



**HAL**  
open science

# Essays on the Role of Information in Job Search Behavior and Demand for Training

Esther Mbih

► **To cite this version:**

Esther Mbih. Essays on the Role of Information in Job Search Behavior and Demand for Training. Economics and Finance. Institut Polytechnique de Paris, 2020. English. NNT : 2020IPPAG014 . tel-03203671

**HAL Id: tel-03203671**

**<https://theses.hal.science/tel-03203671>**

Submitted on 20 Apr 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



INSTITUT  
POLYTECHNIQUE  
DE PARIS

NNT : 2020IPPAG014

Thèse de doctorat



# Essays on the Role of Information in Job Search Behavior and Demand for Training

Thèse de doctorat de l'Institut Polytechnique de Paris  
préparée à l'École Nationale de la Statistique et de l'Administration Économique

École doctorale n°626 Institut Polytechnique Paris (ED IP Paris)  
Spécialité de doctorat : Sciences Économiques

Thèse présentée et soutenue à Paris, le 18 Décembre 2020, par

**ESTHER MBIH**

Composition du Jury :

Arne Uhendorff Researcher, CNRS, CREST	Président
Béatrice Boulu-Reshef Professeur des Universités, Université d'Orléans	Rapporteur
Pierre Cahuc Professeur des Universités, IEP Paris	Rapporteur
Bruno Crépon Professeur, CREST-ENSAE	Directeur de thèse

## Remerciements

La réalisation de cette thèse n'a été possible que grâce aux nombreuses personnes dont j'ai reçu le soutien. Aussi, je tiens à remercier en tout premier lieu Bruno Crépon, pour m'avoir accordé sa confiance et avoir accepté de m'encadrer dans les présents travaux de recherche. Grâce à lui, j'ai eu l'occasion de participer à des projets de recherche passionnants et surtout de m'améliorer en tant qu'économiste et chercheur. Je tiens également à remercier Pierre Cahuc et Stéphane Carcillo, de la Chaire Sécurisation des Parcours Professionnels, pour leur soutien dans l'obtention de mon financement de thèse, ainsi que Francis Kramarz, pour m'avoir chaleureusement accueillie au CREST.

Je remercie aussi très sincèrement, Béatrice Boulu-Reshef et Pierre Cahuc pour avoir accepté d'être les rapporteurs de mon jury de thèse, ainsi qu'Arne Uhlenhorff pour faire partie de mon jury.

Cette thèse comporte des projets de groupe, et à cet égard, je voudrais remercier mes co-auteurs Bruno Crépon, Lucas Girard, Anett John, et Vincent Pons, avec qui j'ai beaucoup apprécié de travailler et dont j'ai beaucoup appris. Je remercie aussi Aïcha Ben Dhia, Louise Paul-Delvaux, et Bertille Picard, pour les longues journées (et parfois soirées) passées à Pôle emploi en ma compagnie pour récupérer et analyser les données nécessaires à nos projets.

Ma reconnaissance va également à Pierre Cahuc, Guillaume Hollard et Arne Uhlenhorff, pour leurs encouragements et leurs précieux conseils dont j'ai bénéficié au CREST, et qui m'ont permis d'avancer y compris dans les moments de doute.

Je tiens aussi à remercier la Direction Générale de Pôle emploi pour son accueil et son ouverture à la recherche et l'expérimentation. Je remercie particulièrement Anita Bonnet, pour m'avoir intégré au sein de son équipe, ainsi que François Aventure, Emmanuel Chion, Stéphane Ducatez, Audrey Perocheau, Cyril Nouveau et Chantal Vessereau, pour avoir permis la réalisation de mes projets de recherche. Je remercie également Guillaume Blache, Yohan Coder, Nicolas Devisme, Alexis Gaignon, Sandra Gaumont, Jérémy Giraudeau, Yannick Galliot, Sophie Hamman, Yves Janeau, Olivier Jugand, Christian Kostrubala, Yang-Ming Lim, Murielle Matus, Rose Morène, Roger Néron, Nicolas Prokovas, Eric Renard, Snjezana Smetisko, Raymonde Toquet, Nicolas Vanni, Jacques Wang, et René Yomelan, avec qui j'ai eu le plaisir d'échanger ou qui m'ont été d'une aide précieuse pour comprendre les données de Pôle emploi. Dylan Glover, ex-doctorant à Pôle emploi et Sciences Po durant presque toute ma thèse, a également été à cet égard d'une aide précieuse. J'adresse également à l'ensemble des conseillers, directeurs d'agences, et demandeurs d'emploi qui ont accepté de me recevoir et de répondre à mes nombreuses questions de sincères remerciements.

Ma reconnaissance va également à Bayes Impact et Paul Duan, pour m'avoir donné l'occasion d'évaluer avec mes co-auteurs le site internet Bob Emploi.

Au CREST, je remercie aussi mes ex "co-bureau" Maddalena Agnoli, Guillaume Bied, Pierre-Louis Bithorel, Guidogiorgio Bodrato, Mélina Hillion, Alicia Marguerie, Elia Perennes et Emilie Sartre, pour leur soutien tout au long de cette thèse et l'amitié qu'ils m'ont témoignée. J'adresse également un remerciement particulier à Reda Aboutajdine, Hélène Benghalem, Antoine Bertheau, Remi Avignon,

Marianne Blehaut, Pauline Carry, Gwen-jiro Clochard, Morgane Cure, Thomas Delemotte, Antoine Ferey, Christophe Gaillac, Manon Garrouste, Germain Gauthier, Etienne Guigue, Malka Guillot, Yannick Guyonvarch, Clémence Lenoir, Jeremy L'Hour, Elio Nimier David, Ivan Ouss, Bérangère Patault, Fabien Perez, Julie Pernaudet, Anasuya Raj, Audrey Rain, Daphné Skandalis, Clémence Tricaud, Jérôme Trinh, Benjamin Walter, Ao Wang et Meryam Zaiem.

Je remercie aussi Fabienne Kunne, qui m'a mentorée durant une année avec une grande bienveillance, et qui m'a apporté un important soutien ainsi que de précieux conseils.

Enfin j'adresse un sincère remerciement à tous les stagiaires et assistants de recherche qui ont contribué à mes projets de recherche : Pierre-Louis Bithorel, pour son sérieux et sa rigueur, mais également Alix Myczkowski, Quiterie Landèche et Marion Richard.

# Contents

List of Figures . . . . .	6
List of Tables . . . . .	7
<b>0 General introduction . . . . .</b>	<b>10</b>
0.1 Information and Job Search behavior . . . . .	10
0.2 Public policy evaluation - Methodolgy . . . . .	14
0.3 My dissertation . . . . .	16
<b>1 Can a Website Bring Unemployment Down? Effects of a French Online Platform on Job Search Efficiency . . . . .</b>	<b>18</b>
1.1 Introduction . . . . .	19
1.2 Setting . . . . .	24
1.3 Data . . . . .	27
1.4 Design . . . . .	30
1.5 Results . . . . .	37
1.6 Conclusion . . . . .	41
1.7 Tables . . . . .	43
1.8 Figures . . . . .	55
Appendices . . . . .	60
1.A Survey questionnaire . . . . .	60
1.B Machine learning algorithms: options and parameters . . . . .	63
<b>2 Do informational frictions affect enrollment in public-sponsored training? Results from an online experiment . . . . .</b>	<b>68</b>
2.1 Introduction . . . . .	69
2.2 Background . . . . .	74
2.3 Design . . . . .	77
2.4 Data and sample description . . . . .	80
2.5 Results . . . . .	84
2.6 Heterogeneity . . . . .	87
2.7 Discussion and conclusion . . . . .	89
2.8 Tables . . . . .	90

2.9	Figures	99
	Appendices	102
2.A	Translation of the baseline questionnaire	102
2.B	Heterogeneous effects without region fixed effects	104
<b>3</b>	<b>Behavioural Barriers to Human Capital Investment: Evidence from a Survey on French Jobseekers</b>	<b>106</b>
3.1	Introduction	107
3.2	Institutional Setting	109
3.3	The survey	112
3.4	Data and sample description	114
3.5	Results	117
3.6	Outlook	121
3.7	Conclusion	126
3.8	Tables	127
3.9	Figures	130
.1	Appendix - Construction of psychological variables analyzed in section 3.5.2 and results	138
.2	Appendix - Survey questionnaire	140
	References	168
	References	174

# List of Figures

0.1	Unemployment rate	11
0.2	Evolution of the French unemployment rate according to :	11
0.3	The randomization process	15
1.1	Timeline of the experiment	55
1.2	Average treatment effects on employment episodes	56
1.3	Average treatment effects on cumulative number of days in unemployment and unemployment benefits	57
1.4	GATES on unemployment durations 6 months after - generic machine learning approach	58
1.5	GATES on unemployment durations 18 months after - generic machine learning approach	59
2.9.1	Email types for each treatment group	99
2.9.2	Randomization design	100
2.9.3	Conversion rate of our emails into calls	100
2.9.4	Answers to baseline questions	101
3.9.1	Random forest - variable importance	130
3.9.2	Predicted probability of training enrollment	130
3.9.3	Informational constraints	131
3.9.4	Beliefs about quality	132
3.9.5	Training returns optimism	133
3.9.6	Financial constraints	134
3.9.7	Beliefs about the share of subsidized training in the <i>Pôle emploi</i> catalogue	135
3.9.8	Difference between the estimated proportion of subsidized training programs and the actual share in the region Ile-de-France	135
3.9.9	Main external barrier face by jobseekers	136
3.9.10	Main internal barrier faced by jobseekers	136
3.9.11	Design of the experiment	137
3.9.12	Power calculations	137
.1.1	Depression index	139
.1.2	Self-efficacy index	139
.1.3	Self-esteem index	139

# List of Tables

0.1	Distribution of training entrants according to the diploma . . . . .	12
1.1	Difference in means between the treatment and the control groups . . . . .	43
1.2	Survey Respondent Attrition . . . . .	44
1.3	Difference in means between survey respondents and non-respondents . . . . .	45
1.4	First Stage . . . . .	46
1.5	Difference in means between takers and non-takers in the treatment group . . . . .	47
1.6	Impact on reemployment expectations, effort and scope of job search . . . . .	48
1.7	Impact on job search methods . . . . .	49
1.8	Impact on well-being, motivation and life balance during job search . . . . .	50
1.9	Effect on unemployment duration 6 months after - results from regressions 1.4 . . . . .	51
1.10	Impacts on unemployment duration 6 months after . . . . .	52
1.10	Effect on unemployment duration 18 months after - results from regressions 1.4 . . . . .	53
1.11	Impacts on unemployment duration 18 months after . . . . .	54
2.8.1	Distribution of job seekers across regions and treatment arms . . . . .	90
2.8.2	Summary Statistics . . . . .	91
2.8.3	Balance table . . . . .	92
2.8.4	Impact on callback . . . . .	93
2.8.5	Impact on enrollment . . . . .	94
2.8.6	Impact on enrollment at Afpa as a fraction of total enrollment . . . . .	95
2.8.7	Summary statistics in the basic email group . . . . .	96
2.8.8	Heterogeneous impact of having a high level of formal education on callback . . . . .	97
2.8.9	Heterogeneous impact of having responded to baseline on callback . . . . .	98
2.B.1	Heterogeneous impact of having a high level of formal education on callback (regressions without region fixed effects) . . . . .	104
2.B.2	Heterogeneous impact of having responded to baseline on callback (regressions without region fixed effects) . . . . .	105
3.8.1	Probability to enroll into training 12 months after registering at <i>Pôle emploi</i> . . . . .	127
3.8.2	Distribution of strata within the sample . . . . .	127
3.8.3	Summary Statistics . . . . .	128
3.8.4	Returns : comparison of average beliefs per sector with true probabilities . . . . .	129



3.8.5 Distribution of jobseekers across barriers . . . . . 129

# Résumé substantiel

Cette thèse explore l'impact de l'information sur comportement de recherche d'emploi et la demande pour les formations professionnelles. Le premier chapitre évalue l'impact du site internet Bob Emploi, qui vise à délivrer de l'information aux demandeurs d'emploi à propos du marché du travail. Les résultats indiquent qu'il n'y a aucun impact sur l'effort de recherche des demandeurs d'emploi et le périmètre géographique et sectoriel de la recherche. Cependant, les demandeurs d'emploi ayant recours à Bob Emploi mobilisent davantage leur réseau personnel ainsi que les services publics de l'emploi. Enfin, il n'y a aucun effet sur le bien-être et sur le retour à l'emploi. Le deuxième chapitre examine le rôle de l'information sur l'entrée en formation professionnelle. Les résultats indiquent que la réception d'un email avec un message mettant l'accent sur les opportunités de retour à l'emploi après la formation fait plus que doubler la probabilité que les demandeurs d'emploi rappellent le centre de formation. Cependant, les taux d'appel sont faibles en valeur absolue (moins de 1%) et il n'y a aucun impact sur l'inscription en formation. Nos résultats suggèrent que l'impact détecté sur les appels est davantage dû à l'augmentation de l'importance accordée aux informations sur la formation plutôt qu'à la mise à jour des croyances des demandeurs d'emploi. Enfin, le troisième chapitre étudie également la demande pour la formation professionnelle, mais prend en compte les contraintes comportementales. Distinguant les croyances "externes" (sur le monde) et les croyances "internes" (sur soi-même), les résultats montrent que les demandeurs d'emploi subissent des contraintes financières les empêchant de rejoindre un programme de formation, et qu'ils sous-estiment la proportion de formations disponibles qui sont financées. Les obstacles internes liés à l'auto-efficacité, aux préférences inter-temporelles, à l'estime de soi et à la capacité d'organisation sont mentionnés à part égale par les demandeurs d'emploi indiquant avoir des obstacles internes à l'inscription en formation. À partir de ce diagnostic, la dernière partie est consacrée au design d'un essai randomisé contrôlé, avec des interventions reposant sur la transmission d'informations via des cours en ligne, et des sessions interactives par groupes de demandeurs d'emploi. Ces cours visent à cibler les croyances externes, internes, ou les deux simultanément.

# Chapter 0

## General introduction

### 0.1. Information and Job Search behavior

Over the past few decades, the unemployment rate has been high in most OECD countries. In France, while it was around 2% in the 1960s, it has gradually increased as a result of technological changes and economic crisis, and has finally exceeded 10% between 2013 and 2016 (see Figure 0.1).

Unemployment has several negative consequences on country's economies and societies: at the macroeconomic level, it represents a loss of production and income. At the microeconomic level, it worsens jobseekers' conditions, by sending a negative signal to recruiters, who are reluctant to hire them as their period of unemployment increases. This in turn discourages jobseekers and reduces their search efforts<sup>1</sup>. Moreover, unemployment is likely to isolate individuals, as work is an important driving force for social integration<sup>2</sup>. An additional issue government has to deal with is the level of the unemployment rate according to sub-groups of the population. Data from the National Institute of Statistics and Economic Studies (*INSEE*) show for instance that young people (under 25) or people without a school diploma are increasingly affected by unemployment (see Figure 0.2). Therefore, specific policies must be developed to meet the needs of each sub-population on the labour market. Reducing unemployment has been a priority objective for many French governments. Several attempts have been made, without any real progress being observed though. Strategies consisted in strengthening labour demand or improving the frequency and quality of matches. Over the last years, a lot of effort have also been invested in jobseekers training, to reduce unemployment. As unemployment is particularly high among people with a low level of education (see Figure 0.2 (b)), training programs can be a way to increase their skills and therefore their employability. However, the demand for vocational training remains low in France<sup>3</sup>. Looking at the characteristics of jobseekers who actually register into training programs, we see that the share of jobseekers without any diploma is quite low, whereas they should be the ones benefiting the most from it : they represent less than 20% of trainees if we consider any type of training, and only 15% if we consider training programs oriented towards a specific job and lasting at least 6 months (see Table 0.1).

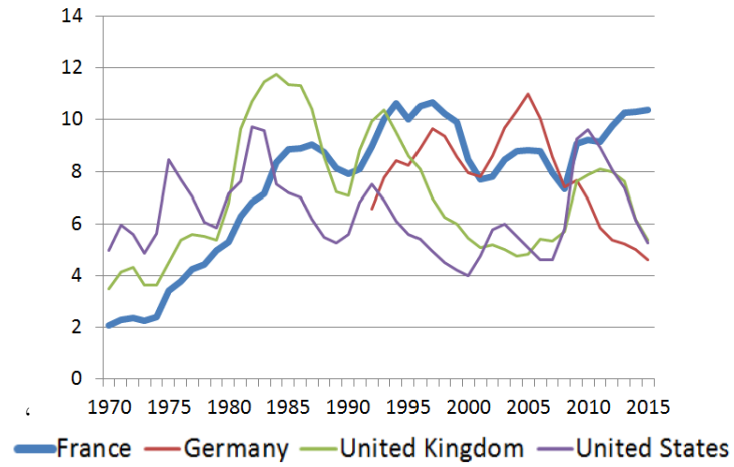
---

<sup>1</sup>See [Krueger et al. \(2011\)](#)

<sup>2</sup>See COE Report, 2011.

<sup>3</sup>Between 2010 and 2015, only 10% of unemployed people have enrolled into a training program each year (source : [Annexe au Projet de Loi de Finances \(2018\)](#) and DARES)

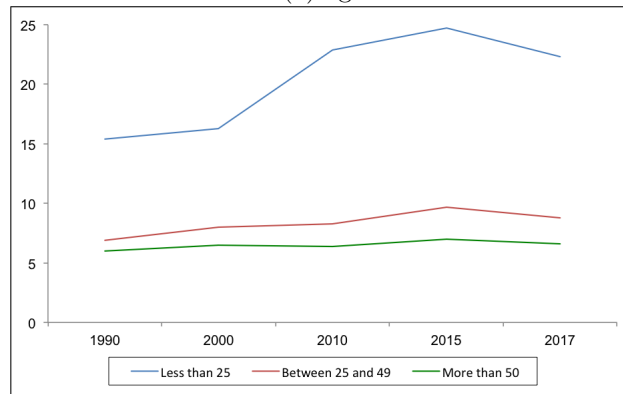
Figure 0.1: Unemployment rate



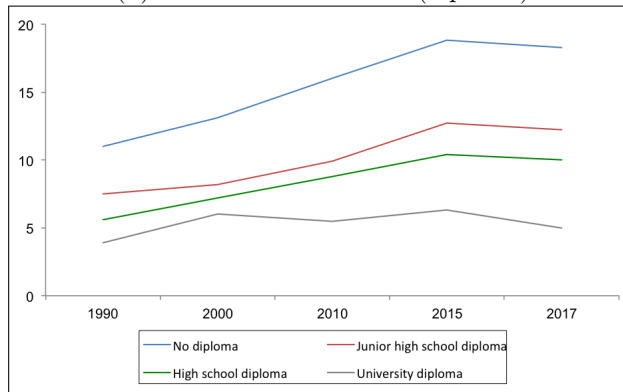
Source : OECD

Figure 0.2: Evolution of the French unemployment rate according to :

(a) age



(b) the level of education (diploma)



Source : the data come from the National Institute of Statistics and Economic Studies (*INSEE*)

Table 0.1: Distribution of training entrants according to the diploma

	Any training	Long lasting training
No diploma	17.40	13.92 %
Junior high school diploma or vocational degree	36.36	34.53 %
High school diploma	26.01	30.79 %
University diploma	20.53	20.76 %

*Notes:* This table presents the distribution of jobseekers enrolling into a training program according to the diploma. The data come from the French public employment agency (*Pôle emploi*). The sample is made of jobseekers registered at *Pôle emploi* who enrolled into a training program between 2015 and 2020 (1,914,148 jobseekers). Long lasting training programs correspond to training lasting at least 6 months which are directly oriented to a specific job. Between 2015 and 2020, it represented 40% of training enrollments.

Overall, due to the lack of rigorous evaluations, it is still difficult to know which policies are really effective in reducing unemployment, and how to learn from past experiences. Therefore, understanding the determinants of job search behaviour, including the demand for training programs, is crucial for designing suitable and effective unemployment policies.

The academic literature has been interested in the determinants of job search for many years. Going back to Stigler's work (Stigler (1962)), standard economic models emphasize the central role of information in job search behavior. These models show that information jobseekers have about the amount of unemployment benefits or the distribution of wages in the economy affects their reservation wage (Mortensen (1970)). However, knowledge about this distribution may be imperfect, and highly depends on jobseekers' needs and what they know about the labour market. Poor knowledge about these parameters can lead to a sub-optimal search strategy and lengthen their unemployment spell by leading them to under-invest in their job search.

With the development of new communication technologies, numerous countries, including France, provide to jobseekers numerous information about the labour market. Information comes from many public and private actors ranging from national employment centers, online job search platforms, training centers or media. However, this profusion of information has not fully resolved informational issues jobseekers face to. On the one hand, they still lack information about crucial aspects of job search (such as the returns of training programs for instance), as some information are hard to collect and highly depend on individual characteristics to be useful. On the other hand, it is now increasingly difficult to select the appropriate information and convert it into a consistent and concrete job search strategy. Several studies indicate that information *per se* is not helpful if not easily understandable and convertible into concrete job search strategies (Hanna et al. (2014), Babcock et al. (2012)).

Results from both the economic and psychological literature have shown that people often deviate from perfectly rational assumption, which creates bias in the decision making process (see for instance Babcock et al. (2012) or DellaVigna and Paserman (2005a), or Dellavigna (2009)). Job search implies important efforts in the short term for hypothetical returns in long-term; and in a context of lack of information and uncertainty, jobseekers are likely to procrastinate and to become discouraged. Meyer (2018) shows that these behavioral mechanisms can have visible consequences on the unemployment spell.

Finally, several papers point out the role of self confidence in the decision making process (see for instance Benabou and Tirole (2002)). On the labour market, this can translate into an under investment in the participation into programs aimed at improving labour market outcomes, such as vocational training.

Several empirical studies have recently shown that providing information to jobseekers about the labour market can mitigate these issues and improve labour market integration. For instance, in a random field experiment, Altmann et al. (2018) have designed a brochure indicating that when unemployed, the likelihood to find a job decreases over time, how useful broaden its scope of search in terms of sector of activity or geographical area is, and finally how long-term unemployment damage life quality. They find that the brochure increases employment and earnings in the year after the intervention by approximately 4% among jobseekers at high risk of unemployment. In another experiment conducted by Belot et al.

(2019), jobseekers were encouraged to search for vacancies on the internet on dedicated computers. Their intervention consisted in exposing jobseekers to a wider set of job offers in terms geographical or sectoral scope. They find that their intervention significantly increases the number of application as well as the number of job interviews.

Regarding the demand for training programs, little is known about how information impacts training enrollment. A notable exception is the paper of [Barr and Turner \(2017\)](#), who find that US unemployment insurance beneficiaries are four percentage points more likely to enroll in a community college program upon receiving a letter with information on the costs and returns of these programs. However, evidence on education investments suggests that information can change investment decisions. For instance, [Conlon \(2018\)](#) finds that delivering information to students about wages in their fields leads students to move to other fields over the following year if they overestimated these wages. [Bleemer and Zafar \(2018\)](#) find that providing information on returns to college education significantly increases reported intentions to enroll.

Overall, these studies tend to demonstrate that information interventions can modify job search behaviour, and that it can be successfully implemented with low costs using online devices. However, some questions remain open : in particular, we need further evidence on the impact of individualized information on labour market outcomes, or on the demand for training programs.

## 0.2. Public policy evaluation - Methodolgy

Very often, fighting against unemployment requires the design of active labour market policies. These policies aim at solving (or at least mitigating) specific issues labour market agents face (qualification issues, informational issues, etc.). However, evaluating the efficiency of these policies is equally important. Indeed, their implementation is based on public funds, and knowledge about their precise impact is crucial to keep learning from past experiences. Nevertheless, knowing how to proceed in order to assess their effects is not straightforward. For instance, one would like to compare individuals' labour market outcomes before and after they benefited from a given policy or program. However, other factors affecting labour market outcomes may have changed while the policy was implemented (for instance a large plant may have moved to the city at the same moment, which may affect return to employment as well), making hard to disentangle the effect of the policy and the effect of these other factors. Another solution could be simply to compare on the one hand individuals who benefited from the program, and on the other hand those who did not, in terms of the labour outcomes. However, this approach can lead to a selection bias, as individuals benefiting from a policy may be unrepresentative of the whole population, and statistically different from those who did not receive the program.

The ideal solution would be to simultaneously observe, for each individual, his labour market outcome with and without the policy we are evaluating. This relates to the notion of “potential outcome”, popularized by Rubin as part of his work on causal inference ([Rubin \(1974a\)](#)). In such settings, a “potential outcome” represents an outcome that we are not able to observe, because individuals cannot simultaneously benefit and not benefit from a policy: therefore, it either corresponds to the policy

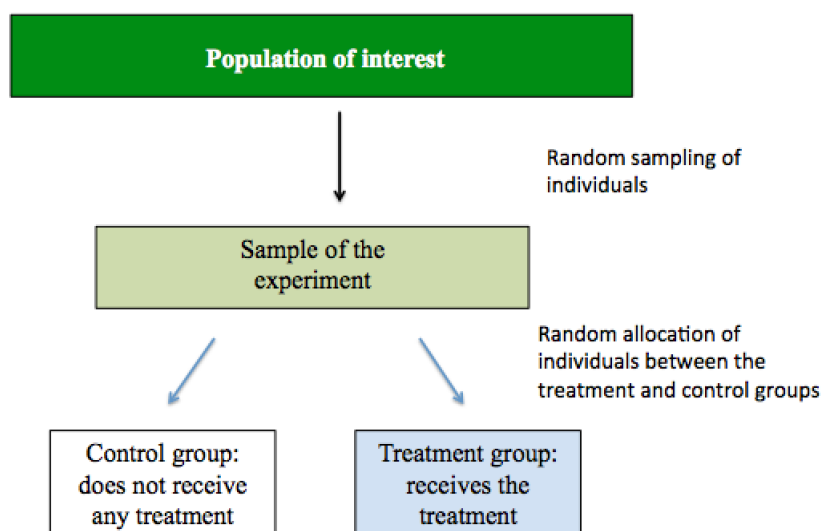
outcome when the individual does not benefit from it, or reversely to the no policy outcome when the individual actually benefit from it. Comparing the observed and the potential outcome would then enable to assess the exact impact of the program. However, and as just mentioned, observing these two outcomes, at the same time, is not possible for a given individual. Therefore, the method of randomized field experiments, on which I heavily rely in this dissertation, has been developed to solve this issue. This method, originally used in medicine, and then popularized in social sciences by economists such as Duflo or Kremer (see for instance [Duflo et al. \(2006\)](#) for an extensive report on this method), first consists in identifying the population of interest, *i.e.* the targeted population of the program. It represents the population to which the results of the experiment are intended to be extended. Then, among this population, a sample of individuals is randomly selected (see [Figure 0.3](#)). These individuals are then randomly allocated into two groups:

- a treatment group, which receives a “treatment”, *i.e.* a program, a public policy, etc.;
- a control group, which does not receive any treatment, but which is still followed over time, as the treatment group.

Comparing these two groups in terms of outcomes at the end of the experimentation finally enables to assess the causal impact of the program. Indeed, when the sample size is sufficiently large, the randomization between the two groups ensures that they are statistically identical (neutralization of the selection bias): any difference in outcome between the two groups is therefore due to the policy we evaluate.

Beyond the assessment of causal impacts, this method offers the possibility to test hypotheses based on theoretical models. Therefore, it is crucial to correctly select, before the experimentation, the outcome variables used to determine the efficiency of the policy being evaluated. A further challenge is also to ensure that the policy or program which is evaluated corresponds to the issues raised by the literature and faced by labour market agents.

Figure 0.3: The randomization process





Econometrically, the method of random field experiments classically relies on the estimation of two main effects:

- *the Intention To Treat effect*, which is the impact of the program on individuals from the treatment group (whether they actually enrolled in the program or not);
- *the Local Average Treatment Effect*, which is the impact of the treatment among compliers, i.e individuals assigned to the treatment group who also accepted to enrol into the program we evaluate.

According to the design of the experimentation, it is possible to include fixed effects in regressions, to control for the effect of specific factors (for instance geographical or time effect, as in chapter 1).

Moreover, several treatments can be evaluated at the same time (see for instance chapter 2): in this case, the number of groups corresponds to the number of treatments (with eventually a pure control group which does not receive any treatment): it is then possible to compare each treatment relative to the other, or to compare any treatment group with the control group.

In any case, the ability to detect an impact when there is one (*i.e* the statistical power of the experiment), increases with the sample size and with the proportion of compliers.

Once the program is evaluated, the goal is to scale it up if it has a positive impact on individuals' labour market outcomes. When this is not the case, the evaluation had avoided to spend a large amount of funds in a program which is not efficient for reducing unemployment. In addition, the analysis of heterogeneous effects according to individuals' characteristics is an opportunity to deepen our knowledge of job search behaviour and of the labour market; moreover, it then becomes possible to efficiently target the beneficiaries of a policy. Overall, and irrespective of the evaluation results, understanding why a given program or policy work or does not works helps the design of improved policies, the derivation of policy recommendations, and contributes to the scientific debate.

### 0.3. My dissertation

My thesis studies the impact of improved information on job search behaviour and on the demand for training programs. The strong collaboration with the French Public Employment Service during my research period gives me the opportunity to run several experiments with particularly rich datasets.

Chapter 1 presents the results of a joint work related to the evaluation of the website Bob Emploi, which aims at delivering news to jobseekers about the labour market. We relied on a large scale random field experiment with 212,277 jobseekers, between April and November 2017. We find no impact on jobseekers' search effort and search scope, whether occupational or geographical. We find modest effects on search methods: jobseekers using the website are more likely to rely on personal networks and to use resources provided by public employment services. However, we do not observe any effect on self-reported well-being and on employment. Our large dataset enable us to run machine learning methods in order to detect heterogeneous impacts across subgroups of the population. However, we find no heterogeneous effects on reemployment.

Chapter 2 presents a joint work focusing on the role of information about training on training enrollment. Partnering with *Pôle emploi*, we sent 50,000 emails advertising training opportunities to jobseekers in four regions of France in late summer 2016. We randomly added short messages on training costs, registration procedures, and training returns to the basic email template. We find that receiving an email with a message emphasizing training returns in terms of employment more than doubles the likelihood that jobseekers call back the training center. However, callback rates are low in absolute value (less than one percent) and we detect no impact on enrollment one to six months after the intervention. We provide suggestive evidence that the effects on callbacks are driven by increasing salience of basic information about training rather than by belief updating.

Finally, chapter 3 also focuses on the demand for vocational training, but incorporating behavioural constraints in the analysis. This joint work is based on a survey we administered to 20,000 jobseekers registered at *Pôle emploi* between March and July 2019. Distinguishing between “external” beliefs (about the world) and “internal” beliefs (about the self), we find that jobseekers experience financial constraints preventing them from joining a training program, and that they underestimate the proportion of subsidized programs available to them. We also find that half of jobseekers lack information about the registration process and training programs in general. Regarding internal barriers, we find that obstacles related either to self-efficacy, self control, self esteem and executive function are equally mentioned among jobseekers reporting internal barriers in training enrollment. In light of this diagnosis, we then propose a design for a randomized controlled trial, with interventions relying on the delivery of information through app-base courses, and interactive sessions involving groups of jobseekers. These courses aim at targeting either external or internal beliefs, or both of them simultaneously. This will enable us to estimate their impact on training enrollment or self reported barriers in absolute terms or relative to each other.

## Chapter 1

# Can a Website Bring Unemployment Down? Effects of a French Online Platform on Job Search Efficiency

*This chapter is joint with Aïcha Ben Dhia, Bruno Crépon, Louise Paul-Delvaux, Bertille Picard and Vincent Pons*

This paper evaluates the impact of an online platform helping job seekers adopt effective job search strategies, using a large-scale randomized experiment. The platform combines labor market data from the French public employment agency and personal data from individual profiles to recommend users occupations and areas with high employment chances and to give them concrete tips to improve their job search methods. The experiment was conducted in collaboration with the French public employment agency on a sample of 212 277 job seekers from April to November 2017. An encouragement design led to a take-up rate of 26.2 % in the treatment group and virtually zero in the control group. Following individual trajectories over 18 months after the intervention, we do not observe any impact on job seekers' search effort and search scope, whether occupational or geographical. We find modest effects on search methods: job seekers using the website are more likely to rely on personal networks and to use resources provided by public employment services. However, we do not find any effect on self-reported well-being and on employment outcomes, both in the short run or in the middle run, indicating that more intensive interventions are required to bring unemployment down.

## 1.1. Introduction

Helping unemployed workers to search for jobs more effectively is central to the agenda of employment agencies. Job seekers face considerable challenges in gathering information on occupations and locations where they are most likely to find employment and in identifying useful search methods and tools.<sup>1</sup> These informational barriers can reinforce behavioral biases and dissatisfaction, generating a vicious circle in which bad search outcomes lower one’s self-esteem, motivation, and search effort (Babcock et al. (2012)). By and large, evaluations of public programs providing job search assistance have not been conclusive on their impact (Card et al. (2017)). Over the last two decades, numerous private websites have been launched to complement the public sector effort by taking advantage of modern digital technologies. These websites primarily work as job boards, bringing together job seekers and job openings. They also frequently contain modules with pieces of advice to help job seekers improve their job search strategy. Even though these websites are increasingly used by job seekers, there is little evidence on their effectiveness. The appeal of online assistance services is that they can be easily implemented at scale, updated and customized to meet users’ needs (Kuhn and Skuterud (2004a), Kuhn and Mansour (2014a), Autor (2009), Horton (2017)). On the other hand, digital tools also come with limitations. Certain sub-groups of the population, especially among the unemployed, are not familiar with digital technologies and may find it difficult to use these websites. Moreover, the content of the recommendations provided by websites may not be sufficiently tailored to each user.

