Formal stuff POSAJE

Arthur Heim

2025-01-31

We model the decision to apply for childcare as a function of perceived eligibility, transaction costs, and potential benefits. Households face uncertainty regarding their eligibility and the costs associated with the application process. The intervention aims to reduce these frictions through information provision and administrative support.

We consider a sequential decision process where at the first period (baseline), a pregnant woman has *ex-ante* unobserved **preference profile** θ_0 with regard to future childcare options. We observe a set of observable attributes $\mathbf{X}$ and intention to use I_0. Among X, there are social group metrics (G) such as SES status (based on education), migration background, and rough metrics of temporal preferences ρ and knowledge of childcare $\mathcal{I}$. There are also cognitive and structural factor that may hinder decisions. Let $\mathcal{C}$ denote the subjective application costs.

In the absence of intervention, the decision to apply can define a latent propensity to apply:

$$I_0^* = \mu_w(\mathbf{Z}) - U_I(\mathcal{I}, \theta, \mathcal{C}), \quad I_0 = 1 \; if \; I^* \geq 0, \quad I_0 = 0 \; otherwise \tag{1}$$

Where $\mathbf{Z}$ is a set of observables and U_I is unobserved and depend on information sets, preference types and subjective costs of application. We assume that $\mathbf{Z} \perp U_I$

.1 Intention gap by social groups

Formally, the gap between two sets of values of covariates $\mathbf{x}$ $\mathbf{x}'$ in intention to use at baseline is:

$$\Delta_{x,x'}(I_0) = \mathbb{E}\Big[I_{0i}|\mathbf{X} = \mathbf{x}\Big] - \mathbb{E}\Big[I_{0i}|\mathbf{X} = \mathbf{x}'\Big] \tag{2}$$

With a random sample, these conditional differences in reported intention to use childcare can be estimated using OLS and heteroskedasticity robust standard errors (Linear probability model) or by first estimating a model with parametric restrictions on I_0^* (such as logit or probit) and computing marginal effects.

However, these parameters and sample analogues are only descriptive difference in means between two groups who also differs in observable characteristics and latent U_I.

We assume that the sampling process of our baseline sample is randomly drawn from the superpopulation These difference in probabiliry

Idée générale : Present oriented –> future benefits and costs Utility of with or without childcare - cost of childcare + subjective costs

How reducing friction costs may affect different people (choice function)

.2   Adaptation of Castell et al. (2024)

Pregnant woman must decide in the future whether they want childcare. Let b be the probability of receiving childcare (i.e. successful application). Application is an action $A \in \{0,1\}$ equals to 1 if the access probability perceived with some error ϵ exceeds application costs c.

$$A = \mathbb{1}(b + \epsilon - c > 0) \tag{3}$$

A mother may not apply if she lack information or underestimate her chance of access ($\square < 0$), she may face transaction costs c which are non-monetary costs going through application process such as complexity, delays, and so on.

A)   Information

Each individual is characterized by the value of the cost c and the error ε they make in the perception of the probability of access. We denote with $F(x) = P(b < x)$ the cumulative distribution function of benefits b given c and ϵ. The probability to apply is:

$$P(A = 1) = P(b > c - \epsilon) = 1 - F(c - \epsilon) \tag{4}$$

Providing full information on application likelihood has the following effect on application:

$$\Delta(c, \epsilon) = P(b > c) - P(b > c - \epsilon) = F(c - \epsilon) - F(c) \tag{5}$$

The effect of the information treatment is thus positive for agents who underestimate their access probability ($\epsilon < 0$) and negative otherwise. The ITT effect sums $\Delta(c, \epsilon)$ over the joint distribution of (c, ϵ).

Therefore, providing information increases application if, under imperfect information, there are more individuals who would benefit from applying but do not apply because they underestimate their eligibility than individuals whose net benefit is negative but who apply because they overestimate their eligibility.

Given our empirical results, it is useful to consider the reasons for why providing information may have no or very small effects on applications. First, it could be that on average agents have accurate beliefs, even though many of them are wrong about their own eligibility. Second, it could be that the effect of erroneous beliefs is small as compared to that of transaction costs, such that for most people with $F(c - \epsilon) = 1$ (no take-up), we also have $F(c) = 1$. Finally, it could simply be that ϵ is uniformly zero, i.e. that there are no information frictions.

