INSTITUTO DE ECONOMÍA

DOCUMENTO de TRABAJO

514

2019

Home alone vs kids club: adult supervision matters for grades

Claudia Martínez A. y Marcela Perticará.

www.economia.puc.cl · ISSN (edición impresa) 0716-7334 · ISSN (edición electrónica) 0717-7593

Home alone vs kids club: adult supervision matters for grades

Claudia Martínez A.
Pontificia Universidad Católica de Chile

Marcela Perticará
Universidad Alberto Hurtado[1*]

January, 2019

Abstract

In this paper, a randomized intervention was used to provide school age children with three hours of after school care in order to explore the causal effects of adult supervision on children's academic outcomes. It was found that there is a positive effect on grades when the after-school program replaces non-parental care, increasing the average GPA and the probability of having a GPA above the median in around 9 percentage points (pp). If the student was alone in the after-school hours at baseline, this value increases to 13 pp. This evidence suggests that the impact of the program is determined by the nature of the alternative care available to them. We find that the program quality, the activities developed, and the characteristics of the personnel do not affect the program effect. We also find no impact on mother's labor force outcomes for the group of students where improvement was observed. These results are consistent with the program working through the provision of adult supervision, instead of through the type of care or an income effect.

Keywords: Childcare; randomized control trial; after-school programs

JEL Codes: J13, I25

[1*] Martínez A.: clmartineza@uc.cl; Perticará M.: mperticara@uahurtado.cl. This project received funding from the Inter-American Development Bank and the Chilean Servicio Nacional de la Mujer (SERNAM). We appreciate the excellent research assistance provided by Antonia Asenjo, María Ignacia Contreras, and Sebastián Otero. We acknowledge financial support from Fondecyt nº1170730

1. Introduction

As the female labor force participation increases from 42.4% in 1990 to 56.6% in Latin America and de Caribbean (LAC),[2] an important question is: Who is taking care of children while parents work and how do these childcare arrangements affect children's wellbeing? Ruiz-Casares et al. (2018) report figures for 61 developing countries of children under five that are unsupervised or under the supervision of another child younger than 10 years of age for at least one hour per day using UNICEF data. In LAC, the figure is between 5.7% and 0.9%, worldwide between 45.3% and 0.5%.

There is limited data to systematically analyze the prevalence of different types of care for school-aged children. In the US, Laughlin (2013) reports that 12.6% of children 5 to 14 years old were under sibling or other relative care with an additional 11.1% under self-care. Self-care increases with age and depending on the mother's working status.[3] In Chile, 64% of first graders are cared for by one of their parents after school (Junaeb, 2017).[4] Therefore, there is a substantial number of children that are unsupervised or inadequately

[2] According to the World Development Indicators.

[3] The fraction of children with working mothers who stay alone at home is 5.6% for 9-11 years old, but 16.1% 12-14 years old children.

[4] 54.6% by the mother, 20.9% by their grandparents, 9.8% by other family members, and 3.4% by non-family members). Retrieved on October 8, 2018 from *https://sistemaencuestas.junaeb.cl/encuestasjunaeb/index.jsp*

supervised, depending on country, family characteristics, and employment status of the mother.

The absence of adult supervision of school-aged children has been associated with antisocial behavior (Azier, 2004), poor school performance (Bettinger et al., 2014), and teenage pregnancy (Dwyer et al., 1990). Previous studies have looked at the potential effect of policies that increase supervised time by extending schooling hours or establishing permanent after school programs (ASP).

In the first case, extra hours or days are devoted to academic activities. The effect of this has been studied in developing countries, finding a decrease in teen pregnancy (Berthelon & Kruger, 2011 for Chile), and improvements in both reading and math scores (Bellei, 2009; Berthelon et al., 2016, both for Chile; Hincapie, 2016 for Colombia; Cerdan, 2007 for Uruguay). In developed countries, longer school days had a positive impact on college enrollment rates (Lavy & Scholoser, 2005) and performance in math, English and science (Lavy, 2012) (Battistin & Meroni, 2015; Jensen, 2013).

On the other hand, ASPs are structured programs that are run under adult supervision in schools throughout the academic year. The extra hours are used heterogeneously depending on the program objectives. They can include a variety of activities, such as homework time, social interaction, snacks, sports, crafts, etc., or can be designated for a specific goal such as a programming camp. As far as the researchers are aware, there are no impact evaluations of ASP programs in developing countries. The evidence in developed countries is mixed (Goerlich et al., 2007; Kremer et al., 2015), but suggests that students *at risk* benefit from ASPs the most (Levine & Zimmerman, 2010) and that expected benefits

depend on the quality the intervention. It is also relevant to consider the alternative care children would have had in the absence of an ASP program: The lower the quality of the alternative care, the greater the impact of the program (Felfe & Zierow, 2012).

This study uses a randomized control trial in which children 6–13 years old were placed in an after-school program in order to estimate the causal effect of providing kids with a supervised environment on student's academic outcomes. The ASP exposed children to playful activities, such as arts, sports, and games. No resources were devoted to any academic activities.

We found that, on average, the program had no impact on academic outcomes. However, consistent with the previous literature, there was a positive effect on academic performance for children that were not under their parent's supervision after school hours. For this group we found that the program had a positive effect on average grades (physical education, average GPA) and on the probability of being above the median of the grade distribution. Moreover, higher effects were found for children that were alone or at the care of "other adults" (not parents or grandparents) at baseline. The effects are not only statistically significant, but also large. The probability of being above the median grade increased 13 percentage points for children that were alone at baseline.

There are at least four mechanisms that are consistent with this positive impact of the ASP. First, the program could have increased female labor force participation, and consequently family income more on this set of parents (Bernal, 2008; Bernal & Keane, 2010; Bernal & Keane, 2011; Black et al., 2014; Brilli 2014). Second, the program itself, could have a direct impact on academic outcomes through the provision of high-quality care

and ludic activities (such as music and art lessons, sports). Exposure to ludic activities could reduce stress levels and improve creativity in children, which could improve academic outcomes (Foster & Jenkins, 2017; See & Kokotsaki, 2016; Winner & Hetland, 2000). However, most papers fail to identify a causal connection between childhood arts participation and cognitive and developmental outcomes. Third, the program might generate higher attendance rates to school hours which could also improve academic outcomes. Finally, supervision of an adult during the after-school hours could generate these results (McCombs et al., 2017).

While we cannot clearly pin down the mechanism, we present evidence consistent with supervision effect and attendance effect, driving most of the impact of the program. On one hand, we observe strong effects on children who were taken care of by "other adults" or were left alone at home before the intervention. For children left alone at baseline, the program increases academic outcomes and attendance rates, and therefore the increase in academic grades could be explained by both the ASP supervision and the increase in school attendance. In this case, attendance is both a mechanism and an outcome. For children taken care of by other adults, the effects are likely driven by the increased supervision. We do not find any evidence that the program effects are driven by either an income effect or the quality of the implementation or the kind of activities taught to children.