In this paper, we study the impact of “Bob Emploi,” an online platform dedicated to help job seekers in France overcome informational barriers and identify effective search strategies. Bob Emploi was launched in November 2016 by Bayes Impact, a private non-profit organization working with tech developers to analyze rich data collected by public administrations with state-of-the-art methods. To design Bob emploi, the tech NGO benefited from a collaboration with the French public service of employment, Pôle emploi. This innovative partnership sets Bob Emploi apart from other private sites. At its creation, Bob Emploi aroused much interest and enthusiasm and it benefited from continued support by policymakers as well as important financial aid. As the French minister of Labor of that time put it, it represented a promising opportunity for “innovators in the tech industry to contribute to social inclusiveness.” Bayes Impact C.E.O. set high expectations by announcing that the website and its algorithm had the potential to lower the unemployment rate by 10%.

When job seekers first log into the Bob Emploi website, they are asked to provide socio-demographic information and to indicate their target job and the status of their current job search (e.g., the number of applications sent and the number of interviews obtained in the past weeks). Bob Emploi then compares user profiles with job market data coming from two main sources. The first dataset contains information about a representative sample of job seekers registered at Pôle emploi on socio-demographic variables,

---

<sup>1</sup>On the difficulty of identifying occupations and geographical areas with the highest chances of reemployment, see e.g. Şahin et al. (2014), Patterson et al. (2016), Papageorgiou (2014), and Herz and van Rens (2019) on occupational mismatch; Altmann et al. (2018) and Gee (2019) on perceived market tightness; Marinescu and Rathelot (2018) on geographical mismatch; Beaman and Magruder (2012), Belot et al. (2019), and Abel et al. (2019) on search methods; and DellaVigna and Paserman (2005a), Caliendo et al. (2015), Spinnewijn (2015), and DellaVigna et al. (2020) on sub-optimal search effort.

career objectives, search methods, and employment outcomes. This dataset includes variables such as age, gender, and highest educational degree, desired job and geographic search zone, and dates and duration of unemployment spells. Second, Bob Emploi uses data on job postings on the Pôle emploi job board and results of an annual survey on employers' expected labor demand. Based on this data, the website provides users with an employment assessment which evaluates how likely their job strategy is to be successful and guides them towards economic sectors and geographic areas that may increase their chances to find a job. In addition, Bob Emploi groups together recommendations of search methods and tools from several job search and recruitment professionals. The website builds on this advice to propose a list of steps that can help users overcome the obstacles and bottlenecks they face in their job search.

We designed a randomized experiment to investigate whether using Bob Emploi modifies job seekers' search strategies and, in turn, accelerates re-employment. The impact evaluation is based on a large-scale encouragement design implemented in 254 Pôle emploi local agencies. Our sample consists of 212 277 job seekers whose registration at Pôle emploi had begun less than a year before the intervention. Randomization was conducted at the individual level, and about half of the sample was assigned to a treatment group. Individuals in this group were invited to attend an information session at their local agency. During this session, Pôle emploi caseworkers introduced them to Bob Emploi.<sup>2</sup> In addition, all job seekers in the treatment group received three advertising emails encouraging them to create an account on Bob Emploi.

We follow the trajectories of individuals in our sample over 18 months after the intervention. We rely on administrative records from Pôle emploi to measure the rate of reemployment, socio-demographic variables, interactions between job seekers, and the use of Pôle emploi's support services (e.g. meeting with a caseworker, attending a job training program or a group resume editing session). The Pôle emploi data also record job seekers' online applications. In addition, we sent an online survey 6 months after the end of the intervention. In this survey, job seekers were asked about their search scope in terms of occupation and geographic area, the search websites they used, and the time spent on job search every week. The survey also contained questions on job seekers' self-assessed well-being, whether they felt support or loneliness in their search, and activities outside of the job search to evaluate their life balance. The response rate to the main questions in the survey was slightly above 15%, which is fairly high for such online surveys among job seekers.<sup>3</sup> We measure a small but statistically significant differential attrition between control and treatment groups.

Our (binary) measure of take-up of the treatment is whether the job seeker created a Bob Emploi account and/or attended an information meeting conducted by Pôle emploi. Take-up reached 26.2 % in the treatment group, and remained close to zero in the control group.<sup>4</sup> In itself, participation in the information meetings gave job seekers an opportunity to find out about other resources at Pôle

---

<sup>2</sup>The meetings occurred at a rate of approximately one per week in each agency between late April and July 2017. Each session lasted approximately an hour and a half and was attended by 13 job seekers on average.

<sup>3</sup>Out of an eligible sample of 86,673 job seekers, DellaVigna et al. (2020) obtain a response rate of less than 8%. In comparable online surveys sent by Pôle emploi, the response rate usually hovers around 10%.

<sup>4</sup>This percentage can be broken down into three types of participation: in the treatment group, 6.6 % of job seekers only created a Bob Emploi account, 14.07 % only attended a meeting, while the remaining 6.53 % did both. As a robustness check, we compare this measure with an alternative definition that also includes survey respondents who reported using Bob Emploi, and we find a correlation between the two take-up measures of 0.78.

emploi and to interact with caseworkers. However, we find that the differential rate of usage between the treatment and the control groups was more than 25 points higher for Bob Emploi than for any other resources presented in the meeting, which indicates that the main effects we measure are driven by Bob Emploi.

Overall, Bob Emploi had limited effects on job seekers' search strategy and reemployment outcomes. We define a modification of a job search strategy as a change in target (i.e. of the desired job or geographic search zone), in the search methods used (e.g. type of contacts solicited in one's social network, internet sites used, etc.), or in the search effort, as measured by the time spent searching and the number of applications sent. First, we find no effect on search effort, as measured by the time spent searching and the number of applications sent. Analyzing the characteristics of the applications sent, we also see no change of job seekers' search scope, whether in terms of occupations or geographic locations.

Secondly, we detect treatment effects along some of the variables measuring the methods job seekers use to search for jobs. Survey respondents in the treatment group report being 1.3 percentage points more likely to rely on their personal network, which is significant at the 5% level. They do not behave differently than the control group regarding other types of networking. Treated job seekers are not more likely to use best practices for the job search frequently recommended by the website (e.g. adapting one's resume to the job listing before applying) either, but Bob Emploi does increase the number of websites they use to look for jobs by 3%. Taking into account the innovative collaboration with Pôle emploi, we pay special attention to the website's impact on job seekers' use of public job search assistance services. The positive effect on the number of websites used is entirely driven by the increased usage of websites that are either created by or closely connected to Pôle emploi. Moreover individuals in the treatment group are 2.4 percentage points (4%) more likely to engage in conversations with their caseworker within the 6 months that followed the intervention, which is significant at the 1% level.<sup>5</sup> The website did not lead to a higher rate of participation to other search assistance programs, such as group workshops or training programs. Consistent with the close partnership between Bob Emploi and Pôle emploi, these results suggest that Bob Emploi acts as a complement to rather than a substitute of public job search services.

Third, to provide a more complete picture of the potential mechanisms by which Bob Emploi may increase the effectiveness of the job search, we turn to survey variables related to job seekers' well-being and life balance. Unemployment spells can be stressful periods for job seekers, especially when they feel isolated, poorly advised, and poorly equipped (Krueger and Mueller (2012)). The failure to obtain interviews and finding a job can generate important psychological costs (McKee-Ryan et al. (2005), Ridley et al. (2020)), drain energy out of social activities, and decrease future search efforts. In spite of Bob Emploi's user-friendly design, uplifting messages, and recommendations to participate in activities outside of the job search, survey respondents in the treatment group do not report increased well-being than in the control group, nor do they feel more supported in their job search. They are also not more likely to engage in sport, cultural, or volunteering activities.

Fourth, and most importantly, we investigate whether the modest effects on upstream variables related

---

<sup>5</sup>An important caveat is that these two effects may be partly driven by the participation in the information meetings.

to the job search translated into effects on employment. As we observe all consecutive exits and entries out of and in unemployment, we can detect whether individuals experienced some employment episodes following the intervention and compute the total number of days spent in unemployment. For any exit out of unemployment, we know whether the individual finds a job and the type of contract she obtains. We do not measure any impact of Bob Emploi on any employment outcome and at any time horizon. Bob Emploi did not increase the likelihood of experiencing some employment episode, irrespective of the form of employment we look at, and did not reduce the duration in unemployment. To rule out the hypothesis that zero mean effects hide simultaneously positive and negative effects depending on individuals characteristics, we perform a heterogeneity analysis. We first consider ITT by subgroups based on age and education level. We then explore more systematically heterogeneity with a larger set of characteristics using the generic machine learning approach (Chernozhukov et al. 2018). In both cases, we do not find heterogeneous impacts on unemployment duration. Lastly, to shed some light on the cost-benefit analysis of this type of assistance programs, we plot the mean cumulative amount of unemployment benefits received by job seekers and observe no significant difference between the two groups. Considering the upper bound of the 95 percent confidence interval, we can reject any effect higher than 0.5 percentage points on experiencing some employment episode within the 18 months following the intervention. To shed light on the cost-benefit analysis of this type of assistance program, we also consider cumulative unemployment benefits received by job seekers and observe no significant difference between the control group and the treatment group.

Unfortunately, the results of the experiment reported in this paper show that Bob Emploi did not fulfill its promise of a breakthrough in job search assistance, in spite of its close partnership with the public employment services and access to the latter’s rich data. Beyond the performances of this particular website, this raises questions about the potential of online labor market intermediaries that aim to orient job seekers by applying modern data analysis techniques.

**Related literature.** A large body of economic work on job search motivates this study. Babcock et al. (2012) insist on the potential for and the existence of information gaps as a source of occupational and geographic mismatch. Biases in perceptions of job arrival rates can also lead individuals to make suboptimal decisions in terms of search intensity (see e.g. DellaVigna and Paserman (2005a), Spinnewijn (2015), Caliendo et al. (2015), McGee (2015), Krueger et al. (2011)). Although the advent of digital technologies in the late nineties gave job seekers access to many online resources with information and advice on labor markets, Kuhn and Mansour (2014a) do not observe significant returns from using the Internet to search for jobs. In fact, finding what will be most relevant and choosing the most appropriate method of search certainly remains a challenge. Even when accurate information is available, Babcock et al. (2012) also mention the difficulty in moving from information to action, particularly during job search periods associated with self-deprecation and demotivation (Falk et al. (2006b), Krueger and Mueller (2012)). In this context, finding useful tools and methods of search may be highly valuable and can help job seekers avoid vicious circles in which failures fuel demotivation which then lowers effort and reemployment chances.

Yet, in their meta-analysis, [Card et al. \(2017\)](#) find that most of public programs providing job search assistance have shown only limited impact, and where some effect is demonstrated, it is not long-lasting. As pointed out by [Belot et al. \(2019\)](#), such programs generally bundle many different forms of assistance, often coupling monitoring and sanctions (e.g. [Blundell et al. \(2004\)](#)). Hence, available empirical evidence is often not sufficient to identify the main bottlenecks in the job search and the precise mechanisms by which interventions could help the beneficiaries.

Our paper adds to a recent strand of the literature that tests more narrowly-designed interventions aiming to solve specific challenges faced during the job search. Addressing the question of the link between misinformation and suboptimal search effort, [Altmann et al. \(2018\)](#) studies the impact of sending a standardized brochure with generic information on job search returns and employment. [Belot et al. \(2019\)](#) follows a sample of job seekers who can use a new search algorithm with recommendations of jobs to expand one’s occupational scope of search. Addressing the challenge of moving from information to action, [Abel et al. \(2019\)](#) help job seekers stick to their intentions by writing down and committing to a search plan. The research question we address in this paper is the impact of information provision (as in [Belot et al. \(2019\)](#) and [Altmann et al. \(2018\)](#)) when combined with actionable advice on search methods (as in [Abel et al. \(2019\)](#)). To detect the impact of Bob Emploi platform, we use a particularly rich set of variables that allows us to document multiple aspects of job search, covering search effort, employment outcomes and detailed information on the search methods used by job seekers. This variety is unusual in a randomized experiment and more common in descriptive studies such as [Krueger et al. \(2011\)](#).

In the three aforementioned randomized experiments, the involvement of the researchers in the design of the program and the intervention itself allows them to reach a high level of precision in the program content, sometimes departing from the real conditions of job search and partly losing external validity. Besides, in most existing studies on private labor market intermediaries, the goal is to compare the provision of counseling services by traditional government agencies and by new entrants from the private sector (e.g. [Behaghel et al. \(2014\)](#), [Krug and Stephan \(2013\)](#)). Instead, this paper studies a platform based on the cooperation between the public services and a private labor market intermediary. As Bob Emploi was developed by Bayes Impact, a non-profit tech start-up, and leveraged data and logistical capacity of Pôle emploi, we explore the potential complementarities between both types of actors.

To the best of our knowledge, this is also the first randomized study of a private website that is entirely dedicated to job search assistance that does not couple it with matching functionalities of job boards. This connects our paper to studies on the potential of online platforms and to the related discussion on profiling in active labor market policy (see, e.g., [Berger et al. \(2001\)](#) for a comprehensive review). In the last two decades, the rise of the Internet offered a promising opportunity to lower labor market search costs. However as pointed out by [Horton \(2017\)](#), this first wave turned out to have only limited impact ([Kuhn and Skuterud \(2004a\)](#)). [Horton \(2017\)](#) suggests that the reason is that providing searchable listings of jobs or applicants does not create enough value-add to significantly improve the functioning of the labor markets. In this paper, as in [Horton \(2017\)](#) and [Belot et al. \(2019\)](#), we test the potential of a more sophisticated website that uses the power of digital technology not only to gather



and make information available for free but also to process and analyze this information, and to deliver customized advice at scale. Therefore we contribute to the literature on how algorithms and digital technologies can solve market frictions ([Resnick and Varian \(1997\)](#); [Adomavicius and Tuzhilin \(2005\)](#); [Varian \(2010\)](#), [Dinerstein et al. \(2018\)](#)).

## 1.2. Setting

### 1.2.1. Online resources for job seekers in France

France’s unemployment system is characterized by a high degree of centralization. The Public Employment Service, Pôle emploi, is not only in charge of distributing unemployment benefits and monitoring job search, but also of providing job search assistance to job seekers and of helping them match with employers. With the rise of the internet in the end of the nineties, new entrants emerged as alternative labor market intermediaries. The first generation of these private websites were often basic job boards such as CareerBuilder or Monster ([Kuhn and Skuterud \(2004a\)](#)). A second wave leveraged the power of social networks: famous international examples include LinkedIn and Viadeo. These online websites may be only at a national scale (e.g. RegionsJob in France) or sector-specific (e.g. dice.com in the United States for tech sector or emploi-environnement.com in France for jobs in environmental fields). At the same time, Pôle emploi started to digitize its traditional services, creating its own platform, called Emploi Store, to gather new online services along with recommendations to private websites that help job seekers in their search.

In fact, private websites frequently augment their basic services with additional resources offered to job seekers. Most often, these resources are free articles presenting what the website considers best practices to search for a job.<sup>6</sup> These extra resources sometimes also include paid services (e.g. getting help from a coach). The advent of artificial intelligence and machine learning over the last decade opened up new opportunities to address job seekers’ and recruiters’ needs. Leveraging large datasets, these techniques work by selecting and analyzing relevant information, allowing for better and customized recommendations. Among labor market intermediaries, many private actors attempt to improve the matching between employers and applicants (e.g. training algorithms on past hiring and résumés to help employers sort applications). However, fewer solutions have been designed to optimize the job search from the perspective of the job seeker. In fact, the business models of these websites are often based on monetizing employers’ postings. As a consequence, most advanced services are oriented towards employers (e.g. boosting job listings to a pool of pre-identified profiles). Moreover, as in other fields of application, big data algorithms are only as good as their number of users and the quality of their data, and it is a significant challenge for all websites to obtain data about job search techniques.

From that perspective, Bayes Impact occupies a unique position: its website is dedicated to improving the job search (and not the matching) and can leverage datasets on job seekers’ trajectories thanks to its partnership with Pôle emploi. At the time of its creation and to the best of our knowledge, no other

---

<sup>6</sup>For example, RegionsJob in France makes recommendations in dedicated posts such as “how to write a good résumé?” or “8 recommendations to perform during an interview”.

private website offers job search assistance with similar technical capacity.

### 1.2.2. Description of Bayes Impact organization

Bayes Impact is a non-profit organization created in 2014 by a small group of developers with the goal of leveraging advanced data analysis techniques to help resolve important public policy challenges. The organization is based in Paris and San Francisco and recruits primarily within the world of tech start-ups. Its approach consists of building partnerships with public institutions in order to access relevant data to find solutions to a public issue. As an example, the organization launched in 2017 a project to document the management of the police’s use of force in partnership with the police services in California. Results of the analyses of these data with artificial intelligence techniques are then shared with citizens via user-friendly, simple interfaces. As stated on its website, Bayes Impact aims to “modernize public services in a transparent manner”.<sup>7</sup> Bayes Impact is financed almost exclusively through public funding. In keeping with this positioning, the organization has taken on a culture of transparency and sharing, notably by keeping the developed codes open-sourced (even though the data generally are not).

### 1.2.3. Description of Bob Emploi website

In France, Bayes Impact formed a partnership with Pôle emploi in 2016 to obtain data on the trajectories of job seekers and the hiring conditions on the job market. Bayes Impact aimed to apply the techniques of big data analysis in order to identify, for each profile of job seekers, the best strategy for rapid re-employment. In November 2016, the organization launched the online platform, Bob Emploi, cross-referencing Pôle emploi data with user profiles to propose an active strategy for job search and to identify priority actions to augment their chances of employment.

More precisely, Bob Emploi cross-referenced two sources of data coming from Pôle emploi and from the users themselves. At the moment of initial login, a user creates her profile with information about standard socio-demographic variables (gender, age, highest educational degree, municipality of residency). The user indicates her desired job, as well as her target geographic zone and salary. Finally, Bob Emploi questions the user about the metrics of her job search over the course of the past weeks: the number of job opportunities identified each week, the number of applications sent, the number of interviews and job offers obtained. The ratio of one step to the next is then used by the site to orient recommendations on specific steps that are proving to be the greatest obstacles in the job search.

Bob Emploi then cross-references each user profile with data on the job market obtained via Pôle emploi. The algorithm uses principally three databases from Pôle emploi:<sup>8</sup>

- a representative sample from a database called Fichier historique, which follows, over the last 10 years, the trajectories of persons enrolled in Pôle emploi throughout their unemployment periods. This base contains the basic socio-demographic information of the job seekers (gender, age, highest educational degree, municipality of residency), as well as the principal job desired (target job) at the

---

<sup>7</sup>See Bayes Impact website <https://www.bayes.org/fr/focus/spc>

<sup>8</sup>Details about the databases used are published by Bayes impact in the github [https://github.com/bayesimpact/bob-emploi/blob/master/data\\_analysis/data/README.md#bmo-bmo-besoin-en-main-doeuvre](https://github.com/bayesimpact/bob-emploi/blob/master/data_analysis/data/README.md#bmo-bmo-besoin-en-main-doeuvre)

moment of enrollment (see details in section 3.2.2) and all the accompanying activities carried out via Pôle emploi in which the job seeker participated (e.g. meetings with caseworkers, job training program). The principal limitations of this base are that it contains little information about periods of employment following unemployment (occupational title of the job, salary obtained, location) and that it does not contain information on the methods used during the job search.

- a database of job offers placed on the Pôle emploi job board by employers. Each observation in the database contains a job title and description as well as the number of applications to the job posted on the job board at the moment of data collection. The compilation of this information allows for the construction of a proxy of market tightness for each career in a given locale. This also permits, by semantic analysis, to retrieve the requirements of job listings.
- data from the annual Pôle emploi survey of a sample of 500,000 employers, called Besoins en Main d’Oeuvre, to know their hiring needs in the six months following the study, and the skills they deem particularly difficult to recruit. This survey allows Pôle emploi to measure market’s labor demand and orient job seekers towards high-demand sectors.

Bob Emploi relies on these assessments to suggest modifications of user’s search targets, in exploring related careers in demand or in widening the geographic zone of her search. Though Bob Emploi’s initial aim was to implement machine learning techniques over the collected parameters affecting job search, the data available limit the predictive power of the algorithm, notably because of the lack of information about reemployment trajectories in the *Fichier historique* (occupation obtained, type of contract) and the absence of precise data on the applications sent and the methods used during the job search. For lack of this information, Bob Emploi relies on recommendations gathered over the course of a series of qualitative interviews with multiple professional recruiters (Pôle emploi caseworkers, recruiters, career coaches) who wish to support the project.

In a broad outline, the recommendation algorithm is based on the following modeling of standard recruitment paths as a funnel. The funnel corresponds to the different stages of the trajectory of a job seeker who manages to find a job : she first defines her search perimeters (occupational and geographical scope of search), identifies offers, applies, and attends an interview to secure, in the end, an offer of employment. The algorithm is based on ratios of reference, arbitrarily defined by the site designers, in relation to which are compared the performance metrics stated by a user during his first login. For example, if the number of applications reported by a user is superior to the algorithm’s reference number for the number of sent applications, but the number of interviews is inferior to the algorithm’s reference number, Bob Emploi concentrates its recommendations on the improvement of the job seeker’s application file, e.g. advising the user to contact professional career coaches to help her prepare for interviews. In the long run, the platform aims to enrich its data through user feedback.

The user landing on the Bob Emploi website begins by creating her account and giving information about her profile and job search activities up until that moment. In return, the site sends him a “diagnostic” of employability in the form of a numerical score deconstructed along different axes (labor demand for the target job, competition, match between the user’s profile and the job requirements). The ensuing

suggestions of key actions are presented in a hierarchical manner. Actions can be to widen one’s search perimeters if the user reported not being able to find many job offers but the data indicate the market is less tight in surrounding municipalities, or to edit one’s resumé if the ratio between the number of applications sent and the number of invitations to interview is particularly low. Each recommendation is matched with concrete advice and signals appropriate external resources, most often online.

Finally, the tone and format used by the interface are deliberately inviting and simple, intentionally distinct from the often austere and complex appearances of administrative sites. A significant portion of the site’s design was created by professional web designers to find ways to render quantified information easily digestible.

## 1.3. Data

Our analysis draws on three data sources: the administrative data collected by Pôle emploi, the user data on connections to the Bob Emploi site shared by Bayes Impact and outcomes collected through online surveys, administered to both control and treatment groups six months after the intervention.

### 1.3.1. Administrative data from Pôle emploi

The administrative data from Pôle emploi cover information about job seekers from their registration at Pôle emploi until the end of their unemployment period.

#### 1.3.1.1. Unemployment records

The Pôle emploi unemployment records, called Fichier historique, contain several standard socio-demographic characteristics including gender, age, municipality of residency and education level along with their unemployment duration at the time the intervention is implemented, and afterwards. They report the individual monthly unemployment benefits, which allows us to compute the cumulative amounts job seekers received over the 18 months following the intervention.

These records also contain information on the job search strategy that the job seeker adopts at the beginning of her unemployment spell, including labor sector, skill level of the desired occupation (e.g. manager), target wage.<sup>9</sup> As these variables are entered only once, at the beginning of the unemployment spell, they cannot be used to detect changes in the job search strategy following the intervention.

#### 1.3.1.2. Employment outcomes

Our main employment outcome captures whether job seekers get reemployed after the intervention. To identify employment episodes, we rely on the ICT01 employment indicator of Pôle emploi.<sup>10</sup> As described in section 1.4.1 and shown in figure 1.1, we drew our sample over three consecutive months in order to

---

<sup>9</sup>Labor sector is recorded as a 5 digit occupation code based on Pôle emploi’s lexicographical classification of occupations (called Répertoire ROME) that is an equivalent of the US O\*NET classification. A 5-digit occupation corresponds to a narrow type of job, such as nurse-anesthetist.

<sup>10</sup>This indicator mainly draws on: (i) employer declarations at the time of hiring and (ii) job seekers’ self-reported declaration of joining the labor force. This second category is necessary to be able to track job seekers who become self-employed.

ensure that the lists of job seekers who would be invited to information meetings were as up-to-date as possible.<sup>11</sup> We compute all our employment outcomes starting from the sampling date of each wave – April 1st for the first wave, May 1st for the second wave and June 1st for the third one. As the intervention ended two months after the sampling date, the employment outcome one month after the end of the intervention for an individual in the first wave is measured by looking at whether she had any employment episode between April 1st and July 1st.

To better characterize these episodes, in terms of duration and stability, we look at two complementary outcomes. First, we look at employment episodes in long-term jobs, that is, whether individuals obtained any work contract of more than 6 months (CDI or CDD longer than 6 months) over the period of interest.<sup>12</sup> Second, we create a variable that counts the cumulative number of days in unemployment since the sampling date.

#### **1.3.1.3. Use of Pôle emploi assistance services**

Having access to both Pôle emploi and Bayes Impact data, we can investigate potential substitution effect between the use of Bob Emploi and Pôle emploi services. Pôle emploi database records all jobsearch actions completed by the job seekers: one-on-one meetings with their caseworkers, participation in workshops at Pôle emploi (e.g. resume writing classes), individual assistance program (e.g. skill assessment) or any information meeting such as the ones that were organized within this experiment. From these records, we create two variables: (i) the number of meetings with caseworker and (ii) the total number of workshops, programs and meetings that the job seeker attends following the intervention.

#### **1.3.1.4. Online applications from Pôle emploi job board**

We use data about online applications on the job board administrated by Pôle emploi. This job board website is one of the most popular search platforms in France. According to [Skandalis \(2019\)](#), it covers about one third of job creation and “vacancies posted on the website are mostly for low-skilled positions but also tend to offer more permanent contracts and relatively better-paid jobs than the average job created in France in the period.”

Since 2014, job seekers can apply online to these job listings, which gives us access to micro-level data on job applications. From this, we can recover the total number of applications and break down into different sub-categories depending on whether the application was initiated by the job seeker herself, by his caseworker or by the employer.

The data also include some of the job listing features such as the job title and its geographical location. Based on this information, we create:

1. a dummy variable capturing the geographical scope of the job search. It is equal to 1 if the job

---

<sup>11</sup>Individuals from the first wave were drawn and added to the sample on April 1st and were invited to attend meetings happening between late April and late May. Individuals from the second wave were sampled on May 1st for meetings between late May and late June while individuals in the third wave were sampled on June 1st and attended meetings between late June and late July.

<sup>12</sup>Employment episodes fall into one of the following categories: long-term contract (called Contrat de Durée Indéterminée, or CDI), short-term contract (called Contrat de Durée Déterminée, or CDD) for more than 6 months, short-term contract of 1 to 6 months, and part-time work.

seeker applies to vacancies outside her own municipality and 0 otherwise.

2. a dummy variable capturing the occupational scope of the job search scope. It is based on Pôle emploi’s lexicographical nomenclature of occupations that is equal to 0 if the individual only applies to jobs within a single 5-digit category and 1 otherwise.

### 1.3.2. Data from Bayes Impact

We use data from Bayes Impact to see who created an account on the Bob Emploi website among the job seekers in our sample.<sup>13</sup>

### 1.3.3. Survey data

To collect additional information on job search behavior, we administered an online survey to the entire sample, 6 months after the end of the intervention.<sup>14</sup> This survey allows us to shed some light on:

- i) job search effort, as measured by the number of hours spent searching over a normal week, along with the frequency at which job seekers send unsolicited applications.
- ii) job search methods, including:
  - relying on social networks: we ask whether job seekers rely on their personal network and on other types of contacts.<sup>15</sup>
  - search websites used. We compute (i) the number of websites created by Pôle emploi that the job seeker used, and (2) the number of other private websites used.<sup>16</sup>
  - geographical scope of job search, that is whether the respondent looks for jobs in his municipality, department or region of residency.
  - job search best practices, that are frequently recommended by Bob Emploi to increase search efficiency.<sup>17</sup>
- iii) well-being: we asked respondents had to choose numbers on a Cantrill scale from 0 to 10 to show their general well-being level, their motivation, and the extent to which they felt supported in their job search. To know about their life balance, respondents were asked if they participated in any sport, art or community activity at least once a month.
- iv) employment expectations: we also asked job seekers when they expected to find a job and create a dummy variable for whether individuals expect to find a job within the following 3 months.

---

<sup>13</sup>For privacy reasons, Bayes Impact could not directly share its list of users’ email addresses, nor could we share email addresses of individuals in our sample from Pôle emploi database. Therefore we used a anonymized matching algorithm to match email addresses from both sources.

<sup>14</sup>The complete questionnaire is available in appendix 1.A.

<sup>15</sup>We asked about 7 different social groups, such as friends, former colleagues or neighbors. Personal networks included friends, friends of friends, former classmates and colleagues. Other contacts included Pôle emploi caseworkers, volunteers from non-governmental organizations, shopkeepers in one’s neighborhood and other job seekers.

<sup>16</sup>We provided a list of 9 websites, including Bob Emploi, for which respondents had to say whether they knew of or used them, and whether they found them useful.

<sup>17</sup>Bob Emploi often suggests to adapt one’s CV or cover letter by reusing words or expressions of the job listing, or to call back to check in with an employer after an interview.

## 1.4. Design

### 1.4.1. Experimental procedure

#### 1.4.1.1. Encouragement design

Our experimental protocol follows an encouragement design. First, we invited the treatment group to information sessions at their local agency. These information meetings gave an introduction to Bob Emploi.<sup>18</sup> Each introduction to Bob Emploi session was hosted by one or two Pôle emploi caseworkers and lasted approximately 1 hour and 30 minutes. It began with an introduction to the challenges associated to job search and the usefulness of widening one’s search methods, notably through personal and professional networking. The caseworkers then presented the Emploi Store, the Pôle emploi platform that brings together digital job search services, which references Bob Emploi. The session concluded with a presentation of the Bob Emploi website using screenshots and, some time, a real-time simulation. These information meetings occurred at a rate of approximately one meeting per week in each participating branch, between April 20 and July 31, 2017.

We reinforced the encouragement design by sending advertising emails to the entire treatment group encouraging job seekers to visit the Bob Emploi site.<sup>19</sup> Three such emails were sent out: the first in the month of July, the second after summer vacation, on September 28th, and the last on November 13th.

#### 1.4.1.2. Sampling and timeline of the intervention

**Selection of agencies** 254 Pôle emploi local agencies were selected to participate in the experiment. The goal of this selection was to use only universal branches with the capacity to implement weekly information sessions throughout the 3 months of the experiment, and which were also as representative as possible of the totality of agencies in terms of size and geographic location.

**Eligibility criteria** The eligibility criteria were defined by a working group of Pôle emploi caseworkers involved in the design and follow-through of the experiment. The objective was to target people deemed the most likely to benefit from Bob Emploi and to be interested in the site. Eligible job seekers registered in participating branches at the moment of the drawings met all of the following criteria:

- has been unemployed for a year or less at the moment of the sampling, and in this respect likely to have not yet explored the entirety of search methods offered by Bob Emploi;<sup>20</sup>

---

<sup>18</sup>Such information sessions are frequently organized by Pôle emploi caseworkers to introduce free tools and programs to job seekers.

<sup>19</sup>In fact, when nearly two-thirds of the information sessions had occurred, a comparison between the list of invited participants and the creation of Bob Emploi accounts had indicated that only a small percentage of participants actually created an account on the website. The reminder emails differed slightly depending on whether the person had attended an information session or not.

<sup>20</sup>Duration of unemployment was defined as the duration of unemployment at the moment of the experiment, meaning from their latest registration at Pôle emploi. It is thus possible that an individual in the sample was unemployed at other periods before the drawing.

- does not work part-time hours exceeding half of full-time employment per month and has declared herself available to begin work immediately. This criterion allows for the protocol to exclude those undergoing a training program or an internship and those not actively searching for employment due to health reasons;
- demonstrates strong self-sufficiency related to job search.<sup>21</sup> The working group of Pôle emploi caseworkers who was involved in the design determined that the Bob Emploi tool was more likely to benefit those already capable of browsing the internet with ease;
- had already defined their job search strategy, especially with a target job in an acceptable geographic zone. Bob Emploi was considered more relevant to reacting to and redirecting a previously defined strategy than to helping users produce initial career objectives;
- and, has a valid email address and agreed to receive emails with information from Pôle emploi.

**Randomization and timeline** The implementation of meetings occurred over a 3 month period, in April, May, and June 2017. The sampling protocol was thus staggered in 3 waves: each month, a sampling was executed in each branch, in order to draw eligible individuals in the treatment group and in the control group. The reference was to form lists of 40 people invited per session, mirrored by 40 people as control group.<sup>22</sup> This resulted in a final sample composed of 212 277 individuals: 119 525 in the treatment group and 92 752 in the control group. Figure 1.1 provides a schematic representation of the timeline of the experiment. Note that for logistical reasons and contrary to information meetings, the reminder emails and the online survey were sent all at once to the entire sample.

## 1.4.2. Description of the sample

### 1.4.2.1. Descriptive statistics and balance checks

Table 1.1 provides descriptive statistics of our sample along with balance tests. Column (1) shows the mean of pre-intervention variables for individuals in the control group and column (2) displays the results of regressions to detect differences in means with the treatment group. This latter column shows that no coefficients are significantly different from zero, which confirms that the randomization was successful at balancing the treatment and the control groups. Thus the composition of the control group reveals the composition of the whole sample.

The composition of the sample is consistent with the eligibility requirements. All job seekers in the sample have less than a year of seniority in unemployment. A majority has been registered at Pôle emploi for more than 6 months (53.20 %) and only about one quarter has been unemployed for less than 3 months. Most individual are aged 26 to 55 although more than a quarter of the sample is in their

<sup>21</sup>The degree of self-sufficiency was estimated based on a variable that indicates the intensity of Pôle emploi assistance. At the time of registration, caseworkers assess job seekers' level of autonomy and assign them to one of three main assistance tracks. Job seekers that are considered to easily handle their job search benefit from a more distant monitoring and assistance from caseworkers.

<sup>22</sup>Small branches sometimes could not fill in 40 eligible job seekers in the control group and hence compiled reduced lists, which explains why the experiment resulted in a slightly higher number of individuals in the treatment group than in the control group.



twenties, which is also in line with the eligibility condition on unemployment seniority. The requirement on self-sufficiency in one’s job search partly explains that more than 80% of the sample earned a formal education degree, with 30.60 % having a vocational degree and as many as 34.55% who own a university degree. The sample is balanced on gender and targeted sectors of search.

