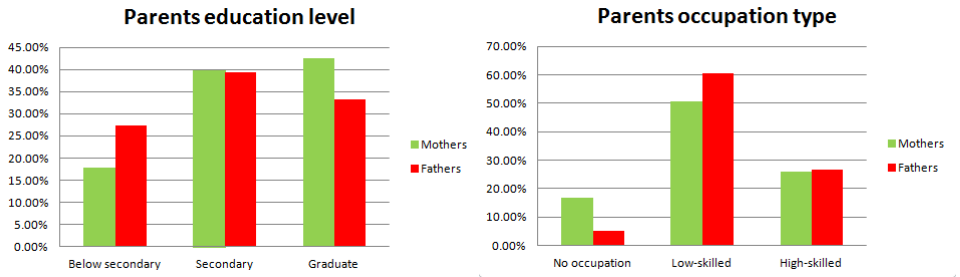
A narrow majority of children was cared for by public facilities, rather than private ones, while a large proportion has grandparents living nearby, and therefore, supposedly, an alternative option to formal care.
The modal value for the ISEE level is the one corresponding to the highest bracket;

however, it needs to be taken into account that the required minimum level to belong to the wealthiest group is quite low. As expected, most of the family nucleuses are composed by both parents of Italian origin, with approximately a quarter of the sample characterized by the presence of at least one foreign parent. Among non-native individuals, we can see that the most represented groups are originally from the Middle-East & Africa and from Eastern Europe.



With respect to parents' participation in the child's educational life, the modal value for mothers is given by the highest in rank, while for fathers by the lowest. This is to be expected, since in general mothers are more present in their children educational (but not only) life. Moreover, a larger occurrence of missing values is registered for fathers.

Looking at parents' educational achievements, the modal value for fathers corresponds to the attainment of a high school degree, while for mothers to a bachelor degree or a higher qualification. The share of parents that do not hold a high school degree is higher for fathers, but for mothers we have more missing values. These proportions do not hold for type of employment: as a matter of fact, for both fathers and mothers, the largest number of observations is identified with low-skills job types. Furthermore, among mothers there is a higher percentage of individuals that do not work.



5 Empirical analysis

5.1 Methodology

The work of Todd and Wolpin (2003) provides a useful insight on the methodological grounds for the estimation of factors that affect child development and achievements and for the adoption of alternative specification methods.

The starting point of the conceptual framework for knowledge acquisition process is based on analogies with firm's production processes. Most of the literature on the education production function focuses on longitudinal information and on cumulative processes. Indeed, the general model provided takes into account a vector for school-supplied inputs history ($\mathbf{S}_{ia}$) and a vector for parents'-supplied inputs history ($\mathbf{F}_{ia}$). A measure for individual mental capacity is included and believed to be constant over time. The function for the measure of achievement, allowing for errors in measurements ε_{ia} is then given by:

$$\mathbf{T}_{ia} = \mathbf{T}_a \left[\mathbf{F}_{ia}, \mathbf{S}_{ia}, \mu_{i0}, \varepsilon_{ia}\right]$$

where i is the individual considered and a is the age to which a specific measurement or input refer to. Although it is present in the model, a measure of fixed mental capacity is non-observable.

Moreover, allowance for inputs provided at different stages of an individual's life do not apply to the type of analysis performed in this study, because the aim is not to measure lifetime achievement. The analysis, instead, focuses on a relatively short life span. For this reason, the specification that is best suited considering the purpose of this research project and the data at hand is the *contemporaneous specification*, whose main assumptions are:

1a) contemporaneous inputs are the only ones relevant for current achievements

or

1b) inputs do not change over time, and therefore current ones are sufficient to capture the entire history of inputs.

Given the short temporal arch considered in this study for capabilities development, it is rather expected that parents inputs such as educational levels and occupation type do not vary over time. In the case of the latter, this is even more likely since it is evaluated

in terms of low-skilled and high-skilled jobs: it is possible to get promoted in a time range of few years, but rather difficult to move across these types. When consider parent's participation in children academic activities, it is quite reasonable to assume that it is constant over time. Finally, regarding ISEE measurement, the assumption is that, even though its level might change over the years, it is the current level that determines a situation of economic advantage or disadvantage.

