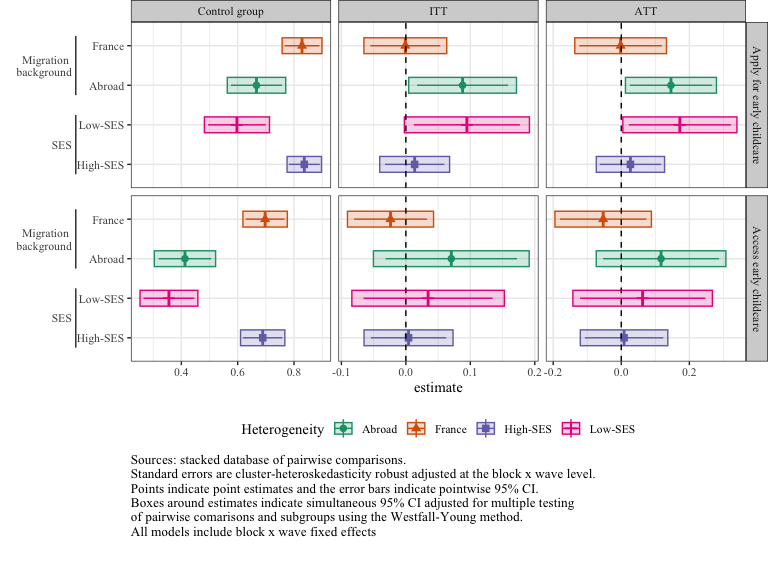
Investigating how administrative burden and search costs affect social inequalities in early childcare access, a randomised controlled trial

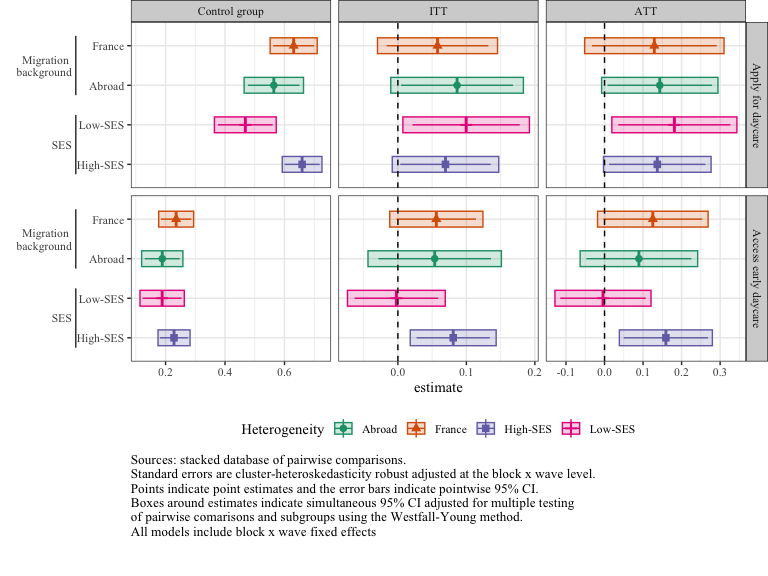
Manuscript submitted to Nature Human Behaviour