Our research offers two main contributions to the literature. First, we measure the causal effect of an ASP on academic outcomes using an RCT in a developing country. Participation in an ASP program increases supervised time and gives children the opportunity to interact with other children outside school hours, but also reduces the time they spend at

home with their siblings and parents. We show that the program does not necessarily have positive effects. Second, taking advantage of baseline information about after-school parental care, we implement a difference-in-difference strategy comparing the effect of the program on children with and without adult supervision at baseline and providing evidence of heterogenous effects. This heterogeneity provides suggestive evidence about the mechanisms over which the program works.

The following sections will describe the interventions and the experimental design, data description, empirical strategy, results, and conclusions.

2. The Intervention and Experimental Design

2.1 The Intervention

The objective of the "4–to–7" ASP program was to increase the labor force participation of mothers or women responsible for taking care of children aged between 6–13 by providing playful activities in after-school hours during the academic year. The ASP is run by the Ministry of Women and Gender Equity in municipalities where a high demand was expected due to the number of children aged 6-13 and female labor force participation. Schools apply to host the program through their municipality and are selected based on three criteria: whether they have an adequate infrastructure, whether they have other ASP programs, and, if possible, whether their standardized test score had improved.

Mothers apply for the program through the schools. The eligibility requirements are all related to mother's characteristics: being economically active, older than 18 years old, working or living in the municipality where the participating school is located, and scoring

low in the socioeconomic targeting scale. Funds are transferred to the municipalities, which then select the managing organizations (NGOs) through a bidding process. Not all the participating children have to be enrolled in the same school that hosts the program. The only eligibility requirement was to either reside or attend another school located in the same municipality.

The Ministry set up terms of reference that established the minimum features of the ASP. The programs were set up in each school with a maximum of 50 or 100 beneficiaries, where the number of slots was defined *ex-ante* based on potential demand. Each program had a coordinator who was required to have formal training in the areas of education, psychology or business. Programs with 50 students must have two monitors, and programs with 100 students must have five. It was recommended that monitors should be teachers of the implementing school. However, this was the case only in 85% of the monitors in the current sample. Instructors were professionals or higher-education students in the areas of education, sports, arts, etc. 77.3% of the participating schools were assigned 50 slots, and the remaining 22.7% of schools were assigned 100 slots.

The intervention consisted of three hours of after-school care during the school week. As the schools did not have the same daily schedule, the program's times varied across schools. However, most schools in the evaluation (18 out of 25) offered the program from 4 to 7 pm.[5]

[5]Only one school offered the program during the morning. In the rest of the schools, the program was run in the afternoon, the starting time varying from 2 pm to 5 pm.

As specified in the terms of reference, the program should follow the following schedule: arrival (10 minutes), motivation (20 minutes), school-work support (30 minutes), a recess with a snack provided (30 minutes), and a thematic workshop (90 minutes). School-work support was structured to help students with their homework, teach study methods, and reinforce lessons. Thematic workshops were related to art, sports, or TICs. Each ASP program decided which thematic workshops were to be implemented based on the students' interest and age. The most common were those related to arts (crafts, theater, dance, music, cinema, circus), followed by TICs and sports.

Concurrent with the impact evaluation, an independent process evaluation was performed in 22 out of the 25 schools participating in the study. Each ASP program was visited twice with the purpose of documenting its implementation. The number of monitors required by protocols was met in 72.7% of the schools, and 94% of schools which had been assigned 50 slots. Attendance was low, reaching an average of 17.5 students. Considering the low attendance rate, the ratio of monitors to students was higher than what was recommended in the terms of reference.

2.2 Experimental Design

The impact evaluation took place in 25 schools where the program was implemented for the first time in 2012. Mothers or legal guardians of 6–13 year–old children were invited to apply for the ASP. In order to do so, they were required to fill out an application form, specifying the number of children to attend, demographic information, and schooling data.

Women were also asked to answer a full questionnaire concerning their individual and family labor and socioeconomic characteristics.

As all schools were overenrolled, available spots were randomized. The unit of randomization was the mother, so that all the children reported in her application form got an invitation to attend the ASP. This was done in order to abide by the main objective of the program, which as to help women find employment. For each available slot there were 1.7 applicants (mothers). Randomization was stratified considering the mothers' baseline work status and whether they had small children (younger than five). The offering process was done by the implementing agency.[6] The evaluation timeline is in Figure 1.

3. Data and Descriptive Statistics

3.1 Data

Administrative data provided by the Ministry of Education on attendance was used along with grades during the implementing year. All outcome variables come from these data and therefore are limited to the students' academic achievements. Monthly attendance is reported by the Ministry of Education as the fraction of the monthly school days each child attended. Grades are reported as the end of the year average by subject and overall GPA.

[6] As participating schools were not selected randomly, we cannot guarantee external validity. However, in our companion paper we report that there were no observable differences in school size, vulnerability, or in the mother's and children's characteristics in comparable schools.

These administrative data are merged with the experimental data (treatment assignment, strata, and baseline characteristics) and self-reported information on baseline childcare use provided by mothers in the ASP application form. It is also merged with a follow-up household survey with the sole purpose of use reported program use. Finally, data from the process evaluation was included to measure the program's quality.

Although the implementing agencies were required to collect attendance data for the ASP, this was not strictly enforced, making the collected data was unreliable. Therefore, we do not use this attendance rate in our impact analysis.

3.2 Baseline Characteristics and Balance

Table 1 presents the data on the outcome of the randomization process. The original sample consisted of 1,358 eligible children in the treatment group and 1,208 in the control group. Twenty-five percent of the children in the control group attended the program (as reported by their mothers); in the treatment group, this figure was 57% (column [4]). The low take-up decreased the power of the experiment, therefore limiting the probability of finding effects.[7]

Descriptive statistics and balance are reported in Table 2. Panels A and B report characteristics for both children and mothers, respectively. For each variable, we show the sample mean, standard deviation and number of observations at baseline (columns [1]-[3]),

[7] In Appendix Table A1 we show that students attending the program are more likely to study in the school where it is implemented. Their mothers have lower per capita income and less years of education.

the treatment and control mean (columns [4] and [5]), and the p-value of the null that the treatment and control group means are equal (column [6]).[8]

Students were on average 9.7 years old and in 4th grade; 47% were female. Only 52% of the students were offered the ASP in their own school. Their average grade in the previous academic year was 5.6 (grades in Chile range between 1-7, 4 being the minimum required to pass), and their average attendance was 89% (85% is required to pass, with some exceptions). Almost 60% of the children were not under their parent's supervision at baseline: 38% were under the care of another adult (grandmother, neighbor, other family members), and 11% were on their own.

