

The cognitive cost of daycare 0–2 for children in advantaged families*

Margherita Fort[§]

Andrea Ichino[¶]

Giulio Zanella[§]

February 8, 2017

Abstract

Exploiting admission thresholds to the highly-reputed daycare system of Bologna, Italy, we show in a RDD that one additional month in daycare at age 0–2 reduces IQ by 0.5% (4.5% of a s.d.) at age 8–14 in a relatively affluent population. This finding is consistent with the hypothesis, suggested in psychology, that children in daycare experience fewer one-to-one interactions with adults, with negative effects in families where such interactions are of particularly high quality. We show in a model that when parents are offered the most preferred daycare program (as opposed to a less preferred one), daycare attendance increases and parents work more or reduce costly market care. At a high earning potential, this increase in family resources is attractive even if it comes at the cost of child IQ. The model lends structure to our RDD, and it is simulated to show that our estimate would be positive in a less advantaged population.

JEL-Code: J13, I20, I28, H75

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

*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 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 (as research assistant), 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 Bologna U., Ben Gurion U., CTRPFP Kolkata, EUI, Florence U., IZA, Hebrew U., Padova U., Princeton U., Siena U., Stockholm U., the 2016 workshop “Interventions during childhood and subsequent child development”, Oslo, and the 2015-2016 Ski and Labor seminars, Laax and S. Anton, as well as to Josh Angrist, Luca Bonatti, Enrico Cantoni, Gergely Csibra, Ricardo Estrada, Søren Johansen, David Levine, Cheti Nicoletti, Giovanni Prarolo and Miikka Rokkanen for very valuable comments and suggestions. Authors’ email addresses: margherita.fort@unibo.it; andrea.ichino@eui.eu; giulio.zanella@unibo.it.

[§]University of Bologna, Department of Economics, and IZA. Fort is also affiliated with CESifo

[¶]European University Institute and U of Bologna, Department of Economics, CEPR, CESifo and IZA

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 is less clear, based on the existing studies of the consequences of alternative childcare arrangements at this very early age.²

We contribute to this literature by studying the causal effect of time spent at age 0–2 in the high-quality public daycare system offered by the city of Bologna, Italy,³ on cognitive outcomes, measured at age 8–14. At this age, the short-lived cognitive 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 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 (from low to high) based on a household size-adjusted function of family income and wealth, which we label “Family Affluence Index” (FAI). The vacant capacity of programs in a given year determines FAI thresholds

¹In the 9 largest non-Scandinavian OECD countries for which data are available the average enrollment rate changed, between 2006 and 2014, from 28.2% to 32.0% in Australia; from 42.4% to 51.9% in France; from 13.6% to 32.3% in Germany; from 28.6% to 24.2% in Italy; from 22.6% to 30.6% in Japan; from 11.2% to 35.7% in South Korea; from 53.9% to 55.9% in the Netherlands; from 42.6% to 38.1% in Spain; from 37.0% to 33.6% in the UK. In the US, this rate increased from 24.1% in 2002 to 28.0% in 2011. In the Scandinavian group the enrolment rate of children under 3 in formal childcare is traditionally large, and yet it rose, between 2006 and 2014, from 61.8% to 65.2% in Denmark; from 26.5% to 27.9% in Finland; from 42.6% to 54.7% in Norway; from 45.7% to 46.9% in Sweden. Daycare 0–2 is also an expensive form of subsidized early education: average public spending across OECD countries was 0.4% of GDP in 2011, or about US \$7,700 (at PPP) per enrolled child (source: OECD Family Database). The 2002 EU council set a target of 33% of children in daycare 0–2 by 2010, but this objective was motivated just as a gender policy.

²The literature on the effects of childcare at age 3–5 is large, but fewer papers study instead what happens at age 0–2, as summarized in [Section 2](#).

³Bologna, 400k inhabitants, is the 7th largest Italian city, as well as the regional capital of Emilia Romagna, a region 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](#)).

such that applicants whose FAI is no greater than the threshold of the most preferred 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 whose parents 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.⁴ The accompanying parent was interviewed by a research assistant, to collect socio-economic information.

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 and to 4.5% of the IQ standard deviation. To interpret this finding and to lend structure to our RDD 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 resources for consumption and improves child outcomes indirectly, and spending time with their offspring (or purchasing market care of comparable quality), which enhances child development directly. The trade-off is complicated by the fact that parental work requires sending children to a daycare program whose quality may be worse than the quality of care at home, particularly in more affluent families. 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.⁵ These interactions are likely to be more effective if complemented by high human capital and high income, which characterize affluent household. In the BDS setting, the adult-to-children ratio at the time our data refer to is 1:4 at age 0 and 1:6 at age 1 and 2, 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-children ratio of about 1.

⁴Wechsler Intelligence Scale for Children (Wechsler *et al.*, 2003). We also collected data on the “Big Five” personality traits, problem behavior, and BMI, finding some weak evidence of favorable health effects and no other statistically significant result. We summarize these results on non-cognitive outcomes in a non-technical report prepared for one of the academic institutions that funded this project (Fort, Ichino, and Zanella, 2016).

⁵See, in particular, Csibra and Gergely (2009, 2011). Other references are reviewed in Section 8.

The central theoretical insight from the model is that when daycare time increases, child IQ decreases in a sufficiently affluent household because of the higher quality of home care (parents, extended family services or care acquired in the market). However, given the high earning potential of an affluent parent and the possibility to substitute other sources of high-quality care with the less expensive daycare provided by the BDS, the loss of child IQ 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 IQ, 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 IQ, because home care is of lower quality than daycare. These different effects of daycare attendance by household affluence induced by the offer of the most preferred program translate into a RD estimand of the IQ effect of daycare around the thresholds that ensure admission to such a program. This estimand may be positive or negative depending on whether the frequency of households attached to each cutoff is skewed towards lower or higher affluence values. A reasonable calibration of the model shows that the negative IQ effect of daycare that we estimate is plausible in the relatively affluent sample that we consider, and that this effect should be even more negative for parents located around thresholds above the median cutoff. This is indeed what we find. The estimated effect would be positive, instead, if the thresholds available for the analysis allowed us to estimate the effect of interest for the universe of applicants, including less affluent households that are far away from the admission cutoffs. Moreover, according to psychologists, one-to-one interactions at age 0–2 should be particularly relevant for girls who, at this early age, are more “mature” than boys, in the sense of being more capable of benefiting from the cognitive stimuli generated by adult-child contacts. Therefore, the IQ impact of daycare for advantaged families should be more negative for girls than boys, and this is again what we find.

After summarizing the economic literature in [Section 2](#), we describe the institutional setting in [Section 3](#). The RD design is constructed in [Section 4](#), and [Section 5](#) describes the interview process. The theoretical model is presented in [Section 6](#). The econometric framework and the results are contained in [Section 7](#). [Section 8](#) reviews the psychological literature providing support for our interpretation of the evidence, 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.⁶ Formal daycare at age 0–2 is an experience of this kind, probably the most important extra-familial one that infants and toddlers can go through during a highly sensitive stage of their life. The economic literature distinguishes between daycare 0–2 (e.g., crèches) and childcare 3–5 (e.g., preschool/kindergarten programs). The case of the older between these two groups has been extensively investigated, often with a special focus on disadvantaged kids,⁷ while we know less about the effects of daycare targeting children in the very first years of their life, especially children from advantaged families.⁸

The few studies in economics that focus on age 0–2 report mixed results. A first group finds, different from us, desirable effects of early daycare attendance for both cognitive and non-cognitive outcomes, concentrated in particular on girls and on children with a disadvantaged family background. In this group, [Felfe and Lalive \(2014\)](#) use administrative data from Schleswig-Holstein to study the effect of daycare 0–2 on language ability and motor skills at age 5–6, instrumenting the probability of attendance with enrollment/children ratios across school districts and exploiting the variability generated by a daycare expansion enacted

⁶See [Borghans *et al.* \(2008\)](#), [Almond and Currie \(2011\)](#), [Heckman and Mosso \(2014\)](#) and [Elango *et al.* \(2015\)](#) for recent surveys.

⁷[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 extremely wide and gains on achievement tests typically fade over time.” (p. 127). Results from the early evaluations of Head Start, the largest randomized study targeting preschoolers, are consistent with these conclusions ([Puma *et al.*, 2012](#)). However, a more careful re-examination of the data, with particular reference to the definition of counterfactuals, reveals positive effects for the disadvantaged population that is targeted by this intervention ([Elango *et al.*, 2015](#)). In line with this finding, [Carneiro and Ginja \(2014\)](#) find persistent health effects of Head Start, using a RD design based on program eligibility rules. [Magnuson *et al.* \(2007\)](#) use the US Early Childhood Longitudinal Study and suggest that pre-kindergarten daycare attendance improves reading and math skills at school entry, but also increases behavioral problems and reduces self-control. [Havnes and Mogstad \(2015\)](#) evaluated a 1975 Norwegian large subsidised expansion of childcare 3–5, concluding that “the benefits of providing subsidized child care to middle and upper-class children are unlikely to exceed the costs.” (p. 101). [Felfe *et al.* \(2015\)](#) reach the same conclusion using data from a similar expansion that took place in Spain during the early 1990s. [Dustmann *et al.* \(2013\)](#) exploit a reform that entitled all German preschoolers to a childcare slot. They find no significant effects for native children and positive effects on school readiness, language and motor skill for children of immigrant parents.

⁸Not so in other disciplines. The early work of Jay Belsky (e.g., [Belsky and Steinberg, 1978](#); [Belsky, 1988](#); [Belsky, 2001](#)) opened the road to studies of the impact of early daycare on the cognitive and non-cognitive outcomes of children, reporting negative consequences that spurred a heated controversy.

in Germany in the early 2000s. They find positive effects which are largest for children whose mothers have attained at most compulsory education as well as for the children of immigrant parents. [Drange and Havnes \(2015\)](#) study the effects of age at entry in daycare 0–2 on language and math test scores at age 7, exploiting the randomization of entry offers at the Oslo public daycare facilities. They find that children who entered daycare at 15 months of age have better test scores than those who entered at 19 months, an effect driven by children from lower income families. With specific reference to Italy, [Del Boca *et al.* \(2015\)](#) show that the benefits of early daycare for children are larger in areas where the rationing system favors more disadvantaged families. 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 \(1968\)](#). They all reached similar conclusions.

On the contrary, studies based on the Quebec universal early daycare extension (a reform that heavily subsidized daycare for children in the age range 0–4 beginning in 1997) typically find undesirable effects on all types of cognitive and non-cognitive outcomes, with losses that are concentrated in particular on boys. A seminal paper in this group is [Baker *et al.* \(2008\)](#), who compare Quebec and the rest of Canada in a Diff-in-Diff design finding that children who benefited from the extension are worse off in terms of behavioral outcomes, social skills and health.⁹ More recently, these authors confirmed the long-run persistence of undesirable effects, with negligible consequences for cognitive test scores ([Baker *et al.*, 2015](#)).

A first important difference between our study and the literature finding positive effects of daycare is that 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

⁹ Three other recent studies, for different countries, provide indirect evidence consistent with this finding for Quebec, exploiting policy changes that alter the amount of maternal care a child receives at 0–2. [Carneiro *et al.* \(2015\)](#) analyze an experiment generated by a 4-month extension of maternity leave enacted in Norway at the end of the 1970s. Looking at very long-run outcomes for treated children — educational attainment and earnings between age 25 and 33 — these authors find positive effects of the extension (i.e., negative effects of less family care at age 0), which are stronger for children of less educated mothers. [Bernal and Keane \(2011\)](#) exploit the 1996 US welfare reform to construct an experiment generating variation in time of maternal care at age 0–2 for children of single mothers. Focusing on children in the 0–5 age range, these authors find a negative effect of less time with mothers on preschool achievement test scores at age 3–6. These effects are larger for children of more educated mothers in this disadvantaged group. [Herbst \(2013\)](#) uses the US Early Childhood Longitudinal Study, Birth cohort (ECLS-B) and estimates negative effects of being in non-parental care at 9 and 24 months on children cognitive scores and motor development. However, outcomes are measured during the treatment and variation of time spent with parents is generated by the comparison between Summer and Winter months.

cities. This is precisely a context in which the quality of one-to-one interactions at home is likely to be better than the analogous 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 making good use of what their families can offer in alternative to daycare, this is the context in which negative effects for girls should emerge more clearly, and in fact they do in our sample.¹⁰

Particularly interesting in relation to our emphasis on advantaged families are the results from another recent evaluation of the Quebec extension by [Kottelenberg and Lehrer \(2017\)](#), who investigate the heterogeneous effects of universal early child care on two outcomes in the short-run: parent-reported motor and social development (at age 0–3) and a Peabody Picture Vocabulary Test (at age 4–5). These authors find, in their own words, “that the Quebec Family Policy significantly boosts developmental test scores for children from single parent households particularly for those who are most disadvantaged and located at the lower quantiles of the distribution. However, children from two-parent families between the 10th and 50th quantile generally receive significant negative impacts from child care. As this group is a large fraction of the sample in the [Baker *et al.* \(2008\)](#) study, it is not surprising that the mean impacts reported are negative in sign.” Our results agree with these conclusions and our theoretical model replicates the pattern they uncover.¹¹

A second possible reason why our results depart from comparable studies pertains to the characteristics of the daycare environment. Both [Felfe and Lalive \(2014\)](#) and [Drange and Havnes \(2015\)](#) study daycare settings (Germany and Norway, respectively) with an adult-

¹⁰Among the studies based on the Quebec extension, [Kottelenberg and Lehrer \(2014a\)](#) and [Kottelenberg and Lehrer \(2014b\)](#) are also of interest because they focus specifically on the heterogeneity of effects at different ages in the 0–4 range and across genders. They show that the negative effects of this intervention are particularly large among kids who start attendance at an earlier age and among boys. More on gender heterogeneity, in their study of the 4-month extension of maternity leave in Norway at the end of the 1970s (see footnote 9), [Carneiro *et al.* \(2015\)](#) report an increase in the school dropout rate of girls who experience less time with their mother at home after birth (p. 403, Table 14), which is consistent with our results and interpretation. However, they do not elaborate on this finding. [Elango *et al.* \(2015\)](#) re-analyze the original data of four demonstration programs (the Perry Preschool Project, PPP, the Carolina Abecedarian Project, ABC, the Infant Health and Development Program, IHDP, and the Early Training Project, ETP) finding more positive effects for boys than for girls, which lead to a substantial gender gap in benefit-to-cost ratios for at least two of them (ABC and PPP). However, they do not seem to find *negative* effects for girls, possibly because these programs are directed to disadvantaged children and are not restricted to the 0–2 age range.

¹¹Both 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 also found by [Noboa-Hidalgo and Urza \(2012\)](#) in Chile for children with disadvantaged backgrounds.

to-child ratio of 1:3. The corresponding ratio at the Bologna daycare facilities during the period that we study was 1:4 at age 0 and 1:6 at ages 1 and 2. From this viewpoint, our study suggests that in order to reconcile parents' child care needs and child development, daycare 0–2 should be designed in a way that ensures a sufficiently high adult-child ratio, if this is cost efficient.

Finally, as far as cognitive outcomes are concerned, most other papers typically focus on math and language test scores, or indicators of school readiness. Our negative estimates refer instead to IQ measured by professional psychologists. There is a general consensus on the fact 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).¹² 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.

Although these are novel patterns in the economic literature, they are not entirely new among psychologists. In a four-decade-old review, [Belsky and Steinberg \(1978\)](#) summarized the findings of daycare research in psychology (some of which employed quasi-experimental methods), 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. A central theme in this early review 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. A specific hypothesis is discussed by [Belsky and Steinberg \(1978\)](#), namely that when negative effects of daycare emerge these are driven by a reduction in maternal involvement and so in children's attachment to their mothers and consequent child insecurity.¹³

¹²The cost of measuring IQ, compared with the increasing availability of almost free administrative data on school outcomes, probably contributes to explain why IQ is used less as an outcome in this literature.

¹³This infant-parent relationship channel occupies a prominent position in subsequent reviews by one of these authors ([Belsky, 1988, 2001](#)).

The maternal channel is at center stage in [Bernal \(2008\)](#), who estimates on NLSY data a dynamic model of the work and child care decisions of the mothers of infants, finding that a mother working full time when her offspring is in the 0–5 age range imposes on the child a cognitive loss quantified in 0.13 standard deviations of the Armed Forces Qualification Test (AFQT) score.

At the level of analysis pursued here, we can’t tell whether the negative IQ effect we uncover is driven by a substitution away from maternal care or from family-based care more generally. However, we show below that in our sample of dual-earner households the counterfactual care mode for the fraction of time children would not have spent in daycare mainly include grandparents and babysitters, in addition to mothers.

3 Institutional setting and administrative data sources

The office in charge of the Bologna Daycare System (BDS) granted us access to the application, admission, and attendance records for all the 66 daycare facilities operating in the City between 2001 and 2005 (of which 8 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 (66 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 2005.¹⁴

The algorithm that matches children to programs is equivalent to a Deferred Acceptance (DA) market design.¹⁵ Parents can apply to as many programs as they wish in the grade-year combination for which their children are eligible. We refer to the set of programs a parent applies to in a given year as the household’s “application set” for that year. Parents are asked by the BDS to provide a preference ordering of the programs in their application set. We show in [Table 1](#) that this ordering favors systematically programs that are closer to

¹⁴See the Online Appendix for descriptive statistics on these programs.

¹⁵See [Gale and Shapley \(1962\)](#) and [Roth \(2007\)](#). [Abdulkadiroglu *et al.* \(2015\)](#) have proposed empirical strategies that exploit the variation induced by a DA mechanism, but they cannot be adapted to our case.

home and, to a lesser extent, also programs with a better reputation.

In the first panel of this table, 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. Statistics are reported by year. 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.¹⁶ The next panel in the same table 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 programs 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 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 ranking is associated with an increased distance of ≈ 0.35 – 0.53 km from home. All these differences are statistically significant at conventional levels.

As for quality, we do not have a reliable objective measure and we have to rely on a reputational indicator that we constructed in the following way.¹⁷ 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 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 $d \in \{1, \dots, 8\}$ denotes the eight resulting 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 .¹⁸ Then the reputation of

¹⁶These results are based on 5,602 children living within the city boundaries. For this analysis we do not consider the remaining households because their preferences over programs are probably affected by commuting patterns on which unfortunately we have no information.