.3   Administrative support

Application costs. We next investigate the impact of reducing application costs. We define c_0, as the application cost one faces when supported by a caseworker, i.e. when one receives treatment T2.

On their own (untreated), agents face a larger cost $c \in [c_0, \infty[$ that is distributed in the population. Given our empirical results, let us consider a situation where imperfect information can be neglected ($\varepsilon = 0$). Then, the effect of this treatment for any agent of type c is:

$$\Delta(c) = P(b > c_0) - P(b > c) = F(c) - F(c_0) \geq 0$$

The population effect sums $\Delta(c)$ over the marginal distribution of c, and is positive. Note that in this case, the probability to apply and the probability to access are identical; the latter is more easily observable. The effect is stronger on agents that face higher transaction costs in the absence of treatment:

$$\frac{\partial \Delta(c)}{\partial c} = f(c) > 0$$

To understand the effect of application costs on access, we also consider the treatment effect on access by the average claimant:

$$\Delta^b(c) = E\left(b \mid b > c_0\right) - E(b \mid b > c) = \frac{\int_{c_0}^{\infty} b f(b) db}{1 - F(c_0)} - \frac{\int_{c}^{\infty} b f(b) db}{1 - F(c)}$$

For any agent of type c this quantity is negative if $c > c_0$ and zero otherwise. Integrating over c yields:

$$\Delta^b = \int_0^{\infty} \Delta^b(c) g(c) dc$$

If b and c are uncorrelated in the population, then Δ^b is negative: lowering application costs extends access to people with lower claims. This corresponds to the idea that application costs are ordeals mechanisms that improve targeting (Kleven and Kopczuk, 2011; Alatas et al., 2016). If b and c are sufficiently positively correlated however, i.e. if those most likely to get access also have higher costs, then Δ^b can be positive. In that case lowering application costs can attract more parents who are more likely to access childcare (Deshpande and Li, 2019; Gupta, 2017).

.4   Compliance

In practice, the intervention is a phone call to which potentially eligible people may not answer or refuse the administrative support. In addition to the decision to apply to childcare, we need to model the decision to comply with the administrative support treatment as we can only identify treatment impact on compliers. Let us assume that mothers decide to take support by comparing the anticipated reduced cost of application $(b - c_0)$ with nuisance from the intervention which we denote with $\kappa > 0$. An agent complies with the invitation if:

$$b > c_0 + \kappa$$

In our analysis, we are interested not only in the average treatment effect, but also in the shape of the marginal treatment effect as a function of the cost of participation κ, following Heckman (2010). We define the marginal treatment effect (MTE) as the effect of the treatment on the compliers, conditional on a value of the cost of administrative support, κ, for any individual with application cost c. The MTE is:

$$\text{MTE}(\kappa; c) = P\left(b > c_0 \mid b > c_0 + \kappa; c\right) - P\left(b > c \mid b > c_0 + \kappa; c\right)$$
$$= 1 - \frac{1 - F(c)}{1 - F(c_0 + \kappa)}$$

if $\kappa < c - c_0$ and zero otherwise. Conditional on c, we have:

$$\frac{\partial MTE(\kappa; c)}{\partial \kappa} \leq 0$$

As κ grows, compliers are more and more selected towards higher chance of access b; they find value in attending the meeting because it reduces their administrative burden, but the share that would apply even in the absence of the treatment grows, so that the treatment effect decreases. This is true for a given c. We take the integral over c to obtain the MTE for the whole population:

$$\text{MTE}(\kappa) = \int_{c_0}^{\infty} MTE(\kappa; c) g(c \mid \kappa) dc$$

When κ increases, every MTE($\kappa; c$) in the integral decreases. If c and κ are uncorrelated, then MTE(κ) will also be a decreasing function of κ. But if c and κ are positively correlated, the MTE can be increasing in κ, since the large

MTE$(\kappa; c)$ with large c receive a higher weight $g(c \mid \kappa)$. Claimants who face large transaction costs and would benefit the most from administrative support also face large costs of accepting it.

