DISCUSSION PAPER SERIES

DP11120
(v. 2)

COGNITIVE AND NON-COGNITIVE COSTS OF DAYCARE 0–2 FOR CHILDREN IN ADVANTAGED FAMILIES

Margherita Fort and Giulio Zanella

LABOUR ECONOMICS AND PUBLIC ECONOMICS

ISSN 0265-8003

COGNITIVE AND NON-COGNITIVE COSTS OF DAYCARE 0–2 FOR CHILDREN IN ADVANTAGED FAMILIES

Margherita Fort and Giulio Zanella

Discussion Paper DP11120
First Published 19 February 2016
This Revision 30 May 2019

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

COGNITIVE AND NON-COGNITIVE COSTS OF DAYCARE 0–2 FOR CHILDREN IN ADVANTAGED FAMILIES

Abstract

Exploiting admission thresholds to the Bologna daycare system, we show using RDD that one additional daycare month at age 0–2 reduces IQ by 0.5% (4.7% of a s.d.) at age 8–14 in a relatively affluent population. The magnitude of this negative effect increases with family income. Similar negative impacts are found for personality traits. These findings are consistent with the hypothesis from psychology that children in daycare experience fewer one-to-one interactions with adults, with negative effects in families where such interactions are of higher quality. We embed this hypothesis in a model that lends structure to our RDD.

JEL Classification: J13, I20, I28, H75

Keywords: daycare, childcare, child development, cognitive skills, non-cognitive skills

Margherita Fort - margherita.fort@unibo.it
University of Bologna

Giulio Zanella - giulio.zanella@unibo.it
University of Bologna

Cognitive and non-cognitive costs of daycare 0–2 for children in advantaged families*

Margherita Fort† Andrea Ichino‡ Giulio Zanella§

January 31, 2019

Abstract

Exploiting admission thresholds to the Bologna daycare system, we show using RDD that one additional daycare month at age 0–2 reduces IQ by 0.5% (4.7% of a s.d.) at age 8–14 in a relatively affluent population. The magnitude of this negative effect increases with family income. Similar negative impacts are found for personality traits. These findings are consistent with the hypothesis from psychology that children in daycare experience fewer one-to-one interactions with adults, with negative effects in families where such interactions are of higher quality. We embed this hypothesis in a model that lends structure to our RDD.

JEL-Code: J13, I20, I28, H75

Keywords: daycare, childcare, child development, cognitive skills, personality.

*A previous version of this paper circulated under the title "Cognitive and non-cognitive costs of daycare 0–2 for girls". An Online Appendix available at the Journal website and at the authors' webpages contains additional material and evidence. We are very grateful to the City of Bologna for providing the administrative component of the data set, and in particular to Gianluigi Bovini, Franco Dall'Agata, Roberta Fiori, Silvia Giannini, Miriam Pepe, and Marilena Pillati for their invaluable help in obtaining these data and in clarifying the many institutional and administrative details of the admission process and the organization of the Bologna Daycare System. We gratefully acknowledge financial support from EIEF, EUI, ISA, FdM, FRDB, HERA, and MIUR (PRIN 2009MAATFS_001). This project would not have been possible without the contribution of Alessia Tessari, who guided us in the choice and interpretation of the psychometric protocols. We also acknowledge the outstanding work of Matteo Escudé, Nurfatima Jandarova, Johanna Reuter, and Zheng Wang (as research assistants), of Valentina Brizzi, Veronica Gandolfi, and Sonia Lipparini (who administered the psychological tests to children), and of Elena Esposito, Chiara Genovese, Elena Lucchese, Marta Ottone, Beatrice Puggioli, and Francesca Volpi (who administered the socioeconomic interviews to parents). Finally, we are grateful to seminar participants at several universities and workshops, as well as to Josh Angrist, Luca Bonatti, Enrico Cantoni, Gergely Csibra, Joe Doyle, Ricardo Estrada, Søren Johansen, David Levine, Salvatore Modica, Cheti Nicoletti, Enrico Rettore, Giovanni Prarolo and Miikka Rokkanen for very valuable comments and suggestions.

†University of Bologna, IZA, and CESifo; margherita.fort@unibo.it.

‡European University Institute, University of Bologna, CEPR, IZA, and CESifo; andrea.ichino@eui.eu.

§University of Adelaide, University of Bologna, and IZA; giulio.zanella@adelaide.edu.au.

1 Introduction

Daycare for infants and toddlers is a convenient solution for parents who need to return to work soon after the birth of a child. Not surprisingly, enrollment rates in center-based daycare are generally growing in countries with a developed labor market.[1] Whether daycare at age 0–2 is also beneficial to children in the long run is less obvious. We study the causal effect of time spent at age 0–2 in the high-quality public daycare system offered by the city of Bologna, one of the richest Italian cities,[2] on cognitive and non-cognitive outcomes measured at age 8–14. At this age, the short-lived effects of daycare 0–2 are likely to have faded away, allowing us to explore longer-term consequences. Identification is based on a Regression Discontinuity (RD) design that exploits the institutional rules of the application and admission process to the Bologna Daycare System (BDS). This strategy allows us to compare similar children attending daycare 0–2 from 4 months of age or older up to 36 months, for periods of different length, including no attendance at all, in a context where private daycare is almost absent and extended family services are the most relevant substitute for daycare.

Applicants to the BDS provide a preference ordering over the programs for which they are eligible, and are assigned to priority groups based on observable family characteristics. Within each priority group, applicants are then ranked based on a household size-adjusted function of family income and wealth (from low to high), which we label "Family Affluence Index" (FAI). The vacant capacity of programs in a given year determines FAI thresholds such that applicants whose FAI is no greater than the threshold of their *most preferred*

[1] In the largest OECD countries for which data are available, between 2005 and 2016 the average enrollment rate changed from 43.9% to 56.7% in France; from 16.8% (year 2006) to 37.3% in Germany; from 27.3% to 35.5% in Italy; from 16.2% to 22.5% (year 2015) in Japan; from 32.7% to 55.3% in Norway; from 38.2% (year 2010) to 53.4% in South Korea; from 14.9% to 34.8% in Spain; from 37.0% to 31.5% in the UK. In the US, this rate increased from 27.4% in 2006 to 28.0% in 2011. For EU countries, the Barcelona European Council had set in 2002 a target of 33% of children in daycare 0–2 by 2010, an objective that was justified as a gender policy. Daycare 0–2 is also an expensive form of subsidized early education: in 2013, average public spending per child aged 0-2 in these same countries was (at PPP) $6,200 in France, $3,400 in Germany, $1,200 in Italy $3,900 in Japan, $9,600 in Norway, $7,000 in South Korea, $1,400 in Spain, $1,000 in the UK, and $700 in the US (source: OECD Family Database, tabulations PF3.1 Public spending on childcare and early education and PF3.2 Enrolment in childcare and pre-school).

[2] Bologna, about 400k inhabitants in 2019, is the 7th largest Italian city and is the regional capital of Emilia Romagna, in the north of the country. The daycare system that we study is a universal crèche system (*asilo nido*) which, in this region, is renowned for its high-quality even outside the country (Hewett, 2001).

program receive an admission offer to that program. Those with a higher FAI are either admitted to a program that they prefer less or, in some cases, are excluded from all programs. The administrative data we received from the City of Bologna contain the daily attendance records of each child but no information on outcomes. Thus, between May 2013 and July 2015 we interviewed a sample of children from dual-earner households with cohabiting parents who applied for admission to the BDS between 2001 and 2005 and who were between 8 and 14 years of age at the time of the interview. Children were tested by professional psychologists using the WISC-IV protocol to measure IQ and the BFQ-C protocol to measure the "Big Five" personality traits. The accompanying parent was interviewed by a research assistant, to collect socio-economic information.

In this affluent population of daycare applicants we find that an additional month in daycare at age 0–2 reduces IQ by about 0.5%, on average. At the sample mean (116.4), this effect corresponds to 0.6 IQ points (4.7% of the IQ standard deviation) and its magnitude increases with family income. We also find that for the better-off families in this population an additional month in daycare at age 0–2 reduces agreeableness and openness by about 1% and increases neuroticism by a similar percentage.

To interpret these findings, we model how children are affected by the decisions of their parents, who face a trade-off between spending time at work, which increases family income and improves child outcomes indirectly, and spending time with their offspring, which enhances child development directly. We allow the trade-off to involve daycare programs that may be of a different quality than home care. Moreover, home care may be of a better quality in more affluent households. This hypothesis is supported by a psychological literature emphasizing the importance of one-to-one interactions with adults in child development during the early years of life, and that these interactions are more effective if complemented by high human capital and high income.[3] In the BDS setting, the adult-to-child ratio is 1:4 at age 0 and 1:6 at age 1–2 (at the time our data refer to), while the most frequent care modes when daycare 0–2 is not available are parents, grandparents, and nannies, all of which imply an adult-to-child ratio close to 1.

The central theoretical insight from the model is that when daycare time increases, child

[3]See, in particular, Csibra and Gergely (2009, 2011). Other references are reviewed in Section 7.

skills decrease in a sufficiently affluent household because of the higher quality of home care. However, given the high earning potential of an affluent parent and the possibility to substitute high-quality informal care with the less expensive daycare provided by the BDS, the loss of child ability is more than compensated by an increase of household consumption. Therefore, the affluent parent takes advantage of the offer of the most preferred program even if it decreases child skills, as long as the parent cares enough about household consumption. For a less affluent household, instead, the offer of the most preferred program increases both household consumption and child skills, because home care is of a lower quality than daycare. The RD estimand around the FAI thresholds determining whether a child is offered her most preferred BDS program, identifies a well-defined weighted average of these heterogenous effects of daycare attendance. Following Card et al. (2015), we show that this estimand can be interpreted as a weighted average of Treatment-on-the-Treated (TT) effects defined by Florens et al. (2008).

After summarizing the relevant literature in Section 2, we present the theoretical model in Section 3 and the institutional setting in Section 4. Section 5 describes the interview process to collect child outcomes. Section 6 shows how the theoretical model maps into the RD framework and presents our results. Finally, Section 7 reviews the psychological literature providing support for our interpretation of the evidence, Section 8 discusses alternative interpretations, and Section 9 concludes.

2 Previous research

This study contributes to the economic literature that investigates how early life experiences shape individual cognitive and non-cognitive skills.[4] This literature typically distinguishes between daycare 0–2 (e.g., crèches) and childcare 3–5 (e.g., preschool/kindergarten programs). Economists devoted considerable attention to the latter, often with a special focus on disadvantaged kids, while paying less attention to the former, especially in more advantaged families.[5]

[4]See Borghans et al. (2008), Almond and Currie (2011), Heckman and Mosso (2014) and Elango et al. (2016) for recent surveys.

[5]Duncan and Magnuson (2013) provide a meta analysis of the large literature on childcare 3–5, concluding that these programs improve children "pre-academic skills, although the distribution of impact estimates is

Not so in other disciplines. In a four-decade-old review, Belsky and Steinberg (1978) summarized the findings of daycare research in psychology, reporting benefits on standardized measures of intelligence for disadvantaged children but no effects on children from advantaged families, and negative effects on non-cognitive outcomes across the board. Subsequent reviews by Belsky (1988, 2001) confirmed negative consequences of daycare. A central theme in Belsky and Steinberg (1978) is that families are affected in different ways by daycare because the latter substitutes for family care of different quality during a developmental stage when adult-child interactions are of paramount importance. Our contribution to this literature is the formalization of this idea in an economic model and its test in a causal framework.

In recent years economists have devoted more attention to the impact of very early childhood interventions on children's outcomes, reporting mixed results. A first group (Felfe and Lalive, 2018, in Germany and Drange and Havnes, 2018, in Norway)[6] reports results that apply to a relatively disadvantaged population, finding desirable effects of early daycare attendance for both cognitive and non-cognitive outcomes, concentrated in particular on girls. On the contrary, Baker, Gruber, and Milligan (2008) found undesirable effects on all types of cognitive and non-cognitive outcomes when studying the 1997 universal early daycare extension in Quebec (a reform that heavily subsidized daycare for 0–4 children in a relatively advantaged population).[7] Kottelenberg and Lehrer (2017) dig deeper into the Quebec data showing that the negative average estimate hides a positive effect for the less advantaged in that population. Similarly, Duncan and Sojourner (2013) use data from the Infant Health and Development Program (IHDP) in the US, finding that it "boosted the cognitive ability of low-income children much more than the cognitive ability of higher-income children" (p. 947). In terms of standardized magnitude (for 1 month of attendance), the positive effects

extremely wide and gains on achievement tests typically fade over time." (p. 127). See also Puma et al. (2012), Elango et al. (2016), Carneiro and Ginja (2014), Havnes and Mogstad (2015) and Felfe, Nollenberger, and Rodrguez-Planas (2015). To the best of our knowledge, only Gormley and Gayer (2005), Cascio and Schanzenbach (2013), and Weiland and Yoshikawa (2013) investigate the heterogeneity of effects of 3–5 programs by family affluence, finding smaller or zero effects for children in advantaged families.

[6]Precursors of these more recent papers are the Carolina Abecedarian Study (Campbell and Ramey, 1994; Anderson, 2008), the Milwaukee Project (Garber, 1988), and Zigler and Butterfield (1968).

[7]More recently, these authors confirmed the long-run persistence of undesirable effects, with negligible consequences for cognitive test scores and with some of the losses concentrated on boys (Baker, Gruber, and Milligan, 2015). Three other recent studies provide indirect evidence consistent with the finding for Quebec, by exploiting policy changes that altered the amount of maternal care a child receives at age 0–2: Carneiro, Løken, and Salvanes (2015) for Norway; Bernal and Keane (2011) and Herbst (2013) for the US.

found for Germany, Norway, and the US are about 0.3%, and the negative effects for Quebec is about 0.2%. These sizes are comparable to ours.

As for the sign, in line with the Quebec studies we find a negative effect because our sample and identification provide estimates for relatively affluent families with employed and cohabiting parents in one of the richest and most highly educated Italian cities. This is precisely a context in which the quality of one-to-one interactions at home is likely to be better than the corresponding quality in daycare 0–2, even if Bologna is renowned for the high standard of its daycare system. Moreover, since girls are more capable than boys of exploiting these interactions in early development (see Section 7), this is a context in which negative effects for girls should emerge more clearly, and in fact they do in our sample.[8] A second possible reason for the negative sign of our estimate pertains to the characteristics of the daycare environment. For instance, both Felfe and Lalive (2018) and Drange and Havnes (2018) study daycare settings with an adult-to-child ratio of 1:3. The corresponding ratio at the BDS facilities during the period that we study was 1:4 at age 0 and 1:6 at ages 1–2, similar to the prevailing ratio in the Quebec context. In this respect, our study suggests that attention should be paid to the adult-to-child ratio when designing daycare 0–2 programs.

Finally, as far as cognitive outcomes are concerned, our negative estimate refers to IQ measured by professional psychologists at age 8–14 (as in Duncan and Sojourner, 2013, who use measures of IQ at age 1–5 and 8), while other studies focus on math and language test scores, or on indicators of school readiness (Drange and Havnes, 2018; Felfe and Lalive, 2018). There is a general consensus that IQ, in addition to being a clinical and standardized indicator, is correlated with a wide set of long term outcomes, including in particular levels of education, types of occupation and income (see, for example, Gottfredson, 1997). Currie (2001) notes that the literature on the effects of childcare has shifted towards the use of learning test scores or indicators of school readiness as outcomes, probably because "gains in measured IQ scores associated with early intervention are often short-lived" (p. 214).[9]

[8] Kottelenberg and Lehrer (2014a,b) focus as well on the heterogeneity of effects by gender (and age) using data from the Quebec expansion. Effects for girls are also studied, with different results, by Carneiro, Løken, and Salvanes (2015) and by Elango et al. (2016). Both positive effects (on emotional regulation, motor skills, and eating) and negative effects (on reasoning and memory) of daycare 0–2 in the short run are found by Noboa-Hidalgo and Urza (2012) in Chile for children with a disadvantaged background.

[9] The cost of measuring IQ, compared with the increasing availability of almost free administrative data on school outcomes, may contribute to explaining why IQ is used less as an outcome in this literature.

From this viewpoint, a contribution of our study is to show that instead daycare 0–2 may have long term negative effects also on IQ. As for non-cognitive outcomes, our results are in line with Baker, Gruber, and Milligan (2008, 2015) even though, different from them, we focus on the Big Five personality traits.

