Can information and assistance increase application and access to formal childcare: Experimental evidence from France

A pre-registration report

Laudine Carbuccia* Carlo Barone† Coralie Chevallier‡ Arthur Heim§

21 octobre, 2023

Abstract

This trial assesses the impact of an intervention aimed at improving demand for and access to early childcare services (ECS) in France. The literature reports that attendance of high-quality ECS has positive effects both on women's employment and on children's development and school achievement, more so for children of low socioeconomic status (SES) families. However, these children are strongly underrepresented in ECS, especially in France. We hypothesize that structural barriers relating to the accessibility of ECS (rationing of places and criteria of access) and to their costs for the families only partly explain the lower demand for and access of low-SES families. Relying on a qualitative preparatory study in France, we expect that a lack of information about these services as well as behavioral and administrative barriers in the application process contribute to unequal access. Therefore, we designed an intervention targeting pregnant mothers and delivering information and support about the availability, costs, eligibility criteria of ECS and the related application procedures. We contact mothers during their visits to maternity wards of hospitals in the metropolitan area of Paris (Ile de France) and administer a baseline questionnaire to collect information on socio-demographic characteristics, knowledge of ECS, and intentions to use them. After this interview, they are randomly assigned to three experimental arms (control, information-only treatment, information plus administrative support treatment). One year later we administer a phone interview on actual recourse and two years later a follow-up

Contents

I Introduction **4**

II Motivations context and intervention **5**

1 Motivations: Early childhood intervention and social inequalities 5

2 The interventions: information and social support to foster access to formal childcare 6

III Research questions **8**

1 Definitions 8

2 Implication for the analysis 8

IV A theory of change and additional research questions **9**

*CRIS/LIEPP - Sciences Po & ENS PSL
†CRIS/LIEPP - Sciences Po
‡ENS PSL
§Paris school of Economics & Cnaf

V Research strategy **11**

1 Data sources . . . 11

A Baseline and endline survey . . . 11

B Power analyses and sample size determination . . . 11

2 Notations . . . 12

A Samples and observables . . . 12

B Treatment variable and potential outcomes . . . 13

3 Experimental design and identification . . . 13

A Assumptions . . . 13

B Target parameters . . . 15

4 Estimation strategy . . . 17

A Roadmap . . . 17

B Inference . . . 18

C Baseline data and non-response . . . 18

D Attrition . . . 18

E Models for intention to treat parameters . . . 19

F Models for T2 against T1 . . . 19

VI Analysis and hypothesis testing **21**

1 Proximate outcomes: application behaviors . . . 21

A Definition of outcomes and set of covariates . . . 21

B Treatment effects on Parents' likelihood to apply . . . 21

C Number of applications . . . 23

D Treatment effects on parents' timing of application . . . 23

2 Final outcomes : Effects on childcare use . . . 23

A Definition of outcomes and set of covariates . . . 23

B Parents' access to ECS . . . 24

3 Intermediate outcomes: parents level of information and perception of ECS . . . 24

A Definition of outcomes and set of covariates . . . 24

B Parents' level of information about ECS . . . 25

4 Cross Hypotheses: Heterogeneity . . . 25

5 Exploratory analyses: . . . 27

VIIData collection **28**

1 Collected variables . . . 28

A Baseline survey . . . 28

B Endline survey . . . 29

2 Exclusion rules . . . 31

VIIData processing and security **31**

IX Research Team **31**

X Bibliography **32**

I Introduction

Early childcare services (ECS) play a critical role in shaping the lives of children and parents. First, they can increase labour market participation of mothers (Kimmel 1998; Nollenberger and Rodríguez-Planas 2015; Morrissey 2017; Simintzi, Xu, and Xu 2022; Hermes et al. 2022; Huber and Rolvering 2023). Second, high-quality early childhood interventions foster cognitive and socio-emotional development, improving educational attainment, and generate long-term positive outcomes in various dimensions (Case, Fertig, and Paxson 2005; Cunha et al. 2006; Camilli et al. 2010; Burger 2010; Nores and Barnett 2010; Johnson and Schoeni 2011; Barnett 2011; Attanasio 2015; Walters 2015; Elango et al. 2016; Kholoptseva 2016).

Recently, van Huizen and Plantenga (2018) systematically synthesized quasi-experimental literature assessing the effects of access to childcare services between 2005 and 2017 across developed countries, analyzing 250 estimates derived from 30 studies that encompass a range of child development outcomes. These outcomes span from cognitive and noncognitive development during early childhood to educational achievement and adult income. This meta-analysis presents important results for childcare policy design. In particular, entry age does not seem to significantly influence outcomes, whereas the intensity of care play a role in certain cases. Childcare services offered through public provision have more pronounced effects on child development than private services. However, the most noteworthy finding is that the positive effects are primarily concentrated among children from socioeconomically disadvantaged backgrounds.

If policymakers want early childcare education to act as a means of reducing inequalities, they must ensure that those with the highest potential benefits can easily access it. Unlike the growing body of research highlighting the advantages of early childhood interventions, the assessment of universal childcare policies doesn't always yield the anticipated outcomes (Cantillon 2011; Ghysels and Van Lancker 2011; Vandenbroucke and Vleminckx 2011; Duncan et al. 2012). In most OECD countries, there is an socio-economic gap in early childcare enrollment, meaning that low-socioeconomic status (SES) parents access these structures the least (OECD 2016).

In this research, we investigate the barriers to formal childcare through a multi-arm experiment testing two easily reproducible policies. In one treatment arm, the policy consists in providing information to parents through simple text-messages emphasizing important features of formal childcare and links to more detailed information. In the second treatment arm, we also offer administrative support to register to formal childcare.

France has a dual system where ECS is separated from preschool before entering elementary school. ECS are accessible to children as early as three months old, and preschool education is mandatory by the age of three. France has a relatively high coverage rate in ECS compared to other OECD countries, but this access is also one of the most unequal according to socio-economic background of families (OECD 2016; Carbuccia et al. 2020). The french ECS landscape is complex, with a mix of public, private, and associative providers, and a variety of ECS types. The most available type of ECS is private childminders, that account for 55.3% of ECS places in 2019 (ONAPE 2020). Private childminders are also more expensive than public and associative daycare for low-SES parents. Public daycare applications are most often centralized at the level of the municipality or the district, while there is no centralization of applications for private or associative structures. Parents have to do one application per center for the latter. In a context of limited availability, parents should apply as early as possible during pregnancy and to as much ECS as possible since structures often have long waiting lists. Furthermore, the allocation of places in daycare centers follows a precise calendar that aligns with the school year. Since daycare centers operate on the school model with groups by age and because maternity leaves end two months and a half after birth, most places are allocated around June for households with a baby under one year old to start care in September. It is much less likely to obtain a place for a child over one year old.

This paper is a detailed pre-analysis plan of the effect of these interventions. It supplement the description provided with the pre-registration of the intervention and main outcomes on the AEA social science registry[1] (www.socialscienceregistry.org/trials/9901)). This experiment has been approved by the IRB of Paris School of Economics[2].

We formally explain under which conditions our design allows to measure each treatment's causal effects and better understand what factor affects decisions making and access to formal childcare. Ultimately, our main objective is to

1. RCT ID AEARCTR-0009901
2. IRB Number: 2022-015

test interventions to influence the demand size and structure, measure their effects on parents' search and application choices and behaviours, use of formal childcare, and uncover some plausible causal mechanisms.

After presenting our motivations and details of the interventions in section II, we define and discuss our main research questions in section III and propose a theory of change in section IV to guide our modelling assumptions, statistical analysis and interpretations. In section V, we formalize our target parameters, identification strategy and econometrics models. In section VII, we define the outcomes and measurement and finally, we pre-define our choices on the statistical analysis in section VI.

II Motivations context and intervention

1 Motivations: Early childhood intervention and social inequalities

Affordability and Accessibility are important supply side factors that can block access to early childcare. On the one hand, the costs of ECS can be too high for low-SES families. On the other hand, other factors may prevent families from accessing ECS, even when parents could afford it. For instance, access to public daycares is heavily rationed and governments are funding more and more facilities. However, increasing supply for one type of childcare can have undesirable consequences. In France, Pora (2020) uses large but staggered increases in daycare provision across municipalities and time to estimate the effect of these apparent supply shocks on mothers' labour market participation. On average, he finds no effect on labour market participation but documents a crowding-out effect on registered childminders. These results have two implications. First, the childcare market offers different services in different segments, and the demand can shift from one segment to another. Second, the absence of an average effect does not mean that childcare has no effect on women's participation in the labour market. If working gives higher priority in daycare assignment mechanisms, then there is no room for an effect on mothers' labour market participation, at least at the extensive margin. Instead, these results suggest a general equilibrium effect in which women attached to the labour market shifts from childminders to daycare and manage to secure access for their children. In France, daycare centers have an objective comparative advantage in that they are, in general, the most affordable ECS solution. They also have a shared perceived comparative advantage over children socialisation, and researchers have documented that parents prefer collective childcare to childminders (Bouteillec, Kandil, and Solaz 2014; Cartier et al. 2017 ; Caenen and Virot 2023). Those who do not work (yet) are less likely to need childminders, and a lower demand brings a new equilibrium with a lower supply.

Surprisingly, this simple supply/demand/equilibrium reasoning is very much overlooked. However, it provides clear guidance on how policymakers may improve welfare and where interventions are likely to make things worse. Economics 101 tells us that going against market forces is not easy, and usually backfires. At the same time, letting markets free when there are strong externalities usually lead to welfare loss. In the childcare market, tax subsidies and transfers already introduce distortions and may not be incentive-compatibles (Lancker and Ghysels 2012). Especially when for-profit companies provide such services (Penn 2007; Noailly and Visser 2009). There may also be *optimisation frictions* because information is not easily accessible, application process and tax-subsidies scheems are complex and so on. Because of heterogeneity in the effects of childcare on different populations, "who gets what and why" matters (Roth 2015). The *quality* of the matching matters. Markets may be noxious because some participants are more vulnerable than others (Satz 2010). However, various interventions can lead to alternative potentially welfare-improving equilibria. One can even design better marketplaces. The recent work by Combe and Heim (2023) models daycare assignment committees as matching markets. Their work shows that policymakers must *define* and *choose* how they consider the problem (with or without days), how they partition the supply across daycares (over weekdays) for different groups of children, and which definition of *fairness* they want to follow. A *fair* assignment is defined over a strict respect of priority rules or versions that tolerate some harmless deviations to increase the number of children assigned. They emphasize the trade-offs between different political objectives and the implication of each choice. *Fair* assignments ensure a form of *procedural justice* (Binns et al. 2018) and allow *transparency* as every decision can be justified individually. Then, they show that there exist an assignment mechanism for each version of the problem that returns the most preferred assignment by all families that respects the definition of fairness, multidimensional constraints and diversity quotas. Well defined priority criteria and diversity constraints offer powerful tools to enforce objectives over the distribution of assignments. However,

as Li (2017) points out, *"Market design have consequences. A consequence is richer than an outcome, it's the effect of the market on the world"*. In particular, the current market structures partly determine the composition of the demand. The lack of transparency and administrative burden may deter some families from even trying to register.

Behavioral sciences 101 suggest that market complexity inversely affects access to Early Childhood Services (ECS) for low-SES populations (see Carbuccia et al. (2023) for an extensive review). Low-SES parents face difficulties on the demand side (Pavolini and Van Lancker 2018), making it less likely for them to secure ECS placements. Several reasons contribute to this phenomenon.

First, low-SES families tend to exhibit a lower intention to apply for ECS for various reasons. Limited peer networks in the use of formal childcare lead to incomplete information about ECS during their decision-making (Carbuccia et al. 2023 ; Lazzari 2012; Vandenbroeck and Lazzari 2014), causing misperceptions about supply, such as overestimating costs and underestimating eligibility (Carbuccia 2023).