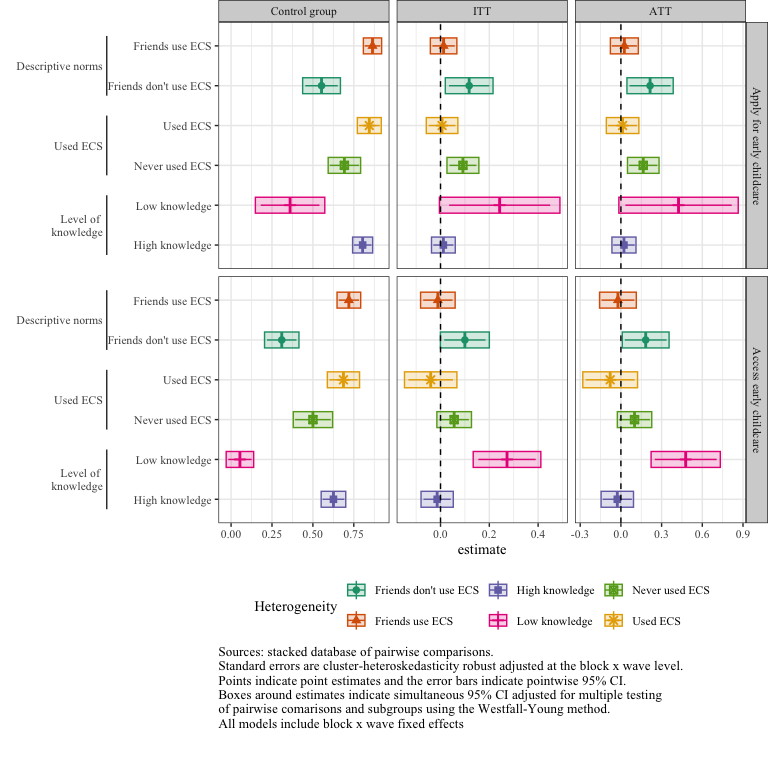
Abstract

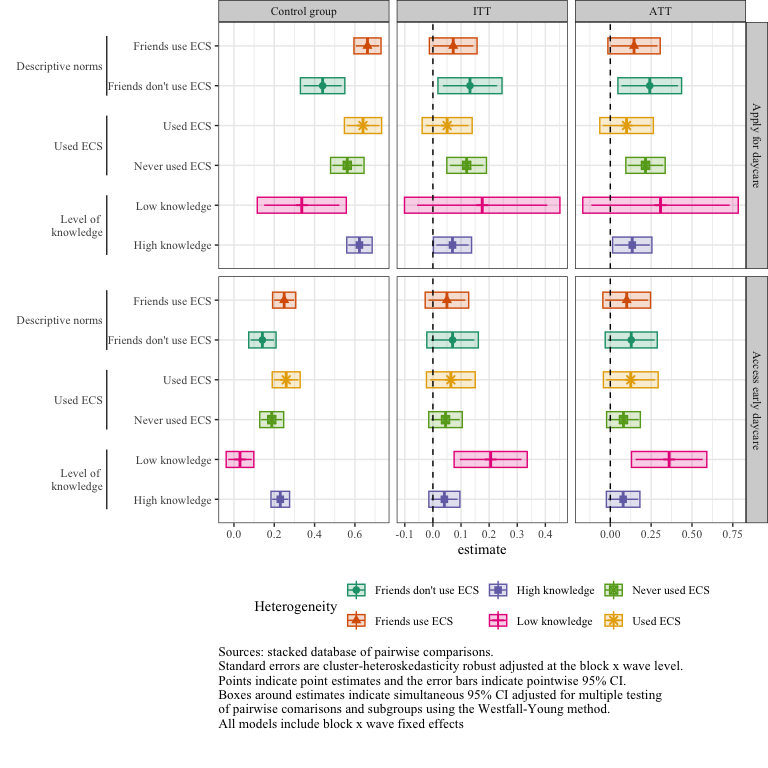
Access to high-quality early childcare for low socioeconomic status (SES) households has the potential to mitigate socioeconomic inequalities. Yet, there is an SES-based gap in early childcare enrolment. While low-SES households would benefit the most from attending early childcare, they access early childcare the least. This study tackles cognitive and behavioural barriers behind this access gap. We test the effectiveness of informational interventions and personalised support to enhance early childcare application and access for low-SES households through a multi-arm experiment. Results reveal that the information-only treatment had minimal impact while adding personalised support significantly bridged the SES-gap in early childcare applications. However, despite large impacts on application rates, we found limited impacts on access rates for low-SES households. By identifying key obstacles to early childcare access for low-SES households, our research underscores the need for effective strategies to promote equal opportunities in early childhood education.