3 Theory

Consider a population of parents who face a trade-off between spending time with their offspring, which enhances child development directly, and spending time at work, which increases household income and so improves child outcomes indirectly.[10] A household is composed of a parent and a child and there are two periods in life: age 0–2 and post age 0–2. A parent values household consumption, c, and the ability of the child, θ.[11] The utility function is

$$v(c, \theta) = c + \alpha\theta, \tag{1}$$

where $\alpha > 0$ is the weight of child ability in parental preferences. Two forms of child care are available: parental child care and a rationed daycare system.[12] Since the parent does not value leisure, she splits the time endowment between work for pay, h, and parental child care, τ_g, so that a parent's time constraint can be written as $h + \tau_g = 1$. The daycare system offers a set of $\mathcal{Z} \geq 2$ programs indexed by $z \in \{0, \frac{1}{\mathcal{Z}-1}, \frac{2}{\mathcal{Z}-1} \cdots, 1\}$. Vacancies are limited and are allocated via a strategy-proof mechanism that will be described below. Each program z is characterized by a combination of quality, $q_d(z)$, and cost of attendance per unit of time, $\pi_d(z)$. This cost is expressed in units of consumption and it reflects two components: a transportation cost $k(z)$ and an attendance fee ϕy_{-1}, with $\phi < 1$, that is identical for all programs and is proportional to past household income $y_{-1} = wh_{-1}$, where $w = w(\theta_g)$ is the

[10] Our framework builds on Becker (1965) as well as on Carneiro, Cunha, and Heckman (2003) and Cunha and Heckman (2007). A similar framework is employed by Bernal (2008).

[11] We are indifferent between treating parental preferences over θ as direct – i.e., the parent values child ability per se – or indirect – i.e, the parent values the future earnings of the child, which increase in the child's cognitive or non-cognitive skills. We also assume that household consumption benefits both the parent and the child.

[12] The more realistic case that allows for a third type of care acquired from babysitters or within the extended family is considered in the Online Appendix.

wage rate (increasing in parental skill θ_g) and $h_{-1} \in [0,1]$ is past labor supply.[13] Therefore, $\pi_d(z) = k(z) + \phi y_{-1}$. Without loss of generality, we assume that daycare programs can be ordered in a way such that the function $s(z) = \alpha q_d(z) - k(z)$ is strictly increasing in z.[14] We later show that, thanks to this assumption, derived utility of parents is also increasing in z and therefore z is the parents' ranking of programs.

Skills are determined at age 0–2 by parental ability, θ_g, household income given by $y = hw$, and the quality of care. Denoting by τ_d time spent by the child in daycare, the technology of skill formation is

$$\theta = \eta(\theta_g) + q_g y \tau_g + q_d(z)\tau_d, \tag{2}$$

where $\eta(.) > 0$ captures inherited parental ability, which is also the baseline child ability, and $q_g y$ represents the quality of child care at home. This specification reflects the idea that while all children attending the same program enjoy the same daycare quality $q_d(z)$, the quality of parental care, $q_g y$, differs among children because parental quality q_g is complemented by the cognitive and economic resources of the household, summarized by y. Such complementarity introduces a convexity which ensures that the parent does not necessarily specialize in producing either child quality or income. Eq. 2 also indicates that a child would benefit from parental ability θ_g directly, even if the parent had zero earnings.

A child requires a fixed amount of care time $b \in (\frac{1}{2}, 1)$.[15] Therefore, the chosen child care arrangement must satisfy $\tau_g + \tau_d = b$, so that parental care and daycare are perfect substitutes at rate 1 in child care time but are substitutes at rate $\frac{q_d(z)}{q_g y}$ in child development. Replacing the time constraints, for each daycare program the parent solves

$$\max_{c,\tau_d} c + \alpha\theta \quad \text{s.t.} \quad \begin{cases} c + \pi_d \tau_d = w(1 - b + \tau_d) \\ \theta = \eta(\theta_g) + q_g w(1 - b + \tau_d)(b - \tau_d) + q_d(z)\tau_d \\ \pi_d = k(z) + \phi y_{-1} \\ 0 \leq \tau_d \leq b \end{cases} \tag{3}$$

[13] In the Bologna context, conditional on a program, parents decide the number of days of attendance but not the number of hours during the day (with few special exceptions). Since every day of attendance requires transportation, total travel cost is proportional to attendance, and this is reflected in our assumption about $k(z)$.

[14] Consistent with the institutional setting described in Section 4, there are no ties.

[15] This restriction means that the child needs active care for at least half the time, but for less than the entire time (for instance, the parent can work while the child sleeps).

The key trade-off in this problem is that if the wage rate is greater than the unit cost of daycare, then increasing τ_d adds resources for consumption and to complement home care. At the same time, however, it reduces parental time with the child, causing a negative direct effect on child ability if the quality of daycare is worse than the quality of parental care.

Let $\underline{z} \geq 0$ be the "reservation program", to be determined below. The parent applies for the subset of programs for which the optimization problem has an interior solution or attains a corner characterized by strictly positive attendance, if the child is offered admission, as well as for the reservation program. Therefore, $A = \{\underline{z}, \ldots, 1\}$ is the application set of the parent. Consider first the interior solution for $z \in A$. This is given by

$$\tau_d^*(z) = \frac{2b-1}{2} + \frac{w + \alpha q_d(z) - k(z) - \phi y_{-1}}{2\alpha q_g w}, \tag{4}$$

and parental utility at the optimum is

$$v^*(z) = \tau_d^*[w - \phi y_{-1} + \alpha q_g w(2b - 1 - \tau_d^*) + \alpha q_d(z) - k(z)] + V, \tag{5}$$

where $V = w(1-b)(1+\alpha q_g b)+\alpha\eta(\theta_g)$. To simplify the analysis, let the number of programs be large enough so that the $[0, 1]$ interval offers a convenient approximation to the set of available programs and z is continuous. Then, using the envelope theorem, $\frac{dv^*}{dz} = [\alpha q_d'(z) - k'(z)]\tau_d^*$ and since we have assumed that the ordering of programs implied by z is such that $s(z) = \alpha q_d(z) - k(z)$ is strictly increasing in z, it follows that also derived utility $v^*(z)$ must be strictly increasing in z. Therefore, the following condition holds,

$$\alpha q_d'(z) - k'(z) > 0, \tag{6}$$

and the ranking z is consistent with derived preferences over programs.[16]

Admission offers are made based on eligibility thresholds, $\mathcal{Y}_z$. If $y_{-1} \leq \mathcal{Y}_z$ then the child qualifies for program z. For a given application set A, the cutoffs $\mathcal{Y}_z$ faced by a given household are random draws from a distribution which has the same support as the distribution of past household income. Two thresholds are of special interest: $\mathcal{Y}^P \equiv \mathcal{Y}_1$, which determines

[16]As shown below in Section 4.1, Eq. 6 is satisfied, on average, in our setting: programs that are ranked higher by parents are typically closer to home, $k'(z) < 0$, and of weakly better quality, $q_d'(z) \geq 0$.

whether a child is offered her most preferred program or not ("Preferred threshold"); and the maximum threshold in A, $\mathcal{Y}^M \equiv \max_{z\in A}\{\mathcal{Y}_z\}$, which determines whether a child is offered any program or not ("Maximum threshold"). If $y_{-1} > \mathcal{Y}^M$ then the child does not qualify for any of the programs in A and so for all these programs the problem has a constrained solution $\tau_d^* = 0$. Otherwise, the child is offered the most preferred program among those for which she qualifies.

Our goal is to model how a parent reacts to the offer of the *most preferred* program $z = 1$ as opposed to the best available alternative. To this end, consider a parent whose y_{-1} is just above $\mathcal{Y}^P$ so that the child barely does not qualify for $z = 1$. The best alternative to the most preferred program may be:

Case (L): a *less* preferred program $z = \ell < 1$, if and only if $\mathcal{Y}^M > y_{-1} > \mathcal{Y}^P$;

Case (N): *no offer*, if and only if $y_{-1} > \mathcal{Y}^M = \mathcal{Y}^P$.

Note that program ℓ is not necessarily in a left neighborhood of 1, and so the ranking difference $1 - \ell$ is a discrete change even if z is continuous. Let $\Delta\tau_d^*$ be the discrete change in optimal daycare time induced by the offer of $z = 1$ instead of its best alternative. Given conditions (4) and (6), starting at an interior solution we can establish the following:

Remark 1 *(First stage). The offer of the most preferred program increases daycare time. The increase differs in cases (L) and (N):*

$$\textit{if (L) then} \quad \Delta\tau_d^* \approx \frac{\alpha q_d'(\ell) - k'(\ell)}{2\alpha q_g w}(1-\ell) > 0, \tag{7}$$

$$\textit{if (N) then} \quad \Delta\tau_d^* = \tau_d^*(1) - 0 = \frac{2b-1}{2} + \frac{w + \alpha q_d(1) - k(1) - \phi y_{-1}}{2\alpha q_g w} > 0. \tag{8}$$

Consider now the corner solutions. Remark 1 allows us to characterize the application set of a parent by modelling the possibility that there exists a program $\underline{z} \in [0, 1]$ for which $\tau_d^*(\underline{z}) = 0$. From Eq. 4, this happens when

$$w - k(\underline{z}) - \phi y_{-1} + \alpha(q_g w(2b-1) + q_d(\underline{z})) = 0. \tag{9}$$

If a program satisfying this condition exists, then $\underline{z}$ is the reservation program. If $\underline{z} = 0$

then the application set $A = [\underline{z}, 1]$ coincides with the entire set of programs. If $0 < \underline{z} < 1$ then a corner solution $\tau_d^*(\zeta) = 0$ exists for any $\zeta \leq \underline{z}$. If there is no program for which utility when $\tau_d^* > 0$ is greater than utility when $\tau_d^* = 0$ then the parent does not apply to any program. Finally, there may exist a program $\overline{z} \in (\underline{z}, 1)$ such that $\tau_d^*(\overline{z}) = b$. In this case, for the subset of programs ranked $\zeta \geq \overline{z}$, the optimal daycare time is at the $\tau_d^*(\zeta) = b$ corner. Then, if $\overline{z} \leq \ell$, $\Delta\tau_d^* = 0$ in case (L) and $\Delta\tau_d^* = b > 0$ in case (N). And if $\ell < \overline{z} \leq 1$ then $\Delta\tau_d^* = b - \tau_d^*(\ell) > 0$ in case (L) and $\Delta\tau_d^* = b > 0$ in case (N).

Focusing on the cases in which the offer of the most preferred program induces a strictly positive increase in daycare attendance, the key prediction of the model concerns the response of child ability at the optimum,

$$\theta^* = \eta(\theta_g) + q_g w(1 - b + \tau_d^*(z))(b - \tau_d^*(z)) + q_d(z)\tau_d^*(z), \tag{10}$$

to an increase in optimal daycare time following the offer of $z = 1$. The response differs in cases (L) and (N):

$$\text{if (L) then } \Delta\theta^* \approx \frac{(w - k(z) - \phi y_{-1})k'(\ell) + \alpha^2 q_d'(\ell)(q_g w(2b-1) + q_d(\ell))}{2\alpha^2 q_g w}(1-\ell); \tag{11}$$

$$\text{if (N) then } \Delta\theta^* = q_g w(1 - b + \tau_d^*(1))(b - \tau_d^*(1)) + q_d(1)\tau_d^*(1) - q_g w(1-b)b. \tag{12}$$

Dividing these expressions by Eqs. 7 and 8 gives the treatment effect of interest:

$$\text{if (L) then } \frac{\Delta\theta^*}{\Delta\tau_d^*} \approx \frac{(w - k(\ell) - \phi y_{-1})k'(\ell) + \alpha^2 q_d'(\ell)(q_g w(2b-1) + q_d(\ell))}{\alpha(\alpha q_d'(\ell) - k'(\ell))}; \tag{13}$$

$$\text{if (N) then } \frac{\Delta\theta^*}{\Delta\tau_d^*} = \frac{q_g w(1 - b + \tau_d^*(1))(b - \tau_d^*(1)) + q_d(1)\tau_d^*(1) - q_g w(1-b)b}{\tau_d^*(1)}. \tag{14}$$

These quantities may be positive or negative. However, the following remark holds:

Remark 2 *(Treatment effect). Under the conditions specified below, there exist values $\widetilde{w}_L$ and $\widetilde{w}_N$ of the parental earning potential w such that*

if (L), and $k'(z) < 0$ and $q'_d(z) \geq 0$ but sufficiently small,[17] *then*

$$\frac{\Delta\theta^*}{\Delta\tau_d^*} \approx \frac{d\theta^*}{d\tau_d^*} < 0 \quad \Leftrightarrow \quad w > \widetilde{w}_L = \frac{k(z)k'(z) - \alpha^2 q'_d(z) q_d(z)}{(1-\phi h_{-1})k'(z) + (2b-1)\alpha^2 q'_d(z) q_g} > 0; \tag{15}$$

if (N), and b is sufficiently small,[18] *then*

$$\frac{\Delta\theta^*}{\Delta\tau_d^*} < 0 \quad \Leftrightarrow \quad w > \widetilde{w}_N = \frac{\alpha q_d(1) + k(1)}{1 - \phi h_{-1} + (1-2b)\alpha q_g} > 0. \tag{16}$$

In other words, in both cases (L) and (N), the skill response to more daycare at the optimum is positive in less affluent households but may be negative in more affluent ones. To see why note that, under the conditions of Remark 2, if the parent is sufficiently affluent then an increase in daycare time generates a skill loss because home care is of better quality than daycare. However, given the high earning potential of the affluent parent, consumption increases enough with the additional working time to compensate for the skill loss in terms of utility, so that the parent trades off child ability for consumption. For a less affluent parent, in addition to an analogous increase in consumption, there is a gain in terms of child skills because in this case the quality of home care is not sufficiently high compared to the quality of daycare.

Whether the skill response to a longer daycare exposure is indeed negative for more affluent children is therefore an empirical question. To answer it, we leverage the BDS institutional setting and the associated data, which we describe next.

[17] Specifically, $0 \leq q'_d(z) < \frac{-(1-\phi h_{-1})k'(z)}{(2b-1)\alpha^2 q_g}$. Section 4.1 suggests that these conditions are satisfied, on average, in our setting. If $q'_d(z) \geq 0$ and sufficiently large, then $\frac{d\theta^*}{d\tau_d^*}$ is positive at all levels of affluence.

[18] Specifically, $b < \frac{1}{2} + \frac{1-\phi h_{-1}}{2\alpha q_g}$. The second term on the RHS of this inequality is strictly positive and it may well be greater than $\frac{1}{2}$, in which case no further restriction is imposed on b beyond the assumption that $b \in (\frac{1}{2}, 1)$.

4 Institutional setting

The BDS granted us access to the application, admission, and attendance records for all the 68 daycare facilities operating in Bologna between 2001 and 2005 (of which 9 are charter). These facilities enroll, every year, approximately 3,000 children of age 0, 1, and 2 in full-time or part-time modules. Henceforth, we refer to these ages as *grades* and we use the term *program* to define a module (full-time or part-time) in a grade (age 0, 1, or 2) of a facility (68 institutions) in a given calendar year (2001 to 2005). There are 941 such programs in our data, and we have information on the universe of 9,667 children whose parents applied for admission to one or more programs of the BDS between 2001 and 2005. Parents can apply to as many programs as they wish in the grade-year combination for which their children are eligible, and they are asked by the BDS to provide a preference ordering of these programs (Section 4.1). Given these preferences, daycare vacancies are allocated by an algorithm that is equivalent to a Deferred Acceptance (DA) market design (Section 4.2).[19]

4.1 Determinants of parents' ranking of programs

Table 1 documents that parental preferences over programs systematically reflect distance from home and, to some extent, program quality as measured by their reputation. In the first panel, geo-referenced information is used to describe the distance in km between each program and the home of the eligible children in the grade-year combination of that program. Mean distance is just above 4 km (s.d. ≈ 2.2), which is also the median distance, and ranges between 100 meters and slightly more than 14 km.[20] The next panel shows that, on average, the ranking of programs is inversely related to their distance from the home of applicants. The most preferred program is typically located at a distance of 1.2 km. The second and third most preferred are located farther away by approximately 200 and 400 additional meters, respectively. The average distance of programs that are explicitly ranked by parents but that are not their most preferred is slightly less than 2 km, while the most distant programs are

[19] See Gale and Shapley (1962) and Roth (2008).

[20] These results are based on 5,602 children with two working parents (i.e., the group on which we focus our study, as explained in Section 4.2) and living within the city boundaries. For this analysis we do not consider households living outside the city boundaries because their preferences over programs are probably affected by commuting patterns on which unfortunately we have no information.