¹⁷We 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.

¹⁸For 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.

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

	2001	2002	2003	2004	2005
<u>Distance statistics:</u>					
mean [s.d.]	4.06 [2.21]	4.13 [2.27]	4.09 [2.24]	4.14 [2.27]	4.15 [2.28]
min / max	0.02 / 13.71	0.07 / 14.25	0.02 / 14.10	0.02 / 13.68	0.01 / 14.12
<u>Mean distance from home to:</u>					
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)
<u>Quality statistics:</u>					
mean [s.d.]	0.00 [0.30]	0.04 [0.42]	0.03 [0.38]	0.00 [0.48]	0.01 [0.47]
min / max	-1.22 / 0.75	-1.16 / 1.18	-1.35 / 1.67	-2.11 / 1.58	-1.54 / 1.31
<u>Mean quality of:</u>					
most preferred	0.07 (0.01)	0.14 (0.01)	0.11 (0.01)	0.12 (0.01)	0.13 (0.01)
second most preferred	0.07 (0.01)	0.12 (0.01)	0.12 (0.01)	0.11 (0.01)	0.11 (0.01)
third most preferred	0.06 (0.01)	0.12 (0.01)	0.11 (0.01)	0.05 (0.01)	0.06 (0.01)
not most preferred and ranked	0.05 (0.00)	0.10 (0.01)	0.09 (0.00)	0.04 (0.01)	0.05 (0.01)
not ranked	-0.00 (0.00)	0.03 (0.00)	0.02 (0.00)	-0.00 (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 following 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 up to a maximum of 4 km. Results are based on 5,602 children located within the city boundaries of Bologna. Standard errors are in parentheses. Standard deviations in brackets.

program j among households in distance cell d is defined as

$$q_{jd} = \bar{r}_{jd} - \bar{r}_{-jd}, \quad (1)$$

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.¹⁹ To capture the overall reputation of program j , we compute the average

$$q_j = \frac{1}{8} \sum_{d=1}^8 \bar{r}_{jd} - \bar{r}_{-jd}. \quad (2)$$

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 1.7 positions better than its alternatives, while the worst program, in 2004, is 2.1 positions worse than its alternatives, on average.

Now consider, as an example, a hypothetical grade-year combination with only three available programs, a , b and c . If all eligible households unanimously ranked these programs in the same way, ($a \succ b \succ c$) at all distances, then their reputation would be ordered as $q_a > q_b > q_c$. In the absence of agreement among households, instead, the reputation of the three programs would be similar: $q_a \approx q_b \approx q_c$. 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

¹⁹The average number of households i , for each combination of program j and distance d , is 138 (s.d. 108) and ranges between 11 (s.d. 10) in distance cell 1 (from 0 to 0.5 km) and 178 (s.d. 89) in distance cell 8 (from 3.5 to 4 km).

2003, there is no statistically significant difference between the average values of q_j for the programs that are ranked in the top positions.²⁰ Only in 2004 and 2005 the reputation of the most preferred program (0.12 and 0.13, respectively) is significantly larger, at conventional levels, than the quality of the average not most preferred but ranked program (0.04 and 0.05, 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 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 the statistics of [Table 1](#).

Demand for admission to the BDS 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 (lowest priority) 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.²¹ Families with a lower value of the index (i.e., less affluent families) have higher priority within a basket.

The DA algorithm determines for each program a “Final” FAI admission threshold defined as the FAI of the most affluent child who receives an offer for that program and accepts it. In [Section 4](#) we show how these Final FAI thresholds can be used to construct a valid RD design. Before doing so, four remarks are in order. First, children can be classified in three mutually exclusive and exhaustive ways: the “admitted and attendants”, who have received

²⁰Out of the total of first-time applicants from the universe to be described below, 61.6% are offered their most preferred program and 89.7% receive an admission offer. Of these, 92.8% are offered one of their first three choices. In the smaller estimation sample (see [Section 5](#)), 47.3% are offered their most preferred program and 75.2% receive an admission offer. Of these, 91.4% are offered one of their first three choices.

²¹The Online Appendix provides details about how this index is constructed.

an admission offer and have accepted it; the “reserves”, who have not received any offer; the “admitted and waivers” (or “waivers” for brevity), who have received an admission offer and have turned it down. It is important to keep in mind that 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. Therefore, since we construct the RD design on the basis of *the first application of each child*, 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 the design.

Second, a child’s FAI is relevant not only for admission, but also for the monthly attendance fee that households must pay if they accept an offer, independently of actual days of attendance during the month. This fee is a function of a child’s FAI, which is well known to potentially interested families before they decide whether to apply.²² As such, it should not pose problems in our analysis, which is conditional on households who have already decided to apply, and it is in any case continuous by construction at the thresholds that we will use in our design.

Third, in [Section 5](#) we illustrate how we match the administrative data with information on family characteristics and children outcomes obtained with interviews. To ensure a greater homogeneity of the interview sample, 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): 6,575 *first applications* to 890 programs originate from this basket in the period 2001-2005. Of these programs, 74 end up with no vacancies for Basket 4 children (i.e., the Final FAI threshold is in Basket 1, or 2, or 3); 271 have sufficient capacity for all Basket 4 applicants (i.e., the final FAI threshold is in Basket 5), and 545 offer admission to some but not all the Basket 4 applicants (i.e., the Final FAI threshold is in Basket 4).²³ The remaining 51 (to reach the total of 941 programs) do not receive applications from Basket 4. Some tables and figures below are based on sub-groups of this sample, for the reasons explained in the respective notes.

²²The 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).

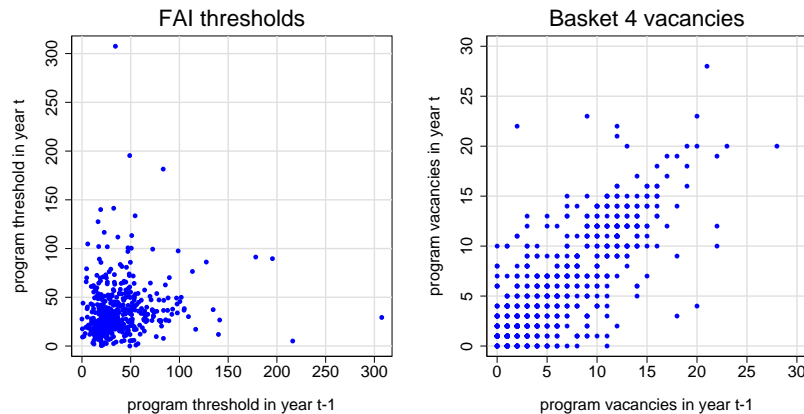
²³These are the programs that we effectively use in the analysis. Descriptive statistics are provided in the Online Appendix.

Fourth, the population of applicants to the BDS is relatively affluent. As detailed in Table 1 of the Online Appendix, the average FAI across the five baskets is about €20k, corresponding to a gross annual household income of about €54k. Households in Basket 4 are even more affluent: the average FAI in this group is about €25k, corresponding to an income of about €67k. This is roughly twice the annual gross household income in Italy at the time the data refer to (all values are in constant 2010 euros).

4 How FAI thresholds can be used for the RD design

We begin by showing that 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. Figure 1 shows that this is unlikely to happen: the two panels plot, for each program, the Final FAI thresholds (left panel) and the Basket 4 vacancies (right panel) in year t against the corresponding thresholds and vacancies in year $t - 1$.

Figure 1: Variability of Final FAI thresholds and of Basket 4 vacancies over time



Notes: Each dot represents a program and the coordinates are either the Final FAI thresholds of that program in two consecutive years (left panel), or the vacant capacity for Basket 4 children in two consecutive years (right panel). FAI stands for Family Affluence Index. Sample: 238 programs with rationing for Basket 4 children in two consecutive years. The lack of persistence of Final FAI thresholds is of course even more evident for programs that are not offered every year, which cannot be represented in this figure.

For both these variables, a prediction based on lags would be highly imprecise and, for an accurate guess, families would need a formidable amount of additional information,

like 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. Thus, even if some parents try to manipulate their FAI, they don't know by how much they should reduce the index in order to receive an offer from a specific program.

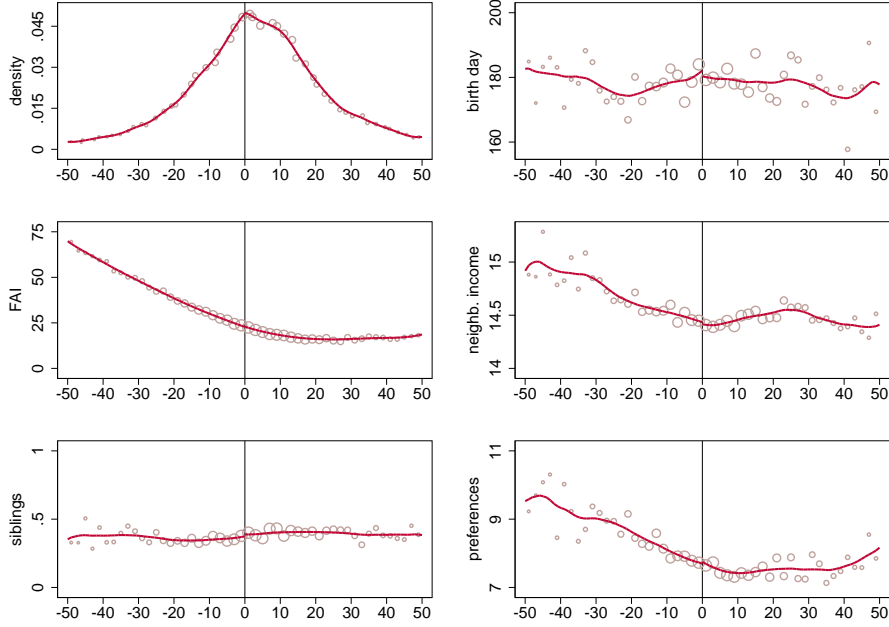
The ultimate proof of this claim is provided by the continuity of the FAI density and of pre-treatment covariates. We assess this continuity in [Figure 2](#), stacking thresholds and centering them at zero so that the FAI distance from each threshold is the running variable. In the top left panel the density of observations is plotted and the [McCrary \(2008\)](#) test rejects the existence of a discontinuity.²⁴ Five relevant pre-treatment covariates are considered in the remaining panels (birth day in the year, FAI, average income in the city neighborhood where the program is located, number of siblings at the first application, and number of programs listed in the application set) and again no discontinuity emerges at the thresholds.²⁵

Given the absence of any evidence of manipulation of the admission process at the BDS, it would then 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 in the analysis. 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. We next show how we circumvent this problem in order to associate every child with one threshold only, so that we can estimate effects of days of attendance in the BDS *independently of the specific program*.

²⁴The log discontinuity of the density is -0.007 with a s.e. of 0.055.

²⁵In these panels, a circle represents the average value of the corresponding covariate in bins of €2k size, and the size of a circle is proportional to the number of observations in this bin. Solid lines represent estimated conditional mean functions smoothed with LLR using all individual observations separately on the two sides of the cutoff. For the reasons explained in [Fort et al. \(2017\)](#), observations with exactly zero distance from FAI thresholds are dropped in the construction of these figures as well as in the related continuity tests. A triangular kernel and optimal bandwidth from [Calonico et al. \(2014b\)](#) are used here and in all of the remaining similar figures below.

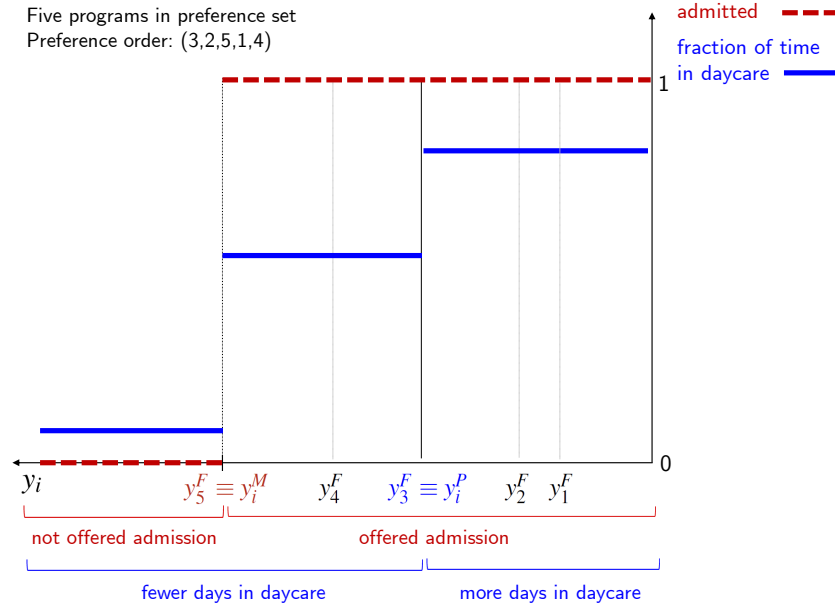
Figure 2: Density of distance and continuity around Final FAI thresholds



Notes: The circles represent the frequency distribution (top-left panel) and the average of five pre-treatment variables (remaining panels) inside €2k bins, plotted as a function of the distance (thousands of real €) of a child’s FAI from her Final FAI threshold. The size of a circle is proportional to the number of observations in the corresponding €2k bin. The bold lines are LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from [Calonico et al. \(2014b\)](#). FAI stands for Family Affluence Index. 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, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in [Fort et al. \(2017\)](#), is different from zero.

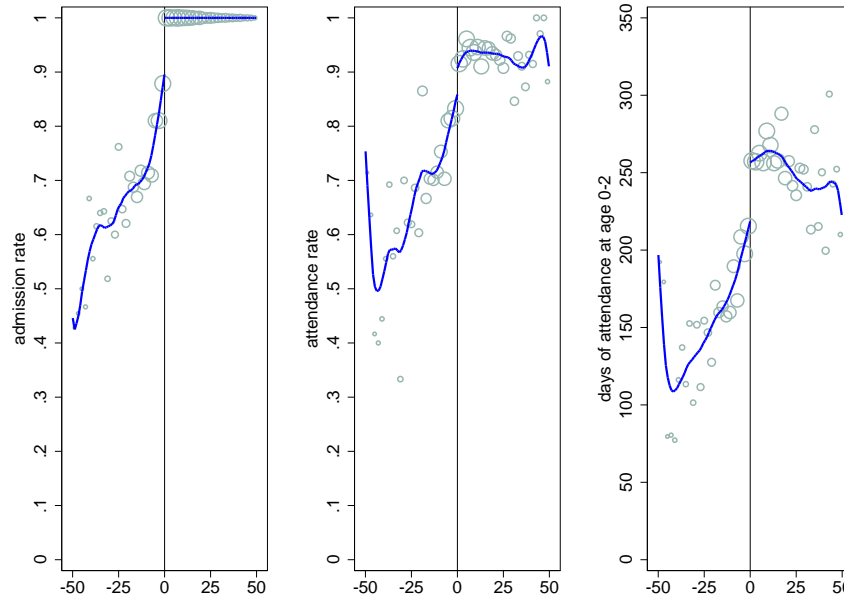
To this end, [Figure 3](#) illustrates the hypothetical situation of parent i who applies for the first time in a given year to five programs out of a total of J programs her child is eligible for. Without loss of generality, $j \in \{1, \dots, 5\}$ denotes these programs, which the parent ranks in the following order: $3 \succ 2 \succ 5 \succ 1 \succ 4$. Let y_i be the FAI of parent i and y_j^F the Final FAI threshold of program j . In [Figure 3](#) these thresholds are ordered along the horizontal axis from the highest on the left to the lowest on the right. $y_5^F \equiv y_i^M$ is the Maximum FAI threshold in i ’s application set. Therefore, if $y_i > y_5^F$ then parent i does not receive any offer at first application because her FAI is too high to qualify for any of the programs in her application set. If, instead, $y_i \leq y_5^F$ then with probability 1 the parent receives at least one offer at first application and possibly qualifies for more than one program if y_i is lower than other thresholds. Thus, the probability of qualifying for at least one program when first applying jumps sharply from 0 to 1 at the y_5^F threshold. This probability is represented by the bold dashed line in [Figure 3](#).

Figure 3: Maximum and Preferred FAI thresholds in the application set



Notes: This figure illustrates the definition of Maximum and Preferred FAI thresholds for a hypothetical child whose parents first apply for admission in a given year listing five programs in the application set. FAI stands for Family Affluence Index.

Figure 4: 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 LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from Calonico *et al.* (2014b). 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 to programs with rationing, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in Fort *et al.* (2017), is different from zero.

The bold solid line, instead, is the fraction of time spent by the child in daycare. This fraction is highest if $y_i \leq y_3^F \equiv y_i^P$, i.e., if the parent receives an offer from the preferred program in her application set — program 3 in this example. However, even if the child qualifies for the program preferred by the parent, this fraction is not equal to 1 because the parent may decide not to fully exploit the offer. We show in [Section 6](#) why these expectations are reasonable in a theoretical model of parental decisions about daycare attendance.²⁶

What matters for our purposes is that, for each child, the fraction of time spent in daycare should jump discontinuously at both the Preferred (y_i^P) and the Maximum (y_i^M) FAI thresholds. In principle, both kinds of thresholds could be used in a RD design, but parents have some control over their Maximum FAI threshold because they can (weakly) increase it by adding more programs to their application set. For this reason, we focus the analysis on Preferred thresholds. Note that these are unique for a child so that, when using them, there are no repeated records and each child is used only once in the analysis.