Mothers were on average 37 years old and had 2.2 children. 53% of them were household heads. Their average years of education were 9.4, and the *per capita monthly* household income was US$116. Finally, 63% of the children were in the stratum characterized by mothers working at baseline and not having children younger than five years old.

The p-values in column [6] show that the groups were balanced in all these variables. Therefore, the randomization allowed for an estimation of the causal impact of the program, and it was not necessary to control any of these variables in our regressions. [9]

[8] Note that some of these variables are missing in some observations. For this reason, the sample size varies in each row of the table.

[9] There is no statistical difference in the reported variables in Table 2 between schools with 50 and 100 slots. However, there is a higher take-up in schools with 100 slots.

3.3 Attrition

The two outcome variables (grades and attendance) are compiled in different data sets by the Ministry of Education, and therefore there are different attrition rates. Regarding attendance, we find approximately 93% of students at baseline. The level of attrition is higher (almost 11%) for grades in 2012. Hence, the final estimation sample comprised 2,284 in the grade data and 2,379 children in the attendance data.

We analyze whether attrition of attendance and grade is correlated with treatment assignment in Table 3. The dependent variable is the probability of being in the administrative data and the parameters of interest are the coefficients of the treatment variable. Columns [1] and [4] report the correlation of treatment assignment with the probability of being in final regressions (for attendance and grades, respectively) without controls. Columns [2] and [5] include control variables (child's age, if child is female, and a dummy indicating if they used childcare at baseline, mother's age, education, household head, *per capita* household income, and the number of children). Finally, in column [3] and [6] we interact them with treatment assignment (not shown). In all cases, the coefficients of treatment assignment are not significant. Furthermore, the full set of interactions is jointly not different from zero. Therefore, there is no difference in attrition by treatment arm.
However, we find that the older the children, the more likely they are to have follow-up data on their grades. Although the estimated coefficient is very small, we control age in all our regressions.

4. Results

4.1 Estimated Equation and Interpretation

Our main equation is as follows:

$$Y_{ij} = \propto_j + \beta T_{ij} + \delta y_{ij,t-1} + \gamma age_{ik} + v_{ij} \quad (1)$$

Where i refers to the individual, j to school strata (defined by the mother's working status and whether they have children younger than 5 at baseline). T_{ij} is an indicator of the treatment assignment, $y_{ij,t-1}$ is the lagged value of the dependent variable, age_{ij} is the student's age, and α_j are school-strata fixed effects. Whenever the baseline value of the dependent variable is missing, we impute a zero, and include a dummy indicating if the value was imputed. Standard errors are clustered at school level.[10] β represents the Intent-to-Treat estimate. As there seems to be substantial imperfect compliance, these estimates might differ from the ATE.

We also indagate the existence of differential effects according to the baseline use of childcare, the program quality and the type of activities taught in the program. To study these heterogeneities for a given subgroup we define a dummy variable $D_{ijk} = 1$ if individual i in school j and strata k belongs to this particular group. Then we estimate the following equation:

[10] We have 25 clusters, which might lead to over rejection of the null. Our results are robust to this correction and are reported in Annex Tables B1-B5. This is consistent with Cameron, Gelbach, and Miller's (2008) simulations that show that with 20 clusters the size of the tests using clustered robust standard errors is close to the nominal one.

$$Y_{ijk} = \propto_{jk} + \beta T_{ijk} + \theta T_{ijk} * D_{ijk} + \pi D_{ijk} + \delta y_{ijk,t-1} + \gamma X_{ijk} + v_{ijk} \quad (2)$$

Where β represents the program impact for students not belonging to the subgroup D_i, and θ represents the heterogeneous impact of the treatment on the subgroup with D=1. The term πD_{ijk} was included in order to control for outcome differences by each specific subgroup. As multiple hypotheses were analyzed on several different outcomes, Annex Tables C1-C4 multiple hypothesis adjusted p-values are presented using Romano and Wolf (2005)'s stepdown hypothesis testing algorithm. The algorithm coded by Clarke (2016) was used.

4.2 Average Effects

Average effects are presented in Table 4. Panel A shows the effects on attendance and Panel B on grade-related outcomes, both in the implementing year (2012). Columns [1], [2], and [3] report the program's impact on attendance rate for the May–November period (implementing months) and the probability of passing the 90% and 95% attendance rate, respectively.[11] We observe that average attendance rates are high (90.7% in 2012,), and that the program had no impact upon any measurement of attendance.[12]

Regarding academic outcomes (Panel B), the point estimates of the program effects are all positive, but small in magnitude and only significant for grades on physical education. However, the effect is not robust to the consideration of the multiple tests used (Table C3).

[11] 85% is the passing rate in Chile, with some exceptions.

[12]We also do not find effects on attendance in 2013, the year after the implementation. Results can be made available upon request.

This result is consistent with the design of the program, which assigned only 30 minutes for homework and offered mostly workshops on arts and sports.

4.3 Heterogeneous Effects

In this section, heterogenous effects are reported to shed light on how the program could have an impact on students' outcomes. First, the literature reports that the ASP program effects depend on the alternative childcare. The effects of the program were then analyzed depending on who takes care of the children at baseline (Table 5). Secondly, as the literature also points out, the program effect depends on the quality of the program. The heterogeneous program impact was reported including various measures of quality in Table 6. This analysis also sheds light on how the program affects outcomes: if the quality is not important, this is consistent with a larger impact of the supervision time rather than the activities in the after school. We also explore whether the activities developed in the program affect its outcome it Table 7.

Third, the ASP program could have an impact on students' outcomes through its effect on female labor supply. In a companion paper, we report the existence of a small average impact on labor force participation and employment of mothers' offered the ASP. In the same paper, we report the existence of heterogeneous impact based on the mother's baseline work status and presence of small children in the household. In order to explore the potential existence of an income effect, Table 8 investigates whether the program is changing employment outcomes of mothers in the group of children whose academic outcomes increase the most with the ASP (children without parental care at baseline).

Baseline Childcare

Regarding baseline childcare, mothers were asked about who was taking care of the child after school hours. For the purposes of analysis, we defined a variable (non-parental care) that takes the value of 1 if the children used any kind of childcare after school or were left alone at baseline, and of 0 if they were taken care of by their parents. Results in Panel A, Table 5 report the interactive effect for children with non-parental care.