Table 1: Distance from home and quality of programs measured by their reputation

	2001	2002	2003	2004	2005
Distance statistics:					
mean [s.d.]	4.06 [2.21]	4.13 [2.27]	4.09 [2.24]	4.14 [2.27]	4.14 [2.28]
min / max	0.02 / 13.71	0.01 / 14.25	0.02 / 14.10	0.02 / 13.68	0.01 / 14.12
Mean distance from home to:					
most preferred	1.22 (0.04)	1.24 (0.04)	1.20 (0.03)	1.22 (0.03)	1.20 (0.03)
second most preferred	1.46 (0.04)	1.46 (0.04)	1.38 (0.04)	1.41 (0.04)	1.42 (0.04)
third most preferred	1.66 (0.05)	1.71 (0.05)	1.57 (0.04)	1.68 (0.04)	1.66 (0.04)
not most preferred and ranked	1.90 (0.02)	1.99 (0.02)	1.95 (0.02)	2.01 (0.02)	1.95 (0.02)
not ranked	4.29 (0.01)	4.38 (0.01)	4.31 (0.01)	4.35 (0.01)	4.36 (0.01)
Quality statistics:					
mean [s.d.]	0.01 [0.26]	0.04 [0.36]	0.03 [0.33]	-0.01 [0.43]	0.00 [0.42]
min / max	-0.86 / 0.65	-1.01 / 1.00	-1.24 / 1.40	-1.92 / 1.02	-1.34 / 1.23
Mean quality of:					
most preferred	0.07 (0.01)	0.12 (0.01)	0.10 (0.01)	0.10 (0.01)	0.11 (0.01)
second most preferred	0.06 (0.01)	0.11 (0.01)	0.11 (0.01)	0.09 (0.01)	0.10 (0.01)
third most preferred	0.06 (0.01)	0.11 (0.01)	0.10 (0.01)	0.03 (0.01)	0.05 (0.01)
not most preferred and ranked	0.05 (0.00)	0.08 (0.01)	0.08 (0.00)	0.03 (0.01)	0.04 (0.00)
not ranked	0.00 (0.00)	0.03 (0.00)	0.02 (0.00)	-0.01 (0.00)	-0.00 (0.00)

Notes: For each year, the Table reports statistics on program distance (in km) from the home of applicants and on program quality. The distance measure is based on geo-referenced information. Quality is a reputational indicator constructed in the folllowing way. First, we compute the difference between the average ranking of a program and the average ranking of its alternatives in each grade-year combination, for all the households located in a given distance cell from the program and its alternatives. The overall quality of a program is then the average of the distance-specific qualities. Each distance cell is an interval of 0.5 km (an annulus, effectively). Results are based on 5,602 children with two working parents and living within the city boundaries of Bologna. Standard errors are in parentheses. Standard deviations in brackets.

those that parents do not rank even if available in their grade-year combination. On average over all programs, moving one position down in the preference ordering is associated with an increased distance of $\approx$ 0.35–0.53 km from home. All these differences are statistically significant.

As for quality, we do not have an objective measure and we rely on a reputational indicator that we constructed in the following way.[21] Consider a set of programs, denoted by j, for which some households, denoted by i and located in a cell of distance d from all these programs, are eligible for. Each distance cell (an annulus, effectively, i.e., a region bounded by two concentric circles) is an interval of 0.5 km up to a maximum of 4 km (the distance beyond which parents typically do not rank programs), so that $d \in \{1, ..., 8\}$ denotes the eight resulting, non-overlapping cells of 0.5 km size. Let r_{ijd} be the rank of program j in the application set of household i in distance cell d.[22] The reputation of program j among households in distance cell d is defined as

$$q_{jd} = \bar{r}_{jd} - \bar{r}_{-jd}, \tag{17}$$

where $\bar{r}_{jd}$ is the average ranking of program j in distance cell d, while $\bar{r}_{-jd}$ is the average ranking of the programs different from j in the same cell. Therefore, q_{jd} measures the difference between the average ranking of program j and the average ranking of its alternatives in each grade-year combination, for all the households located in the same distance cell d from j and its alternatives. Considering different distance cells, note that each program j is compared with partially different alternatives and by different households in each of these cells. So it may be preferred in some cells but not in others. However, larger values of q_{jd} in different cells imply that j has in general a positive reputation among different groups of households and with respect to different alternatives for given distance.[23] To capture the

[21] We do not have information on program-specific teacher-to-children ratios. However, guidelines for programs in the BDS are set at the central level (Comune di Bologna, 2010), with little autonomy left to the different facilities. Specifically, the BDS strictly enforces standards concerning goals and daily planning of educational activities, and the number of teachers and square meters per child. While programs may still differ, these guidelines suggest a relatively uniform quality across programs. This uniformity is in line with the evidence based on the reputational indicator described below.

[22] For all programs that were not explicitly ranked by a parent, we impute the ranking position that follows the rank of the least preferred among the explicitly ranked programs. This imputation captures the idea that programs not ranked are all indifferently less preferred than the ranked ones. The average fraction of programs not ranked by a parent is about 90% and is constant across years.

[23] The average number of households i, for each combination of program j and distance d, is 138 (s.d. 108)

overall reputation of program j, we compute the average

$$q_j = \frac{1}{8}\sum_{d=1}^{8} \bar{r}_{jd} - \bar{r}_{-jd}. \tag{18}$$

Positive values of q_j indicate a better reputation, meaning that j is systematically more likely to beat its alternatives at all distances. Given the way it is constructed, this measure of quality is centered around zero (third panel of Table 1: s.d. $\approx 0.3 - 0.4$), but it differs across programs. For example, in 2003 the best program according to this reputational indicator, is ranked 1.4 positions better than its alternatives, while the worst program, in 2004, is 1.92 positions worse than its alternatives, on average.

Now consider, as an example, a hypothetical grade-year combination with only three available programs, $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. If all eligible households unanimously ranked these programs in the same way, ($\mathcal{A} \succ \mathcal{B} \succ \mathcal{C}$) at all distances, then their reputation would be ordered as $q_{\mathcal{A}} > q_{\mathcal{B}} > q_{\mathcal{C}}$. In the absence of agreement among households, instead, the reputation of the three programs would be similar: $q_{\mathcal{A}} \approx q_{\mathcal{B}} \approx q_{\mathcal{C}}$. The evidence in the last panel of Table 1 suggests that there is little agreement, at least at the top of the rankings. In 2001, 2002 and 2003, there is no statistically significant difference between the average values of q_j for the programs that are ranked in the top positions. Only in 2004 and 2005 the reputation of the most preferred program (0.10 and 0.11, respectively) is significantly larger than the quality of the average less preferred but ranked program (0.03 and 0.04, respectively).

On the basis of this evidence, we conclude that in every year parents certainly prefer programs that are closer to home. As for quality, the revealed reputation of ranked programs shows some convergence of preferences on specific programs in later years, but differences among programs, if they exist, are unlikely to play a major role when parents rank them. Had these differences been of first order importance they would have showed up in Table 1. Moreover, in the population that we study, 61.6% of applicants are offered their most preferred program and 89.7% receive an admission offer; of these, 84.1% are offered one of their first three choices. In light of these facts, in the remainder of the paper we assume that $q'_d(z) \approx 0$ among the top ranked programs, and in particular at the best alternative $z = \ell$ to the most preferred program $z = 1$.

and ranges between 11 (s.d. 10) in cell 1 (from 0 to 0.5 km) and 178 (s.d. 89) in cell 8 (from 3.5 to 4 km).

4.2 Admission process

Demand for admission systematically exceeds supply and there are, on average, about 1,500 vacancies for about 1,900 applicants each year. The rationing mechanism is based on a lexicographic ordering of applicants. At a first level, applicants to each program are assigned to priority groups based on observable family characteristics. First (highest priority), children with disabilities. Second, children in families assisted by social workers. Third, children in single-parent households, including those resulting from divorce or separation. Fourth, children with two cohabiting and employed parents. Fifth, children in households with two cohabiting parents of whom only one is employed. For brevity, we refer to these priority groups as "Baskets" 1 to 5. At a second level, within each of these five baskets children are ranked according to a Family Affluence Index (FAI). This is an index of family income and net wealth, adjusted for family size.[24] Families with a lower value of the index (i.e., less affluent families) have higher priority within a basket. The Deferred Acceptance algorithm determines for each program a "Final" FAI admission threshold, which is defined as the FAI of the most affluent child who receives an offer for that program and accepts it. Given the application set A, these thresholds are effectively random numbers for an applicant. A child qualifies for all programs in A whose Final thresholds are greater than the child's FAI, and is offered the most preferred program among these.

At the end of the admission process, children can be classified in three mutually exclusive and exhaustive ways: the "admitted and attendants", who have received an admission offer and have accepted it; the "reserves", who have not received any offer; the "waivers", who have received an admission offer and have turned it down. Children who are "reserves" or "waivers" in a given year may re-apply, be offered admission, and attend daycare in later years, as long as they are not older than 2. We consider only *the first application of each child*. Thus, the possibility to turn down an offer (or to be rejected) and to re-apply and attend later is one of the reasons of fuzziness in our RD design. Attending children are charged by the BDS a monthly fee that depends on their FAI but that is independent of actual days of attendance during the month. The fee schedule is well known to potentially

[24] The Online Appendix to Section 4 provides details about how this index is constructed.

interested families before they decide whether to apply,[25] and it is continuous by construction at the admission thresholds that we will use in our design.

4.3 How FAI thresholds can be used for the RD design

To ensure a greater homogeneity of the interview sample (to be described in Section 5), we restrict the entire analysis to children in "Basket 4" (i.e., children with both parents employed and cohabiting at the time of the application), which is the largest group of applicants (about 70% of the total): 6,770 *first applications* to 911 programs originate from this basket in the period 2001-2005. Of these programs, 80 end up with no vacancies for Basket 4 children (i.e., the Final FAI threshold is in Basket 1, or 2, or 3); 285 have sufficient capacity for all Basket 4 applicants (i.e., the final FAI threshold is in Basket 5), and 546 offer admission to some but not all the Basket 4 applicants (i.e., the Final FAI threshold is in Basket 4). The remaining 30 (to reach the total of 941 programs) do not receive applications from Basket 4 households. Some tables and figures below are based on sub-groups of this sample, for the reasons explained in the respective notes. The population of applicants is relatively affluent, particularly in Basket 4. The average FAI across the five baskets is about €20k (constant 2010 prices), corresponding to a gross annual household income of about €54k. In Basket 4 the average FAI is about €25k, corresponding to an income of about €67k. This is roughly twice the average annual gross household income in Italy at the time the data refer to.[26]

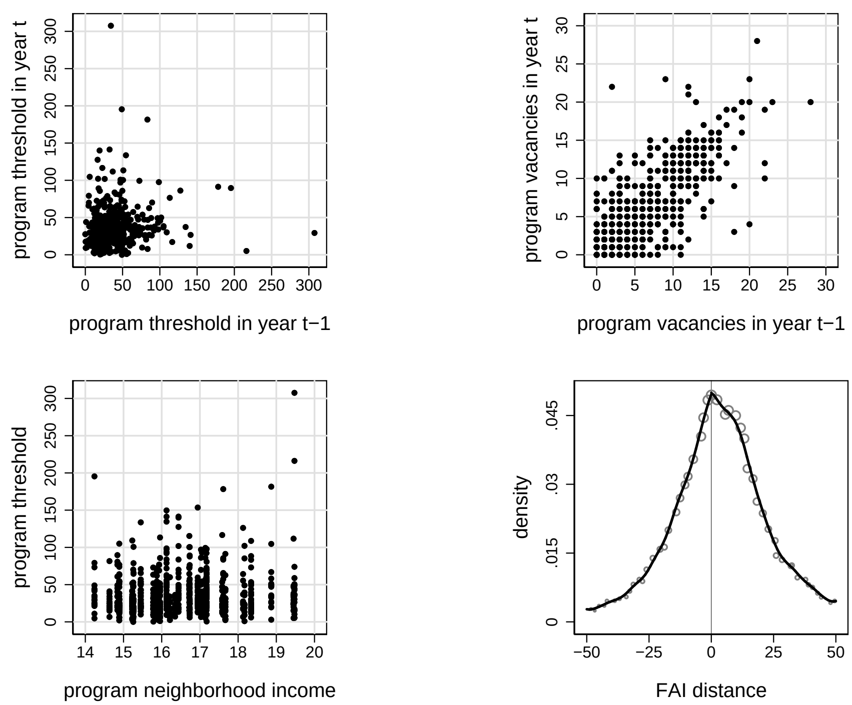
In this institutional setting, parents cannot predict Final FAI thresholds and thus cannot manipulate their FAI to secure an admission offer. If FAI thresholds were persistent across years, it would be easy for them to find out the final thresholds of the programs they wish to apply for. This is not the case: in a regression of the threshold in year t on the threshold in year $t-1$ (or the most recent previous threshold) for each program, the slope is estimated with high precision to be low (0.14; s.e. 0.04) and the R^2 is just 0.02. The top-left panel of Figure 1 illustrates this lack of persistence. For an accurate guess of FAI thresholds, families

[25] The fee is an increasing step function of the FAI. This function increases stepwise along brackets that are about €500 wide, with an initial step (from a FAI of zero) of €17 per month and then constant steps of about €6, before reaching the maximum fee of €400 per month independently of household income. The kink at which the daycare fee becomes regressive is located at a FAI of about €30k, roughly corresponding to a gross annual family income of about €80k (all these values are expressed in 2010 euros).

[26] Additional descriptive statistics are provided in the Online Appendix to Section 4.

would need a formidable amount of additional information.[27] The top-right panel similarly shows that, for each program, given the number of Basket 4 vacancies in year $t-1$ there is substantial variability in the number of vacancies in year t. Moreover, the bottom-left panel shows that, given the low degree of socioeconomic segregation across neighborhoods in Bologna, the distribution of thresholds within neighborhoods is almost invariant to changes in average neighborhood income and that the dispersion of these thresholds is large within each neighborhood. Thus, even if parents tried to manipulate their FAI, they would not know by how much the index should be reduced in order to receive an offer from a specific program.

Figure 1: Final FAI thresholds, Basket 4 vacancies over time, and Final FAI density



Notes: Top panels: each dot represents a program and the coordinates are either the Final FAI thresholds of that program in two consecutive periods (left panel), or the vacant capacity for Basket 4 children in two consecutive periods (right panel). Sample: 238 programs with rationing for Basket 4 children in two consecutive periods. Bottom-left panel: each dot represents a program and the coordinates are the program threshold on the vertical axis and the average income of the program neighborhood on the horizontal axis. Bottom-right panel: the circles represent the frequency distribution inside €2k bins (circle size is proportional to population size in the bin), plotted as a function of the distance (thousands of real €) of a child's FAI from her Final FAI threshold ("FAI distance"). The bold lines are from separate LLR on the underlying individual observations on each side of the cutoff, with a triangular kernel and a bandwidth of €5k. Sample: 5,861 children with two working parents, born between 1999 and 2005, whose parents first applied for admission between 2001 and 2005 to programs with rationing, and whose FAI distance from the Final FAI thresholds is at most €50k and is different from zero. FAI stands for Family Affluence Index.

[27]For example: the vacant capacity of the programs they wish to apply for, the number of applicants to these programs, the FAI of each applicant, how other applicants rank programs, and how many admitted children in each program turn down the offer they receive.

Additional support for this claim is provided by the continuity of the FAI density and of the distribution of pre-treatment covariates. In the bottom-right panel of Figure 1, the density of observations is plotted stacking thresholds and centering them at zero so that the FAI distance from each threshold is the running variable.[28] The McCrary (2008) test rejects the existence of a discontinuity: the gap in the (log) density at the cutoff is -0.007, with a standard error of 0.055. As for pre-treatment covariates, the continuity of the distribution of five relevant ones that we observe in the universe (birth day, FAI, average income in the city neighborhood where the program is located, number of siblings at the first application, and number of programs in the application set) is assessed using the test of Canay and Kamat (2018). Results are reported in the first column of Panel A of Table 2 and the test never rejects the null that the distribution of any covariate is continuous at the Final cutoff in the Basket 4 universe.[29]

Given the absence of any evidence of manipulation of the admission process, it would seem natural to use observations around each Final FAI threshold for the RD design, but this would be problematic because children applying to many programs would be over-represented. Specifically, reserve children would appear as many times as the number of programs they apply for, while admitted children and waivers would appear as many times as the number of programs they qualify for. Mapping the model of Section 3 into this institutional setting allows us to circumvent this problem by associating every child with one Final FAI threshold only, i.e., the unique threshold of her most preferred program, $\mathcal{Y}^P$.[30]

Consider, as an example, a group of households who apply for the first time in a given year to the same set of five programs whose identity is denoted by $j \in \{\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{D}, \mathcal{E}\}$. All these parents rank program $\mathcal{C}$ as their most preferred one, so that $z = 1$ for $j = \mathcal{C}$ and $\mathcal{Y}_{j=\mathcal{C}} = \mathcal{Y}_{z=1} = \mathcal{Y}^P$, but they may rank the remaining four programs in different ways. The comparison between the children in this group for whom y_{-1} is barely below $\mathcal{Y}^P$ (and so are

[28]The higher probability mass around the stacked thresholds is due to the fact that all programs have children immediately to the right and to the left of the cutoff, while children farther away from the cutoff are observed only for programs with a larger number of applications.