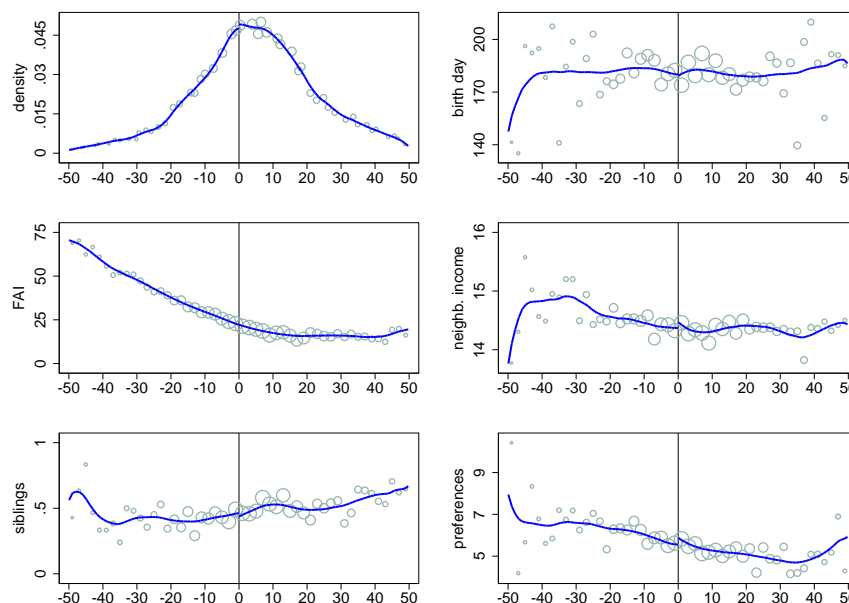
[Figure 4](#) shows how probability and days of attendance change in a discontinuous way around Preferred thresholds. The running variable is the FAI distance from the cutoff, with positive values indicating a FAI lower than the threshold. In the left and middle panels the admission and the attendance rates increase sharply (by 10.4 and 4.9 percentage points, respectively) as the FAI crosses the cutoff from higher to lower values, with some fuzziness due to the reasons discussed in the comment to [Figure 3](#). These discontinuities translate into a jump of nearly two months (38 working days) of total daycare time in the right panel.

In [Figure 5](#) we show that the frequency of observations and pre-treatment covariates are all continuous around Preferred FAI thresholds in the universe, supporting the validity of a RD design constructed around them.²⁷ We next describe how we collected information on cognitive outcomes.

²⁶ The fraction of time spent in daycare does not change if the household has a low enough FAI to qualify in programs 2 and 1, because the parent would still be offered the strictly preferred program 3. If $y_i > y_3^F$ (but still $y_i \leq y_5^F$) then the fraction is lower (relative to the $y_i \leq y_3^F$ case) because the parent receives offers but not from the preferred program. In this example, daycare time is constant for FAI levels between y_3^F and y_5^F because in all these cases the parent will receive an offer from program 5, which she strictly prefers to program 4. It may appear surprising that daycare time is positive for FAI values above y_5^F . This fuzziness occurs because a parent may not qualify for any program at her *first* application but re-apply, be offered admission, and accept it in later years (if the child is not older than 2).

²⁷Using the [McCrary \(2008\)](#) test, the log discontinuity of the density is 0.022 with a s.e. of 0.13.

Figure 5: Density of distance and continuity of covariates around Preferred FAI thresholds



Notes: The circles represent the frequency distribution (top-left panel) and the average of five pre-treatment variables (remaining panels) 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 LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from [Calonico et al. \(2014b\)](#). 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 to programs with rationing, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in [Fort et al. \(2017\)](#), is different from zero.

5 The interview sample

The administrative records that we have received from the City of Bologna do not contain children outcomes at any stage of their development, nor they contain pre-treatment family characteristics beyond the few ones mentioned above. Therefore, we have 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 1,379 households with a FAI sufficiently close to Final FAI thresholds and which 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. IQ was measured at this age because we are not interested in short-lived effects of daycare. It would be of great interest to explore outcomes at an even later age, but the available BDS administrative data do not allow us to

go back to application and attendance records before the 2001-2005 period. 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 phone.

Upon arrival at the interview site (a dedicated space at the University of Bologna), the child was administered an IQ test by a professional psychologist, and the accompanying parent was interviewed in a separate room by a research assistant to collect socioeconomic information. The test we used is the “Wechsler Intelligence Scale for Children” (WISC-IV), which measures Full Scale IQ.²⁸ The average IQ of interviewed children is 116.4 on a scale normalized by age and with mean equal to 100 for the Italian population of children in the same age range who took the WISC-IV.²⁹ The standard deviation is about 12.4. Overall, each child and the accompanying parent spent about 3 hours at the interview site.

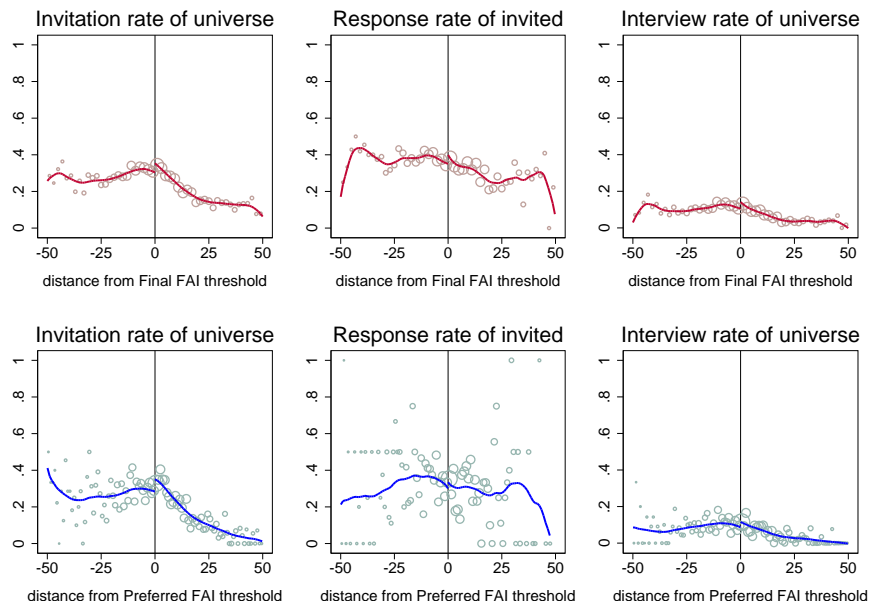
We obtained information on 458 children, corresponding to a response rate of 33.2% of the invited (about 40% in proximity of Final FAI thresholds, as shown below). Of these interviews, only 444 provided a complete set of variables to be used in the econometric analysis.³⁰ The left panels of Figure 6 show that the invitation rate of households in the Basket 4 universe exhibits some discontinuity around Final and Preferred thresholds. Given the essential continuity of the response rates of the invited in the middle panels, the jump in the invitation rate induces a small discontinuity in the interview rate with respect to the universe (right panels). These small jumps are not a source of concern in the light of the continuity documented for all the observable covariates in Figures 2 and 5 for the universe, as well as in Section 7 for the interview sample.

²⁸Wechsler *et al.* (2003). Table 6 in the Online Appendix contains descriptive statistics for the Full Scale IQ and the four underlying sub-scales (verbal ability, working memory, perceptual reasoning and processing speed) in the interview sample. Tables 7-14 in the same appendix replicate the econometric analysis of the main text for the four sub-scales. With different degrees of intensity, the results for the full scale hold similarly for the sub-scales.

²⁹The top-left panel of Figure 9 in the Online Appendix confirms the absence of differences by age in the IQ score produced by the WISC-IV.

³⁰In 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 socio economic questions, thus generating missing values in some relevant pre-treatment variables.

Figure 6: Invitation, response, and interview rates around Final and Preferred FAI thresholds



Notes: The circles represent the invitation rate for the universe (left), the response rate of the invited (middle), and the interview rate for the universe (right) inside €2k bins, plotted as a function of the distance (thousands of real €) of a child’s FAI from either her Final FAI thresholds (top) or her Preferred FAI threshold (bottom). The size of a circle is proportional to the number of observations in the corresponding €2k bin. The bold lines are LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from [Calonico et al. \(2014b\)](#). FAI stands for Family Affluence Index. Sample: 5,937 (top row) and 5,363 (bottom) children with two working parents, born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005 to programs with rationing, whose FAI distance from the Final (top) or Preferred (bottom) FAI thresholds is at most €50k.

To gauge the representativeness of this sample with respect to the Basket 4 universe, it is important to keep in mind that 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. This is indicated by the size of the circles in the upper part of [Figure 6](#), which are proportional to the fraction of observations in the corresponding bins. The consequences of this choice are reflected in the descriptive statistics of key administrative variables for the Basket 4 universe, the invited and the interviewed samples that we illustrate in [Table 2](#). The p-values reported in the last column refer to tests of the equality of means for the Basket 4 universe and the invited (first row), for the invited and the interviewed (second row, in square brackets) and for the Basket 4 universe and the interviewed (third row, in curly brackets). The general pattern suggests, as expected, that there are no significant differences between the interviewed and the invited, while both these groups differ in some dimensions with respect to the Basket 4 universe.

Table 2: 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 (20.50)	26.50 (19.70)	27.10 (17.55)	0.007 [0.547] {0.010}
Number of preferences at first application	5.42 (3.66)	5.29 (3.42)	5.59 (3.53)	0.222 [0.120] {0.341}
Siblings at first application	0.48 (0.66)	0.49 (0.65)	0.54 (0.70)	0.755 [0.151] {0.079}
Day of birth in the year	182.8 (104.1)	186.6 (106.)	180.5 (111.1)	0.222 [0.310] {0.673}
Offered admission at first application	0.897 (0.303)	0.777 (0.417)	0.752 (0.432)	0.000 [0.297] {0.000}
Offered preferred program at first application	0.616 (0.486)	0.511 (0.500)	0.473 (0.432)	0.000 [0.169] {0.000}
Waiver at first application	0.124 (0.330)	0.075 (0.263)	0.068 (0.251)	0.000 [0.607] {0.000}
Year first applied	2003.1 (1.43)	2003.4 (1.42)	2003.5 (1.38)	0.000 [0.086] {0.000}
Year child born	2002.0 (1.58)	2002.5 (1.63)	2002.6 (1.62)	0.000 [0.086] {0.000}
Grade first applied for	0.882 (0.786)	0.568 (0.673)	0.540 (0.676)	0.000 [0.459] {0.000}
Days in	212.2 (143.3)	223.6 (151.4)	230.5 (156.3)	0.010 [0.417] {0.017}
Ever attended (share with days in >0)	0.847 (0.360)	0.784 (0.411)	0.782 (0.414)	0.000 [0.916] {0.001}
<i>N</i>	6,575	1,379	444	

Notes: This 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), in the sample invited for an interview (1,379 children from this universe), and in the interview sample (444 children interviewed from the universe). The p-values in the last column refer to tests of the equality of means for the Basket 4 universe and the invited (first row), the invited and the interviewed (second row, in square brackets), the Basket 4 universe and the interviewed (third row, in curly brackets). FAI stands for Family Affluence Index.

For instance, the invited and the interviewed children have a slightly higher FAI than the universe. This is not surprising given how we invited families and the fact that the evolution of the admission process pushes Final FAI thresholds towards higher FAI values (see [Section 3](#)). We also see, in this table, that the offer rate is substantially higher in the universe than in the interview/invited samples. This happens because 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. These are all consequences of the way we selected the invited families, that we traded off to gain homogeneity and comparability at the FAI thresholds. The number of preferences and the number of children in the household at first application are all similar across the three samples.

Other variables in [Table 2](#) exhibit significant differences across groups because of the sampling design. Children in the Basket 4 universe are slightly younger, have first applied for lower grades, and have spent more days in daycare. There are twice as many children turning down offers in the Basket 4 universe as in the invited/interview samples. However, once again, these differences are not a threat to the internal validity of our RD design, given the continuity of covariates and densities at the thresholds.

As for external validity, [Table 3](#) compares the means of selected socioeconomic variables, that are available only for the interview sample, with corresponding means for representative samples of the population of families with two employed parents in Northern Italy.³¹ The comparison reveals that the interview sample is, by and large, representative of the corresponding Italian population in terms of demographics. However, parents are slightly more educated and less frequently self-employed. The higher educational attainment of the parents in the interview sample is relevant for the interpretation of our results, because it

³¹For age and educational characteristics we used the Bank of Italy Survey of Household Income and Wealth (SHIW), a biennial survey that can be weighted to represent the Italian population. From the waves of this survey, we selected observations to mimic the Basket 4 universe of the BDS administrative files in 2001-2005. Specifically, we restricted the analysis to households with two 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 at least one child between 8 and 14 years of age. For parental occupation we used the ISTAT Labor Force Statistics, selecting workers of the 2005 wave, in the 25–44 age range (i.e., the age range of parents in our sample when they first applied for daycare admission 8 to 12 years prior to the interview).

is one of the reasons why, different from other studies, our estimated effects of daycare 0–2 refer to children who, at home, can enjoy a relatively richer cultural environment by Italian standards.

In the next section we discuss the economic interpretation of the causal estimand that we can identify in this interview sample around Preferred FAI thresholds.

Table 3: The interviewed sample in comparison to the Northern Italian population

	Interview sample	Northern Italy
Child age	10.7 (1.6)	11.1 (1.7)
Father age	47.3 (4.8)	47.0 (4.7)
Mother age	44.9 (4.1)	45.0 (4.8)
Years education father	14.2 (3.8)	13.1 (3.0)
Years education mother	15.5 (3.2)	14.4 (2.5)
Father self-employed	0.236 (0.425)	0.276 —
Mother self-employed	0.106 (0.308)	0.173 —
Observations	444	93

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 – age, education) and in the Labor Force Statistics (ISTAT – parental occupation).

6 The economics of RD based on preferred thresholds

The economics of our RD design is presented using a framework that accommodates three variants of the same model. We begin with a baseline version that is stripped down to the essentials in order to isolate the fundamental driving forces in a transparent way. This baseline model, presented in [Section 6.1](#), shows how children are affected by the decisions of their parents who face a trade-off between spending time with their offspring, which enhances child development directly, and spending time at work, which increases family resources for

consumption and improves child outcomes indirectly.³² This trade-off is complicated by the fact that parental work may require sending children to a daycare program whose quality may be worse than the quality of care at home, particularly for affluent families.

Different parents resolve the trade-off in different ways and we are interested in showing how their choice is affected by just qualifying for their preferred daycare program, versus just being excluded from it. The analysis is conducted conditioning on the fact that parents have applied for daycare and therefore must derive positive utility from being offered at least the program that they most like. Such an offer induces parents to increase daycare attendance of their child, with respect to the case in which qualification is only for a less preferred program, because the most preferred one is closer to home, and thus less costly, or it is of better quality, or both (see [Section 3](#)).

The key theoretical insight of this baseline model is that in a sufficiently affluent household child IQ decreases, because of the higher quality of home care, if daycare time increases and the parent works more. However, given the high earning potential of an affluent parent, the loss of child IQ is more than compensated by the increase of household consumption net of daycare cost. Therefore, the affluent parent takes advantage of the offer of the most preferred program even if it decreases child IQ, as long as she cares enough about household consumption; otherwise, she would not have applied for daycare. For a less affluent household, instead, the offer of a more preferred program generates both an increase of consumption and an increase of child IQ, because home care is of lower quality than daycare. These differential effects of daycare attendance by household affluence, which are induced by the offer of the most preferred program, translate into an ITT-RD estimand of the effect of daycare on IQ around Preferred thresholds. This may be positive or negative depending on whether the frequency of households attached to each RD cutoff is skewed towards lower or higher affluence values. The baseline model allows us to shed light also on the possibility of differential effects of daycare by gender.

We then present in [Section 6.2](#) results from a numerical solution of a more general model embedding the actual features of the BDS setting, like the presence of other types of child

³²Our framework builds on the theory of the allocation of time ([Becker, 1965](#)) and on the economic theory of human development ([Carneiro *et al.*, 2003](#); [Cunha and Heckman, 2007](#)). A similar framework is employed by [Bernal \(2008\)](#).

care providers beyond parents and daycare, the existence of other sources of subsistence consumption in case of no labor income, the specific schedule relating daycare fees to the FAI, and the coincidence of Maximum and Preferred thresholds in the application set. This more realistic model confirms the existence of a level of affluence above which qualification for the preferred program induces an IQ loss for the child, which is nevertheless compensated by a sufficiently large increase in consumption. Differently than in the baseline model, the increase in consumption originates from both an increase in labor supply and a reduction of expenditures in market care. Using the observed density of the FAI, we use this extended model also to show that while the ITT-RD estimand predicted for the entire population of applicants in all baskets is positive, the corresponding estimand for the more affluent applicants in Basket 4 is negative.

Finally, we extend the baseline model in [Section 6.3](#) to take into account the intertemporal aspect of the problem faced by parents, who can apply for daycare immediately after the birth of their child or later. The relevant trade-off is analogous to the one just described: in order to spend time with the child when it is relatively more effective (i.e., immediately after birth), the parent must leave the labor market for a longer period, with detrimental consequences on future earnings and thus on future resources for child development and consumption. The novel implication of this dynamic analysis concerns the interpretation of the differential response of child IQ to the offer of the most preferred program by household affluence. Under reasonable assumptions, affluent parents obtain the maximum increase in utility by applying for daycare immediately after birth, in order to avoid the depreciation of their high earning potential deriving from a prolonged interruption of labor market attachment. This generates a larger increase in consumption that compensates the IQ loss of the child who attends daycare “too early”, when home care is of particularly higher quality than daycare and dynamic complementarities in skill formation cannot be fully exploited. Less affluent households, instead, delay the application for daycare (a prediction indirectly supported by our evidence) because the increase in consumption deriving from continued labor market attachment is lower, and daycare becomes desirable only if it does not have detrimental effects on child IQ.

6.1 Baseline model

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 cognitive skill of the child, θ , which we refer to as “IQ”.³³ The utility function is

$$u(c, \theta) = c + \alpha\theta, \quad (3)$$

where $\alpha > 0$ captures the relative weight of IQ in parental preferences.

Two forms of child care are available: family care and a daycare system offering a set of programs indexed by z . 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. The first one is a transportation cost $k(z)$ incurred by the parent to reach the facility. The second one is an attendance fee ϕy_{-1} , with $\phi < 1$, that is identical for all programs and is proportional to past household income $y_{-1} = w(\theta_g)h_{-1}$, where $w(\theta_g)$ is the wage rate (increasing in parental IQ) and $h_{-1} \in [0, 1]$ is past labor supply. Therefore, $\pi_d = k(z) + \phi y_{-1}$. Note that y_{-1} is the theoretical counterpart of the FAI. For a given past labor supply, it correlates positively with the permanent component of household affluence which is captured by the wage rate.

Without loss of generality, we assume that the programs for which parents are eligible can be ordered in a way such that the function $s(z) = \alpha q_d(z) - k(z)$ is strictly increasing in z . We will later show that, thanks to this assumption, derived utility of parents is also increasing in z and therefore z is the ordering of programs provided by parents to the BDS.