Regarding school-supplied inputs, besides the main variable of interest, i.e. the choice of a formal childcare arrangement, inputs provided by pre-primary schools are controlled by clustering individuals for type of institution. The dimension of the sample at hand and of the schools involved do not allow for a clusterization at the facility level.

Lastly, a vector describing early home environment ($\mathbf{Z}_i$) is included.

Therefore, the final modified version of the model is the following:

$$\mathbf{Y}_i^k = \mathbf{Y}\left[\mathbf{F}_i^k, \mathbf{S}_i^k, \mathbf{Z}_i^k, \varepsilon_i^k\right]$$

where

- $k \in (C, N)$, with $C =$ cognitive skills and $N =$ non cognitive skills, that can both assume value 1 or 2

- $\mathbf{S}_i$ indicates the choice of care in the early years

- $\mathbf{F}_i$ includes parents' participation, preference of type of care, educational level, occupation type and level of ISEE

- $\mathbf{Z}_i$ is composed by variables indicating the child's sex, parents' civil status, existence of siblings in the household, family origin and presence of grandparents living nearby.

5.2 Empirical strategy

The nature of this study is not experimental. As a matter of fact, the treatment is not assigned randomly to the sample, but it is a decision determined by specific factors regarding family features and background. This problem can be overcome by controlling for all of the inputs believed to have an impact on such decision, which is the reason why an extensive amount of information was collected through the questionnaire, taking into consideration previous work and pertinent literature that provide an indication on what

are the factors deemed relevant.

Another issue that could be raised is that the very same elements that influence whether an individual belongs to the treatment group have an impact on the outcome. Put it more simply, it is reasonable that a common input, such as mother's education level for instance, has an effect both on the choice to get the treatment and on child's capabilities, there might be identification issues and results would be biased. Given the lack of an appropriate instrumental variable that captures the effect of early childcare, the Bivariate Probit Model is chosen to obviate erroneous estimations. This kind of regression has the feature of fitting maximum–likelihood two-equation Probit Models of ordinal variables on independent variables, allowing the two lists of regressors to be different one from the other. Errors are normally distributed with correlation ρ.

In particular, a Semi-Ordered Bivariate Probit is the extension of the model that best fits the analysis implemented in this study, since we are dealing with ordered outcome variables and the dummy variable for treatment.

The two equations, built normally as a Probit and as an Ordered Probit, are then evaluated simultaneously with a cluster control for the kind of pre-primary institution (public, council, private and schools outside the town) attended between the age of 3 and 5. The controls used for both regressions include family nucleus structure and parents' characteristics; level of ISEE and family origin impact on outcome will be controlled for in separate specifications, in order to be able to isolate the effects of these two different indicators of disadvantage.

5.3 Results

Before proceeding with the presentation of main results regarding children outcomes, it is relevant to show two preliminary findings obtained through the evaluation of the entire number of questionnaires available.

The first one concerns the analysis of the factors that might have an impact on the type of care chosen for a child's early years. In this case, the outcome is identified with the treatment variable, which is a dummy that assumes value 1 in case a formal care

arrangement was chosen, and value 0 otherwise. The determinants of such decision are believed to be characteristics regarding the family nucleus structure and parents inputs. The presence of siblings within the family is considered differently with respect to what will be done in the following analysis on child cognitive and non-cognitive outcomes, by restricting the evaluation to brothers and sisters born before the beginning of the pre-primary education period. Children born afterwards, indeed, could not have an impact on parents decision. The estimation model chosen for this identification is a probit, because of the dummy nature of the treatment variable.