Second, decision-making factors exacerbate the information deficit. Parents, in part, rely on fast and frugal heuristics for childcare decisions, and the heuristic of adhering to social norms yields different outcomes based on social class (Chaudry, Henly, and Meyer 2010; Sandstrom and Chaudry 2012). In low-SES communities, where ECS usage is less common, following social norms often results in not using ECS (Abrassart and Bonoli 2015). Conversely, this is the opposite for high-SES populations where using ECS is the norm.

Additionally, as the opportunity cost of staying at home and the career impact is lower for low-SES individuals, they are more likely to stick to the default situation of self-caring for their child. Lastly, ECS entails a cost in the present for uncertain future benefits, especially for those not currently employed. Low-SES populations have been shown to be more present-biased and risk-averse, and tend to be more hesitant in such situations (Tversky and Kahneman 1974 ; Pepper and Nettle 2017).

On the other hand, even when low-SES parents do intend to apply for a place at a ECS, they may also face more behavioral barriers in their application process. First, because low-SES parents are less familiar with administrative procedures and vocabulary, they may face higher friction costs when applying and therefore be less likely to complete the application process (Sunstein 2022). Secondly, because their cognitive resources may be more focused on more pressing issues during pregnancy and the early years (e.g. paying bills, housing) and they are less informed about the tips and tricks, they may apply less efficiently, i.e. *later* and to fewer structures, which is detrimental in the context of low-supply. Since low-SES parents are also less likely to benefit from influential social resources, the more complex and obscure the system is, the higher the barriers for low-SES populations. In this context, helping low-SES parents navigate the system by delivering essential information and by directing help in the application process should reduce at least the SES gap in ECS application, if not in ECS enrollment.

2 The interventions: information and social support to foster access to formal childcare

To address this gap in knowledge and address these inequalities, we conduct a randomized control trial in France with three arms: a control group and two treatment groups. The timeline of the experiment is presented in Figure 1. The first treatment group (T1) receives information about early childcare service options through an inexpensive and easily scalable design, while the second treatment group (T2) receives the same information, complemented by administrative support via phone calls. From a policy perspective, distinguishing between the effects of these two treatments is of paramount importance. T1 represents a cost-effective intervention that any policymaker would consider when confronted with the issue of the ECS access gap, which could be readily expanded if its effectiveness is demonstrated. However, given the potential presence of additional behavioral barriers preventing low-SES parents from taking action, even when they express an intention to use ECS, this treatment may yield minimal to negligible effects in closing the access gap, despite the reminders. This underscores the need to compare T1 with T2, a treatment designed to address in-depth behavioral barriers.

The aim of this study is, therefore, to evaluate the impact of these interventions on

(i) application behavior,
(ii) access to an ECS slot.

Ultimately, follow-up surveys would allow us to provide causal evidence of the impact of access to ECS on

(iii) mothers' labor market outcomes and
(iv) children's development.

Between September and December 20233, we contacted pregnant mothers (4 to 9 months of pregnancy) during their visits to maternity wards of hospitals in the metropolitan area of Paris (Ile de France) and administered a baseline questionnaire to collect information on socio-demographic characteristics, knowledge of ECS, and intentions to use them. After this interview, they were randomly assigned to three experimental arms (control, information-only treatment, information plus administrative support treatment) within blocks of pre-registered covariates. One year later we administer a phone interview on ECS use and two years later a follow-up to measure long-term impacts on child development and labor market outcomes.

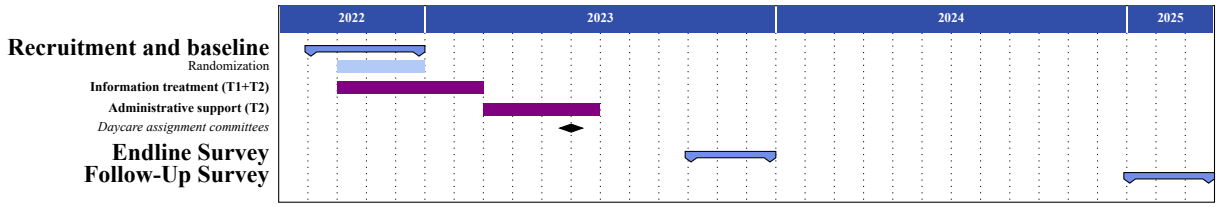
The treatment arms are as follows:

1. **Information-only** (T1) (*October-December 2022*): this involves two clusters of contents
   (a) **Cluster 1**: information aimed at helping families to identify a type of ECS that fits well with their preferences and constraints. Treated families receive a text message providing access to a short video presenting information on the availability and characteristics of different types of ECS in France and how they may correspond to the different preferences and needs of families. One text message reminder is sent to encourage mothers to watch this video. In the following days, a second text message with related reminders gives access to a second short video on eligibility to and costs of ECS in France and a third video presents detailed information on a specific type of childcare service (halte-garderies) which may be particularly well-suited to low-SES families (less intensive care, less restrictive criteria of access, easier application procedures and more flexible time schedules).
   (b) **Cluster 2**: information aimed at understanding the application process. Like for 1a, this involves sending text messages and related reminders with links to two videos presenting information on the calendar of applications, the procedures to apply, and tips to maximize the chances of success, such as applying to multiple ECS. The content of the first video of this cluster will vary depending on the district where families live, as each one has a specific application process.
   (c) **Reminders:** Reminders on Cluster 2 is sent by text messages after birth (February 2023) in order to maximize applications for the June commission, where most of the slots get allocated. We send also generic reminders to apply shortly before the deadline for the 2023 applications to ECS (April 2023).
2. **Information then administrative support treatment** (T2) (*February-April 2022*): This involves one or several phone calls with parents to deliver personalized assistance, as well as personalized application reminders. We randomly assign the sample over groups of two weeks of intervention and research assistants. During this period, we call each parent as many time as needed to reach them. We present our services to parents according to a systematized procedure. When parents show interest, we first establish a diagnosis of their choices, intentions, and needs after birth. Parents are at very different stage of their decision making and we adapt our intervention accordingly. When they have not decided yet, we help them identify the ECS solution that best fits their needs, including how accessible each solution is given their situation, and how affordable each solution is through cost simulations. When they have identified the type of ECS they want, we assist them according to their demands. Some just need help to spot the ECS structures they could apply to, while others need us fill the application forms with them. We code each type of assistance in a systematic way to qualify the intensity of assistance provided, and create a typology of the types of situation parents were in. Parents also receive the same information as T1 according to the same timeline. We call this treatment "administrative support".
3. **Control group** Parents assigned to the control group receive a *placebo* treatment. We wanted to maintain some contact to minimize attrition at endline. Theses messages were also sent to the two intervention groups for the same reason. These text messages are about
   (a) general events throughout the year (e.g., welcoming text message, winter and summer holidays, new year), and

(b) some useful tips not affecting our outcomes of interest (e.g., flea markets around Paris). Importantly, participants in the control group also receive videos, but not about ECS. The content of the videos are about emotions during pregnancy.

Importantly, neither the information nor the administrative support interventions are prescriptive about childcare choices. Our goal is to help people be informed, make choices on their own with or without guidance and assistance with formal applications.

Figure 1: Timeline of the experiment



III Research questions

1 Definitions

We designed this intervention to answer the following research questions:

1. **Can information and assistance increase applications and access to formal childcare?**
   (a) *Is this effect heterogeneous across socio-economic background ?*
2. **Does providing administrative support boost application and access compared to providing only information?**
   (a) *Is this effect heterogeneous across socio-economic background ?*

Consistent with the motivations presented in section II above, these research questions are policy oriented. Our primary objective is 1: to test if we can increase application and access to formal childcare using two easily reproducible interventions and 2: measure how much active assistance for search and application increase access compared with no intervention or simply additional information.

2 Implication for the analysis

Question 1 translates in an analysis of the average effect of our intensive treatment T2 on a measure of access to formal childcare. Question 1a assesses overall treatment effect heterogeneity by socioeconomic background. It can be estimated by comparing pooled treatment effects between blocks of low and high SES. We used a block-randomization design to ensure balance between treatment arms for this purpose.

Question 2 relates to the treatment effect heterogeneity between treatment arms. With a three arms randomized control trial, there are several comparisons we can make. T2 is an "expensive", labor-intensive policy and T1 is a "low-cost" alternative. Therefore,

- Comparisons between T1 and control tests how "*unsolicited provision of information*" (as implemented) affects application behaviors and enrollment.

- Comparisons between T2 and T1 tests how, *after the provision of unsolicited information*, being *offered optional assistance and support* to apply for formal childcare affect outcomes.
- Comparing outcomes of T2 against control tests the overall effect of information and administrative support.

The last comparison captures the effect of both treatments and their interaction. We discuss it hereafter more precisely.

Each choice regarding childcare is an important decision made under binding constraints. Informed decision making requires good knowledge of all options, benefits, costs, and also opportunity costs and benefits. The information treatment aims at helping parents defining their choice set and adjust their expectations. It makes potential benefits and costs (e.g., direct cost of a place, opportunity costs, flexibility) of using ECS clearer. It also make administrative burden more salient and potentially increase the need for help. However, some parents may not react to the information. After all, the information treatment are unsolicited which can be seen as a nuisance. They may not trust the source, or may simply missed our messages.

Those in T2 are offered an additional intermediary choice: receiving help. Receiving help is observed. We know who complies with assistance among those in T2. However, since they received information before this choice, the relevant counterfactual for T2 compliers can only be identified in comparison with the T1 sample. With random assignment within blocks, we can estimate the characteristics of those compliers based on comparisons of baseline observables with T1 (A. Abadie 2003; J. Angrist, Hull, and Walters 2023). In particular, Question 2a is tightly linked to compliance heterogeneity and the distribution of potential outcomes across high and low socio-economic background. All of which are identified using treatment assignment as instrument of receiving assistance among T2 and T1 as a control group.

However, our conclusions on questions 1, 1a, 2, and 2a cannot be solely based on the magnitude and significance of these treatment effects. There are many intermediary channels through which the interventions can increase access to childcare, or not. This research project also aim at understanding *how* and *why* we may observe some effects and when we may not. We define a simple theory of change to guide the definition of additional research questions.

IV A theory of change and additional research questions

Figure 2 represent our simple theory of change. It focus on the paths through which the treatment effects are likely mediated. At every level, the comparison between different treatment arms provide informative metrics on causal mechanisms fostering or preventing access to formal childcare. We formally define the different set of outcomes as followed:

Definition IV.1 (Intermediary outcomes)**.** The intermediary outcomes are parents level of information and perceptions of ECS. It should be noted that an increase in the level of information and/or a change in perceptions are not a necessary condition for the proximate outcomes to happen. Indeed, there is another, non measurable, channel that is behavioral barriers alleviation. Therefore, if we observe a change in our proximate outcomes but no change in our intermediate outcomes, it will indirectly imply that this change is a result of behavioral barriers alleviation only.

Definition IV.2 (Proximate outcomes)**.** The proximate outcomes are parents' intention to apply to ECS and application behaviors as measured by their attempts to access ECS, the timing and number of applications.