I Setting

I.1 Baseline

Pregnant woman i at time t_0 have a set of observable attributes bundled in a matrix $\mathbf{A}_0$ and an unobservable set of information $\mathcal{I}$. In addition, Let $\mathbf{X} \subset \mathbf{A}$ denote fixed attributes (e.g. measures of social groups) and let $\mathbf{K} \subset \mathbf{A} \backslash X$ some observed metrics for preferences, knowledge, norms, values, and so on. In particular, we observe women intention to use childcare and their most favourite childcare arrangement.

$\mathbf{X}$ and $\mathbf{K}$ may be correlated as unobserved factor determine both observed characteristics such as social groups and preferences, norms, intention to use childcare…

A) Ex-ante preference profiles

Woman must choose whether she applies to formal childcare and more precisely which childcare types. Assume for simplicity that the choice set is sparse such that she can choose to apply to Daycare D, Childminder C and if she doesn't apply to one of the two, the default is parental care P. Woman i has $ex-ante$ preference order over childcare solutions we denote $\succeq_i$ which can be:

$$\succ_i \equiv \begin{cases} \left.\begin{array}{l} a := D \succeq_i C \succeq_i P \\ b := C \succeq_i D \succeq_i P \end{array}\right\} \text{childcare enthusiasts} \;\; ; \\[1em] \left.\begin{array}{l} c := D \succeq_i P \succeq_i C \\ d := C \succeq_i P \succeq_i D \end{array}\right\} \text{Selective parents} \;\; ; \\[1em] \left.\begin{array}{l} e := P \succeq_i C \succeq_i D \\ f := P \succeq_i D \succeq_i C \end{array}\right\} \text{Parental care} \;\; ; \end{cases} \tag{6}$$

We can also bundle preference profiles by "daycare enthusiasts", "childminder enthusiasts".

$$\succ_i \equiv \begin{cases} \left.\begin{array}{l} a := D \succeq_i C \succeq_i P \\ c := D \succeq_i P \succeq_i C \end{array}\right\} \text{Daycare enthusiasts} \;\; ; \\[1em] \left.\begin{array}{l} d := C \succeq_i P \succeq_i D \\ b := C \succeq_i D \succeq_i P \end{array}\right\} \text{Childminder enthusiasts} \;\; ; \\[1em] \left.\begin{array}{l} e := P \succeq_i C \succeq_i D \\ f := P \succeq_i D \succeq_i C \end{array}\right\} \text{Parental care} \;\; ; \end{cases} \tag{7}$$

At t_0, childcare enthusiasts and selective parents reveal their intention to use and most preferred childcare. therefore, we observe the following preference types in the data. Let W denote intention to use.

- those who intend to use childcare and prefer daycare have preference types $\succeq_i \in \{a, c\}$: the daycare enthusiasts
- those who intend to use childcare and prefer childminders have preference types $\succeq_i \in \{b, d\}$: the childminder enthusiasts
- Those who do not intend to use childcare have preference types $\succeq_i \in \{e, f\}$: those who would rather care for their children themselves.

If we knew if the parent would accept another childcare and assume strict preference order, we could classify parents of all 6 preference types at baseline. But we don't. Therefore, these types are unknown.

Moreover, these preference types evolve over time as parents' situation, information sets, needs etc. change. There are also cognitive and structural factor that may hinder decisions. Let $\mathcal{C}$ denote the subjective application costs.

In the absence of intervention, the decision to apply can define a latent propensity to apply:

$$W^*(0) = \mu_w(\mathbf{Z}) - U_w(\mathcal{I}, \succeq, \mathcal{C}), \quad W = 1 \; if \; W^* \geq 0, \quad W = 0 \; otherwise \tag{8}$$

Where $\mathbf{Z}$ is a set of observables and U_w is unobserved and depend on information sets, preference types and subjective costs of application.

I.2   Endline data

The endline survey includes the different sets of outcomes including some metrics of knowledges, application behaviours and access to formal childcare. We leave attrition issues for now and focus on application behaviours. We come back to these points later.