As all reported results correspond to ITT, it is relevant to look at the program's take-up in these two groups. Although take-up is slightly higher in children with non-parental care at baseline (which is consistent with families substituting the ASP for other forms of childcare), the difference is not statistically different from zero (Table 5, column [1]). Therefore, results are not mechanically driven by differences in use but could be driven by differences in alternative care.

The first row of Table 5, panel A shows that the program's assignment for the base category (parental care at baseline) has a negative impact on attendance, although the coefficients are not always significant. On the contrary, for students under non-parental care at baseline, coefficients are always positive and significant when the outcome is attendance rate over 95%. Regarding grades, the coefficients for students under parental care at baseline are insignificant in all grade outcomes, but always negative, suggesting that substituting institutional care for parental care does not necessarily improve children's outcomes. On the other hand, the coefficients for students under non-parental care at baseline are positive and significant for 90% attendance, art, language, average GPA, and the probability of having a GPA over the median. The overall average GPA increased by 1.2 (column [10]), the average

grade in art increased 1.5 (column [5]), and the grade in language and literature increased 1.3 (column [7]). The program also increased the probability of being above the median in 8.6 percentage points (column [11]). The coefficients on other grade outcomes were also positive, but not significant. Therefore, the ASP program had a positive impact on these students.

In panel B, Table 5, we open the "non-parental care at baseline" into its main components: care by other adults, siblings, and alone. Again, column [1] shows that although take-up is higher for some of these groups with different types of childcare at baseline, all these coefficients are not statistically different from zero.

Furthermore, it shows that the positive effects in Panel A are mostly observed in children that are either left alone at home or placed under the care of another adult (relatives and non-relatives). In fact, the greater effects are for children left alone, which could be related to the fact that the program provides them with a safe environment. For these children, there is a strong impact (3 pp) on attendance rates, suggesting that the program might have had a deterrent effect on absenteeism. This effect is relatively large, considering that attendance rates are high (around 91% for the control group). There is also a positive effect in grades: The effect on the average GPA is 0.19 points (column [10]) and the effect on the probability of having a GPA above the median is 13.4 pp. All these effects are too robust to wild cluster, and some of them are not significant when considering multiple hypotheses (Table C1). Still, the general conclusion does not change (Table B2).

The positive impact of ASP being restricted to non-adult supervision or non-parental adult is consistent with previous studies on the importance of the counterfactual care for ASP

program impact. Note that given equation (2), θ measures outcomes differences between treated children with and without parental care at baseline. For these children, the program increases formal supervision. And in this sense, a positive θ could be interpreted as the value of substituting informal care for institutional care.

Program Quality and activities

The quality of the program, in relation to alternative care can also play an important role when considering its potential impact. If relative quality does affect the program impact, it would be consistent with the hypothesis that adult supervision is driving the program impacts.

In Table 6, the interactive impact of several measures of program quality is reported. Each coefficient reported corresponds to θ in equation (2), while controlling for the treatment dummy and the quality measure dummy. Each row of Table 6 corresponds to a different measure of quality. Above median quality index is a dummy that takes the value of 1 if the program is above the median quality, where quality was measured by an index that captures the quality of infrastructure, teachers, and materials, as measured in the process evaluation. In the second and third row, measures of monitor's quality are presented: a dummy that takes value of 1 if at least 25% of monitors are school teachers and if monitors have at least an average of 3 years of experience. The next two quality measures (rows 4 and 5) are related the planning of the program components. We define two indicators. The first reports whether the program components were defined by the beginning of the school year. The second index measures if the planning also fixed the 30-minute time slot that should be dedicated to

homework time. Finally, we study the existence of an interactive effect with the activities planning (if activities found in the process evaluation were those planned at the beginning of the year), and the observed student/monitor ratio.

It is striking that independent on how quality is measured, it does not seem to have an impact on attendance or grade outcomes. Some coefficients in Table 6 are significant, but when the multiple hypotheses are considered in Table C2, this significance disappears.[13]

No evidence was found that the program quality had a differential impact on the ASP effects. These results should be taken with caution. First, our quality measures might not capture all relevant differences in quality. This study was constrained by what was observed in the process evaluation to define quality measures. Second, the average quality was high. For example, 70% of schools had program components defined in March and 80% had the homework time defined in the same period.

The impact that the type of activities in the ASP had on the program's outcome was also investigated. If there was no impact, this would also be consistent with the dominant impact coming from the supervision effects. Table 7 shows the interactive effect of different measures of activities if a program has at least one course of TIC, social science, personal care, and sports. All the schools offer at least one course in arts (dance, drama, and painting

[13] It is important to note that take-up does not differ by quality (column [1], table 6), and that therefore the effects are again not mechanically driven by differences in use.

classes), so it is not possible to distinguish the specific effect of these disciplines. We do not find that there are differential effects depending of the workshop theme.

Mothers' Labor Market outcomes

As previously mentioned, the program can have a direct effect on academic outcomes through the curriculum or by providing better care than the alternative available to them. The effect could also be indirect by increasing the family's disposable income through female employment and decreasing childcare cost.

Martínez A. and Perticará (2017) found that the ASP had a positive impact on mothers' employment. To understand if these effects drive the students' outcomes, we analyze the existence of heterogeneous effects on female labor market outcomes (labor force participation, employment, and income) according to the children's care at baseline. Results are presented in Table 8. There is no systematic impact of the ASP by baseline parental care on any of the labor market outcomes. Given that, the stronger effects are found in the group that were not cared by their parents at baseline, and particularly on those who were left alone. These results are not consistent with labor market effects driving the program's impact on student outcomes.

5. Conclusions

The impacts of an after-school program on children's academic outcomes in Chile was studied using an experimental strategy. It was found that the program had no average impact on academic outcomes (grades and attendance).

There are, however, large and statistically significant effects for children who were not taken care of by their parents during the program's hours at baseline. Moreover, the stronger effects were found in children left alone at home at baseline and in children who were taken care of by non-relative adults. Although the data of its prevalence is limited, informal care arrangements for children are a common strategy. The results of this study show that providing children with a safe environment might increase attendance rates and academic achievements. The most robust results were found in art grades, overall GPA, and the probability of being above the median of the grade distribution.

Although the design does not allow for the direct testing of the mechanisms driving the effects, the program characteristics and the heterogeneous analysis reported can shed light on them. First, the program only had 30 minutes for homework time per day, and most of the time was devoted to playful activities. Therefore, it was not set up as an academic ASP. Also, the quality of the program and the activities did not have a differential impact on the ASP effects. The short time devoted for academic activities and the lack of importance of the quality and type of activities show that a supervision channel might be more relevant than the activities themselves. Furthermore, the effect is most substantial when children are left alone at home at baseline, which is again consistent with the importance of adult supervision. Finally, no evidence was found to suggest that the effect is driven by the program's labor market effects on their mothers.