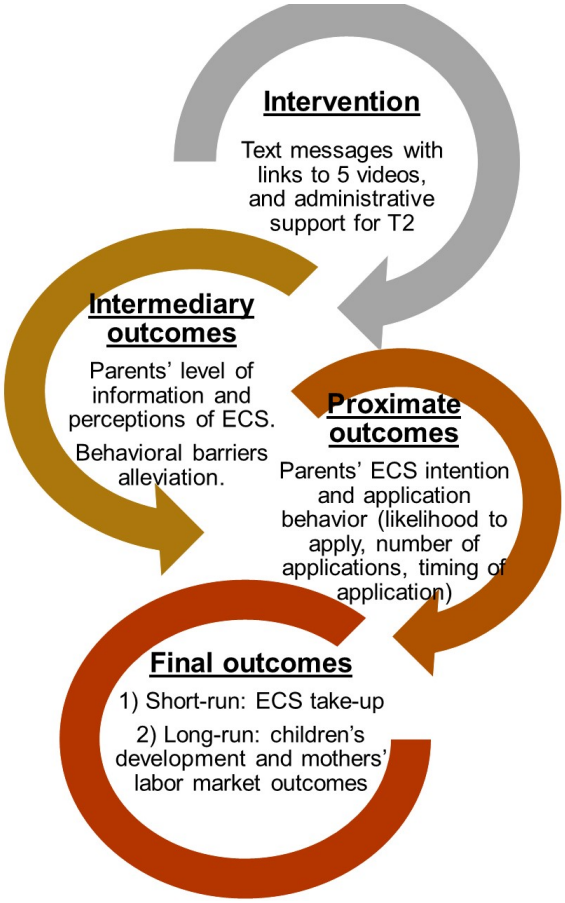
Definition IV.3 (Final outcomes)**.**

The final outcomes are divided into two sets:

1. **short run:** access to ECS
2. **long run:** children development and parents' labour market participation, incomes, family structure.

We wanted to make clear the nested relationships between different sets of outcomes. For our intervention to affect the final outcomes, it needs to change proximate outcomes. We think that this can only happen if there is an effect

Figure 2: Theory of change: illustration



on intermediary outcomes. Furthermore, if there are important treatment effects on proximate outcomes but no final outcomes, it shows evidence of rationing and mismatch on the childcare market. In other words, we use a deductive approach to understand causal mechanisms with pre-specified interpretations of our econometric analysis.

Additionally, we link question 1 to wider *consequences* which we include as *long run final outcomes*. We think of these consequences as reduced-form measures of the shock created by our intervention which *potentially* jointly affected parents decision in many dimensions and through different channels. The main difference between the long term outcomes and the others is that we do not have all the relevant information to interpret the results. However, we expect the results on the other set of outcomes to be informative and guide a more formal set of hypotheses. Therefore, these analyses will be conducted in another study with a dedicated pre-analysis plan.

We now move to the formal definitions of our empirical strategy.

V Research strategy

We start by briefly describing the data source, timing and their main contents. Then, we define the notations to define models and our identification and estimation strategies. We discuss inference together with the presentation of the models.

1 Data sources

A Baseline and endline survey

The experimental design embeds a baseline survey which generates an initial database containing a set of parents fitting our inclusion criteria and constitute the experimental sample. Undocumented migrants were excluded. We also excluded participants under 18, participants that do not have a cell phone or did not agree to leave their phone number, and participants living outside of our targeted areas (Paris, Seine Saint Denis, and Val de Marne) or about to move out of these regions. The experimental sample was then randomly split into the three treatment arms shortly after recruitment. We answer our main empirical questions for this analysis from surveys:

- **baseline** (September - December 2022)
- **endline survey** (October - December 2023)

Both are collected through computer assisted interviews by the research team with trained research assistants (in-person interviews for the baseline and CATI for the endline). All questionnaires contain information about parents' socio-demographic characteristics, intended ECS use (or actual use for Endline), ideal childcare arrangements (type, intensity), perception of the supply (monetary and non-monetary costs of using childcare vs. staying at home, accessibility), perceived social norms and beliefs about children's development during the early years. Importantly, both questionnaires also cover additional set of outcomes (e.g., mental well-being, breastfeeding, smoking) so that the participants remain unaware of the goal of study.

If we manage to secure additional funding, we also plan an follow-up survey one to one and a half years after the endline survey and matching with administrative records from the National family allowance fund (Cnaf) and employment records. These analysis will have their own pre-analysis plan.

We present the details of each survey in section VII.

B Power analyses and sample size determination

See the "sample size" section within the "Experimental design" section of the main information for registration document.

2 Notations

A Samples and observables

At baseline, we observe a sample of N_b pregnant mothers we call *individuals* indexed $i \in \{1 : N_b\}$. This sample has a nested structure by maternity ward and lottery. Individuals i recruited in a given ward and subject to a given lottery belong to a *wave* $w \in \{1 : W\}$ [3]. We note W_i the wave of individual i. We collect a lot of information on each mothers, some of which have a specific role in our research strategy.

First, we define blocks of individuals based on the cross product of three variables:

- **Education** E ("tertiary"/"secondary or lower")
- **Intention to use ECS** I ("no"/"yes but never has"/"yes and already has")
- **Supply** S ECS coverage rate higher/lower than the average in the department.

Let $b_i \in \{1 : B\}$ denotes the block of individual i defined as the product set of the blocking variables and waves (E_i, I_i, S_i, W_i). Let $B_i = \mathbb{1}(b_i = b)$ denotes block-dummies and **B** the matrix of blocks.

Block-randomization serves two purposes.

1) We want to reduce noise in our estimations and
2) be able to compare treatment effects across subgroups.

Our choice is guided by the discussion in section II which gives a simple rationale: parents' level of education is strongly associated with childcare use, parents who intend to use ECS are more likely to apply, but this may be more or less effective depending of the ECS coverage rate. Education is linked to intention to use and parents strategy may be different by education groups if supply is high or low. Therefore accounting for these interacted characteristics is expected to reduce the residual variance thus, improving precision of our treatment effect estimators. Since one of our main research questions is about treatment effect heterogeneity between high and low SES, blocking also ensures balance between groups and independence between education and treatment. Furthermore, treatment effect heterogeneity across different sets of blocks can be informative about mechanisms.

All variables measured at baseline constitute a set of attributes bundled in a matrix **A** of size (N_b, A). Attributes may vary with time but their measurement at baseline are informative of the parents situation at that time. This matrix includes E_i, I_i, S_i, W_i and B_i and other variables, among which we select $\mathbf{X} \subset \mathbf{A}$, a set of relevant attributes for each analysis. The method for choosing variables in **X** depend on the problem. We observe individuals i again in the endline survey with $N_p \leq B_b$ observations. Note that the difference between the two sample size must be investigated and may require statistical adjustment in case of differential attrition (See section D below).

This survey generates the set of outcome variables. In our general presentation, we follow the usual conventions and denote an outcome as a random variable Y. Y follows different definitions across sets of outcomes (intermediary, proximate and final) and which measurement of the outcome is used. We only assume that all distributions have finite first moments and that each observation Y_i is a realization of the random variable Y drawn from the same data generating process. We denote **Y** the matrix of outcomes. Some outcomes variables are related and we denote $\mathbf{Y}^f \subset \mathbf{Y}$ for a *family* of outcomes. We abuse of notations and also denote $\mathcal{H}^f \in \mathcal{H}$ a family of hypotheses with associated statistical tests to define the sets over which we adjust the Family-wise error rate.

a) Notations for the main families of outcomes The family of intermediary outcomes is noted $\tilde{I} \subset \mathbf{Y}$ and contain variables that captures the levels of information, perception of ECS and so on. The family of proximate outcomes measure application behavior. We let $\tilde{\mathbf{D}} \subset \mathbf{Y}$ be the matrix of measurements of application behavior. Finally, we let $\tilde{\mathbf{Y}} \subset \mathbf{Y}$ be the matrix of measurements of final outcomes.

3. There were 7 waves. Every two weeks, we proceed to random assignment.

B Treatment variable and potential outcomes

We denote Z^g the random variable that maps treatment assignment status such that:

$$g = \begin{cases} 2 := \textbf{(T2)}: Offered\ assistance \\ 1 := \textbf{(T1+T2)}: Offered\ information \\ 0 := \textbf{(T0)}: Offered\ placebo \end{cases}$$

and $Z_i^g = \mathbb{1}(g_i = g), g \in \{0,1,2\}$ denote treatment offered to individual i. These variables do not define treatment arms but which intervention they receive. Therefore, those in T2 have both $Z_i^2 = Z_i^1 = 1$.

We also note $\bar{Z}^g = \mathbb{1}(g_i \neq g)$ to note groups who did *not* receive treatment g and conveniently restrict subsamples using this notation.

a) Potential outcomes with information treatment Following Rubin (1974) causal model based on potential outcomes, we assume that treatment assignment reveal potential outcomes and SUTVA (*stable unit treatment value assumption*).

Let $Y_i(1)$ and $Y_i(0)$ denote potential outcomes when parents received information or not respectively. Note that these potential outcomes are defined as *intention to treat* parameters. $Y_i(1) - Y_i(0)$ is the individual gain of having access to our information treatment.

Among T1 and T0, treatment status link observed outcome Y_i with potential outcomes through this switching equation :

$$Y_i = Z_i^1 Y_i(1) + (1 - Z_i^1) Y_i(0) \quad \forall i\ s.t.\ \bar{Z}_i^2 = 1 \tag{1}$$

In words, we assume that observed outcomes in randomly assigned treatment arms reveal potential outcomes. Let $Y_i(2)$ denote potential outcomes when parents **received assistance**. It depends on parents choice to accept help.

3 Experimental design and identification

A Assumptions

We explicitly define the standard assumptions that our setting is meant to satisfy by design (Athey and Imbens 2017). Treatment status Z_i was assigned using block randomisation and therefore implies:

Hypothesis V.1 (Independance by block randomisation)**.**

$$(Y_i(0), Y_i(1), Y_i(2), \mathbf{X}_i, D_i(1), D_i(2)) \perp Z_i | B_i \tag{2}$$

Assuming our randomization went well, this assumption is verified which is sufficient for a causal interpretation of all intention to treat parameters (see below). It also gives a causal interpretation to the *first stage* effect on support take-up.

a) Compliance with support Let $D_i = \mathbb{1}(accept\ assistance)$ denote the parsimonious binary variable coding for receiving our support. By design, $D_i = 0\ \forall i\ s.t.\ \bar{Z}_i^2 = 1$; nobody in T1 or T0 could access our support. Those with $D_i = 1$ are those in T2 who accepted our support, but only after receiving information, like T1.

D measures who accept help to navigate administrative burden of ECS enrollment after receiving low-cost information. This compliance rate is a policy relevant parameter for our experiment. It captures the *demand for help* (when offered, and among those informed).

In the instrumental variable framework, Participation D_i is also the realization of potential participation revealed by random assignment. Let $D_i(0)$ $D_i(1)$, $D_i(2)$ denote potential participation *switched* by Z_i^1 and Z_i^2 by the following equation:

$$D_i = \underbrace{D_i(0) + Z_i^1(D_i(1) - D_i(0))}_{=0} + Z_i^1 Z_i^2(D_i(2) - D_i(1) - D_i(0)) \tag{3}$$

- When $Z_i^1 = Z_i^2 = 0$, $D_i = D_i(0)$.
- When $Z_i^1 = 1$ and $Z_i^2 = 0$, $D_i = D_i(1)$
- When $Z_i^2 = 1$, $D_i = D_i(2)$

Now let us restrict this equation to comparisons between T2 and T1, and T2 and T0:

$$D_i = D_i(1) + Z_i^2 \times (D_i(2) - D_i(1)) \quad \forall i \ s.t. \ \bar{Z}_i^0 = 1 \tag{4}$$

$$D_i = D_i(0) + Z_i^2 \times (D_i(2) - D_i(1) - D_i(0)) \quad \forall i \ s.t. \ \bar{Z}_i^1 = 1 \tag{5}$$

Equation (4) is standard and meaningful. The instrument Z^2 reveals the right potential outcomes. However, equation (5) does not. A clear causal parameter can only be defined over the sample of informed parents if we think that information did not change compliance with our support. Intuitively, we don't know who would have accepted our help in group T1 and T0. The instrumental variable strategy uses the share of compliers in T2 to infer the share of compliers in the comparison group. If information had an effect at least on some families, then T0 is simply not a good comparison group. For instance, some families in $T1$ may have found enough resources in the information to register early and refused our help. But they might have accepted our help had they not receive our information treatment. When we consider T2 and T1, such families are considered *never-takers* in the terminology of Joshua D. Angrist, Imbens, and Rubin (1996). This is relevant, and offering help likely has no other effect on outcomes but through participation. But in comparison with T0, such families would be a form of *defiers* and an IV estimators would not recover causal impacts, unless we assume strong additional exclusions restrictions in this first stage:

$$D_i(z) = D_i(0) \ \forall \ z \in \{0,1\}$$

and in the second stage.