Table 1: Decision on treatment

sex	0.155
	(0.183)
civil_status	0.022
	(0.146)
siblings1	-0.093
	(0.198)
grandparents	-0.549
	(0.307)
origin	0.143
	(0.099)
participation_mother	0.339**
	(0.127)
participation_father	-0.040
	(0.099)
education_mother2	0.412**
	(0.172)
education_father2	0.204
	(0.149)
employment_mother1	0.400**
	(0.183)
employment_father1	-0.340
	(0.202)
isee	-0.018
	(0.051)
constant	-1.314*
	(0.051)
Observations	229

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1 shows the results of this preliminary analysis. We observe that all the inputs provided by the mother of the child, i.e. participation to his academic life, education level and type of occupation, have a positive and significant effect on the decision to adopt a

formal care arrangement. Thus, mothers that are more attentive on educational processes and that are more educated themselves seem to prefer childcare to care provided by family members. The same can be stated about mothers with high-skilled jobs, who might have less time to stay at home and take care of the baby. Inputs provided by fathers and presence of other children in the household do not seem to have an impact, as well as the availability of grandparents living nearby. It is relevant to notice that ISEE level does not seem to be relevant, suggesting that it is possible that fee reductions provided to poorer families ensure service affordability, and that coming from an immigrant family does not influence the decision either. It seems likely, then, that choice of care arrangement is not induced by economic possibility nor cultural belief.

Table 2: Period length of formal childcare

sex	-0.898
	(1.430)
civil_status	-1.479
	(1.498)
siblings	-1.479
	(1.498)
grandparents	-5.035**
	(1.917)
origin	1.036
	(0.714)
participation_mother	2.743**
	(0.893)
participation_father	-0.721
	(0.785)
education_mother2	2.149
	(1.429)
education_father2	2.003
	(1.170)
employment_mother1	3.148**
	(1.406)
employment_father1	-2.595
	(1.602)
isee	0.240
	(0.385)
constant	-2.422
	(3.813)
Observations	229

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The second important result concerns only children using a formal care arrangements, and aims at understanding which factors affect for how many months children benefit from childcare facilities. Considering that children are entitled to attend pre-primary school facilities when the 3rd year of age is reached, this is basically an evaluation on what drives parents to turn to formal care earlier than others. The outcome variable, duration, expresses in months the amount of time spent by the child in a childcare facility, and same regressors for the evaluation on type of care decision are considered. Given the continuous nature of the duration variable, an OLS regression was more appropriate for this analysis. Results shown in Table 2 indicate that duration is affected only by factors related to families' organizational needs. As a matter of fact, children use a formal facility earlier if the mother has a high skilled job, which is likely to involve a shorter period of absence due to maternity (especially for professionals, self-employed, or those holding responsibility positions their workplace), and if there are no grandparents living nearby that could represent an alternative option in children's earliest months. As it can be seen from the Table above, the former coefficient is positive, the latter is negative and both are significant.

Main results As explained in the previous section, when evaluating the impact of formal childcare on children outcomes, there could have been problems regarding identification, with the treatment being affected by the very same inputs that have an impact on the dependent variable. These potential shortcomings have instead been overcome through the adoption of a Bivariate Ordered Probit model. For the following specifications, the Wald test for independent equations has always been rejected, thus suggesting that corresponding Univariate Ordered Probit results were biased. Moreover, the cut points generated by the Bivariate Order Probit are all significant, ensuring that categories reported by the dependent variable are significantly different one from the other.

Cognitive skills

The first focus of this study was centred on cognitive skills developed by the child.

The outcome regarding linguistic and mathematical competences is of particular importance because is the one linked to measurements of future years school performance. Gaps

in this area of capabilities are in fact the ones that can strongly undermine the child's initial learning process and his prospect academic achievements.

As previously mentioned, I will control separately for ISEE and origin in the second equation, in order to gain a clearer idea of how these two components that represent indications of socio-economic advantages might have an impact on children outcomes.