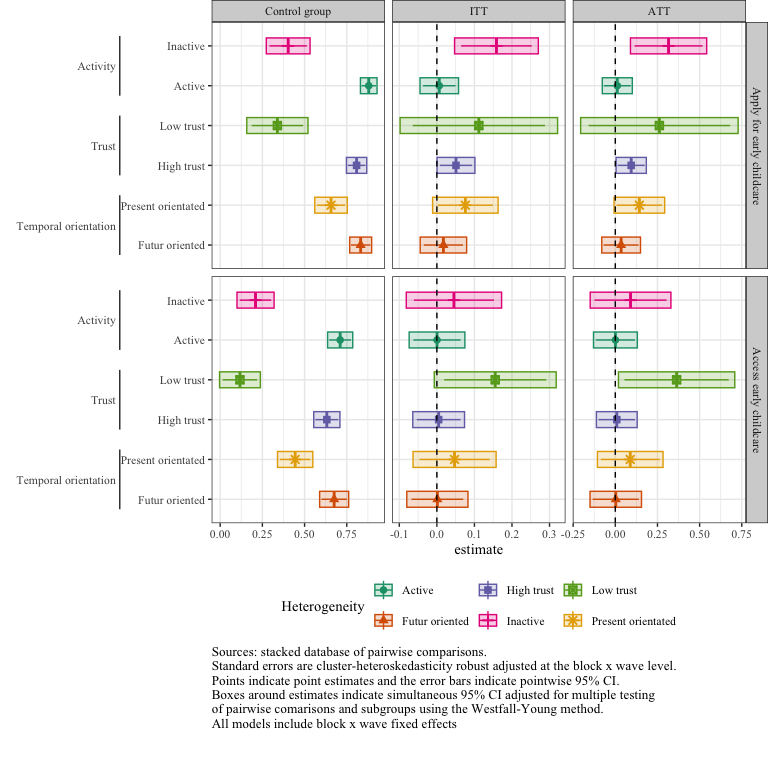
# 1 New plots











# 2 Methods

## 2.1 Analysis plan

### 2.1.1 Randomisation protocol

To improve precision and ensure balance on a set of characteristics likely to affect our main outcomes, we used a block-random assignment procedure based on the cross product of: i) Education of the mother (“tertiary”/“secondary or lower”), ii) Intention to use early childcare (“no”/“yes but has never used early childcare before”/“yes and already has used early childcare before”), and iii) Supply (early childcare coverage rate higher/lower than the average in the department).

We collected data over a three-month period and defined “waves” of two-weeks periods over which we performed the assignment procedure. Ultimately, our design was built on 6 randomisation waves with three assignment conditions in the 12 blocks of each lottery.

### 2.1.2 Assignment probabilities

Each individual from the baseline sample had a 1/3 probability of being assigned to one of three groups: control, information-only, or information + support. Since our analyses compared pairs of assignment conditions (e.g. information-only vs. control), our analyses relied on conditional assignment probabilities that exclude one treatment arm. However, some blocks had a small or odd number of observations, potentially resulting in slight variations in assignment probabilities across groups. To account for this, we estimated a probit model of assignment based on block-fixed effects in subsamples of pairwise comparisons. We then used the predicted probabilities from these models as propensity scores, denoted as , where subscripts denote the household of block in sub-sample of pairwise comparison . With these, we define inverse propensity score weights and centred assignment where is a dummy that equals 1 when the household of block has been assigned to the treatment group in sub-sample .

## 2.2 Intention to treat : main average effects

Because we have two treatment arms, we followed Goldsmith-Pickham (2024) and used stacked regressions to estimate the average difference between each pair of assignment condition , while avoiding contamination bias in our estimates. We used the following equation:

Where Y denotes the outcome of individual of block in sub-sample ; denotes block-wave dummies and are interacted with , a factor variable of sub-sample pairs of treatment arms. is a dummy that equals 1 when the household of block has been assigned to the treatment group in sub-sample . The estimates of correspond to the average intention to treat effect1. Following Chaisemartin-Ramirez (2022) and Abadie Et. Al. (2022), we used cluster-heterosckedasticity-robust standard errors adjusted at the block level2,3.This adjustment is very conservative as shown by Abadie Et. Al. (2022), and the detected effects correspond to the lower bound one could expect given these treatments3.

### 2.2.1 Robustness checks

As pre-registered, we assessed the robustness of our results using a data-driven selection of potential confounders with post-lasso and estimated both pooled and fully-interacted regressions following Negi Wooldrige (2021) and Lin (2013), allowing heterogeneous treatment effects within blocks4,5. The pooled regression use OLS to estimate the following equation:

And Lin regressions:

With , the matrix of covariates selected by the lasso method and previously centred.