Let us make clear how observed outcomes relates to potential outcomes, instruments and compliance.

b) Potential outcomes with support Potential outcomes are defined over the product set of potential participation and instruments and we denote them $Y_i(d,z)$ with d and $z \in \{0,1,2\}$, so there are 9 potential outcomes. However, our design allows to eliminate several :

$$Y_i(2,0) = Y_i(0,1) = Y_i(2,1) = Y_i(1,0) = \emptyset$$

The first four cases do not exist because of one-sided non-compliance which eliminates these potential defiers, the fifth case is excluded because all those who were offered assistance had information first and the last case is excluded by our information protocol[4].

- $Y_i(2,2)$ is the outcome of compliers when treated
- $Y_i(1,2)$ is the outcome of never-takers among those who received information.
- $Y_i(1,1)$ is the outcome of T1 and, by SUTVA: $Y_i(1,1) = Y_i(1)$
- $Y_i(0,0)$ is the outcome of T0 and, by SUTVA: $Y_i(0,0) = Y_i(0)$

For the two first potential outcomes, we need an additional hypothesis to retrieve $Y_i(2)$:

Hypothesis V.2 (Exclusion restrictions)**.** Among those who receive the information treatment, offering support has no other effect on outcomes but through the assistance provided. Formally,

$$Y_i(d,1) = Y_i(d,2) = Y_i(d) \quad \forall i \in \bar{Z}_i^0 = 1, \forall \ d \in \{1,2\} \tag{6}$$

With these notations, we can define our target parameters.

4. Only those in T1 received information.

B Target parameters

a) Intention to treat parameters A first parameter of interest is the intention to treat effect of offering additional support to those who already received information as:

Definition V.1 (Average effects of offering additional support)**.**

$$ITT(2) = \mathbb{E}[Y_i|Z_i^2 = 1] - \mathbb{E}[Y_i(1)|Z_i^2 = 0] \quad \forall i \ s.t. \ \bar{Z}_i^0 = 1$$

$ITT(2)$ is defined for individuals that are not in the control group T0. Therefore, $Z_i^1 = 1$ for everyone in this subsample. Note that we can also compare T2 with T0 in intention to treat to capture the effect of the bundled intervention.

We are also interested in conditional intention to treat effects :

$$ITT_x(2) = \mathbb{E}[Y_i|Z_i^2 = 1, \mathbf{X} = \mathbf{x}] - \mathbb{E}[Y_i(1)|Z_i^2 = 0, \mathbf{X} = \mathbf{x}] \quad \forall i \ s.t. \ \bar{Z}_i^0 = 1$$

Estimating these parameter provide a first answer to question 1 and 1a. But it does not show the effect on those who receive our support.

To answer question 2, we need to compare the effect on T2 vs T1 with T1 vs T0.

We define the average effect of information on any outcome $Y \in \mathbf{Y}$ as:

Definition V.2 (Average effects of information)**.**

$$ITT(1) = \mathbb{E}[Y_i(1) - Y_i(0)] \quad \forall i \ s.t. \ \bar{Z}_i^2 = 1$$

$ITT(1)$ is defined for individuals that are not in T2.

These two target parameters constitute a family of 2 results for which we want to test $\mathcal{H}^f$:

1) if any of the two is different from 0
2) if T2 is different from T1.

We formally test these hypothesis accounting for the FWER using generalized linear hypothesis testing following Hothorn, Bretz, and Westfall (2008). We also compute confidence interval of the difference and interpret the set of plausible values.

Note that contrary to definition V.2, only one potential outcome is revealed in definition V.1. The left conditional expectation is indeed the average for compliers and never-takers.

We are interested in the average direct effect of the support offered to parents on those who accepted it. It requires the estimation of the first stage which we define now.

b) First stage: The demand for help and decision making

Definition V.3 (First stage)**.**

- **Average demand**: Let $\pi_i = D_i(2) - D_i(1)$ and let

  $$\pi \equiv \mathbb{E}[\pi_i] = \mathbb{E}[D_i(2) - D_i(1)]$$

  be the average effect of proposing help on receiving help.

- **Block-specific demand**: let $\pi_b = \mathbb{E}[D_i(2) - D_i(1)|B_i = b]$ be the average effect of proposing help on receiving help within block b.

$$\pi = \sum_b pr(B_i = b)\mathbb{E}[D_i(2) - D_i(1)|B_i = b]$$

The average effect is the weighted average of the within-block effect with proportion of observations in each block as weights.

The first stage effect captures the demand for help in choosing and applying for formal childcare solutions. The population who accept our help are called *compliers*.

Remarks The first stage effect could be more formally analyzed if we think of it as a threshold choice function in a generalized Roy Model (J. J. Heckman and Vytlacil 2005). In the basic framework, the agent selects into treatment if the net utility from doing so is positive. This choice depends on the information set, subjective expected costs, benefits and so on. Importantly, such formalization is actually numerically equivalent to the LATE framework of Joshua D. Angrist, Imbens, and Rubin (1996; See Mogstad, Santos, and Torgovitsky 2018; Kline and Walters 2019). However, generalized Roy models can recover more general causal parameters under additional hypotheses. In our framework, participation decision reflects subjective cost-benefit evaluations and the joint decision of accepting help and choosing childcare arrangements. Therefore, modelling the *conditional average encouragement effect function* $\pi(\mathbf{x}) = \mathbb{E}[D_i(2) - D_i(1)|\mathbf{X} = \mathbf{x}]$ with a set of covariates that capture information set, perceived benefits, costs or views on formal childcare, in a more structured model can be very informative. Second, Generalized Roy models allow to estimate marginal treatment effects and treatment effect heterogeneity over the distribution of the propensity to accept help.

This formalisation is left for futur work and we focus on the identification of the LATE parameters.

c) Effects of support on those who accepted it

Definition V.4 (Average treatment effect on assistance compliers)**.**

$$ATT(1) = \mathbb{E}[Y_i(2) - Y_i(1)|D_i(2) > D_i(1)] \tag{7}$$

$ATT(1)$ is the average treatment effect on those who accepted support and information but who would otherwise receive information: the low-cost intervention.

This target parameter allows to answer our main research questions and document mechanisms when estimated over intermediary outcomes.

We also use the results proved by A. Abadie (2003) and recently revisited by Słoczyński, Uysal, and Wooldridge (2022a) and Słoczyński, Uysal, and Wooldridge (2022b).

Under hypotheses V.1 and V.2, we recall the following lemma holds:

Lemma V.1 (Abadie 2003, pp. 236-237)**.**

Let $g(\cdot)$ *be any measurable real function of* (Y, D, X) *such that* $\mathrm{E}|g(Y, D, X)| < \infty$. *Define*

$$\kappa_0 = (1 - D)\frac{(1 - Z) - (1 - p(X))}{p(X)(1 - p(X))},$$

$$\kappa_1 = D\frac{Z - p(X)}{p(X)(1 - p(X))}$$

$$\kappa = \kappa_0(1 - p(X)) + \kappa_1 p(X) = 1 - \frac{D(1 - Z)}{1 - p(X)} - \frac{(1 - D)Z}{p(X)}.$$

Under Assumptions V.1 and V.2,

1. $\mathrm{E}[g(Y, D, X) \mid D_1 > D_0] = \frac{1}{\mathrm{P}(D_1 > D_0)}\mathrm{E}[\kappa g(Y, D, X)]$. *Also,*

2. $\mathrm{E}\left[g\left(Y_0, X\right) \mid D_1 > D_0\right] = \frac{1}{\mathrm{P}(D_1 > D_0)} \mathrm{E}\left[\kappa_0 g(Y, X)\right]$, *and*

3. $\mathrm{E}\left[g\left(Y_1, X\right) \mid D_1 > D_0\right] = \frac{1}{\mathrm{P}(D_1 > D_0)} \mathrm{E}\left[\kappa_1 g(Y, X)\right]$.

4. *Moreover,* $(\mathrm{a}-\mathrm{c})$ *also hold conditional on* X.

This Lemma is very powerful, and has many useful applications to recover interesting parameters. For instance, by setting $g\left(X_i, Y_i(d)\right) = X_i$, an IV procedure yields the average of any predetermined covariate X_i for compliers. Therefore, we can characterize the complier population and see if they differ from other comparison groups. Similarly, we can estimate the mean for never-takers by regressing $X_i(1-D_i)Z_i$ on $(1-D_i)Z_i$.

More importantly, we can estimate the distribution of the missing potential outcomes for compliers. When Y is a continuous variable such as information scores and so on, by setting $g\left(X_i, Y_i(d)\right) = \mathbb{1}(Y_i \leq y)$ for a constant y and each value of d, we obtain the complier cumulative distribution functions of $Y_i(2)$ and $Y_i(1)$ evaluated at y.

d) Selection bias We simply estimate average (conditional) differences between families in T2 by treatment status *i.e.* as-treated analysis to gauge selection bias (J. Heckman et al. 1998). We also compare compliers and never takers characteristics estimated with the previous method.

We now move to our estimation strategy for these parameters.

4 Estimation strategy

We need to define our empirical strategy to estimate:

1) intention to treat parameters
2) treatment effects on compliers
3) heterogeneous treatment effects

And inference for hypothesis testing. We explain our general method to deal with non-response in the baseline. Our sample may suffer from attrition between baseline, endline and follow-up. We define our strategy to test and correct for differential attrition.

A Roadmap

A lot of our target parameters are identified and can be estimated by simple comparisons of weighted averages between treatment arms across across blocks. Regressions allow to estimate such weighted average, compute appropriate standard errors and add additional control variables.

For that, we define a simple roadmap that we can use with each family of outcomes:

1. We define the main measurement and alternative definitions and which comparison groups should be retained.
2. We run a baseline estimation with treatment and only block fixed effect.
3. We run additional estimations with control variables to improve precision.
4. We run separate estimations on predefine subgroups.

With this roadmap and the numerous research questions to test, we also need to control for familywise error rates associated with our statistical tests.

B Inference

Although individual observations are likely correlated within clusters such as maternity wards, our decision regarding inference depends primarily on our design (Alberto Abadie et al. 2020, 2022). We expect heterogeneous treatment effect within blocks and therefore, follow de Chaisemartin and Ramirez-Cuellar (2020) and use cluster-robust standard errors adjusted at the block level. Within each family of outcomes the chances of type-II error increase and the test statistics based on point-estimate standard error will over-reject the null hypothesis of no treatment effect. Instead, we consider testing the K null hypotheses. $H_0^1, \dots, H_0^K$ individually and require that the familywise error rate, i.e., the probability of falsely rejecting at least one true null hypothesis, is bounded by the nominal significance level $\alpha = .05$. We use adjusted p-values to describe the decision rules. Adjusted p-values are defined as the smallest significance level for which one still rejects an individual hypothesis H_0^j, given a particular multiple test procedure (Hothorn, Bretz, and Westfall 2008). By construction, we can reject an individual null hypothesis $H_0^k, k = 1, \dots, K$, whenever the associated adjusted p-value is less than or equal to the pre-specified significance level α, i.e., $p_k \leq \alpha$. The adjusted p-values.

Consistent with our pre-registration and power calculations, we use $\alpha = .05$.

C Baseline data and non-response

Our baseline database includes 1849 mothers assigned in each treatment arms. All important question for the design were mandatory. For the other variables, parents were usually given the opportunity to not answer or say they don't know. When we use these variables in the analysis, we include dummies for such cases. When variables are quantitative but values are missing, we include a dummy for missing and impute the mean value within block.

Importantly, all continuous variables used as a regressors are demeaned although we avoid using continuous regressors and we generally use dummified variables (through quantiles for instance).