Table 3: Linguistic and mathematical competences on origin

treatment	0.923**
	(0.304)
origin	0.438***
	(0.060)
treat_origin	-0.123**
	(0.056)
sex	0.104
	(0.175)
civil_status	-0.204**
	(0.071)
siblings_older	0.046
	(0.100)
siblings_younger	-0.041
	(0.127)
grandparents	0.171
	(0.380)
participation_both	0.146**
	(0.348)
education_mother	0.174**
	(0.066)
education_father	-0.014
	(0.051)
employment_mother	0.015
	(0.164)
employment_father	0.241*
	(0.109)
Observations	147

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Starting with the specification for origin, we can see from Table 3 that treatment is significant and has a positive impact. Also the coefficient for origin is positive at the highest level of significance, validating the hypothesis that children with both parents coming from Italy have a higher chance of a better school performance. It is interesting to also notice that the interaction term is significant with negative sign. This implies

that the treatment has a higher impact for children with a lower score in the origin variable, i.e. for those born in households with foreign parents. Hence, this result shows that not only using a formal care arrangement has a positive effect on linguistic and mathematical competences, but also that it can represent a valuable tool for the reduction of gaps generated by background settings. Of the other significant results, it is found that mother's level of education has a positive coefficient while parents' civil status is negative. The latter might be explained by a higher level of attentiveness towards the child, in cases of single parents or split-up parents.

Table 4: Linguistic and mathematical competences on isee

treatment	1.231***
	(0.193)
isee	0.084***
	(0.017)
treat_isee	-0.084***
	(0.040)
sex	0.082
	(0.124)
civil_status	-0.201***
	(0.016)
siblings_older	0.132
	(0.132)
siblings_younger	-0.130
	(0.126)
grandparents	0.409
	(0.380)
participation_both	0.146
	(0.191)
education_mother	0.174**
	(0.066)
education_father	-0.014
	(0.051)
employment_mother	0.015
	(0.164)
employment_father	0.241*
	(0.109)
Observations	147

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Turning to the specification for the ISEE level, I find similar results (shown in Table 4). Once again treatment is significant and positive, as well as the coefficient for ISEE.

The coefficient for the interaction term constructed on this two variables is significant at 1% level and has a negative sign, implying a higher impact of formal care arrangements on children coming from poorer families. Regarding other inputs, education level of the mother of the child and occupation type of the father are positive and significant, while parents' civil status is significant but with a negative coefficient.

The outcome regarding graphical competences has been included in this analysis thanks to the presence of this evaluation element in the records provided by all of the pre-primary school institutions. Its relevance in relation to prospect academic achievements is unclear, although it can be considered as a proxy for child's maturity (ability to visualize and represent the human figure and level of understanding of space and proportion) and precision (ability to cut and paste figures of different degree of complexity, for instance). Prior to the implementation of the analysis, it was necessary to recode the outcome variable into 3 categories rather than 4, due to the an extremely low occurrence of children receiving a valuation of 1. Hence, the 2 lowest classes were grouped together. For this outcome, controlling for the ISEE level yields relevant results, while the specification obtained using the origin factor does not provide any further information on possible factors influencing the dependent variable. The former identification (see Table 5), in particular, reports a significant and positive effect of being cared for by childcare facilities, while income level identified with ISEE doesn't appear to be relevant. Furthermore, the coefficient for the child's sex is significant, indicating that girls show better results than boys, while the presence of younger siblings in the household has a negative impact. Lastly, once again the educational level of the mother is confirmed to have a positive and significant effect.

Non-cognitive skills

The second part of the analysis implemented in this study is related to non-cognitive children outcomes, identified with degree of autonomy on one hand, and social relationships on the other.

For the former capability, a re-codification of the outcome variable was necessary for the same issues encountered in the specification for the variable measuring graphical compe-

tences. It seems rather logical, indeed, that a certain minimum degree of autonomy and of graphical ability is reached by a very large majority of 5-years-old children. Following this slight fine-tuning procedure, the analysis was performed for both origin and ISEE level (see Table 6 and Table 7). Taking into account the effect exercised by other inputs on treatment, the only variable with a significant coefficient in both identifications is treatment itself, which implies that having benefited from a formal type of care arrangements has a positive impact on the child degree of autonomy. The coefficient for the origin indicator is significant and positive as well, with children coming from native families being more self-confident and autonomous. An attempt to control for the interaction of these two effects has been carried out, but it brought no evidence on whether treatment has a stronger impact on disadvantaged children.