IQ is determined at age 0–2 by household resources and the quality of care. Household resources are summarized by parental income $y = hw(\theta_g)$ where $h \in [0, 1]$ is labor supply. Denoting, respectively, by τ_g and τ_d time spent by the child in parental care and in daycare, the technology of IQ formation is given by

$$\theta = q_g y \tau_g + q_d(z) \tau_d, \quad (4)$$

³³We are indifferent between treating parental preferences over child IQ as direct – i.e., the parent values child’s θ per se – or indirect – i.e., the parent values the future earnings of the child, which increase in IQ. To simplify the analysis we do not separate the consumption of the parent from that of the child, who benefits like the parent from household consumption.

where $q_g y$ 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 differs between children because it is complemented by the cognitive and economic resources of the household, summarized by y .

A child requires a fixed amount of care time, which we normalize to 1. Therefore, the chosen child care arrangement must satisfy³⁴

$$\tau_g + \tau_d = 1. \quad (5)$$

Eqs. 4 and 5 indicate that parental care and daycare are perfect substitutes at rate 1 in child care time, but are substitutes at rate $\frac{q_d}{q_g y}$ in child development. The parent does not value leisure and splits the time endowment (assumed equal to the care time required by the child) between work for pay and parental child care, so that a parent's time constraint is $h + \tau_g = 1$,³⁵ and the budget constraint is $c + \pi_d \tau_d = wh$. Therefore, the parent solves:

$$\max_{c, \tau_d} c + \alpha \theta \quad \text{s.t.} \quad \begin{cases} c = (w - \pi_d) \tau_d \\ \theta = q_g w \tau_d (1 - \tau_d) + q_d(z) \tau_d \\ \pi_d = k + \phi y_{-1} \\ 0 \leq \tau_d \leq 1 \end{cases} \quad (6)$$

The key trade-off in this problem is that increasing τ_d adds resources for consumption and the home care of children, if the wage rate is greater than the unit cost of daycare, but it reduces parental time with children, with a negative direct effect on IQ if the quality of daycare is smaller than the quality of parental care.

Let $A = [\underline{z}, 1]$ be the application set of a parent, i.e., the subset of programs for which the child is eligible and for which the optimization problem has an interior solution. The “reservation program”, $\underline{z} \geq 0$, will be determined below. This interior solution is given by

³⁴The more realistic possibility that allows for the existence of a third type of care acquired in the market or within the extended family will be considered in [Section 6.2](#)

³⁵In the corner solution in which $\tau_g = 1$, the parent does not work ($h = 0$) and the child does not attend daycare ($\tau_d = 0$). Given [Eq. 4](#), θ would be equal to 0 in this case, which can be taken as a normalization for IQ corresponding to the level of the cognitive skills of a child entirely cared by a parent with zero earnings. Consumption c would also be zero in the absence of other sources of income or wealth in the household, which we do not consider here for simplicity. [Section 6.2](#) will take care of these issues in a more realistic way. Note, incidentally, that our data refer to dual-earner households, to which this corner solution does not apply.

$$\tau_d^*(z) = \frac{1}{2} + \frac{w + \alpha q_d(z) - k(z) - \phi y_{-1}}{2\alpha q_g w} \in (0, 1), \quad (7)$$

where $w - k(z) - \phi y_{-1} = w - \pi_d > 0$ because $c^* > 0$, and parental utility at the optimum is

$$u^*(z) = \tau_d^*[w - \phi y_{-1} + \alpha q_g w(1 - \tau_d^*) + \alpha q_d(z) - k(z)]. \quad (8)$$

To simplify the presentation of results, let z take real values in the $[0, 1]$ interval. Then, using the envelope theorem, $\frac{du^*}{dz} = [\alpha q'_d(z) - k'(z)]\tau_d^*$ 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 $u^*(z)$ must be strictly increasing in z . Therefore, the condition

$$\alpha q'_d(z) - k'(z) > 0 \quad (9)$$

holds and the ranking z is consistent with derived preferences over programs. Note that [Eq. 9](#) is satisfied, on average, in our setting, as shown by the evidence presented in [Table 1](#): programs that are ranked higher by parents are typically closer to home, $k'(z) < 0$, and of weakly better quality, $q'_d(z) \geq 0$.

6.1.1 Key predictions of the model for a specific household

Consider a parent who compares the offer of a more preferred program vs. a less preferred one, i.e., an increase in z . Using [Eq. 9](#) we can derive a first unambiguous prediction, which follows from differentiating the optimal daycare time in [Eq. 7](#) with respect to parental ranking:

Remark 1 *The offer of a more preferred program increases daycare time*

$$\frac{d\tau_d^*}{dz} = \frac{\alpha q'_d(z) - k'(z)}{2\alpha q_g w} > 0. \quad (10)$$

This inequality holds, in particular, around the threshold of the most preferred program ($z = 1$) and is in line with the evidence reported in the right panel of [Figures 4](#) (for the Basket 4 universe) and [10](#) (for the interview sample), which is the first stage of our RDD. This happens because, given the sign of the derivative in [Eq. 9](#), the possibility to attend a more preferred program weakly increases the utility derived from daycare quality and proximity,

thus making a longer daycare time more desirable. Moreover, Eq. 10 implies a smaller first stage for more affluent (high w) households. This prediction, too, is supported by the data. In the Basket 4 universe, the offer of the most preferred program increases daycare attendance by 64.1 days (robust s.e. 8.3) for children whose FAI is below the median FAI, and by 46.2 days (robust s.e. 11.8) for children whose FAI is above the median.³⁶ The intuition here is that the desirability of a longer daycare time following the offer of a more preferred program is weaker for more affluent parents because of the higher quality of their home care.

Using Remark 1 we can characterize the application set of a parent by considering the possibility that $\tau_d^*(z) = 0$ for some program with rank $z \in [0, 1]$. This happens when

$$w - k(z) - \phi y_{-1} + \alpha(q_g w + q_d(z)) = 0. \quad (11)$$

If a program satisfying this condition exists, then, given Eq. 10, the application set of the parent is $A = [z, 1]$. If $z = 0$, the application set coincides with the entire set of programs for which the child is eligible. If $0 < z < 1$ then a corner solution with $\tau_d^* = 0$ exists for any program ranked $z < z$. If $z = 1$, the parent lists only the most preferred program in the application set, while if there is no program for which $u^*(z) > 0$ with $\tau_d^* > 0$, the parent does not apply to any program. This non-participation condition occurs when $u^*(z = 1) < 0$. Finally, there may exist a program $\bar{z} \in (z, 1]$ such that $\tau_d^*(\bar{z}) = 1$. In this case, for the subset of programs ranked $z \geq \bar{z}$, the optimal daycare time is at the $\tau_d^*(z) = 1$ corner and the offer of a more preferred program does not induce a change in daycare attendance.

To complete the characterization of the application set note that the model predicts larger sets for more affluent households,

$$\frac{dz}{dw} = -\frac{1 - \phi h_{-1} + \alpha q_g}{\alpha q'_d(z) - k'(z)} < 0, \quad (12)$$

and the data confirm this prediction. Regressing the number of preferences on the FAI as well as on grade and year fixed effects in the Basket 4 universe, the estimated coefficient on FAI is 0.015 (robust s.e. 0.003).

³⁶These estimates are obtained by regressing days in daycare on a dummy for whether the child qualifies for the preferred program and a second degree polynomial in the FAI, as well as grade and year fixed effects.

The key prediction of the model concerns the response of child IQ at the optimum,

$$\theta^* = q_g w \tau_d^* (1 - \tau_d^*) + q_d(z) \tau_d^*, \quad (13)$$

to the offer of a more preferred daycare program. Differentiating with respect to z we obtain

$$\frac{d\theta^*}{dz} = \frac{(w - k(z) - \phi y_{-1})k'(z) + \alpha^2(q_g w + q_d(z))q'_d(z)}{2\alpha^2 q_g w}, \quad (14)$$

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

Remark 2 *If $k'(z) < 0$ and $q'_d(z) \geq 0$ but sufficiently small,³⁷ then there exists a value \tilde{w} of the parental earning potential w such that*

$$\frac{d\theta^*}{dz} < 0 \quad \Leftrightarrow \quad w > \tilde{w}, \quad (15)$$

where

$$\tilde{w} = \frac{k(z)k'(z) - \alpha^2 q'_d(z)q_d(z)}{(1 - \phi h_{-1})k'(z) + \alpha^2 q'_d(z)q_g} > 0. \quad (16)$$

That is, the optimal IQ response to the offer of a more preferred program is positive in less affluent households and negative in more affluent ones.

The conditions on transportation costs and daycare quality under which Remark 2 holds are satisfied, on average, in our setting given the evidence in Table 1. To see why, under these conditions, the child IQ response to the offer of a more preferred program is governed by Eq. 15, note that when the parent is offered a more preferred program she increases daycare time (Remark 1). If she is sufficiently affluent ($w > \tilde{w}$), this generates an IQ loss for the child 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 IQ loss in terms of utility. To put it another way, the offer of a more preferred program allows the affluent parent to increase utility by trading a sufficiently large increase in household consumption for less child IQ. For a less affluent parent ($w < \tilde{w}$), in addition to an analogous increase in consumption, there is a gain in

³⁷Specifically, $q'_d(z) < \frac{-(1 - \phi h_{-1})k'(z)}{\alpha^2 q_g}$. If $q'_d(z) \geq 0$ and sufficiently large, then the IQ response is positive at all levels of affluence.

terms of child IQ because in this case the quality of home care is not sufficiently high with respect to the quality of daycare. If, instead, the offer of a more preferred program exposed the child to a longer daycare time of sufficiently better quality, it does not come as a surprise that the IQ response would be positive for all children.

6.1.2 The ITT-RD estimand at a specific Preferred FAI threshold

Having analyzed the outcomes deriving from the optimal choices of one household, we now study the ITT-RD estimand that can be identified around *one specific* Preferred FAI threshold. We therefore concentrate on a subset of parents who all rank the same specific program as their most preferred one ($z = 1$), omitting the household index to simplify notation.

Qualification for this program depends on whether a household's FAI is just to the left (superscript l) or just to the right (superscript r) of a cutoff y_{-1}^P . Parents in the right neighborhood of this cutoff (i.e., the slightly less affluent in the conventional ordering of our figures) qualify for this program and have (approximately) a FAI equal to $y_{-1}^{P,r}$. For those in the left neighborhood, the FAI approximates instead $y_{-1}^{P,l}$ and they qualify for programs that they prefer less ($z < 1$).

Using $H = \{\alpha, q_g, A, k(\cdot), q_d(\cdot)\}$ to denote the set of household and application set characteristics, let $\theta^* \equiv \theta^*(z, y_{-1}, H)$ denote potential child IQ as a function of its determinants (Eq. 13). If this potential outcome is continuous around the y_{-1}^P cutoff for given z , we can write that:

$$\mathbb{E}[\ln \theta^* | z = 1, y = y_{-1}^{P,l}] \equiv \bar{\vartheta}^*(z = 1, y_{-1}^{P,l}) = \bar{\vartheta}^*(z = 1, y_{-1}^{P,r}) \equiv \mathbb{E}[\ln \theta^* | z = 1, y = y_{-1}^{P,r}], \quad (17)$$

when the child is offered her most preferred program. Note that in Eq. 17 only the expectation on the right is observable. Similarly, we can write that:

$$\mathbb{E}[\ln \theta^* | z < 1, y = y_{-1}^{P,l}] \equiv \bar{\vartheta}^*(z < 1, y_{-1}^{P,l}) = \bar{\vartheta}^*(z < 1, y_{-1}^{P,r}) \equiv \mathbb{E}[\ln \theta^* | z < 1, y = y_{-1}^{P,r}], \quad (18)$$

when the child is not offered her most preferred program, and in this case only the expectation on the left is observable. The evidence on the continuity of covariates provided in Section 4 for the Basket 4 universe supports the assumption of continuity of potential outcomes (at least as far as observables in H are concerned). The analogous evidence for the interview

sample will be described in [Section 7](#).

Letting the size of the right and left neighborhoods approach zero so that $y_{-1}^{P,l} = y_{-1}^{P,r} = y_{-1}^P$, the second of these two continuity conditions, [Eq. 18](#), is sufficient to identify the response of child IQ for children who are offered their most preferred program with respect to the counterfactual situation in which they are not:

$$\begin{aligned}\beta(y_{-1}^P) &= \bar{\vartheta}^*(z = 1, y_{-1}^{P,r}) - \bar{\vartheta}^*(z < 1, y_{-1}^{P,r}) \\ &= \bar{\vartheta}^*(z = 1, y_{-1}^{P,r}) - \bar{\vartheta}^*(z < 1, y_{-1}^{P,l}) \\ &= \bar{\vartheta}^*(z = 1, y_{-1}^P) - \bar{\vartheta}^*(z < 1, y_{-1}^P),\end{aligned}\tag{19}$$

where the third line is the ITT-RD estimand containing quantities that are observable.

Note that if [Remark 2](#) holds, then

$$\beta(y_{-1}^P) < 0 \quad \Leftrightarrow \quad y_{-1}^P > \tilde{y} \equiv \tilde{w}h_{-1},\tag{20}$$

that is, the estimand of the IQ effect of qualifying for the most preferred program at a specific threshold y_{-1}^P is negative if this threshold is sufficiently high, which implies that the households around this threshold are relatively more affluent.

6.1.3 The ITT-RD estimand pooling all thresholds

Our setting, however, is characterized by multiple cutoffs like y_{-1}^P , one for each program that has been listed as the most preferred by a group of households.³⁸ At each one of these cutoffs we compare children who just qualified for the corresponding program with children who barely did not. We then aggregate the estimates for the different cutoffs integrating across them to obtain

$$\beta = \mathbb{E}_{F_{z=1}^P}[\bar{\vartheta}^*(z = 1, y_{-1}^P)] - \mathbb{E}_{F_{z<1}^P}[\bar{\vartheta}^*(z < 1, y_{-1}^P)],\tag{21}$$

where $F_{z=1}^P \equiv F^P(y_{-1}^P | z = 1)$ is the frequency distribution of the cutoffs over the population of households located just to the right of them, while $F_{z<1}^P \equiv F^P(y_{-1}^P | z < 1)$ is the corresponding distribution for the households located just to the left.

³⁸ See [Cattaneo *et al.* \(2016\)](#) for a more general discussion of RD designs with multiple cutoffs.

In the presence of multiple cutoffs, identification requires that

$$F^P(y_{-1}^P|z = 1) = F^P(y_{-1}^P|z < 1) = F^P(y_{-1}^P), \quad (22)$$

where $F^P(y_{-1}^P)$ denotes the distribution of households attached to the different cutoffs. To test this condition, 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 hypothesis that the observations immediately at the right and at the left of each Preferred FAI threshold come from the same distribution cannot be rejected, both in the Basket 4 universe (p-value: 0.21) and in the interview sample (p-value: 0.41).³⁹

Therefore, the ITT-RD estimand that we can identify around Preferred thresholds is

$$\begin{aligned} \beta &= \mathbb{E}_{F^P}[\bar{\vartheta}^*(z = 1, y_{-1}^P)] - \mathbb{E}_{F^P}[\bar{\vartheta}^*(z < 1, y_{-1}^P)] \\ &= \mathbb{E}_{F^P}[\beta(y_{-1}^P)], \end{aligned} \quad (23)$$

which may be positive or negative depending on whether the frequency of households attached to each cutoff y_{-1}^P is skewed towards lower or higher values. For instance, it is negative if the cutoffs at which $\beta(y_{-1}^P) < 0$ have a sufficiently large weight in Eq. 23, which is the case of our estimates described in Section 7.

This interpretation of the estimand sign crucially hinges on the validity of the conditions defined by Eq. 22, which may fail in a small sample. Suppose for example that at the Preferred cutoffs with a high FAI, households were disproportionately frequent on the left (where they would not be offered their preferred program), while at the Preferred cutoffs with a low FAI, households were disproportionately frequent on the right (where they would be offered their preferred program). In this case, the distribution $F_{z=1}^P$ would first-order stochastically dominate $F_{z<1}^P$ (i.e., $F_{z=1}^P \leq F_{z<1}^P$). Since income and IQ are positively correlated, the ITT-RD estimand β could be negative even if the IQ effect of qualifying for the preferred program were positive for all households. However, this possibility is ruled out in our case by the evidence that the condition in Eq. 22 holds not only in the Basket 4 universe but also in the smaller interview sample.⁴⁰

³⁹The “right and left neighborhoods” are approximated by bins of size €2k for the Basket 4 universe and €4k in the interview sample. Figure 4 in the Online Appendix plots the two pairs of CDF’s.

⁴⁰We thank an anonymous referee for raising this point. The evidence displayed in Figure 4 of the Online

The estimand β in Equation 23 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 affluent parents would have access to their preferred program, with negative effects on the IQ of their children that may not be socially optimal even if the utility of these parents increases. Moreover, in Section 3 we have shown that parents value the proximity of programs and, to a lesser extent, their quality. A public investment in daycare expansion would allow parents to receive offers from programs that they prefer more. Our estimates speak precisely about the effect of such a policy, which may have undesirable consequence on the cognitive skills of more affluent children.

6.1.4 Heterogeneity of effects across gender

The psychological literature mentioned in the introduction, to be discussed in greater detail in Section 8, suggests that gender differences in the effect of daycare time may be expected if girls are better equipped than boys at exploiting one-to-one interactions with adults for the development of their IQ. To introduce this possibility in the basic model, we allow the technology of IQ formation to differ across genders in the following way:

$$\theta = (1 + \lambda(f))(q_g y \tau_g + q_d(z) \tau_d) + \chi(f), \quad (24)$$

where $f = 1$ if the child is a female while $f = 0$ otherwise, $\lambda(1) > \lambda(0) = 0$, and $\chi(f)$ is an unrestricted IQ shifter.⁴¹ We also assume that parents make daycare decisions based on a belief $\lambda_b(1) \geq 0$ about $\lambda(1)$.

Using Eq. 13, the gender gap in the IQ effect of a variation in daycare time induced by the offer of a more preferred program can be written as

$$\left. \frac{d\theta^*}{d\tau_d^*} \right|_{f=1} - \left. \frac{d\theta^*}{d\tau_d^*} \right|_{f=0} = -2q_g w (\tau_d^*|_{f=1} - \tau_d^*|_{f=0}) + \frac{\lambda(1)}{1 + \lambda(1)} \left(\left. \frac{d\theta^*}{dz} / \frac{d\tau_d^*}{dz} \right|_{f=1} - q'_d(z) \frac{\tau_d^*}{d\tau_d^*/dz} \right) \Big|_{f=0}. \quad (25)$$

This gender gap has three components. The sign of the first one depends on the gender difference in the optimal daycare time chosen by a parent, which, using the interior solution

Appendix indicates that, if anything, we are in the opposite case.