D Attrition

Outcomes are measured in the second survey (called endline survey) between mid-October and late December through CATI. The research team will exert different efforts to maximize the response rate and reach the highest response rate. For instance, respondents to the endline survey are offered a voucher of 10€ to compensate for their time loss. In our pre-registration, we anticipated an average answer rate of 75 %. We follow standard practices to correct for imbalance (see for instance Doyle 2020).

Let $R_i = \mathbb{1}(Answer)$ denote the individual survey response status.

1) We first run a sanity check and estimate the following equation with OLS:

$$R_i = \sum_b \gamma_b B_i + \delta_1 Z^1 + \delta_2 Z^2 + \varepsilon_i$$

2) We jointly test $\hat{\delta_1} = \hat{\delta_2} = 0$
3) We run a fully saturated version of the previous regression and
4) We predict response status to construct corrective weights using standard functions of the inverse predicted probability and treatment arms.
5) We select baseline attributes that predict response using double lasso
6) We generate an alternative set of weights from the model with covariates selected using double lasso and more individual characteristics as a robustness check.

We present unweighted results and estimations with both sets of weights.

We also collect informations on how many calls we make to before reaching a household etc. This might be used to correct for endogenous non-response following Behaghel et al. (2015).

E Models for intention to treat parameters

a) Baseline models ITT parameter can be simply estimated by OLS with block fixed effects :

$$Y_i = \sum_b \alpha_b B_i + \beta Z_i + \varepsilon_i, \quad i \in \bar{Z} \tag{8}$$

Where α_b estimates average difference in outcome across blocks and β estimates a treatment-variance weighted average of block average intention to treat. We restrict over the relevant comparison groups $\bar{Z}$ corresponding to the ITT we define.

In most settings, the treatment-variance weighting of OLS is no issue since we compare to groups with the same treatment probability. However, some blocks may be small and slightly imbalanced. We therefore adopt a doubly-robust approach where :

1) We first regress Z_i^1 on block fixed effects
2) predict assignment probabilities for each block from the previous model
3) construct inverse probability weights for individuals of T2 and T1
4) run a weigthed version of the previous equation.

b) Robustness The main motivation for adjusting for covariates in RCTs is that the precision of the estimated average treatment effect can be improved if the covariates are sufficiently predictive of the outcome (Lin 2013). Randomisation ensures there are no correlation in the population between Z_i and the covariates $\mathbf{X}_i$, which is sufficient for the lack of bias from including or excluding the covariates (Athey and Imbens 2017). We follow Negi and Wooldridge (2021) recommendation and use pooled regression. In particular, including baseline outcomes (centred or in dummies) may increase precision while leaving the p-limit of the ITT coefficients unchanged. We then estimate the following equation using OLS:

$$Y_i = \mathbf{B}'\alpha + \beta Z_i + \mathbf{X}'\rho + \varepsilon_i \tag{9}$$

We use double lasso to select variables.

c) Heterogeneity To assess heterogeneity, we simply run separate regression over subgroups of interest. To account for simultaneous inference and test equalities of treatment across groups, we *stack* the separate models and run a single estimation. We use the method described by Hothorn, Bretz, and Westfall (2008) which assume joint normality of the treatment effect and use the variance covariance estimated from the stacked equation to correct for simultaneous inference.

F Models for T2 against T1

a) First stage effects The average effect of offering support on take-up can be obtained estimating the following equation using OLS:

$$D_i = \sum_b \alpha_b B_i + \pi Z_i^2 + \varepsilon_i, \quad i \ s.t. \ \bar{Z}_i^0 = 1 \tag{10}$$

Where π is the average demand for assistance among those who received information $i.e\ \pi = E_b\Big(E\left(D_i(2)|B_i = b\right)\Big)$. It estimates the average effect of encouragement on participation as a weighted average of the within-block average effects, with weights proportional to the product of the fraction of observations in the block and the probabilities of receiving and not receiving the treatment.

The following equation can also be estimated using OLS:

$$D_i = \sum_b Z_i^2 B_i \pi_b + \varepsilon_i, \quad i \ s.t. \bar{Z}_i^0 = 1 \tag{11}$$

The recent work of Blandhol et al. (2022), Słoczyński, Uysal, and Wooldridge (2022a) and Słoczyński, Uysal, and Wooldridge (2022b) revisit the use of 2SLS to estimate the LATE and show that the only specifications that have a LATE interpretation are "saturated" specifications that control for covariates non-parametrically or with restrictive parametric assumption.

Equation (11) is fully saturated[5] and estimate one effect of Z by block such that $E(\pi_b) = E(D_i(2)|B_i = b)$. The OLS estimation is equivalent to a non parametric estimation of the first stage.

b) Average effects on compliers With independence, a first stage effect with one sided non-compliance and exclusion, the average treatment effects on compliers are identified as the ratio of the expected difference in outcome by treatment arms over the expected first stage effect (Joshua D. Angrist, Imbens, and Rubin 1996):

$$\begin{aligned} LATE(Y) &= \frac{(E_b(Y_i|Z_i^2=1,B_i=b) - E_b(Y_i|Z_i^2=0,B_i=b))}{(E_b(D_i|Z_i^2=1,B_i=b)) - \underbrace{\left(E_b\left(D_i|Z_{ijc}^2=0,B_i=b\right)\right)}_{=0 \text{ one sided non-compliance}}} \\ &= \frac{(ITT(2))}{E_b(\pi_b|B_i=b)} \end{aligned} \tag{12}$$

Because of one-sided non compliance, these parameters correspond to the average treatment effect on the treated (Frölich and Melly 2013). These parameters can be estimated via the following 2SLS system of equation:

$$2SLS \begin{cases} Y_i = & \sum_b \alpha_b B_i + \delta D_i + \mu_i \\ D_i = & \sum_b \gamma_b B_i + \sum_b B_b' Z_i^2 \pi_b + \epsilon_i \end{cases} \tag{13}$$

adjusting standard errors for clustering at the block level.

c) Compliers characteristics and distribution of potential outcomes J. Angrist, Hull, and Walters (2023) conveniently summarize the use of instrumental variable in economics of education and recall many important results. In our setting comparing T2 and T1, the population contributing to the instrumental variable analysis contains only never-takers and compliers in this setting. The population with $D_i = 0$ and $Z_i^2 = 1$ are never-takers, $D_i = Z_i^2 = 1$ are compliers, while the group with $D_i = Z_i^2 = 0$ is a mixture of never-takers and compliers. However, their relative sizes are identified as the complier share equals the size of the first stage.

Under assumptions V.1 and V.2, A. Abadie (2003) propose a simple 2SLS procedure for characterizing compliers, described by:

$$g(X_i, Y_i) \times \mathbb{1}(D_i = d) = \sum_b \alpha_b B_i + \gamma_d \mathbb{1}(D_i = d) + \mu_i \quad d \in \{0,1\} \tag{14a}$$

$$\mathbb{1}(D_i = d) = \sum_b \eta_b B_i + \sum_b Z_i^2 \times B_i \pi_b + \epsilon_i \tag{14b}$$

where $g(X_i, Y_i)$ is any function of family baseline characteristics X_i and outcomes Y_i. Note that the first stage is fully saturated and π_w recover the block-specific first stage effects so that the projection of the first stage estimate correspond to the conditional expectation function and the TSLS estimate correspond to a LATE-like parameter (see Theorem 3 in J. D. Angrist and Imbens (1995)). Setting $d = 1$ in (14b) and (14a) means using Z_i^2 as an instrument for D_i in an IV procedure with $g(X_i, Y_i)$ multiplied by D_i as the outcome. Setting $d = 0$ means we use Z_i to

5. This first stage is saturated because it includes a parameter for each value of $\mathbb{E}[D_i|Z_i, b_w]$.

instrument $(1-D_i)$ in an equation with $g(X_i,Y_i)(1-D_i)$ as the outcome which recover characteristics of treated and untreated compliers:

$$\gamma_d = E\left[g(X_i,Y_i(d)) \mid D_i(2) > D_i(1)\right], \quad d \in \{0,1\} \tag{15}$$

By setting $g(X_i,Y_i(d)) = X_i$, the IV procedure produces the average of any predetermined covariate X_i for compliers within a block, and I estimate never-taker means by regressing $X_i(1-D_i)Z_i$ on $(1-D_i)Z_i$ with block fixed effect.

With this transformation, I can also estimate the distribution of the missing potential outcomes for compliers. By setting $g(X_i,Y_i(d)) = \mathbb{1}(Y_i \leq y)$ for a constant y and each value of d, I obtain the complier cumulative distribution functions of $Y_i(1)$ and $Y_i(0)$ evaluated at y.

More specifically, Abdulkadiroğlu, Pathak, and Walters (2018) propose to use a symmetric kernel function $g(X_i,Y_i(d)) = \frac{1}{h}K(\frac{Y_i-y}{h})$ in equation (14a) to estimate the density of the potential outcomes for compliers at $Y = y$. $K(\cdot)$ is a symmetric kernel function maximized at zero and h is a bandwidth that shrinks to zero asymptotically. We follow Abdulkadiroğlu, Pathak, and Walters (2018) and evaluates complier densities at a grid of 100 points with kernel bandwidth defined as $h = 1.06 \times N^{\frac{-1}{5}}\sigma_d$, where N is the sample size and σ_d is the standard deviation of the potential outcome.

VI Analysis and hypothesis testing

Now that our empirical strategy is well defined and all causal parameters introduced, we can explain how we can answer our main research questions.

Question 1 translates in an analysis of the average effect of our treatments on a measure of demand for and access to ECS. Question 1a assesses overall treatment effect heterogeneity by socio-economic background. It can be estimated by comparing pooled treatment effects between blocks of low and high SES.

Question 2 relates to the treatment effect heterogeneity between treatment arms. Question 2a assesses the differential impact of T2 compared to T1 by socioeconomic background.

1 Proximate outcomes: application behaviors

A Definition of outcomes and set of covariates

We expect our intervention to change to components of application behaviors previously noted $\tilde{\mathbf{D}}$:

- the decision to use ECS (and which ones)
- the number of applications.

For this set of outcomes, we use a double-LASSO to choose vectors in $\mathbf{X}_i$ among baseline control variables at the parent level (mother and father level of education, employment status, country of birth, baseline level of information, household composition, intention to use ECS, past ECS use, fluency in french, number of children, coverage rate, adherence with traditional norms, amount of family and friends using ECS, access to a computer, age of the mother, whether ECS opening hours fit with parents' working hours, beliefs in ECS returns, gender of the child, level of trust in daycare centers, perceived costs, perceived opportunity costs and costs on well-being and on the career).

B Treatment effects on Parents' likelihood to apply

Following our analysis roadmap, we first define measurement for our main outcome which is simply a dummy if parent is registered to any formal childcare.

We use the combination of answers to the endline survey and known applications for those who received administrative support to define a binary variable that equals one when a household applied to any ECS and 0 otherwise. We will assess the veracity of parents' answers before running these analyses.

Because dates of birth are heterogeneous in our sample, some parents may not have registered yet but intend to. Therefore, we will also use an alternative measurement that takes value 1 if parents intent to apply to ECS or have already applied, and 0 otherwise as a robustness check.

a) Intention to treat estimates -**Hypothesis 1** (subset $\bar{Z}^2$): We hypothesize that the information treatment increases the likelihood that parents apply to at least one ECS.

- We test the null $\mathcal{H}_1 : \hat{\beta}_1 = 0$ with a one sided test of positive treatment effect.

-**Hypothesis 2**(subset $\bar{Z}^2$: We hypothesize that offering information and administrative support increases the likelihood that parents applied to at least one ECS.

- We test the null $\mathcal{H}_2 : \hat{\beta}_0 = 0$ with a one sided test of positive treatment effect of the OLS regression of equation (8).

This effect is estimated in intention to treat and simply measures the average differences between T2 and T0. We then estimate this model with covariates.

-**Hypothesis 3** (subset $\bar{Z}^0$): We hypothesize that offering administrative support increase ECS applications compared to information only.