Let $\tilde{Y}_i = \mathbb{1}(\text{applied to any childcare})$ be the endline measure of application behaviours which derive from:

$$\tilde{Y}_i = \underbrace{\mathbb{1}(\text{applied for childminders})}_{Y^c} + \underbrace{\mathbb{1}(\text{applied for daycare})}_{Y^d}$$

In practice, $\tilde{Y}$ also include other childcare arrangement.

The relationships between the observed variables from Baseline and Endline can be represented using Directed Acyclic Graphs such as Figure 1. This model is not identified. No causal relationship can be inferred.

Figure 1: How observed baseline characteristics relate with the main outcome.
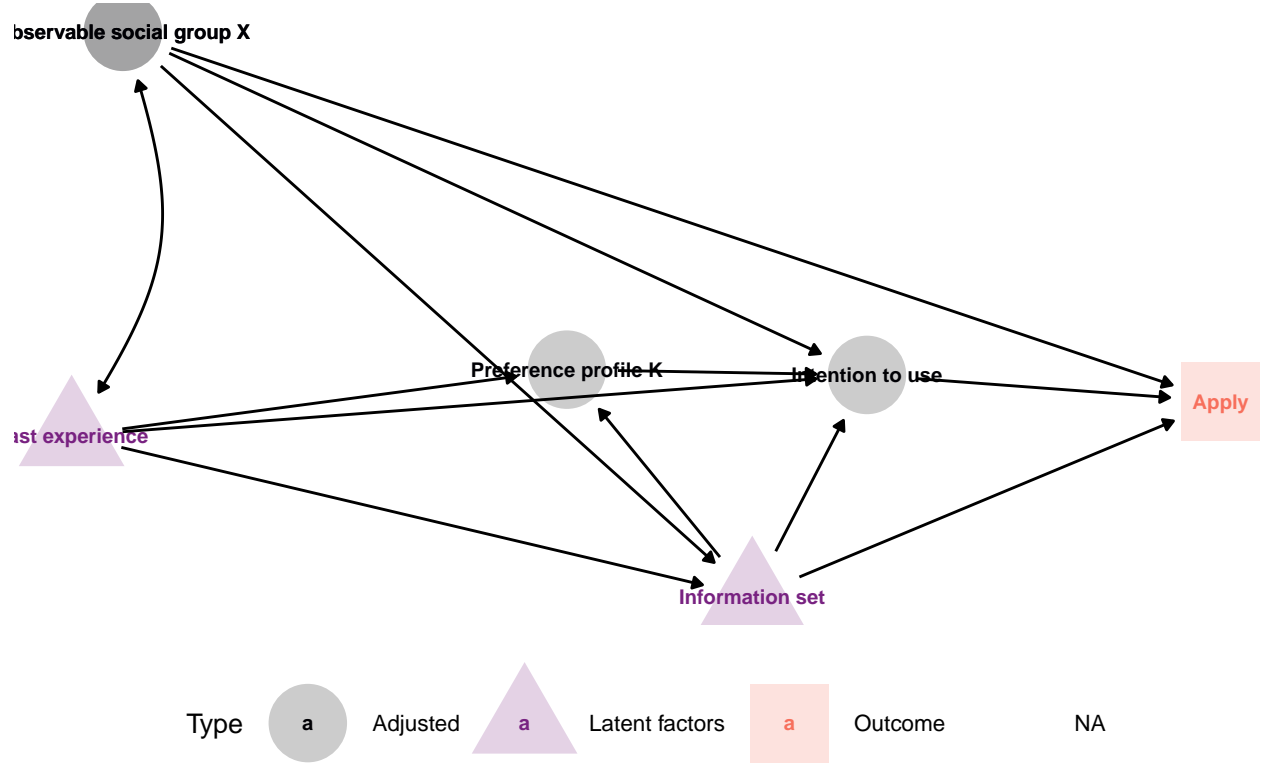
I.3   Intention to action gap

We analyse the intention-to-action gap between different social groups (High/low SES, migration background) (**X**) and subjective attributes (Information, temporal orientation) (**K**).