Table 8 and Table 9 show results for the impact of formal childcare on social relationships, once again controlling for both the ISEE level and the origin indicator. In this case, the two identifications provide the same outcome, with treatment being significantly positive, i.e. treated children have less behavioural problems and better interaction with other children and adults. Moreover, also presence of younger children within the household is found to be significant and positive. Thus, having younger siblings improves relational capabilities. It seems, however, that income level and family origin do not have any impact on this type of outcome.

6 Conclusions

The purpose of this research project was to analyse the impact of early years formal childcare arrangements on educational objectives, including both cognitive and non-cognitive skills, in the short term. The focus was on capabilities acquired prior to the beginning of compulsory schooling, that in Italy corresponds to primary school attendance. The hypothesis behind this study was that children cared by formal types of arrangements would have a higher score in their educational outcomes evaluation, once other influential inputs had been controlled for. Furthermore, a particular interest was placed on the possibility that this impact would be larger for children coming from socio-economical disadvantaged backgrounds, identified mainly with income levels and family origin (native or foreign).

The analysis was carried out separately for different types of outcome, in order to account for potential discrepancies in the effects produced. Moreover, the two identified indicators for underprivileged backgrounds were controlled for in separate specifications. The results show that formal type of childcare arrangements in the early years, i.e. under 3 years of age, has a positive and significant impact on all of the outcomes considered, even though some differences emerged. The effect on linguistic and mathematical capabilities is found to be not only relevant per se, but also in terms of assuming a "closing the gap" function for children coming from disadvantaged backgrounds, with respect to both income level and family origin identifications. This result is particularly noteworthy, given that this is the outcome that more than others foresees prospective academic achievements. The impact on graphical capabilities is relevant only when controlling for income levels, and there seem to be no differences in performances determined by socio-economical background. Regarding outcomes that measure non-cognitive skills, results show that children coming from native families have a higher degree of autonomy; however, the positive effect generated by the treatment is not proved to be different (in any direction) for those belonging to underprivileged households.

The overall conclusion that can be drawn from this study is that formal type of care affects positively all the capabilities achievements taken into account, contributing to close the existing gaps endured by children coming from disadvantages backgrounds for

the most relevant cognitive outcome, while impacting indistinctively achievements that are more related to nature, rather than nurture, features. Therefore, the consequential policy suggestion that can be derived is that we should invest more in the provision of formal childcare facilities.

One major positive spill over of this analysis is the creation of an original and comprehensive dataset, that represents a starting point for the process of filling the information gap in this research area. Given the amount and kind of data provided, this dataset can be used for further analysis on care arrangements at an early age. An improvement on this type of study can undoubtedly be boosted by a more standardised format for measuring children outcomes at the end of the pre-primary educational path. Indeed, information obtained would be easier to collect and more user-friendly during the codification process. With less tightened time and money constraint and a higher organizational capacity, other types of information could be included in the analysis, such as longitudinal (backwards) records on parents achievements and inputs, and measures for schooling and teachers quality. Moreover, future research on long-term outcomes for the analysed sample is recommended, in order to provide a more comprehensive analysis of the impact of this educational policy.

This study has been carried on with the aim of representing a pilot project for future research in this area. Therefore, the final recommendation is to implement this study in other settings of larger dimensions, in such a way to supply more evidence in support of the findings reported above. Furthermore, it would be interesting to investigate results obtained in different areas of the country, conducting this kind of research in areas were formal childcare facilities are less spread and accounting for potential regional differences that characterize the Italian territory.

References

[1] Berlinski, S., Galiani, S., and Gertler, P. (2009). The effect of pre-primary education on primary school performance. *Journal of public Economics*, 93(1), 219-234.

[2] Brilli, Y., Del Boca, D., and Monfardini, C. (2013). Child care arrangements: determinants and consequences (No. 18). *Centre for Household, Income, Labour and Demographic Economics (CHILD)-CCA*.

[3] Carneiro, P., and Heckman, J.J. (2003). Human Capital Policy. In Heckman, J.J., Krueger, A.B. and Friedman B.M. (eds) *Inequality in America: what Role for Human Capital Policies?* Cambridge: MIT press, 77-239

[4] Cunha, F., and Heckman, J. J. (2008). Formulating, identifying and estimating the technology of cognitive and noncognitive skill formation. *Journal of Human Resources*, 43(4), 738-782.