#### 1.4.2.2. Survey respondents and attrition

Table 1.2 shows that there was low but significant unbalanced attrition on survey response between the treatment and the control groups. Response rate to the survey is 0.4 percentage points lower in the treatment group than in the control group. Respondents were more likely to be male than non-respondents, as shown in table 1.3. They are significantly older and are 12 percentage points more likely to have earned a university degree than non-respondents. They also have been unemployed for longer at the time of the survey.

### 1.4.3. Empirical strategy

#### 1.4.3.1. Intent-to-Treat estimation

In the empirical analysis in Section 2.5, we aim at identifying the causal impact of Bob Emploi on job search strategy and employment outcomes of unemployed individuals. Given the randomized design of the experiment with balanced treatment and control groups, we primarily rely on Intent-to-Treat estimations. We estimate equations of the following kind :

$$Y_{it} = \alpha_1 + \beta_1 T_i + \mu_{i1}^{a,m} + \epsilon_{it1} \quad (1.1)$$

where  $Y_{it}$  measures the outcome of interest, e.g., job search effort or employment, for individual  $i$  at time  $t$ .  $T_i$  is the dummy equal to one if the individual  $i$  is in the treatment group. To account for this multi-local multi-timing sampling, we add fixed effects  $\mu_{i1}^{a,m}$  for each strata defined by the month  $m$  of inclusion in the sample and the local Pôle emploi agency  $a$  of individual  $i$ .<sup>23</sup> As explained in the previous subsection, we drew job seekers to invite to information meetings from monthly lists of local Pôle emploi agencies over three consecutive months. In small agencies and to respect a standard procedure for the meetings, we had to sample more individuals in the treatment group than in the control group. At the aggregate level, this implies that treatment and control groups are not balanced across agencies, with an over-representation of small agencies in the treatment group.

$T_i$  is a dummy for treatment status and  $\beta_1$  thus captures the average effect of the treatment on outcome  $Y$  in month  $t$  after treatment.<sup>24</sup> Throughout the analysis, we cluster the robust error term  $\epsilon_{it1}$  by local agency.

In section 2.5, we mainly present results for  $t = 6$  months after the intervention and we show similar results for  $t = 18$  months in the appendix. The model lends itself to graphical presentations which we

<sup>23</sup> $\mu_{i1}^{a,m}$  is a dummy variable equals to 1 if the individual  $i$  is in the agency  $a$  and included in the sample at month  $m$  (either April, May or June)

<sup>24</sup>Controlling for agency and month strata, the random treatment assignment ensures that  $T_i$  is orthogonal to  $\epsilon_{it1}$ . Hence  $\beta_1$  identifies the average causal effect of the treatment even without controlling for any other covariate.

use to document treatment effects over time on employment outcomes.

### 1.4.3.2. Definition of take-up and LATE estimation

To estimate the individual effects of Bob Emploi on compliers, we turn to a LATE. Considering the experimental protocole, compliers are individuals who either created an account on Bob Emploi website or participated in an information meeting. Information meetings were an opportunity for job seekers to find out about other resources at Pôle emploi and to interact with their caseworker (who was often the person who invited them to the meeting). We cannot rigorously rule out that meetings had an effect on our outcomes in and by themselves and therefore we count as treated those who participated in a meeting. The list of created accounts is obtained by matching email addresses of Bob Emploi users recorded by Bayes Impact with email addresses of individuals in our sample. Survey data give us a second possible measure of Bob emploi usage on the sample of respondents, which we use as a robustness check on survey outcomes.<sup>25</sup>

Table 1.4 compares the two take-up definitions, either including survey responses on Bob Emploi usage or not, to verify that they are consistent and that our results are robust to using either one. Column (1) shows the take-up differential between the treatment and the control group based on meeting participation and Bob Emploi connection only. Column (2) runs the same regression restricting the sample to survey respondents, which allows a comparison with column (3) where we count as additional takers respondents who report to use Bob Emploi. We see in column (1) that the experiment was successful at creating a substantial take-up differential of 26.8% while keeping a participation in the control group to virtually zero. Reassuringly, take-up differentials among respondents only, showed in columns (2) and (3), are very close and highly correlated, although the mean in the control group is slightly higher when using survey responses.<sup>26</sup>

In tables of section 2.5, we therefore choose to show the results of LATE estimation defining as takers individuals who either attended an information meeting or connected to Bob Emploi website. The equation is similar to 1.1:

$$Y_{it} = \alpha_2 + \beta_2 \widehat{B}_i + \mu_{i2}^{a,m} + \epsilon_{it2} \quad (1.2)$$

where  $Y_{it}$  measures the outcome of interest for individual  $i$  at time  $t$ .  $\widehat{B}_i$  is a dummy for Bob Emploi take-up status, which we estimate by the appropriate first stage regression with  $T_i$  the treatment dummy :

$$B_i = \gamma + \delta T_i + \mu_{i0}^{a,m} + \eta_i \quad (1.3)$$

<sup>25</sup>Among other websites, respondents were asked whether they knew, used and liked Bob Emploi (the complete questionnaire is available in appendix 1.A). Hence we can count respondents who reported that they used Bob Emploi as takers of the treatment.

<sup>26</sup>In column (2), the differential increases to 42%. This is very close to what we obtain when we add survey responses in the take-up definition in column (3). Yet we see in column (3) that this definition identifies 9% of the control group as takers. The correlation between column (2) and (3) is 0.78. Discrepancies could be due to individual mistakes in survey responses (Bob Emploi being easily confounded with any Pôle emploi website). Job seekers might also have used a different address to connect to Bayes Impact. Finally, the 9% takers in the control group in column (3) might reflect a natural contamination due to information spreading during the three months when information meetings were organized in local agencies.

This ensures that controlling for  $\mu_{i2}^{a,m}$ ,  $\widehat{B}_i$  is orthogonal to the heteroskedastic error term  $\epsilon_{it2}$ , so that  $\beta_2$  captures the individual effect of the treatment on outcome  $Y$ .

### 1.4.3.3. Description of takers

Table 1.4 shows that the take-up differential between the treatment and the control groups is 26.8%. As the take-up is virtually zero in the control group, this take-up is only made of compliers of the treatment group. Compliers fall into three distinct categories: 14.07 % only attended an information meeting, 6.6 % did not attend a meeting but created an account on Bob Emploi, and the last 6.53 % both attended a meeting and created an account. The take-up regularly increased over time during the three months when the meetings were organized, followed by discrete bumps corresponding to the days when reminder emails were sent out. Using survey data, we observe that overall the intervention led to a significantly higher differential in Bob emploi usage compared to other websites presented during the meetings.<sup>27</sup> Thus we can interpret our results as mostly driven by Bob Emploi usage.

Table 1.5 shows the selection pattern of compliers along the same variables as table 1.1. Perhaps surprisingly, take-up markedly increases with age. Take-up is highest for job seekers above 50 years old, and lower for the youngest job seekers. This is not only due to older job seekers participating more to meetings as the correlation still holds if we restrict to creations of Bob Emploi accounts only. While younger job seekers tend to be more familiar with digital technologies and could have used Bob Emploi more easily, it seems on the contrary that the website rather appealed to older job seekers.

### 1.4.3.4. Heterogeneity analysis

#### Heterogeneity analysis by subgroups

Zero effects on average may hide heterogeneous effects for certain sub-populations. Since Bob Emploi is a numerical device, it could be of greater benefit to younger or more highly educated people. Conversely, to the extent that its role is to provide information accessible through other online means, it could help individuals who are less informed and less familiar with these tools. While it may benefit only certain groups of people, it could also be a waste of time for other individuals or provide them with bad advice.

Let  $G_i$  be a binary variable indicating whether the individual belongs to the subgroup under consideration. Notations are similar to those in section 1.4.3.1, except that we add  $\gamma_g$ , a coefficient neutralizing the effect of being in group  $G = 1$  on  $Y_{it}$  and  $\delta_g$  indicating an additional effect of the treatment when being in group  $G = 1$ . We test the significance of  $\delta_g$  to assess whether or not there are heterogeneous effects.

$$Y_{it} = \alpha_1 + \beta_1 T_i + \gamma_g G_i + \delta_g T_i \times G_i + \mu_{i1}^{a,m} + \epsilon_{it1} \quad (1.4)$$

There are, however, limits to this approach. Examining a large number of subgroups implies repeatedly testing the significance of  $\delta_g$  coefficients and thus possibly making false discoveries. However, Bob

<sup>27</sup>We measure that the differential in Bob emploi usage between the respondents in the treatment group and the control group reached 28%, whereas it is only 4% for the main other website that was presented during information meetings, called Emploi Store.

Emploi’s project is to provide individual advice based on individual information. It is then possible that the algorithm of the Bob Emploi site has helped individuals with very complex profiles that are not captured by these estimates within coarse subgroups. A finer exploration of heterogeneity can be carried out thanks to machine learning techniques.

### Generic machine learning (Chernozhukov et al. 2018)

Therefore, a second approach consists of estimating the impacts in the most personalized way possible. The individual characteristics considered are: age, gender, years spent working in the sector of interest, whether the job seeker looks for a part-time job or not, regional variables<sup>28</sup>, nationality, reason why the job seeker is on Pôle emploi’s lists<sup>29</sup>, readiness to take up a new position<sup>30</sup>, job sector, type of the desired contract (short or long term), level of qualification, type of accompaniment received at Pôle emploi agency, a proxy of the accepted home-work distances, the baseline benefits received, the reservation wage and the number of children. Following Rubin’s framework (1974b), let  $Y_i(1)$  be the individual’s unemployment duration if the individual is invited to use Bob Emploi, and  $Y_i(0)$  the unemployment duration if not. A given individual has a set of characteristics  $X_i$ . The impact of the invitation for a specific individual  $i$  relates his or her set of specific characteristics to the impact of the invitation. Formally this is the Conditional Average Treatment Effect (CATE) function:

$$\text{CATE}(X_i) = \mathbb{E}[Y_i(1) - Y_i(0)|X_i]$$

The relationships between these variables and the individual impact, and thus the  $\text{CATE}(\cdot)$  function can be non-linear and complex. In addition, not all the information available is necessarily relevant for calculating this function, and variable selection techniques may be required. The use of machine learning techniques is therefore promising.

The generic machine learning procedure requires bi-partitioning the data into two sets of equal size:  $A$  and  $M$ . The first stage proceeds on  $A$ . It consists in approximating  $\text{CATE}(\cdot)$  and a second function  $B(X_i) = \mathbb{E}[Y_i(0)|X_i]$ . To estimate the function  $\text{CATE}(\cdot)$  we consider two options, following the typology of Künzel et al. (2017):

- using a two-learner and predict separately  $\mathbb{E}[Y_i(1)|X_i]$  among the invited people and  $\mathbb{E}[Y_i(0)|X_i]$  among the control group. Three classical algorithms are used to perform this type of estimation: gradient boosting, random forest (Breiman 2001) and elastic net (Zou and Hastie 2005). When implementing elastic net, age squared is added to the list of covariates, and gender and nationality are interacted together and with every other covariates. No interaction is added for the tree-based methods since relevant ones should automatically be found when growing the forest.
- using a single-learner and directly estimating  $\text{CATE}(\cdot)$  with an algorithm designed for causal estimation. The chosen algorithm is the causal forest with local centering (Wager and Athey 2017,

<sup>28</sup>The french department (country subdivision) and whether the individual lives in a sensitive urban zone or not.

<sup>29</sup>It could be the end of a short-term contract or of an internship for example.

<sup>30</sup>The job seeker may not be available at the time if he or she is undergoing training, is on maternity leave or is on sick leave, for example.

[Athey et al. 2019](#)).

To estimate  $B(\cdot)$ , we keep the prediction of  $\mathbb{E}[Y_i(0)|X_i]$  among the control group in the first case, and we estimate it separately among the control group in the second case, using a random forest.

Once both  $\widehat{\text{CATE}}(\cdot)$  and  $B(\cdot)$  have been approximated, one can proceed to the second step of the generic machine learning which is done on the second half of the sample,  $M$ . Let  $\widehat{\text{CATE}}(\cdot)$  and  $\hat{B}(\cdot)$  be the corresponding approximations, and  $\overline{\text{CATE}}_M = \frac{1}{\#\{i \in M\}} \sum_{i \in M} \widehat{\text{CATE}}(X_i)$ , i.e the mean predicted impact on sample  $M$ . Following [Chernozhukov et al. \(2018\)](#), we first run the linear regression bellow, weighted by  $w(X_i) = \frac{1}{p(X_i)(1-p(X_i))}$  where  $p(\cdot)$  is the propensity score:

$$Y_{it} = \theta_0 + \theta_1 \hat{B}(X_i) + \theta_2 \widehat{\text{CATE}}(X_i) + \beta^{\text{ITT}}(T_i - p(X_i)) + \beta^{\text{HET}}(T_i - p(X_i)) \left( \widehat{\text{CATE}}(X_i) - \overline{\text{CATE}}_M \right) + \epsilon_{it1} \quad (1.5)$$

Authors showed that (a)  $\beta^{\text{ITT}}$  corresponds then to the mean effect of being invited, and that (b)  $\beta^{\text{HET}}$  is non-zero if  $\widehat{\text{CATE}}(\cdot)$  actually captured heterogeneous impacts. Note that  $\beta^{\text{HET}} = 0$  would always have two possible interpretations: either there is no heterogeneous impacts, or  $\widehat{\text{CATE}}(\cdot)$  fits to badly the true  $\text{CATE}(\cdot)$  to allow us to detect heterogeneity<sup>31</sup>.

We also run a second regression on sample  $M$ , aiming at evaluating quintiles of impacts. The idea is to order the  $\widehat{\text{CATE}}(X_i)$  and divide the individuals into 5 groups of equal sizes: from the lowest approximate impact to the highest. Let  $l_k$  be the  $k$ -th quintile of  $\widehat{\text{CATE}}(\cdot)$  evaluated on sample  $M$ . Let  $(Q_{ki})_{k \in \llbracket 1,5 \rrbracket}$  be the binary variables indicating the individual's quintile group, based on the predicted impacts, i.e.  $Q_{ki} = \mathbb{1}\{l_{k-1} \leq \widehat{\text{CATE}}(X_i) < l_k\}$ , with  $l_0$  simply being the lowest value of  $\widehat{\text{CATE}}(X_i)$ .

The objective is to estimate the average effects within these quintile groups. These impacts are called Sorted Grouped Average Treatment Effects (GATES). The following linear regression, weighted again by  $w(X_i)$ , gives unbiased estimates of the GATES  $(\gamma_k)_{k \in \llbracket 1,5 \rrbracket}$ :

$$Y_{it} = \theta_0 + \theta_1 \hat{B}(X_i) + \theta_2 \widehat{\text{CATE}}(X_i) + \sum_{k=1}^K \gamma_k \cdot (T_i - p(X_i)) \cdot Q_{ki} + \epsilon_{it1} \quad (1.6)$$

Not only can we determine whether the impacts within quintile groups are significant, but we can also test the difference between them. In particular, we can examine whether the group with the highest predicted impacts does have a significantly different effect from the group with the lowest predicted impacts.

Thus, essentially, we use this approach to (1) approximate individualized impacts of the invitation in a very fine and data-driven manner using machine learning techniques, (2) assess heterogeneity thanks to the estimation of  $\beta^{\text{HET}}$  and test for its significance, (3) investigates differences in impacts among different groups.

To reduce the effect of sample splitting on our estimates, we repeat the bi-partition in  $A$  and  $M$ , the

---

<sup>31</sup> $\theta_0$ ,  $\theta_1$  and  $\theta_2$  are coefficient of secondary importance that will not be commented here. Note that if the regressor  $\widehat{\text{CATE}}(X_i)$  interacted with  $\theta_2$  is not explicitly included in the regression of the theoretical part in [Chernozhukov et al. \(2018\)](#), it is included in the description of the algorithm and in the authors' code. This is why we choose to follow the implementation of [Chernozhukov et al. \(2018\)](#) and include this control.

training of the algorithms and the regressions 100 times and correct p-values and confidence intervals following the authors.

## 1.5. Results

In this section, we summarize the main results of our experiment. We first consider the effects of Bob Emploi on job search strategy, namely search effort, target and methods. We then analyze the treatment effects on employment and well-being. All regressions in this section show the results of regressions on outcomes measured 6 months after the intervention both in ITT (panel A) and LATE (panel B). They are replicated with outcomes 18 months after the intervention in the appendix.

### 1.5.1. Effort and scope of job search

Table 1.6 shows how Bob Emploi affected job seekers' reemployment expectations, search effort and target. Column (1) first presents the impact of Bob Emploi on individuals' expectations of reemployment as reported in the survey. On one hand, Bob Emploi advice can improve job seekers' search strategy, and thus increase their effort and expectations. On the other, information may make them realize that their reemployment chances are low, which could either make them look harder or discourage them. In the control group, we see that 57% respondents believed that they would find a job within three months after the date of the survey. There appears to be no significant difference in the treatment group. This rules out a predominant demotivation impact which is a legitimate worry in any information intervention.

By identifying reachable targets and providing concrete methods of search, Bob Emploi may also increase the returns to time spent on the job search. Combined with a general motivation effect, this can lead users to increase the number of hours spent searching. In extreme cases, if it sufficiently increases the returns to each unit of effort, providing good information can in theory have a net decreasing impact on search effort. To measure job seekers' search effort, we compute both the total number of hours respondents declare to spend on job search (column (2)) and the number of online applications they usually send (columns (3) and (5)). Column (2) shows that people invited to use Bob Emploi do not spend significantly more time searching for a job, nor do they send more online applications.<sup>32</sup>

If Bob Emploi does not affect time spent on job search, it may improve its efficiency. As a first lever to make the job search more efficient, Bob Emploi may lead the job seeker to widen the scope of her search by providing new information on labor demand. In fact, when the algorithm identifies that the job seeker aims for an occupation or a region where the labor market is particularly tight, it encourages her to consider jobs in neighboring regions or to apply to related occupations. We test for such widening effects using survey questions on applications as well as observations of applications in the administrative dataset from Pôle emploi job board. More precisely, column (4) shows whether treated respondents apply more often to jobs outside their municipality. In column (6) we investigate for an

---

<sup>32</sup>In the control group, respondents declare to spend on average 8.5 hours per week searching, which is in line with other surveys (e.g. [Krueger and Mueller \(2012\)](#), [DellaVigna et al. \(2020\)](#)) and half of them report to send an unsolicited application every week. This translates in only 2.4 applications on average according to the administrative application dataset. Apart from reporting mistake, this discrepancy may reflect the selection bias among respondents or the limited coverage of Pôle emploi application website.

effect on the total number of observed applications outside of one’s municipality. Both variables yield no significant impact. Similarly, in column (7) we see no change in the occupational scope of search, as measured by the number of observed online applications in different occupations.

### 1.5.2. Job search methods

Bob Emploi not only offers a diagnosis of reemployment perspectives but connects the information with actionable advice of how to improve one’s search strategy. Table 1.7 leverages our particularly rich set of variables to explore the extent to which using Bob Emploi modified job seeker’s search techniques. These variables reflect frequent recommendations given on the website, such as leveraging one’s personal and professional networks or adapting one’s application to the specifics of each job listing. Bob Emploi also sorted online tools that it often referred job seekers to and it is therefore interesting to know if this led job seekers to use more assistance websites. These different techniques are analyzed using survey data and presented in columns (1) to (7) of table 1.7. We see in column (1) that Bob Emploi increases the reported use of personal network by 1.3 percentage point in the ITT estimation, and by 3.1 in the LATE estimation. However, the effect does not hold for professional network. Looking at columns (3) to (5), we detect no significant effect on either of the most frequent suggestions made by Bob Emploi (adapting one’s CV or cover letter to the job listing, or calling back a recruiter after an interview). Turning to columns (6) and (7), we see that Bob Emploi significantly increases the number of websites used, although the effect is limited to websites created and hosted by Pôle emploi. This effect is perhaps not surprising as Bob Emploi often refers to less well-known Pôle emploi websites and so did caseworkers who led the information meetings. Finally, considering the connections between Bob Emploi and Pôle emploi, it is of particular interest to take a closer look at the impact of Bob Emploi on job seekers’ interactions with Pôle emploi. In fact, one may worry that such private website discourage job seekers from engaging in conversations with their caseworkers or from participating in the many programs that are offered by Pôle emploi. If anything, we observe the opposite effect. Column (8) shows that Bob Emploi increases the fraction of job seekers who had had at least one interaction with their caseworker within the 6 months that followed the intervention, and column (9) shows that there is no impact on the number of attended programs and workshops at Pôle emploi. These last results might be partly driven by participation to information meetings by itself.<sup>33</sup>

### 1.5.3. Well-being and life balance

Table 1.8 shows the impact of using Bob Emploi on job seekers’ reported well-being. Job seekers’ demotivation and dissatisfaction is an important concern in economics as well as in sociology or psychology.<sup>34</sup> Along with a direct loss of material wealth and the unrewarding aspect of job search, unemployment

---

<sup>33</sup>During the meetings, caseworkers put an emphasis on the main platform called ‘Emploi Store’ where Pôle emploi gathers several search websites including Bob Emploi and other less well-known search online tools created by Pôle emploi. Besides, the data on recorded interactions with caseworkers did not allow us to properly distinguish with interactions that were only due to invitations to the information meetings.

<sup>34</sup>The importance of discouragement and unhappiness is supported by empirical evidence: while remaining longer unemployed, workers express more dissatisfaction and unhappiness with their lives and are particularly sad during episodes of job search (Krueger et al. 2011).



is often associated with social bounds dislocation, loss of meaning and of social identity.<sup>35</sup> With time, individuals who stay unemployed may get discouraged and drop out of the labor market. This can become a vicious circle when job search itself does not lead to positive outcomes because of a bad job search strategy: demotivation and self-deprecation lower individuals' energy and might preclude them from grasping promising job opportunities. This in turn would lead to more disappointments and unsuccessful searches. By providing actionable advice to optimize one's search effort in a user-friendly manner, Bob Emploi has the potential to give job seekers a feeling of purpose, support and achievement. If this leads to better search outcomes, it can increase motivation and optimism.

Columns (1), (2) and (3) report the answers to the three questions of the survey that relate to well-being. In these questions, respondents had to provide a number between 0 and 10 to indicate their level of overall well-being, of motivation during their search and the extent to which they felt supported. Means in the control group reveal that respondents estimate to have an average well-being but feel significantly more motivated than they feel supported. This would provide a rationale to create more advice tools such as Bob Emploi. Nevertheless, as can be seen from both panels, Bob Emploi does not seem to affect positively any of these measures.

Column (4) of table 1.8 summarizes three questions asking about job seekers' activities outside of job search, namely whether they take part in sport, arts, culture or community activities at least once a month. In fact, Bob Emploi often recommends to its users that they keep a balanced life and get involved in some social activities such as volunteering. Column (4) shows the maximum value of the three dummies, so that a respondent who engages in either sport, arts or community activities is counted as 1. As with previous outcomes, we don't observe any significant difference between the treatment and the control groups.

#### 1.5.4. Employment outcomes

The main objective of Bob Emploi is to help job seekers get reemployed faster. To identify any difference between the treatment and the control groups, in particular any acceleration in reemployment, we look at individual paths in and out of unemployment over time.

As explained in section 3.2.2, we first measure whether job seekers experience any employment episode over the months that follow the intervention. As this variable does not take into account the quality of the job obtained, especially regarding their stability, we also look more specifically at episodes in long-term jobs. This is represented in figure 1.2. Panel (a) shows the fractions of individuals who have an employment episode over the 18 months following the intervention: the top graph indicates the levels in both treatment and control groups, while the bottom graph represents the treatment effect, along with confidence interval. The concave pattern of the evolution of levels is in line with models of job search frictions and other empirical evidence. Strikingly, 18 months after the end of the intervention, still 40% of the sample have no employment episode at all. Yet, as can be seen clearly from the bottom graph, the treatment appears to have no effect on the likelihood to have an employment episode faster. Panel (b)

---

<sup>35</sup>Psychological studies show that unemployed individuals have lower psychological and physical well-being (see McKee-Ryan et al. 2005 for a meta-analysis).



works similarly for long-term jobs. Consequently, levels are lower than in panel (a) and only reach 25% after 18 months. We do not see any impact of the treatment on this variable either. Considering the upper bound of the 95 percent confidence interval, we can reject any effect higher than 0.5 percentage points on experiencing some employment episode within the 18 months following the intervention.

To detect marginal changes in the duration of employment, we complement these variables by counting the cumulative number of days in unemployment, that is, the number of days registered at Pôle emploi. Structured in a similar fashion as the previous figures, panel (a) of figure 1.3 shows that this does not enable us to see any effect on Bob Emploi.<sup>36</sup>

Lastly, from a cost-benefit analysis, it may be interesting to see if Bob Emploi leads to any reduction in the total amount of unemployment benefits. This is what we show in panel (b) of figure 1.3, plotting the mean cumulative benefits received by job seekers in the sample. Mean levels hide some heterogeneity, as about 30% of job seekers in the sample receive no benefits. Once again, the bottom graph makes it clear that being invited to use Bob Emploi made no difference in the final amount of benefits received.

### 1.5.5. Heterogeneity analysis on unemployment duration

One of the main expectation from Bob Emploi project was to reduce the duration of unemployment by accompanying the unemployed. The heterogeneity analysis below allows us to go beyond the mean analyses conducted so far and examine whether the invitation to use the Bob Emploi website has an effect for a specific sub-sample. According to section 1.4.3.3, the individuals invited to register on the site who actually participated have specific characteristics. The site may work better for some job seekers than others.

A first, traditional level of analysis will consist of examining average effects for subgroups that we constructed beforehand. A second approach is to look for the most affected subgroups by using a personalized prediction of the effects of the treatment by machine learning algorithms.

#### 1.5.5.1. Impacts for a small number of subgroups

The regression estimates of unemployment duration 6 months after the start of the experiment are detailed in table 1.9 and those, 18 months after are in table 1.10. A negative coefficient would correspond to a reduction in unemployment duration in days. The "group effect" coefficient corresponds to the  $\gamma_g$  estimate, and the "treatment effect" coefficient corresponds to the  $\beta_1$  estimate, i.e the effect of being invited when the individual is in group  $G = 0$ . The third row presents the estimate of the  $\delta_g$  coefficient, the term in front of the interaction between the treatment dummy and the group dummy in the column.

Four categories of age and four levels of education are considered. None of the coefficients in the third row are significant, after 6 months as well as after 18 months. This suggest that there is no effect of the invitation to register on the site for these specific subgroups. However, one coefficient is significant at level 10%: the estimate of  $\beta_1$  in column (1) with  $G = 1$  when the individual is less than 25 years old. This could suggest a reduction of 0.59 day of unemployment in the period of six months for people older than 25. However, the non-significance of  $\delta_G$  does not argue in favour of an effect specific to this

---

<sup>36</sup>Due to a scaling effect, levels are virtually identical in the treatment and control groups all over the period of observation.

group. Moreover there is no specific effect for older groups. It is thus hard to conclude that there are heterogeneous impacts. The significance of this coefficient may only be due to multiple hypothesis testing.

### 1.5.5.2. Machine learning approach

Results of the generic machine learning approach are reported in table 1.10 and figure 1.4. Performance of the algorithms is measured in table 1.11(a) using Chernozhukov et al. (2018)’s criteria. According to these, random forest and causal forest give the best approximations of individualized impacts. We then report estimations using these two best algorithms. Table 1.11(b). contains estimations of  $\beta^{\text{ITT}}$  and  $\beta^{\text{HET}}$  in regression 1.5. Both are non-significant. The absence of effect in mean suggested by the non-rejection of  $\beta^{\text{ITT}} = 0$  is consistent with results in section 1.5.4: ITT estimates on unemployment duration in days is not significant. Moreover, we cannot reject  $\beta^{\text{HET}} = 0$ , meaning that heterogeneity has not been detected. Rigorously, this could have three sources: (a) either approximations of the individual treatment effect by the algorithms are bad, (b) or there is no heterogeneity, (c) or sample splitting has caused a lack of statistical power. Considering explanation (a), trying different algorithms relying on different assumptions should preserve us from totally missing heterogeneity if there is any. One could nonetheless try supplementary algorithms to fit better  $\text{CATE}(\cdot)$ . Regarding (c), the lack of statistical power due to sample splitting can be a problem when using the generic machine learning approach. However, considering the very large sample size (212 277 individuals) and the take-up here, this should not be a serious issue. Therefore, (b) the absence of heterogeneity appears to be the most plausible explanation.

We also report GATES estimates in table 1.11(c) and compare the difference between both extreme quintiles of treatment effects. These corresponds to estimations of coefficients  $\gamma_1$  and  $\gamma_5$  in regression 1.6. We also test for their difference. This step is not mandatory since no heterogeneity seems to have been detected in the first stage, but they confirm that the groups do not show different average effects. No matter what machine learning algorithm we use, neither coefficients nor their differences are statistically significant, suggesting that there is no group of individuals being more affected than another. For more details, we report graphically all the GATES in figure 1.4. Once again, these are consistent with the suspicion of absence of heterogeneity.

## 1.6. Conclusion

This evaluation offers a unique opportunity to measure the potential of private websites dedicated to job searches. Partnering with the French public employment services, we implemented a large-scale randomized experiment to test the impact of Bob emploi, a private website using data analysis to provide job search assistance. We find that the website did not have any significant impact on employment. The limited effects we detect on some search variables appear not to be sufficient to augment job seekers’ likelihood of reemployment. The large sample included in the experiment and the significant take-up rate ensure that these null effects are precisely estimated and cannot be the result of the intervention

not taking place. We did not identify any subgroups by age or education level that would experience a greater reduction in unemployment duration thanks to the website. The more personalized exploration of treatment effects by machine learning algorithms also fails to detect heterogeneous effects.

Our findings provide evidence that the standard job search assistance services offered by private websites are unlikely to help job seekers. Bob emploi, which was designed by experienced tech developers and benefited from significant public funding and rich administrative data on labor markets, was in a better position than the majority of existing private websites to have a positive impact. We cannot rule out that better data algorithms coupled with matching functionalities may be more effective. Yet, the story of Bob emploi makes it clear that these improvements cannot be considered low-hanging fruit. Overall, our results suggest that the enthusiasm around digital platforms like Bob emploi may have to be toned down.

## 1.7. Tables

Table 1.1: Difference in means between the treatment and the control groups

	Control group mean (1)	Regression coefficient on treatment dummy (2)
Female	50.72 [ 50.00 ]	-0.128 ( 0.236 )
<b>Age</b>		
Less than 25 years old	23.11 [ 42.15 ]	-0.286 ( 0.173 )
Between 25 and 39 years old	45.42 [ 49.79 ]	0.265 ( 0.216 )
Between 40 and 54 years old	23.33 [ 42.29 ]	0.068 ( 0.192 )
More than 55 years old	8.14 [ 27.35 ]	-0.047 ( 0.127 )
<b>Seniority in unemployment</b>		
Less than 3 months	25.25 [ 43.44 ]	0.100 ( 0.181 )
Between 3 and 6 months	21.31 [ 40.95 ]	-0.207 ( 0.179 )
More than 6 months	53.20 [ 49.90 ]	0.105 ( 0.223 )
<b>Education Level</b>		
No high-school diploma	16.19 [ 36.83 ]	-0.022 ( 0.169 )
Vocational degree	30.60 [ 46.08 ]	-0.123 ( 0.190 )
High school diploma	18.66 [ 38.96 ]	0.146 ( 0.163 )
University degree	34.55 [ 47.55 ]	-0.000 ( 0.198 )
N	92752	212277

**Notes:** This table reports summary statistics for the control group as well as the results of balance checks testing equal means in the treatment and the control groups over a set of pre-intervention covariates. Column (1) shows average values of characteristics for the control group, with standard deviations in brackets. Column (2) reports regression coefficients on treatment group dummy with standard errors in parentheses. All regressions include agency x month fixed effects. Standard errors are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.2: Survey Respondent Attrition

	Survey Respondent (1)
Treatment	-0.004 (0.002)**
Control Mean	0.149
N	212277

**Notes:** This table shows the coefficient of a regression where the dependent variable is a survey response dummy and the independent variable is a treatment group dummy, reflecting the differential attrition to the survey between the treatment and the control groups. The regression includes agency x month fixed effects. Standard errors, in parentheses, are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.3: Difference in means between survey respondents and non-respondents

	Non-respondent mean (1)	Regression coefficient on respondent dummy (2)
Female	49.38 [ 50.00 ]	8.590 ( 0.308 ) ***
<b>Age</b>		
Less than 25 years old	24.35 [ 42.92 ]	-9.147 ( 0.270 ) ***
Between 25 and 39 years old	46.01 [ 49.84 ]	-4.529 ( 0.372 ) ***
Between 40 and 54 years old	22.09 [ 41.48 ]	9.122 ( 0.338 ) ***
More than 55 years old	7.55 [ 26.43 ]	4.554 ( 0.237 ) ***
<b>Seniority in unemployment</b>		
Less than 3 months	25.51 [ 43.59 ]	-2.263 ( 0.279 ) ***
Between 3 and 6 months	21.53 [ 41.10 ]	-1.015 ( 0.266 ) ***
More than 6 months	52.73 [ 49.93 ]	3.241 ( 0.309 ) ***
<b>Education Level</b>		
No high-school diploma	16.89 [ 37.47 ]	-4.819 ( 0.233 ) ***
Vocational degree	31.86 [ 46.59 ]	-6.319 ( 0.294 ) ***
High school diploma	18.99 [ 39.22 ]	-1.466 ( 0.226 ) ***
University degree	32.26 [ 46.75 ]	12.601 ( 0.312 ) ***
N	181072	212277

**Notes:** This table reports some summary statistics for individuals who did not respond to the survey as well as the results of tests of equal means with the respondents over a set of pre-intervention covariates. Column (1) shows average values of characteristics for non-respondents, with standard deviations in brackets. Column (2) reports regression coefficients on survey response dummy with standard errors in parentheses. All regressions include agency and month fixed effects, weighted by the empirical probability of being assigned to the treatment group for each agency x month strata. Standard errors are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.4: First Stage

	Bob account	Attended meeting	Bob account + Attended meeting	Bob account + Attended meeting	Bob account + Attended meeting + Reported Use
	(1)	(2)	(3)	(4)	(5)
Treatment	0.130 (0.002)***	0.205 (0.006)***	0.269 (0.006)***	0.421 (0.007)***	0.441 (0.007)***
Control Mean	0.002	0.000	0.002	0.004	0.091
N	212052	212052	212052	31171	31171
Correlation	0.662	0.851	1.000	1.000	0.788
Sample	All	All	All	Survey	Survey

**Notes:** This table reports regression coefficients on a treatment group dummy estimating equation 1.3, for different types of take-up. The dependent variables in columns (1) to (3) are dummy variables for the creation of a Bob emploi account, the attendance to an information meeting, the creation of a Bob emploi account or the attendance to an information meeting, respectively. The sample used in columns (1) to (3) is the whole sample. Column (4) shows the results of the same regression as (3) on the restricted sample of survey respondents. Column (5) uses as dependent variable a dummy for having either created a Bob emploi account, attended a meeting or reported using Bob emploi in the survey. The correlation coefficient shows the correlation with the dependent variable used in column (4). Means in the control group are computed separately. All regressions include agency x month fixed effects. Standard errors in parentheses are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.5: Difference in means between takers and non-takers in the treatment group

	Non-taker mean (1)	Regression coefficient on taker dummy (2)
Female	50.40 [ 50.00 ]	1.204 ( 0.359 ) ***
<b>Age</b>		
Less than 25 years old	25.13 [ 43.38 ]	-11.345 ( 0.342 ) ***
Between 25 and 39 years old	46.28 [ 49.86 ]	-5.503 ( 0.413 ) ***
Between 40 and 54 years old	21.78 [ 41.27 ]	9.044 ( 0.357 ) ***
More than 55 years old	6.81 [ 25.20 ]	7.804 ( 0.278 ) ***
<b>Seniority in unemployment</b>		
Less than 3 months	25.38 [ 43.52 ]	-1.117 ( 0.347 ) ***
Between 3 and 6 months	21.51 [ 41.09 ]	-1.111 ( 0.317 ) ***
More than 6 months	52.88 [ 49.92 ]	2.197 ( 0.396 ) ***
<b>Education Level</b>		
No high-school diploma	16.24 [ 36.88 ]	-1.012 ( 0.281 ) ***
Vocational degree	31.15 [ 46.31 ]	-0.472 ( 0.381 )
High school diploma	19.35 [ 39.51 ]	-2.256 ( 0.298 ) ***
University degree	33.25 [ 47.11 ]	3.740 ( 0.425 ) ***
N	94906	119525

**Notes:** This table reports some summary statistics for individuals in the treatment group who did not take up the treatment as well as the results of tests of equal means with the takers over a set of pre-intervention covariates. Column (1) shows average values of characteristics for non-takers, with standard deviations in brackets. Column (2) reports regression coefficients on take-up dummy with standard errors in parentheses. All regressions include agency and month fixed effects, weighted by the empirical probability of being assigned to the treatment group for each agency x month strata. Standard errors in parentheses are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.