An after-school program can then have a positive impact on students' outcomes when the quality of the alternative care is low. Additionally, adult supervised care seems to dominate other mechanisms in the Chilean setting. Both conclusions could be considered in the

program design, targeting, and planning. As the program has a larger impact on students that are home alone, this characteristic could be included in the program targeting.

This paper points out the necessity of further research in this area. On the one hand, institutional care could, in theory, expose children to more stressful situations (long school hours, absence of free time, bullying) that in turn could generate negative impact, as reported by Baker (2008). Studying the potential effect of institutional care on socioeconomic outcomes might put at rest concerns regarding the welfare of children who spend long hours at school. On the other hand, research designs that could directly address the mechanisms underlying the program's impacts would also be relevant. Although in this paper we provide evidence consistent with a dominant supervision effect, from an academic and policy point of view it could be extremely important to directly test this hypothesis by designing, for example, a program that offers alternative ASP.

6. References

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

Bellei, Cristian (2009). "Does Lengthening the School Day Increase Students' Academic Achievement? Results from a Natural Experiment in Chile." *Economics of Education Review*, 28: 629–640

Berthelon, Matias E. and Diana I. Kruger (2011). "Risky Behavior among Youth: Incapacitation Effects of School on Adolescent Motherhood and Crime in Chile." *Journal of Public Economics*, 95(1-2) 41-5

Bettinger, Eric, Torbjørn Hægeland and Mari Rege (2014). "Home with Mom: The effects of stay-at-home parents on children's long-run educational outcomes." *Journal of Labor Economics*, 2014, vol. 32, issue 3, 443–467

Cascio, Elizabeth U. (2009). "Maternal Labor Supply and the Introduction of Kindergartens into American Public Schools." *Journal of Human Resources*, 44 (1): 140-170

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

Clarke, Damian (2016). "RWOLF: Stata module to calculate Romano-Wolf stepdown p-values for multiple hypothesis testing", Statistical Software Components, Boston College Department of Economics.

Dwyer, K.M., Richardon, J.L., Daley, K.L. et. al. (1990). "Characteristics of eighth grade students who initiate self-care in elementary and junior High School." *Pediatrics*, 86: 448– 454.

Cornelissen, T., C. Dustmann, A. Raute, and U. Schönberg (2017). "Who Benefits from Universal Childcare? Estimating Marginal Returns to Early Childcare Attendance." Forthcoming *Journal of Political Economy*.

Felfe, Christina & Zierow, Larissa (2014). "After-School Care and Children's Cognitive and Non-Cognitive Skills." *The B.E. Journal of Economic Analysis and Policy*, 14 (4): 1299-1336.

Gelbach, J.B (2002). "Public Schooling for Young Children and Maternal Labor Supply." *American Economic Review*, 92 (1):307-322.

Goerlich Susan, Sherri Lauver and Rebecca A. Maynard (2006). "Impacts of After-School Programs on Students Outcomes" *Campbell Systematic Reviews*.

Havnes, T. and M. Mogstad (2011). "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy*, 3 (2):97–129.

Havnes, T. and M. Mogstad (2015). "Is Universal Child Care Leveling the Playing Field?" *Journal of Public Economics,* 127:100–114.

Kremer, Kristen P., Brandy R. Maynard, Joshua R. Polanin, Michael G. Vaughn and Christine M. Sarteschi (2015). "Effects of After-School Programs with At-Risk Youth on Attendance and Externalizing Behaviors: A Systematic Review and Meta-Analysis" *Journal of Youth Adolescence*, March; 44(3): 616-636.

Lefebvre, P. and Merrigan, P (2008). "Child-Care Policy and the Labor Supply of Mothers with Young Children: A Natural Experiment from Canada." *Journal of Labor Economics*, 26(3): 519-48.

Levine, Phillip B. and David J. Zimmerman (2010) "After-School Care", in *Targeting Investments in Children: Fighting Poverty When Resources are Limited*. University of Chicago Press

Laughlin Lynda (2013) "Who's Miniding the Kids? Child Care Arrangements: Spring 2011". U.S. Department of Commerce, Economics and Statistics Administration, U.S. Census Bureau, April, pp. 70-135.

Martínez A., Claudia and Marcela Perticará (2017). "Childcare Effects on Maternal Employment: Evidence from Chile", *Journal of Development Economics*, 126:127-137.

McCombs, Jennifer, Anamarie Whitaker and Paul Yoo (2017). "The Value of Out-of School Time Programs". *RAND. Perspectives. Expert Insights on a Timely Policy Issue*.

Romano, Joseph and Michael Wolf. (2005). "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica*, 73(4): 1237–1282.

Ruiz-Casares, Mónica, José Ignacio Nazif-Muñoz, René Iwo and Youssef Oulhote (2018) "Nonadult Supervision of Children in Low and Middle-Income Countries: Results from 61 National Population-Based Surveys". *International Journal of Environmental Research and Public Health*, 15:1564.

Shaeffer, Sheldon (2015). "The Demand for and the Provision of Early Childhood Services since 2000: Policies and Strategies", Paper commissioned for the EFA Global

Monitoring Report 2015, Education for All 2000-2015: Achievements and Challenges, 35 p.

Winner, E., and Hetland, L. (2000). The Arts in Education: Evaluating the Evidence for a Causal Link. Journal of Aesthetic Education, 34(3/4), 3-10. Retrieved from http://www.jstor.org/stable/3333

Table 1: Compliance Rates

	Base Line	In follow-up	Participating	Participation Rate
	[1]	[2]	[3]	[4]=[3]/[2]
Control	1.208	1.073	267	0,25
Treatment	1.358	1.184	668	0,56
Total	2.566	2.257	935	

Note: Columns [1] and [2] indicate the number of applicants who were surveyed at the baseline and follow-up. Column [3] presents the number of applicants who report having participated in the program (take-up).