[29]In the Online Appendix we provide graphical evidence of the continuity of means of these covariates. A similar graphical analysis is provided in the same Online Appendix for all the remaining instances in which the continuity of covariates is tested in Table 2.

[30]Of the 911 programs receiving applications from parents in Basket 4, only 890 are listed as the most preferred by at least one family out of 6,575 households with non-missing FAI information.

offered their preferred program) vs. those for whom y_{-1} is barely above $\mathcal{Y}^P$ (and so are not offered their preferred program) provides a quasi-experimental variation that allows us to quantify the predictions of the model. For some households in this group, the best alternative if they do not qualify for program $\mathcal{C}$ is qualification for a less preferred program. This is Case (L) in the model. For the remaining households, $\mathcal{Y}_{\mathcal{C}}$ is also the Maximum threshold in their application set and so not qualifying for the preferred program implies not qualifying for any program at this first application. This is Case (N) in the model. However, since Final FAI thresholds are random draws from the viewpoint of applicants, conditional on the number of applications there is no self-selection of households in cases (L) or (N).[31]

This example applies to each program that is most preferred by a group of applicants and to the corresponding $\mathcal{Y}^P$ threshold. Figure 2 is constructed by normalizing to zero and pooling these different Preferred cutoffs and shows how offer rates, attendance rates, and average days of attendance change in a discontinuous way at these thresholds. The running variable is the FAI distance from the Preferred cutoff, with positive values on the right indicating a FAI lower than the threshold. This convention is maintained in all the analogous RD figures that follow. In the left and middle panels the admission and the attendance rates increase sharply (by 10.1 and 4.8 percentage points, respectively) as the FAI crosses the cutoff from higher to lower values, with some fuzziness due to the possibility of reapplying and being offered admission in a later year. These discontinuities translate into a jump of nearly two months of attendance (38 working days) in the right panel.[32] On the contrary, the frequency of observations around the Preferred FAI thresholds is continuous. Using again the McCrary (2008) test, the gap in the (log) density at the cutoff is 0.022 with a standard error of 0.13. Similarly, the second column of panel A of Table 2 shows that the Canay and Kamat (2018) test never rejects the null that the distribution of any

[31] Among the 6,575 children in the Basket 4 universe with non-missing FAI information, 4,716 ($\approx$ 72%) are in case (L).

[32] As expected, given the discussion in Section 3, children in cases (L) and (N) attend daycare for approximately the same number of months (12.3 and 11.7, respectively), if offered their preferred program. If not offered their preferred program, instead, children in case (L) attend for about 9 months in a less preferred program, while children in case (N) attend for about 5.4 months, which is an average between those who re-apply and attend in a later year and those who never attend any BDS program. The respective unconditional differences in attendance at the Preferred threshold are 3.24 (s.e. 0.20) in case (L) and 6.30 (s.e. 0.33) in case (N).

pre-treatment covariate is continuous at the Preferred threshold in the Basket 4 universe.

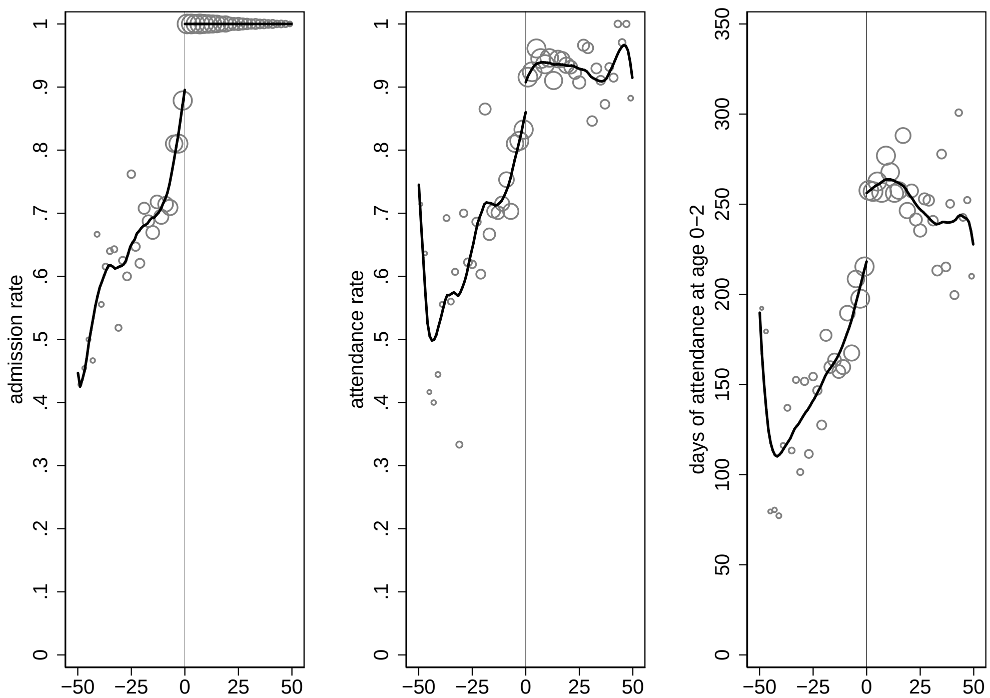
Before formalizing this approach to the identification and estimation of the effects of daycare 0–2 (Section 6), we describe in the next section how we collected the cognitive and non-cognitive outcomes that we investigate.

Table 2: Continuity of the distribution of pre-treatment covariates at the FAI thresholds.

Covariate	Final thresholds, B4 universe	Preferred thresholds, B4 universe	Preferred thresholds, interview sample
Panel A			
FAI	0.38	0.11	0.59
Siblings	0.62	0.96	0.23
Preferences	0.37	0.38	0.30
Birth day	0.23	0.32	0.49
Neighborhood income	0.72	0.41	0.92
Father's years education			0.28
Mother's years education			0.53
Father's year of birth			0.24
Mother's year of birth			0.01
Father self-employed			1.00
Mother self-employed			1.00
Cesarean delivery			0.54
Joint Test - CvM statistic	0.23	0.66	0.56
Joint Test - Max statistic	0.18	0.27	0.35
Panel B			
Invitation of universe	0.26	0.01	
Response of the invited	0.46	0.70	
Interview of universe	0.08	0.64	
Joint Test - CvM statistic	0.46	0.70	
Joint Test - Max statistic	0.44	0.69	

Notes: The table reports, in column 1, the p-values from the Canay and Kamat (2018) test of the continuity of the distribution of pre-treatment covariates at the Final FAI thresholds in the Basket 4 (B4) universe. In the remaining columns, the p-values are reported for the same test at the Preferred FAI thresholds (see Section 4.3), both in the Basket 4 universe (column 2) and in the interview sample (column 3; see Section 5). The null hypothesis is that the distribution of the covariate is continuous at the cutoff. Panel A considers pre-treatment covariates, and Panel B the invitation, response, and interview rates to be described in Section 5. Samples: 6,086 children (5,061 children for neighborhood income) in column 1; 6,300 children (5,247 for neighborhood income) in column 2; 414 children interviewed out of the invited from the Basket 4 universe with no missing values in the covariates included in Table 6 (375 for neighborhood income) in column 3. All children are born between 1999 and 2005, their parents first applied for admission between 2001 and 2005, and their FAI distance from the Final FAI thresholds is different from zero. The test is implemented using the `rdperm` package provided by Canay and Kamat (2018) using the default values chosen by these authors for the number of effective observations used from either side of the cutoff and the number of random permutations.

Figure 2: Admission offers and attendance around Preferred FAI thresholds



Notes: The circles represent offer rates (left), attendance rates (middle) and average days of attendance at age 0–2 (right) inside €2k bins, plotted as a function of the distance (thousands of real €) of a child's FAI from her Preferred FAI threshold. The size of a circle is proportional to the number of observations in the corresponding €2k bin. The bold lines are from separate LLR on the underlying individual observations on each side of the cutoff, with a triangular kernel and the optimal bandwidth selection developed in Calonico, Cattaneo, and Titiunik (2014), Calonico, Cattaneo, and Farrell (2018) and Calonico, Cattaneo, and Titiunik (forthcoming) and implemented in `STATA` by these same authors. FAI stands for Family Affluence Index. Sample: 5,101 children with two working parents, born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005, whose FAI distance from the Final FAI thresholds is at most €50k and is different from zero.

5 The interview sample

The administrative records do not contain children outcomes at any stage of their development, nor do they contain pre-treatment family characteristics beyond the few ones mentioned above. Therefore, we organized interviews in the field to collect information on outcomes and socioeconomic background for the children included in our final sample.

Between May 2013 and June 2015 we sent invitation letters via certified mail to 1,383 households (of whom 1,379 have non-missing information) with a FAI sufficiently close to Final FAI thresholds and who first applied for admission to a program of the BDS during the period 2001-2005. At the time of the invitation, children were between 8 and 14 years of age. In these letters, families were given a brief description of the research project and were invited to contact us (either via e-mail or using a toll-free phone number) to schedule an appointment for an interview. Families were informed that participants would receive a gift card worth €50 usable at a large grocery store and bookstore chain. After a few weeks from receipt of the letter, families who had not yet responded were sent a reminder via e-mail or were contacted by telephone. Upon arrival at the interview site (a dedicated space at the University of Bologna), the child was administered an IQ test (the "Wechsler Intelligence Scale for Children", WISC-IV) and a personality test (the "Big Five Questionnaire for Children", BFQ-C) by a professional psychologist, and the accompanying parent was interviewed in a separate room by a research assistant to collect socioeconomic information. Overall, each child and the accompanying parent spent about 3 hours at the interview site. Table 3 reports summary statistics of the resulting cognitive and non-cognitive test scores. Both scales are normalized by age. The relatively high average IQ of interviewed children (the normalized IQ scale has a mean of 100 for the Italian population of children in the same age range who took the WISC-IV) is in line with the high socioeconomic status of the population under study.

Table 3: IQ and personality traits, summary statistics

	Mean	Std. dev.	Median	Min	Max
IQ	116.4	12.4	116	75	158
Openness	47.7	9.0	48	18	68
Conscientiousness	47.6	10.0	47	16	71
Extraversion	47.4	9.7	48	17	72
Agreeableness	53.4	9.1	54	5	71
Neuroticism	48.3	8.2	47	28	74

Notes: The table reports summary statistics of test scores from the "Wechsler Intelligence Scale for Children" (WISC-IV, an IQ test) and from the "Big Five Questionnaire for Children" (BFQ-C, a personality test) in the interview sample. As explained in the text, the observations are 444 for IQ and 447 for the personality traits (446 for Agreeableness). For IQ, only the summary score (Full Scale IQ) is considered here; descriptive statistics for the four underlying sub-scales (verbal ability, working memory, perceptual reasoning, and processing speed) are reported in the Online Appendix to Section 5.

We obtained information for 458 children, corresponding to a response rate of 33.2% of the invited (about 40% in the proximity of Final FAI thresholds, as shown in Figure A8 of the Online Appendix). Of these interviews, only 444 provided a complete set of variables to be used in the econometric analysis when IQ is the outcome, and 447 (446 for Agreeeableness) when the Big Five personality traits are the outcome.[33] Panel B of Table 2 shows, using the Canay and Kamat (2018) test, that our sampling design produces distributions of household invitations, responses, and interviews that are all continuous at the Final FAI thresholds. Only for the invitation rate from the universe we see evidence of a discontinuity at the Preferred thresholds. This is not a source of concern in light of the other results reported in Table 2, in particular the continuity of the interview rate in the Basket 4 universe.

In order to increase the comparability of children on the two sides of the cutoffs, families were invited starting from those closer to Final FAI thresholds. The consequences of this choice are reflected in Table 4, which displays the descriptive statistics of key administrative variables for the Basket 4 universe, the invited, and the interview samples. The general pattern is, as expected, that there are no significant differences between the interviewed and the invited, while both these groups differ in some dimension with respect to the Basket 4 universe. For instance, the invited and the interviewed have a higher FAI than the universe. This is not surprising given how we invited families. The table also shows that the offer rate is higher in the universe than in the interview/invited samples. This happens because, given a large admission rate in the BDS, sampling around Final FAI thresholds implies oversampling reserves. As a result, the attendance rate is somewhat unbalanced too. Similarly, the rate at which parents are offered the preferred program is higher in the universe than in the invitation and interview samples, where it is roughly balanced. Moreover, children in the Basket 4 universe are slightly younger, have first applied for higher grades, have spent less days in daycare, and turn down admission offers at a higher rate than in the invited/interview samples. These are all consequences of the way we selected the invited families, in an attempt to increase the homogeneity of the sample and the comparability around FAI thresholds. However, these differences are not a threat to the internal validity of our RD design, given the continuity of the distribution of covariates and of the density at the thresholds. The number of preferences and the number of children in the household at first application are instead all similar across the three samples.

[33] In 7 cases, parents informed us that their children had already been tested recently using the WISC-IV, and this test does not provide reliable information if replicated. In 7 additional cases, parents did not answer all of the socioeconomic questions, thus generating missing values in some relevant pre-treatment variables. For Agreeableness, an outlier is not used in the analysis.

Table 4: Descriptive statistics for the basket 4 universe, the invited and the interview samples

Variable	Universe Basket 4	Invited	Interview	p-value
FAI at first application	24.87	26.50	27.10	0.005
	(20.50)	(19.70)	(17.55)	[0.547]
				{0.010}
N. of preferences at first application	5.42	5.29	5.59	0.199
	(3.66)	(3.42)	(3.53)	[0.120]
				{0.341}
Siblings at first application	0.48	0.49	0.54	0.749
	(0.66)	(0.65)	(0.70)	[0.151]
				{0.079}
Day of birth in the year	182.8	186.6	180.5	0.226
	(104.1)	(106.6)	(111.1)	[0.310]
				{0.674}
Offered admission at first application	0.897	0.777	0.752	0.000
	(0.304)	(0.417)	(0.432)	[0.297]
				{0.000}
Offered pref. program at first application	0.616	0.511	0.473	0.000
	(0.486)	(0.500)	(0.500)	[0.169]
				{0.000}
Waiver at first application	0.124	0.075	0.068	0.000
	(0.330)	(0.263)	(0.251)	[0.607]
				{0.000}
Year first applied	2003.1	2003.4	2003.5	0.000
	(1.43)	(1.42)	(1.38)	[0.135]
				{0.000}
Year child born	2002.0	2002.5	2002.6	0.000
	(1.58)	(1.63)	(1.62)	[0.086]
				{0.000}
Grade first applied for	0.882	0.568	0.541	0.000
	(0.786)	(0.673)	(0.676)	[0.459]
				{0.000}
Total days of attendance	212.2	223.6	230.5	0.010
	(143.3)	(151.4)	(156.3)	[0.417]
				{0.017}
Ever attended (share with days in >0)	0.847	0.784	0.782	0.000
	(0.360)	(0.411)	(0.414)	[0.916]
				{0.001}
N	6,575	1,379	444	

Notes: The table compares the means of variables from the administrative records in the Basket 4 universe (6,575 children born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005; 6,572 children for "Number of preferences at first application" due to missing values), in the sample invited for an interview (1,379 children from this universe), and in the interview sample (444 children interviewed from the universe with non-missing IQ score and covariates). The p-values in the last column refer to tests of the equality of means for the Basket 4 universe vs the invited (first row), the invited vs the interviewed (second row, in square brackets) and the Basket 4 universe vs the interviewed (third row, in curly brackets). FAI stands for Family Affluence Index.