⁴¹The shifter is unrestricted because the existence and sign of gender differences in IQ is controversial and our analysis is not affected by this issue.

for τ_d^* in Eq. 7 and the parental belief about gender differences, is

$$\tau_d^*|_{f=1} - \tau_d^*|_{f=0} = \frac{-\lambda_b(1)}{1 + \lambda_b(1)} \frac{(w - k(z) - \phi y_{-1})}{2\alpha q_g w} \leq 0, \quad (26)$$

because $w - k(z) - \phi y_{-1} > 0$ at the interior solution for consumption. That is, the parent chooses a weakly shorter daycare attendance for girls than for boys. This happens because if $\lambda_b(1) > 0$ and the child is a girl, the marginal unit of parental time is more valuable at producing child IQ than at consumption, therefore labor supply decreases, home care time increases, and daycare time decreases, relative to the case in which the child is a boy. The sign of the second component, instead, depends on the sign of the IQ effect for a girl, $\frac{d\theta^*}{dz}|_{f=1}$, given that $\frac{d\tau_d^*}{dz}|_{f=1} > 0$ because of Remark 1. The sign of the third component is non-positive if daycare quality does not decrease with the ranking.

A particularly relevant case that is supported by the data is that parents perceive no gender difference in the technology of IQ formation, i.e., $\lambda_b(1) = 0$, even if $\lambda(1) > 0$. In this case, the optimal levels of daycare time do not differ between boys and girls ($\tau_d^*|_{f=1} = \tau_d^*|_{f=0}$), nor do parents' responses to the offer of a more preferred program ($\frac{d\tau_d^*}{dz}|_{f=1} = \frac{d\tau_d^*}{dz}|_{f=0}$).⁴² Therefore, the gender gap in the IQ effects of being offered a program with a higher z reduces to

$$\frac{d\theta^*}{d\tau_d^*}\Big|_{f=1} - \frac{d\theta^*}{d\tau_d^*}\Big|_{f=0} = \frac{\lambda(1)}{1 + \lambda(1)} \left(\frac{d\theta^*}{dz}\Big|_{f=1} / \frac{d\tau_d^*}{dz} \right) - q'_d(z) \frac{\tau_d^*}{d\tau_d^*/dz}\Big|_{f=0}. \quad (27)$$

The sign of this expression depends on the sign of $\frac{d\theta^*}{dz}|_{f=1}$, which is negative for affluent households if Remark 2 holds. This implies a larger IQ loss for a girl than for a boy in an affluent population. The findings described in Section 8 are precisely consistent with this prediction and thus with the hypothesis that $\lambda_b(1) = 0$ and $\lambda(1) > 0$: the offer of the most preferred program induces the same increase of daycare time for both genders, but in affluent families the IQ loss is larger for girls than for boys.

⁴² This because

$$\frac{d\tau_d^*}{dz}\Big|_{f=1} - \frac{d\tau_d^*}{dz}\Big|_{f=0} = \frac{k'(z) - \frac{k'(z)}{(1+\lambda_b(1))}}{2\alpha q_g w},$$

which is zero if $\lambda_b(1) = 0$. The intuition is similar to the one for levels: the offer of a more preferred program weakly increases daycare quality, thereby making the marginal unit of daycare time more valuable to the parents of girls than to the parents of boys, provided they are aware of gender differences in the technology of IQ formation.

6.2 A more general model with specific features of the BDS

We now relax several major restrictions of the baseline model, namely absence of: curvature in the utility and IQ production functions; non-parental child care alternatives to daycare; other sources of subsistence consumption in case of no labor income; two key features of the BDS, i.e., the nonlinear regressive fee and the possibility that the Maximum and Preferred thresholds coincide, in which case not qualifying for the preferred program means not qualifying for any program.

6.2.1 Setup

Let parents' preferences be represented by $\ln c + \alpha \ln \theta$, and the IQ production function by $\theta = \tau(y)^\eta + \underline{\theta}$, where τ is a nonlinear aggregator (to be specified below) of time spent in 1:1 interaction with an adult in alternative child care modes (weighted by the quality of the interaction, which depends on household affluence), $\eta > 0$, and $\underline{\theta}$ is a constant minimum IQ level. Now a parent can also acquire child care time from the market, τ_m , at price π_m per unit of time. Although for brevity we refer to τ_m as “market” child care, we include in this category both extended family caregivers (e.g., grandparents and other relatives, whose services have some cost as well) and market services strictly defined (e.g., babysitters, nannies, and private daycare).⁴³ Assume that there are two types of daycare programs: the most preferred, labeled P (program $z = 1$ in the baseline model), and the less preferred, labeled L ($z < 1$ in the baseline model). As before, the price of daycare reflects a transportation cost, k , and an income-based fee, $\phi(y_{-1})$, which is now nonlinear, so that $\pi_d^j = k^j + \phi(y_{-1})$, $j = \{P, L\}$. We assume $\pi_d^P \leq \pi_d^L$ because of the weakly lower transportation cost associated with the preferred program (see Table 1).

Daycare is rationed, and offers are made based on eligibility cutoffs relative to past income, y_{-1} . Using y^P and y^L to denote the thresholds for admission to programs P and L , consider a neighborhood of y^P and define $y^M \equiv \max\{y^L, y^P\}$. If $y_{-1} \leq y^P$, the ordering of y^L and y^P is irrelevant and the child is offered P . If $y_{-1} > y^P$, instead, the outcome depends on this ordering. Let $\mu(y_{-1})$ denote the probability that $y^M = y^L \geq y^P$ for a household with

⁴³We assume for simplicity that π_m is an average price not changing with the composition of τ_m . See below for the details on the calibration of this parameter.

past income y_{-1} . In this case the child is offered L . If $y^M = y^P \geq y^L$, which occurs with probability $1 - \mu(y_{-1})$, then the child does not qualify for any daycare program. This case is labeled N . Once an outcome in $\{P, L, N\}$ is determined, qualified households choose their optimal daycare time τ_d . For not qualified households, $\tau_d = 0$.

Parental, market, and daycare time are aggregated into a single input by a CES function,

$$\tau = (q_g(y)\tau_g^\rho + q_m(y)\tau_m^\rho + \mathbb{I}[y_{-1} \leq y^M]q_d^j\tau_d^\rho)^{\frac{1}{\rho}}, \quad j = \{P, L\}, \quad (28)$$

where $q_g(y)$ and $q_m(y)$ – the quality of parental and market care – are increasing functions of household income. This formulation captures the idea that market child care, being chosen by parents, is complemented by the same resources used in parental care.

Like in the baseline model, the parent chooses working time h , consumption c , and the child care arrangement (τ_g, τ_m, τ_d) so to maximize utility, subject to the technology of IQ formation, the budget constraint, $c + \pi_m(1 - \tau_g - \tau_d) + \pi_d^j\tau_d = wh + b\mathbb{I}[h = 0]$, where b represents a capped non-employment benefit in case of no labor income, the time constraint, $h + \tau_g = 1$, the child care requirement constraint, $\tau_g + \tau_m + \tau_d = 1$, and the daycare availability constraint. The model has solutions that can be grouped into three relevant cases for the theoretical interpretation of our RD estimand: the household is offered the preferred program (case P, associated with an IQ of θ^P), the less preferred program (case L, θ^L), or no daycare (case N, θ^N).

In this setting, the percentage change in child IQ induced by the offer of the preferred daycare program, at each level of y , is approximated by

$$\Delta \ln \theta(y) = \ln \theta^P(y) - \mu(y_{-1}) \ln \theta^L(y) - (1 - \mu(y_{-1})) \ln \theta^N(y), \quad (29)$$

and the ITT-RD estimand around Preferred thresholds is, under the same continuity conditions discussed above,

$$\beta = \mathbb{E}_{FP} [(\bar{\vartheta}^P(y_{-1}^P) - \mu(y_{-1})\bar{\vartheta}^L(y_{-1}^P) - (1 - \mu(y_{-1}))\bar{\vartheta}^N(y_{-1}^P))], \quad (30)$$

where $\bar{\vartheta}^P$, $\bar{\vartheta}^L$, and $\bar{\vartheta}^N$ are the population averages of the logs of θ^P , θ^L , and θ^N in a neighborhood of a Preferred threshold.

6.2.2 Calibration

We solve the model numerically after calibrating the parameters as follows. For preferences, we set $\alpha = 0.25$, a value taken from estimates for Italy of the degree of intergenerational altruism provided by [Bellettini, Taddei, and Zanella \(2017\)](#). As for the IQ production function, we set $\eta = 0.9$ and $\rho = 0.48$. These values are chosen to illustrate that it is possible to observe a positive average IQ effect of qualifying for the preferred program in the universe of applicants to the BDS and, at the same time, a negative effect in the sample of more affluent dual-earner households that is the focus of our analysis. This same logic guides our choice of $\underline{\theta}$, which is set to reflect the IQ level (expressed in model units) of the child from the poorest model household who is offered the less preferred program (0.6). The $q_g(y)$ and $q_m(y)$ functions are assumed to be logistic and such that, for each parent, the quality of market daycare is 90% the quality of own parental care. Specifically, we set $q_g(y) = (1 + 15 \exp(-2y - 0.5))^{-1}$, so that maximum parental quality is 1, and $q_m(y) = 0.9q_g(y)$.

Turning to institutional parameters, the probability that Preferred and Maximum thresholds coincide, $1 - \mu(y_{-1})$, is predicted for the Basket 4 universe by a logistic regression as a function of the FAI and its square. The estimated probability is increasing in the FAI (indicating that Maximum and Preferred thresholds are more likely to coincide at higher levels of affluence, as one should expect), ranging from 0.04 at a FAI of 2k, to 0.58 at a FAI of 70k. Similarly, we input into the model the actual daycare fee schedule $\phi(y_{-1})$ described in footnote 22.

The transportation cost component of the daycare price is assumed to be zero for the most preferred program, which on average is the one closest to home (see [Table 1](#)). For the less preferred program, we assume that it takes 30 extra minutes to reach the facility,⁴⁴ and the value of this time is set equal to 1/16 (i.e., half an hour in a 8-hour working day) the wage of the provider of market daycare. The price of market daycare services, in turn, is calibrated to match the average annual wage of a babysitter in the city of Bologna, as calculated from [jobpricing.it](#). This average is €20k per year, or about 37% the average household income

⁴⁴As shown in [Table 1](#), the difference in the distance from home between the the most preferred program and the average of the ranked less preferred programs is about 750 meters, which, according to Google Maps, in Bologna can be covered by an adult in approximately 8 minutes, so that 30 minutes is about the total time for delivery and pick-up of the child.

among the universe of applicants to the BDS in our data, which is about €54k (both values are expressed in constant 2010 euros). Therefore, because in the model average household income is normalized to 1, we set $\pi_m = 0.37$. The non-employment benefit b is instead set at 0.1 of the average income, reflecting the prevailing levels in Italy at the time of the analysis.⁴⁵

Finally, the quality of daycare is calibrated to reflect the difference in one-to-one interactions between daycare and parental care in a household with average income. Based on our calibration of $q_g(y)$, the former is about 0.45. Assuming that the BDS complements interactions in daycare with the same resources as the average household, then moving from an adult to child ratio of 1:1 at home to an adult to child ratio of 1:4/1:6 in daycare should reduce by 4/5 child care quality with respect to the average household. Therefore, based on the evidence reported earlier in the paper that the preferred facilities are, on average, of slightly better quality than less preferred ones, we set $q_d^P = 0.11$ and $q_d^L = 0.08$.

The results of the numerical solution are plotted in Figures 7 and 8. The first three panels of Figure 7 plot the optimal child care arrangement chosen by the parent when the child is offered the preferred program, the less preferred one, and no program, respectively, as a function of the FAI. These panels exhibit the following patterns. First, conditional on being offered admission, more affluent households use less daycare, because of the higher quality of the two home-based care modes (the daycare lines in the top two panels of Figure 7 are downward sloping). This prediction can be tested and is confirmed by our data: regressing the number of days spent in daycare on FAI as well as on grade and year fixed effects in the group of 5,897 children in Basket 4 who were offered admission at their first application, the estimated coefficient on FAI is -0.81 (robust s.e. 0.10).⁴⁶

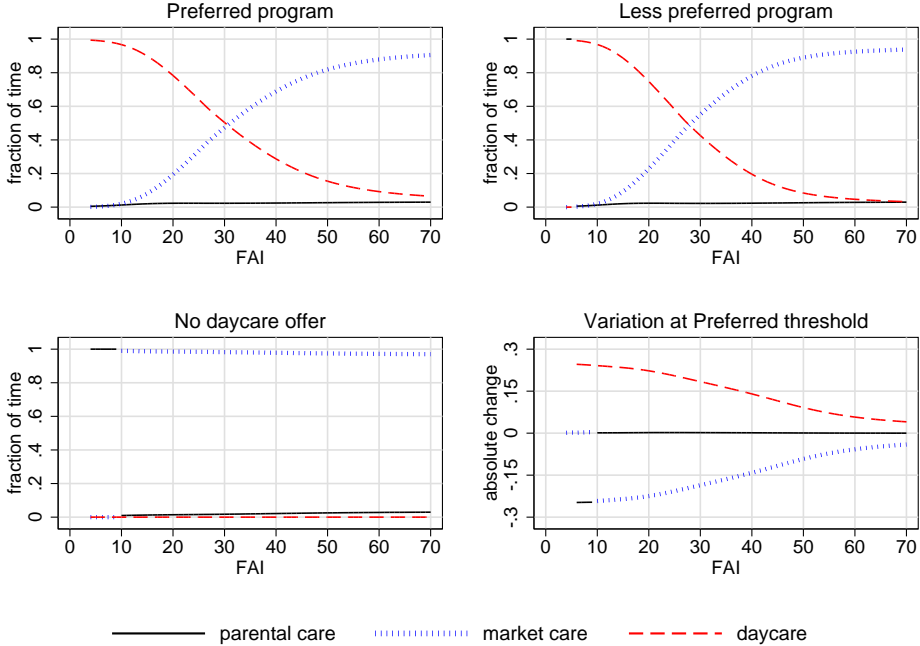
Second, a comparison of the vertical height of the daycare lines between the top two panels shows that parents use more daycare at any level of the FAI when offered the preferred program, because of the lower transportation cost and the weakly higher quality. For the universe of children in Basket 4, this was shown in the left panel of Figure 4; for the interview sample, the corresponding evidence is in Figure 10. How this variation changes at different

⁴⁵See the “Decreto Legislativo” n. 151 of 26/03/2000.

⁴⁶The remaining 678 children to reach the total of 6,575 in Basket 4 were not offered admission at their first application because they were relatively more affluent. If we include them in the sample for this test, they mechanically induce a negative relation between the FAI and days of attendance. Indeed, when they are included, the estimated coefficient on FAI is -1.43 (robust s.e. 0.13).

levels of the FAI is shown by the daycare line in the fourth panel, which describes the change in the optimal child care arrangement when the child crosses the threshold for the preferred program at each level of affluence. As in the baseline model and in the data (see Remark 1), we see that the change in optimal daycare time is positive but smaller at higher levels of the FAI. We also see in this panel that the offer of the preferred program allows the sufficiently affluent household to economize on market care (the market line indicates negative changes after a FAI level of about €9k, corresponding to a gross annual family income of approximately €24k). This reduction is smaller for households that are progressively above the €9k level because they can access a market care of increasingly higher quality.

Figure 7: Child care arrangement and its variation at the Preferred threshold, by FAI



Notes: The figure shows the child care arrangement optimally chosen by the parent when the child is offered the preferred program (top-left), the less preferred program (top-right), and no program (bottom-left), as well its variation at the preferred threshold (i.e., when the child is offered the preferred program, bottom-right) as a function of the FAI. The data are generated by a numerical solution of the calibrated model.

At low levels of the FAI, below €9k, the patterns are influenced by the fact that the cost of market care exceeds the earning potential of the parent, who therefore spends all her time with the child in case of no daycare offer (bottom left panel). As a results, in this range of FAI levels, qualification for the preferred program induces no change of market care

usage and a decrease of time spent by parents with their children (bottom right panel). At the €9k FAI level we observe a discontinuity in the behaviour of parents: above this level of affluence the parent is always employed, parental care does not change with qualification for the most preferred program, and the parent just substitutes market care with daycare. Another discontinuity is observed at a FAI of €5k (approximately €13k of annual family income), a level below which a parent who is offered the less preferred program prefers to turn down the offer, provide full-time parental care, and live off the unemployment benefit (top-right panel). Below this level, the parent is at the same corner solution both in the L and in the N cases, and so the offer of the preferred program induces a downward jump of nearly 100 percentage points in the fraction of time the child is in parental care, fully substituted by an increase in daycare time.⁴⁷

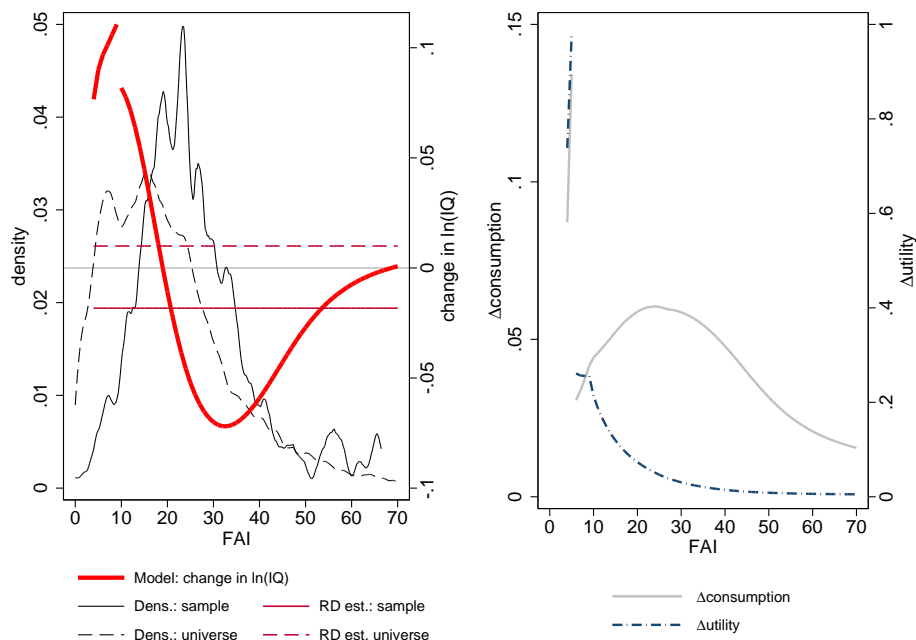
The percentage variation in child IQ when the child is offered the preferred program (Eq. 29) is given by the thick line shown in the left panel of Figure 8. For each level of the FAI, this is the effect for a child with that FAI and whose preferred program has a hypothetical threshold exactly equal to that same FAI. Like in the baseline model, there exists a FAI level such that the effect is positive for less affluent households and negative for more affluent ones. Our calibration implies that this sign reversal occurs at a FAI of about €18k, roughly equivalent to a gross annual family income of €48k. We also see in this figure that at very high levels of affluence the negative IQ effect decreases in absolute size after reaching a minimum at a FAI of about €33k (gross annual family income of about €88k). The reason is that very affluent parents are relatively less inclined to increase daycare time following the offer of the preferred program (fourth panel of Figure 7). As a consequence, the negative ITT-RD estimand approaches zero at very high levels of the FAI.