Table 1.6: Impact on reemployment expectations, effort and scope of job search

	Survey Data:				Administrative Data:		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Expects to find a job in less than 3 months	Hours per week spent on job search	At least one unsolicited application per week	Applies to job beyond own municipality	Number of online applications	Number of online applications beyond own municipality	Number of online applications beyond past occupation
<b>Panel A: ITT</b>							
Treatment	-0.001 (0.007)	-0.097 (0.098)	0.005 (0.006)	-0.001 (0.006)	0.023 (0.067)	0.034 (0.063)	-0.012 (0.033)
<b>Panel B: LATE</b>							
Taker	-0.002 (0.015)	-0.231 (0.229)	0.011 (0.014)	-0.003 (0.013)	0.086 (0.247)	0.125 (0.234)	-0.044 (0.122)
Control Mean	0.576	8.517	0.456	0.523	2.439	2.117	2.057
N	19523	31171	31171	31171	212052	212052	212052

**Notes:** This table reports regression coefficients on a treatment group dummy estimating ITT equation 1.1 in panel A and LATE equation 1.2 in panel B. Means in the control group are computed separately. Columns (1) to (4) use dependent variables from survey data, columns (5) to (8) use dependent variables from administrative data which are available for the entire sample. All regressions include agency x month fixed effects. Standard errors in parentheses are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.7: Impact on job search methods

	Survey Data:						Administrative Data:		
	(1) Uses personal network	(2) Uses professional network	(3) Adapts CV to job listing	(4) Adapts cover letter to job listing	(5) Follows up with recruiting firms	(6) Number of used Pôle emploi websites	(7) Number of used private websites	(8) At least one meeting with caseworker	(9) Number of workshops attended at Pôle emploi
<b>Panel A: ITT</b>									
Treatment	0.012 (0.006)**	0.006 (0.006)	0.009 (0.007)	-0.004 (0.006)	0.011 (0.006)*	0.070 (0.009)***	-0.009 (0.011)	0.024 (0.003)***	-0.001 (0.010)
<b>Panel B: LATE</b>									
Taker	0.029 (0.014)**	0.013 (0.014)	0.020 (0.015)	-0.008 (0.012)	0.026 (0.014)*	0.167 (0.022)***	-0.022 (0.026)	0.090 (0.012)***	-0.004 (0.036)
Control Mean	0.691	0.571	0.633	0.760	0.430	1.446	0.855	0.572	0.653
N	31171	31171	22970	23106	23044	31171	31171	212052	212052

**Notes:** This table reports regression coefficients on a treatment group dummy estimating ITT equation 1.1 in panel A and LATE equation 1.2 in panel B. Means in the control group are computed separately. Columns (1) to (7) use dependent variables from survey data, columns (8) use dependent variables from administrative data which are available for the entire sample. All regressions include agency x month fixed effects. Standard errors in parentheses are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

Table 1.8: Impact on well-being, motivation and life balance during job search

	Overall well-being	Feels motivated during jobsearch	Feels supported during jobsearch	Partici- pates in non-job related activities
	(1)	(2)	(3)	(4)
<b>Panel A: ITT</b>				
Invité	0.026 (0.025)	-0.036 (0.027)	0.051 (0.033)	0.008 (0.005)
<b>Panel B: LATE</b>				
Participant	0.061 (0.059)	-0.085 (0.064)	0.121 (0.078)	0.018 (0.011)
Control Mean	5.182	7.069	3.884	0.779
N	31171	31171	31171	30276

**Notes:** This table reports regression coefficients on a treatment group dummy estimating ITT equation 1.1 in panel A and LATE equation 1.2 in panel B. Means in the control group are computed separately. All columns use dependent variables from survey data. All regressions include agency x month fixed effects. Standard errors in parentheses are clustered at the agency level and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively.

### 1.7.1. Impacts on unemployment duration 6 months after: heterogeneity analysis

#### 1.7.1.1. Mean difference for a small number of subgroups

Table 1.9: Effect on unemployment duration 6 months after - results from regressions 1.4

	Less than 25 years old (1)	Between 25 and 39 years (2)	Between 40 and 54 years (3)	More than 55 years old (4)	No high school diploma (5)	Vocational degree (6)	High school diploma (7)	University degree (8)
<b>Group effect</b>	-21.46 0.59***	-2.72 0.47***	15.60 0.52***	18.77 0.75***	1.02 0.67	1.15 0.53**	-3.20 0.61***	0.51 0.56
<b>Treatment effect</b>	-0.59 0.34*	-0.35 0.39	-0.44 0.34	-0.40 0.31	-0.50 0.33	-0.44 0.35	-0.47 0.34	-0.44 0.36
<b>Treatment x Group effect</b>	0.28 0.78	-0.25 0.61	-0.11 0.68	-0.73 1.01	0.21 0.82	-0.08 0.65	0.06 0.82	-0.08 0.65

**Notes:** Standard errors are in parentheses and are clustered at the Pôle emploi agency level. Statistical significance is computed based on the robust p-value and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively. Each column reports the results from an OLS regression estimating on each subgroup the unemployment duration on the treatment dummy. The OLS specification includes month and agency fixed effects and are weighted with the empirical probability of being assigned to the treatment group for each agency x month cell.

### 1.7.1.2. Generic machine learning (100 splits)

Table 1.10: Impacts on unemployment duration 6 months after

(a) Evaluation criteria of algorithms' performance

	Elastic Net with expansions	Boosting	Random Forest	CF with LC
Best BLP ( $\Lambda$ )	0.31	0.24	0.36	0.24
Best GATES ( $\bar{\Lambda}$ )	0.86	0.86	1.00	0.91

**Notes:** Medians over 100 splits in half. The best ML method can be chosen by maximizing  $\Lambda$  and  $\bar{\Lambda}$ . Maximizing  $\Lambda$  is similar to maximizing the correlation between the approximation of the individual treatment effect and its true value. The machine learning method maximizing this criterion can be considered as the most relevant to estimate coefficients in equation 1.5, reported in table 1.11(b). The machine learning method maximizing  $\bar{\Lambda}$  is the most relevant to estimate regression 1.6 whose results are reported in table 1.11(c). We follow the authors and combine both choices to present the two best methods. See Chernozhukov et al. (2018) for more details.

(b)  $\beta^{ITT}$  and  $\beta^{HET}$  estimates from regression 1.5

	Random Forest		CF with LC	
	$\beta^{ITT}$	$\beta^{HET}$	$\beta^{ITT}$	$\beta^{HET}$
Estimate	-0.57	0.04	-0.58	0.01
Confidence interval (90%)	[-1.37, 0.22]	[-0.05, 0.12]	[-1.38, 0.21]	[-0.23, 0.26]
Adjusted p-value	(0.31)	(0.75)	(0.30)	(1.00)

**Notes:** Medians over 100 splits in half.  $\beta^{ITT}$  corresponds to the mean effect of being invited on the unemployment duration.  $\beta^{HET}$  is non-zero if heterogeneity has been detected. Following Chernozhukov et al. (2018), confidence intervals are adjusted and so are p-values.

(c) GATES and difference between extreme estimated quintile treatment effects

	Random Forest			CF with LC		
	Q5	Q1	Difference	Q5	Q1	Difference
Estimate	-0.14	-1.12	1.11	-0.70	-0.66	0.11
Confidence interval (90%)	[-1.91, 1.63]	[-2.89, 0.66]	[-1.40, 3.62]	[-2.48, 1.09]	[-2.44, 1.12]	[-2.40, 2.62]
Adjusted p-value	(1.00)	(0.44)	(0.38)	(0.84)	(0.85)	(0.59)

**Notes:** Medians over 100 splits in half. Quintiles Q5 and Q1 are estimations of  $\gamma_5$  and  $\gamma_1$  in regression 1.6. We also test for the difference between both. Following Chernozhukov et al. (2018), confidence intervals are adjusted and so are p-values.

## 1.7.2. Impacts on unemployment duration 18 months after: heterogeneity analysis

### 1.7.2.1. Mean difference for a small number of subgroups

Table 1.10: Effect on unemployment duration 18 months after - results from regressions 1.4

	Less than 25 years old (1)	Between 25 and 39 years (2)	Between 40 and 54 years (3)	More than 55 years old (4)	No high school diploma (5)	Vocational degree (6)	High school diploma (7)	University degree (8)
<b>Group effect</b>	-63.57 1.58***	-12.11 1.40***	51.14 1.54***	52.86 2.56***	10.35 1.83***	7.26 1.47***	-8.44 1.57***	-7.49 1.58***
<b>Treatment effect</b>	-1.05 0.94	0.63 1.14	-0.22 0.94	-0.45 0.87	-0.35 0.93	-0.78 0.98	-0.12 0.96	-0.26 1.04
<b>Treatment x Group effect</b>	2.14 1.97	-2.24 1.71	-0.74 2.02	1.12 3.18	-0.16 2.30	1.33 1.85	-1.31 2.18	-0.37 1.90

**Notes:** Standard errors are in parentheses and are clustered at the Pôle emploi agency level. Statistical significance is computed based on the robust p-value and \*\*\*, \*\*, and \* indicate significance at 1, 5, and 10% respectively. Each column reports the results from an OLS regression estimating on each subgroup the unemployment duration on the treatment dummy. The OLS specification includes month and agency fixed effects and are weighted with the empirical probability of being assigned to the treatment group for each agency x month cell.

### 1.7.2.2. Generic machine learning (100 splits)

Table 1.11: Impacts on unemployment duration 18 months after

(a) Evaluation criteria of algorithms' performance

	Elastic Net with expansions	Boosting	Random Forest	CF with LC
Best BLP ( $\Lambda$ )	0.73	0.89	0.61	0.65
Best GATES ( $\bar{\Lambda}$ )	5.04	5.48	4.91	5.44

*Notes:* Medians over 100 splits in half. The best ML method can be chosen by maximizing  $\Lambda$  and  $\bar{\Lambda}$ . Maximizing  $\Lambda$  is similar to maximizing the correlation between the approximation of the individual treatment effect and its true value. The machine learning method maximizing this criterion can be considered as the most relevant to estimate coefficients in equation 1.5, reported in table 1.13(b). The machine learning method maximizing  $\bar{\Lambda}$  is the most relevant to estimate regression 1.6 whose results are reported in table 1.13(c). We follow the authors and combine both choices to present the two best methods. See Chernozhukov et al. (2018) for more details.

(b) ITT ( $\beta^{ITT}$ ) and heterogeneity ( $\beta^{HET}$ ) estimates

	Boosting		CF with LC	
	$\beta^{ITT}$	$\beta^{HET}$	$\beta^{ITT}$	$\beta^{HET}$
Estimate	-0.24	-0.05	-0.42	-0.03
Confidence interval (90%)	[-2.53, 2.04]	[-0.17, 0.09]	[-2.71, 1.88]	[-0.25, 0.21]
Adjusted p-value	(1.00)	(0.89)	(1.00)	(1.00)

**Notes:** Medians over 100 splits in half.  $\beta^{ITT}$  corresponds to the mean effect of being invited on the unemployment duration.  $\beta^{HET}$  is non-zero if heterogeneity has been detected. Following Chernozhukov et al. (2018), confidence intervals are adjusted and so are p-values.

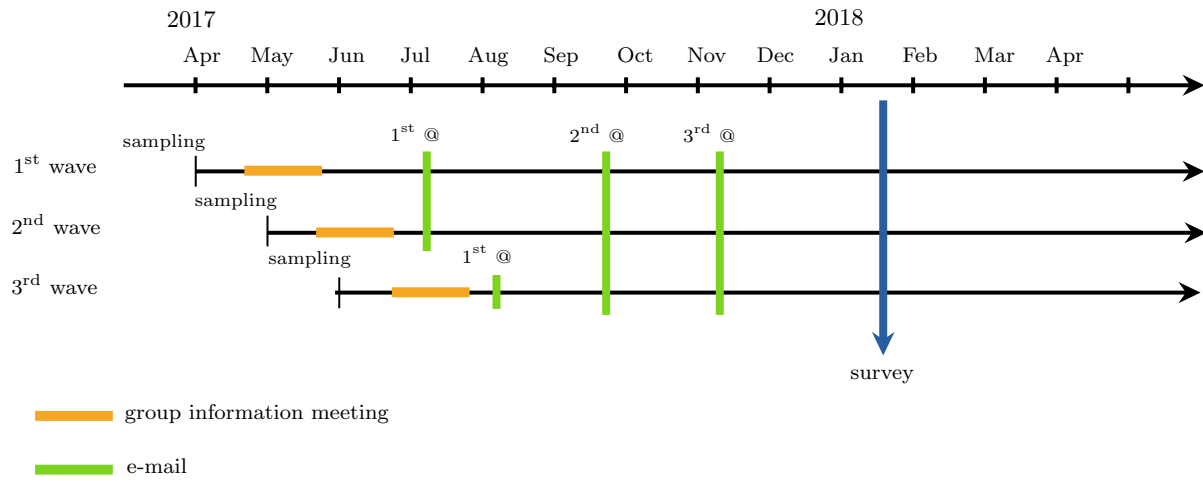
(c) GATES and difference between extreme estimated quintile treatment effects

	Boosting			CF with LC		
	Q5	Q1	Difference	Q5	Q1	Difference
Estimate	-1.86	0.02	-2.05	-0.63	0.35	-0.71
Confidence interval (90%)	[-6.98, 3.27]	[-5.10, 5.14]	[-9.29, 5.18]	[-5.77, 4.52]	[-4.78, 5.47]	[-7.98, 6.55]
Adjusted p-value	(0.86)	(1.00)	(0.47)	(1.00)	(1.00)	(0.64)

**Notes:** Medians over 100 splits in half. Quintiles Q5 and Q1 are estimations of  $\gamma_5$  and  $\gamma_1$  in regression 1.6. We also test for the difference between both. Following Chernozhukov et al. (2018), confidence intervals are adjusted and so are p-values.

## 1.8. Figures

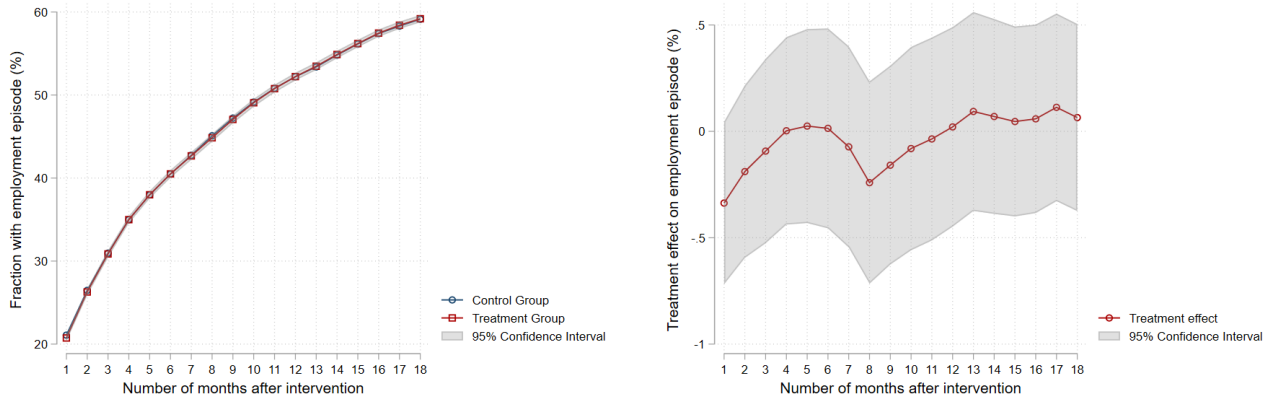
Figure 1.1: Timeline of the experiment



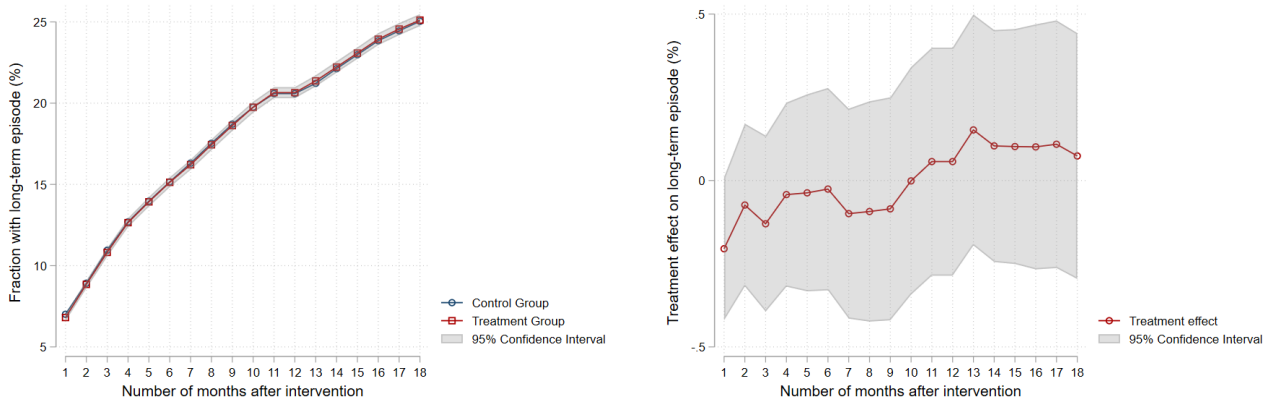
**Notes:** This figure shows the timeline of the experiment. The sample was drawn over three consecutive waves at the beginning of April, May and June. Information meetings for each wave started within the month of sampling and last one month. Three reminder emails were sent out in late July, late September and early November. The online survey was sent six months after the last wave of information meetings in January 2018.



Figure 1.2: Average treatment effects on employment episodes



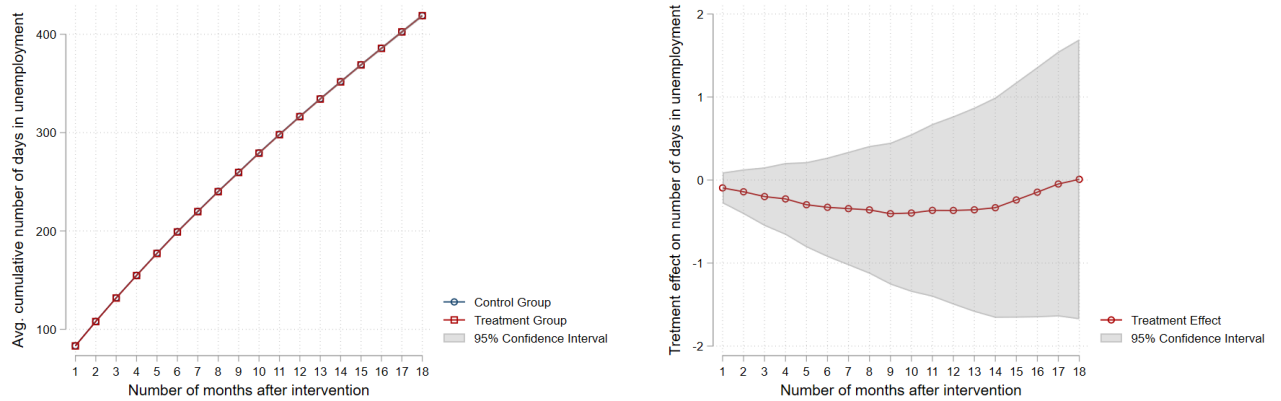
(a) Fraction with employment episode



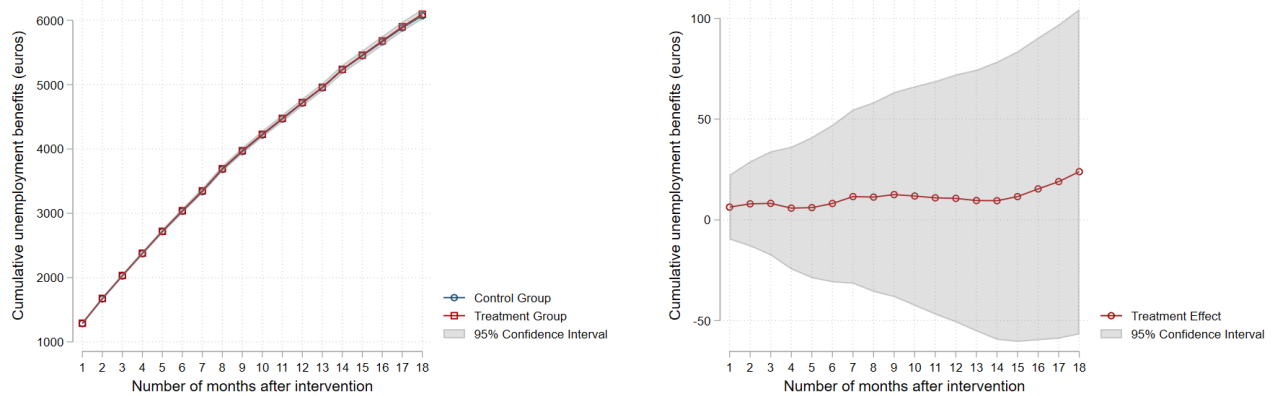
(b) Fraction with long-term employment episode

**Notes:** Figure (a) in the top panel shows the fraction of individuals who experienced an employment episode over the period. The left graph plots the evolution in levels over time, in blue and red for the control group and the treatment group, respectively. The right graph shows the ITT treatment effect as estimated by equation 1.1. Figure (b) in the bottom panel works similarly for the fraction of individuals with an employment episode corresponding to a long-term contract. Shaded grey areas correspond to symmetric 95% confidence intervals around values in the treatment group. For all graphs, the x-axis represents the number of months since the individual was included to the experimental sample.

Figure 1.3: Average treatment effects on cumulative number of days in unemployment and unemployment benefits



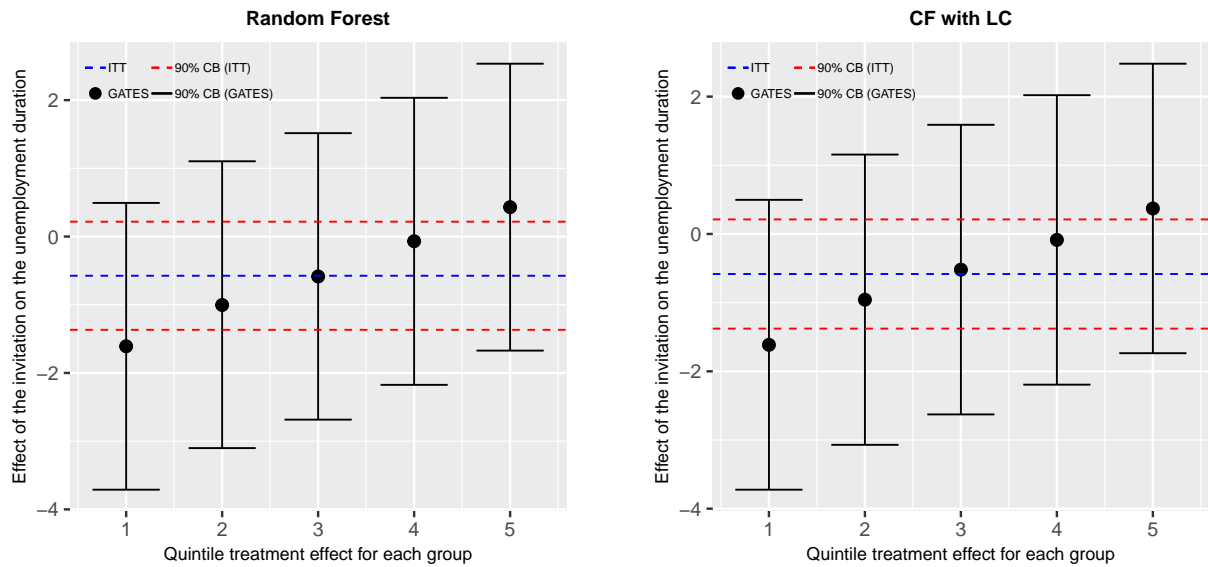
(a) Cumulative number of days in unemployment



(b) Cumulative unemployment benefits

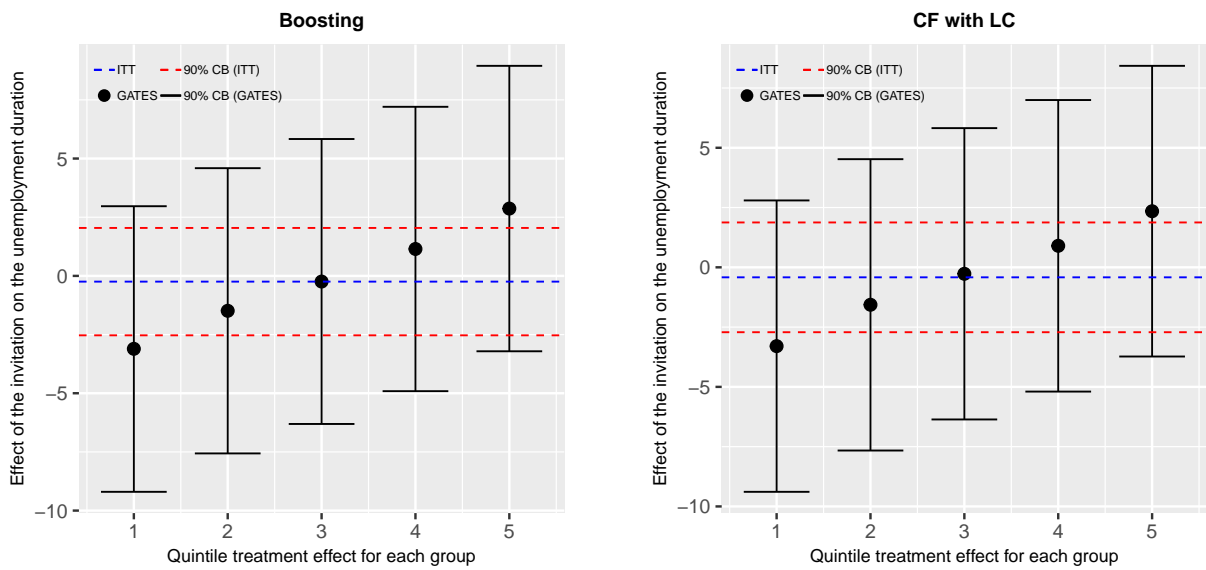
**Notes:** Figure (a) in the top panel shows the mean cumulative number of days in unemployment over the period of observation. The left graph plots the evolution in levels over time, in blue and red for the control group and the treatment group, respectively. The right graph shows the ITT treatment effect as estimated by equation 1.1. Figure (b) in the bottom panel works similarly for the mean cumulative unemployment benefits received over the period. Shaded grey areas correspond to symmetric 95% confidence intervals around values in the treatment group. For all graphs, the x-axis represents the number of months since the individual was included to the experimental sample.

Figure 1.4: GATES on unemployment durations 6 months after - generic machine learning approach



**Note:** Sample size = 212 277 individuals. Number of repetitions: 100. According to selection criteria in table 1.11(a), random forest (left) and causal forest (right) with local centering are the two best algorithms to approach heterogeneity. Thus, we report GATES estimates for these methods. Following Chernozhukov et al. (2018), GATES have been reordered after estimation, to make results more readable on the graph. Thus, since all GATES are closed to zero with relative large variance, estimates for quintiles Q1 and Q5 in table 1.11(c) do not coincide necessarily with estimates for Q1 and Q5 on the graph, due to reordering. The group heterogeneity score (x-axis) corresponds to an estimated quintile of the ITT. 90% confidence bounds computed following Chernozhukov et al. (2018) are reported.

Figure 1.5: GATES on unemployment durations 18 months after - generic machine learning approach



*Notes:* Sample size = 212 277 individuals. Number of repetitions: 100. According to selection criteria in table 1.13(a), boosting (left) and causal forest (right) with local centering are the two best algorithms to approach heterogeneity. Thus, we report GATES estimates for these methods. Following Chernozhukov et al. (2018), GATES have been reordered after estimation, to make results more readable on the graph. Thus, since all GATES are closed to zero with relative large variance, estimates for quintiles Q1 and Q5 in table 1.13(c) do not coincide necessarily with estimates for Q1 and Q5 on the graph, due to reordering. The group heterogeneity score (x-axis) corresponds to an estimated quintile of the ITT. 90% confidence bounds computed following Chernozhukov et al. (2018) are reported.

# Appendix

## 1.A. Survey questionnaire

In this section, we report the survey questionnaire. Note that the first question was directly included in the email that invited job seekers to respond to the survey, while the other questions were gathered on a dedicated webpage and accessible via a hyperlink in the email. The questions were adapted to the employment status of the respondent, and asked employed individuals to refer to their latest job search period. We do not report the exact format of each question (e.g. whether respondents could fill an empty box or move a cursor to pick a number).

[q1] **What is your current employment status?**

*Note : The “employment” category includes all types of contracts (short term/ seasonal (CDD), long-term/ salaried (CDI), temporary work and subsidized employment contracts).*

- I am unemployed and searching for a job
- I am employed but I am still searching for another job
- I am employed and I am not currently searching for a job
- I intend to create my own business or I am self-employed
- I am doing an internship or a job training
- Other answer (please specify)

*This survey takes on average 5 minutes. Your responses are kept anonymous and will never be shared with your caseworker. Your responses are very useful! They help us improve our support services!*

*The following questions are about your job search.*

[q2] **During a typical week, over 7 days, how many days are you active in your job search?**

[q3] **On an average day that you look for work, how much time do you spend on your job search?**

- Between 0 and 30 minutes
- Between 30 minutes and 1 hour
- Between 1h and 2h
- Between 2h and 4h
- More than 4h
- I would prefer not to say

[q4] **In your opinion, what is the most useful action to move forward in your job search?**

- Responding to job offers online
- Networking
- Sending unsolicited applications
- I would prefer not to say

[q5] **Over the past month, how many times have you been invited to interview for a job?**

[q6] **In general, you send unsolicited applications...**

- Multiple times per week
- Once per week
- 1-3 times per month
- Less than once a month
- I have never sent an unsolicited application
- I would prefer not to say

*The following 4 questions are about your global outlook on daily life.*

[q7] **On a scale of 1-10, where would you put your life right now?**

*Note: 0 indicates that you feel you are living the worst possible life and 10 indicates the best possible life for you. You may slide the cursor to select your response.*

[q8] **On a scale of 1-10, how motivated do you feel in your job search?**

*Note: 0 represents a total absence of motivation and 10 a very strong motivation. You may slide the cursor to select your response.*

[q9] **On a scale of 1-10, how much support do you feel you have during your job search?**

*Note: 0 indicates a total lack of support and 10 a robust support system. You may slide the cursor to select your response.*

[q10] **For the following sites, say if: you've never heard of it / you've heard of it / you've used it and find it useful / you've used it and don't find it useful**

- La Bonne Boîte
- L'Emploi Store
- Le Bon Coin
- LinkedIn
- Bob Emploi
- Jobi Joba
- La Bonne Formation

- CV Designer
- Jobeggs
- Pôle-emploi.fr

[q11] **Think back to your latest job applications. How frequently did you do the following? Systematically / Often / Sometimes / Rarely / Never**

- You added key words from the job offer to tailor your resumé
- You modified your cover letter to fit the job offer
- You followed-up with the recruiter a few days after sending your application

[q12] **Do you rely on the following persons to help you in your job search? Absolutely / More or less / Not really / Not at all**

- Friends and family
- Friends of friends
- Former classmates or college alumni
- Former coworkers
- Your Pôle emploi caseworker
- Volunteers from organizations that assist job seekers
- Local business owners
- Persons found online (via LinkedIn, etc.)