Table 2: Balance between treatment and control group at baseline

Variables	Average [1]	SD [2]	Nº [3]	Treatment [4]	Control [5]	P-value T=C [6]
			Panel A: Students			
Age	9.72	2.26	2,566	9.76	9.68	0.424
Female	0.47	0.50	2,566	0.47	0.47	0.352
Grade	4.04	2.03	2,557	4.06	4.03	0.775
=1 if attend school where the program is given	0.52	0.50	2,566	0.5	0.53	0.553
GPA (previous year)	5.59	0.65	2,014	5.58	5.6	0.564
GPA (previous year) is missing	0.22	0.41	2,566	0.22	0.21	0.671
Attendace rate (previous year)	0.89	0.13	2,379	0.89	0.89	0.656
Attendance rate (previous year) is missing	0.07	0.26	2,566	0.07	0.07	0.911
=1 if uses non parental childcare at baseline	0.57	0.50	2,105	0.55	0.59	0.732
=1 if child is taken care of by an adult at baseline	0.38	0.49	2,105	0.38	0.38	0.408
=1 if child is left alone at home at baseline	0.11	0.31	2,105	0.10	0.11	0.492
=1 if child is left with siblings at baseline	0.09	0.28	2,105	0.07	0.09	0,32
			Panel B: Mothers			
Age	36.89	8.55	2,561	36.92	36.87	0.821
=1 if household head	0.53	0.50	2,566	0.52	0.54	0.867
# of children	2.19	1.16	2,566	2.19	2.18	0.950
Years of education	9.37	3.22	2,482	9.35	9.39	0.822
Per capita income of household (US$)	116	86	2,544	117	116	0.287
Works and children <5 years old	0.20	0.40	2,566	0.20	0.20	0.246
Does not work and children <5 years old	0.06	0.23	2,566	0.06	0.06	0.679
Works and children >5 years old	0.63	0.48	2,566	0.63	0.62	0.343
Does not work and children >5 years old	0.11	0.32	2,566	0.11	0.12	0.680

Note: Baseline survey data collected from March to May 2012. The sample size varies according to the amount of data without observations for each respective variable. Income variable is measured in U$ dollars (march 2013). Columns [1], [2] and [3] show the variable mean for the total of the sample, the standard deviation and the number of observations, respectively. Column [4] and [5] show the variable mean for the treatment and control group, respectively. Column [6] the p-value of the null hypothesis that Treatment=Control.

Table 3: Attrition and Baseline Characteristics

	In final regressions (attendance)			In final regressions (grades)		
	[1]	[2]	[3]	[4]	[5]	[6]
Treatment (T)	-0.001	-0.005	0.030	-0.008	-0.008	-0.086
	(0.012)	(0.011)	(0.053)	(0.012)	(0.011)	(0.072)
Gender		-0.008	-0.018		-0.010	-0.004
		(0.013)	(0.017)		(0.015)	(0.024)
Age		0.004	0.005		0.017***	0.012***
		(0.003)	(0.003)		(0.003)	(0.004)
Mother's Age		-0.002*	-0.002		-0.002	-0.001
		(0.001)	(0.002)		(0.001)	(0.002)
Household Head		-0.002	0.009		-0.020	-0.017
		(0.016)	(0.019)		(0.015)	(0.023)
# of kids		-0.005	-0.010		-0.012	-0.017
		(0.008)	(0.010)		(0.010)	(0.021)
=1 if any kind of childcare		-0.008	0.002		0.000	-0.001
		(0.017)	(0.017)		(0.019)	(0.027)
Per-capita income in household		-0.000	-0.000		-0.000	-0.000
		(0.000)	(0.000)		(0.000)	(0.000)
Constant		1.005***	0.994***		0.855***	0.898***
		(0.073)	(0.084)		(0.075)	(0.106)
Observations	2,566	2,014	2,014	2,566	2,014	2,014
R-squared	0.104	0.135	0.136	0.078	0.125	0.126
F-test: all interactions with treatment are zero (p-value)			0.552			0.320
Controls	No controls	Controls	Controls and interactions of controls and treatment variable	No controls	Controls	Controls and interactions of controls and treatment variable

Note: The dependent variable takes a value of 1 if the individual was found on either attendance data (columns [1]-[3]) or grades data (columns [4]-[6]). The sample are all students participating in the study (with baseline). The sample size varies according to the missing covariate data. Standard error in brackets. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Table 4: Intent-to-Treat Effects in Attendance and Grade (2012)

	Panel A: Attendance			Panel B: Grades						
	Attendance rate May-November	=1 if attendace rate is >0.90	=1 if attendace rate is >0.95	Art	Physical Education	Language and Literature	Math	Science	GPA	=1 if above the median
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]
Treatment	0.006	-0.002	0.015	0.043	0.055**	0.010	0.030	0.012	0.020	0.016
	(0.005)	(0.024)	(0.026)	(0.029)	(0.026)	(0.032)	(0.032)	(0.027)	(0.022)	(0.023)
Observations	2,379	2,379	2,379	2,280	2,277	2,280	2,280	2,280	2,284	2,284
R-squared	0.276	0.231	0.218	0.309	0.276	0.372	0.348	0.396	0.488	0.359
Control group mean	0.907	0.677	0.365	5.926	6.250	5.134	5.149	5.231	5.532	0.494

Note: Columns [1] - [9] report the intent-to-treat (ITT) estimates and standard errors (in parenthesis) of program assignment. The sample size varies according to the number of observations with missing values in the respective outcome variable. This sample is obtained merging both baseline and administrative data. All regressions include school-strata fixed effects and control for age. Cluster standard errors at school level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Heterogenous Effects by Childcare use at baseline (2012)

		Attendance			Grades						
	First Stage Program Participation	Attendance rate May-November	=1 if attendace rate is >0.90	=1 if attendace rate is >0.95	Art	Physical Education	Lenguage and Literature	Math	Science	GPA 2012	=1 if above the median
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
	Panel A: Parental versus non-parental care										
Treatment	0.255***	-0.001	-0.067*	-0.070**	-0.030	-0.014	-0.076	-0.029	-0.051	-0.050	-0.032
	(0.056)	(0.006)	(0.036)	(0.027)	(0.052)	(0.045)	(0.050)	(0.069)	(0.063)	(0.038)	(0.029)
Treatment * Non-parental care at baseline	0.058	0.012	0.114*	0.110**	0.147**	0.100	0.128*	0.073	0.106	0.123**	0.086**
	(0.059)	(0.009)	(0.059)	(0.052)	(0.054)	(0.069)	(0.071)	(0.082)	(0.080)	(0.053)	(0.037)
Observations	2,131	2,379	2,379	2,379	2,280	2,277	2,280	2,280	2,280	2,284	2,284
R-squared	0.230	0.277	0.234	0.223	0.313	0.280	0.373	0.349	0.398	0.489	0.362
Control group mean	0.254	0.907	0.677	0.365	5.926	6.250	5.134	5.149	5.231	5.532	0.494
p-value full effect	0.000	0.105	0.293	0.349	0.001	0.069	0.239	0.347	0.174	0.044	0.121