- We test the null $\mathcal{H}_3 : \hat{\beta}_2 = 0$ with a one sided test of positive treatment effect.

We follow the road map and estimate these models with covariates and adjust p-value for the FWER over each pair of test (with or without covariates).

If we reject $\mathcal{H}_1$, we conclude that the our information treatment is effective in increasing the likelihood of applying to any ECS. For this set of outcomes and all the other below, we do not expect sizeable effects of information only because the informational intervention boils down to a few simple messages. This especially if behavioral and administrative barriers are involved. However, given the very low cost of the information-only treatment, even a small positive effect is interesting to detect.

If we reject $\mathcal{H}_2$, we conclude that the offering information and administrative support increases the likelihood that parents applied to at least one ECS.

If we reject $\mathcal{H}_3$, we conclude that offering administrative support further increases this probability compared to information only.

Importantly, this effect is largely driven by the compliance rate. Indeed, almost all those who accepted the administrative support registered to at least one ECS. It also depends on compliers' behaviour in the absence of treatment.

b) Average treatment effect on compliers To estimate the effect on compliers, we use random assignment as an instrument for participation. We estimate equation 13 by 2SLS to retrieve the average treatment effect on compliers. This model also answers hypothesis 3 but retrieve the average effect on those who accepted administrative support. We also test the null $\mathcal{H}_\delta : \delta = 0$ with a one sided test.

Finally, we apply the same strategy in exploratory analysis over types of childcare they registered to and see if treatments fosters registrations in specific ECS types.

C Number of applications

We want to see if our treatment affected the total number of applications sent per households. We define the total number of applications sent per households as sum of applications sent across type of applications. For instance, if parents applied to municipal daycare and 5 associative daycares, the total number of applications sent will be 6).

We follow the exact same strategy as the previous one although we have different hypotheses to test.

- **Hypothesis 4**: We hypothesize that information increases the total number of applications sent and we test $\mathcal{H}_4 : \hat{\beta}_0 = 0$ with a one sided test.
- **Hypothesis 5**: We hypothesize that offering information and administrative support increases the total number of applications sent. We test the null $\mathcal{H}_5 : \hat{\beta}_1 = 0$ with a one sided test of positive treatment effect of the OLS regression of equation (8).

  This effect is estimated in intention to treat and simply measures the average differences between T2 and T0. We then estimate this model with covariates.
- **Hypothesis 6**: We hypothesize that offering support increases the total number of applications sent compared to information only treatment. We test $\mathcal{H}_6 : \hat{\beta}_2 = 0$ with a one sided test.

Like before, we estimate the effect on compliers.

Finally, we apply the same strategy in exploratory analyses over the different ECS type they parents applied to to see if treatments fosters applications to specific ECS type (see below). In particular, we expect increased access to public daycare, associative daycare, and drop-in daycare because our videos gave more details about these types of care.

D Treatment effects on parents' timing of application

We define another family of outcomes to measure the timing of applications. In the endline survey, we ask parents at which date they started to apply for ECS and we have the child(ren)'s date of birth. In the baseline, we also have the month of pregnancy. From these, we construct a relative time measure and model the duration from the month of conception to first registration. For these outcomes, the models presented earlier do not apply. Instead, we use survival models to account for censoring and estimate the difference in restricted mean survival times (Nemes, Bülow, and Gustavsson 2020). For the rest, we follow the same strategy as the previous outcome and we register the following hypotheses.

- **Hypothesis 7**: We hypothesize that information made parents apply earlier.
- **Hypothesis 8**: We hypothesize that offering information and administrative support made parents apply earlier.
- **Hypothesis 9**: We hypothesize that offering support may have more positive effect on timing of application than the information-only treatment.

2 Final outcomes : Effects on childcare use

A Definition of outcomes and set of covariates

We expect our intervention to improve to ECS access, conditional on the fact that supply-side factors are not too strong.

For this set of outcomes, we use a double-LASSO to choose vectors in $\mathbf{X}_i$ among baseline control variables at the parent level (mother and father level of education, employment status, country of birth, baseline level of information, household composition, intention to use ECS, past ECS use, fluency in french, number of children, coverage rate, adherence with traditional norms, amount of family and friends using ECS, access to a computer, age of the mother, whether ECS opening hours fit with parents' working hours, beliefs in ECS returns, gender of the child, level of trust in daycare centers, perceived costs, perceived opportunity costs and costs on well-being and on the career).

B Parents' access to ECS

Following our analysis roadmap, we first define measurement for our short-term final outcomes. We use answers to the endline survey to define a binary variable that equals one when a person accessed to any ECS type and 0 otherwise. We will assess the veracity of parents' answers prior running these analyses.

Our baseline strategy consists in estimating equation (8) with OLS over $\bar{Z}^2$ and $\bar{Z}^0$ i.e.~comparing T1 and T0 and T2 and T1 to retrieve the intention to treat parameters. We index the estimations of the ITT parameters β_g with $g \in \{1,2\}$, the corresponding treatment arms. Note that β_2 is computed against parents who also received the information treatment.

-**Hypothesis 10** (subset $\bar{Z}^2$): We hypothesize that the information treatment increases the likelihood that parents access a place in ECS.

- We test the null $\mathcal{H}_{10} : \hat{\beta}_0 = 0$ with a one sided test of positive treatment effect.

-**Hypothesis 11**(subset $\bar{Z}^1$: We hypothesize that offering information and administrative support increases the likelihood that parents access a place in ECS.

- We test the null $\mathcal{H}_{11} : \hat{\beta}_1 = 0$ with a one sided test of positive treatment effect of the OLS regression of equation (8).

This effect is estimated in intention to treat and simply measures the average differences between T2 and T0. We then estimate this model with covariates.

-**Hypothesis 12** (subset $\bar{Z}^0$): We hypothesize that offering administrative support increases the likelihood that parents access a place in ECS compared to information only.

- We test the null $\mathcal{H}_{12} : \hat{\beta}_2 = 0$ with a one sided test of positive treatment effect.

We follow the road map and estimate these models with covariates chosen among the relevant set of variables presented above. We adjust p-value for the FWER over each pair of test (with or without covariates).

We estimate also the average treatment effect on compliers with a 2SLS of equation (13). We also test the null $\mathcal{H}_\delta : \delta = 0$ with a one sided test.

Finally, we apply the same strategy in exploratory analysis over the different ECS type they registered to see if treatments fosters access to specific ECS type. In particular, we expect increased access to public daycare, associative daycare, and drop-in daycare because our videos gave more details about these types of care.

3 Intermediate outcomes: parents level of information and perception of ECS

A Definition of outcomes and set of covariates

We define another family of outcomes to measure how the intervention affected parents level of information about ECS.

For this set of outcomes, we use a double-LASSO to choose vectors in $\mathbf{X}_i$ among baseline control variables at the parent level (mother and father level of education, employment status, country of birth, baseline level of information, household composition, intention to use ECS, past ECS use, fluency in french, number of children, coverage rate, adherence with traditional norms, amount of family and friends using ECS, access to a computer, age of the mother, whether ECS opening hours fit with parents' working hours, beliefs in ECS returns, gender of the child, level of trust in daycare centers, perceived costs, perceived opportunity costs and costs on well-being and on the career).

B Parents' level of information about ECS

Following our analysis roadmap, we first define measurement for our intermediate outcomes. We want to know whether we increased their general level of information, and the accuracy of perceived ECS costs. Our main measurements are:

a) Number of ECS types known,
b) Whether parents know that ECS prices vary according to resources,
c) Parents' subjective level of information coded ex-post by the interviewers,

We will have three separated information variables on which we will estimate the effect of our interventions, testing hypotheses 13, 14 and 15. We use stacked regression and jointly estimate the effect on the three outcomes and test the null that they are all equals to 0 against at least one is positive.

We follow the exact same strategy as before although we have different hypotheses to test. For each of these variables, we make the following hypotheses:

- **Hypothesis 13**: We hypothesize that information made parents more informed about ECS and we test $\mathcal{H}_{13}$: $\hat{\beta_0} = 0$ with a one sided test.

- **Hypothesis 14**: We hypothesize that offering information and administrative support increases the likelihood that parents access a place in ECS. We test the null $\mathcal{H}_{14} : \hat{\beta}_1 = 0$ with a one sided test of positive treatment effect of the OLS regression of equation (8).

This effect is estimated in intention to treat and simply measures the average differences between T2 and T0.

- **Hypothesis 15**: We hypothesize that offering administrative support+information have more positive effect on parents' information level than information only. We test $\mathcal{H}_{15} : \hat{\beta}_2 = 0$ with a one sided test.

We follow the road map and estimate these models with covariates chosen among the relevant set of variables presented above. We adjust p-value for the FWER over each pair of test (with or without covariates).

We estimate also the average treatment effect on compliers with a 2SLS of equation 13. We also test the null $\mathcal{H}_\delta : \delta = 0$ with a one sided test.

4 Cross Hypotheses: Heterogeneity

For the heterogeneity analysis, we focus only on T2, the group with highest treatment intensity that is most likely to yield large enough effect size for us to detect. We test heterogeneity in intention to treat and with instrumental variable using again the stacked-regression trick to recover the variance covariance of the error and jointly test equality between treatment effects (Oberfichtner and Tauchmann 2021).

a. *Social background*: We are willing to answer research question 1a and research question 2a. Indeed, we expect low-SES parents to benefit more from the intervention because they are more likely to misperceive the ECS offer at Baseline due to informational barriers, and to experience behavioral barriers during the application process. Our intervention eases all those constraints so we expect that they benefit from it.

High-SES may also benefit from the intervention. For example, it may be easier for high SES parents to use the information provided or take advantage from our help. However, we expect high-SES parents to navigate the system very well, even without our intervention.

Therefore, we posit: **Hypothesis 16**: We hypothesize that T2 will have a larger impact on low-SES parents. To measure SES, we will use a dummy that takes the value one if the mother has attended tertiary education, and 0 otherwise.

Robustness: we will test the robustness of these findings using alternative measurements of SES:

i) *Occupation*: we will convert mothers' occupation in ISCO 08 classification, then we will convert ISCO 08 in ISEI (score of occupational prestige), and finally we will dichotomize this score in a dummy taking value 1 if the mother is above the median ISEI score of our sample, and 0 otherwise;

ii) *Income*: we will create a dummy that takes the value 1 if the household is above median income of our sample and 0 otherwise;

iii) *Activity*: we will create a dummy that takes the value 1 if the mother was employed at baseline, and 0 otherwise.

For all these measurements, including education, we will start by taking the situation of the mother (except for income). Then, we will test alternative specifications by taking a dominance criteria: we will take the highest value of the household. For instance, if the mother has no tertiary education but the other parent does, we will code this household as having tertiary education.

b. *Past ECS usage*:

Hypothesis 17: We hypothesize that T2 will have a larger impact on parents who have never used ECS.

We define a dummy that take the value 1 if parents had already used ECS at baseline, and 0 otherwise.

c. *Level of information*: We expect relatively less-informed parents at baseline to benefit more from the intervention compared to more informed ones. After the end of the baseline questionnaire, surveyor marked parents knowledge of ECS and we use this variable to define groups of High/low information at baseline.

Hypothesis 18: We hypothesize that T2 will have a larger impact on parents who had low level of information at baseline.

d. *Migration background*:

Hypothesis 19: We hypothesize that T2 will have a larger impact on foreign-born parents.

Indeed, they are more likely to misperceive the ECS offer at Baseline due to informational barriers, and to experience behavioral barriers during the application process. We will introduce an interaction term with a dummy that take the value 1 if the mother is not born in France, and 0 otherwise.

- We test $\mathcal{H}_{19} : \widehat{\beta}_1 = 0$ with a one sided test.

5 Exploratory analyses:

The effect of several additional variables will be relevant to explore:

1) *Coverage rate in ECS*: We expect the effect of our intervention to vary according to the cover rate in ECS. We will compare parents living in areas with a cover rate in ECS below the median cover rate of the sample with parents living in areas above this median cover rate. We will introduce an interaction term with a dummy that take the value 1 if leave in an area with a below-median ECS coverage rate, and 0 otherwise.