As for external validity, Table 5 compares the means of selected socioeconomic variables available only for the interview sample with the corresponding means for a representative sample of the population of families with two employed parents of young children in large cities of Northern Italy. The comparison reveals that the interview sample is, by and large, representative of the corresponding Italian population in terms of demographics. However, parents in our sample are slightly more educated and more frequently self-employed. The higher educational attainment of these parents is relevant for the interpretation of our results because it is one of the reasons why, different from other studies, our estimated effects of daycare 0–2 refer to children who can enjoy at home a relatively richer cultural environment by Italian standards.

Table 5: The interviewed sample in comparison to the Northern Italian population

	Interview sample	Northern Italy
Child age	11.1	11.1
	(1.6)	(1.9)
Father age	47.5	46.9
	(4.8)	(4.7)
Mother age	45.1	44.9
	(4.1)	(4.7)
Years education father	14.2	13.7
	(3.7)	(2.5)
Years education mother	15.5	13.9
	(3.2)	(2.4)
Father self-employed	0.236	0.145
	(0.425)	(0.355)
Mother self-employed	0.106	0.087
	(0.308)	(0.284)
Observations	444	69

Notes: The table compares the means of variables in the interview sample with the corresponding means in the Bank of Italy Survey of Household Income and Wealth (SHIW). From the SHIW, we selected observations to mimic the Basket 4 universe of the BDS administrative files in 2001-2005. Specifically, we restricted the sample to households with two cohabiting employed parents from the 2000–2006 waves, living in cities of Northern Italy with a population of at least 200,000, and who, between 2013 and 2015, had children aged between 8 and 14. The average child age reported in the table from the SHIW is the average age of the youngest child in these households.

6 A RD design for the effect of daycare 0–2

6.1 The estimand

Our goal is to identify and estimate the average effect of additional daycare attendance on the log of child ability for children attending for τ_d days, which in our context is the average effect of Treatment on the Treated (TT) defined by Florens et al. (2008),

$$TT_{\tau_d|y_{-1}}(\tau_d, y_{-1}) \equiv \int \frac{\partial \ln \theta(\tau_d, y_{-1}, u)}{\partial \tau_d} dF_{U|\tau_d, y_{-1}}(u), \tag{19}$$

where $F_{U|\tau_d,y_{-1}}(u)$ is the c.d.f. of individual heterogeneity U conditional on attendance equal to τ_d and FAI equal to y_{-1}.[34] We follow Appendix A.2 and Remark 3 of Card et al. (2015)[35] to show that a fuzzy RD design around a specific Preferred FAI threshold $\mathcal{Y}^P$ can be used to identify a weighted average of the causal effect of interest on the set of children whose most preferred program has this admission threshold, who react to the offer of their most preferred program vs. the best alternative, and who all apply for a given number of programs, as exemplified in Section 4.3.

Our setting is characterized by unobserved determinants of attendance and by the possibility that a child is in cases (L) or (N), as discussed in Section 3. To accommodate these features, we write $\tau_d = \tau_d(y_{-1}, \omega, e)$, where e is the realization of the determinants of non-compliance E (possibly correlated with U) and ω is the realization of $\Omega = \mathbb{I}(\mathcal{Y}^P \neq \mathcal{Y}^M)$, where $\mathbb{I}(\cdot)$ is the indicator function. Thus, the RD estimand at this cutoff can be written as

$$\beta(\mathcal{Y}^P) = \frac{\lim\limits_{y_{-1}\to\mathcal{Y}^{P,r}} \mathbb{E}\left[\ln \theta(\tau_d(y_{-1}, \omega, e), y_{-1}, u)|y_{-1}\right] - \lim\limits_{y_{-1}\to\mathcal{Y}^{P,l}} \mathbb{E}\left[\ln \theta(\tau_d(y_{-1}, \omega, e), y_{-1}, u)|y_{-1}\right]}{\lim\limits_{y_{-1}\to\mathcal{Y}^{P,r}} \mathbb{E}\left[\tau_d(y_{-1}, \omega, e)|y_{-1}\right] - \lim\limits_{y_{-1}\to\mathcal{Y}^{P,l}} \mathbb{E}\left[\tau_d(y_{-1}, \omega, e)|y_{-1}\right]}, \tag{20}$$

where $y_{-1} \to \mathcal{Y}^{P,r}$ indicates that the FAI approaches the Preferred cutoff from the right, and

[34] Under our assumptions, around the Preferred FAI threshold and conditioning on observable covariates (most notably the number of applications), child ability at the optimum can be written in compact form as

$$\theta = \eta(\theta_g) + q_g \frac{y_{-1}}{h_{-1}}(1 - b + \tau_d(z))(b - \tau_d(z)) + q_d \tau_d(z) = \theta(\tau_d, y_{-1}, u),$$

where here and in what follows we omit the * that in Section 3 denotes values at the optimum.

[35] See https://www.econometricsociety.org/sites/default/files/ECTA11224SUPP.pdf for the supplementary material of Card et al. (2015).

analogously for $y_{-1} \to \mathcal{Y}^{P,l}$ from the left.[36] The Online Appendix to Section 6 shows that this estimand is a weighted average (over Ω, E and U) of the causal effect of interest, i.e.,

$$\beta(\mathcal{Y}^P) = \int \frac{\partial \ln \theta(\widetilde{\tau}_d(\mathcal{Y}^P, \omega, e), \mathcal{Y}^P, u)}{\partial \tau_d} \psi(\omega, e, u, \mathcal{Y}^P) dF_{\Omega,E,U}(\omega, e, u), \tag{21}$$

where $\widetilde{\tau}_d(\mathcal{Y}^P, \omega, e)$ is a value of daycare attendance between $\tau_d^r(\mathcal{Y}^P, \omega, e) \equiv \lim\limits_{y_{-1} \to \mathcal{Y}^{P,r}} \tau_d(y_{-1}, \omega, e)$ and $\tau_d^l(\mathcal{Y}^P, \omega, e) \equiv \lim\limits_{y_{-1} \to \mathcal{Y}^{P,l}} \tau_d(y_{-1}, \omega, e)$, i.e., the levels immediately to the right and immediately to the left of $\mathcal{Y}^P$ for given realizations of Ω and E, and where

$$\psi(\omega, e, u, \mathcal{Y}^P) = \frac{(\tau_d^r(\mathcal{Y}^P, \omega, e) - \tau_d^l(\mathcal{Y}^P, \omega, e)) \frac{f_{Y_{-1}|\omega,e,u}(\mathcal{Y}^P)}{f_{Y_{-1}}(\mathcal{Y}^P)}}{\int (\tau_d^r(\mathcal{Y}^P, \omega, e) - \tau_d^l(\mathcal{Y}^P, \omega, e)) \frac{f_{Y_{-1}|\omega,e}(\mathcal{Y}^P)}{f_{Y_{-1}}(\mathcal{Y}^P)} dF_{\Omega,E}(\omega, e)}. \tag{22}$$

These weights imply that the children contributing to the estimand are the treated whose attendance changes at the cutoff when they are offered their most preferred program.

To make explicit the presence of cases (L) and (N), Eq. 21 can be written as

$$\begin{aligned}
\beta(\mathcal{Y}^P) \;&=\; \int \Big[\omega \frac{\partial \ln \theta(\widetilde{\tau}_d(\mathcal{Y}^P, 1, e), \mathcal{Y}^P, u)}{\partial \tau_d} \psi(1, e, u, \mathcal{Y}^P) + \\
&\quad (1-\omega) \frac{\partial \ln \theta(\widetilde{\tau}_d(\mathcal{Y}^P, 0, e), \mathcal{Y}^P, u)}{\partial \tau_d} \psi(0, e, u, \mathcal{Y}^P)\Big] dF_{\Omega,E,U}(\omega, e, u) \\
&=\; \mu \mathbb{E}\left[\frac{\partial \ln \theta(\widetilde{\tau}_d(\mathcal{Y}^P, 1, e), \mathcal{Y}^P, u)}{\partial \tau_d} \psi(1, e, u, \mathcal{Y}^P)\right] + \\
&\quad (1-\mu) \mathbb{E}\left[\frac{\partial \ln \theta(\widetilde{\tau}_d(\mathcal{Y}^P, 0, e), \mathcal{Y}^P, u)}{\partial \tau_d} \psi(0, e, u, \mathcal{Y}^P)\right],
\end{aligned} \tag{23}$$

where $\mu \equiv \mathbb{E}(\Omega|\mathcal{Y}^P)$, i.e., the probability that a child is in case (L). This follows from the fact that Ω is stochastically independent of (U, E) given $\mathcal{Y}^P$ because, as argued in Sections 3 and 4.3, FAI thresholds are random draws from the viewpoint of applicants, conditioning on the number of applications. Therefore, there is no self-selection in cases (L) and (N) so that

[36] Superscripts r and l in $\mathcal{Y}^{P,r}$ and $\mathcal{Y}^{P,l}$ are chosen so to be consistent with the convention adopted in the RD figures above, where we assume that y_{-1} is ordered from higher values on the left to lower values on the right, so that admission to the Preferred program occurs to the *right* of the cutoff $\mathcal{Y}^P$. Note that $\tau_d(.)$ in Eq. 20 also depends on what is offered to the parent on the two sides of the cutoff, i.e., $z = 1$ on the right $(y_{-1} \to \mathcal{Y}^{P,r})$ and $z = \ell$ or no offer on the left $(y_{-1} \to \mathcal{Y}^{P,l})$. To simplify the notation we do not make this dependence explicit in $\tau_d(\cdot)$, although it is taken into account in the derivations that follow, as indicated by the notation $\tau_d^r(\cdot)$ and $\tau_d^l(\cdot)$ introduced below.

μ is an exogenous probability at a given $\mathcal{Y}^P$.

In words, our estimand at a specific Preferred FAI threshold can also be interpreted as an average, weighted by the probability that a child is in case (L) or (N), of the weighted averages of the causal effects of interest for children in each of these two cases. These averages are evaluated at the levels of daycare attendance $\widetilde{\tau}_d(\mathcal{Y}^P, 1, e)$ and $\widetilde{\tau}_d(\mathcal{Y}^P, 0, e)$ that are located between the values of attendance immediately to the right and immediately to the left of $\mathcal{Y}^P$ for each realization of e in cases (L) and (N).

Integrating over the different Preferred FAI thresholds that characterize our institutional setting, we obtain the overall average of the $\beta(\mathcal{Y}^P)$'s, weighted by the frequency of observations attached to each cutoff:

$$\beta = \int \beta(\mathcal{Y}^P) d\mathcal{F}(\mathcal{Y}^P), \tag{24}$$

where $\mathcal{F}$ is the distribution of Preferred FAI thresholds.[37] This solution to the aggregation problem is in the spirit of Cattaneo et al. (2016) and is also implemented by Card et al. (2015) where multiple cutoffs arise from pooling different years.[38]

The estimand β in Eq. 24 is not only relevant for parents but also for a policy maker interested in expanding vacancies in the existing facilities. As a consequence of this expansion, a larger number of households would have access to their preferred program. Our estimates speak precisely about the effect of such a policy, which may have undesirable consequence on the ability of more affluent children.[39]

[37] Eq. 24 assumes implicitly that the frequency distribution of the cutoffs over the population of households located just to the right of them is identical to the analogous distribution to the left. To support this assumption, in the Online Appendix to Section 6 we consider the empirical distribution functions of the preferred FAI thresholds in the right and left neighborhoods and we test their similarity using the Wilcoxon rank-sum test. The null hypothesis of similarity cannot be rejected, both in the Basket 4 universe (p-value: 0.21) and in the interview sample (p-value: 0.41). Figure A-14 in the Online Appendix plots the two pairs of CDF's.

[38] The application in Card et al. (2015) is the effect of unemployment benefits on unemployment duration in Austria. In their setting the running variable is annual earnings, and the earnings cutoff at which the benefit schedule has a kink varies by year. Moreover, different from us, their running variable measures with error the underlying assignment variable used to determine unemployment benefits. In our application, instead, the FAI used by the BDS to allocate daycare is exactly observed. As shown in the Online Appendix to Section 6, the absence of measurement error in our setting simplifies the econometric model with respect to Card et al. (2015).

[39] It is true that other weighting schemes, as for example those discussed by Bertanha (2017), would be informative on more general counterfactual scenarios in other institutional settings. An analysis in this spirit is presented in the Online Appendix to Section 3, where we simulate a calibrated version of the theoretical model under alternative weighting schemes.

6.2 Empirical model

Let θ_i be the skill trait of child i, observed at age 8–14, and let $\tau_{d,i}$ denote the treatment intensity, measured as months spent in daycare over the entire age 0–2 period.[40] The running variable is the FAI, y_i, at first application, and the equation we estimate is[41]

$$\ln \theta_i = \alpha + \beta \tau_{d,i} + m(y_i) + \gamma A_i + \delta X_i + \epsilon_i, \tag{25}$$

where β is an empirical counterpart of the theoretical estimand derived in Eq. 24, $m(y_i)$ is a second-order polynomial in the running variable, A_i is a vector of variables describing the application set of a child (specifically, the number of programs included in the application set and dummies for the city neighborhood of the preferred program), and X_i is a vector of pre-treatment variables (parents' education, parents' year of birth, number of siblings at the first application, whether parents were self-employed – as opposed to employees – during the year preceding the first application, birth day, and a dummy for cesarean delivery of the child).[42] Finally, ϵ_i captures unobserved determinants of ability.

As usual in RD designs, the inclusion of pre-treatment variables is not strictly necessary for identification but it may increase efficiency and, most important, similar estimates of β when observables are included or not support the validity of the identifying assumption that pre-treatment covariates are continuous at the thresholds (Imbens and Lemieux, 2008; Lee and Lemieux, 2010). More direct evidence on the validity of this assumption in the interview sample is provided in the third column of Panel A of Table 2, which shows that the Canay and Kamat (2018) test does not reject the null that the distributions of pre-treatment covariates are jointly continuous at the Preferred FAI threshold.[43]

We estimate equation (25) by IV using as an instrument a dummy P_i indicating whether

[40] In the administrative data we observe the precise daily attendance of children in daycare. For convenience, we rescale days of attendance in months defined as 20 working days.

[41] In this parametric specification we do not center and stack thresholds, different from what we do in the continuity test, thus avoiding the problems generated by observations located precisely at the thresholds. In the Online Appendix to Section 6 we also report non-parametric estimates that confirm the general pattern of the results described below, although those pertaining to non-cognitive outcomes are less precise.

[42] Descriptive statistics for these variables are reported in Tables A-7 and A-8 of the Online Appendix.

[43] Thanks to the information acquired from parents in the interviews, here we can assess continuity for a set of 11 covariates, which is larger than the one observed in the universe. In one case only, mother's year of birth, the p-value is smaller than 5%.

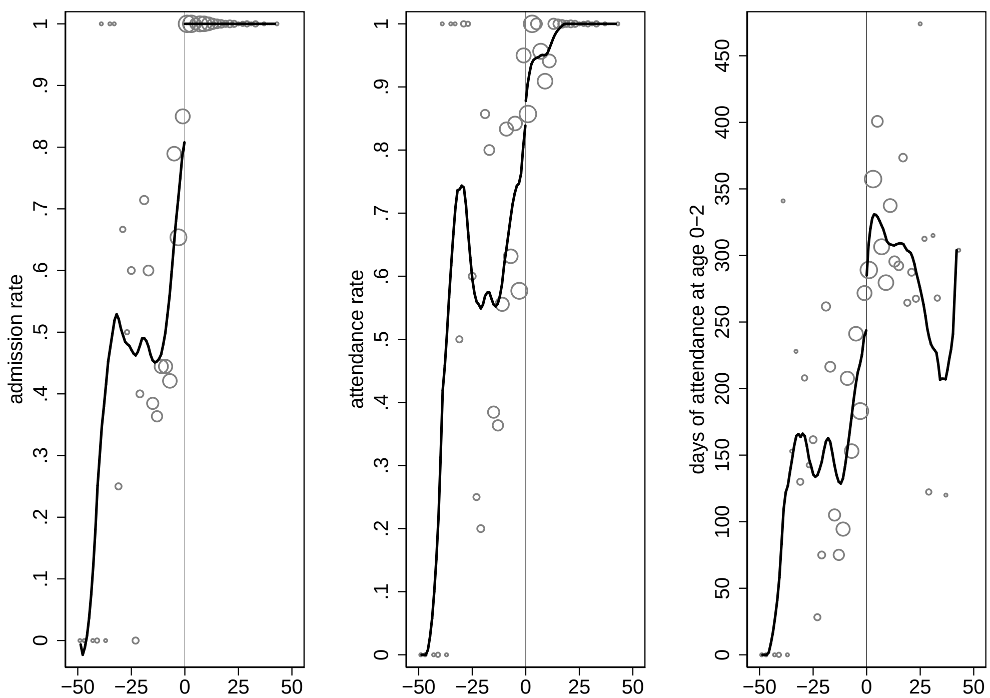
a child qualifies for her most preferred program at the first application or not,

$$P_i = \mathbb{I}(y_i \leq \mathcal{Y}_i^P), \tag{26}$$

where $\mathcal{Y}_i^P$ denotes the FAI threshold of child i's most preferred program.