Formally, the gap between two sets of values of covariates **x x**$'$ in intention to use at baseline is:

$$\Delta_{x,x'}(W_i) = \mathbb{E}\Big[W_i|\mathbf{X}=\mathbf{x}\Big] - \mathbb{E}\Big[W_i|\mathbf{X}=\mathbf{x}'\Big] \tag{9}$$

The application gap is :

$$\Delta_{x,x'}(\tilde{Y}_i) = \mathbb{E}\Big[\tilde{Y}_i|\mathbf{X}=\mathbf{x}\Big] - \mathbb{E}\Big[\tilde{Y}_i|\mathbf{X}=\mathbf{x}'\Big] \tag{10}$$

Denoting $\tilde{Y}*$ the variable for access to formal childcare, The access gap is :

$$\Delta_{x,x'}(\tilde{Y}*_i) = \mathbb{E}\Big[\tilde{Y}*_i|\mathbf{X}=\mathbf{x}\Big] - \mathbb{E}\Big[\tilde{Y}*_i|\mathbf{X}=\mathbf{x}'\Big] \tag{11}$$

In our setting, the covariates are discrete. These three parameters can therefore be estimated from the control group by regressing the outcome on group dummies using OLS and heterokedasticity robust standard errors, or by computing marginal effects from a logit or probit model. OLS estimates are reported in Figure 2.

Figure 2: The intention-to-action gap in early childcare application and access in the control group across four dimensions of heterogeneity. Intention to use early childcare is measured at the baseline survey during pregnancy. Early childcare application and early childcare access are measured at the endline survey one year after.
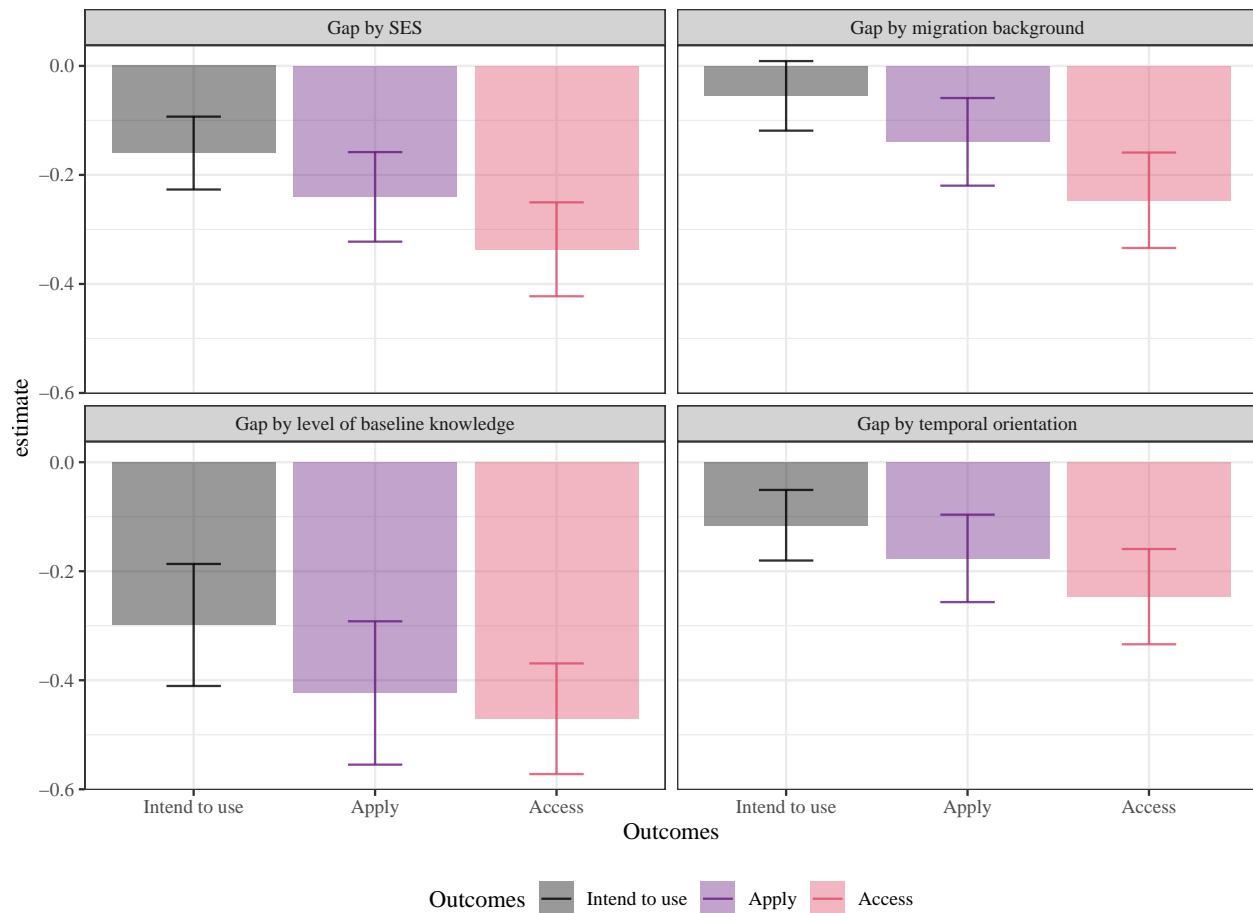
Sources: Control group only. Intention are measured at the baseline survey during pregnancy (Q4 2022); application and access are measured at the endline survey one year after (Q4 2023).
Notes: Coefficients of OLS regressions of the outcomes on dummies for the group variable of interest. Error bars indicate pointwise 95% CI based on heteroskedasticity–robust standard errors (HC1). Intention–to–action gap in early childcare application and access gap in the control group across four dimensions of heterogeneity:
– Gap by SES compares households (HH) in which the mother did not attend any kind of post–secondary education with those in which the mother did.
– Gap by migration background compares HH in which the mother was born abroad with HH in which the mother was born in France.
– Gap by level of baseline knowledge compares parents with very low knowledge with those wih higher knowledge of early childcare.
– Gap by temporal orientation compares HH in which the mother prefers a reward of 50. in three days (defined as present–oriented) to HH in which the mother prefers 80. in 3 months.