[q13] **What was your monthly salary (after taxes) from your latest job?**

[q14] **You are searching a job...**

- In your municipality and in neighboring municipalities
- Throughout your county
- Throughout your region
- Throughout the country
- I would prefer not to say

[q15] **What is your current monthly salary (after taxes)?**

## 1.B. Machine learning algorithms: options and parameters

### 1.B.1. List of covariates

- female
- nationality
- age
- age squared - `age_square`
- years spent working in the sector of interest - `exper_tirage`
- whether the job seeker looks for a part-time job or not - `temps_part`
- whether the individual lives in a sensitive urban zone or not - `zus`
- reason why the job seeker is on Pôle emploi's lists - `motif_inscription`
- job sector - `famille_prof`
- type of the desired contract (short or long term) - `contrat1`
- country subdivision - `departement_de`
- readiness to take up a new position - `catregr1`
- level of qualification - `qualif`
- type of accompaniment received at Pôle emploi agency - `dc_modalitesuiviaccomp_groups`
- a proxy of the accepted home-work distances - `distance_quantiles`
- the baseline benefits received - `tot_amount_groups`
- the reservation wage - `salmt1_groups`
- the number of children - `nenf1_groups`



## 1.B.2. Models

### 1.B.2.1. Elastic net with expansions

---

- R formula:

```
y ~ female * nationality * ( age + age_square + exper_tirage + temps_part + zus +  
  motif_inscription + famille_prof + contrat1 + departement_de + catregr1 +  
  qualif + dc_modalitesuiviaccomp_groups + distance_quantiles + tot_amount_  
  groups + nenf1_groups + salmt1_groups )
```

---

- Number of individual characteristics: 17 + squared age
  - Number of regressors including each interaction: 2 128
- 

- Rules to fit the algorithm - caret options

- Method: cross-validation
  - Pre-processing method: range
  - Optimality criteria: default - smallest Root Mean Square Error (RMSE)
  - Number of folds: 2
- 

- Grid of parameters

- Two parameters:
    - \*  $\alpha$  = mixing percentage between lasso and ridge: 5 values, default choice using grid search methods from the package
    - \*  $\lambda$  = regularization parameter: 5 values, default choice
    - \* Grid size: 25 distinct combinations of parameters
- 

- R Package: glmnet
-

### 1.B.2.2. Random forest

---

- R formula:

```
y ~ age + exper_tirage + female + temps_part + zus + nationality + motif_
  inscription + famille_prof + contrat1 + departement_de + catregr1 + qualif +
  dc_modalitesuiviaccp_groups + distance_quantiles + tot_amount_groups +
  nenf1_groups + salmt1_groups
```

---

- Number of individual characteristics: 17
- 

- Rules to fit the algorithm - caret options

- Method: cross-validation
  - Pre-processing method: range
  - Optimality criteria: default - smallest Root Mean Square Error (RMSE)
  - Number of folds: 2
- 

- Grid of parameters

- Two parameters:
    - \* Mtry: 5, 10 or 15
    - \* Minimal node size: 20, 30 or 40
    - \* Grid size: 9 distinct pairs of parameters
  - Splitting rule: variance reduction
- 

- R Package: ranger
-

### 1.B.2.3. Stochastic Gradient Boosting

---

- R formula:

```
y ~ age + exper_tirage + female + temps_part + zus + nationality + motif_
  inscription + famille_prof + contrat1 + departement_de + catregr1 + qualif +
  dc_modalitesuiviaccomp_groups + distance_quantiles + tot_amount_groups +
  nenf1_groups + salmt1_groups
```

---

- Number of individual characteristics: 17
- 

- Rules to fit the algorithm - caret options

- Method: cross-validation
  - Pre-processing method: range
  - Optimality criteria: default - smallest Root Mean Square Error (RMSE)
  - Number of folds: 2
- 

- Grid of parameters

- Two parameters:
    - \* Maximum depth of each tree: 5 values, default choice using grid search methods of the package
    - \* Number of trees 5 values, default choice
    - \* Grid size: 25 distinct combinations of parameters
  - Fixed parameters:
    - \* Shrinkage parameter applied to each tree in the expansion: fixed, = 0.1
    - \* Minimum number of observations in the terminal nodes: fixed, = 10
- 

- R Package: gbm
-

#### 1.B.2.4. Causal Forest with 500 trees

---

- R formula:

```
y ~ age + exper_tirage + female + temps_part + zus + nationality + motif_
  inscription + famille_prof + contrat1 + departement_de + catregr1 + qualif +
  dc_modalitesuiviaccomp_groups + distance_quantiles + tot_amount_groups +
  nenf1_groups + salmt1_groups
```

---

- Number of individual characteristics: 17
- 

- Detailed rules to fit the algorithm can be found here:

<https://grf-labs.github.io/grf/REFERENCE.html#parameter-tuning>.

- By default, 100 distinct sets of parameter values are drawn in the space of possible parameter values.
- 

- Tunable parameters

- Tree-growing options:

- \* `Mtry`<sup>37</sup>: number of variables tried for each split.
- \* A target for the minimum number of observations in each tree leaf.
- \* Fraction of the data used to build each tree

- Honesty behavior:

- \* The fraction of data that will be used for determining splits to get honest estimates
- \* Enable or disable pruning - two possible values

- Split balance parameters:

- \* A tuning parameter that controls the maximum imbalance of a split
  - \* A tuning parameter that controls how harshly imbalanced splits are penalized.
- 

- R Package: `grf`
- 

---

<sup>37</sup>Note that when using the `grf` package, all factor variables are turned to dummies. Thus, `mtry` can take value from 1 to 158 with only 17 different characteristics - including factor variables with several categories.

## Chapter 2

# Do informational frictions affect enrollment in public-sponsored training? Results from an online experiment

*This chapter is joint with Aïcha Ben Dhia*

Despite massive and increasing public spending in training for the unemployed, little is known about how job seekers decide to enroll in a training program. Decisions related to job training might be undermined by informational gaps, especially about program costs, enrollment procedures, and expectations of reemployment chances. This paper reports the results of a low-cost intervention aimed at testing for the existence of misinformation about training costs and returns, and its impact on enrollment. Partnering with the French Public Employment Services and the largest training provider in France, we sent 50,000 emails advertising training opportunities to job seekers in four regions of France in late summer 2016. We randomly added short messages on training costs, registration procedures, and training returns to the basic email template. We find that receiving an email with a message emphasizing training returns in terms of employment more than doubles the likelihood that job seekers call back the training center. However, callback rates are low in absolute value (less than one percent) and we detect no impact on enrollment one to six months after the intervention. We provide suggestive evidence that the effects on callbacks are driven by increasing salience of basic information about training rather than by belief updating. Overall, this suggests that public services need to invest in improving relevant knowledge among job seekers rather than relying exclusively on digital communication to modify behavior.

## 2.1. Introduction

Government-sponsored vocational training plays a leading role in public policies used to combat structural unemployment and to mitigate the negative employment effects of business cycle downturns (McCall et al. (2016)). In France, more than 4 billion euros of public expenditure are devoted annually to training for the unemployed.<sup>1</sup> To maximize the impact of these investments, policymakers target public funding towards sectors perceived as having high labor demand and towards job-seekers most likely to benefit from the program. However, the decision to participate in a training program ultimately takes place at the individual level and remains in the job seeker's hands. Information gaps regarding the pecuniary and non-pecuniary costs and returns from training may hinder the ability of job seekers to make optimal decisions. The efficiency of the whole job training system hence relies heavily on job seekers having access to information.

This paper presents the results of an experiment testing the effects of information frictions on job seekers' training demand by measuring the impact of online information provision on enrollment decisions. The experiment took place in late summer 2016, in the context of a large-scale public investment increase in vocational training targeted at the unemployed in France. The French government sought to nearly double the number of trainees, amounting to an additional 500,000 job seekers enrolled within a year. We partnered with *Pôle emploi*, the French Public Employment Service, and *Afpa* (Agence nationale pour la formation professionnelle des adultes), the largest training provider in France. We collaborated on a large-scale emailing campaign addressed to more than 50,000 job seekers, which advertised a list of 24 standard training programs in 4 regions of France, to boost enrollment. Emails were sent on August 30 and 31, 2016. Reminders were sent ten days after and programs started within the following three weeks. The experiment built on a similar campaign run earlier in the year by our partners using its target sample, operational schedule and email template. In its basic version, the email contained the list of programs offered in the region, and interested recipients were directed to the webpage with full program information if they clicked on the link provided in the email. The email also included a phone number to call the training center for additional information and to enroll.

Our intervention slightly varied the content of the messages that were sent out. We randomly sampled a *Control group* that received no email at all, and formed 5 different treatment groups. The *Basic email group* received a basic version of the email, allowing us to measure the impact of receiving an email on enrollment. This email contained the list of programs offered in the region, with hyperlinks directed to the webpage with full program information for each training program. The email also included a phone number to call the training center for additional information and to enroll. To test for the existence of specific information barriers, the four additional treatment groups received emails based on the same template but augmented with short sentences emphasizing different key information about training participation. All emails also included a hyperlink leading to a webpage with more detailed information. Job seekers in the *Cost email group* were reminded that training participation was entirely subsidized and would entitle participants to a stipend. In the *Simplicity email group*, additional sentences

---

<sup>1</sup>See [Annexe au Projet de Loi de Finances \(2018\)](#) (Appendix of the 2018 Draft Budget Bill).

emphasised the simplicity of registration procedures, stressing the availability of assistance from call operators. Messages received by the *Returns email group* provided job seekers with information about the potential returns from training: a short sentence mentioned the numerous job opportunities opened up by the training, and a hyperlink led to a webpage with rich metrics on wages and recruitment rate for the relevant jobs. The last *All info email group* received an email combining all three additional sentences.

We tested the impact of the intervention on two main outcomes: callback rates to *Afpa* and enrollment in a training program within the six months following the experiment. Along with these variables, we measured intermediate outcomes, including whether recipients opened the email and clicked on one of the links. This helps us shed some light on the degree to which recipients interacted with the information provided. Furthermore, three days before sending the emails, we sent a short baseline survey to the entire sample in order to capture prior beliefs about training cost and registration procedures.<sup>2</sup> We also asked respondents to estimate their expected wages and employment probabilities over a period of six months with and without training.

The survey reveals important information gaps about basic aspects of training costs and suggests that many respondents are uncertain or skeptical about training returns.<sup>3</sup> One third of respondents believe their unemployment benefits will decrease or get suspended if they participate in a training program, while nearly half of respondents believe training is not fully subsidized. 14% expect to pay more than 1,000 euros out of pocket. In addition, half of respondents perceive the registration procedure to be complicated or very complicated, which may act as a strong deterrent when considering whether to participate in a program. Finally, 26% do not expect training to increase their re-employment probability and up to 68% do not report any difference in expected wage with or without training. These results provide motivation for the intervention.

While the overall rates are low (around 0.5% overall), our results concerning the callback rates confirm the importance of information provision. As expected given the design of the campaign, all callbacks came from email recipients, who were significantly more likely to call back training centers over the month that followed the intervention than job seekers of the *Control group* who did not receive any email. Our modifications to the information content of the emails highlight important heterogeneity in the nature of information. Emails emphasizing training returns had the highest impact and almost tripled the callback rate compared to the group that received the basic email. More precisely, the callback rates of the *Return email group* and the *All info email group* increased by 0.4 and 0.36 percentage points from a mean of 0.27% in the *Basic email group*, respectively, and these increases are significant at the 1% level. Receiving an email on registration simplicity also increased callbacks by 70% compared to the level in the *Basic email group*, and the effect is significant at the 5% level. Perhaps surprisingly, given the results from the baseline survey, we detect no additional impact of emails with messages on cost compared to the basic email.

---

<sup>2</sup>The delay between the survey and the intervention was imposed by our partners' logistical constraints and prevented us from sending reminders to increase the response rate.

<sup>3</sup>The response rate for this survey is relatively low (13%). However, those who responded to the survey are on average more educated than the rest of the sample, they have more work experience and benefit less from assistance from Pôle emploi. It is thus plausible that such misinformation among the remaining job seekers might be even *more* pronounced.

Contrary to the results on callback rates, enrollment six months after the experiment in both our listed programs or any public-sponsored program was not affected by the intervention. This null effect is unlikely to be entirely due to mistargeting as enrollment in any training program six months after the intervention hovers around 6% in all groups, including the *Control group* who received no email. This indicates that training was an option that the targeted population of the campaign was considering.

As discussed in [Bleemer and Zafar \(2018\)](#), information interventions may have an impact through two main mechanisms: (1) by updating people’s beliefs, or (2) by making information more salient and acting as a reminder. Disentangling these two mechanisms is important as they have different policy implications. In the case of belief updating, the efficiency of interventions is determined by the precision with which uninformed individuals are targeted with tailored messages. On the contrary, if effects are mainly due to salience, no such targeting is needed as all individuals benefit from regular reminders.

Our data only allow us to provide suggestive evidence on these two mechanisms. We look for heterogeneous effects along individual observable characteristics that indicate individuals’ misbeliefs. Following [Bleemer and Zafar \(2018\)](#), the rationale of these tests is that if the impact of the emails is due to belief updating, it should mainly affect individuals with wrong beliefs. On the contrary, under the salience scenario, emails could have an effect irrespective of individuals’ initial beliefs. The highest impact would be obtained on individuals for whom the message is the most salient, that is, on individuals who pay most attention to their emails. Since the low response rate to the baseline survey prevents us from using baseline answers in heterogeneity analysis, we propose an alternative method leveraging callers in the *Basic email group* and survey respondents. The method relies on two assumptions. First we assume that individuals who call back in the *Basic email group* are the least misinformed and that additional messages convince marginally less informed job seekers to call back. Secondly, we assume that respondents to non-mandatory online surveys are individuals who pay most attention to their emails. In fact, responding to the baseline survey is highly correlated with opening the intervention email. This supports the interpretation that this variable is a sign of digital literacy and easiness to handle online communication with *Pôle emploi*. Under these assumptions, variables that correlate with callbacks in this group may be used as a proxy to identify misinformed individuals and responding to baseline may be used as a proxy for attention. We use these proxies in a standard heterogeneity analysis framework and we observe whether they increase or decrease the effect of the treatment. For example, in the updating scenario, characteristics of the callers in the *Basic email group* should decrease the effect of additional email messages, because additional messages convince less-informed job seekers.

We observe that callback rates in the *Basic email group* are correlated with having an educational degree higher than high school diploma (the *baccalauréat*), which generally corresponds to better informed individuals. When interacted with the treatment dummies, high education turns out to significantly reinforce the impact of receiving an additional message (it more than doubles the effect). The results are less consistent when we look at the impact of each additional message separately but the pattern is consistent with an increase of the effect of the messages on returns. As the education variable is correlated with many other individual characteristics, results should be taken with caution but suggest that the effects we observe are rather due to information salience among attentive readers rather than updating.



Had it been mainly updating, we would expect the coefficients of the interactions to be negative. Running the same regressions with a dummy for responding to baseline as a proxy for attentive job seekers reveals that baseline respondents systematically react more to emails and to each email message separately. The incremental effect on the sub-group of baseline respondents is even larger than on the high education sub-group: the impact of receiving an additional message and the impact of receiving a message on returns are tripled in all regressions, with all coefficients being significant at the 1% level.<sup>4</sup>

Overall, our findings suggest important information gaps that can deter job seekers from starting a training program. They reveal the existence of misinformation on very basic features of training programs and that marginal modifications of messages can affect at least some real-world behaviors, although the effects do not translate into increased enrollment in job training. These results offer several interesting takeaways from a policy perspective and for future research. Considering that baseline respondents are likely to be better informed than the average population, it encourages public services to improve on information systems, even to communicate simple basic messages. Our intervention also shows that message content matters, even when it is delivered in a very simple manner and that job seekers seem to be particularly sensitive to employment returns. Yet online messages alone do not have long-lasting effects on significant outcomes such as enrollment and they seem to work primarily on individuals that are most informed.

**Related literature.** While there is a growing literature on the determinants of enrollment in formal education, especially at the primary and secondary levels (Dynarski and Scott-Clayton (2008) Barr and Turner (2017), Abdulkadiroğlu et al. (2018)), much less is known about determinants of demand for vocational training (Barnow and Smith (2015)). As with other educational investments, participation decisions depend on individuals' beliefs about the pecuniary cost of the program and its expected returns in terms of future earnings and employment probabilities (Jacobson and Davis (2017)). Yet there are reasons to believe that these parameters are particularly hard to know in the context of job training. In France, as in other developed countries, public-sponsored training is a complex institutional system that involves many different stakeholders, including the public administration at both the national and regional levels, public employment services, and private training providers. This results in a highly diverse landscape of programs, funding opportunities, and training providers. Existing empirical evidence provides only estimates of average returns to large classes of training programs and highlights important heterogeneity across individuals and institutional settings (Card et al. (2017); Barnow and Smith (2015)).<sup>5</sup> Moreover, job seekers enter training programs in the course of their professional lives, at very different ages, in different labor markets, and with very different backgrounds. This vast heterogeneity can generate high levels of uncertainty for job seekers regarding the returns of different programs for their employment chances and take-home pay.

---

<sup>4</sup>To see whether the intervention durably affected beliefs, we also sent out an endline survey two months after the intervention. We got a low response rate as with the baseline survey, with a slightly unbalanced attrition across groups. Along with an additional sample size reduction, this prevented us from measuring any potential change on individuals' beliefs due to the intervention.

<sup>5</sup>This is one of the important take-aways of Card et al. (2017) meta-analysis of active labor market policies and Barnow and Smith (2015) review of U.S. programs. For example, Andersson et al. (2016) look at returns to two major public programs of vocational training in the United States. Despite the many similarities between both programs, they find moderately positive returns for one program but no significant returns for the second one. This is all the more puzzling as many job seekers are eligible for both streams.

As noted by [Barnow and Smith \(2015\)](#), despite the long tradition of evaluating training programs, there is only limited evidence on how information impacts training enrollment, with the notable exception of [Barr and Turner \(2017\)](#). Our paper starts to fill this gap. [Barr and Turner \(2017\)](#) find that US unemployment insurance beneficiaries are four percentage points more likely to enroll in a community college program upon receiving a letter with information on the costs and returns of these programs. The authors attribute this strikingly large effect, a 40% increase relative to the baseline enrollment rate, to the efficient complementarity of well-coordinated institutional support and endorsement from the White House.

An important puzzle that emerges from the existing training literature is unexplained returns heterogeneity, both across sites and participants (see e.g. [Andersson et al. \(2016\)](#), or [McCall et al. \(2016\)](#) for a review). [Jacobson and Davis \(2017\)](#) dig further in that direction by exploiting a particularly rich dataset in Florida allowing them to compare training returns by training program and participant socio-demographic characteristics. Their findings show that women select higher-returns fields and suggest that there is considerable room to increase their gains by altering their choice of field. Such informational barriers might slow down desirable re-allocation and a follow-up to this study could improve on information targeting leveraging similar individual-level data as in [Jacobson and Davis \(2017\)](#). In Germany, [Altmann et al. \(2018\)](#) run a similar experiment to ours, sending a brochure to a vast sample of job seekers informing them of observed returns to job strategies and consequences of unemployment. While the intervention has no significant average effect, they also find that it increases employment and earnings for a specific group of individuals, namely those at higher risk of long-term unemployment. For this group, the brochure increases employment and earnings in the year after the intervention by roughly 4%, which is remarkable considering the low cost of the intervention.

Recent works on the impact of information on education investments (e.g. [Dizon-Ross \(2019\)](#) and [Conlon \(2018\)](#)) showed in different contexts that changing parents and students' beliefs about educational outcomes could change investment decisions. Our study is closest to the experiment reported in [Bleemer and Zafar \(2018\)](#), where the authors provide information on cost and returns to college education in two separate treatment arms. Measuring the impact on intended enrollment decision, they find the cost intervention to have no effect, while information on returns significantly increased reported intentions to enroll and with lasting effects on beliefs two months after the experiment. These results align well with the findings of our study.

From a policy perspective, it is equally important to know whether such information interventions can be successfully implemented at scale and with low costs. Public services increasingly rely on digital communication tools, which provide a low-cost and highly scalable means to spread information, which can be continuously modified and individually tailored ([Kuhn and Skuterud \(2004b\)](#); [Kuhn and Mansour \(2014b\)](#); [Autor \(2009\)](#); [Horton \(2017\)](#)). However, digital communication also comes with limitations. Certain sub-groups of the population, especially among the unemployed, are not familiar with digital technologies and do not have access to online information. Online messages might also not be as convincing as a discussion with another individual. Several papers underline the need to combine online messages with offline assistance ([Castleman and Page \(2015\)](#); [Carrell and Sacerdote \(2013\)](#)). [Finkelstein](#)

and Notowidigdo (2018) confirm this complementarity in the context of SNAP enrollment, *i.e.* the food assistance program in the US. They find that delivering information on SNAP eligibility almost doubles enrollment but when complemented with assistance from public servants, information triples enrollment. Despite poor targeting properties of the interventions, their computations suggest that these interventions are a cost-effective policy. However, given the importance of job training, the complexity of the programs, and uncertainty about the impact of training, online communications have the potential to decrease misinformation and other obstacles that might limit the effectiveness of job training programs.

Finally our study relates to recent work on behavioral obstacles in job search (DellaVigna and Paserman (2005a); Babcock et al. (2012); Spinnewijn (2015); Caliendo et al. (2015); McGee (2015); DellaVigna et al. (2020)). Babcock et al. (2012) argue that complex institutional systems might deter job seekers from optimal decisions, notably in the context of training enrollment. Insofar as people are not perfectly rational, barriers in terms of financing cost or administrative hassles may be especially salient for the most vulnerable among the unemployed, resulting in an exacerbation of inequality in long-term outcomes (Bertrand et al. (2004), Schilbach et al. (2016)). Our paper adds some more evidence to this and suggests that reminders and repetition of simple information can help job seekers.

The rest of the paper is structured as follows. Section 2 gives background on how the French training system works. Section 3 explains the design of the experiment while section 4 provides an overview of our data. In section 5 we present our main results and we conclude in Section 6.

## 2.2. Background

Training to the unemployed is one of the main labor market policies in France. From 2014 to 2018, an average of 10% of all job seekers registered at *Pôle Emploi* participated in some form of training program, amounting to nearly 3 million trainees over the period. Pursuing this trend, in 2018, the French government launched a massive 5-year plan, channeling 15 billion euros into training towards uneducated youth and low-skilled job seekers. In this section, we briefly describe the French training system for the unemployed, and then provide more details on the context of the intervention and on our partners.

### 2.2.1. Public-sponsored training in France

The public-sponsored training system for the unemployed is managed by three main protagonists: administrative regions, the Public Employment Service (*Pôle emploi* hereafter), and the State. Esther <sup>6</sup> These three players jointly account for more than 80% of the nearly 5 billion euros going annually to fund training to the unemployed.<sup>7</sup> Thus they play a crucial role in the type of programs and sectors where job seekers can train. Although *Pôle emploi* aims primarily at quickly reducing unemployment

<sup>6</sup>Firms and other third-party institutions called “OPCA” (Organisme Paritaire Collecteur Agréé) are in charge of training for employed individuals and only play a minor role in job seekers’ training.

<sup>7</sup>Out of the 4.91 billions euros spent in 2015 on training for the unemployed, 1.47 billion were contributed for by regions, 1.94 billion by *Pôle emploi*, 82 millions were spent by firms, 37 millions by the State and 31 millions by beneficiaries themselves. This is reported in the See [Annexe au Projet de Loi de Finances \(2018\)](#) (Appendix of the 2018 Draft Budget Bill).

through short programs while regions fund longer training delivering professional qualifications, all three protagonists prioritize sectors with high labor demand in each region. The largest share of their subsidies are allocated by regions to group programs through a system of public auctions. This process allows them to set a number of requirements that training providers have to meet and to select providers that offer the best trade-off between price and program quality. The remaining share of their funding is available in the form of individual grants to fund individual training.

**Training costs.** When a program is funded by regions, *Pôle emploi* or the State, its direct cost is entirely covered. Additional grants may be given to cover transportation or housing. Only if either the desired program is not a sponsored group program or if the job seeker does not obtain an individual grant must she pay the entire cost of the program out of her pocket. By definition such programs are outside of the list of subsidized programs and they are very rarely advertised by *Pôle emploi* or proposed by caseworkers. In total, only 6% of all training programs are paid by beneficiaries.<sup>8</sup> Job seekers are formally not allowed to complement the maximum stipend they can get with their own money to pay for a program. Hence, by and large, participating to a public-sponsored training can be considered as free, aside from transportation and accommodation costs.

Upon enrollment, job seekers under unemployment insurance will keep the exact same amount of unemployment benefits. If their benefits exhaust before the end of the training, they get extended. Job seekers who are not eligible to unemployment benefits can also receive a special subsidy provided either by *Pôle emploi* or the State. Such subsidies vary across individuals but typically range between 300 and 500 euros per month. Enrolling in a training program can therefore only increase unemployment benefits.

**Enrollment procedures.** Enrollment processes vary across job seekers and largely depend on how they hear about the program, which is generally either on the Internet or by discussing with their *Pôle Emploi* caseworker.<sup>9</sup> If a job seeker wishes to enroll in a training program, her caseworker is asked to make sure that the training is consistent with her professional project and that she is committed enough to pursue training until the end.

If these conditions are met, caseworkers often help job seekers to look for available and financially supported programs within their sector of interest, prioritizing group programs that are subsidized by regions or *Pôle emploi*.

If job seekers wish to enroll in a group training, they have access to a limited number of providers, which are the ones that have been selected through public auctions. Because of the requirements set in the public auction, program contents (at least from what job seekers can read on brochures and learn on websites) tend to be fairly similar across selected providers. Nonetheless, there can remain important

---

<sup>8</sup>See [Annexe au Projet de Loi de Finances \(2018\)](#)

<sup>9</sup>When a job seeker first registers at *Pôle Emploi*, she gets assigned to a caseworker that will assist her throughout her job search. One important mission of the caseworker is to make sure that the job seeker meets the requirements to receive her unemployment benefits by attending mandatory workshops and actively pursuing her job search. They also help job seekers navigate the administrative system, e.g to participate to a job training program. However, caseworkers get assigned to a minimum of 150 job seekers at the same time, which limits the assistance they can provide. A caseworker who assists the most autonomous job seekers (who are most at ease with the Internet and other job search tools) follows from 300 to 600 individuals at the same time.

differences across training centers in terms of size, staff or educational methods. Given the absence of standardized and easily understandable systems of certification in France, it is quite complex to get a complete picture of the available training supply and to use reliable information to choose the best program and the most efficient provider.

Once they have obtained funding and identified a potential training center, job seekers attend an informational meeting at the training center that gives them more detailed information about the program. They might also have to take some selection tests at entry. To complete their enrollment, they finally need to get an enrollment form with signatures from the training center and their caseworker. This back-and-forth process was digitized in 2016, which considerably simplified the procedure. Registrations are now recorded on an online platform that both training providers and caseworkers can access in real time.

**Information available on training.** There is no centralized platform that aggregates all the information on training programs available to job seekers. Information is spread across different websites, often at the regional level. *Pôle emploi* hosts two websites that help job seekers look for training programs in their sector of interest and geographical area. Other institutions provide larger catalogs that are not limited to public-sponsored programs and also include training for workers.

Importantly, none of these websites include precise information on training returns. On their own websites, providers often post quantitative performance rates but those do not come from any rigorous evaluation. Information websites generally include only the dates and location of the program, along with a short description of the program content.

### 2.2.2. Context of the intervention

The experiment took place within the *Plan 500,000*, a national program to massively increase training participation among job seekers. The program had set the ambitious goal to increase by 500,000 the number of job seekers enrolled in a training program, corresponding to a 50% increase compared with previous years.<sup>10</sup> This inevitably required an important effort of recruitment and advertising and it was crucial for training providers and public services to communicate intensively about available programs.

Given the short timeline, public services focused on expanding existing supply rather than promoting new programs : they primarily funded programs that were similar to the ones funded before the *Plan 500,000* and collaborated with already existing providers. Thus, training programs advertised in emailing campaigns such as the one of this study were not different from standard programs. For these reasons, our experiment does not study programs with particularly high employment returns and low demand from job seekers. The context of the national plan, however, means that advertising campaigns were meant to recruit marginal job seekers, i.e. job seekers who would not have trained in absence of a national program. It also means that job seekers in all groups, including the *Control group*, were exposed to other, regular communication campaigns promoting training.

The campaign of this study was run jointly by regional offices of *Pôle emploi* and *Afpa*, the largest

---

<sup>10</sup>This target was reached, and the rate of job seekers enrolled in a training increased from 10% in 2015 to 15% in 2016.

training provider in France. *Afpa* has a unique status and history as a training provider in France: it was created in 1944 within the Ministry of Labor as a public institution in charge of professional training and it has played a central role in training job seekers since then. Although several recent reforms changed its status, to stimulate competition with other providers, *Afpa* has kept an important market share as well as massive equipment and numerous centers all over the territory with special connections with *Pôle emploi*. Not surprisingly, it participated actively in the *Plan 500,000* and ran several advertising campaigns to boost enrollment throughout the year 2016. The implementation of the campaign of this study followed the same procedure as earlier campaigns run by *Afpa* and *Pôle emploi*. In an emailing campaign launched in June 2016 in two regions, 37,000 emails were sent, resulting in 347 callbacks and 71 job seekers who agreed to register and participate to an information meeting.

## 2.3. Design

### 2.3.1. The campaign

The campaign was originally planned and designed by *Afpa* and *Pôle emploi*, as part of a larger advertising plan within the national training program *Plan 500,000*. It was targeted at four administrative regions.<sup>11</sup> In each region, a list of 5 to 7 programs were offered, which added up to 24 programs advertised in total. Overall, these are certifying training programs, with an average duration of 6 months, aimed at making persons involved in those programs directly operational for low and medium-skilled occupations. Messages were sent out on August 30 and 31, with reminder emails sent on September 9 and 10 2016, for programs that were starting within the first three weeks of September. This timing was decided by our partners.

As our partners had run a similar emailing campaign three months before the experiment, they chose to reuse the same email template. This basic email is showed in Appendix 2.9.1, Figure 2.9.1(b). It includes several motivational slogans about training and a short introduction sentence encouraging job seekers to enroll. It also lists the available programs in the region selected in the campaign.<sup>12</sup> Email recipients could click on one of the programs to open the *Afpa* webpage with more detailed information on the program and the jobs it may lead to. Finally, job seekers were provided with a phone number and were invited to call *Afpa* centers to get more detailed information and enroll.

### 2.3.2. The intervention

Our intervention consisted in introducing small variations to the basic email template. In collaboration with our partners, we designed three additional messages, all of which were not longer than a sentence or two, with a hyperlink to a webpage containing more detailed explanations.

---

<sup>11</sup>These regions are namely Auvergne-Rhône-Alpes, Centre, Hauts-de-France and Nouvelle-Aquitaine. They are geographically spread out over the French metropolitan territory and represent nearly one third of the French population.

<sup>12</sup>Each job seeker had been selected in the sample because she was searching in the same sector as one of the available programs. Displaying the entire list of available programs may have created some confusion, which in turn may have lowered the average response rate.

**Message on training costs.** A first modification consisted in adding a message on training costs, which was sent to the *Cost email group*. More specifically, we added a short sentence that reminded job seekers that training participation was fully subsidized and that they could be entitled to benefits. This is shown in Appendix 2.9.1, Figure 2.9.1(b). A hyperlink at the end of the sentence pointed to an external webpage hosted by *Pôle emploi* with more information about the type of benefits job seekers could be entitled to if they enrolled. The amount of the benefits could not be directly displayed in the email as it depended on each individual situation.

**Message on registration procedures.** In the emails sent to the *Simplicity email group*, we added to the basic template a sentence explaining that registration procedures had been simplified and that job seekers could get assistance from *Afpa* staff members. This is shown in Appendix 2.9.1, Figure 2.9.1(c). The additional webpage provided detailed explanations on the different steps to enroll.

**Message on training returns.** This third type of message, sent to the *Returns email group*, was meant to convey high training returns, primarily in terms of reemployment. As it was difficult to provide detailed statistical information in the email, the sentence simply said that training would “lead to many job opportunities”. In addition, job seekers could click to open a webpage from the *Pôle emploi* information website with several metrics including seasonal recruitment rates and average wages in the job as well as some indicators of market tightness based on *Pôle emploi* database. Wage and recruitment information were computed at the regional level and using observational data from employment administrative records. The email with information on returns is shown in Appendix 2.9.1, Figure 2.9.1(d). It illustrates an example for a program for a job of a network administrator. As visible in the figure, providing job-specific information required that only one training be displayed in the email. It is possible that this may have made the email easier to read and more impactful, independently from the information on returns.

**Message with all information.** Lastly, we gathered the three messages into a single email, to test for possible complementarities. This is shown in Appendix 2.9.1, Figure 2.9.1(e). If adding all three messages does not have crowd out the job seekers’ attention, it is certainly the most policy-relevant email as it addresses all types of information gaps.

### 2.3.3. Eligibility criteria and sampling

*Pôle emploi* and *Afpa* established eligibility criteria in order to target job seekers with potential interest in the listed training programs. Using comprehensive unemployment records in the four regions of the experiment, we first sampled job seekers who had agreed to receive advertising emails from *Pôle emploi*. We then restricted the list to job seekers seeking jobs in a professional sector that was related to one of the campaign programs. More precisely, two types of job seekers were selected. First, we sampled job seekers for whom one of the campaign programs matched their desired job. Of those, we only kept individuals who had reported less than 3 years of experience when they registered at *Pôle emploi* as



our partners considered that job seekers with more work experience would not be interested in getting trained.