At very low levels instead (below the €9k FAI level), qualification for the preferred program allows the parent to move from non-employment to work and thus to increase resources that complement the infra-marginal home care time in the production of IQ. This increase in resources is larger at higher levels of earning potential and this explains why the thick line is upward sloping in this range, up to a discontinuity point which corresponds to

⁴⁷These extreme changes are omitted from the bottom-right panel to preserve a readable scale of the graph.

the one observed in the bottom-left panel of Figure 7. In the range between the €9k and the €33k FAI levels, the thick line is downward sloping because the increase in resources for infra-marginal home care time triggered by the offer of the preferred program does not compensate the effect of decreasing parental time of progressively higher quality.

Figure 8: Variation of IQ, consumption, and utility at the Preferred threshold, by FAI



Notes: The thick line in the left panel shows the change in $\ln(\text{IQ})$ of the child at the preferred threshold (i.e., when the child is offered the preferred program) as a function of the FAI. This is generated by a numerical solution of the calibrated model. Superimposed on this figure are the empirical densities of the FAI in the interview sample (dens.: sample) and in the universe of applicants to the BDS across all baskets (dens.: universe), obtained via kernel density estimation with a triangular kernel and a bandwidth of 5. Applying these empirical weights to the change in $\ln(\text{IQ})$ produced by the model yields the two horizontal lines, which represent the ITT-RD estimands of the IQ effect of qualifying for the preferred daycare program in the interview sample of Basket 4 (RD est.: sample) and in the universe of applicants to the BDS across all baskets (RD est.: universe). The right panel shows the variation in household consumption and parental utility at the preferred threshold as a function of the FAI, as generated by the numerical solution.

Superimposed on this figure are the empirical densities of the FAI in the interview sample and in the universe of applicants to the BDS. By integrating the changes in $\ln(\text{IQ})$ generated by the model with respect to these distributions, it is possible to obtain quantitative predictions of the RD effect of interest in these two samples.⁴⁸ The result is given by the two

⁴⁸This exercise is in the spirit of Bertanha (2016), who suggests an estimation procedure to extrapolate from the average treatment effect on the observed distribution of subjects at the available cutoffs, to a more general average effect based on the entire distribution of subjects. This procedure cannot be applied in our case, due to the small sample size, but we aim for a similar goal with the calibration described here.

horizontal lines in the left panel of [Figure 8](#), which represent ITT-RD estimands of the IQ effect of being offered the preferred program. In our sample, which is shifted towards higher levels of the FAI, the model predicts an average negative effect of about -1.8% . However, the model also predicts a positive average effect of about $+1.0\%$ in the universe of applicants to the BDS, where the incidence of less affluent households is higher. The right panel of [Figure 8](#) shows the variation in household consumption and parental utility following the offer of the preferred program. These changes are always positive.

6.3 Dynamic model

The parental decision to send a child to daycare has intertemporal dimensions that are relevant for the interpretation of our estimates. First, as suggested by [Cunha, Heckman, and Schennach \(2010\)](#) there is evidence of dynamic complementarities in cognitive skill formation: an early parental investment in the production of these skills increases the return to later investment. Second, the psychological literature (see [Section 8](#)) indicates that parental time with children is relatively more crucial for IQ formation when they are very young, while at older ages interactions with other adults and with peers acquire more relevance. Third, there is evidence (see, for instance, [Lalive and Zweimüller, 2009](#)) that delaying the return to work after the birth of a child is costly for a parent in terms of future wages and career prospects. A longer delay would not only reduce household consumption, but also family resources that could be later devoted to complement parental interactions with children for a more effective investment in their IQ. Therefore, a household faces a dynamic trade-off, which is illustrated below keeping only the relevant features of the baseline model.

We assume that the first three years of life of a child (period “age 0–2”), can be divided in two sub-periods, $t \in \{0, 1\}$. The parent decides whether to apply for daycare in sub-period 0 and then again in sub-period 1. Denoting with s_t an indicator taking value 1 if an application is filed in sub-period t , there are four possible combinations defined by $\{s_0, s_1\}$. A parent does not apply for daycare in a sub-period when, even if the child is admitted to her preferred program, her utility from daycare attendance is lower than the utility of staying at home with the child. Therefore, to analyze the participation decision we focus on the preferred program only, $z = 1$, which is assumed to have a quality $q_d(1) = q_d$ and a cost

of attendance $\pi_d = k(1) = k$. Note that this cost of attendance does not depend on family affluence (i.e., $\phi = 0$) and relates only to the distance of the preferred program from home. This assumption simplifies the analysis at no loss of generality and is in line with the low cap on attendance fees that effectively characterizes the BDS (see footnote 22).

Daycare attendance is treated as a discrete choice in each sub-period: $\tau_{dt} \in \{0, 1\}$. That is, we abstract from the within-sub-period decision concerning days of attendance, and focus on the intertemporal variation across sub-periods, which goes from a minimum of 0 in the combination $\{0, 0\}$ to a maximum of 2 in the combination $\{1, 1\}$. The problem faced by the parent is, therefore:

$$\max_{c, \tau_{d0}, \tau_{d1}} c + \alpha\theta \text{ s.t. } \begin{cases} c = (w - k)(\tau_{d0} + \tau_{d1}) + \gamma\tau_{d0}\tau_{d1} \\ \theta_0 = q_{g0}(1 - \tau_{d0}) + q_d\tau_{d0} \\ \theta_1 = q_{g1}(1 - \tau_{d1}) + q_d\tau_{d1} \\ \theta = \theta_0 + \theta_1 + \theta_0\theta_1 + w(\tau_{d0} + \tau_{d1}) + \gamma\tau_{d0}\tau_{d1} \\ \tau_{d0} \in \{0, 1\} \\ \tau_{d1} \in \{0, 1\} \end{cases} \quad (31)$$

where we set $q_{g0} > q_{g1}$ to reflect the assumption that the quality of parental time with a child is higher in the first sub-period. The term γ captures instead the wage premium for labor market attachment, which gives more resources for both consumption and IQ formation in addition to baseline earnings $w(\tau_{d0} + \tau_{d1})$.

Utility at the optimum, V_{s_0, s_1} , derived by the parent in the four possible combinations is:

$$\begin{aligned} V_{0,0} &= \alpha(q_{g0} + q_{g1} + q_{g0}q_{g1}), \\ V_{0,1} &= w - k + \alpha(q_{g0} + q_d + q_{g0}q_d + w), \\ V_{1,0} &= w - k + \alpha(q_d + q_{g1} + q_dq_{g1} + w), \\ V_{1,1} &= 2(w - k) + \gamma + \alpha(q_d + q_d + q_d^2 + 2w + \gamma). \end{aligned}$$

A comparison of these values reveals that the decisions about whether and when to apply depend on household affluence in the way summarized by the following remark.

Remark 3 *Under the assumption that the quality of parental care is sufficiently higher in*

sub-period 0 than in sub-period 1,⁴⁹ less affluent families are more likely to delay daycare application or to not apply at all. More precisely, let T_{0100} be the affluence level at which the parent switches from $\{s_0, s_1\} = \{0, 0\}$ to $\{s_0, s_1\} = \{0, 1\}$, and similarly for T_{1101} . These values are:

$$T_{0100} \equiv \frac{k + \alpha(q_{g1} - q_d)(1 + q_{g0})}{1 + \alpha} \quad (33)$$

$$T_{1101} \equiv \frac{k - \gamma(1 + \alpha) + \alpha(q_{g0} - q_d)(1 + q_d)}{1 + \alpha}. \quad (34)$$

If

$$w < T_{0100} \quad (35)$$

the parent never applies for daycare. If

$$T_{0100} < w < T_{1101} \quad (36)$$

the parent stays with the child in sub-period 0 and applies for daycare only in sub-period 1.

If

$$T_{1101} < w \quad (37)$$

the parent applies in both periods.

We cannot test empirically the predictions of Remark 3 because we do not observe potential applicants who did not apply to the BDS. However, indirect evidence is offered by the comparison of the average FAI of the households who first apply at age 0, which is €24.7k, or at age 1, which is instead €23.8k. Although not statistically significant at conventional levels (p-value: 0.11), this difference indicates that on average the parents who delay by one year after birth their first application are less affluent, while those who first apply immediately after birth tend to be more affluent.⁵⁰ Note that this finding does not contradict Remark

⁴⁹Specifically, it must be that

$$q_{g0} - q_{g1} > \gamma \frac{(1 + \alpha)}{\alpha} + q_d^2 + q_{g0}(q_{g1} - 2q_d). \quad (32)$$

⁵⁰If q_{g0} were not sufficiently higher than q_{g1} (i.e., if condition 32 were not satisfied), we would not be able to rank T_{0100} and T_{1100} and the relationship between affluence and the decision about whether and when

1 or Figure 7: affluent parents prefer to anticipate the application for the reasons discussed here, but this is compatible with a smaller reaction to the offer of a more preferred program or with a shorter daycare attendance conditional on positive attendance.

Given that the continuity conditions defined in Eq. 22 are satisfied in our empirical application, the finding that affluence induces parents to apply as early as possible after birth does not constitute a threat for the identification of the ITT-RD estimand in Eq. 23.⁵¹ This finding, however, is relevant for the interpretation of Remark 2 and thus for the sign of the ITT-RD estimate in the case of relatively more affluent parents. If these parents apply for daycare earlier than the less affluent ones, then the negative IQ effect for the more affluent induced by qualification for the preferred program may reflect early attendance, i.e., the deprivation of valuable home resources when these are most effective.

Under different hypotheses, the three theoretical settings that we have analyzed lead to similar predictions. When offered the most preferred daycare program, as opposed to a less preferred one, relatively affluent parents take advantage of this opportunity to increase daycare attendance of their children and so work more or reduce costly market care. Even if this increase in daycare attendance is smaller than the one occurring in a less affluent household, it generates an increase of family resources that is large enough to become attractive even at the cost child IQ.

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

Let θ_i be the IQ of a child observed at age 8–14 and denote with $\tau_{d,i}$ the treatment intensity, measured as months spent in daycare over the entire 0–2 age period.⁵² The running variable is the FAI, y_i , at first application and the estimated equation is⁵³

to apply for daycare would be more blurred. The indirect evidence reported above suggests this is not a concern in our setting.

⁵¹The reason is that this estimand compares the IQ of children whose parents have the same level of affluence and who differ only by whether they are offered their preferred program or not.

⁵²In the administrative data at our disposal we observe the precise daily attendance of children in daycare. For convenience, we rescale days of attendance in months defined as 20 working days.

⁵³In this parametric specification we do not center and stack thresholds, different from what we do in the continuity figures, thus avoiding the problems generated by observations located precisely at the thresholds. For a discussion of the perils of stacking thresholds in RD designs see Fort *et al.* (2017).

$$\theta_i = \alpha + \beta\tau_{d,i} + g(y_i) + \gamma A_i + \delta X_i + \epsilon_i, \quad (38)$$

where β is the empirical counterpart of the theoretical estimand derived in Eq. 23, $g(y_i)$ is a second order polynomial in the running variable, A_i is a vector of variables describing the application set of a child (dummies for the city neighborhood of the preferred program and the number of programs included in the application set), and X_i is a vector of pre-treatment personal and family 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, birthday in the year and a dummy for cesarean delivery of the child). Finally, ϵ_i captures other unobservable covariates.

As usual in RD designs, the inclusion of pre-treatment observables is not strictly necessary for identification but it may increase efficiency and, most important, similar estimates of the treatment effect β when observables are included or not supports 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 by Figure 9 that replicates the analogous evidence offered in Figure 5 for the Basket 4 universe. Thanks to the information acquired from parents in the interviews, here we can assess continuity for the larger set of 11 covariates that is included in Eq. 38. Table 4 presents results of a formal joint test of the continuity of these covariates that confirms the validity of our design.⁵⁴

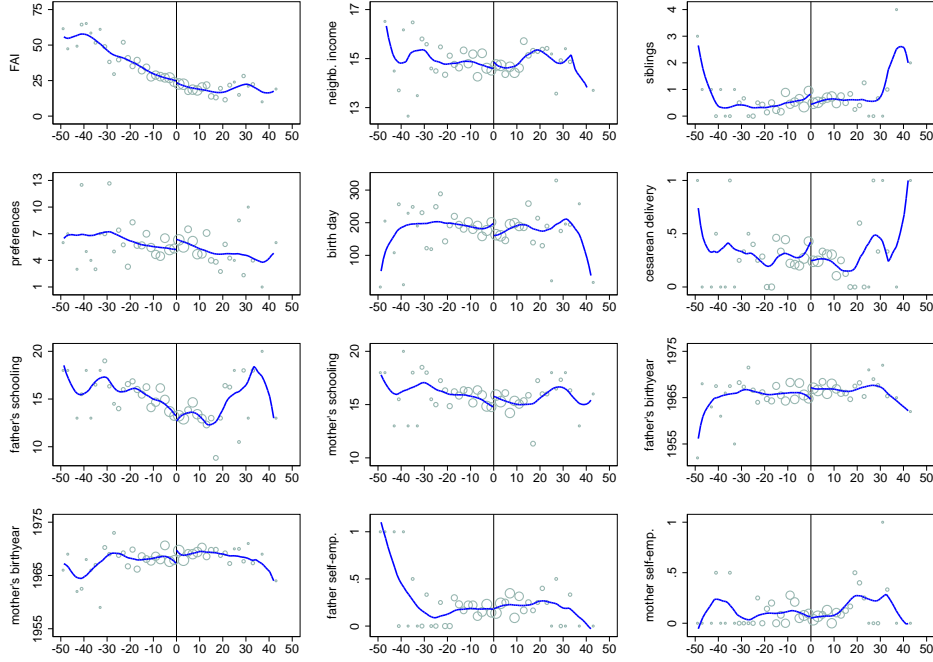
We estimate equation (38) by IV using as an instrument the dummy P_i which indicates whether a child qualifies for her preferred program at her first application or not,

$$P_i = \mathbb{I}(y_i \leq y_i^P). \quad (39)$$

Figure 10 replicates for the interview sample the evidence of Figure 4, which was based on the Basket 4 universe. The admission rate, the attendance rate and days of attendance all jump discontinuously at the preferred thresholds (by 19 percentage points, 10 percentage points and 64.2 days, respectively).

⁵⁴For this test, we follow Abdulkadiroglu *et al.* (2014) and estimate a system of 11 equations (one for each covariate $x_{it} \in A_i$ or $\in X_{it}$ of Eq. 38) via SUR and then we test the joint significance of the instrument defined below by equation Eq. 39 across the system of equations: $x_i = \alpha + \beta P_i + g(y_i) + \epsilon_i$.

Figure 9: Continuity of covariates around Preferred FAI thresholds, interview sample



Notes: The circles represent the average of eight pre-treatment variables 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 LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from [Calonico et al. \(2014b\)](#). 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 to programs with rationing, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in [Fort et al. \(2017\)](#), is different from zero.

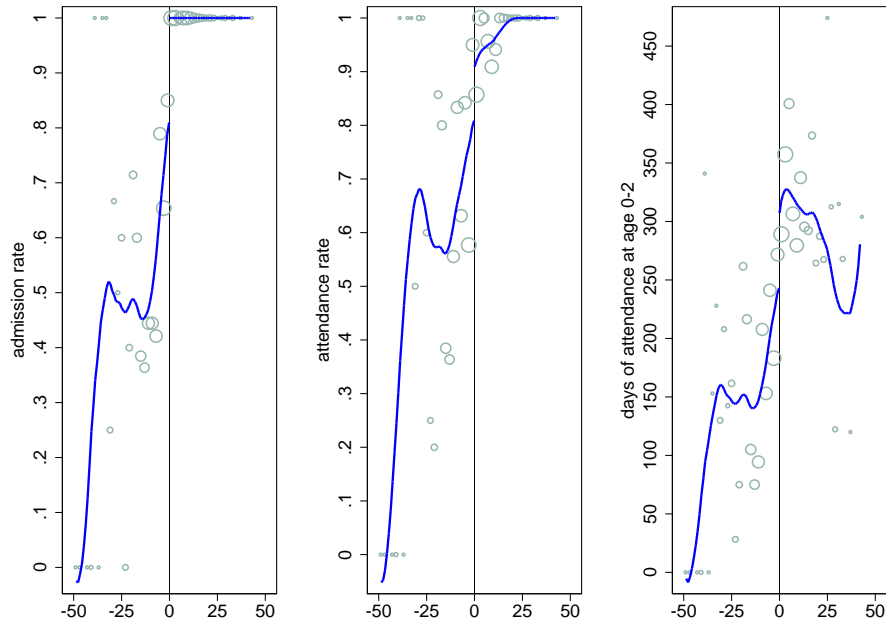
Table 4: Continuity of covariates, regression-based test in the interview sample

var1	var2	var3	var4	var5	var6	var7	var8	var9	var10	var11
0.03	0.06	-0.04	-0.17	-0.02	-0.11	0.06	-0.01	0.12	0.05	-0.01
(0.09)	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.11)	(0.11)	(0.11)	(0.05)	(0.03)

H_0 : coefficients jointly zero across the 11 equations; $\chi^2(11) = 8.98$; p-val = 0.62

Notes: Results from Seemingly Unrelated Regression of a system of 11 equations like [Eq. 38](#), where 11 pre-determined (i.e., at the time of the first application) covariates act as dependent variables. The reported figures are estimates of the coefficient on the instrument P_{it} , which indicates whether a child's FAI is below the Preferred FAI threshold or not; the running variable is the Family Affluence Index (FAI), and second-order polynomials in FAI and application set controls are included on the RHS. Robust standard error in parentheses, clustered at the facility level. The table also reports the results from a test that the 11 coefficients on P_{it} are jointly zero across the 11 equations. Legend: var1 = neighborhood median income; var2 = number of siblings; var3 = number of programs in the application set; var4 = birth day in the year (1-366); var5 = whether Cesarean delivery of child; var6 = father education, years; var7 = mother year of birth; var9 = mother year of birth siblings; var10 = whether father was self-employed at the time of the first application; var11 = whether mother was self-employed at the time of the first application; Sample: 444 interviewed children with two working parents, born between 1999 and 2005 and whose parents first applied for admission between 2001 and 2005. * significant at 5%; ** significant at 1%.