If one think that because sampling was done in different maternity wards which may be used by different mothers such that there are intra-maternity ward correlation of outcomes, inference should rely on cluster-robust standard errors adjusted at the maternity ward level. Controlling for maternity ward dummies would remove the correlation with the social group variable and the between maternity ward difference in outcomes (by the FWL theorem).

Another parameter of interest is the intention to action gap, and differences between social groups. We can compute

1) the share of parents who applied by baseline intention
2) the shares of parents who applied by baseline intention and SES group

The first one shows how many parents switch preference types by moving from intention to no application and vice versa. There are 29% who didn't want childcare but did apply ; and 85% who intended to use but didn't apply.

	Simple dif	Residualised MW	Fixed Effect
(Intercept)	0.864	0.000	
	(0.018)	(0.016)	
Educ2Low-SES	−0.160		
	(0.034)		
resEduc		−0.153	
		(0.035)	
LowSES			−0.153
			(0.035)
Num.Obs.	623	623	623
R2	0.038	0.034	0.041
R2 Adj.	0.037	0.032	0.028
R2 Within			0.034
R2 Within Adj.			0.032
AIC	603.7	602.1	616.1
BIC	612.5	610.9	656.0
RMSE	0.39	0.39	0.39
Std.Errors	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robust
Mean of DV	0.80	-0.00	0.80
FE: MaternityWardBaseline			X

The second column separate these averages by SES group. Interestingly, Low SES groups who did not intend to use are 24% to apply while High SES groups are 38% to switch. However, 90% High SES group who intended to use did apply while only 74% of Low SES did. High SES parents who already intended to apply are more likely to commit to this intention and those who didn't intend to use are more likely to have changed their mind. On the contrary, Low-SES who intended to use daycare are less likely to commit to their choice and those who didn't intend to use are less likely to change their mind.