Figure 3 replicates for the interview sample the evidence of Figure 2, which was based on the Basket 4 universe. Also in this sample the admission rate, the attendance rate and days of attendance all jump discontinuously at the preferred thresholds (by 19.2 percentage points, 3.9 percentage points and 41.3 days, respectively).

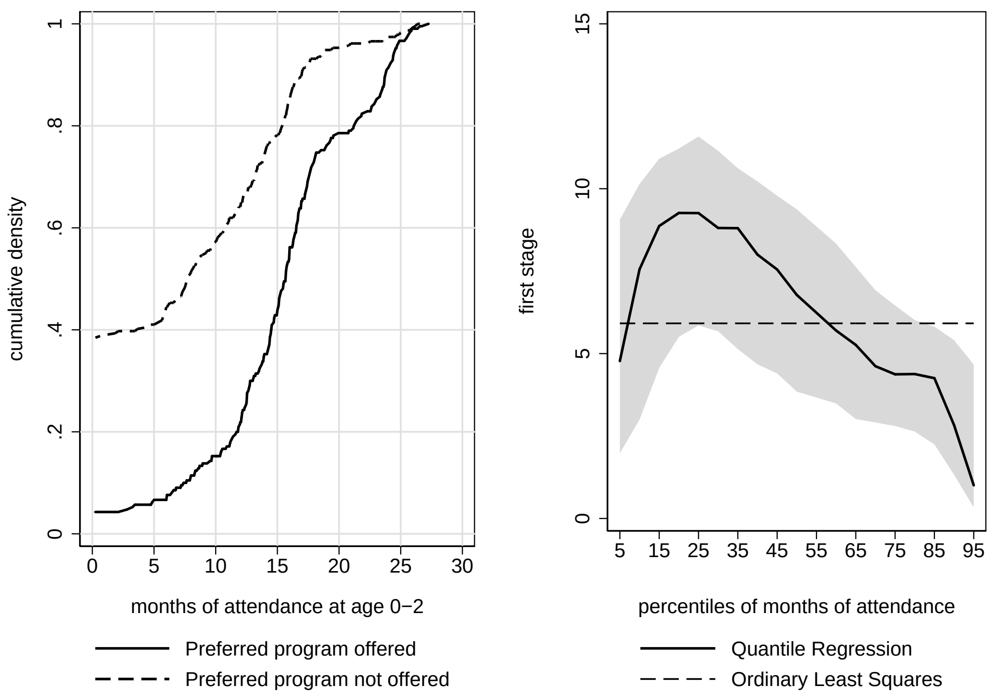
Figure 3: Admission offers and attendance around Pref. FAI thresholds, interview sample



Notes: The circles represent offer rates (left), attendance rates (middle) and average days of attendance at age 0–2 (right) inside €2k bins, plotted as a function of the distance (thousands of real €) of a child's FAI from her Preferred FAI threshold. The size of a circle is proportional to bin size. The bold lines are from separate LLR on the underlying individual observations on each side of the cutoff, with a triangular kernel and the optimal bandwidth selection developed in Calonico, Cattaneo, and Titiunik (2014), Calonico, Cattaneo, and Farrell (2018) and Calonico, Cattaneo, and Titiunik (forthcoming) and implemented in `STATA` by these same authors. FAI stands for Family Affluence Index. Sample: 373 interviewed children with two working parents, born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005, whose FAI distance from the Final FAI thresholds is at most €50k and is different from zero.

Monotonicity of the treatment in the instrument must also be satisfied for β to identify the estimand in Eq. 24. Remark 1 shows that we should expect monotonicity to hold in our setting: the offer of the most preferred program implies unambiguously that daycare attendance increases weakly for all parents and strictly so for at least some. This prediction is supported by the evidence of Figure 4. In the left panel, following Angrist and Imbens (1995), we plot the c.d.f. of days of attendance for the two groups of children defined by the instrument.

Figure 4: Monotonicity of the treatment in the instrument



Notes: The left panel shows the CDF of days of attendance in daycare 0–2 for the two groups of children defined by the instrument (whether the child qualifies for the preferred program). The right panel plots the coefficients from quantile regressions of total days of attendance in daycare 0–2 on the instrument and the same controls included in the estimation of Eq. 25. The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. The shaded areas represent the 95% percentile confidence intervals based on 1,000 block-bootstrap replications (so to preserve dependence within programs). Each coefficient is obtained by running a separate quantile regression for the 19 quantiles from 0.05 to 0.95. The dashed, horizontal line is the corresponding first-stage OLS estimates. Sample: 444 interviewed children with two working parents and non-missing IQ score and covariates, born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005.

Visual inspection indicates that the distribution of days of attendance for those who are offered their most preferred program (continuous line) first-order stochastically dominates the corresponding one for those who are not (dashed line). Under local “type independence” (Fiorini and Stevens, 2017), this is a necessary condition for monotononicity.[44] We use

[44]In our context, local “type independence” requires that the joint distribution of days of attendance in

the procedure developed by Barrett and Donald (2003) to test formally this ordering and we cannot reject the null (p-value: 0.9998; see Table A–10 of the Online Appendix for full details). The right panel of Figure 4 plots the estimates of the effect of being offered the most preferred program at different quantiles of months of attendance, based on our preferred specification with all the controls. These estimates are always positive and statistically significant, suggesting no violation of monotonicity also conditional on covariates.

6.3 Results: cognitive skills

The first row of the left panel in Table 6 ("All FAI thresholds") reports estimates of the effect of just qualifying for the most preferred program on IQ, which we refer to as the Intention To Treat (ITT) effect.[45] The first specification includes only the polynomial in the running variable, the second adds the application set characteristics, and the third one includes all controls, yielding similar estimates. Taking the third column as the preferred specification, the estimated ITT indicates that barely qualifying for the preferred program reduces IQ by 3% (p-value: 0.005). First stage estimates are reported in the second row of the table,[46] and indicate that just qualifying for the preferred program increases attendance by about six months. The F-test statistic on the excluded instrument indicates that weak instruments are not a concern.

Rescaling the ITT effect by the first stage gives the IV estimate of the effect of one additional month in daycare. In our preferred specification (and similarly in the others) this is a statistically significant loss of about 0.5% (p-value: 0.004) which, at the sample mean

case of qualification for the most preferred program or in case of no qualification (potential treatments) is independent of the running variable (FAI) locally at the cutoff. This assumption is not needed for the identification of the TT in Eq. 19 (see the Online Appendix to Section 6 and the discussion in Cattaneo, Frandsen, and Titiunik, 2015 and de la Cuesta and Imai, 2016), but we maintain it, as conventional in the literature, to test for monotonicity.

[45]Specifically, we estimate the following equation,

$$\ln \theta_i = \tilde{\alpha} + \tilde{\beta} P_i + \tilde{m}(y_i) + \tilde{\gamma} A_i + \tilde{\delta} X_i + \tilde{\epsilon}_i,$$

where $\tilde{m}(y_i)$ is a second-order polynomial in the FAI and $\tilde{\beta}$ is the ITT effect.

[46]In this case, we estimate the first stage equation,

$$\tau_{d,i} = \bar{\alpha} + \bar{\beta} P_i + \bar{m}(y_i) + \bar{\gamma} A_i + \bar{\delta} X_i + \bar{\epsilon}_i,$$

where $\bar{m}(y_i)$ is a second order polynomial in the FAI and $\bar{\beta}$ is the first stage estimate.

Table 6: Effects of daycare 0–2 attendance on log IQ, for all children and by level of the Preferred FAI threshold

	All FAI thresholds (mean threshold: €24.6k)			FAI thresholds $\leq$ median (mean threshold: €16.4k)			FAI thresholds $>$ median (mean threshold: €33.0k)		
ITT effect of qualifying for the preferred program	-0.026*	-0.030**	-0.030**	-0.010	-0.018	-0.018	-0.044*	-0.045*	-0.048**
	(0.010)	(0.011)	(0.010)	(0.016)	(0.015)	(0.014)	(0.017)	(0.018)	(0.017)
First stage: effect of qualifying on months of attendance	6.3**	6.4**	5.9**	5.6**	5.6**	5.1**	5.6**	5.8**	5.5**
	(0.9)	(0.9)	(0.9)	(1.4)	(1.5)	(1.5)	(1.3)	(1.3)	(1.2)
IV effect of one month of daycare attendance	-0.004**	-0.005**	-0.005**	-0.002	-0.003	-0.003	-0.008*	-0.008*	-0.009**
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
F-stat on excluded instruments	49.1	46.9	44.8	15.2	14.3	11.6	19.6	19.0	19.3
Number of observations	444	444	444	224	224	224	220	220	220
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pre-treatment controls			Yes			Yes			Yes

Notes: The table reports parametric estimates of the effect of one month of daycare 0–2 on log IQ, and the associated ITT and first stage, at all levels of the Preferred FAI threshold and separately for Preferred FAI thresholds below or above the median Preferred FAI threshold. ITT coefficients are from regressions of log IQ on the instrument (whether the child qualifies for the preferred program) and controls (see footnote 46). First-stage coefficients are from regressions of months (1 month = 20 days of attendance) spent in daycare 0–2 on the instrument and controls (see footnote 45). IV coefficients are from regressions of log IQ on months of attendance and controls using a dummy for qualification in the preferred program as the instrument (see Eq. 25). The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. Sample: 444 interviewed children with two working parents, born between 1999 and 2005, with non-missing outcome or covariates and whose parents first applied for admission to daycare between 2001 and 2005. Robust standard errors in parentheses, clustered at the facility level. * significant at 5%; ** significant at 1% or better.

(116.4), corresponds to 0.6 IQ points and to 4.7% of the IQ standard deviation. As argued in Section 5, the interview sample is characterized by relatively affluent and educated parents in one of the richest Italian cities. Therefore, in light of Remark 2, it should not come as a surprise that the IQ effect of daycare turns out to be negative in this population.

To further illustrate the empirical relevance of Remark 2, we split children in two groups according to whether the Preferred FAI threshold they are associated with is above or below the median of all Preferred thresholds. Results are reported in the middle and right panels of Table 6.[47] For the less affluent group, the estimates refer to the effect of daycare 0–2 around a Preferred FAI threshold of €16.4k on average (corresponding to a gross annual family income of about €43k), while in the more affluent group above the median the average threshold is €33.0k (annual family income of about €88k). Parametric estimates of the IQ response to more daycare in these two groups are reported in the middle and right panels of Table 6. In the less affluent group, the ITT effect of just qualifying for the most preferred program is estimated to be negative but relatively small (less than 2%) and statistically insignificant. In the more affluent group the estimated loss is nearly three times larger (almost 5%) and precisely estimated. The second row in these same panels displays the first stage effect, which is similar for more and less affluent households in our sample. Rescaling the ITT effect by the first stage gives a statistically significant IV estimate for the IQ loss of between 0.8% and 0.9% in the more affluent group, while for less affluent households the estimate is an insignificant 0.2%–0.3%. A similar pattern emerges from the non-parametric estimates for the two groups reported in the Online Appendix to Section 6, where we also summarize the analysis for the four sub-scales that compose the total IQ score considered here. With different degrees of intensity, these results hold similarly for the sub-scales.

As indicated by Eq. 23, the IV estimates reported in Table 6 are averages of the effects of one additional month in daycare for children in case (N): $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1 - \Omega$ (32% in the interview sample) and for the remaining children, who are in case (L): $\mathbb{I}(\mathcal{Y}^P \neq \mathcal{Y}^M) = \Omega$. Table 7 reports results from a specification that allows the IV effects to be different in the two cases. Starting with column 1, which is based on the entire interview sample, the IV

[47] The Online Appendix reproduces the figure and tables of the main text related to the validity of our identification strategy separately for the two affluence groups.

estimate of the IQ loss induced by one additional month in daycare attendance is 0.7% for children in case (L) with no significant difference detected for the remaining children in case (N).[48] When we split the interview sample by household affluence as in Table 6, for the less advantaged children the effect is again small in both cases (L) and (N), while for the more advantaged the IQ loss is confirmed to be large in magnitude (about 1.4% and statistically significant), again with no difference between cases (L) or (N).

Table 7: IV effects of daycare 0–2 attendance on IQ, by cases (L) and (N)

Mean FAI threshold:	All thresholds €24.6k	$\leq$ median €16.4k	$>$ median €33.0k
Daycare attendance	-0.007**	-0.003	-0.014**
	(0.003)	(0.004)	(0.005)
Daycare attendance$\times\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$	0.004	-0.000	0.009
	(0.003)	(0.005)	(0.006)
$\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) \equiv (1-\Omega)$	-0.028	0.019	-0.098
	(0.040)	(0.049)	(0.077)
Number of observations	444	224	220

Notes: The table reports parametric estimates of the effect of one additional month of daycare 0–2 on log IQ, at all levels of the Preferred FAI thresholds and separately for Preferred FAI thresholds below or above their median. Coefficients are from regressions of log IQ on months of attendance, months of attendance interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$ and the full set of controls A_i and X_i using the dummy P_i for qualification in the preferred program and the same dummy interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1-\Omega$ as the instruments (see footnote 48). The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. Sample: 444 interviewed children with two working parents, born between 1999 and 2005, with non-missing information, whose parents first applied for admission to daycare between 2001 and 2005. Robust standard errors in parentheses, clustered at the facility level. * significant at 5%; ** significant at 1% or better.

[48] Specifically, we estimate by IV

$$\ln\theta_i = \alpha_L + [\alpha_N - \alpha_L](1-\Omega_i) + \beta_L\tau_{d,i} + [\beta_N - \beta_L]\tau_{d,i}(1-\Omega_i) + m(y_i) + \gamma A_i + \delta X_i + \epsilon_i,$$

where β_L is the estimate for case (L) and $[\beta_N - \beta_L]$ measures how much the estimate for case (N) differs from that of case (L). The instruments are P_i and $P_i(1-\Omega_i)$. Like in the Basket 4 universe (see footnote 32) and as expected given the discussion in Section 3, children in cases (L) and (N) attend daycare for approximately the same number of months (16 and 15, respectively), if offered their preferred program. If not offered their preferred program, instead, children in case (L) attend for about 9 months in a less preferred program, while children in case (N) attend for about 5 months, which is an average between those who re-apply and attend in a later year and those who never attend any BDS program. The respective first stages for Daycare attendance at the Preferred threshold are 3.93 (s.e. 0.82) in case (L) and 9.91 (s.e. 0.62) in case (N). Additional descriptive statistics for the interview sample children in the two cases are reported in Table A-21 of the Online Appendix.

6.4 Results: non-cognitive skills

The corresponding results in the personality domain are reported in Tables 8 and 9, where the Big Five personality traits are the outcome variables and where estimates are reported for all children and by level of the Preferred FAI threshold. First-stage estimates are essentially the same as those reported in Table 6 and so are not repeated here.[49]

Referring to the preferred specification in the third column of the third panel, Table 8 shows that qualifying for the preferred daycare program at age 0–2 reduces openness at age 8–14 by 8%, agreeableness by 6.8%, and increases neuroticism by 5.1% in the more affluent group. No significant effects are detected in this group for either conscientiousness or extraversion, although the point estimates are similary negative. In the less affluent group, instead, no effects on personality emerge: point estimates are generally closer to zero and often positive (negative for neuroticism). These magnitudes and patterns are in agreement with the analogous ITT effect estimated for IQ in the more affluent group (a significant -4.8%) and in the less affluent one (an insignificant -1.8%).

The finding that for more affluent children sufficient precision is attained only for three of the Big Five personality traits may be due to the fact that in a relatively small sample the measurement of IQ, being task-based, may be more precise than the measurement of personality, which is based on a questionnaire. An alternative explanation is suggested by the sample correlation between IQ and, respectively, openness (0.294), conscientiousness (0.001), extraversion (0.005), agreeableness (0.021), and neuroticism (-0.042). The personality traits for which we detect effects of additional daycare attendance are only those that are more correlated with IQ.[50]

The ITT results translate into the estimates of the effect of an additional month in daycare presented in Table 9. For the more affluent, this treatment decreases openness and agreeableness by 1.4% and 1.2% and increases neuroticism by 0.9%, with no effect for the less affluent and for the other traits. A comparison with the IQ effect for the more affluent

[49] They are not numerically identical because for these outcomes we have 447 (446 for Agreeeableness) children instead of 444 (see Section 5 and, specifically, footnote 33). All the descriptive statistics are essentially unchanged for this slightly larger sample.