[5] Cunha, F., and Heckman, J. J. (2009). The economics and psychology of inequality and human development. *Journal of the European Economic Association*, 7(2-3), 320-364.

[6] Del Boca, D., Pasqua, S., and Suardi, S. (2012). Childcare, family characteristics and child outcomes: an analysis on Italian data (No. 9). *Centre for Household, Income, Labour and Demographic Economics (CHILD)-CCA.*

[7] Del Boca, D., Piazzalunga, D., and Pronzato, C. D. (2014). Early child care and child outcomes: the role of grandparents. Evidence from the Millennium Cohort Study (No. 24). *Centre for Household, Income, Labour and Demographic Economics (CHILD)-CCA*.

[8] European Commission/EACEA/Eurydice/Eurostat (2014.) *Key data on Early Childhood Education and Care in Europe. 2014 Edition*. Eurydice and Eurostat Report. Luxembourg: Publications Office of the European Union.

[9] OECD (2014). *OECD Family Database*. Paris: OECD (www.oecd.org/social/family/database.htm)

[10] Felfe C., Lalive R. (2012). Early child care and child development: for whom it works and why. *IZA* Discussion Papers 7100.

[11] Gormley, W. T. (2008). The Effects of Oklahoma's Pre-K Program on Hispanic Children. *Social Science Quarterly*, 89(4), 916-936.

[12] Gormley, W. T., and Gayer, T. (2005). Promoting school readiness in oklahoma an evaluation of tulsa's pre-k program. *Journal of Human resources*, 40(3), 533-558.

[13] Greene, W. (1990). *Econometric Analysis*. New York: Macmillan.

[14] Goodman, A. and Sianesi, B. (2005) Early education – how long do the impacts last? *Fiscal Studies*, 26(4): 513–548

[15] Hansen, K., and Hawkes, D. (2009). Early childcare and child development. *Journal of Social Policy*, 38(02), 211-239.

[16] 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.

[17] Heckman, James J. (2006) Skill Formation and the Economics of Investing in Disadvantaged Children. *Science*, 312(5782): 1900-1902

[18] Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A., & Yavitz, A. (2010). The rate of return to the HighScope Perry Preschool Program. *Journal of public Economics*, 94(1), 114-128.

[19] Istat (2014) L'offerta comunale di asili nido e altri servizi socio-educativi per la prima infanzia.

[20] 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.

Appendix

Table 5: Graphical competences on isee

treatment	0.736**
	(0.282)
isee	0.0344
	(0.052)
sex	0.381***
	(0.093)
civil_status	0.139
	(0.116)
siblings_older	0.184
	(0.156)
siblings_younger	-0.356*
	(0.176)
preference	-0.0400
	(0.262)
education_mother	0.262*
	(0.076)
education_father	-0.0267
	(0.105)
employment_mother	-0.0165
	(0.172)
employment_father	0.0637
	(0.114)
Observations	159

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Degree of autonomy on origin

treatment	0.772**
	(0.272)
origin	0.124***
	(0.031)
civil_status	0.0136
	(0.091)
siblings_older	-0.0500
	(0.121)
siblings_younger	0.241
	(0.213)
preference	-0.722
	(0.528)
education_mother1	0.0339
	(0.266)
education_father1	-0.471
	(0.302)
employment_mother1	0.0547
	(0.182)
employment_father1	0.0957
	(0.086)
Observations	126

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Degree of autonomy on isee

treatment	0.976**
	(0.321)
isee	0.0160
	(0.074)
civil_status	-0.0235
	(0.090)
siblings_older	-0.0226
	(0.141)
siblings_younger	0.217
	(0.200)
preference	-0.817
	(0.498)
education_mother1	0.0682
	(0.309)
education_father1	-0.462
	(0.331)
employment_mother1	0.0648
	(0.205)
employment_father1	0.128
	(0.067)
Observations	126