$$\Delta_{x,x'}(\tilde{Y}_i - W_i) = \qquad\qquad \mathbb{E}\left[\tilde{Y}_i - W_i | \mathbf{X} = \mathbf{x}\right] - \mathbb{E}\left[\tilde{Y}_i - W_i | \mathbf{X} = \mathbf{x}'\right] \qquad (12)$$

$$= \qquad \underbrace{\mathbb{E}\left[\tilde{Y}_i | \mathbf{X} = \mathbf{x}\right] - \mathbb{E}\left[\tilde{Y}_i | \mathbf{X} = \mathbf{x}'\right]}_{\Delta_{x,x'}(\tilde{Y}_i)} - \underbrace{\left(\mathbb{E}\left[W_i | \mathbf{X} = \mathbf{x}\right] - \mathbb{E}\left[W_i | \mathbf{X} = \mathbf{x}'\right]\right)}_{\Delta_{x,x'}(W_i)} \qquad (13)$$

This parameter can be estimated either by regressing the difference between application and intention on social group dummies or by:

1) stacking baseline and endline databases
2) defining an outcome Y which equals W in the baseline and $\tilde{Y}$ in the endline.
3) regressing Y on an endline dummy, an interaction with the group indicator and mother fixed effects and clustered SE

The coefficient of the interaction is an estimate of $\Delta_{x,x'}(\tilde{Y}_i - W_i)$.

The following table show these estimates by high/low education.

	(1)	(2)	(3)
factor(ECSPlanToBaseline)0	0.295		
	(0.047)		
factor(ECSPlanToBaseline)1	0.850		
	(0.018)		
ECSPlanToBaseline = 0 × Educ2 = High-SES		0.375	
		(0.077)	
ECSPlanToBaseline = 0 × Educ2 = Low-SES		0.236	
		(0.058)	
ECSPlanToBaseline = 1 × Educ2 = High-SES		0.903	
		(0.018)	
ECSPlanToBaseline = 1 × Educ2 = Low-SES		0.742	
		(0.038)	
i(Educ2, ECSPlanToBaseline, ref = "High-SES")			−0.022
			(0.105)
ECSPlanToBaseline			0.528
			(0.079)
Educ2High-SES			0.375
			(0.077)
Educ2Low-SES			0.236
			(0.058)
Num.Obs.	494	494	494
R2	0.249	0.278	0.278
R2 Adj.	0.247	0.273	0.273
AIC	445.7	430.4	430.4
BIC	454.1	447.2	447.2
RMSE	0.38	0.37	0.37
Std.Errors	Heteroskedasticity-robust	Heteroskedasticity-robust	Heteroskedasticity-robus
Mean of DV	0.74	0.74	0.74

Table 1: Intention to application gap by SES status

	Group average	OLS	DID
Educ2High-SES	−0.036*		
	(0.021)		
Educ2Low-SES	−0.112***	−0.076*	
	(0.036)	(0.041)	
(Intercept)		−0.036*	
		(0.021)	
Baseline			−0.036*
			(0.021)
Baseline × Educ2Low-SES			−0.076*
			(0.041)
Num.Obs.	494	494	1117
R2	0.006	0.008	0.781
R2 Adj.	0.004	0.006	0.503

* p <0.1, ** p <0.05, *** p <0.01

II   The interventions

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

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

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.

Our intervention use intention to use and SES among the blocking variable. They are interacted together and with cover rate and randomisation waves.

II.1   The information treatment

We consider the average intention-to-treat effect of information:

$$ITT(1) = \mathbb{E}\Big[\tilde{Y}(1)|Z_i = 1\mathbf{B}\Big] - \mathbb{E}\Big[\tilde{Y}(0)|Z_i = 0\mathbf{B}\Big] \tag{14}$$

We estimate this model running an OLS of application on a dummy for $Z = 1$, block x wave fixed effects and clustered SE at the block w wave level. adding an interaction of the treatment with a group variable X (where X is also in the B fixed effects) We can also run a similar model for the

10

Table 2: Intention to application gap by SES status

	ITT	Interaction
Z	−0.006	0.014
	(0.021)	(0.045)
Z × High_SES		−0.031
		(0.049)
Num.Obs.	960	960
R2	0.350	0.350
R2 Adj.	0.288	0.287
RMSE	0.35	0.35

* p <0.1, ** p <0.05, *** p <0.01