We also sampled individuals who were seeking jobs in professional sectors that were close to the listed programs. As an example, a carpenter could be interested in getting some training in brick laying. This procedure was intended to help individuals complete their skill sets and expand the range of jobs they could apply to. Such individuals were selected only if they had reported more than 3 years of experience in their own professional sector. Finally, we removed from the list all job seekers who had ever participated in a training program since the beginning of their unemployment period. There were no specific criteria on the type of program and this may have left out many job seekers who had enrolled in a short job-search-related program and who would have been interested in a longer training with professional skill content. In total, 63,246 job seekers were sampled through this procedure, out of which 6.5% were looking for the same job as one of the listed programs.

#### 2.3.4. Experimental design and randomization

As can be seen in Figure 2.9.2, we randomly assigned all job seekers to one of six groups. The first group served as a *Control group*: these individuals received no email at all. A second group (*Basic email group*) received the basic email showed in Appendix 2.9.1, Figure 2.9.1(b). It was based on the template that our partners had used in their previous campaign, with the appropriate list of programs. We then formed four groups corresponding to our four different messages. The *Cost email group* received an email with the training cost message, while the *Simplicity email group* had an email with a message on training registration. Job seekers in the *Returns email group* received the email with a message on training returns. A last email combined all three messages (*All info email group*).

We created groups of equal sizes across treatment arms in each region. Due to logistical constraints and institutional differences across regions, we could not have all six groups in each region, which explains why treatment groups end up having different sizes. The distribution of the sample by treatment arm and region is summarized in Table 2.8.1, which shows that emails with information on training costs could only be sent in two regions.

We stratified the randomization at three levels: first, we split the sample by region as program lists were region-specific. Considering that listed programs were fairly heterogeneous, we also avoided imbalances across groups by stratifying the assignment by training program. The last strata was created based on whether job seekers were looking for the same job as the training.

We use the comparison between *Control group* that received no email to all other groups to study the effect of receiving an email, irrespective of its content. By comparing recipients of the basic email to job seekers in other email treatment groups, we can then identify which messages boost the impact of the basic email, thereby identifying potential information gaps.



## 2.4. Data and sample description

### 2.4.1. Data

This section provides a description of the data we use in the experiment.

#### 2.4.1.1. Email opening and click rates

*Pôle emploi* emailing software allows us to partially track job seekers' activity upon receiving an email. For each email recipient, we can see whether she opened the email, clicked on one of the hyperlinks in the email or if there was an error in the email address and the email bounced back.

#### 2.4.1.2. Callback data

In the emails, job seekers were invited to call back *Afpa* training centers to get more information about training programs and enroll. The phone number was specific to this emailing campaign. Call operators had to give some information about the programs, confirm job seekers' interest and invite them to participate to a first information meeting at the training center. At the end of the call, call operators had to indicate whether job seekers confirmed their interest in the program and whether they were available to this information meeting. Using names and first names we could match 269 names to our initial sample. <sup>13</sup> <sup>14</sup>

#### 2.4.1.3. Training enrollment

One of the key variables in our data records job seekers' enrollment into a training program after the intervention. *Afpa* training centers provided us with a list of job seekers who had enrolled in one of the listed programs one month after the intervention. As emails might have raised interest in training more generally and boosted participation in programs outside the campaign, we also leverage unemployment records to measure enrollment in *any* training program. We focus on enrollment one and six months after the intervention.

#### 2.4.1.4. Unemployment records

Finally, we use administrative data from *Pôle emploi*, as it provide individual socio-demographic characteristics such as age, gender, education level and family situation.

It also gather information on job seekers' past work history : this includes the professional category in their last job, and the number of months of experience in their professional sector of interest.

Administrative data give information about job seekers' current unemployment spells as well, such as the duration of unemployment, or the targeted jobs, which is a key variable for the experiment. Professional

---

<sup>13</sup>These data were manually recorded, with frequent typos, and many job seekers did not remember their unemployment ID, which is used as unique identifiers. This made the matching with our lists less precise and cumbersome. We manually corrected typos on names, first names, and unemployment IDs. We then tested the robustness of matching on names by comparing individuals' gender and region in both datasets and by using semi-automatic matching methods that did not rely on manual editing of typos.

<sup>14</sup>Only one observation belonged to the *Control group*. All other callbacks came from individuals in one of the email groups.

sectors are fairly narrow and correspond to 5-digit US occupational categories. They are organized in a hierarchical way with lexicographical classification and links to similar sectors. As training programs are also matched with those 5-digit professional sectors, we could sample job seekers who were searching in the same professional sector or in a closely related sector as the training programs advertised in the campaign. We then created a variable indicating whether the job seeker was searching in the exact same sector as the training or in a closely related one.

In addition, unemployment records contain information about the job seeker's *assistance track*, which corresponds to the level of assistance they need from the caseworker. It ranges from "low", which is the least intensive track concerning people who are familiar with digital tools and who are fairly autonomous in their job search, to "high" for job seekers with very little autonomy in their job search, who might not be comfortable using a computer.

Finally we have access to variables capturing job seekers' preferences, such as the type of contract the jobseekers is looking for, e.g. part-time work or short-term contracts.<sup>15</sup>

We use these administrative data to both sample job seekers as reported in subsection 2.3.3, and to describe the final sample, as these characteristics might be correlated with the impact of the intervention.

## 2.4.2. Sample description and balance checks

Table 3.8.3 provides key summary statistics of our sample and provides balance checks to assess whether the randomization was successful. Column 1 of Table 3.8.3 describes the sample along individual characteristics. The average age is 41, and about 60% of our sample are men. Just like jobs, training programs are highly segregated by gender and the gender imbalance in the sample is explained by the type of training programs in the campaign, which mainly attract men. At the time when we drew the sample, individuals had been unemployed on average for 13 months, although this average hides important dispersion that is typical of skewed duration distributions such as the ones of unemployment spells.

In line with *Pôle emploi* targets, most job seekers in the sample are low-skilled people, as indicated by the fact that 57% have less than a high-school diploma (baccalaureate). Yet 66% have a formal degree in their professional sector and individuals report an average of about 10 years of work experience in their desired jobs. Because of eligibility criteria (see previous section), only 6% are looking for a job in the same sector as one of the listed programs. Finally, only 13% of our sample benefit from intensive assistance from *Pôle emploi*, meaning that most individuals in our sample are considered to be fairly autonomous in their job search and familiar with online communication.

Panel II of Table 3.8.3 shows that only 56.4% of email recipients opened the email sent to them as part of our study. Columns 5 to 7 show how email openers differ from other email recipients. Not surprisingly, we see that they are more educated, as illustrated by a higher share with a formal degree in their job and an over-representation of employees and managers and educational levels higher than a

---

<sup>15</sup>Variables capturing job seekers' preferences need to be taken with caution as they are reported by job seekers only once at the time of registration and they rarely get cross-checked by caseworkers. Some of these variables such as desired wage or maximum acceptable distance to home are also recorded as point measures whereas one would want functions to describe indifference curves. Hence, in Table 3.8.3 we only keep two preference variables indicating if the job seeker said she was looking for part-time work or short-term contracts as we believe that these two variables are easier to interpret as stand-alone dummies and plausibly less likely to change with time as they tend to depend on family situation.

high school diploma. Women were significantly more likely to open their emails, but this is likely to be driven by a strong correlation between gender and education in the list of advertised programs.

Figure 2.9.3 gives details on the conversion rate of our emails into calls (*i.e* the rate of phone calls after receiving emails) through the study of clicks rates. It indicates that among job seekers who opened our emails, only 22.35% clicked on at least one training link, and only 2.63% of clickers called the training center. However, when considering people who called the training center, 65% of them had also clicked on one of the training link, which suggests that overall, training in the type of programs we promoted was an option that the callers were at least considering.

Table 2.8.3 shows that the randomization was successful at balancing groups along most individual characteristics. Column 1 displays the mean value of each characteristic, along with its standard deviations in brackets. In columns 2 to 6, we report the  $\beta$  coefficients of several regressions of the following type:

$$X_i = \alpha + \beta G_i^j + \epsilon_i.$$

In these regressions,  $X_i$  is an observable characteristic (e.g. female gender),  $\alpha$  is constant and  $G_i^j$  is a dummy for belonging to treatment group  $j$ . We run each regression only on individuals in the *Control group* and treatment group  $j$  with  $j$  ranging from 2 to 6, which means that the  $\beta$  coefficient is significant if and only if the *Control group* and the treatment group  $j$  are not balanced along  $X_i$ . We see in the table that only a few coefficients are significant, as should be statistically expected from the multiplicity of the tests we run on balanced groups.

### 2.4.3. Baseline survey

Three days before the intervention, we sent out a short online survey to measure existing misinformation about training. The complete questionnaire in both French and English is shown in Appendix 2.A. The survey was sent from *Pôle emploi* servers and had seven short questions related to the information gaps that the intervention targeted, asking the following:

- Question 1 asked job seekers whether they would have to pay to enroll in training programs “offered” by *Pôle emploi*. The wording of this question explicitly excluded supplemental costs such as transportation or housing, which are generally not subsidized.
- Question 2 asked whether enrolling in a training program has an impact on one’s unemployment benefits. By default unemployment benefits remain unchanged if a job seeker receiving benefits enrolls in a professional training program.<sup>16</sup>
- Question 3 asked about people’s perceptions of how easy it is to register. As described in section 2, the training system as a whole is hard to navigate. At the individual level, the main challenge is to identify a relevant training program, obtain funding and get one’s caseworker’s approval. However, once the program has been identified and validated, individual registration itself is fairly straightforward and often facilitated by caseworkers and training center staff members.

<sup>16</sup>Unemployment benefits can only get extended in case they are exhausted before training ends. This question only referred to the *amount* of benefits received.

- Questions 4 to 7 aimed at capturing people’s expectations of training returns. In questions 4 and 6, respondents had to estimate their re-employment probability within the following six months, with and without training, while in questions 5 and 7 they had to do the same exercise but for re-employment wages.

The response rate to the survey was 12.8%. While this response rate is low in absolute terms, it is fairly standard for such online surveys sent by *Pôle emploi*.<sup>17</sup> As can be seen in Panel I of Table 3.8.3, survey respondents are slightly older, more likely to be female, and significantly more educated than the rest of the sample (based on their highest formal degree and whether or not they have a degree in the job they search). They are also generally in the least intensive assistance track.<sup>18</sup> Interestingly, they seem to be selected along the same variables as people who opened the email (see Panel II of Table 3.8.3, discussed above). This evidence suggests that responding to baseline and opening emails do not depend on people’s intrinsic interest in the message but rather their internet fluency and how easily they communicate with *Pôle emploi* by email.

Despite this fairly advantageous selection in terms of education and other covariates, responses to the survey reveal some important information gaps. The left panel of figure 2.9.4(a) shows significant misinformation regarding direct training costs. About 45% of respondents to question 1 think training is not fully subsidized by *Pôle emploi*: almost 20% think that a 6-month training program “offered by *Pôle emploi*” will cost them up to 500 euros while another 20% estimate this cost to be higher than 1000 euros. Such priors about training costs must be an important barrier to enrollment, even for job seekers who believe training to be relevant for their professional skills.

The right panel of figure 2.9.4(a) shows further misinformation regarding how participation in the training impacts unemployment benefits. About 30% of respondents to question 2 think getting trained will modify their unemployment benefits. Among those, roughly 30% think their benefits will decrease and nearly 10% think they will entirely lose their benefits.

Turning to people’s subjective perception of administrative procedures, more than half of respondents to question 3 report that registering to a training program is complicated or very complicated. Figure 2.9.4(b) shows that up to 14% choose the latter option. Finally, Figures 2.9.4(c) and 2.9.4(d) illustrate respondents’ answers to questions 4 to 7 estimating people’s expectations about training returns, which show a more complicated picture.<sup>19</sup> The mode value of answers to questions 4 and 6 is at 50%, which may reveal people’s uncertainty about their baseline re-employment probability and future wage. It is nevertheless striking that 68% of respondents to questions 5 and 7 expect training to make zero difference in their future earnings. Respondents to questions 4 and 6 believe that training would increase their reemployment probability by 8 percentage points on average, but 26% of them expect no change at all and 18% think getting trained will instead lower their reemployment chances.

<sup>17</sup>This rate could have been pushed up had we been able to send reminders. However this was not possible due to the very short delay between the survey and the intervention itself.

<sup>18</sup>As explained in section 2.4.1.4, when they first register at *Pôle emploi*, job seekers get assigned to one of three assistance tracks that determines how closely their assigned caseworker will assist them, depending on their Internet fluency and how easily they handle their job search.

<sup>19</sup>Importantly, we do not know people’s past work histories and returns are likely to vary across individuals, all the more given that we did not specify the training program respondents had to imagine.

## 2.5. Results

### 2.5.1. Methodology

To estimate the impact of the intervention, we run several regressions. Equation (2.1) estimates the effect of receiving an email compared with no email:

$$Y_i = \beta E_i + \gamma' X_i + r_i + \epsilon_i \quad (2.1)$$

where  $Y_i$  is the outcome dummy (e.g. callback or enrollment) and  $E_i$  is a dummy for having received any email.<sup>20</sup> The coefficient of interest is  $\beta$ . In other words, equation (2.1) compares individuals in the *Control group* to the pooled sample of all other groups.  $X_i$  are individual covariates including gender, age, assistance track at *Pôle emploi* and education levels.  $r_i$  are region fixed effects accounting for the fact that regions did not have the same number of treatment groups and did not have the same listed programs.<sup>21</sup> Lastly  $\epsilon_i$  is a heteroskedastic error term.

To know if additional messages increased the impact of the basic email, we run a similar regression, simply separating individuals in the *Basic email group* from all other email groups. The corresponding regression is showed in equation (2.2). It is similar to (2.1) but we add a dummy  $M_i$  for having received any additional message in the email:

$$Y_i = \beta_1 E_i + \beta_2 M_i + \gamma' X_i + r_i + \epsilon_i \quad (2.2)$$

To compare email treatment groups and test which message is the most impactful, we remove the *Control group* and restrict the sample to emailed individuals only. The *Basic email group* is used as the reference group and we introduce one dummy per email group with additional message, as showed in equation (2.3) below:

$$Y_i = \beta^{cost} T_i^{cost} + \beta^{simp} T_i^{simp} + \beta^{ret} T_i^{ret} + \beta^{all} T_i^{all} + \gamma' X_i + r_i + \epsilon_i \quad (2.3)$$

where  $T_i^{cost}$ ,  $T_i^{simp}$ ,  $T_i^{ret}$   $T_i^{all}$  are dummies for messages on training cost, registration simplicity, training returns and email with all information, respectively.

Finally, we report the results of the same regression for the sample of individuals who not only received an email but also opened it. Email subjects were identical across treatment groups, so the inference is still valid. Focusing on people who open the email might increase power since we remove individuals who did not even open the email, which just add noise to the regressions. At the same time, excluding who did not open their email can reduce the precision of the estimation of the coefficients of covariates, which reduced statistical power.

---

<sup>20</sup>Note that there is no constant term as it would be colinear with the four constant fixed effects.

<sup>21</sup>If we focus on the *Basic email group* alone, we observe that callback rates are different across regions which confirms the relevance of region fixed effects beyond differences in the number of treatment arms.

## 2.5.2. Impact on callback rates

Our main outcome is whether job seekers called back *Afpa* training center. As it is the email's call-to-action and the first step to enroll, we interpret callback as evidence that the email raised interest in participating in a training program. Panel III of Table 3.8.3 shows that the average callback rate was overall very low, barely reaching half of a percentage point. This low number is of comparable magnitude to the previous campaign run in June 2016 by *Afpa* and *Pôle emploi*. Looking at individual characteristics, we observe that people who called back are significantly more educated than the rest of the sample - a selection bias that is twice stronger than for baseline respondents or email openers. Those who called back are also twice more often seeking jobs that directly match the employment opportunities of one of the campaign programs. Because of the eligibility criteria, they consequently have less professional experience.<sup>22</sup>

Table 2.8.4 shows the impact of the intervention on callback rates, using the regression specifications outlined above. Column (1) displays the results of regression (2.1). The mean in the *Control group* is virtually zero and confirms that all callbacks came from people who had received an email. Column (2) corroborates these results adding a set of covariates including gender, age, dummies for assistance and education levels. Columns (3) and (4) show that the effect is amplified by additional messages. On average, the callback rate upon receiving an email with any additional message is more than twice as large than with a basic email, which is the case with and without covariates.

Columns (5) to (8) show that the largest impact on callback is obtained by giving information on returns. Emails emphasizing returns (with and without the other messages on cost and simplicity) more than double the callback rate.<sup>23</sup> Messages on registration simplicity increase callback by 70%: in columns (5) and (6), it adds 0.19 percentage points to a mean of 0.27 in the *Basic email group*. However, although we found visible information gaps on training costs in the baseline survey, emphasizing that training was free and could entitle participants to a stipend did not appear to significantly increase callback.<sup>24</sup>

Results are similar in columns (7) and (8), which report regressions using the restricted sample of individuals who opened the email they received. Since email subjects were the same in all treatment groups, restricted treatment groups remain statistically balanced. Out of the 6,503 individuals who opened the basic email they received, we read that 0.44% called back the training center. This percentage is not significantly higher for individuals who received and opened an email on training cost. However it increases by 81% among people who received the email on registration simplicity and it is multiplied

---

<sup>22</sup>As explained in section 2.4, individuals looking for jobs that corresponded to one of the listed programs were eligible only if they had less than 3 years of experience, whereas individuals in close sectors were eligible if they had more than 3 years of experience.

<sup>23</sup>A possible explanation of the differential impact of the email with information on returns is that emails are easier to read because they display only the relevant training program for the job seeker instead of a list. Job seekers with low bandwidth might have identified their program of interest more quickly than among a list of 4 to 6 other programs (see figure 2.9.1). In order to capture this low bandwidth effect, we had agreed with our partners to add one additional group in the smallest region of the experiment. In this group, emails were identical to the basic email and did not contain any additional message, but they only displayed the most relevant training program for the job seeker instead of a list of 5 programs as in Figure 2.9.1(b). Had return emails generated an increase in callback simply by raising attention to the most relevant training program, we would in theory have observed a similar effect in the target email group. Unfortunately, the sample sizes are likely too small to detect any significant effect and conclude on this hypothesis.

<sup>24</sup>Even though this treatment was done in regions 1 and 2 only, the null effect does not seem to be due to power limitation. In fact, running the same regression in these two regions only, we see that the effects of other email treatments remain statistically significant.

by 2.5 for emails emphasizing training returns. These effects are of the similar magnitude in all regions even though regions had very different listed programs and callback rates upon receiving the basic email.

### 2.5.3. Enrollment

The final outcome variable is training enrollment after the intervention. This outcome is the most policy-relevant of our study but also the one that is most difficult to change in the short run. We measure enrollment using two different sources of data. A first dataset created by *Afpa* operators lists all individuals who enrolled in one of the listed programs after calling back the center. This dataset only contains 11 individuals, a strikingly low number given that 269 individuals had called back.<sup>25</sup> It is hard to know whether this bad performance is due to usual low turnout for the programs advertised in the experiment, or to the timing of the campaign, or for other reasons.

As most programs were to start only two weeks after the emails were sent, we could have missed individuals who needed more time to make their decision and finalize their enrollment. Hence we turn to the more comprehensive training dataset compiled by *Pôle emploi* to measure participation to *any* training program (in any *Afpa* center or in other training center). Emails could indeed have raised interest in training beyond the listed programs of the campaign. To allow for short- and long-term effects, we measure enrollment one month and six months after the intervention.

Panels A and B of Table 2.8.5 show that our intervention had no visible impact on training enrollment after one and six months. Receiving any email (compared to not receiving an email) had no effect on training participation, nor did any of the additional messages taken separately.

To further investigate the reasons for this null effect, we turn to Panel IV of Table 3.8.3, which shows the distribution of individual characteristics among job seekers who participated in a training within the six months that followed the intervention. We see that more than 6% of the sample enrolled in a training program, with two thirds of them participating to a program that was longer than two weeks. To correctly interpret this number, one must remember that the experiment took place in the middle of a vast national program to boost training participation among job seekers. Therefore, even job seekers in the *Control group* were exposed to multiple information campaigns promoting training.

The selection pattern of trainees in our sample, that we show in panel IV of Table 3.8.3, looks somewhat different than the one for email openers in panel II. Although both trainees and email readers are positively selected on education levels, trainees are more often male, slightly younger than non-trainees and in more intensive assistance tracks, with about a year less experience in the job they search. This selection is absent or reversed if we look at openers, who are more often female and slightly older than non-openers. These differences in selection may suggest that emailing does not reach the job seekers who are most likely to be interested in training. This mis-targeting might have contributed to the low impact of the intervention.

Finally, as our interventions could have increased enrollment to programs in *Afpa* centers beyond the listed programs, possibly at the expense of other centers. Unemployment records contain a dummy

---

<sup>25</sup>However, this number is in line with previous performances of *Afpa* campaigns. In a previous campaign run by *Afpa* and *Pôle emploi* three months before our intervention, 71 job seekers had enrolled out of an initial sample of 37 000 email recipients.



variable that is meant to indicate whether job seekers who participated in a training enrolled in an *Afpa* center. We use it to test whether the fraction of enrolled job seekers in an *Afpa* center increased upon receiving emails. Looking at the fraction of enrolled job seekers is valid from a statistical inference point of view, as the intervention had no effect on the total number of job seekers enrolled. Results are showed in Table 2.8.6. We see no robust impact of any treatment on this outcome.

## 2.6. Heterogeneity

In the final part of our paper, we now explore whether the treatment had heterogeneous effects to shed some light on whether the impact on callbacks was due to increasing salience or to belief updating.<sup>26</sup> Following [Bleemer and Zafar \(2018\)](#), in a scenario where the main mechanism driving the effects on callback is information updating, then additional messages in the treatment groups should have encouraged marginally misinformed or less-informed job seekers to call back.

A second possible mechanism that is considered in [Bleemer and Zafar \(2018\)](#) is that messages increase information salience. If additional messages work through updating, we should see a negative correlation between these variables and the treatment effects. In the salience scenario and with no specific assumption on the joint distribution of attention and beliefs, we do not predict any particular correlation between treatment effect and initial beliefs. Nevertheless, if attention is also a driving mechanism that explains why certain individuals call back upon receiving the basic email, then it would come as no surprise that similarly attentive recipients are also more sensitive to additional messages in the treatment groups and we would detect a positive correlation between treatment effects and variables that characterize callers in the *Basic email group*.

To apply this test, we could in principle use the answers from the baseline survey to identify less-informed job seekers. Unfortunately the low response rate to the survey only allows us to run regressions on a fifth of the initial sample. Using this reduced sample size does not ensure sufficient power to detect any heterogeneity pattern. We propose an alternative method. In the updating scenario, those who called back in the *Basic email group* are individuals who already had accurate information whereas additional callers in the other treatment groups were less informed. We can hence use variables that characterize callers in the *Basic email group* to identify well-informed job seekers.

Following this reasoning, we first characterize individuals who called back in the *Basic email group*. We do it in a similar fashion as we do for the whole sample in panel III of Table 3.8.3. We find that people who called back in the *Basic email group* are more educated and in less intensive assistance tracks, as is showed in Table 2.8.7. We interpret these two patterns as indicating a high education level. Hence we group the two first education levels to create a dummy for having at least a high school degree and we run the following regression:

$$Y_i = \beta_1 Z_i + \beta_2 T_i + \beta_3 T_i \times Z_i + \gamma'_1 X_i + \gamma'_2 X_i \times Z_i + r_i^1 + r_i^2 \times Z_i + \epsilon_i \quad (2.4)$$

<sup>26</sup>Given that we could not observe any average impact on enrollment, we show heterogeneity tests for callback only. Similar tests on enrollment variables show no heterogeneous patterns.



In this equation,  $Z_i$  is a dummy for having an education level above the *baccalauréat*,  $T_i$  is any treatment dummy,  $X_i$  are the same covariates as in previous regressions and  $r_i^j$  are region fixed effects. The coefficient of interest is  $\beta_3$ , which will be positive if treatment effects are higher on people with high  $Z_i$ .

To test for heterogeneous effects of each additional message, we adapt equation 2.4 and look at coefficients on interactions between  $Z_i$  and four treatment dummies:<sup>27</sup>

$$Y_i = \beta_1 Z_i + \sum_j \beta_2^j T_i^j + \sum_j \beta_3^j T_i^j \times Z_i + \gamma_1' X_i + \gamma_2' X_i \times Z_i + r_i^1 + r_i^2 \times Z_i + \epsilon_i \quad (2.5)$$

The results presented in Table 2.8.8 confirm that high education is associated with a higher impact of receiving an email on callback. This can be seen by looking at columns (1) and (2) of Table 2.8.8. Having a high school diploma almost doubles the impact of calling back after receiving an email. The effect of receiving an email remains significant for those with low education (the coefficient on the simple dummy of receiving an email is positive and significant). This result shows that the average impact found in Table 2.8.4 is not entirely driven by those with higher education. The same pattern remains when we split the *Basic email group* and additional message groups, as showed in columns (3) and (4).

When we look at each treatment separately, as we do in the four last columns of Table 2.8.8, high education appears to significantly increase the impact of both email treatments with information on returns. The interaction with the *Cost email group* also turns positive and significant in columns (5) and (7) although the effect does not remain once we add additional covariates. Not all these effects remain significant in the restricted sample of openers only.

A possible interpretation for such pattern is that educated people are more familiar with emails and internet communication, and thus more likely to react to interventions that are sent by email. Attention in this context might be strongly correlated with digital literacy. To further explore this hypothesis, we run the same regressions from equations (2.4) and (2.5), this time interacting the treatments with a dummy for responding to baseline. Answering to online surveys is indeed correlated with how familiar individuals are with online communication in professional contexts and formal institutions like *Pôle emploi*.<sup>28</sup> Results from these regressions are remarkably consistent and strong. Columns (1) and (2) of Table 2.8.9 show that responding to baseline predicts an effect of receiving an email almost three times larger than for non-respondents. It also largely improves the efficiency of each separate message, especially for the email treatment with information on returns only.

These results should naturally be taken with caution. In these simple heterogeneity tests, variables that are interacted with treatment dummies are correlated with many other individual characteristics. It is therefore impossible to rigorously identify one main driving factor. However, available evidence suggest that the effects on callbacks are due to a salience effect benefiting those who are most familiar

<sup>27</sup>Equation 2.4 has twice more coefficients to estimate than equation 2.1 because of interacted terms. This might prevent us from detecting an effect on the variable of interest. As a robustness check, we also run the same regressions dropping the *Cost email group*. This allows us to remove region fixed effects along with their interacted terms as all regions have then the same number of treatment groups. Results are showed in Table 2.B.1 and 2.B.1 of appendix section 2.B. We find very similar results are to those with region fixed effects.

<sup>28</sup>Answering to the baseline survey may also have raised individuals' attention to their emails, especially when they related to training programs and independently of pre-existing digital literacy. We cannot rule out this interpretation.

with digital communication.

## 2.7. Discussion and conclusion

We provided with a low-cost intervention embedded in an advertising campaign for public-sponsored training programs. A baseline survey suggests that there exist important information gaps on training that might affect job seekers' enrollment. This might not come as a surprise if we consider the complexity of the training system: the high diversity of programs as well as participants' and providers' heterogeneity make it almost impossible for any job seeker to gather all the information she might need to make an optimal decision. The existing literature itself remains puzzled by the persistent heterogeneity of training effects across participants and institutional settings. Nevertheless, this study focused on arguably simple features of training participation which one would assume to be common knowledge among job seekers. Yet even for such basic information, the baseline survey reveals that a significant fraction of job seekers hold incorrect beliefs. Taken at face value, these biased beliefs would be sufficient to strongly deter individuals from enrolling. An important question for future research is to better characterize those who are misinformed and exploit this information to design targeted intervention.

This work also shows that very simple messages can modify people's behaviors. Treatments only consisted in adding one sentence and a hyperlink to standard emails. Such light modifications are virtually costless and prove that details can make a difference. As these email campaigns are daily routine for public employment services, such marginal and cheap improvements can help to significantly raise communication efficiency.

Yet, the intervention did not have any measurable impact on actual enrollment in training programs. While we face some statistical power constraints, we can rule out any large effect on training participation and the effects we obtain on callback are also very low in absolute value. As suggested by other studies on college enrollment (e.g. [Carrell and Sacerdote \(2017\)](#)), a fruitful avenue for future research and efficient policy could be to mix such online interventions with offline assistance and better targeting. The importance of caseworkers throughout the enrollment process suggests that information interventions can have a stronger impact if they are also targeted at *Pôle emploi* caseworkers themselves.

Overall, this study could be a first step to better understand the determinants of training participation. The inexpensive and policy-grounded aspect of our experiment makes it very easy to replicate, improve upon, and scale. More research could be undertaken to confirm the robustness of our results, by testing similar interventions on different samples, at different timings and advertising more specifically training programs with low demand and high returns. Similar messages could also be spread out through different communication channels to reach out to other types of job seekers. Finally, this study offers an interesting example of collaboration between researchers and administrative services, making results directly policy-relevant as both the setting and the methodology are grounded in existing practices.

## 2.8. Tables

Table 2.8.1: Distribution of job seekers across regions and treatment arms

Region	Control group	Basic email group	Cost email group	Simplicity email group	Return email group	All info email group
Auvergne-Rhône-Alpes	2,870	2,871	2,870	2,870	2,870	2,870
Centre	1,684	1,685	1,685	1,685	1,685	1,685
Hauts de France	2,078	2,079	-	2,079	2,079	2,079
Nouvelle Aquitaine	5,104	5,105	-	5,105	5,104	5,104

*Notes:* This table shows the distribution of job seekers across treatment groups and regions. One can check that the sample size is similar across groups within each region. For administrative reasons, there was no *Cost email group* in Hauts de France and Nouvelle Aquitaine.

Table 2.8.2: Summary Statistics

	All	I. Baseline respondents			II. Opened the email			III. Called Afpa center			IV. Enrolled in a training		
	1	Non-resp 2	Resp 3	[3 - 2] 4	Not open 5	Open 6	[6 - 5] 7	No call 8	Call 9	[9 - 8] 10	No train 11	Train 12	[12 - 11] 13
Female (%)	38.0	35.9	52.3	8.17***	34.7	41.4	6.77***	38.5	36.4	-2.09	38.4	32.8	-5.65***
Age	41.8	41.3	45.3	2.00***	41.3	42.3	1.06***	41.9	43.4	1.49**	41.9	40.5	-1.43***
Foreigner (%)	8.5	9.0	5.5	-1.75***	9.0	8.1	-0.84***	8.5	17.8	9.30***	8.5	8.6	0.07
Married (%)	50.6	49.8	55.9	3.05***	48.4	52.3	3.93***	50.6	46.3	-4.32	50.5	50.8	0.22
Number of children	1.0	1.0	1.0	-0.02***	1.0	1.0	-0.03**	1.0	1.0	-0.03	1.0	1.0	0.03*
Duration (months, capped at 18)	13.0	13.0	12.8	-0.13***	13.3	12.7	-0.60***	13.0	13.0	-0.02	13.1	11.5	-1.61***
Looking for short-term contract (%)	5.5	5.4	5.6	0.09	5.5	5.4	-0.10	5.5	5.4	-0.09	5.7	2.6	-3.05***
Looking for part-time work (%)	9.7	9.2	12.8	1.79***	9.1	10.5	1.42***	9.9	9.9	0.05	10.1	3.9	-6.18***
Formal training in desired job (%)	65.5	64.5	72.6	4.05***	61.9	68.2	6.36***	65.4	63.9	-1.56	65.3	69.3	4.05***
Programs match with desired job (%)	6.6	6.6	7.0	0.20	6.4	6.8	0.45**	6.6	12.3	5.71***	6.5	8.7	2.22***
Experience in desired job (months)	128.5	125.4	149.5	12.01***	123.9	132.2	8.25***	128.6	128.1	-0.57	129.3	117.6	-11.71***
<i>Assistance track</i>													
Low (%)	41.2	40.9	43.3	1.18***	41.3	41.6	0.31	41.5	37.7	-3.77	42.1	28.7	-13.43***
Moderate (%)	43.8	44.0	42.6	-0.65**	44.0	43.4	-0.57	43.6	44.8	1.13	43.6	46.6	3.00***
Intensive (%)	13.3	13.4	12.5	-0.45**	13.0	13.3	0.34	13.2	15.3	2.12	12.7	21.8	9.05***
<i>Education level</i>													
No high school nor vocational degree (%)	12.8	13.4	8.4	-2.52***	15.1	11.3	-3.74***	13.0	5.6	-7.39***	13.1	8.1	-5.00***
Vocational degree (%)	43.3	44.7	33.8	-5.45***	47.4	39.5	-7.84***	43.0	30.6	-12.41***	43.4	41.8	-1.59**
High school diploma or GED (%)	23.5	22.7	29.0	3.18***	21.0	25.5	4.56***	23.5	34.3	10.84***	23.3	26.5	3.27***
Bachelor degree or more (%)	18.3	17.2	25.5	4.13***	15.0	21.0	6.07***	18.4	22.4	4.02	18.1	20.8	2.72***
<i>Professional status</i>													
Unskilled worker (%)	2.7	2.9	1.2	-0.87***	3.4	2.1	-1.32***	2.6	3.7	1.09	2.7	2.5	-0.17
Skilled worker (%)	32.6	34.4	20.6	-6.90***	37.4	28.7	-8.69***	32.5	24.3	-8.29***	32.7	32.4	-0.27
Employee (%)	55.1	53.4	66.5	6.55***	50.5	58.9	8.35***	55.2	60.1	4.84	54.7	60.8	6.06***
Manager (%)	3.2	2.8	5.8	1.50***	2.5	3.8	1.29***	3.2	1.9	-1.35	3.2	3.2	0.02
N =	63246	55175	8071		22362	29049		51242	268		59136	4110	
		87.2%	12.8%		43.5%	56.5%		99.5%	0.5%		93.5%	6.5%	

*Notes:* This table presents baseline summary statistics for outcome and control variables used in the main regression tables and the appendix tables, as well as other background variables mentioned in the paper. Column 1 displays the variable means in the whole sample. Panel I compares individuals who did not respond to baseline (column 2) to those who did (column 3). Column 4 shows the coefficient we obtain by regressing a response dummy on the covariate of the row. Stars reflect the significance of the coefficient with robust standard errors. Panel II works similarly, comparing individuals who did not open the email they received (column 5) to those who did. This comparison is done among individuals who received an email, that is in all groups but the control group. Panel III compares individuals who called back Afpa and those who did not among people who received an email. Panel IV compares individuals who did and did not enroll in a training within the 6 months that followed the experiment in the whole sample. Variables are extracted from unemployment records (see appendix for table references). Formal degree in desired job means that the job seeker has a formal educational degree in the job he is looking for. Email training in desired job means that the training that is advertised in the email that the job seeker receives leads to the same job as the one he is looking for. Past experience in the same job refers to job seekers who worked in the job that the offered training leads to. The two last rows refer to sample sizes and their percentage as a share of the relevant group of comparison (whole sample for panel I and IV, sub-sample of individuals who received an email for panel II and III). \*\*\* p-value < 0.01, \*\* p-value < 0.05, \* p-value < 0.1.