[50] With specific reference to conscientiousness, this pattern is in line with a remark in Elango et al. (2016): this personality trait is "a non-cognitive skill that is of interest due to its low correlation with cognition and high correlation with important later-life outcomes." (p. 254)

Table 8: ITT effect of qualifying for the preferred program on personality, for all children and by level of the Preferred threshold

	All FAI thresholds (mean threshold: €24.7k)			FAI thresholds $\leq$ median (mean threshold: €16.4k)			FAI thresholds $>$ median (mean threshold: €33.0k)		
Openness	-0.032	-0.029	-0.027	-0.002	-0.005	0.004	-0.076*	-0.069*	-0.080*
	(0.021)	(0.021)	(0.021)	(0.040)	(0.040)	(0.038)	(0.029)	(0.029)	(0.031)
Conscientiousness	0.001	0.006	0.003	0.033	0.039	0.038	-0.009	-0.007	-0.006
	(0.023)	(0.024)	(0.023)	(0.035)	(0.032)	(0.033)	(0.037)	(0.039)	(0.038)
Extraversion	-0.032	-0.032	-0.039^{+}	-0.052	-0.054	-0.055	-0.030	-0.030	-0.037
	(0.023)	(0.023)	(0.022)	(0.040)	(0.039)	(0.037)	(0.027)	(0.030)	(0.032)
Agreeableness	-0.034	-0.030	-0.024	0.001	0.007	0.017	-0.060*	-0.068*	-0.068*
	(0.021)	(0.022)	(0.020)	(0.033)	(0.033)	(0.033)	(0.026)	(0.033)	(0.032)
Neuroticism	0.018	0.014	0.013	-0.011	-0.013	-0.026	0.045	0.046^{+}	0.051^{+}
	(0.019)	(0.019)	(0.020)	(0.031)	(0.033)	(0.033)	(0.027)	(0.027)	(0.027)
Number of observations	447	447	447	225	225	225	222	222	222
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pre-treatment controls			Yes			Yes			Yes

Notes: The table reports parametric estimates of the effect of qualifying for the most preferred daycare program on the log of scores in the Big Five Questionnaire for Children, at all levels of the Preferred FAI threshold and separately for Preferred FAI thresholds below or above the median Preferred FAI threshold. The coefficients are from distinct regressions of each outcome on a dummy for qualification in the preferred program and controls. The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. Sample: 447 (446 for Agreeeableness) interviewed children with two working parents , born between 1999 and 2005, with non-missing outcome or covariates and whose parents first applied for admission to daycare between 2001 and 2005. Robust standard errors in parentheses, clustered at the facility level. $^{+}$ significant at 10%; * significant at 5%; ** significant at 1% or better.

Table 9: Effects of daycare 0–2 attendance on personality, for all children and by level of the Preferred FAI threshold

	All FAI thresholds (mean threshold: €24.7k)			FAI thresholds ≤ median (mean threshold: €16.4k)			FAI thresholds > median (mean threshold: €33.0k)		
Openness	-0.005	-0.005	-0.004	-0.000	-0.001	0.001	-0.013*	-0.011*	-0.014**
	(0.003)	(0.003)	(0.003)	(0.007)	(0.007)	(0.007)	(0.006)	(0.005)	(0.005)
Conscientiousness	0.000	0.001	0.000	0.006	0.007	0.007	-0.002	-0.001	-0.001
	(0.004)	(0.004)	(0.004)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Extraversion	-0.005	-0.005	-0.006^{+}	-0.009	-0.009	-0.011	-0.005	-0.005	-0.006
	(0.004)	(0.003)	(0.004)	(0.007)	(0.007)	(0.007)	(0.005)	(0.005)	(0.005)
Agreeableness	-0.005	-0.005	-0.004	0.000	0.001	0.003	-0.010*	-0.011*	-0.012*
	(0.003)	(0.003)	(0.003)	(0.006)	(0.005)	(0.006)	(0.005)	(0.006)	(0.006)
Neuroticism	0.003	0.002	0.002	-0.002	-0.002	-0.005	0.008	0.007	0.009^{+}
	(0.003)	(0.003)	(0.003)	(0.005)	(0.006)	(0.006)	(0.005)	(0.005)	(0.005)
Number of observations	447	447	447	225	225	225	222	222	222
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes		Yes	Yes		Yes	Yes
Pre-treatment controls			Yes			Yes			Yes

Notes: The table reports parametric estimates of the effect of one month (20 days of attendance) of daycare 0–2 on the log of scores in the Big Five Questionnaire for Children, at all levels of the Preferred FAI threshold and separately for Preferred FAI thresholds below or above the median Preferred FAI threshold. The coefficients are from distinct regressions of each outcome on months of attendance and controls using a dummy for qualification in the preferred program as the instrument. The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. Sample: 447 (446 for Agreeeableness) interviewed children with two working parents, born between 1999 and 2005, with non-missing outcome or covariates and whose parents first applied for admission to daycare between 2001 and 2005. Robust standard errors in parentheses, clustered at the facility level. $^{+}$ significant at 10%; * significant at 5%; ** significant at 1% or better.

(-0.9%) suggests that cognitive and non-cognitive skills respond similarly at age 8–14 to a substitution of informal care with formal care in households characterized by different socioeconomic status and therefore different quality of informal care. The similarities in the effects for the cognitive and non-cognitive domains is not surprising given that these skills are similarly sensitive to early influences in life, as discussed in Section 7.

Table 10 replicates for the Big Five personality traits the IV analysis by cases (L) and (N) reported in Table 7 for IQ, with qualitatively similar results. Specifically, we observe no statistically significant difference between the (L) and (N) cases.

Table 10: Effects of daycare 0–2 attendance on personality, by cases (L) and (N)

	All thresholds		≤ median		> median	
	β_L	$\beta_N - \beta_L$	β_L	$\beta_N - \beta_L$	β_L	$\beta_N - \beta_L$
Openness	-0.004	0.000	0.003	-0.010	-0.019*	0.009
	(0.006)	(0.007)	(0.009)	(0.011)	(0.009)	(0.010)
Conscientiousness	-0.001	0.002	0.011	-0.004	-0.007	0.012
	(0.006)	(0.006)	(0.008)	(0.012)	(0.010)	(0.010)
Extraversion	-0.007	-0.000	-0.014	0.010	-0.003	-0.007
	(0.007)	(0.007)	(0.010)	(0.012)	(0.009)	(0.011)
Agreeableness	-0.006	0.003	0.004	0.007	-0.015	0.006
	(0.006)	(0.008)	(0.008)	(0.008)	(0.009)	(0.013)
Neuroticism	0.002	-0.001	-0.009	0.010	0.016^{+}	-0.012
	(0.006)	(0.007)	(0.010)	(0.009)	(0.009)	(0.009)
N	447	447	225	225	222	222

Notes: The table reports parametric IV estimates of the effect of one month of daycare 0–2 on the log scores in the Big Five Questionnaire for Children, at all levels of the Preferred FAI threshold and separately for Preferred FAI thresholds below or above the median Preferred FAI threshold. Coefficients are from regressions of each outcome on months of attendance, months of attendance interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M)$ and the full set of controls A_i and X_i, using the dummy P_i for qualification in the preferred program and the same dummy interacted with $\mathbb{I}(\mathcal{Y}^P = \mathcal{Y}^M) = 1 - \Omega$ as the instruments (see footnote 48). The running variable is the Family Affluence Index (FAI), and the polynomial in the running variable is of second order. Sample: 447 (446 for Agreeeableness) interviewed children with two working parents, born between 1999 and 2005, with non-missing outcome or covariates and whose parents first applied for admission to daycare between 2001 and 2005. Robust standard errors in parentheses, clustered at the facility level. $^{+}$ significant at 10%; * significant at 5%; ** significant at 1% or better.

7 Suggestions from the psychological literature

Psychologists have produced persuasive empirical evidence that during the first three years of life one-to-one interactions with adults (more than interactions with peers) are a crucial input for both the cognitive and non-cognitive development of a child. For instance, in an empirical field study of 42 American families, Hart and Risley (1995) have recorded one full hour of words spoken at home every month for 2.5 years by parents with their children at age 0–2. They conclude that "the size of the children's recorded vocabularies and their IQ scores were strongly associated with the size of their parents' recorded vocabulary and their parents' scores on a vocabulary pre-test" (p. 176). Along the same lines, Rowe and Goldin-Meadow (2009) and Cartmill et al. (2013) show that the quality of parental inputs in the first three years of life (e.g. in terms of parental gesture and talking) improves children's vocabulary before school entry. Similarly, Gunderson et al. (2013) finds that parental praise directed to 1- to 3-years-old children predicts their motivation five years later.[51]

A fascinating theory explaining why early one-to-one interactions with adults are so important has been proposed by Csibra and Gergely (2009, 2011). According to these authors, the communication between a trusted adult and a child allows the latter to understand more rapidly if an experience has a general value or only a specific one. Lacking such communication, the child must repeat an experience many times in order to assess its general or particular validity (very much like statistical inference requiring a large sample). An adult, instead, can quickly inform the child about the nature of what he or she is experimenting. If the adult can be trusted, then the child can save time and move on to other experiences, thus gaining a developmental advantage.

The focus on one-to-one interactions in our context is relevant because, as noted by Clarke-Stewart, Gruber, and Fitzgerald (1994), infants and toddlers generally experience less one-to-one attention in daycare than at home because at home they are typically taken care of by a parent, a grandparent, or a nanny. Under these care modes a child receives

[51] Related to these results, some psychologists have estimated negative effects of increasing parental (in particular maternal) working time on cognitive and non-cognitive outcomes of children. See, for example, Brooks-Gunn, Han, and Waldfogel (2002), Adi-Japha and Klein (2009), McPherran Lombardi and Levine Coley (2014) and the meta-analysis in Li et al. (2013). Different from the economic literature, however, most of these studies are observational and do not exploit quasi-experimental identification strategies.

attention in a 1:1 ratio, possibly somewhat less if, for example, siblings are present. This is precisely the case for the children in our sample. When we asked their parents which options were available at the time of the first application as an alternative to daycare during the workday, 50.5% checked "the mother", 11% "the father", 44.8% "the grandparents", 4.5% "other family members", 18.9% "a babysitter or a nanny", and only 12.1% checked "some other daycare center" (multiple answers were possible).[52] The adult-to-child ratio in daycare 0–2 depends instead on the specific institutional setting. At the BDS, during the period under investigation, this ratio was 1:4 at age 0 and 1:6 at age 1–2. This may be part of the reason why both Felfe and Lalive (2018) and Drange and Havnes (2018) find positive effects of daycare 0–2 in Germany and Norway. In their institutional setting, the adult-to-child ratio is about 1:3.

A related hypothesis emphasized by Belsky and Steinberg (1978) and Belsky (1988, 2001) is that the negative effects of daycare are driven by decreased interactions with the mother, leading to a drop in maternal involvement, children's attachment to their mothers, and consequent child insecurity. This maternal channel is at center stage in Bernal (2008). We cannot tell whether the negative effects of daycare that we uncover are driven by a substitution away from maternal care or from family-based care more generally. However, the aforementioned statistics indicate that in our sample the counterfactual care mode for the fraction of time children would not have spent in daycare is not just mothers.

Psychologists have also supported with persuasive empirical evidence the idea that girls are more "mature" than boys, in the sense of being more capable of absorbing cognitive stimuli at an early age. For example, Fenson et al. (1994) study 1,800 toddlers (16-30 months of age) finding that girls perform better in lexical, gestural, and grammatical development.

[52] In Bologna there are very few private daycare facilities outside the public system. The reason is that Bologna is one of the Italian cities with the largest and most highly-reputed public daycare systems, which leaves little room for independent private providers, relative to other cities. The BDS includes 9 charter facilities that are privately managed but strictly regulated by the BDS. According to the reputational indicator of quality described in Section 4.1, these charter programs are perceived by parents as worse than the non-charter ones. On average, for given distance, charter programs are ranked 1.6 positions lower than the non-charter ones if they are included in the application set of a parent. Moreover, the probability that a charter program is not ranked by parents in their application set is higher than for non-charter ones (0.95 vs 0.91). The odds that a charter program is ranked first by a parent is 0.007 while the odds that a program is charter is 0.046. Therefore, it is unlikely that the worse quality of these charter programs is responsible for the negative effects that we estimate – which derive from the offer of the most preferred program to a parent. If anything, their presence should reduce the absolute size of our estimates.

Galsworthy et al. (2000) examine about 3,000 2-year-old twin pairs and show that girls score higher on verbal and non-verbal tests. A longitudinal study by Bornstein, Hahn, and Haynes (2004) involving 329 children observed between age 2 and 5 reaches similar conclusions for an age range partially overlapping with ours: they show that "girls consistently outperformed boys in multiple specific and general measures of language" (p. 206).

If girls at age 0–2 are relatively more capable of making good use of stimuli that improve their skills, then their development is hurt (more than for boys) by an extended exposure to daycare because it implies fewer one-to-one interactions with adults which are more valuable for them than for boys as inputs in the technology of skill formation. In the Online Appendix to Section 3 we extend our theoretical framework to model how parental decisions concerning child care are compatible with these gender differences in the effects of daycare 0–2. A replication of our analysis by gender for both IQ and the Big Five personality traits is also reported in the Online Appendix to Section 7 and can be summarized as follows: for girls, the ITT effects of just qualifying for the most preferred program consist of a reduction in IQ by 3.9% (p-value: 0.017) and extraversion by 7.2% (p-value: 0.036), with a decrease in openness of 10.5% (p-value: 0.053) that emerges for more affluent girls. For boys, the corresponding coefficients are smaller and we cannot reject the hypothesis that they are equal to zero.[53] A similar gender difference emerges also from the IV estimates, which indicate that for girls one additional month in daycare 0–2 reduces IQ by 0.7% (p-value: 0.016) and, when restricting to more affluent girls, openness by 1.8% (p-value: 0.067) and extraversion by 2% (p-value: 0.097), with smaller and insignificant effects for boys.[54] This gender heterogeneity in the cognitive and non-cognitive losses induced by daycare attendance supports the relevance of one-to-one interactions with adults as an explanation of our results.

[53]This gender gap is not due to differences in the first stage. Similarly, we can easily dismiss the possibility that this gender heterogeneity in the effects of daycare 0–2 reflects differences in pre-treatment characteristics of boys and girls in our sample. The Online Appendix reproduces the figure and tables of the main text related to the validity of our identification strategy separately for the two gender groups.

[54]Interestingly, in a longitudinal study of 113 first-born preschool children, 58 girls and 55 boys, Bornstein et al. (2006) find, in line with our results, that "Girls who had greater amount of non-maternal care from birth to 1 year scored lower on the Spoken Language Quotient at preschool" (pag. 145).

8 Alternative mechanisms

We have also explored some alternative interpretations of our results. A first possibility is that daycare 0–2 may impact negatively on IQ because it increases children's exposure to infections (Baker, Gruber, and Milligan, 2008), which have been shown to harm human capital accumulation and cognitive development (see Eppig, Fincher, and Thornill, 2010, John, Black, and A., 2017 and the review in Bleakley, 2010). However, since boys are more vulnerable than girls to infection exposure at a young age (Muenchhoff and Goulder, 2014), this explanation is at odds with the gender difference in the effects of daycare 0–2 that we uncover.

Second, in line with the early results reported in Belsky and Steinberg (1978) and additional findings in Baker, Gruber, and Milligan (2008), it is also possible that daycare 0–2 induces a disengagement of parents from the education of children, which may impact negatively on their development. We do not have direct evidence to dismiss this interpretation but, again in light of the gender heterogeneity of our results, it is not clear why parental disengagement should be more pronounced for girls than for boys.

A third alternative mechanism that might specifically explain the gender difference that we estimate in the effects of daycare 0–2 refers to the possibility that the loss suffered by girls depends on the sex ratio within each program. Psychologists have observed that in early education "(T)eachers spend more time socializing boys into classroom life, and the result is that girls get less teacher attention. Boys receive what they need ... Girls' needs are more subtle and tend to be overlooked." (Koch, 2003, p. 265). However, we do not find any evidence that sex ratios affect the size of the effects for girls and boys, perhaps because the variation in these ratios is quite small for the children in our sample.