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Social relationships on origin

treatment	1.259**
	(0.463)
origin	0.0469
	(0.095)
civil_status	-0.0439
	(0.040)
siblings_older	0.0340
	(0.069)
siblings_younger	0.306*
	(0.132)
preference	-0.418
	(0.422)
participation_both	-0.228
	(0.221)
education_mother1	0.136
	(0.294)
education_father1	-0.104
	(0.073)
employment_mother1	-0.311*
	(0.143)
employment_father1	0.190
	(0.134)
Observations	129

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 9: Social relationships on isee

treatment	1.257**
	(0.485)
isee	0.0325
	(0.054)
civil_status	-0.0595
	(0.038)
siblings_older	0.0633
	(0.051)
siblings_younger	0.304*
	(0.145)
preference	-0.424
	(0.443)
participation_both	-0.239
	(0.203)
education_mother1	0.136
	(0.287)
education_father1	-0.0929
	(0.066)
employment_mother1	-0.288
	(0.148)
employment_father1	0.175
	(0.128)
Observations	129

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

In the tables presented, notation referring to education levels and occupation type are as follows: education_mother and education_father include all the possible education levels, education_mother1 and education_father1 categorize information for those who hold at least a graduate degree, those with a secondary degree and those with qualifications below secondarym education_mother2 and education_father2 distinghuish individuals who hold at least a graduate degree from the rest; employment_mother and employment_ father distinguish for no occupation, low-skilled occupation and high-skilled occupation, while employment_mother1 and employment_father1 distinguish only individuals with an high-skilled job from the rest.

Questionario per la famiglia del bambino M1AA01

1. **Composizione nucleo familiare**
   a. Indicare lo stato dei genitori

   Sposati / Conviventi ❑
   Separati / Divorziati ❑
   Genitore single ❑
   Vedovo/a ❑

   b. Indicare età e sesso dei componenti del nucleo familiare

Componente	Ruolo	Sesso	Data di nascita
1	Padre		
2	Madre		
3	Figlio iscritto alla 1 elementare nell'A.A. 2014/15		
4			
5			
6			
7			
8			
9			
10			

   a. Indicare se la famiglia è residente presso il comune di Mantova — Si ❑ No ❑

   b. Indicare se vi sono dei nonni presenti sul territorio — Si ❑ No ❑

2. **Informazioni sul bambino iscritto alla prima elementare nell'A.A. 2014/2015**
   a. Indicare la lingua parlata in casa con il bambino

   Italiano ❑
   Albanese ❑
   Arabo ❑
   Cinese ❑
   Filippino ❑
   Francese ❑
   Inglese ❑
   Rumeno ❑
   Russo ❑
   Altro (specificare_____) ❑

   b. Indicare se il bambino ha frequentato una scuola materna
   Si ❑ No ❑
   In caso di risposta negativa, specificarne il motivo:
   _____
   _____

   c. Tipologia di cura ricevuta la bambino nel periodo precedente alla scuola materna (è possibile contrassegnare più di una opzione)
   Asilo nido pubblico ❑
   Asilo nido privato ❑
   Baby sitter dedicata esclusivamente alla cura del bambino ❑
   Baby sitter dedicata anche alla cura della casa ❑
   Nonni ❑
   Genitori ❑
   Altro (specificare_____) ❑

   d. Al di là della cura indispensabile da parte dei genitori, ed escludendo valutazioni legate ai costi ed alla comodità, quale opzione ritenete migliore per lo sviluppo del bambino? (selezionare solo una opzione)
   Asilo nido ❑
   Baby sitter dedicata esclusivamente alla cura del bambino ❑
   Baby sitter dedicata anche alla cura della casa ❑
   Nonni ❑
   Altro (specificare_____) ❑

   e. Se il bambino ha frequentato un asilo nido, indicare la durata dell'utilizzo di questo servizio in termini di età del bambino:
   da _____ a _____ .

   f. Se il bambino non ha frequentato un asilo nido, indicare l'eventuale presentazione di una domanda per la frequenza — Si ❑ No ❑

   g. Chi accompagna il bambino a scuola?
   _____

   h. Chi va a prendere il bambino a scuola?
   _____

   i. Indicare il livello di partecipazione agli incontri previsti dagli insegnanti

	Padre	Madre
Meno del 50%	❑	❑
Tra il 50% e il 75%	❑	❑
Superiore al 75%	❑	❑

   j. Quanto pagate per un pasto presso la mensa scolastica?