Figure 10: 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 LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from [Calonico et al. \(2014b\)](#). 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 to programs with rationing, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in [Fort et al. \(2017\)](#), is different from zero.

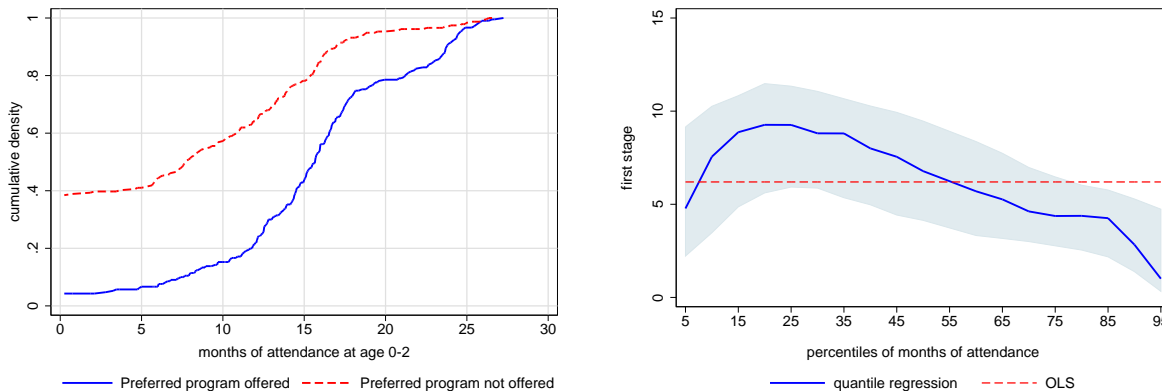
Because of how we invited families into the project, starting from those closer to the Final FAI thresholds, the interview sample differs from the Basket 4 universe in a way such that these jumps are larger than those observed in [Figure 4](#), resulting in a stronger first stage of our Instrumental Variable estimates.

Since the model of [Section 6](#) indicates that we are in the presence of “essential heterogeneity” ([Heckman et al., 2006](#)), our RD design cannot identify the ATE or the ATT of daycare attendance, but if monotonicity is satisfied it can identify the average effect of an additional month of daycare attendance on the IQ of children who attend for a longer period because their parents have been offered their most preferred program as opposed to a less preferred one.⁵⁵ [Remark 1](#) shows that, in our setting, we should expect monotonicity to hold: the offer of the most preferred program (as represented by $z = 1$ as opposed to $z < 1$)

⁵⁵Therefore, β is an average of the Local Average Treatment Effects (LATE) at the different Preferred thresholds (see [Hahn et al., 2001](#)). [Felfe and Lalive \(2014\)](#), thanks to a continuous instrument, can analyze a similar causal estimand using a Marginal Treatment Effect (MTE) approach.

implies unambiguously an increase of daycare attendance for all parents. This prediction is indeed supported by the evidence of Figure 11. In the left panel we follow Angrist and Imbens (1995) and plot the Cumulative Distribution Function (CDF) of days of attendance for the two groups of children defined by our instrument. Visual inspection indicates that days of attendance for those who are offered their most preferred program (continuous line) first-order stochastically dominate days of attendance for those who are not (dashed line), which is a necessary condition for monotonicity (under independence). As suggested by Fiorini and Stevens (2014), we use 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 15 of the Online Appendix for full details). The right panel of the same figure plots estimates of the first stage effect of being offered the preferred program on different quantiles of days 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.⁵⁶

Figure 11: Monotonicity of 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. 38. 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, born between 1999 and 2005 whose parents first applied for admission between 2001 and 2005.

⁵⁶Since there is no evidence that monotonicity is violated, we do not need to follow the approach suggested by DeChaisemartin (2016) who proposes a weaker condition for the identification of the the LATE in the presence of defiance. Incidentally, also the Kitagawa (2015) or Mourifié and Wan (2015) testing procedures are not applicable in our case, because the treatment is continuous.

The first row of the left panel in [Table 5](#) reports estimates of the Intention To Treat (ITT) effect of just qualifying for the most preferred program in the interview sample.⁵⁷ The specification in the first column includes only the polynomial $g(y_i)$. The second column adds the application set characteristics, and the third one includes all controls. Note that these ITT estimates are similar across columns. Taking the third column as the preferred specification, the estimated ITT reveals that crossing the Preferred FAI threshold (i.e., having a FAI barely sufficient to qualify for the preferred program), reduces total IQ by 3.0%. This estimate is statistically different from zero (p-value: 0.005) and it is negative as the numerical prediction (-1.8%) of the model in [Section 6.2](#), when the observed distribution of the FAI in the interview sample is used to weight the solutions computed for each specific FAI level. Assuming that the model is a correct representation of reality, the two numbers would coincide if the distribution of FAI levels at each Preferred thresholds corresponded to the distribution of FAI levels in the interview sample.

First stage estimates are reported in the second row of [Table 5](#).⁵⁸ Just qualifying for the preferred program increases daycare 0–2 attendance by about six months. The F-test statistic on the excluded instrument is sufficiently large in all specifications and samples, suggesting that weak instruments are not a concern.⁵⁹ Rescaling the ITT effect by the first stage gives the IV estimate of the effect of one month of daycare 0–2 attendance. In our preferred specification (third column), and similarly in the others, this is a statistically significant loss of about 0.5% (p-value: 0.004) which, at the sample mean (116.4), corresponds to 0.6 IQ points and to 4.5% of the IQ standard deviation.

[Table 6](#) offers a check on the parametric assumptions underlying [Table 5](#), using the

⁵⁷ Specifically, we estimate the following reduced form equation,

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

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

⁵⁸ In this case, we estimate the first stage equation,

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

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

⁵⁹In the Basket 4 universe described by the right panel of [Figure 4](#), crossing the preferred threshold from higher to lower FAI implies an increase of only about three months of daycare attendance. The fact that in the interview sample, as shown also by [Figure 10](#), the first stage is larger is due to the differences between this sample and the universe discussed in [Section 5](#).

Table 5: Effects of daycare 0–2 attendance on IQ, for all children and by level of the Preferred FAI threshold

	Dependent variable: $\ln(IQ)$			
	All FAI thresholds (mean threshold: €24.7k)	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* (0.010)	-0.030** (0.011)	-0.030** (0.010)	-0.056** (0.018)
First stage: effect of qualifying on months of attendance	6.3** (0.9)	6.4** (0.9)	5.9** (0.9)	4.6** (1.3)
IV effect of one month of daycare attendance	-0.004** (0.002)	-0.005** (0.002)	-0.005** (0.002)	-0.012** (0.004)
F-stat on excluded instruments	49.1	46.9	44.8	12.9
Number of observations	444	444	444	216
Polynomial in FAI	Yes	Yes	Yes	Yes
Application set controls		Yes	Yes	Yes
Pre-treatment controls			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. First-stage coefficients are from regressions of months (1 month = 20 days of attendance) spent in daycare 0–2 on the instrument and controls. 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. 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.

methodology suggested in [Calonico *et al.* \(2014b\)](#).

Column 1 shows non-parametric estimates for the whole interview sample, based on a triangular kernel and a local polynomial of degree zero with optimal bandwidth selection.⁶⁰ Results are in line with those of the left panel in [Table 5](#) although less statistically significant given the small sample size. The ITT effect of just qualifying for the preferred program is estimated to be a loss of 2.8% (it is 3.0% in [Table 5](#)). The non-parametric first stage estimate is slightly lower than the parametric one as well (4.6 instead of 6.3 months). As for the IV estimate, both the conventional and the bias-corrected non-parametric estimators suggested by [Calonico *et al.* \(2014b\)](#) imply a 0.6% IQ loss induced by one additional month of daycare 0–2 (which is slightly higher than the 0.5% parametric loss reported in [Table 5](#)).

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 effect of daycare, for children who increase attendance because their parents are offered their most preferred program as opposed to a less preferred one, is estimated to be negative in this population. To further probe the empirical validity of the prediction in [Remark 2](#), we separate children in two groups according to whether the Preferred threshold to which they are associated is above or below the median of all Preferred thresholds. Results are reported in the middle and right panels of [Table 5](#).⁶¹ For the less affluent group, below the median, the estimates refer to the effect of daycare 0–2 around a Preferred 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). The numerical solution of the model described in [Section 6.2](#) suggests that the level of FAI above which the IQ effect induced by the offer of the most preferred program turns from positive to negative is about €18k. Therefore, we expect the estimates for the less affluent group to be close to zero or even slightly positive, while those for the more affluent group should be unambiguously negative.

⁶⁰Estimates based on a higher order local polynomial, which according to [Imbens and Lemieux \(2008\)](#) should be preferable in this context, are unstable and not reliable. Since the non-parametric analysis is implemented on stacked and centered thresholds, here we drop observations located at zero distance from thresholds for the reasons discussed in [Fort *et al.* \(2017\)](#). See the note to the table for further details.

⁶¹ The Online Appendix reproduces the figure and tables of the main text supporting the validity of our identification strategy separately for the two affluence groups. All the available covariates are perfectly balanced and continuous at the thresholds in both cases.

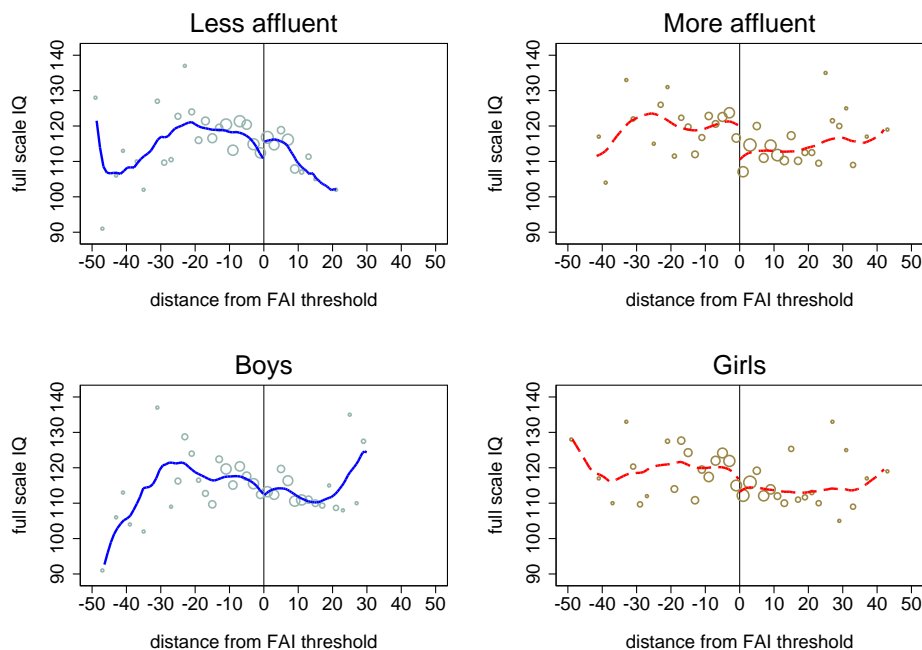
Table 6: Effects of daycare 0–2 attendance on IQ: nonparametric estimates.

Preferred thresholds	All	\leq median	$>$ median	All	All
Gender	All	All	All	Girls	Boys
<i>ITT of just qualifying</i>	-0.028 (0.020)	0.0324 (0.033)	-0.0791 ⁺ (0.032)	-0.037 (0.026)	-0.011 (0.041)
<i>First stage</i>	4.6** (1.3)	3.8* (1.8)	4.9** (2.2)	3.8* (1.7)	6.2* (2.4)
robust p-value	0.045	0.213	0.289	0.485	0.055
<i>Effect of 1 month (conventional)</i>	-0.006 (0.005)	0.009 (0.010)	-0.016* (0.008)	-0.010 (0.008)	-0.002 (0.007)
<i>Effect of 1 month (bias-corrected)</i>	-0.006 (0.005)	0.014 (0.010)	-0.023** (0.008)	-0.015 ⁺ (0.008)	-0.001 (0.007)
<i>Effect of 1 month (robust)</i>	-0.006 (0.006)	0.014 (0.013)	-0.023* (0.011)	-0.015 (0.011)	-0.001 (0.009)
Bandwith for Loc. Poly (h)	6.544	3.094	7.691	7.076	5.088
Bandwith for bias (b)	16.351	7.039	16.780	16.297	10.820
Number of observations	150	45	78	91	46

Notes: The table reports nonparametric estimates of the effect of one month of daycare 0–2 on the log of IQ, with related ITT and first stage. The methodology used is detailed in [Calonico et al. \(2014b\)](#), and it is implemented using the software described in [Calonico et al. \(2014a\)](#). The grade of local polynomials is zero and the kernel is triangular. The table also reports the optimal bandwidths for the local polynomial (h) and for the bias (b) as well as the p-value from a test for the null hypothesis that the first stage coefficient is zero, obtained using a robust and bias corrected estimator for the first stage. The ITT and first stage estimates are obtained using the conventional nonparametric estimator. The effects of one month of daycare 0–2 are obtained using three distinct RD estimators: the local polynomial estimator (conventional), the bias-corrected estimator proposed by [Calonico et al. \(2014b\)](#) and the bias-corrected estimator with robust standard errors. The running variable is the Family Affluence Index (FAI). Sample: 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. ⁺ significant at 10%; * significant at 5%; ** significant at 1% or better.

The top panel of Figure 12 offers a graphical confirmation of these predictions. While the LLR fit for the more affluent group shows a negative jump of child IQ when the FAI crosses Preferred thresholds from higher to lower values (left to right), for the less affluent group, the jump is positive and smaller in size.

Figure 12: Full scale IQ around Preferred FAI thresholds, by family affluence and by gender



Notes: The dots represent average IQ inside €2k bins, plotted as a function of the distance (thousands of real €) of a child’s FAI from her Preferred FAI threshold, by family affluence (top panel) and by gender (bottom panel). The size of a circle is proportional to the number of observations in the corresponding €2k bin. “Less affluent” and “More affluent” are observations associated with Preferred FAI thresholds below or above the median Preferred FAI threshold, respectively. The bold lines are LLR on the underlying individual observations, with a triangular kernel and optimal bandwidth selection, from Calónico *et al.* (2014b). 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 to programs with rationing, whose FAI distance from the Final FAI thresholds is at most €50k and, for the reasons discussed in Fort *et al.* (2017), is different from zero.

Parametric estimates of these jumps are reported in the top row of the middle and right panels of Table 5. In the less affluent group, the ITT effect of just qualifying for the most preferred program is estimated to be negative but very small. In the more affluent group the estimate is similarly negative, but about 5 time larger (-5-6%) and very precise. The second row in the same panels of Table 5 displays the first stage effect and confirms another prediction of the model described in Section 6. As implied by Remark 1 the first stage is smaller for more affluent parents (about 4.7 vs. 5.7 months). Rescaling the ITT effect by the

first stage gives a statistically significant IV estimate for β of about -1.2–1.4% in the more affluent group, while for less affluent households the estimate is essentially null. A similar pattern emerges in the nonparametric estimates for the two groups reported in columns 2 and 3 of Table 6.

8 Suggestions from the psychological literature

Psychologists have produced persuasive empirical evidence that one-to-one interactions with adults (more than interactions with peers) are a crucial input for cognitive development in the first three years of life 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 two and a half 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 input 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 1-3 years old children predicts their motivation five years later.⁶²

What is perhaps even more interesting is why, according to psychologists, these one-to-one interactions with adults early in life are so important. A fascinating theory 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 value. In the absence of such communication the child has to repeat and confirm the experience many times in order to assess its general or particular validity (very much like in a sort of statistical inference requiring a large sample). An adult,

⁶²Related to these results, psychologists (like economists, see footnote 9), have estimated negative effects of increasing parental working time (in particular maternal) on cognitive, non-cognitive and behavioral outcomes of children. See, for example, Brooks-Gunn 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.

instead, can quickly inform the child on 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 an advantage in terms of cognitive development.

The focus on one-to-one interactions in our context is relevant because, as noted by [Clarke-Stewart *et al.* \(1994\)](#), 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 attention in a 1:1 ratio, possibly somewhat worse 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).⁶³ 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, different from us as explained in [Section 2](#), both [Felfe and Lalive \(2014\)](#) and [Drange and Havnes \(2015\)](#) find positive effects of daycare 0–2 in Germany and Norway. In their institutional setting, the adult-to-child ratio is about 1:3, much closer to a family environment.

A related claim that psychologists have supported with persuasive empirical evidence is 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. [Galsworthy *et al.* \(2000\)](#) examine about 3,000 2-year-old twin pairs and show

⁶³In 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 8 charter facilities (offering 40 programs) that are privately managed but strictly regulated by the BDS. According to the reputational indicator of quality described in [Section 3](#), these charter programs are perceived by parents as worse than the non-charter ones. On average, for given distance, charter programs are ranked 0.43 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 IQ effect that we estimate – which derives from the offer of the most preferred program to a parent. If anything, their presence should reduce the absolute size of our estimates.

that girls score higher on verbal and non-verbal cognitive ability. The longitudinal study of [Bornstein *et al.* \(2004\)](#) on 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, and these are more valuable for their cognitive development. In [Section 6.1.4](#), we have proposed a model of how parental decisions concerning child care are compatible with these gender differences in the effects of daycare 0–2 .