Table 2.8.3: Balance table

	(1)		(2)	(3)	(4)	(5)	(6)
	Control (C)		Basic	Cost	Simplicity	Returns	All info
	Mean	[S.D.]	C = Basic	C = Cost	C = Simp	C = Ret	C = All
Female (%)	36.26	[ 48.08 ]	0.70	0.339	0.504	0.179	0.319
Age	41.62	[ 10.85 ]	0.20	0.084	0.192	0.167	0.171
Foreigner (%)	8.46	[ 27.83 ]	0.49	-0.138	-0.173	0.041	0.159
Married (%)	50.46	[ 50.00 ]	-0.14	0.428	-0.294	0.128	-0.247
Number of children	1.01	[ 1.23 ]	-0.01	-0.027	-0.017	-0.005	0.001
Duration (months, capped at 18)	13.06	[ 5.61 ]	0.00	-0.150	-0.089	-0.080	-0.020
Looking for short-term contract (%)	5.51	[ 22.81 ]	-0.44	0.014	-0.093	0.019	0.166
Looking for part-time work (%)	8.91	[ 28.48 ]	0.51	1.373*	0.600	0.373	0.304
Formal training in desired job (%)	66.08	[ 47.35 ]	-0.46	-0.624	-0.758	-1.769***	-0.269
Programs match with desired job (%)	6.49	[ 24.64 ]	0.08	-0.046	0.058	-0.018	-0.001
Experience in desired job (months)	127.96	[ 100.62 ]	0.38	2.286	0.760	0.194	1.238
<i>Assistance track</i>							
Low (%)	40.27	[ 49.05 ]	0.28	1.769*	0.961	1.084*	1.041
Moderate (%)	44.37	[ 49.68 ]	0.10	-1.129	-0.752	-0.928	-0.076
Intensive (%)	13.77	[ 34.46 ]	-0.44	-0.837	-0.395	-0.266	-1.025**
<i>Education level</i>							
No high school nor vocational degree (%)	12.18	[ 32.71 ]	1.17***	-0.354	0.729*	1.370***	0.458
Vocational degree (%)	44.94	[ 49.75 ]	-1.63**	-0.710	-0.719	-1.132*	-0.919
High school diploma or GED (%)	23.25	[ 42.25 ]	0.37	-1.038	-0.049	-0.651	-0.464
Bachelor degree or more (%)	17.70	[ 38.17 ]	0.10	2.147**	0.055	0.431	0.943*
N =	11736		11740	4555	11739	11738	11738

*Notes:* This table shows balance tests across treatment arms to check that the randomization was successful at creating statistically comparable groups. The first two columns show variable means in the control group that received no email, with standard deviations in brackets. Column (2) shows the coefficients of regressions testing the effect on each variable of belonging to the basic email group compared to the control group. Columns (3) to (6) proceed similarly for each treatment group. We use robust standard errors for all regressions and three stars indicate a p-value < 0.01; two stars indicate a p-value < 0.05; one star indicates a p-value < 0.1. We observe that the randomization was successful at balancing groups along observable characteristics. A few significant and small differences emerge, as is expected from such statistical procedure. The last row of the table shows the sample size in each treatment group. The cost group is smaller as it was only implemented in region Auvergne-Rhone-Alpes and Centre.

Table 2.8.4: Impact on callback

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed only	Emailed only	Openers only	Openers only
Received any email	.496*** (.0321)	.45*** (.0322)						
Received basic email			.264*** (.0489)	.238*** (.0486)				
Received basic email and message			.565*** (.0387)	.514*** (.0388)				
- <i>Cost</i>					.0563 (.132)	-.035 (.127)	.132 (.218)	-.0556 (.212)
- <i>Simplicity</i>					.196** (.0793)	.194** (.0795)	.359*** (.138)	.305** (.141)
- <i>Returns</i>					.409*** (.0899)	.352*** (.0879)	.699*** (.154)	.546*** (.153)
- <i>All info</i>					.366*** (.0878)	.371*** (.089)	.572*** (.148)	.547*** (.154)
Mean in the control group	.0085	.0085	.0085	.0085	-	-	-	-
Mean in the basic email group	-	-	-	-	.2726	.2726	.4459	.4459
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	58314	63246	58314	51510	47498	29049	26644

*Notes:* This table shows the effect of receiving emails on calling back *Afpa* center. All regressions use a callback dummy as their outcome. In column (1) we group all individuals who received an email and compare them to those who received no email (the control group), as per equation (2.1). Column (2) adds to this regression a set of covariates including gender, age, assistance intensity at Pôle emploi and educational levels as covariates. Column (3) splits emailed individuals into two groups: the first explaining variable is a dummy for being in the basic email group and the second is a dummy for all other email groups, as per equation (2.2). Column (4) adds covariates. In column (5), we remove the control group and regress callback on five separate dummies for each email treatment group, using the basic email group as the reference group as per equation (2.3). Column (6) adds covariates. Finally columns (7) and (8) display the results of the same regression as (5) and (6) on a restricted sample with only individuals who opened the email they received. All regressions include region fixed effects. Means in the reference groups are computed separately. Standard errors are in parenthesis: \*\*\* p-value < 0.01, \*\* p-value < 0.05, \* p-value < 0.1.

Table 2.8.5: Impact on enrollment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed Only	Emailed Only	Openers only	Openers only
<b>Panel A: Enrollment 1 month after the intervention</b>								
Received any email	.0157 (.142)	.0311 (.152)						
Received basic email			.0845 (.183)	.0935 (.195)				
Received basic email and additionnal message			-.0049 (.145)	.0123 (.155)				
- <i>Cost</i>					-.135 (.235)	-.146 (.254)	-.0402 (.338)	-.0171 (.368)
- <i>Simplicity</i>					-.085 (.183)	-.09 (.195)	.176 (.277)	.196 (.295)
- <i>Returns</i>					-.144 (.181)	-.152 (.193)	-.191 (.268)	-.188 (.285)
- <i>All info</i>					-.0252 (.184)	.0151 (.197)	.196 (.278)	.307 (.299)
Mean in the control group	1.9598	1.9598	1.9598	1.9598	-	-	-	-
Mean in the basic email group	-	-	-	-	2.0443	2.0443	2.4912	2.4912
<b>Panel B: Enrollment 6 months after the intervention</b>								
Received any email	.282 (.25)	.341 (.265)						
Received basic email			.322 (.321)	.386 (.341)				
Received basic email and additionnal message			.271 (.257)	.327 (.273)				
- <i>Cost</i>					.12 (.434)	.0221 (.464)	.617 (.629)	.485 (.673)
- <i>Simplicity</i>					-.0846 (.324)	-.111 (.343)	.108 (.48)	.11 (.51)
- <i>Returns</i>					.129 (.326)	.13 (.346)	.141 (.48)	.141 (.509)
- <i>All info</i>					-.246 (.322)	-.217 (.342)	-.0934 (.477)	.0409 (.508)
Mean in the control group	6.3309	6.3309	6.3309	6.3309	-	-	-	-
Mean in the basic email group	-	-	-	-	6.6525	6.6525	8.1808	8.1808
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	58314	63246	58314	58314	47498	29049	26644

*Notes:* This table shows the effect of receiving emails on enrolling to a training program. It replicates the regressions and the format of table 3 with a different dependant variable. In panel A we measure enrollment 1 month after the intervention; in panel B we measure enrollment 6 months after the intervention. See more detailed explanations in the footnotes of table 2.8.4.

Table 2.8.6: Impact on enrollment at Afpa as a fraction of total enrollment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed Only	Emailed Only	Openers only	Openers only
<b>Panel A: Enrollment 1 month after the intervention</b>								
Received any email	-.0741 (.831)	.148 (.846)						
Received basic email			-.647 (1.02)	-.597 (1.04)				
Received basic email and additionnal message			.101 (.857)	.378 (.874)				
- <i>Cost</i>					1.31 (1.58)	1.74 (1.78)	1.33 (1.97)	1.9 (2.25)
- <i>Simplicity</i>					.53 (1.01)	.895 (1.05)	.171 (1.25)	.379 (1.33)
- <i>Returns</i>					.208 (.98)	.454 (1.02)	.654 (1.27)	.902 (1.35)
- <i>All info</i>					1.22 (1.06)	1.21 (1.09)	.757 (1.29)	.496 (1.32)
Mean in the control group	4.0480	4.0480	4.0480	4.0480	-	-	-	-
Mean in the basic email group	-	-	-	-	3.3382	3.3382	3.7895	3.7895
<b>Panel B: Enrollment 6 months after the intervention</b>								
Received any email	.189 (1.42)	.515 (1.46)						
Received basic email			-2.4 (1.73)	-2.15 (1.76)				
Received basic email and additionnal message			.984 (1.47)	1.34 (1.5)				
- <i>Cost</i>					2.66 (2.64)	2.57 (2.77)	1.08 (3.14)	.972 (3.3)
- <i>Simplicity</i>					2.37 (1.72)	2.75 (1.76)	.813 (2.11)	1.18 (2.18)
- <i>Returns</i>					3.62** (1.75)	3.74** (1.8)	3.77* (2.19)	3.83* (2.26)
- <i>All info</i>					4.16** (1.79)	4.02** (1.82)	2.05 (2.15)	1.75 (2.18)
Mean in the control group	12.8936	12.8936	12.8936	12.8936	-	-	-	-
Mean in the basic email group	-	-	-	-	10.3048	10.3048	11.7895	11.7895
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	3621	3357	3621	3357	3357	2738	2110	1939

*Notes:* Same regressions as in table 2.8.5, restricting the sample to enrollees only and replacing the enrollment outcome with the fraction of enrollees in an *Afpa* program. Standard errors are in parenthesis: \*\*\* p-value < 0.01, \*\* p-value < 0.05, \* p-value < 0.1.



Table 2.8.7: Summary statistics in the basic email group

	All	Called Afpa center		
	1	No call 2	Call 3	[3 - 2] 4
Female (%)	37.0	37.0	37.9	0.98
Age	41.8	41.8	44.3	2.50
Foreigner (%)	9.0	8.9	20.7	11.77
Married (%)	50.3	50.4	37.5	-12.85
Number of children	1.0	1.0	1.0	-0.04
Duration (months, capped at 18)	13.1	13.1	13.6	0.50
Looking for short-term contract (%)	5.1	5.0	13.8	8.75
Looking for part-time work (%)	9.4	9.4	13.8	4.38
Formal training in desired job (%)	65.6	65.6	56.7	-8.98
Programs match with desired job (%)	6.6	6.6	12.5	5.94
Experience in desired job (months)	128.3	128.3	147.8	19.55
<i>Assistance track</i>				
Low (%)	40.5	40.6	31.2	-9.32
Moderate (%)	44.5	44.4	53.1	8.68
Intensive (%)	13.3	13.3	15.6	2.30
<i>Education level</i>				
No high school nor vocational degree (%)	13.4	13.4	9.4	-3.99
Vocational degree (%)	43.3	43.3	34.4	-8.96
High school diploma or GED (%)	23.6	23.6	43.8	20.18**
Bachelor degree or more (%)	17.8	17.8	6.2	-11.58***
<i>Professional status</i>				
Unskilled worker (%)	2.7	2.7	3.1	0.45
Skilled worker (%)	33.2	33.2	18.8	-14.45**
Employee (%)	54.7	54.7	65.6	10.93
Manager (%)	3.3	3.3	0.0	-3.31***
N =	11740	11708	32	
		99.7%	0.3%	

*Notes:* This table presents summary statistics for the basic email group. It is structured in a similar fashion as table 3.8.3. Column 1 displays the variable means in the basic email group. Columns 2, 3 and 4 compare individuals who called back Afpa and those who did not. See more detailed explanations in the footnotes of table 3.8.3.

Table 2.8.8: Heterogeneous impact of having a high level of formal education on callback

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed only	Emailed only	Openers only	Openers only
High educ	.242**	.206	.235**	.154	.428**	.292	.445	.722
	(.098)	(.188)	(.0978)	(.192)	(.17)	(.302)	(.276)	(.44)
Received any email	.318***	.299***						
	(.0372)	(.0376)						
Received any email X High educ	.412***	.394***						
	(.0672)	(.068)						
Received basic email			.195***	.17***				
			(.0582)	(.0574)				
Received basic email X High educ			.163	.194*				
			(.102)	(.104)				
Received basic email and additionnal message			.355***	.338***				
			(.0436)	(.0442)				
Received basic email and additionnal message X High educ			.484***	.451***				
			(.0811)	(.0819)				
- Cost					-.162	-.184*	-.293	-.331
					(.116)	(.11)	(.217)	(.206)
- Cost X High educ					.439*	.261	.755*	.478
					(.251)	(.243)	(.398)	(.387)
- Simplicity					.118	.141	.204	.216
					(.0897)	(.091)	(.168)	(.172)
- Simplicity X High educ					.179	.118	.321	.222
					(.167)	(.171)	(.278)	(.287)
- Returns					.236**	.228**	.456**	.423**
					(.0988)	(.0982)	(.188)	(.188)
- Returns X High educ					.406**	.312*	.511	.341
					(.191)	(.189)	(.313)	(.314)
- All info					.212**	.224**	.353*	.375**
					(.0974)	(.0977)	(.181)	(.187)
- All info X High educ					.355*	.364*	.451	.417
					(.185)	(.192)	(.299)	(.315)
Mean of calls in the control group	.0085	.0085	.0085	.0085	-	-	-	-
Mean of calls in the basic email group	-	-	-	-	.2726	.2726	.4459	.4459
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	59568	63246	59568	51510	48543	29049	27342

Notes: This table shows the heterogeneous effect of receiving emails on calling back *Afpa* center depending on education level. All regressions use a callback dummy as their outcome and include region fixed effects, as per equation (2.5). Columns are structured in a similar fashion as table 2.8.4. All variables are interacted with the high education dummy. See more detailed explanations in the footnotes of table 2.8.4.

Table 2.8.9: Heterogeneous impact of having responded to baseline on callback

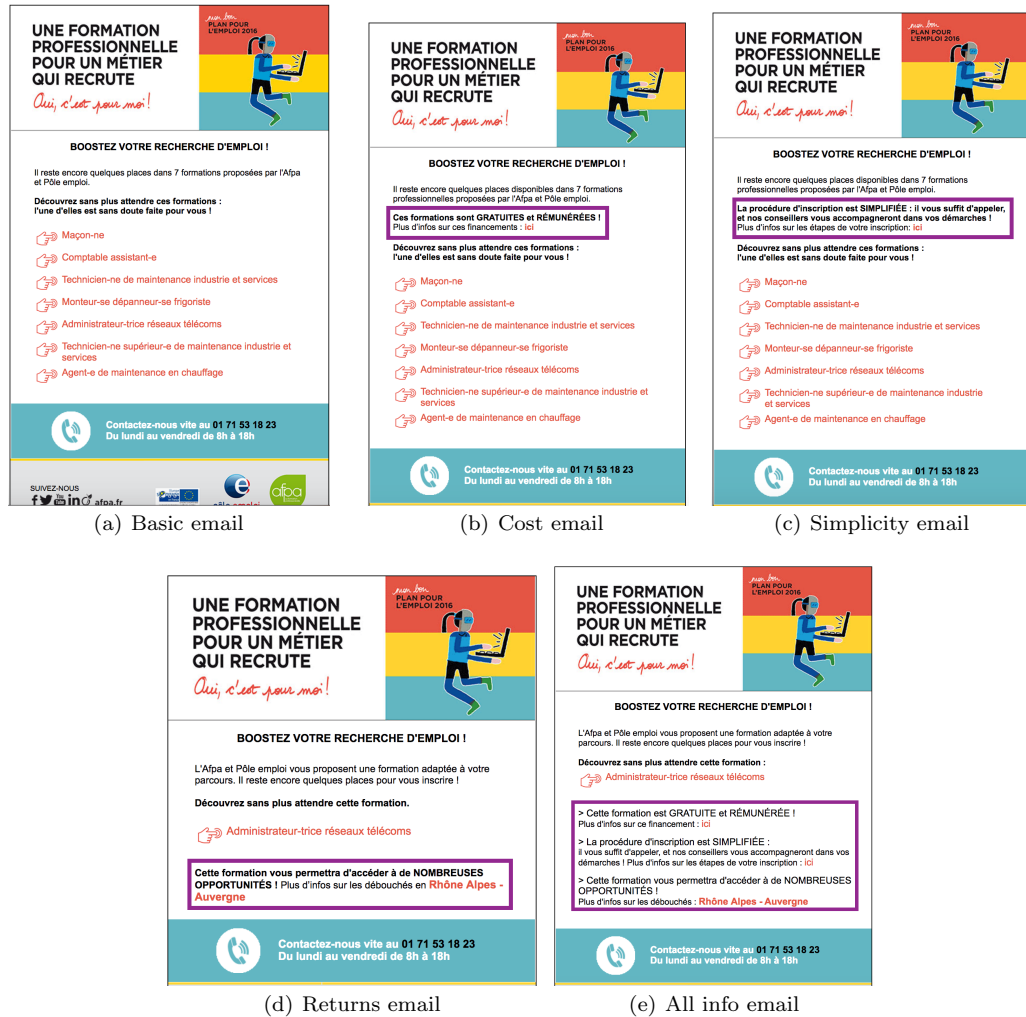
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed only	Emailed only	Openers only	Openers only
Answered baseline	.13	.11	.12	-.0356	.304	.409	.111	-.137
	(.169)	(.351)	(.168)	(.353)	(.281)	(.527)	(.349)	(.763)
Received any email	.401***	.383***						
	(.0316)	(.0321)						
Received any email X Answered baseline	.767***	.695***						
	(.136)	(.135)						
Received basic email			.244***	.229***				
			(.0506)	(.0509)				
Received basic Email X Answered baseline			.156	.167				
			(.172)	(.183)				
Received basic email and additionnal message			.448***	.429***				
			(.0375)	(.0381)				
Received basic email and additionnal message X Answered baseline			.947***	.851***				
			(.168)	(.167)				
- <i>Cost</i>					-.0236	-.115	.0061	-.173
					(.128)	(.123)	(.23)	(.224)
- <i>Cost X Answered baseline</i>					.681	.636	.67	.706
					(.485)	(.474)	(.585)	(.576)
- <i>Simplicity</i>					.135*	.143*	.289**	.271*
					(.0789)	(.0805)	(.147)	(.153)
- <i>Simplicity X Answered baseline</i>					.49	.433	.366	.308
					(.323)	(.329)	(.394)	(.403)
- <i>Returns</i>					.25***	.218***	.458***	.363**
					(.0856)	(.0847)	(.156)	(.156)
- <i>Returns X Answered baseline</i>					1.32***	1.2***	1.32***	1.23**
					(.412)	(.418)	(.498)	(.501)
- <i>All info</i>					.289***	.33***	.469***	.528***
					(.0877)	(.0908)	(.157)	(.167)
- <i>All info X Answered baseline</i>					.647*	.488	.572	.361
					(.357)	(.364)	(.437)	(.446)
Mean of calls in the control group	.0085	.0085	.0085	.0085	-	-	-	-
Mean of calls in the basic email group	-	-	-	-	.2726	.2726	.4459	.4459
Region FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	59568	63246	59568	51510	48543	29049	27342

*Notes:* This table shows the heterogeneous effect of receiving emails on calling back *Afpa* center depending on whether individuals responded to baseline. All regressions use a callback dummy as their outcome, with region fixed effects, as per equation (2.5). Columns are structured in a similar fashion as table 2.8.4. All variables are interacted with the high education dummy. See more detailed explanations in the footnotes of table 2.8.4.

## 2.9. Figures

### 2.9.1. Email types for each treatment group

Figure 2.9.1: Email types for each treatment group

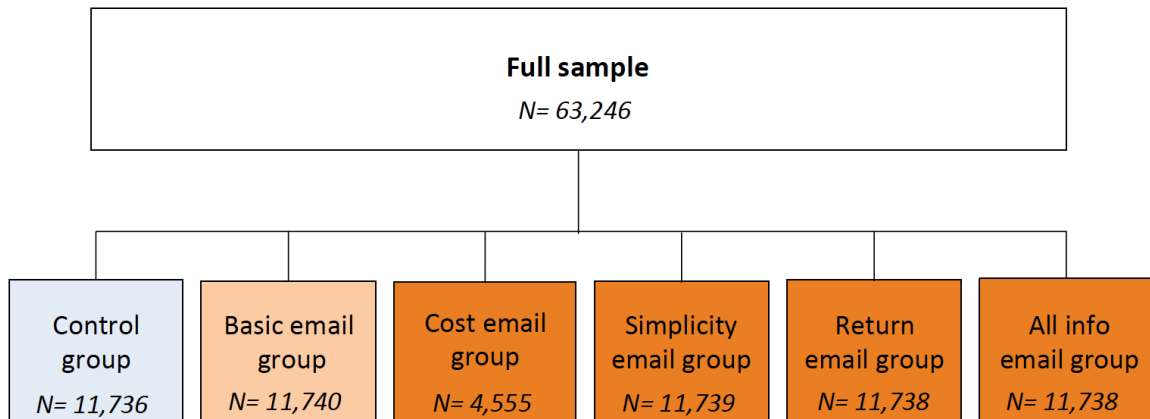


*Notes:* This figure shows the five emails that have been sent to jobseekers belonging to our treatment groups. Figure 2.9.1(b) displays the basic email that was sent to the *Basic email* group. This email served as template for all other emails. In the top left corner, the text says “A training program leading to many job opportunities: yes, it’s for me!”. The main text in the email says “BOOST YOUR JOB SEARCH! There are still some seats left in one of the 7 training programs offered by *Afpa* and *Pôle emploi*. Take a look: there surely is one for you!”. This text is followed by the list of programs in the region (here region 1). The bottom banner gives contact information to call *Afpa* centers.

Figures 2.9.1(b) to 2.9.1(e) shows the four emails in the additional message groups. The sections that differ from the basic email are in purple boxes:

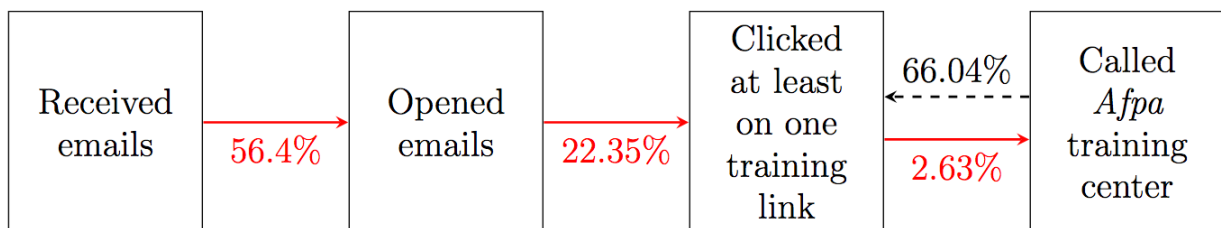
- The email on cost (Figure 2.9.1(b)) adds one sentence saying “This training is FREE and SUBSIDIZED! More info on funding options: here.”
- The email on simplicity (Figure 2.9.1(c)) adds one sentence saying “The registration procedure is SIMPLIFIED : you just need to call and our caseworkers will help you throughout the process ! More info on the steps towards enrollment : here.”
- Figure 2.9.1(d) is an example of the returns email sent to the *Returns email* group. One can note that only one training program is showed, that is most adapted to the job seeker work trajectory. An additional sentence at the bottom of the email says “This training will help you get numerous job opportunities! More info on these opportunities in [REGION].”
- Finally, the email sent to the *All info email* group adds all these additional messages (see as an example figure 2.9.1(e)).

Figure 2.9.2: Randomization design



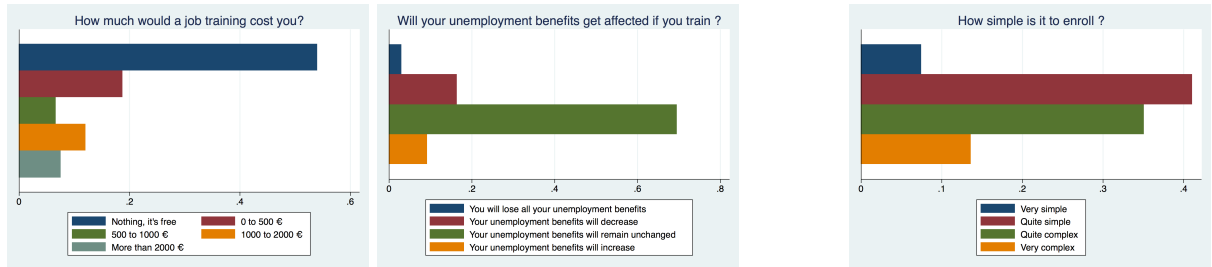
Notes: This figure illustrates the randomization design of the experiment, with the different treatment arms and their corresponding sample sizes.

Figure 2.9.3: Conversion rate of our emails into calls



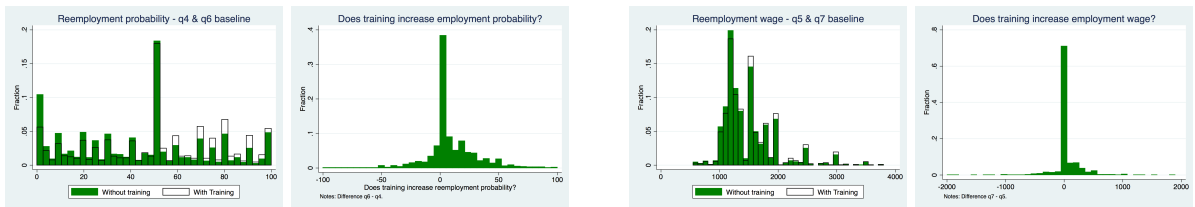
Notes: This figure illustrates the conversion rate of our emails into calls. Starting from the left, it appears that among job seekers who received our emails, 56.40% opened our emails. Out of these 56.40%, 22.35% clicked on at least one of the hyperlink redirecting to the associated training page. The last figure in red indicates that among those who clicked, only 2.63% decided to call the training center. Symmetrically, the percentage in black, above the dashed arrow, indicates that among those who called, 66.04% clicked on at least one training link.

Figure 2.9.4: Answers to baseline questions



(a) Answers on training cost

(b) Answers on enrollment simplicity



(c) Reemployment wages with and without training

(d) Expected reemployment likelihoods with and without training

Notes:

- This figure shows the distribution of answers to our baseline survey about training programs promoted by *Pôle emploi*. Figure 2.9.4(a) displays the distribution of answers to the first two questions about the cost of these programs. The horizontal axis shows the fraction of each answer from 0 to 1. While about half of respondents think training is free, the remaining fraction believe that it is costly. Similarly, while nearly seventy percent of respondents don't think unemployment benefits get affected upon enrollment in a training program, more than 20 percent either think that they decrease or get removed.
- Figure 2.9.4(b) shows the distribution of answers to the third question about the registration process. The horizontal axis on these two figures displays the fraction of each answer from 0 to 1. While about 40% respondents consider that it is quite simple to enroll, 35% view enrollment as quite complex and nearly 15% consider it to be very complex.
- Figure 2.9.4(c) presents the distribution of wage expectations with and without training. The first histogram reports the answers of both questions 4 and 6, asking respondents to estimate their future wage assuming they get reemployed within the following 6 months, with and without training. The second graph computes the difference: a positive result means that the respondent believes training would increase her future wage upon reemployment. Strikingly, more than 70% respondents expect training to make no difference for their future wage.
- Figure 2.9.4(d) shows the distribution of reemployment expectations with and without training. The first histogram reports the answers of both questions 5 and 7, asking respondents to estimate the probability of getting reemployed with and without training. The second graph computes the difference: a positive result means that the respondent believes training participation would increase her chances to get reemployed. Both histograms make it visible that respondents rather believe that training may help them getting reemployed, although nearly 40% expect training to make no difference on their reemployment chances.

## Appendix

### 2.A. Translation of the baseline questionnaire

Thank you for participating in this survey.

It will only take 3 minutes to answer !

[q1] **(In your opinion, if you decide to participate in a 6 month vocational training offered by Pôle emploi, how much would it cost you? (apart from indirect costs such as transportation or childcare expenses)?** *(Selon vous, si vous décidez de suivre une formation professionnelle de 6 mois proposée par Pôle emploi, combien cela vous coûtera-t-il (en dehors des frais indirects comme les transports ou la garde des enfants) ?)*

- Nothing, it's free
- Between 0 and 500
- Between 500 and 1000
- Between 1000 and 2000
- More than 2000

[q2] **In your opinion, if you participate in a 6 month vocational training offered by Pôle emploi:** *(Selon vous, si vous suivez une formation professionnelle de 6 mois proposée par Pôle emploi:)*

- You will lose all your unemployment benefits
- Your unemployment benefits will decrease
- Your unemployment benefits will remain unchanged
- Your unemployment benefits will increase

[q3] **In your opinion, the steps to enrol into a 6 month vocational training offered by Pôle emploi are:** *(Vous pensez que les démarches pour s'inscrire dans une formation professionnelle de 6 mois proposée par Pôle emploi sont:)*

- Very simple
- Quite simple
- Quite complex
- Very complex

[q4] **What are the chances that you find a full-time job within the next 12 months?**

**Set a percentage between 0 and 100 using the cursor.**

**0 (very low) 100 (very high)** *Quelles sont vos chances de retrouver un emploi à temps plein dans les 12 prochains mois? Indiquez un pourcentage entre 0 et 100 à l'aide du curseur. 0 (très faibles) 100 (très fortes))*

[q5] **If you find a full-time job in your professional sector or in a closely related one within the next 12 months, how much will be your net monthly wage?**

**Set an amount between 0 and 100000 euros.** *(Si vous obtenez un emploi à temps plein dans votre secteur d'activité ou dans un secteur proche dans les 12 prochains mois, de combien sera votre salaire mensuel net ? Indiquez un montant entre 0 et 100000 euros.)*

Imagine from now on that you have participated in a 6 month vocational training for a job in your professional sector or a closely related sector. *(Imaginez à présent que vous avez suivi une formation professionnelle de 6 mois dans un métier de votre secteur d'activité ou dans un secteur proche.)*

[q6] **What are the chances that you find a full-time job within the 12 months following the training ?**

**Set a percentage between 0 and 100 using the cursor.**

**0 (very low) 100 (very high)** *(Quelles sont vos chances de retrouver un emploi à temps plein dans les 12 mois qui suivent la formation? Indiquez un pourcentage entre 0 et 100 à l'aide du curseur. 0 (très faibles) 100 (très fortes))*

[q7] **If you find a full-time job in your professional sector or in a closely related one within the 12 months following the training, how much will be your net monthly wage?**

**Set an amount between 0 and 100000 euros.** *(Si vous obtenez un emploi à temps plein dans votre secteur d'activité ou dans un secteur proche dans les 12 mois qui suivent la formation, de combien sera votre salaire mensuel net ? Indiquez un montant entre 0 et 100000.)*



## 2.B. Heterogeneous effects without region fixed effects

Table 2.B.1: Heterogeneous impact of having a high level of formal education on callback (regressions without region fixed effects)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed only	Emailed only	Openers only	Openers only
High educ	-.0149 (.0149)	-.256** (.111)	-.0149 (.0149)	-.316*** (.118)	.143 (.1)	-.181 (.224)	.13 (.166)	.0507 (.336)
Received any email	.322*** (.0373)	.305*** (.038)						
Received any email X High educ	.431*** (.0684)	.413*** (.0696)						
Received basic email			.196*** (.0581)	.172*** (.0573)				
Received basic email X High educ			.158 (.102)	.189* (.104)				
Received basic email and additionnal message			.36*** (.0437)	.346*** (.0447)				
Received basic email and additionnal message X High educ			.508*** (.0824)	.476*** (.0835)				
- <i>Cost</i>					-.0638 (.102)	-.0544 (.0912)	-.106 (.192)	-.0994 (.175)
- <i>Cost X High educ</i>					.627*** (.231)	.456** (.22)	.969*** (.367)	.695** (.352)
- <i>Simplicity</i>					.118 (.0896)	.14 (.0909)	.203 (.168)	.213 (.172)
- <i>Simplicity X High educ</i>					.184 (.167)	.124 (.171)	.327 (.278)	.227 (.287)
- <i>Returns</i>					.235** (.0987)	.228** (.0981)	.453** (.188)	.418** (.188)
- <i>Returns X High educ</i>					.41** (.191)	.321* (.19)	.5 (.313)	.335 (.314)
- <i>All info</i>					.211** (.0973)	.223** (.0977)	.347* (.18)	.367** (.186)
- <i>All info X High educ</i>					.36* (.185)	.37* (.192)	.451 (.299)	.419 (.315)
Mean of calls in the control group	.0085	.0085	.0085	.0085	-	-	-	-
Mean of calls in the basic email group	-	-	-	-	.2726	.2726	.4459	.4459
Region FE	No	No	No	No	No	No	No	No
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	59568	63246	59568	51510	48543	29049	27342

Notes: This table shows the heterogeneous effect of receiving emails on calling back *Afpa* center depending on education level. All regressions use a callback dummy as their outcome, for the restricted sample where we remove the cost email group. This allows to remove region fixed effects. Columns are structured in a similar fashion as table 2.8.4. All variables are interacted with the high education dummy. See more detailed explanations in the footnotes of table 2.8.4.