We don't make any prediction about the direction of the effect. Indeed, our intervention aims at easing process, and make registration earlier, which is crucial to gain priority points to get a place (Combe and Heim 2023). Therefore it may be easier to get apply and get access to ECS in high coverage rate cities. However, the program may also increase parents relative competitivity in highly competitive context, that is, for instance, they apply earlier, make more applications and navigate the ECS market more easily than non-treated parents. Hence, it may increase their probability of getting a daycare slot in higly competitive contexts.

2) *ECS types*: We will apply the same strategy as in our main analyses to explore the impact of our interventions on the different ECS type. We want to see if treatments fosters demand for and access to specific ECS types. In particular, we expect increased access to public daycare, associative daycare, and drop-in daycare because our videos gave more details about these types of care. We will define dummies for applications to each type of childcare.

3) *Intention*: The baseline questionnaire collects respondents' *intention to apply*. It will be relevant to explore if treatment impacts are larger for parents that already intended to apply before child birth, or for those that were undecided, or even for those that were opposed to ECS at baseline.

4) *Temporal preferences*: The baseline questionnaire collects respondents' temporal preferences. Previous behavioral interventions among low-SES parents showed that information + behavioral tools (e.g., reminders) delivery have larger effects on more present-orientated parents Mayer et al. (2019). Hence, it will be relevant to explore if treatment impacts are larger for more present-orientated subjects.

5) *Intervention to birth date timing*: Combe and Heim (2023) showed that one major factor that drives ECS access inequalities among parents that applied is that high-SES parents apply very early during pregnancy. Therefore, it will be relevant to explore if the impact of treatment is larger when parents received the intervention early during pregnancy.

6) *Single-parent families*: It will be relevant to explore the heterogeneity of the treatment effect according to whether the family is a single-parent family or not.

7) *Date of birth of the child*: As the time at which the children are born during the year is known to affect childcare arrangements , it will be relevant to explore the heterogeneity of the treatment effect according to whether the date of birth of the child.

8) *Effect on other siblings*: There may be spill-overs of the treatment on other siblings that are less than 3 years-old. It will be relevant to explore them.

9) *Beliefs in returns*: We have a measure of whether parents had high or low beliefs in ECS returns. We expect parents having high beliefs in return to benefit more from the intervention.

10) *Norms*: We measure adherence to traditional norms in the questionnaire. It will be relevant to explore the heterogeneity of the treatment effect according to adherence to traditional norms.

11) *Perception of the supply*: We measure parents perception of the supply (perceived difficulty of accessing a place, perceived fit with working hours, perceived accessibility of information about ECS). It will be relevant to explore the heterogeneity of the treatment effect according to parents' perception of the supply.

VII Data collection

See the "Data collection procedure" section of the main information for registration document.

1 Collected variables

A Baseline survey

Questionnaires are administered on a tablette with the support of the interviewer. The following variables are collected in the questionnaires:

- Socio-demographic characteristics of the family (level of education; migratory background; mothers and household income; activity)
- Intention to use ECS (yes, no, don't know yet)

- Level of information on ECS (e.g., types of ECS known, perceived costs)
- Planned activity after child delivery (i.e., whether the mother is planning to go back to work/look for a job after child delivery, and if so when)

Some other information is also collected to understand the determinants of the application to ECS:

- Deprivation/salience of vital needs

- Temporal preferences

- 2 measurements of mothers' investment and health behaviors during the pregnancy i) smoking behavior during the pregnancy ii) breastfeeding intention

- Informal care available

- Social resources available (e.g., friends to help parents fill the forms)

- Perceived impact of staying at home on the mother's professional career and wages

- Perceived impact of staying at home on the mother's wellbeing

- Parents' beliefs about children's development: how long should a child be (exclusively) cared after by her parents; benefits of ECS on children's development; skills one child should acquire during early childhood
- Parents' level of trust in several communities to take care of their children

i) parents and brothers and sisters
ii) the rest of the family
iii) Friends
iv) Childminders (assistante maternelle)
v) childcare centers (crèches)

- Social norms regarding childcare

- Access to a computer

B Endline survey

CATI Questionnaires: questionnaires administered on the phone to participants by an interviewer assisted with a computer. The following variables are collected in the questionnaires:

1) **Information about the baby**:

- Multiple pregnancy: No, Twins, Triplets, etc.
- Prematurity: YN, number of weeks
- Birth Date(s) of the babies
- Assigned sex at birth of the baby
- First Name of the baby
- Birth Weight of the baby
- Health outcomes:

  i) whether the baby has any particular chronic health condition or disability,
  ii) nights spent in the hospital since home return,
  iii) subjective health

- Breastfeeding: YN, length
- Whether the family have a Family Allowance fund account (Caf) for the baby, and if not why? All outcomes are collected for every baby in case of multiple pregnancy

2) **Mother health and well-being**:

- Health outcomes: night spent in the hospital at birth; nights spent in the hospital since the mother and the baby came out of the maternity hospital for the first time.
- Well-being: we will use the french version of the WHO 5 well-being index

3) **Treatment-related outcomes**:

- Whether the family received the text messages with links to videos (Y/N)
- Compliance: whether the family watched the videos (Y/N), how many (/5), watched until the end (Y/N). If not watched, why?
- Perceived usefulness: 0 to 10 scale
- General satisfaction: 0 to 10 scale

4) **Situation of the household**:

- Family composition at the time of the interview
- Labor market outcomes: employment status and/or plans, date of return to work, full-time or part-time employment, relative intensity of labor market situation compared to before pregnancy, and occupation. The questions will be asked for the two parents if the parents are still together.
- Housing situation: moving since baseline questionnaire (Y/N), where, and when if so.
- Second-generation migration status

5) **Childcare related outcomes**:

a) *ECS behavior*:

- ECS use (Y/N)
- Applied to ECS (Y/N)
- Ideal type of childcare wanted, and reason for not using the ideal childcare type if so.
- If they have another baby less than 3yo: ECS-related outcome for this baby (see below)

If the household uses ECS:

- Type of ECS used as the main childcare solution
- ECS use start date
- Intensity of use: Number of hours per weeks
- Is the household satisfied with the current number of hours of use (Y/N)
- Secondary childcare solution: does the household combine several childcare solutions? If so, type of ECS used

If the household doesn't access ECS:

- Current childcare solution use (e.g., maternal or paternal care, informal care)

Then for those who applied to at least one ECS type:

- Type of ECS and number of applications
- Status of the application (in progress, rejected, or refused). Reason of refusal if so. If rejected, did they renew the application, and if not, why?
- Timing of application: when did they start to apply to ECS
- Initial starting date wanted
- Application behavior: did they call ECS? Did they go on-site? Did they use their personal or social network?

For those who did not apply to any ECS: - ECS plans (Y/N)
- Did they look for information (Y/N), and if so, when?

If they plan to use ECS:
- ECS type
- Starting date wanted

b) *ECS perceptions*:

- Information level: perceived cost of childcare, knows that the price is reduced through subsidies (Y/N), number of ECS known, subjective level of information (coded ex-post by the interviewer)
- Perceived ECS accessibility
- Perceived accessibility of information about ECS
- Perceived compatibility with working hours
- Perceived control over the ECS outcome: 0 to 100 scale
- Level of stress due to ECS: 0 to 100 scale

- Perceived complexity of the application process: 0 to 100 scale
- Access to informal care (as much as needed, from time to time, for emergency cases, never)
- Ideal entry age in ECS for the child
- Perceived norms: a) Perceived proportion of mothers using ECS, b) Perceived proportion of mothers opposed to ECS, c) adherence to traditional norms
- Non-monetary costs of using ECS or not: a) on well-being, b) on career, c) opportunity costs.
- Perceived beliefs in ECS returns

Questionnaires are available in 3 languages: French, English, and Arabic.

2 Exclusion rules

For Baseline, see the "Targeting" section within the "Experimental Design" section of the main information for registration document.

For Endline, the following exclusion rules will be applied: i) If the Premiers Pas child died since Baseline interview, ii) If the parents have no Family Allowance fund account for the baby (because parents are undocumented), iii) If the parents don't understand a simple level of French, Arabic, or English. Some participants may have gone through the inclusion process at Baseline because a translator translated the Baseline survey for them in the maternity wards, and the Research Assistant thus included them in the trial.

VIII Data processing and security

- **Data protection officier**: The project has been conducted with a close collaboration with Sciences Po's Data Protection Officer to ensure compliance with data protection regulations throughout the project.
- **Secure Cloud Storage**: Data is securely stored on an encrypted cloud platform with very restricted access to prevent unauthorized access and data breaches.
- **De-identification and persona information storage**: All personally identifiable information (PII) has been removed from the dataset. Any remaining identifying data is stored in a separate VeraCrypt-encrypted directory, adding an extra layer of security.
- **Secure Cloud Storage**: Data is securely stored on an encrypted cloud platform with very restricted access to prevent unauthorized access and data breaches.

These measures collectively safeguard data integrity and protect the privacy of individuals involved in the research.

IX Research Team

The research team is as follows:

- Laudine Carbuccia, PhD student at the Center for Research on Social Inequalities (Sciences Po) & at the Department of Cognitive Science at ENS/PSL. She works full-time on the project.
- Arthur Heim, PhD candidate,at Paris School of Economics and research and evaluation officer at the National family allowance fund (Cnaf)
- Carlo Barone, full professor of Sociology at the Center for Research on Social Inequalities (Sciences Po).
- Coralie Chevallier, cognitive psychologist at the Department of Cognitive Science at ENS/PSL.

X Bibliography

Abadie, A. 2003. "Semiparametric Instrumental Variable Estimation of Treatment Response Models." *Journal of Econometrics* 113 (2): 231–63.

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling-Based Versus Design-Based Uncertainty in Regression Analysis." *Econometrica* 88 (1): 265–96.

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge. 2022. "When Should You Adjust Standard Errors for Clustering?" *The Quarterly Journal of Economics* 138 (1): 1–35.

Abdulkadiroğlu, Atila, Parag A. Pathak, and Christopher R. Walters. 2018. "Free to Choose: Can School Choice Reduce Student Achievement?" *American Economic Journal: Applied Economics* 10 (1): 175–206.

Abrassart, Aurélien, and Giuliano Bonoli. 2015. "Availability, Cost or Culture? Obstacles to Childcare Services for Low-Income Families." *Journal of Social Policy* 44 (4): 787–806.

Angrist, J. D., and Guido W Imbens. 1995. "Two-Stage Least Squares Estimation of Average Causal Effects in Models With Variable Treatment Intensity." *Journal of the American Statistical Association* 90 (430).

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91 (434): 444–55.

Angrist, Joshua, Peter Hull, and Christopher Walters. 2023. "Methods for Measuring School Effectiveness." In *Handbook of the Economics of Education*. Elsevier.

Athey, S., and G. W. Imbens. 2017. "Chapter 3 - The Econometrics of Randomized Experiments." In *Handbook of Economic Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 1:73–140. Handbook of Field Experiments. North-Holland.

Attanasio, Orazio P. 2015. "The Determinants of Human Capital Formation During the Early Years of Life: Theory, Measurement, and Policies." *Journal of the European Economic Association* 13 (6): 949–97.

Barnett, W. S. 2011. "Effectiveness of Early Educational Intervention." *Science* 333 (6045): 975–78.

Behaghel, Luc, Bruno Crépon, Marc Gurgand, and Thomas Le Barbanchon. 2015. "Please Call Again: Correcting Nonresponse Bias in Treatment Effect Models." *Review of Economics and Statistics* 97 (5): 1070–80.

Binns, Reuben, Max Van Kleek, Michael Veale, Ulrik Lyngs, Jun Zhao, and Nigel Shadbolt. 2018. "'It's Reducing a Human Being to a Percentage'; Perceptions of Justice in Algorithmic Decisions." *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–14.