   Tra 0 e 1 € ❑
   Tra 1 e 2 € ❑
   Tra 2 e 3 € ❑
   Più di 4 € ❑
   Più di 5 € ❑

   k. Ritenete che sia un costo eccessivo? Si ❑ No ❑

   l. Vi sono altri figli a carico che usufruiscono di una mensa scolastica? Se si, indicare a quale anno scolastico sono iscritti.
   _____
   _____

3. **Informazioni sui genitori**

   a. Grado di istruzione (selezionare solo un quadratino per genitore)

	Padre	Madre
Non applicabile	❑	❑
Licenza elementare	❑	❑
Licenza media	❑	❑
Diploma di scuola secondaria superiore	❑	❑
Laurea	❑	❑
Post laurea	❑	❑

   b. Tipo di occupazione (selezionare solo un quadratino per genitore)

	Padre	Madre
Non ha mai avuto un lavoro retribuito fuori di casa	❑	❑
Non ha un lavoro retribuito a partire dalla nascita del bambino (o del primo figlio)	❑	❑
Proprietario di piccole imprese (comprende i proprietari di piccole imprese, con meno di 25 dipendenti, come ad esempio negozi al dettaglio, servizi, ristoranti)	❑	❑
Impiegato (comprende impiegati d'ufficio, segretarie, dattilografe, programmatori, impiegati di servizio)	❑	❑
Addetto ai servizi o alle vendite (comprende personale di viaggio, addetti alla ristorazione, addetti all'assistenza personale, addetti alla protezione civile, venditori)	❑	❑
Lavoratore nel settore agricolo o ittico (comprende agricoltori e lavoratori agricoli, forestali e zootecnici, addetti alla pesca e alla caccia)	❑	❑
Artigiano o operaio specializzato (comprende muratori, carpentieri, idraulici, elettricisti, ecc., fabbri, meccanici, artigiani)	❑	❑
Operaio di fabbrica o operatore di macchinari (comprende operai di stabilimento, addetti alle macchine, alle catene di montaggio, conduttori di veicoli)	❑	❑
Lavoratore non qualificato (comprende collaboratori domestici e personale di pulizia, custodi, fattorini, facchini, portieri, braccianti, addetti alla pesca e manovali edili)	❑	❑
Dirigente o funzionario (comprende dirigenti di grandi imprese, con almeno 25 dipendenti, o dirigenti di settori di grandi compagnie, membri dei corpi legislativi, dirigenti amministrativi e giudiziari e di organismi collettivi, ufficiali delle forze armate)	❑	❑
Professionista (comprende scienziati, matematici, laureati in informatica, architetti, ingegneri, biologi e medici, insegnanti, avvocati, commercialisti, sociologi, scrittori e artisti, membri del clero)	❑	❑
Tecnico o diplomato (comprende diplomati o tecnici nei settori scientifico, ingegneristico, informatico, tecnici e assistenti in scienze biologiche e mediche, tecnici nelle scuole, nei settori finanziari e di vendita, commerciali e amministrativi)	❑	❑
Altro (specificare _____)	❑	❑

   c. Se la madre ha correntemente un lavoro retribuito, indicare a quale età del bambino è rientrata a lavorare
   _____

   d. Livello di reddito (selezionare un solo quadratino per genitore)

	Padre	Madre
Meno di 10 000 €	❑	❑
10 001 € - 15 000 €	❑	❑
15 001 € - 30 000 €	❑	❑
30 001 € - 70 000 €	❑	❑
Più di 70 000 €	❑	❑

   e. Paese d'origine

	Padre	Madre
Italia	❑	❑
Albania	❑	❑
Algeria	❑	❑
Bangladesh	❑	❑
Filippine	❑	❑
Ghana	❑	❑
India	❑	❑
Marocco	❑	❑
Nigeria	❑	❑
Repubblica Moldova	❑	❑
Repubblica Popolare Cinese	❑	❑
Altro (specificare _____)	❑	❑