Lasso selection and estimations were performed separately for each comparison pair and inference was based on point-wise standard errors clustered at the block-level.

We also reproduced our main analyses using a linear probability model with one dummy for each treatment arm and strata fixed effects, and logistic regressions.

## 2.3 Local average treatment effect of administrative support

In our setting, administrative support was only offered to those assigned to the information + support treatment group, making non-compliance one-sided. Out of the households that were offered the support, 52% opted for it on average (See Extended Data Table 2). Thus, under exclusion restriction, an instrumental variable strategy using assignment to the information + support treatment group to predict compliance with support can retrieve the Local average treatment effect (LATE), which can have, an “average treatment effect on the treated” interpretation6,7.In this setting, there were two pairs of comparison: Information + support control or Information + support information only. The former identifies the treatment effects of Denoting the compliance status, we estimated the average treatment effect on the treated through the following system of equations using weighted TSLS:

The instrument was the dummy for being in the information + support group demeaned within block using the estimated instrument propensity score. We interpreted the coefficient as the average first stage effect. As Borusyak Et Al (2024) showed, this model retrieved a weighted average of block-specific treatment effects**BorusyakEtAl2024?**. We used cluster robust standard errors adjusted at the block level.

## 2.4 Conditional average treatment effects

Our strategy was to use subgroup analysis to see whether different populations respond to the treatment differently in accordance with our theories. However, these differences in conditional average treatment effects could only be interpreted as a descriptive measure of association between the group and the treatment effect, but would not represent the causal effect of a change in the group value on the ATE since the group is not randomly manipulated. We estimated average treatment effects by pre-specified subgroups following the same estimation strategy as the main models presented above. We retrieved conditional average effects using fully saturated stacked regressions interacting all right hand side variables with subgroup dummies (see Goldsmith-Pinkham et Al. (2022))8. In the two stage least squares models, we estimated one first stage for each sub-group.

# 3 Inference and tests

We adjusted p-values and confidence intervals to account for the Family-wise error rates using the Young-Westfall method using the R package multcomp9. Our multiple testing procedure adjusted, for each outcome, the p-values and confidence intervals for simultaneous inference on the 3 comparisons (information - control, information + support - control, information + support - information only).

All analyses were performed using R 4.3.0 and R studio 2023.12.1.

# 4 Mechanisms: decomposing the intention-to-action gap and treatment effects

## 4.1 Intention gap by social groups

We analysed the differences in baseline intention to use childcare, application and access at endline across different subgroups. Formally, the gap between two sets of values of covariates in intention to use at baseline is:

Another parameter of interest is the average intention to action gap the expected individual difference , and differences between social groups. To be clear, this variable equals 0 when people are consistent in their reported intention and application behaviours, equals 1 when the person applied but didn’t intend to (*Switchers*) and -1 when they intended to but did not apply (*Quitters*).

The conditional gaps in intention and application can thus be rewritten:

In words, difference between de average gap in application and the average gap in intention between groups is the difference in the expected intention to action gap between groups. The latters being weighted average of the share of switchers and quitters, we could decompose the change in application behaviours across groups by the changes in the share of switchers and quitters.

In Table , we tested the difference in the share of quitters and switchers by social groups. For switchers, there were no difference by SES status. However for quitters, Low-SES parents - who were already less likely to intend to use childcare – were twice more likely to quit.

Intention to application gap by SES status

|  | Group lhs: Switcher | Group lhs: -Quitter | Diff lhs: Switcher | Diff lhs: -Quitter |
| --- | --- | --- | --- | --- |
| Educ2High-SES | 0.049\*\*\* | -0.085\*\*\* |  |  |
|  | (0.012) | (0.016) |  |  |
| Educ2Low-SES | 0.070\*\*\* | -0.182\*\*\* | 0.021 | -0.097\*\*\* |
|  | (0.019) | (0.028) | (0.022) | (0.032) |
| (Intercept) |  |  | 0.049\*\*\* | -0.085\*\*\* |
|  |  |  | (0.012) | (0.016) |
| Num.Obs. | 494 | 494 | 494 | 494 |
| R2 | -0.000 | 0.019 | 0.002 | 0.021 |
| R2 Adj. | -0.002 | 0.017 | -0.000 | 0.019 |
| * p < 0.1, \*\* p < 0.05, \*\*\* p < 0.01 | | | | |

# 5 Treatment effects on quitter and switchers.

Now that we had seen that there are more quitters than switchers and that the gap in quitters were twice as large among low SES group, we looked at the effect of information and administrative support on these probabilities.