		Attendance			Grades						
	First Stage Program Participation	Attendance rate May-November	=1 if attendace rate is >0.90	=1 if attendace rate is >0.95	Art	Physical Education	Lenguage and Literature	Math	Science	GPA 2012	=1 if above the median
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
	Panel B: By type of care										
Treatment	0.254***	-0.001	-0.068*	-0.070**	-0.031	-0.014	-0.076	-0.030	-0.053	-0.051	-0.032
	(0.057)	(0.006)	(0.036)	(0.027)	(0.052)	(0.045)	(0.050)	(0.069)	(0.063)	(0.038)	(0.029)
Treatment * Other adults	0.028	0.007	0.089	0.056	0.147**	0.125	0.135	0.109	0.159*	0.143**	0.090**
	(0.066)	(0.009)	(0.068)	(0.059)	(0.071)	(0.077)	(0.079)	(0.089)	(0.091)	(0.061)	(0.039)
Treatment * Siblings	0.083	0.009	0.114	0.095	0.040	-0.017	-0.020	-0.127	-0.110	-0.044	0.005
	(0.098)	(0.016)	(0.083)	(0.068)	(0.092)	(0.100)	(0.145)	(0.171)	(0.134)	(0.105)	(0.079)
Treatment * Alone	0.127	0.032***	0.203**	0.307***	0.227***	0.110	0.217*	0.119	0.104	0.187***	0.134*
	(0.090)	(0.011)	(0.080)	(0.074)	(0.068)	(0.090)	(0.114)	(0.097)	(0.096)	(0.056)	(0.077)
Observations	2,131	2,379	2,379	2,379	2,280	2,277	2,280	2,280	2,280	2,284	2,284
R-squared	0.232	0.280	0.235	0.227	0.313	0.281	0.374	0.351	0.400	0.491	0.363
Control group mean	0.254	0.907	0.677	0.365	5.926	6.250	5.134	5.149	5.231	5.532	0.494
P-values full effect											
Other Adult	0.000	0.416	0.689	0.764	0.022	0.037	0.323	0.195	0.052	0.047	0.163
Alone	0.000	0.002	0.061	0.003	0.002	0.243	0.150	0.441	0.635	0.041	0.184

Note: Columns [2] - [10] report the intent-to-treat (ITT) estimates and standard errors (in parenthesis) of program assignment. Column [1] reports the firts stage por program participation. The sample size varies according to the number of observations with missing values in the respective outcome variable. In panel A, non-parental care is a dummy variable that takes value of 1 for all the kids who weren`t taken care of by their parents at baseline, zero otherwise. In panel B, the base cateogory is taken care of by parents. All regressions include school-strata fixed effects and control for age. Cluster standard errors at school level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Heterogenous Effects by program quality

	Panel A: Attendance				Panel B: Grades						
	First Stage Program Participation	Attendance rate May-November	=1 if attendace rate is >0.90	=1 if attendace rate is >0.95	Art	Physical Education	Lenguage and Literature	Math	Science	GPA	=1 if above the median
	[1]	[2]	[3]	[4]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
(1) Above median quality index (Mean= 0.579)	0.032	0.003	0.075*	0.034	0.051	-0.012	-0.016	0.040	0.032	0.030	0.054
	(0.075)	(0.011)	(0.042)	(0.052)	(0.056)	(0.049)	(0.072)	(0.066)	(0.052)	(0.045)	(0.044)
p-value full effect	0.000	0.235	0.546	0.525	0.123	0.211	0.902	0.417	0.650	0.349	0.288
(2) At least 25% of monitors are school teachers (Mean= 0.226)	-0.021	0.013	-0.017	-0.056	-0.013	0.034	-0.029	-0.073	0.024	-0.008	-0.001
	(0.071)	(0.014)	(0.064)	(0.048)	(0.084)	(0.072)	(0.080)	(0.062)	(0.058)	(0.057)	(0.075)
p-value full effect	0.000	0.221	0.688	0.373	0.731	0.219	0.766	0.493	0.573	0.817	0.945
(3) Monitors with above the median experience (4 yrs) (Mean= 0.497)	0.045	-0.002	-0.012	-0.014	-0.035	0.041	-0.048	-0.037	0.004	0.006	0.001
	(0.080)	(0.010)	(0.049)	(0.055)	(0.057)	(0.054)	(0.067)	(0.067)	(0.057)	(0.046)	(0.046)
p-value full effect	0.000	0.422	0.666	0.932	0.659	0.089	0.637	0.949	0.786	0.592	0.878
(4) Program components defined by march (Mean= 0.697)	-0.019	0.006	0.063	0.012	0.096*	0.011	0.054	0.085	0.013	0.050	0.099**
	(0.095)	(0.006)	(0.037)	(0.055)	(0.052)	(0.046)	(0.066)	(0.068)	(0.046)	(0.041)	(0.043)
p-value full effect	0.000	0.501	0.792	0.754	0.052	0.161	0.773	0.257	0.966	0.319	0.087
(5) Fixed time slot devoted to study (Mean= 0.801)	0.004	-0.023	-0.046	-0.069	-0.011	-0.062	-0.080	0.061	-0.061	-0.028	0.050
	(0.099)	(0.019)	(0.039)	(0.052)	(0.050)	(0.051)	(0.087)	(0.085)	(0.071)	(0.053)	(0.073)
p-value full effect	0.000	0.518	0.791	0.758	0.053	0.162	0.774	0.259	0.967	0.320	0.091
(6) Planification is closely followed (Mean= 0.420)	-0.009	-0.004	-0.040	0.015	0.029	-0.005	0.014	-0.020	-0.019	0.001	0.011

	(0.087)	(0.010)	(0.047)	(0.056)	(0.055)	(0.053)	(0.067)	(0.066)	(0.057)	(0.044)	(0.045)
p-value full effect	0.001	0.585	0.218	0.645	0.031	0.171	0.851	0.799	0.947	0.501	0.718
(7) Students/Monitor ratio is below the median (Mean= 0.560)	-0.203**	-0.002	-0.049	-0.048	-0.008	-0.013	-0.038	-0.089	0.016	-0.030	-0.115***
	(0.074)	(0.006)	(0.051)	(0.053)	(0.061)	(0.052)	(0.066)	(0.070)	(0.058)	(0.046)	(0.039)
p-value full effect	0.000	0.899	0.323	0.724	0.422	0.204	0.442	0.593	0.915	0.803	0.124

Note: Columns [2] - [10] report the intent-to-treat (ITT) estimates and standard errors (in parenthesis) of program assignment interacted with different measures of the program quality. Column [1] reports the firts stage por program participation. The sample size varies according to the number of observations with missing values in the respective outcome and quality variable.

The different hight quality dummies are defined in the following way:
(1) Above median quality index. The quality index is defined including measures of infrastructure, materials and monitor quality as reported in the process evaluation. The index does not include measures related to children's behavior. (2) At least 25% of monitors are school teachers. (3) On average, monitors have at least 4 years (median) of experience. (4) By march, the program components were already defined. (5) There was a fixed time slot devoted to study. (6) Planification (as describe in the original proposal) is closely followed (all the observed activities are described in the original plan). (7) Students/Monitor ratio is below the median.