The lower panel of [Figure 12](#) provides a graphical description of how IQ changes around Preferred FAI thresholds, separately by gender. While the LLR fit for boys is continuous when the FAI crosses the Preferred threshold from higher to lower values (left to right), for girls we observe a negative jump. Girls with a FAI small enough to just qualify for their most preferred program have a lower IQ with respect to girls who barely do not.⁶⁴

This finding is confirmed by the parametric estimation of [Eq. 38](#), separately for boys and girls. Results are reported in [Table 7](#), for the same three specifications already considered in [Table 5](#). In our preferred specification, which includes all controls in the third and sixth columns, the ITT effect of just qualifying for the most preferred program is a statistically significant IQ loss of about 3.9% for girls (p-value: 0.017), while for boys it is smaller and we cannot reject the hypothesis that it is equal to zero. 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), while the effect for boys is almost half this size and is again not statistically different from zero.⁶⁵ The middle panel shows that the gender gap

⁶⁴We 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, separately by gender, the figures and tables reported here supporting the validity of our identification strategy. All the available covariates as well as treatment intensity are balanced and continuous at the thresholds for both genders. Similar across genders are also the answers given by parents to the question concerning alternative modes of care.

⁶⁵Interestingly, 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).

in the ITT is not due to differences in the first stage. For both boys and girls, attendance in daycare 0–2 increases by approximately 6 months when the FAI crosses the Preferred threshold from higher to lower values. A qualitatively similar gender difference in the effects of IQ emerges from the non-parametric estimates reported in columns 4 and 5 of Table 6 where, if anything, the losses for girls appear to be even larger.

Table 7: Gender heterogeneity of the IQ effects of daycare 0–2

	Dependent variable: ln(IQ)					
	Boys			Girls		
ITT effect of qualifying for the preferred program	-0.013 (0.016)	-0.018 (0.016)	-0.023 (0.016)	-0.040** (0.015)	-0.043** (0.016)	-0.039** (0.016)
First stage: effect of qualif. on months of attendance	6.11** (0.98)	6.68** (0.93)	6.32** (0.89)	6.54** (1.12)	6.42** (1.19)	5.60** (1.19)
IV effect of one month of daycare attendance	-0.002 (0.003)	-0.003 (0.002)	-0.004 (0.002)	-0.006* (0.002)	-0.007* (0.003)	-0.007** (0.003)
F-stat on excluded instrum.	38.9	51.1	50.6	33.9	29.0	22.1
Number of observations	215	215	215	229	229	229
Polynomial in FAI	Yes	Yes	Yes	Yes	Yes	Yes
Appl. set controls		Yes	Yes		Yes	Yes
Pre-treat. controls			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, by gender. First-stage coefficients are from regressions of months (1 month = 20 days of attendance) spent in daycare 0–2 on the instrument and controls, as in equation (58). IV coefficients are from regressions of log IQ on months of attendance and controls, as in equation (38), 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: 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.

This gender difference in the cognitive losses induced by daycare attendance supports the relevance of one-to-one interactions with adults as an explanation of our results.⁶⁶

⁶⁶We have also explored the possibility that the loss suffered by girls depend on sex ratios within each

9 Conclusions

This paper joins a growing literature studying the effects of time spent in daycare 0–2 for children in advantaged families. We study 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 IQ loss at age 8–14. This loss is even more pronounced when, within this population, we look at children with relatively more affluent parents. These are typically the relevant marginal subjects to be considered in an evaluation of daycare expansions for the worldwide increasing community of families in which both parents want to work.

We interpret this findings in a theoretical model showing that, when offered the most preferred daycare program as opposed to a less preferred one, relatively affluent parents take advantage of this opportunity to increase daycare attendance of their children and so work more or reduce costly market care. This increase in attendance is smaller than the analogous one occurring in a less affluent household, but due to a higher earning potential it generates an increase of family resources that is large enough to become attractive even if it comes at the cost of a decrease in child IQ.

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, according to which 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.

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, possibly because the variation in these ratios is quite small for the children in our sample. Moreover, the data do not support another possible hypothesis according to which 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 *et al.*, 1999; Borra *et al.*, 2012; Fitzsimons and Vera-Hernandez, 2013), and early daycare enrollment or 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.

The latter claim is supported by the finding that daycare 0–2 has a more negative effect on the IQ of girls, in combination with the psychological evidence suggesting that girls are developed enough at this young age to exploit high 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 under their jurisdiction. Our estimates speak precisely about the effect of such a policy, which would allow more affluent children to attend for a longer time in programs that they prefer more, with negative effects on their IQ that may not be socially optimal even if the utility of their parents increases.

References

- Abdulkadiroglu, A., Angrist, J. D., Narita, Y., Pathak, P. A., 2015, Research design meets market design: Using centralized assignment for impact evaluation, Working Paper 21705, National Bureau of Economic Research.
- Abdulkadiroglu, A., Angrist, J., Pathak, P., 2014, The Elite Illusion: Achievement Effects at Boston and New York Exam Schools, *Econometrica* 82, 137–196.
- Adi-Japha, E., Klein, P. S., 2009, Relations Between Parenting Quality and Cognitive Performance of Children Experiencing Varying Amounts of Childcare, *Child Development* 80, 893–906.
- Almond, D., Currie, J., 2011, Human Capital Development Before Age Five, volume 4, Part B of *Handbook of Labor Economics*, 1315 – 1486, Elsevier.
- Anderson, J. W., Johnstone, B. M., Remley, D. T., 1999, Breast-feeding and cognitive development: a meta-analysis, *The American Journal of Clinical Nutrition* 70, 525–535.
- Anderson, M. L., 2008, Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects, *Journal of the American Statistical Association* 103, 1481–1495.
- Angrist, J., Imbens, G., 1995, Two-Stage Least Squares Estimation of Average Causal Effects in Models with Variable Treatment Intensity, *Journal of the American Statistical Association* 90, 431–442.
- Baker, M., Gruber, J., Milligan, K., 2008, Universal Child Care, Maternal Labor Supply, and Family Well-Being, *Journal of Political Economy* 116, 709–745.
- Baker, M., Gruber, J., Milligan, K., 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., Donald, S. G., 2003, Consistent Tests For Stochastic Dominance, *Econometrica* 71, 71–104.
- Becker, G. S., 1965, A Theory of the Allocation of Time, *The Economic Journal* 75, 493–517.
- Bellettini, G., Taddei, F., Zanella, G., 2017, Intergenerational altruism and house prices: Evidence from bequest tax reforms in Italy, *European Economic Review* 92, 1 – 12.
- Belsky, J., 1988, The effects of infant day care reconsidered, *Early Childhood Research Quarterly* 3, 235 – 272.
- Belsky, J., 2001, Emanuel Miller Lecture: Developmental risks (still) associated with early child care, *Journal of Child Psychology and Psychiatry* 42, 845–859.
- Belsky, J., Steinberg, L. D., 1978, The effects of day care: A critical review, *Child Development* 49, 929–949.
- Bernal, R., 2008, The effect of maternal employment and child care on children’s cognitive development, *International Economic Review* 49, 1173–1209.

- Bernal, R., Keane, M. P., 2011, Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers, *Journal of Labor Economics* 29, pp. 459–512.
- Bertanha, M., 2016, Regression Discontinuity Design with Many Thresholds, Mimeo.
- Borghans, L., Duckworth, A. L., Heckman, J. J., ter Weel, B., 2008, The Economics and Psychology of Personality Traits, *Journal of Human Resources* 43, 972–1059.
- Bornstein, M., Hahn, C.-S., Gist, N., Haynes, M., 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, 129–156.
- Bornstein, M. H., Hahn, C.-S., Haynes, O. M., 2004, Specific and general language performance across early childhood: Stability and gender considerations, *First Language* 24, 267–304.
- Borra, C., Iacovou, M., Sevilla, A., 2012, The effect of breastfeeding on children's cognitive and noncognitive development , *Labour Economics* 19, 496 – 515.
- Brooks-Gunn, Jeanne, W.-J. H., Waldfogel, J., 2002, Maternal Employment and Child Cognitive Outcomes in the First Three Years of Life: The NICHD Study of Early Child Care, *Child Development* 73, 1052–1072.
- Calonico, S., Cattaneo, M. D., Titiunik, R., 2014a, Robust Data Driven Inference in the Regression-Discontinuity Design, *The Stata Journal* 14, 909–946.
- Calonico, S., Cattaneo, M. D., Titiunik, R., 2014b, Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs, *Econometrica* 82, 2295–2326.
- Campbell, F. A., Ramey, C. T., 1994, Effects of Early Intervention on Intellectual and Academic Achievement: A Follow-up Study of Children from Low-Income Families, *Child Development* 65, 684–698.
- Carneiro, P., Cunha, F., Heckman, J., 2003, Interpreting The Evidence of Family Influence on Child Development, Mimeo.
- Carneiro, P., Ginja, R., 2014, Long-Term Impacts of Compensatory Preschool on Health and Behavior: Evidence from Head Start, *American Economic Journal: Economic Policy* 6, 135–73.
- Carneiro, P., Løken, K. V., Salvanes, K. G., 2015, A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children, *Journal of Political Economy* 123, pp. 365–412.
- Cartmill, E., Armstrong, B., Gleitman, L., Goldin-Meadow, S., Medina, T., Trueswell, J., 2013, Quality of early parent input predicts child vocabulary 3 years later, *PNAS* 110, pp. 11278–11283.
- Cattaneo, M., Keele, L., Titiunik, R., Vazquez-Bare, G., 2016, Interpreting Regression Discontinuity Designs with Multiple Cutoffs, Mimeo.
- Clarke-Stewart, K., Gruber, C., Fitzgerald, L., 1994, *Children at home and in day care* , L. Erlbaum Associates, Hillsdale, NJ.

- Comune di Bologna, 2010, Progetto pedagogico del nido d'infanzia del Comune di Bologna, Report: Comune di Bologna- Settore Istruzione.
- Csibra, G., Gergely, G., 2009, Natural Pedagogy, *Philosophical Transactions of the Royal Society B* 13, 148–153.
- Csibra, G., Gergely, G., 2011, Natural Pedagogy as evolutionary adaptation, *Trends in Cognitive Sciences* 366, 1149–1157.
- Cunha, F., Heckman, J., 2007, The Technology of Skill Formation, *American Economic Review* 97, 31–47.
- Cunha, F., Heckman, J. J., Schennach, S. M., 2010, Estimating the Technology of Cognitive and Noncognitive Skill Formation, *Econometrica* 78, pp. 883–931.
- Currie, J., 2001, Early Childhood Education Programs, *Journal of Economic Perspectives* 15, 213–238.
- DeChaisemartin, C., 2016, Tolerating Defiance? Local Average Treatment Effects Without Monotonicity, *Quantitative Economics*, forthcoming .
- Del Boca, D., Pronzato, C., Sorrenti, G., 2015, When rationing plays a role: selection criteria in the Italian early child care system, *Carlo Alberto Notebooks* 399, Collegio Carlo Alberto.
- Drange, N., Havnes, T., 2015, Child Care Before Age Two and the Development of Language and Numeracy: Evidence from a Lottery, *IZA Discussion Papers* 8904, Institute for the Study of Labor (IZA).
- Duncan, G. J., Magnuson, K., 2013, Investing in Preschool Programs, *Journal of Economic Perspectives* 27, 109–32.
- Dustmann, C., Raute, A., Schönberg, U., 2013, Does Universal Child Care Matter? Evidence from a Large Expansion in Pre-School Education, *Mimeo*.
- Elango, S., Garca, J. L., Heckman, J. J., Hojman, A., 2015, Early childhood education, *Working Paper* 21766, National Bureau of Economic Research.
- Felfe, C., Lalive, R., 2014, Does Early Child Care Help or Hurt Childrens's Development?, *IZA Discussion Papers* 8484, Institute for the Study of Labor (IZA).
- Felfe, C., Nollenberger, N., Rodrguez-Planas, N., 2015, Cant Buy Mommys Love? Universal Childcare and Children's Long-Term Cognitive Development, *Journal of Population Economics* 28, 393–422.
- Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., Stiles, J., 1994, Variability in Early Communicative Development, *Monographs of the Society for Research in Child Development* 59, pp. i+iii–v+1–185.
- Fiorini, M., Stevens, K., 2014, Assessing the Monotonicity Assumption in IV and Fuzzy RDD Designs, *Economics Working Paper Series- University of Sidney* 2014-13.
- Fitzsimons, E., Vera-Hernandez, M., 2013, Food for Thought? Breastfeeding and Child Development, *IFS Working Papers* (W13/31).

- Fort, M., Ichino, A., Rettore, E., Zanella, G., 2017, On the Perils of Stacking Thresholds in RD Designs, Mimeo.
- Fort, M., Ichino, A., Zanella, G., 2016, Daycare 0-2 for Children: Does Early Daycare Improve Children's Outcomes?, Report prepared for the Fondazione Rodolfo DeBenedetti.
- Gale, D., Shapley, L., 1962, College Admissions and the Stability of Marriage, *American Mathematical Monthly* 69, pp. 9–15.
- Galsworthy, M. J., Dionne, G., Dale, P. S., Plomin, R., 2000, Sex differences in early verbal and non-verbal cognitive development, *Developmental Science* 3, 206–215.
- Garber, H., 1988, The Milwaukee Project: Preventing Mental Retardation in Children At Risk, Technical report, American Association on Mental Retardation.
- Gottfredson, L. S., 1997, Mainstream science on intelligence: An editorial with 52 signatories, history, and bibliography, *Intelligence* 24, 13 – 23, special Issue *Intelligence and Social Policy*.
- Gunderson, E. A., Gripshover, S. J., Romero, C., Dweck, C. S., Goldin-Meadow, S., Levine, S. C., 2013, Parent praise to 1- to 3-year-olds predicts children's motivational frameworks 5 years later, *Child Development* 84, 1526–1541.
- Hahn, J., Todd, P., Van der Klaauw, W., 2001, Identification and Estimation of Treatment Effects with a Regression Discontinuity Design, *Econometrica* 69, 201–209, notes and Comments.
- Hart, B., Risley, T., 1995, *Meaningful Differences in the Everyday Experience of Young American Children*, Brookes Publishing Co., Baltimore MD.
- Havnes, T., Mogstad, M., 2015, Is Universal Child Care Leveling the Playing Field?, *Journal of Public Economics* 127, 100 – 114.
- Heckman, J. J., Mosso, S., 2014, The Economics of Human Development and Social Mobility, *Annual Review of Economics* 27, 109–32.
- Heckman, J. J., Urzua, S., Vytlačil, E., 2006, Understanding Instrumental Variables in Models with Essential Heterogeneity, *The Review of Economics and Statistics* 88, 389–432.
- Herbst, C. M., 2013, The Impact of Non-Parental Child Care on Child Development: Evidence from the Summer Participation Dip, *Journal of Public Economics* 105, 86 – 105.
- Hewett, V., 2001, Examining the Reggio Emilia Approach to Early Childhood Education, *Early Childhood Education Journal* 29, 95–100.
- Imbens, G. W., Lemieux, T., 2008, Regression discontinuity designs: A guide to practice, *Journal of Econometrics* 142, 615–635.
- Kitagawa, T., 2015, A test for instrument validity, *Econometrica* 83, 2043–2063.
- Koch, J., 2003, Gender Issues in the Classroom, in: Reynolds, W. M., Miller, G. E. (eds.), *Handbook of Psychology*, volume 7 (Educational Psychology), 259–284, Wiley.

- Kottelenberg, M. J., Lehrer, S. F., 2014a, Do the perils of universal childcare depend on the child's age?, *CESifo Economic Studies* 60, 338–365.
- Kottelenberg, M. J., Lehrer, S. F., 2014b, The Gender Effects of Universal Child Care in Canada: Much ado about Boys?, Mimeo.
- Kottelenberg, M. J., Lehrer, S. F., 2017, Targeted or universal coverage? assessing heterogeneity in the effects of universal childcare, *Journal of Labor Economics* 35.
- Lalive, R., Zweimüller, J., 2009, How does parental leave affect fertility and return to work? evidence from two natural experiments, *The Quarterly Journal of Economics* 124, 1363.
- Lee, D. S., Lemieux, T., 2010, Regression discontinuity designs in economics, *Journal of Economic Literature* 48, 281–355.
- Li, J., Johnson, S. E., Han, W.-J., Andrews, S., Kendall, G., Strazdins, L., Dockery, A., 2013, Parents' nonstandard work schedules and child well-being: A critical review of the literature, *The Journal of Primary Prevention* 35, 53–73.
- Magnuson, K. A., Ruhm, C., Waldfogel, J., 2007, Does prekindergarten improve school preparation and performance?, *Economics of Education Review* 26, 33–51.
- McCrary, J., 2008, Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test, *Journal of Econometrics* 142, 698–714.
- McPherran Lombardi, C., Levine Coley, R., 2014, Early Maternal Employment and Children's School Readiness in Contemporary Families, *Developmental Psychology* 50, 2071–2084.
- Mourifié, I., Wan, Y., 2015, Testing local average treatment effects assumptions, Mimeo, University of Toronto.
- Noboa-Hidalgo, G. E., Urza, S. S., 2012, The Effects of Participation in Public Child Care Centers: Evidence from Chile, *Journal of Human Capital* 6, 1 – 34.
- Puma, M., Bell, S., Cook, R., Heid, C., Broene, P., Jenkins, F., Mashburn, A., Downer, J., 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, A. E., 2007, Deferred acceptance algorithms: History, theory, practice, and open questions, Working Paper 13225, National Bureau of Economic Research.
- Rowe, M. L., Goldin-Meadow, S., 2009, Differences in early gesture explain SES disparities in child vocabulary size at school entry, *Science* 323, 951–953.
- Wechsler, D., Kaplan, E., Fein, D., Kramer, J., Morris, R., Delis, D., Maelender, A., 2003, Wechsler intelligence scale for children: Fourth edition (WISC-IV), Assessment instrument, San Antonio, TX; Pearson .
- Zigler, E., Butterfield, E. C., 1968, Motivational Aspects of Changes in IQ Test Performance of Culturally Deprived Nursery School Children, *Child Development* 39, pp. 1–14.