## 5.1 Average estimations

Table 5.1: ITT on Switching and quitting status

|  | Gap | Quitter | Switchers |
| --- | --- | --- | --- |
| Information-only vs Control | -0.01 (0.02) | -0.00 (0.02) | -0.01 (0.01) |
|  | [-0.06, 0.04] | [-0.05, 0.04] | [-0.03, 0.01] |
|  | adj.p.val. = 0.762 | adj.p.val. = 0.829 | adj.p.val. = 0.553 |
| Information + Support vs Control | 0.04\*\* (0.02) | -0.04\*\* (0.02) | 0.00 (0.01) |
|  | [-0.01, 0.09] | [-0.08, 0.00] | [-0.03, 0.03] |
|  | adj.p.val. = 0.099 | adj.p.val. = 0.059 | adj.p.val. = 0.719 |
| Information + support vs Information-only | 0.05\*\* (0.02) | -0.04\* (0.02) | 0.01 (0.01) |
|  | [-0.01, 0.10] | [-0.08, 0.01] | [-0.02, 0.04] |
|  | adj.p.val. = 0.099 | adj.p.val. = 0.094 | adj.p.val. = 0.672 |
| Mean control group | -0.07 (0.02) | 0.12 (0.02) | 0.05 (0.01) |
|  | [-0.12, -0.01] | [0.08, 0.16] | [0.02, 0.09] |
| Num.Obs. | 2906 | 2906 | 2906 |
| R2 | 0.272 | 0.144 | 0.354 |
| R2 Adj. | 0.202 | 0.062 | 0.292 |
| Fixed effects | X | X | X |
| Chi 2 | 6.92 | 6.43 | 4.45 |
| P-value | 0.074 | 0.092 | 0.217 |
| Sources: stacked database of pairwise comparisons.   \*= p<.1, \*\*= p<.05, \*\*\*= p<.01 based on point-wise p-value.  Standard errors are cluster-heteroskedasticity robust adjusted at the block x wave level.  Adjusted p-value and confidence intervals account for simultaneous inference using the Westfall method.   Each column estimates jointly the effects of the program using fully-saturated stacked regressions. Control means estimated separately by OLS.  Joint significance test of null effect using Chi-2 test and p-value are reported at the bottom of the table. | | | |

1. Athey, S. & Imbens, G. W. [Chapter 3 - the econometrics of randomized experiments](https://doi.org/10.1016/bs.hefe.2016.10.003). in *Handbook of field experiments* (eds. Banerjee, A. V. & Duflo, E.) vol. 1 73–140 (North-Holland, 2017).

2. de Chaisemartin, C. & Ramirez-Cuellar, J. At What Level Should One Cluster Standard Errors in Paired Experiments, and in Stratified Experiments with Small Strata? *SSRN Electronic Journal* (2020).

3. Abadie, A., Athey, S., Imbens, G. W. & Wooldridge, J. M. [When Should You Adjust Standard Errors for Clustering?](https://doi.org/10.1093/qje/qjac038) *The Quarterly Journal of Economics* **138**, 1–35 (2022).

4. Negi, A. & Wooldridge, J. M. Revisiting Regression Adjustment in Experiments with Heterogeneous Treatment Effects. *Econometric Reviews* **40**, 504–534 (2021).

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

6. Frölich, M. & Melly, B. Identification of Treatment Effects on the Treated with One-Sided Non-Compliance. *Econometric Reviews* **32**, 384–414 (2013).

7. Imbens, G. W. & Angrist, J. D. Identification and Estimation of Local Average Treatment Effects. 12 (1994).

8. Goldsmith-Pinkham, P., Hull, P. & Kolesár, M. *Contamination bias in linear regressions*. (2022).

9. Bretz, F., Hothorn, T. & Westfall, P. *Multiple Comparisons Using R*. (Chapman and Hall/CRC, 2010). doi:[10.1201/9781420010909](https://doi.org/10.1201/9781420010909).