Blandhol, Christine, Magne Mogstad, John Bonney, and Alexander Torgovitsky. 2022. "When Is TSLS Actually LATE?" Becker Friedman Institute for Economics at UChucago. February 9, 2022.

Bouteillec, Nathalie Le, Lamia Kandil, and Anne Solaz. 2014. "L'accueil En Crèche En France : Quels Enfants y Ont Accès ?" *Population Societes* N° 514 (8): 1–4.

Burger, Kaspar. 2010. "How Does Early Childhood Care and Education Affect Cognitive Development? An International Review of the Effects of Early Interventions for Children from Different Social Backgrounds." *Early Childhood Research Quarterly* 25 (2): 140–65.

Caenen, Yann, and Pauline Virot. 2023. "La Part Des Enfants de Moins de 3 Ans Confiés Principalement à Une Assistante Maternelle Ou Une Crèche a Presque Doublé Entre 2002 Et 2021." *DREES*, Études et Résultats, no. 1257 (February).

Camilli, Gregory, Sadako Vargas, Sharon Ryan, and W Steven Barnett. 2010. "Meta-Analysis of the Effects of Early Education Interventions on Cognitive and Social Development." *Teachers College Record* 112 (3): 579–620.

Cantillon, Bea. 2011. "The Paradox of the Social Investment State: Growth, Employment and Poverty in the Lisbon Era." *Journal of European Social Policy* 21 (5): 432–49.

Carbuccia, Laudine. 2023. "Behind the Gap: Exploring Determinants of the SES Gap in Early Childcare Enrolment in France Through Qualitative Insights." *CRIS Working Paper*.

Carbuccia, Laudine, Carlo Barone, Gregoire Borst, Angela Greulich, Lidia Panico, and Maxime Tô. 2020. "Revue de Littérature Sur Les Politiques d'accompagnement Au Développement Des Capacités Des Jeunes Enfants." Dossier d'\'etude 215. CNAF.

Carbuccia, Laudine, Valentin Thouzeau, Carlo Barone, and Coralie Chevallier. 2023. "Unequal Access to Early Childcare: What Role Do Demand-Side Factors Play? A PRISMA Systematic Review." LIEPP Working Paper. Paris. January 2023.

Cartier, Marie, Anaïs Collet, Estelle Czerny, Pierre Gilbert, Marie-Hélène Lechien, and Sylvie Monchatre. 2017. "Pourquoi Les Parents Préfèrent-Ils La Crèche ? Les Représentations Hierarchisées Des Modes de Garde Professionnels." *Revue Française Des Affaires Sociales*, no. 2: 250–65.

Case, Anne, Angela Fertig, and Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics* 24 (2): 365–89.

Chaisemartin, Clément de, and Jaime Ramirez-Cuellar. 2020. "At What Level Should One Cluster Standard Errors in Paired Experiments, and in Stratified Experiments with Small Strata?" Working Paper 27609. NBER.

Chaudry, Ajay, Julia Henly, and Marcia Meyer. 2010. "Conceptual Frameork for Child Care Decision Making." ACF-OPRE White Paper. Washington, DC.: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Combe, Julien, and Arthur Heim. 2023. "The Plumber, the Engineer, the Economists : Theoretical Solutions to the Daycare Assignment Problem with Real-Life Applications in France." In *Social Investment & The Feminization Of Poverty: Essays on The Design And Evaluation Of Social Investment Policies In France*, Paris School of Economics. Paris, France: PhD Thesis.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, 1:697–812. Elsevier.

Doyle, Orla. 2020. "The First 2,000 Days and Child Skills." *Journal of Political Economy* 128 (6): 2067–2122.

Duncan, Greg J., Katherine Magnuson, Ariel Kalil, and Kathleen Ziol-Guest. 2012. "The Importance of Early Childhood Poverty." *Social Indicators Research* 108 (1): 87–98.

Elango, Sneha, Jorge Luis Garcìa, James J. Heckman, and Andres Hojman. 2016. "Early Childhood Education." In *Economics of Means-Tested Transfer Programs in the United States*, Moffitt, Robert A. Vol. II. University of Chicago Press.

Frölich, Markus, and Blaise Melly. 2013. "Identification of Treatment Effects on the Treated with One-Sided Non-Compliance." *Econometric Reviews* 32 (3): 384–414.

Ghysels, Joris, and Wim Van Lancker. 2011. "The Unequal Benefits of Activation: An Analysis of the Social Distribution of Family Policy Among Families with Young Children." *Journal of European Social Policy* 21 (5): 472–85.

Heckman, James J., and Edward Vytlacil. 2005. "Structural Equations, Treatment Effects, and Econometric Policy Evaluation." *Econometrica* 73 (3): 669–738.

Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica* 66 (5): 1017–98.

Hermes, Henning, Marina Krauß, Philipp Lergetporer, Frauke Peter, and Simon Wiederhold. 2022. "Early Child Care and Labor Supply of Lower-Ses Mothers: A Randomized Controlled Trial." *SSRN Electronic Journal*.

Hothorn, Torsten, Frank Bretz, and Peter Westfall. 2008. "Simultaneous Inference in General Parametric Models." *Biometrical Journal* 50 (3): 346–63.

Huber, Katrin, and Geske Rolvering. 2023. "Public Child Care and Mothers' Career Trajectories."

Huizen, Thomas van, and Janneke Plantenga. 2018. "Do Children Benefit from Universal Early Childhood Education and Care? A Meta-Analysis of Evidence from Natural Experiments." *Economics of Education Review* 66 (October): 206–22.

Johnson, Rucker C, and Robert F Schoeni. 2011. "The Influence of Early-Life Events on Human Capital, Health Status, and Labor Market Outcomes Over the Life Course." *The B.E. Journal of Economic Analysis & Policy* 11 (3).

Kholoptseva, Evgenia. 2016. "Effects of Center-Based Early Childhood Education Programs on Children's Language, Literacy, and Math Skills: A Comprehensive Meta-Analysis," May.

Kimmel, Jean. 1998. "Child Care Costs as a Barrier to Employment for Single and Married Mothers." *The Review of Economics and Statistics* 80 (2): 287–99.

Kline, Patrick, and Christopher R. Walters. 2019. "On Heckits, LATE, and Numerical Equivalence." *Econometrica* 87 (2): 677–96.

Lancker, Wim Van, and Joris Ghysels. 2012. "Who Benefits? The Social Distribution of Subsidized Childcare in Sweden and Flanders." *Acta Sociologica* 55 (2): 125–42.

Lazzari, Dr Arianna. 2012. "Literature Review of the Participation of Disadvantaged Children and Families in ECEC Services in Europe," 127.

Li, Shengwu. 2017. "Ethics and Market Design." *Oxford Review of Economic Policy* 33 (4): 705–20.

Lin, Winston. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7 (1): 295–318.

Mayer, Susan E., Ariel Kalil, Philip Oreopoulos, and Sebastian Gallegos. 2019. "Using Behavioral Insights to Increase Parental Engagement: The Parents and Children Together Intervention." *Journal of Human Resources*

54 (4): 900–925.
Mogstad, Magne, Andres Santos, and Alexander Torgovitsky. 2018. "Using Instrumental Variables for Inference About Policy Relevant Treatment Parameters." *Econometrica* 86 (5): 1589–619.
Morrissey, Taryn W. 2017. "Child Care and Parent Labor Force Participation: A Review of the Research Literature." *Review of Economics of the Household* 15 (1): 1–24.
Negi, Akanksha, and Jeffrey M. Wooldridge. 2021. "Revisiting Regression Adjustment in Experiments with Heterogeneous Treatment Effects." *Econometric Reviews* 40 (5): 504–34.
Nemes, Szilárd, Erik Bülow, and Andreas Gustavsson. 2020. "A Brief Overview of Restricted Mean Survival Time Estimators and Associated Variances." *Stats* 3 (2): 1–13.
Noailly, Joëlle, and Sabine Visser. 2009. "The Impact of Market Forces on Child Care Provision: Insights from the 2005 Child Care Act in the Netherlands." *Journal of Social Policy* 38 (3): 477–98.
Nollenberger, Natalia, and Núria Rodríguez-Planas. 2015. "Full-Time Universal Childcare in a Context of Low Maternal Employment: Quasi-experimental Evidence from Spain." *Labour Economics* 36 (October): 124–36.
Nores, Milagros, and W. Steven Barnett. 2010. "Benefits of Early Childhood Interventions Across the World: (Under) Investing in the Very Young." *Economics of Education Review* 29: 271–82.
Oberfichtner, Michael, and Harald Tauchmann. 2021. "Stacked Linear Regression Analysis to Facilitate Testing of Hypotheses Across OLS Regressions." *The Stata Journal* 21 (2): 411–29.
OECD. 2016. "Who Uses Childcare? Background Brief on Inequalities in the Use of Formal Early Childhood Education and Care (ECEC) Among Very Young Children." OECD.
ONAPE. 2020. "L'accueil Du Jeune Enfant En 2019." Paris, France: CNAF.
Pavolini, Emmanuele, and Wim Van Lancker. 2018. "The Matthew Effect in Childcare Use: A Matter of Policies or Preferences?" *Journal of European Public Policy* 25 (6): 878–93.
Penn, Helen. 2007. "Childcare Market Management: How the United Kingdom Government Has Reshaped Its Role in Developing Early Childhood Education and Care." *Contemporary Issues in Early Childhood* 8 (3): 192–207.
Pepper, Gillian V, and Daniel Nettle. 2017. "The Behavioural Constellation of Deprivation: Causes and Consequences." *BEHAVIORAL AND BRAIN SCIENCES*, 1–66.
Pora, Pierre. 2020. "Keep Working and Spend Less? Collective Childcare and Parental Earnings in France." Working Paper. INSEE.
Roth, Alvin E. 2015. *Who Gets What–and Why: The New Economics of Matchmaking and Market Design*. Houghton Mifflin Harcourt.
Rubin, Donald B. 1974. "Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies." *Journal of Educational Psychology* 66(5): 688–701.
Sandstrom, Heather, and Ajay Chaudry. 2012. "'You Have to Choose Your Childcare to Fit Your Work': Childcare Decision-Making Among Low-Income Working Families." *Journal of Children and Poverty* 18 (2): 89–119.
Satz, Debra. 2010. "Noxious Markets." In *Why Some Things Should Not Be for Sale: The Moral Limits of Markets*, edited by Debra Satz, 0. Oxford University Press.
Simintzi, Elena, Sheng-Jun Xu, and Ting Xu. 2022. "The Effect of Childcare Access on Women's Careers and Firm Performance." SSRN Scholarly Paper. Rochester, NY. November 20, 2022.
Słoczyński, Tymon, S. Derya Uysal, and Jeffrey M. Wooldridge. 2022a. "Abadie's Kappa and Weighting Estimators of the Local Average Treatment Effect." *SSRN Electronic Journal*.
———. 2022b. "Doubly Robust Estimation of Local Average Treatment Effects Using Inverse Probability Weighted Regression Adjustment." August 2, 2022.
Sunstein, Cass R. 2022. "Sludge Audits." *Behavioural Public Policy* 6 (4): 654–73.
Tversky, Amos, and Daniel Kahneman. 1974. "Judgment Under Uncertainty: Heuristics and Biases" 185.
Vandenbroeck, Michel, and Arianna Lazzari. 2014. "Accessibility of Early Childhood Education and Care: A State of Affairs." *European Early Childhood Education Research Journal* 22 (3): 327–35.
Vandenbroucke, Frank, and Koen Vleminckx. 2011. "Disappointing Poverty Trends: Is the Social Investment State to Blame?" *Journal of European Social Policy* 21 (5): 450–71.
Walters, Christopher R. 2015. "Inputs in the Production of Early Childhood Human Capital: Evidence from Head Start." *American Economic Journal: Applied Economics* 7 (4): 76–102.