All regressions include school-strata fixed effects and control for age. Cluster standard errors at school level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Heterogenous Effects by workshop theme

	Panel A: Attendance				Panel B: Grades						
	First Stage Program Participation	Attendance rate May-November	=1 if attendace rate is >0.90	=1 if attendace rate is >0.95	Art	Physical Education	Lenguage and Literature	Math	Science	GPA	=1 if above the median
	[1]	[2]	[3]	[4]	[6]	[7]	[8]	[9]	[10]	[11]	[12]
(1) At least one TICs course (Mean= 0.286)	-0.121	-0.006	0.042	0.005	0.075	0.015	0.023	0.110	0.048	0.047	0.032
	(0.073)	(0.009)	(0.061)	(0.060)	(0.045)	(0.061)	(0.054)	(0.074)	(0.052)	(0.045)	(0.042)
p-value full effect	0.001	0.748	0.634	0.730	0.003	0.228	0.405	0.113	0.283	0.155	0.182
(2) At least one science (including social sciences) course (Mean= 0.543)	0.013	0.006	-0.030	0.025	0.052	0.028	0.069	0.020	0.022	0.026	0.052
	(0.078)	(0.009)	(0.051)	(0.052)	(0.063)	(0.053)	(0.061)	(0.069)	(0.057)	(0.048)	(0.045)
p-value full effect	0.000	0.232	0.545	0.439	0.011	0.035	0.327	0.288	0.544	0.186	0.195
(3) At least one personal care course (Mean= 0.425)	-0.008	-0.005	0.015	0.008	-0.002	0.034	-0.090	0.069	0.051	0.003	0.009
	(0.077)	(0.008)	(0.049)	(0.050)	(0.063)	(0.050)	(0.058)	(0.066)	(0.051)	(0.046)	(0.049)
p-value full effect	0.000	0.222	0.883	0.576	0.450	0.036	0.277	0.213	0.196	0.539	0.604
(4) At least one sport course (Mean= 0.771)	-0.025	0.000	0.013	-0.042	-0.019	-0.043	-0.088	0.032	-0.044	-0.038	0.009
	(0.121)	(0.017)	(0.041)	(0.055)	(0.062)	(0.041)	(0.070)	(0.065)	(0.053)	(0.046)	(0.052)
p-value full effect	0.000	0.160	0.998	0.823	0.268	0.151	0.838	0.353	0.918	0.626	0.503

Note: Columns [2] - [10] report the intent-to-treat (ITT) estimates and standard errors (in parenthesis) of program assignment interacted with workshop themes dummies. Column [the firts stage por program participation. The sample size varies according to the number of observations with missing values in the respective outcome and workshop theme.

Note that at each school children could attend different kind of workshops. Workshop themes dummies are defined to indicate whether at each school at least one workshop was of seven different areas: sciences or more academic subjects; personal care; sports, arts, drama/dance/circus; arts; music.

All regressions include school-strata fixed effects and control for age. Cluster standard errors at school level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Mother's outcomes. Heterogenous Effects by Childcare use at baseline

	Labor Force Participation			Employment			Working Hours	Income	
	Participates (at least one month during May-Dec)	Participates (always)	Months Participating (May-Dec)	Works (at least one month during May-Dec)	Works (always)	Worked Months		Monthly Income	Hourly Income
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
	Panel A: Parental versus non-parental care								
Treatment	-0.043	-0.003	-0.180	-0.002	-0.017	0.193	1.167	26.018*	0.192
	(0.029)	(0.036)	(0.236)	(0.028)	(0.033)	(0.287)	(1.536)	(13.885)	(0.164)
Treatment * Non-parental care at baseline	0.058	0.014	0.291	0.003	0.017	0.145	-1.848	-15.589	0.148
	(0.037)	(0.036)	(0.219)	(0.035)	(0.039)	(0.313)	(1.961)	(21.177)	(0.329)
Observations	1,874	1,874	1,874	1,874	1,874	1,655	1,594	1,567	1,525
R-squared	0.150	0.163	0.173	0.160	0.161	0.225	0.184	0.174	0.122
Control group mean	0.682	0.546	4.884	0.648	0.475	6.245	28.396	263.863	1.828
p-value full effect	0.654	0.794	0.686	0.977	0.999	0.263	0.687	0.489	0.178

	Panel B: By type of care								
	Labor Force Participation			Employment			Working Hours	Income	
	Participates (at least one month during May-Dec)	Participates (always)	Months Participating (May-Dec)	Works (at least one month during May-Dec)	Works (always)	Worked Months		Monthly Income	Hourly Income
	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]
Treatment	-0.043	-0.003	-0.182	-0.002	-0.017	0.192	1.167	25.791*	0.193
	(0.029)	(0.036)	(0.235)	(0.028)	(0.033)	(0.285)	(1.527)	(13.922)	(0.166)
Treatment * Other adults	0.043	0.009	0.229	-0.022	-0.006	0.009	-2.931	-10.420	0.093
	(0.052)	(0.043)	(0.330)	(0.047)	(0.046)	(0.393)	(2.015)	(20.693)	(0.322)
Treatment * Siblings	0.167**	0.070	0.906	0.111	0.069	0.652	3.695	14.682	0.485
	(0.079)	(0.099)	(0.625)	(0.078)	(0.098)	(0.701)	(4.608)	(44.505)	(0.478)
Treatment * Alone	0.027	-0.006	0.076	0.005	0.057	0.239	-2.445	-50.316	0.073
	(0.055)	(0.081)	(0.515)	(0.063)	(0.104)	(0.741)	(3.753)	(35.698)	(0.481)
Observations	1,874	1,874	1,874	1,874	1,874	1,655	1,594	1,567	1,525
R-squared	0.154	0.166	0.177	0.164	0.163	0.227	0.186	0.176	0.122
Control group mean	0.682	0.546	4.884	0.648	0.475	6.245	28.396	263.863	1.828
P-values full effect									
Other Adult	0.993	0.919	0.795	0.647	0.681	0.480	0.388	0.427	0.546
Alone	0.725	0.903	0.922	0.955	0.644	0.694	0.691	0.362	0.316

Note: Columns [2] - [10] report the intent-to-treat (ITT) estimates and standard errors (in parenthesis) of program assignment. Column [1] reports the firts stage por program participation. The sample size varies according to the number of observations with missing values in the respective outcome variable. In panel A, non-parental care is a dummy variable that takes value of 1 for all the kids who weren`t taken care of by their parents at baseline, zero otherwise. In panel B, the base cateogory is taken care of by parents. All regressions include school-strata fixed effects and control for age. Cluster standard errors at school level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.