Finally, the data do not support the hypothesis that gender differences in breastfeeding explain the gender gap in the effects of daycare. The duration of breastfeeding has been shown to be positively associated with cognitive outcomes (Anderson, Johnstone, and Remley, 1999; Borra, Iacovou, and Sevilla, 2012; Fitzsimons and Vera-Hernandez, 2013), and early daycare attendance may shorten it. However, we find no effect (and specifically no differential effect by gender) of days in daycare on the duration of breastfeeding.

9 Conclusions

This paper contributes to a growing literature studying the effects of time spent in daycare at age 0–2 on child ability. We studied the offspring of dual-earner households with cohabiting parents in Bologna, one of the most educated and richest Italian cities with a highly reputed public daycare system. For the children in this population, our results indicate a quantitatively and statistically significant loss in IQ and in some non-cognitive traits at age 8–14. This loss is even more pronounced when we focus on children with relatively more affluent parents within this population. These are typically the relevant marginal subjects to be considered in an evaluation of daycare expansions as a response to the growing incidence of dual-earner households in advanced countries.

We interpret these findings in a theoretical model showing that when offered their most preferred daycare program (as opposed to a less preferred one or no offer), relatively affluent parents take advantage of this opportunity to increase labor supply and child's daycare attendance. Due to the higher earning potential of more affluent parents, this increase in attendance generates an increase of family resources that is large enough to become attractive even if it comes at the cost of child ability.

These results seem relevant not only because of their novelty with respect to the literature, but also because they implicitly support the hypothesis, suggested by psychologists, that the sign and size of the effects of daycare 0–2 are mostly driven by three factors. First, whether this early life experience deprives children of one-to-one interactions with adults at home. Second, by the quality of these interactions, which is likely to be higher in more affluent households. And, third, by whether children can make good use of them. The latter claim is supported by the finding that daycare 0–2 has a more negative effect on the ability of girls, in combination with the psychological evidence suggesting that girls are developed enough at this young age to exploit higher quality interactions with adults that for boys are less valuable.

Our identification strategy exploits affluence thresholds that discriminate between similar parents whose children attend daycare 0–2 for longer versus shorter periods because they are barely admitted to their preferred program instead of being just excluded from it. This strategy makes our results valuable not only for parents but also for policy makers interested in expanding vacancies in the daycare systems that they manage. The estimates we presented speak precisely about the effect of such a policy, which would allow more affluent children to attend for a longer time in programs that their parents prefer more, with negative effects on their skills that may not be socially optimal even if the utility of parents increase.

References

Adi-Japha, Esther and Pnina S. Klein. 2009. "Relations Between Parenting Quality and Cognitive Performance of Children Experiencing Varying Amounts of Childcare." *Child Development* 80 (3):893–906.

Almond, Douglas and Janet Currie. 2011. "Human Capital Development Before Age Five." In *Handbook of Labor Economics*, vol. 4 Part B, edited by Stephen Machin Eric A. Hanushek and Ludger Woessmann. Elsevier, 1315–1486.

Anderson, James W, Bryan M Johnstone, and Daniel T Remley. 1999. "Breast-feeding and cognitive development: a meta-analysis." *The American Journal of Clinical Nutrition* 70 (4):525–535.

Anderson, Michael 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.

Angrist, Joshua and Guido Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430):431–442.

Baker, Michael, Jonathan Gruber, and Kevin Milligan. 2008. "Universal Child Care, Maternal Labor Supply, and Family Well-Being." *Journal of Political Economy* 116 (4):709–745.

———. 2015. "Non-Cognitive Deficits and Young Adult Outcomes: The Long-Run Impacts of a Universal Child Care Program." Working Paper 21571, National Bureau of Economic Research.

Barrett, G. F. and S. G. Donald. 2003. "Consistent Tests For Stochastic Dominance." *Econometrica* 71 (1):71–104.

Becker, Gary S. 1965. "A Theory of the Allocation of Time." *The Economic Journal* 75 (299):493–517.

Belsky, Jay. 1988. "The Effects of infant day care reconsidered." *Early Childhood Research Quarterly* 3 (3):235 – 272.

———. 2001. "Emanuel Miller Lecture: Developmental risks (still) associated with early child care." *Journal of Child Psychology and Psychiatry* 42 (7):845–859.

Belsky, Jay and Laurence D. Steinberg. 1978. "The Effects of Day Care: A Critical Review." *Child Development* 49 (4):929–949.

Bernal, Raquel. 2008. "The effect of maternal employment and child care on children's cognitive development." *International Economic Review* 49 (4):1173–1209.

Bernal, Raquel and Michael P. Keane. 2011. "Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers." *Journal of Labor Economics* 29 (3):pp. 459–512.

Bertanha, M. 2017. “Regression Discontinuity Design with Many Thresholds.” Available at SSRN: https://ssrn.com/abstract=2712957 or http://dx.doi.org/10.2139/ssrn.2712957, University of Notre Dame, Department of Economics.

Bleakley, H. 2010. “Health, human capital, and development.” *Annual Review of Economics* 2:283–310.

Borghans, Lex, Angela Lee Duckworth, James J. Heckman, and Bas ter Weel. 2008. “The Economics and Psychology of Personality Traits.” *Journal of Human Resources* 43 (4):972–1059.

Bornstein, Marc, Chun-Shin Hahn, Nancy Gist, and Maurice Haynes. 2006. “Long term cumulative effects of childcare on children’s mental development and socioemotional adjustment in a non-risk sample: the moderating effects of gender.” *Early Child Development and Care* 176 (2):129–156.

Bornstein, Marc H., Chun-Shin Hahn, and O. Maurice Haynes. 2004. “Specific and general language performance across early childhood: Stability and gender considerations.” *First Language* 24 (3):267–304.

Borra, Cristina, Maria Iacovou, and Almudena Sevilla. 2012. “The effect of breastfeeding on children’s cognitive and noncognitive development .” *Labour Economics* 19 (4):496 – 515.

Brooks-Gunn, Jeanne, Wen-Jui Han, and Jane Waldfogel. 2002. “Maternal Employment and Child Cognitive Outcomes in the First Three Years of Life: The NICHD Study of Early Child Care.” *Child Development* 73 (4):1052?1072.

Calonico, Sebastian, Matias D. Cattaneo, and M.H. Farrell. 2018. “On the Effect of Bias Estimation on Coverage Accuracy in Nonparametric Inference.” *Journal of the American Statistical Association* 113 (522):767–779.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs.” *Econometrica* 82 (6):2295–2326.

———. forthcoming. “Regression Discontinuity Design Using Covariates.” *Review of Economics and Staistics* .

Campbell, Frances A. and Craig T. Ramey. 1994. “Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families.” *Child Development* 65 (2):684–698.

Canay, Ivan A and Vishal Kamat. 2018. “Approximate permutation tests and induced order statistics in the regression discontinuity design.” *The Review of Economic Studies* .

Card, David, David S. Lee, Zhuan Pei, and Andrea Weber. 2015. “Inference on Causal Effects in a Generalized Regression Kink Design.” *Econometrica* 83 (6):2453–2483.

Carneiro, Pedro, Flavio Cunha, and James Heckman. 2003. "Interpreting The Evidence of Family Influence on Child Development." Mimeo.

Carneiro, Pedro and Rita Ginja. 2014. "Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start." *American Economic Journal: Economic Policy* 6 (4):135–73.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes. 2015. "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children." *Journal of Political Economy* 123 (2):pp. 365–412.

Cartmill, Erica, Benjamin Armstrong, Lila Gleitman, Susan Goldin-Meadow, Tamara Medina, and John Trueswell. 2013. "Quality of early parent input predicts child vocabulary 3 years later." *PNAS* 110 (28):pp. 11278–11283.

Cascio, Elizabeth U and Diane Whitmore Schanzenbach. 2013. "The Impacts of Expanding Access to High-Quality Preschool Education." *Brookings Papers on Economic Activity, Economic Studies Program, The Brookings Institution* 47 (2):pp. 127–192.

Cattaneo, Matias D., Brigham R. Frandsen, and Roco Titiunik. 2015. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." *Journal of Causal Inference* 3 (1):1–24.

Cattaneo, Matias D., Luke Keele, Roco Titiunik, and Gonzalo Vazquez-Bare. 2016. "Interpreting Regression Discontinuity Designs with Multiple Cutoffs." *The Journal of Politics* 78 (4):1229–1248.

Clarke-Stewart, K., C. Gruber, and L. Fitzgerald. 1994. *Children at home and in day care* . Hillsdale, NJ: L. Erlbaum Associates.

Comune di Bologna. 2010. "Progetto pedagogico del nido d'infanzia del Comune di Bologna." Report: Comune di Bologna Settore Istruzione, Comune di Bologna Settore Istruzione.

Csibra, Gergely and Gyorgy Gergely. 2009. "Natural Pedagogy." *Philosophical Transactions of the Royal Society B* 13 (4):148–153.

———. 2011. "Natural Pedagogy as evolutionary adaptation." *Trends in Cognitive Sciences* 366:1149–1157.

Cunha, Flavio and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2):31–47.

Currie, Janet. 2001. "Early Childhood Education Programs." *Journal of Economic Perspectives* 15 (2):213–238.

de la Cuesta, Brandon and Kosuke Imai. 2016. "Misunderstandings About the Regression Discontinuity Design in the Study of Close Elections." *The Annual Review of Political Science* 19:375–396.

Drange, Nina and Tarjei Havnes. 2018. "Child care before age two and the development of language and numeracy: Evidence from a lottery." *Journal of Labor Economics* Forthcoming.

Duncan, Greg J. and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27 (2):109–32.

Duncan, Greg J and Aaron J Sojourner. 2013. "Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps?" *Journal of Human Resources* 48 (4):945–968.

Elango, Sneha, Jorge Luis Garca, James J. Heckman, and Andrs Hojman. 2016. "Early Childhood Education." In *Economics of Means-Tested Transfer Programs in the United States, volume 2*, edited by Robert A. Moffitt. University of Chicago Press, 235–297.

Eppig, C., C. L. Fincher, and R. Thornill. 2010. "Parasite prevalence and the worldwide distribution of cognitive ability." *Proceedings of the Royal Society, Series B 277* :3801–3808.

Felfe, Christina and Rafael Lalive. 2018. "Does Early Child Care Help or Hinder Childrens's Development?" *Journal of Public Economics* Forthcoming.

Felfe, Christina, Natalia Nollenberger, and Nria Rodrguez-Planas. 2015. "Cant Buy Mommys Love? Universal Childcare and Children's Long-Term Cognitive Development." *Journal of Population Economics* 28 (2):393–422.

Fenson, Larry, Philip S. Dale, J. Steven Reznick, Elizabeth Bates, Donna J. Thal, Stephen J. Pethick, Michael Tomasello, Carolyn B. Mervis, and Joan Stiles. 1994. "Variability in Early Communicative Development." *Monographs of the Society for Research in Child Development* 59 (5):pp. i+iii–v+1–185.

Fiorini, M. and K. Stevens. 2017. "Assessing the Monotonicity Assumption in IV and Fuzzy RD Designs." Mimeo.

Fitzsimons, Emla and Marcos Vera-Hernandez. 2013. "Food for Thought? Breastfeeding and Child Development." IFS Working Papers (W13/31), Institute of Fiscal Studies.

Florens, J. P., J. J. Heckman, C. Meghir, and E. Vytlacil. 2008. "Identification of Treatment Effects Using Control Functions in Models With Continuous, Endogenous Treatment and Heterogeneous Effects." *Econometrica* 76 (5):1191–1206.

Gale, David and Lloyd Shapley. 1962. "College Admissions and the Stability of Marriage." *American Mathematical Monthly* 69:pp. 9–15.

Galsworthy, Michael J., Ginette Dionne, Philip S. Dale, and Robert Plomin. 2000. "Sex differences in early verbal and non-verbal cognitive development." *Developmental Science* 3 (2):206–215.

Garber, Howard. 1988. “The Milwaukee Project: Preventing Mental Retardation in Children At Risk.” Tech. rep., American Association on Mental Retardation.

Gormley, William T and Ted Gayer. 2005. “Promoting school readiness in Oklahoma an evaluation of Tulsa’s pre-k program.” *Journal of Human resources* 40 (3):533–558.

Gottfredson, Linda S. 1997. “Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography.” *Intelligence* 24 (1):13 – 23. Special Issue Intelligence and Social Policy.

Gunderson, Elizabeth A., Sarah J. Gripshover, Carissa Romero, Carol S. Dweck, Susan Goldin-Meadow, and Susan C. Levine. 2013. “Parent Praise to 1- to 3-Year-Olds Predicts Children’s Motivational Frameworks 5 Years Later.” *Child Development* 84 (5):1526–1541.

Hart, Betty and Todd Risley. 1995. *Meaningful Differences in the Everyday Experience of Young American Children* . Baltimore MD: Brookes Publishing Co.

Havnes, Tarjei and Magne Mogstad. 2015. “Is Universal Child Care Leveling the Playing Field?” *Journal of Public Economics* 127:100 – 114.

Heckman, James J. and Stefano Mosso. 2014. “The Economics of Human Development and Social Mobility.” *Annual Review of Economics* 27 (2):109–32.

Herbst, Chris M. 2013. “The Impact of Non-Parental Child Care on Child Development: Evidence from the Summer Participation Dip.” *Journal of Public Economics* 105:86 – 105.

Hewett, ValarieMercilliott. 2001. “Examining the Reggio Emilia Approach to Early Childhood Education.” *Early Childhood Education Journal* 29 (2):95–100.

Imbens, Guido W. and Thomas Lemieux. 2008. “Regression discontinuity designs: A guide to practice.” *Journal of Econometrics* 142 (2):615–635.

John, C. C., M. M. Black, and Nelson C. A. 2017. “Neurodevelopment: The Impact of Nutrition and Inflammation During Early to Middle Childhood in Low Resource Settings.” *Pediatrics* 139:S59–S71.

Koch, Janice. 2003. “Gender Issues in the Classroom.” In *Handbook of Psychology*, vol. 7 (Educational Psychology), edited by William M. Reynolds and Gloria E. Miller. Wiley, 259–284.

Kottelenberg, Michael J. and Steven F. Lehrer. 2014a. “Do the Perils of Universal Childcare Depend on the Childs Age?” *CESifo Economic Studies* 60 (2):338–365.

———. 2014b. “The Gender Effects of Universal Child Care in Canada: Much ado about Boys?” Queen’s University Working Paper, Queen’s University.

———. 2017. “Targeted or universal coverage? Assessing heterogeneity in the effects of universal childcare.” *Journal of Labor Economics* 35 (3).

Lee, David S. and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2):281–355.

Li, Jianghong, Sarah E. Johnson, Wen-Jui Han, Sonia Andrews, Garth Kendall, Lyndall Strazdins, and Alfred Dockery. 2013. "Parents' Nonstandard Work Schedules and Child Well-Being: A Critical Review of the Literature." *The Journal of Primary Prevention* 35 (1):53–73.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2):698–714.

McPherran Lombardi, Caitlin and Rebekah Levine Coley. 2014. "Early Maternal Employment and Childrens School Readiness in Contemporary Families." *Developmental Psychology* 50 (8):2071–2084.

Muenchhoff, M. and P. J. Goulder. 2014. "Sex differences in pediatric infectious diseases." *The Journal of Infectious Diseases* 209 (3):S120–S126.

Noboa-Hidalgo, Grace E. and Sergio S. Urza. 2012. "The Effects of Participation in Public Child Care Centers: Evidence from Chile." *Journal of Human Capital* 6 (1):1 – 34.

Puma, Mike, Stephen Bell, Ronna Cook, Camilla Heid, Pam Broene, Frank Jenkins, Andrew Mashburn, and Jason Downer. 2012. "Third Grade Follow-up to the Head Start Impact Study Final Report." OPRE Report 2012-45, Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Roth, Alvin. 2008. "Deferred acceptance algorithms: history, theory, practice, and open questions." *International Journal of Game Theory* 36 (3):537–569.

Rowe, Meredith L. and Susan Goldin-Meadow. 2009. "Differences in Early Gesture Explain SES Disparities in Child Vocabulary Size at School Entry." *Science* 323 (5916):951–953.

Weiland, Christina and Hirokazu Yoshikawa. 2013. "Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills." *Child Development* 84 (6):2112–2130.

Zigler, Edward and Earl C. Butterfield. 1968. "Motivational Aspects of Changes in IQ Test Performance of Culturally Deprived Nursery School Children." *Child Development* 39 (1):pp. 1–14.