Table 2.B.2: Heterogeneous impact of having responded to baseline on callback (regressions without region fixed effects)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	ITT	ITT	ITT	ITT	Emailed only	Emailed only	Openers only	Openers only
Answered baseline	-.0097 (.0097)	-.285 (.242)	-.0097 (.0097)	-.439* (.251)	.151 (.172)	.0223 (.442)	.0502 (.218)	-.688 (.59)
Received any email	.413*** (.0321)	.397*** (.0328)						
Received any email X Answered baseline	.781*** (.138)	.71*** (.138)						
Received basic email			.244*** (.0506)	.229*** (.0509)				
Received basic Email X Answered baseline			.161 (.172)	.175 (.183)				
Received basic email and additionnal message			.463*** (.0381)	.447*** (.0389)				
Received basic email and additionnal message X Answered baseline			.964*** (.171)	.868*** (.17)				
- <i>Cost</i>					.168 (.116)	.107 (.109)	.329 (.211)	.201 (.201)
- <i>Cost X Answered baseline</i>					.753 (.462)	.702 (.458)	.7 (.557)	.687 (.554)
- <i>Simplicity</i>					.135* (.0789)	.143* (.0805)	.291** (.147)	.271* (.153)
- <i>Simplicity X Answered baseline</i>					.496 (.323)	.435 (.329)	.377 (.394)	.309 (.403)
- <i>Returns</i>					.251*** (.0856)	.22*** (.0848)	.452*** (.157)	.358** (.157)
- <i>Returns X Answered baseline</i>					1.32*** (.413)	1.18*** (.419)	1.32*** (.498)	1.22** (.502)
- <i>All info</i>					.289*** (.0878)	.331*** (.0909)	.464*** (.157)	.523*** (.167)
- <i>All info X Answered baseline</i>					.642* (.357)	.481 (.364)	.573 (.437)	.362 (.447)
Mean of calls in the control group	.0085	.0085	.0085	.0085	-	-	-	-
Mean of calls in the basic email group	-	-	-	-	.2726	.2726	.4459	.4459
Region FE	No	No	No	No	No	No	No	No
Covariates	No	Yes	No	Yes	No	Yes	No	Yes
Sample	All	All	All	All	Emailed	Emailed	Openers	Openers
N =	63246	59568	63246	59568	51510	48543	29049	27342

*Notes:* This table shows the heterogeneous effect of receiving emails on calling back *Afpa* center depending on whether individuals responded to baseline. All regressions use a callback dummy as their outcome, for the restricted sample where we remove the cost email group. This allows to remove region fixed effects. Columns are structured in a similar fashion as table 2.8.4. All variables are interacted with the high education dummy. See more detailed explanations in the footnotes of table 2.8.4.

## Chapter 3

# Behavioural Barriers to Human Capital Investment: Evidence from a Survey on French Jobseekers

*This chapter is joint with Bruno Crépon, Lucas Girard and Anett John*

### **Abstract**

Despite potentially high gains in labor market outcomes, the demand for vocational training programs in France remains low. We design a survey administered to 20,048 jobseekers that studies the barriers which prevent individuals from taking up vocational training. In addition to “external” barriers, such as lack of information or inadequate supply, we also investigate “internal” psychological barriers, including self-efficacy, executive function, and self-control problems. The first part of the paper presents the survey and describes its main results, contrasting the prevalence of external and internal barriers for different categories of jobseekers. In the second part of the paper, we use the insights gained from the survey to inform the design of a randomized controlled trial. We develop interventions to target either external beliefs (about the world) or internal beliefs (about the self). Interventions consist of app-based courses with interactive sessions involving groups of jobseekers.

### 3.1. Introduction

In France, vocational training is one of the main policies consistently used by public authorities to combat unemployment. Since the 2008 economic crisis, the French government implemented numerous campaigns to increase training, with increasingly ambitious objectives: In 2013, the “Plan 30,000 formations” was initiated, with the aim to train 30 000 jobseekers. One year later, the “Plan 100,000 formations” focused on sectors experiencing shortages of skilled workers. In 2016, the “Plan 500,000” was launched, with the aim to double the annual number of jobseekers joining a training program, bringing the total to 1 million. Finally, the “Plan d’Investissement dans les Compétences” (hereafter *PIC*) was launched in 2018, pledging one billion euros to help train 2 million jobseekers. While previous plans had focused on the long-term unemployed, the *PIC* shifted the emphasis squarely to young jobseekers with few or no qualifications.

Each of these campaigns consisted of concerted public efforts to provide and finance high-quality training places, inform jobseekers, and motivate them to participate. However, even a massive increase in allocated public funds did not translate into an increase in training enrollment: The entry rate into training programs for jobseekers has been consistently around 10% since 2010.<sup>1</sup> From an external viewpoint, training programs should be appealing to jobseekers: they receive skills training for positions which are strongly demanded by firms, with funding provided for the registration fees. The low take-up of training programs suggests that the provision of skills and financial support are not sufficient to generate demand for vocational training.

The goal of this paper is to shed light on this discrepancy. Using the database of jobseekers registered with the French public employment service *Pôle emploi*, we surveyed 20,048 jobseekers to understand the determinants of the demand for training programs. Instead of relying solely on supply-side factors (information, quality of supply, financing), we take inspiration from evidence pointing to the relevance of behavioral dimensions in the demand for training (Babcock et al. 2012).

In the first part of our paper, we present the results of our survey, whose aim is to identify barriers preventing jobseekers from taking up vocational training. To this end, we distinguish between “external” (neoclassical) barriers and “internal” (psychological) barriers. External barriers relate to obstacles originating from the environment, over which jobseekers have little control. In line with both the academic literature (Jensen (2010), Barr and Turner (2017)) and qualitative research we conducted among jobseekers, we choose to study four external barriers: the lack of information jobseekers may have about training programs (informational barriers), barriers related to the quality and supply of training programs, barriers related to the returns to training (employment opportunities after completing a training program), and finally the financing of training programs. Conversely, internal barriers relate to obstacles jobseekers face due to their preferences, beliefs, or cognitive constraints. Following the economic and psychological literature, we consider four classical internal barriers: self-efficacy (see e.g. McKelway (2018), as well as Caliendo et al. (2015) on the related concept of locus of control), executive function (Berger et al.

---

<sup>1</sup>This percentage represents the enrollment rate within 12 months after registering at *Pôle emploi* (Source : *DARES*). There was an increase in 2016 following the implementation of the “Plan 500,000”. Yet, that rise was largely initiated by caseworkers rather than jobseekers (see *Cour des Comptes* (2018)).

(2019)), self-control (Paserman (2008, DellaVigna and Paserman (2005b)), and self-esteem (Ghosal et al. (ming), Falk et al. (2006a)). Furthermore, we investigate what we refer to as “external beliefs”, namely jobseekers’ beliefs about their environment, such as the returns to training or the availability of funding. If these beliefs are biased, they may in themselves constitute internal barriers.

Across all barriers studied, we find strong evidence for the importance of beliefs. This concerns both “internal” beliefs about one’s own ability to achieve desirable outcomes (self-efficacy) and “external beliefs” about the availability of financial support. We show that the latter is highly pessimistic, suggesting that simple information to correct beliefs may be effective.

In the last part of the paper, we use the survey’s insights to inform the design of a randomized controlled trial whose interventions aim at changing jobseekers’ beliefs about themselves or their environment.

This design will allow us to assess the impact of each intervention relative to each other, as well as their heterogeneous impact according to the group composition (in terms of jobseekers type) jobseekers belong to. In addition, we plan to run a short baseline survey during the online collective meeting, aimed at identifying the most likely barrier jobseekers face. This will allow us to examine the heterogeneity of treatment effects with respect to these measured barriers.

**Related literature.** Most studies on vocational training focus on its effectiveness in terms of reemployment and wage, with very heterogeneous results (see for instance Greenberg et al. (2003), or Card et al. (2017) for a review). However, little is known about the demand for such programs. Yet, understanding the determinant of the demand for training is crucial in order to attract the relevant population and to ensure that they benefit from training. Our paper starts to fill this gap. The literature on educational investment traditionally point out the role of cost and perceived returns in schooling participation (see for instance Dizon-Ross (2019) or Jensen (2010)). As with investment decision in formal education, individuals’ participation into training programs also depends on these parameters (Jacobson and Davis (2017)). However, their value might be harder to know in the context of vocational training. In France in particular, the training system for jobseekers involves several actors who contribute to its financing. Consequently, rules for obtaining training grants are not straightforward, while direct costs *per se* are hard to assess. Bertrand et al. (2004) also detail in their work the impact of complex administrative procedures : people experiment psychological aversion to program, which tend to lower their propensity to join such program. Therefore, and as explained by Babcock et al. (2012), this complex institutional setting may prevent individuals from joining such programs. In addition, due to the large heterogeneity of training programs, in terms of duration, sector of activities, or targeted population, training returns are particularly hard to assess. Results from the behavioral economics have demonstrated that people often fail in making optimal decision under uncertainty contexts (see Dellavigna (2009), Spinnewijn (2015)). Our paper also relates to the literature on the role of internal. For instance, Benabou and Tirole (2002) highlight the crucial role of motivation and self-confidence in one’s own abilities and in the decision making process, while Caliendo et al. (2020) emphasizes the role of locus of control (self-efficacy) in training participation.

Finally, our paper is built on the literature on the impact of information on job search behavior, as

the interventions we describe in the last part of our paper relies on the delivery of news to jobseekers. Several experiments have successfully demonstrated the impact on information provision on labor market outcomes. For instance, [Altmann et al. \(2018\)](#) find that delivering news about to jobseekers has a positive impact on employment and earnings. Regarding educational investment, [Dizon-Ross \(2019\)](#) shows that delivering information about cost and returns can affect schooling participation Regarding vocational training. Turning more specifically to vocational training, [Barr and Turner \(2017\)](#) find that providing news about the costs and the returns to US unemployment insurance beneficiaries increase their probability to enroll. Thus, and as stated by [Dasgupta et al. \(2012\)](#), proper information about training programs can increase enrollment in such programs.

The rest of the paper is structured as follows. Section 2 provides information about the institutional setting of the French vocational training system. Section 3 details our survey questionnaire. Section 4 reviews the data we use in our paper. In section 5 we present our main survey results. Section 6 uses the diagnosis we make in section 5 to propose an experimental design with interventions aimed at targeting external or internal barriers. Section 7 concludes.

## 3.2. Institutional Setting

This section reviews the features and the objectives of the *PIC* since we run our study as part of this training plan. We also provide some general information about both the institutional setting of vocational training in France and the enrollment process of jobseekers into training programs.

### 3.2.1. Plan d’investissement dans les compétences (*PIC*)

The *PIC* is a government plan which is part of a broader investment plan aimed at initiating several structural reforms and meeting economic challenges experienced in France. Its implementation started in 2018 and is scheduled until 2022. In that overall context, the objective of the *PIC* is to promote access to training for jobseekers, particularly those isolated from the labor market. The objective is to improve their labor market outcomes (employment, wages). Two types of jobseekers are specifically targeted:

- young people, *i.e.*, individuals under 26 years old;
- low-skilled people *i.e.*, individuals with a level of qualification below the high school degree.<sup>2</sup>

Indeed, these populations experience a higher level of unemployment compared to the whole population: the global unemployment rate reached 9,4% in 2017, whereas it amounted to 22% for young people<sup>3</sup> and to 17% for low-skilled people<sup>4</sup>. The objective of the *PIC* is to train a million of young jobseekers and an additional million low-skilled jobseekers.

<sup>2</sup>Individuals who completed the last year of high school but did not pass the diploma certifying the completion of high school (“baccalauréat”) are also considered as low-skilled.

<sup>3</sup>This percentage corresponds to the unemployment rate among people who are less than 25 years old. Source: “Enquête Emploi” survey, [Insee \(2018\)](#).

<sup>4</sup>This percentage corresponds to the unemployment rate among people who have at most a level of education that corresponds to the last year of junior high school, be it with or without obtaining the “baccalauréat”. Source: “Enquête Emploi” survey, [Insee \(2018\)](#).

Since this plan's final purpose is to promote access to sustainable employment, training programs promoted are extended intensive ones. The *PIC* thus focuses on long training courses, lasting at least two months, intending to make jobseekers directly employable in jobs the training program leads to (*i.e.*, boilermaker, miller, etc.). Besides, refresher training in reading, writing, arithmetic, IT skills, etc. can precede these long training programs, as these shorter training courses may be prerequisites before entering into long training courses or may, by themselves, facilitate jobseekers labor market integration.

Public authorities have allocated 14.6 billion euros to the realization of this plan: 6.7 billion are dedicated to the training of young jobseekers, 7.1 billion for low-skilled jobseekers. In addition to the funding of training programs, 10% of the entire budget are devoted to testing new tools designed to improve the training system and ease the entry into training. Compared to the previous training plan (*Plan 500 000*), the global budget allocated to the *PIC* is much higher (14.6 billion versus 1.5 billion), although the *Plan 500 000* took place over only two years (2016 and 2017)<sup>5</sup>. This increase reports a willingness to improve access to training and to support the most vulnerable people in terms of reemployment.

### 3.2.2. Enrollment process into a vocational training program

To implement the *PIC*, all actors involved in the training of jobseekers – the State, *Pôle emploi*, and regions - are mobilized. Indeed, these three actors are the main funders of vocational training for jobseekers. Together, they form an institutional setting contributing each year to 5 billion euros in the training of jobseekers<sup>6</sup>. In the context of the *PIC*, the role of the State is to define a training policy with its objective, its targeted population, and its financial means, while the role of *Pôle emploi* is to reinforce the support of jobseekers, directing them to training programs when appropriate, according to the State's guidelines. *Pôle emploi* and regions, which usually buy training positions<sup>7</sup> have also to be careful regarding the training positions they buy, in terms of reemployment opportunities<sup>8</sup> but also in terms of novelty of pedagogical approaches. Regarding the training supply, training centers are thus encouraged to favor innovative digital tools aimed at making teaching easier. Overall, this multiplicity of actors, whose roles are sometimes identical and poorly separated, makes the general system of vocational training unclear and complex, which can affect jobseekers enrollment.

As regards registration processes, jobseekers registered at *Pôle emploi* and wishing to join a training program must, above all, gain validation from their caseworker<sup>9</sup> in order to get a funding for the training program. The caseworker reviews the consistency of the training project of the jobseeker and checks that the program the jobseeker decided to enroll into leads to jobs offering a sufficient level of reemployment opportunities<sup>10</sup>. The jobseeker also has to write a cover letter: the goal is to ensure her commitment to

---

<sup>5</sup>See report [Cour des Comptes \(2018\)](#).

<sup>6</sup>See [Annexe au Projet de Loi de Finances \(2018\)](#) (Appendix of the 2018 Draft Budget Bill).

<sup>7</sup>Traditionally, *Pôle emploi* buys positions for training courses lasting on average 300 hours whose goal is to make jobseekers quickly find a job. In comparison, regions buy positions for longer training programs (on average 750 hours) whose goal is to train the most vulnerable people (see [Cour des Comptes \(2018\)](#)).

<sup>8</sup>They rely on field studies such as "Enquete Besoin en Main d'Oeuvre" conducted each year towards firms by *Pôle emploi*.

<sup>9</sup>When they register at *Pôle emploi*, each jobseeker is assigned to a caseworker according to her level of autonomy regarding job search

<sup>10</sup>To our knowledge, there is no reliable resource or report showing the returns of training programs: at best, jobseekers and caseworkers can obtain information about reemployment opportunities for a given sector or job.

the training program and to make her aware of the constraints associated with enrollment (transportation and childcare costs, transportation duration, etc.), to avoid dropouts during the training. Simultaneously, she has to look for funding. To do so, she receives help from her caseworker, who orients her in priority towards training programs that are already subsidized. Such training programs are training positions already bought by Pôle emploi or regions through public auctions, in various sectors chosen according to the opportunities of reemployment they offer. Such training programs are therefore entirely free for jobseekers (in terms of training fees; it does not include potential ancillary costs due to transportation for instance). However, if none of these subsidized training programs correspond to the job seeker's training project, she then has to look for a training center offering it and request individual funding. As a last resort, *Pôle emploi* can offer such funding. It entirely covers the costs of the training program but requires a cover letter from the job seeker, the caseworker approval regarding the consistency of the training project, and the name of the training center the job seeker has chosen. These documents are then examined by an independent commission that decides to allocate either all of the funding or to not allocate it at all, as jobseekers are not allowed to supplement their funding training by themselves. All in all, obtaining funding for a non-subsidized training program is therefore quite selective, much more than for subsidized programs. However, in 2017, out of all training programs funded by Pôle emploi, individual training funding represented more than 68%, versus only 8.4% for already subsidized training courses. Thus, the financing of training represents a major challenge both for jobseekers (as most of them have to request an individual funding) but also for the budgetary balance of public administrations. During these steps, the job seeker receives help from her caseworker. However, while *Pôle emploi* displays on its website the list of subsidized training available for each sector, information about non subsidized training programs is not centralized but instead dispatched to several external websites. Despite the regular organization of information workshops about training at *Pôle emploi*, very often, caseworkers redirect jobseekers to existing online resources. Having access to consistent information about training programs and vocational training can thus be difficult, especially for jobseekers who are uncomfortable with digital technologies.

Once the job seeker has found both her training center and funding, she has to attend an information meeting in the training center. She may also have to take entrance tests, particularly for extended training programs requiring prerequisites, for instance, in French, arithmetic, or knowledge of digital resources. The training center must then confirm the entry into training of jobseekers, so does the caseworker. Before 2016, it was up to the jobseeker to report the entrance documents to her caseworker and her training center. This procedure was fully digitalized in 2016, which considerably simplified the registration procedure for jobseekers.

Overall, the entry into training therefore includes several steps, which may be complex for some of them. Indeed, in addition to the decision to join a training program, jobseekers have to look for information about reemployment opportunities, funding, training centers, training timetables, etc. They also have to deal with the administrative steps and might have to do entrance tests. Self-esteem, about the ability to succeed in these tests and in the training itself, as well as the ability to organize efficiently are additional issues involved in the enrollment process. Furthermore, as information does not always



exist or may be difficult to access, beliefs and perception play a key role in the decision to join a training program. In our questionnaire, we thus systematically explore all of these dimensions, starting from the beginning of the entry to the training process (construction of the training project) until the actual enrollment into training.

In order to identify the obstacles jobseekers encounter not only when it comes to decide whether or not to undertake training, but also during the enrollment process, we built our survey so that it systematically reviews the steps of entry to training. As we conducted our survey as part of the *PIC*, our questions relate to jobseekers' perceptions towards training programs which are promoted under this plan. As mentioned in subsection 3.2.1, the ultimate purpose of this plan is to increase jobseekers' skills and to help them to quickly reach better labour market outcomes. Therefore two types of training programs are highlighted: (i) training courses of at least two months to acquire the skills required for a specific occupation (for example: boilermaker, milling machine, kitchen clerk, etc.); (ii) refresher training in basic knowledge (written and oral language proficiency, numeracy skills) which are prerequisites for training described in (i). We thus explicitly focus on these two types of training programs in our survey.

Overall, we surveyed 100,041 jobseekers registered at *Pôle emploi*, between March and July 2020. Four reminders were sent at three weeks of interval each. The questionnaire was administered online, and we invited jobseekers to respond by email. We relied on the software *Pôle emploi* usually used to administer our questionnaire. Therefore, surveyed individuals received our emails as if they were directly sent from *Pôle emploi*. We end up with a response rate of roughly 20% (with a complete response rate of approximately 9%).

### 3.3. The survey

#### 3.3.1. Survey questionnaire

This section provides a description of the survey we administered to jobseekers. In total, our survey contains 34 questions, distributed into seven modules.

The first module (Introduction - questions 1 to 6) collects general information about jobseekers' preferences towards job search, in terms of types of contract, the professional sector and the job targeted. We also ask them to estimate, among the set of jobseekers looking for a job in the same sector as them, the proportion finding a sustainable job (a permanent contract, or a fixed-term contract of at least six months) within three months from registration at *Pôle emploi*. The aim of this question is to collect information about jobseekers' beliefs regarding the probability of return to employment without any training program. By comparing these responses with those about return to employment *after* a training program (see next module), we obtain the expected added value (or loss) of training programs from jobseekers. Finally, comparing this added value with the empirical added value from *Pôle emploi*'s historical data allows us to compute the distance between jobseekers' beliefs and the real situation, and to know whether jobseekers are rather pessimistic or optimistic regarding training programs' returns.

The second module ("*Attentes vis-à-vis des formations et de leur intérêt*" - questions 7 to 14) collects

data on jobseekers expectations towards training programs, and is divided into two parts. The first part (questions 7 to 9) collects information about jobseekers vocational training background. After asking them whether they have already received a training of at least two months in any job or sector, the rest of the questionnaire is explicitly interested in their opinion about training courses promoted under the *PIC* (mainly training courses lasting at least two months, aimed at making jobseekers involved in those programs directly operational in the sector or jobs they are interested in). The purpose of this first part is therefore to obtain information about the perceived use of such a training within their sector, as well as the type of training (upgrade training, certifying training, refresher training, or training for a new job) which would be, in their opinion, the most useful.

The second part of this module asks them to estimate different proportions regarding training enrollment and return to employment after ending a training program. In particular, we ask them to estimate the proportion of jobseekers ending and validating a training course after registering for it. We also ask them to estimate the proportion of jobseekers finding a sustainable job three months after the end of the training, whether for a training of at least two months, leading to a job corresponding to their sector of activity or for a refresher training, or for the combination of both. This information is crucial as it will be compared to jobseekers' answer in the first module on the proportion of jobseekers who find a sustainable job in their sector of activity without any training (see question 6). This will enable us to obtain information about their beliefs in terms of training returns, and to know whether these beliefs are correct by comparing them to *Pôle emploi*'s empirical data. Finally, this module ends with a question about the proportion of free training places among those offered and listed by *Pôle emploi*. As training funding is an important issue for jobseekers (see subsection 3.2.2), the comparison with field data will again allow us to highlight any gaps in the knowledge of training funding systems promoted by *Pôle emploi*.

The third module ("*Information à propos des formations*" - questions 15 to 17) examines information jobseekers have about vocational training. We collect data on their knowledge about the different types of training, in terms of certification, funding, and administrative procedures, but also about their ability to obtain practical information about the training they are interested in (starting date, duration, location, possible entrance tests, etc.). We also review the information sources they use to learn about vocational training.

The fourth module ("*Offre et qualité des formations*" - questions 18 to 20) examines how jobseekers perceive the quality of available training programs, as well as their opinion about training offer. In particular, we examine their opinion about the quantity of available training courses, their geographic accessibility, and their relevance for finding a job and prosper in it. We are also interested in their opinion and perception of training content (ability to follow the theoretical level of training courses, balance between theory and practice, etc.) and the quality of speakers (in human and technical terms). Finally, we measure the importance jobseekers attach to certification.

The fifth module ("*Capacité à s'inscrire et suivre une formation*" - questions 21 to 26) collects data about jobseekers behavioural beliefs towards their intrinsic ability to join and validate a training program. In particular, we ask them whether they think they have both the technical and social skills to find a job

in the sector of activity they are interested in, and to validate a training of at least two months, leading to a job corresponding to their sector of interest (questions 21 et 22). The goal of the whole module is to explore the extent to which internal barriers may affect jobseekers training enrollment. In this module, we focus on depression (see question 23), self-efficacy (see question 24), and self-esteem (see question 25), using questions from the psychological and behavioural literature.<sup>11</sup>

The sixth module (“*Accès aux formations - sélection et financement*”) collects data about jobseekers’ perception regarding the selection process of training programs, particularly in terms of funding access (question 27 to 29). It also assesses the importance jobseekers attach to funding to pursue a training program (q29 to 30).

The seventh module (“*Plans et intentions*” - questions 31 to 32) collects data about jobseekers’ intentions towards training. The goal is to assess how close the job seeker is to entering training. This will enable us to design heterogeneous intervention according to this distance.

Finally, the last module (“*Obstacles à l’entrée en formation*” - questions 33 to 34) collects data about the main internal barriers (depression, lack of self-esteem, or self-efficacy) as well as the main external barriers (lack of information, of training supply, of funding, or of reemployment opportunities after training programs) jobseekers face to when they consider the entry into training courses. These barriers are all explored in detail throughout the questionnaire, within the different modules dedicated to them. Thus, each barrier pointed out in this last module makes it possible, by referring to the module dedicated to it, to understand precisely the reasons why this obstacle hinders the entry into training of the job seeker. It will therefore be possible in a second stage to design interventions to remove this obstacle.

## 3.4. Data and sample description

In this section, we describe the data we use in our study, and provide a description of our sample.

### 3.4.1. Survey Data

The structure of our survey allows us to directly extract the information from job seekers answers to examine the main barriers (either external or internal) they face to (see section above). In addition to these data, answers collected about the employment rate after completing a training program and the proportion of available subsidized programs will be compared with administrative records (see next section) in order to assess the degree of pessimism or optimism of jobseekers, which will help us design appropriate interventions.

### 3.4.2. Administrative data

We rely on administrative data to collect information about training returns. More precisely, we measure the proportion of jobseekers finding a job within three months from the end of a training

---

<sup>11</sup>Questions on related to depression are from [Manos et al. \(2011\)](#), and are adapted in French by [Wagener et al. \(2014\)](#). Questions on self-efficacy are from [Schwarzer \(1995\)](#), and are adapted in French by Michelle Dumont, Ralf Schwarzer and Matthias Jerusalem in 2000. Questions on self-esteem are from [Rosenberg \(1965\)](#), and are adapted in French by [Vallieres and Vallerand \(1990\)](#).

program, among those who registered and completed one at least once. Historical data from *Pôle emploi* list all training programs jobseekers have enrolled into, with their respective sector of activity and completion rate. These data also provide detailed information about the objective and duration of each training program. This enables us to distinguish between refresher training programs, long training programs, and the combination of both (which we define as a refresher training program, followed by a long training course within three months). In addition, we also have access to historical data reporting all reemployment episodes jobseekers have experienced while they were registered at *Pôle emploi*. These data give detailed information about the type of contract and their start and end dates, which enables us to only focus the type of jobs we are interested in when looking at reemployment probabilities, *i.e.* sustainable jobs (permanent contracts or fixed term contract lasting at least six months). By combining these two datasets, we are able to assess an average reemployment and completion probability for each type of training program, and to disaggregate this information per sector.

We also use a representative sample of jobseekers from a database called *Fichier historique*, host by *Pôle emploi*, to compute the reemployment probabilities of jobseekers within three months from their registration at *Pôle emploi*, whether they completed or not a training program. This database reports all unemployment spells of jobseekers who registered at least once at *Pôle emploi*. As these data contain information about the sector of activity in which jobseekers are looking for a job, combining this dataset with historical data about reemployment episodes enables us to compute the average reemployment probability within three months from registering at *Pôle emploi* for each sector. Finally, we use unemployment records from *Pôle emploi* to collect socio-demographic information about jobseekers, such as age, gender, level of education, sector of activity of interest, etc. These data also indicate the degree of assistance they receive from their caseworker (assistance track)<sup>12</sup>. Overall, using these data allows us have reliable information about jobseekers we do not have to obtain through our survey, and to constitute our total sample of jobseekers along these variables.

### 3.4.3. Sample description

As our survey is conducted as part of the *PIC*, we target jobseekers corresponding to the target of the government as a priority. As mentioned in subsection 3.4.2, administrative records from *Pôle emploi* allow us to precisely identify these jobseekers, as it contain detailed information on both individuals' educational background, and age. As explained in subsection 3.2.1, in the context of the *PIC*, low and unskilled people are defined as having at most a level of education corresponding to the high school degree (without graduating), while young people are defined as being strictly less than 26. However, in order to be able to distinguish between the specific barriers experienced by *PIC* jobseekers and the ones experienced by the general population of jobseekers, we have also included in our sample individuals who are outside the scope of the *PIC*. In the end, our sample comes from a stratified sampling along age (less than 26 versus more than 26) and educational background (low and unskilled people vs medium and high skills jobseekers), where the proportion of young on the one hand and low and unskilled jobseekers

---

<sup>12</sup>When they register to *Pôle emploi*, jobseekers are assigned to a caseworker according to their degree of autonomy towards job search (high/medium/low)

on the other hand are overrepresented. This sampling process ensures to have in our sample enough jobseekers from each strata to run our analysis. Our final sample amounts 100 041 jobseekers, with 60% of young jobseekers and 40% of low or unskilled jobseekers, distributed as follows: 40% of the entire sample consists of young jobseekers who are also low or unskilled jobseekers; 20% of the sample are not young but are low or unskilled jobseekers; 20% are young but are low or unskilled jobseekers while the remaining 20% are neither young nor low or unskilled jobseekers (see Table 3.8.2)<sup>13</sup>.

Finally, since our survey is conducted online, we only kept in our sample jobseekers who agreed to receive information by email.

Table 3.8.3 provides summary statistics of socio-demographic variables for our entire sample, and according to the response status to our survey. In the whole sample, the average age is approximately 30, and about half of the sample consists of women. The average unemployment spell is 20 months, and 23% of jobseekers receive intensive assistance from their caseworker, which indicates a weak degree of autonomy regarding job search. The table also indicates that approximately 9% of our sample integrally responded to our survey. This rate is low, but fairly standard for such online surveys targeting jobseekers registered at *Pôle emploi*. Finally, the table shows that socio-demographic characteristics are, for the majority of them, unbalanced between respondents and non respondents. Respondents are on average 5 years older than non respondents and are 10 percentage points more likely to have a level of education equivalent or above the high school degree. They are also in less often in an intensive assistance track from the caseworker. This may suggest that respondents to our survey have a better access to sustainable jobs, as these variables are correlated to labour market integration. Therefore, in the next section, we will decompose results for the whole sample and for young and low educated people, in order to see whether they all face the same barriers regarding training enrollment, and to verify the external validity.

#### 3.4.4. Training enrollment

As mentioned above, administrative data from *Pôle emploi* contain historized information on training programs completed by jobseekers. This enables us to study training enrollment and to explore whether the probability to join a training program differs across socio demographic characteristics. Table 3.8.1 displays the probability to join a training program within 12 months after registering at *Pôle emploi*. This table indicates that this probability is approximately 10%. This probability is lower for the *PIC* population, but is much higher for young jobseekers.

To deepen our analysis, we rely on machine learning methods to identify the population which is the most likely to join a training program within 12 months after registering at *Pôle emploi*. Using a random sample of individuals who registered at *Pôle emploi* between 2015 and 2016, we build a predictive model on 75% of the whole database, and we test it on the remaining sample to avoid over fitting. The model we use to predict entry to training is the Random Forest, from (Breiman 2001). The associated algorithm randomly draws observations as well as predictors associated with covariates. At each step, the algorithm only keeps the most efficient predictors and starts again. This model does not have an econometric interpretation. However, it provides the relative importance of each variable without,

---

<sup>13</sup>These percentages are approximately the same among job seekers registered at *Pôle emploi*

however, indicating the precise link between this variable and the predicted one. Results are shown in Figure 3.9.1: the most important variables involved in the prediction of training enrollment are the age of the jobseeker, her number of experience in her previous job (variable “exper”), or the number of children (variable “nenf”). Finally, Figure 3.9.2 displays the distribution of the predicted probability to join a training program. In line with results from Table 3.8.1, we see that the spontaneous training enrollment probability is low.

## 3.5. Results

In this section, we review our survey results. In particular, using jobseekers answers, we identify the main barriers they face when considering the entry into training programs. To this end, we distinguish between external and internal barriers. We also explore heterogeneity in responses between the *PIC* population and the whole sample, as our sampling process allows us to statistically compare responses for young and low educated people on the one hand, and for the whole sample on the other hand. The diagnosis we make in this section will then enable us to design interventions aimed at lessening the obstacles we identify.

### 3.5.1. External barriers

#### Informational barriers

The first barrier we explore in our survey relates to information access. Indeed, and as explained in the introduction, the lack of information about the available programs or the operational steps to join a training program may severely affect enrollment (see Babcock et al. (2012), Bertrand et al. (2004)). Figure 3.9.3 displays the proportion of jobseekers feeling well informed about training program in general (question 15), as well as the proportion knowing how to find information about training centers, training characteristics (speakers, number of hours, eventual entrance tests, etc) and registration steps (question 16). Results are presented for the whole sample (in green) and the *PIC* population (in gray). We observe that about half of jobseekers lack information about training programs, centers, and registration steps and do not know where to find it. These percentages do not statistically differ for the *PIC* and non *PIC* population, except for the registration steps, for which the *PIC* population reports being slightly better informed.<sup>14</sup> Thus, information can be a first obstacle to training enrollment, as not only jobseekers lack information but do not know how to access it.

#### Barriers related to supply and quality

The perceived quality of training may also affect training enrollment: if jobseekers believe that existing programs do not fit their needs, because they are of low quality in terms of content, speakers or reemployment opportunities, they are unlikely to consider it as an option to improve their labor market outcomes. It may also be the case that they think that training programs are of good quality overall, but that the supply is not sufficient or too far from their home. Figure 3.9.4 gives some insight about these two

---

<sup>14</sup>We performed a ttest comparing percentages in for the *PIC* population and the non *PIC* population.

dimensions (questions 18 and 19). Among respondents, roughly 75% think that training programs are of good quality: more than 60% think that it could help them finding a job, more than 80% think that speakers are competent while three quarters think that administrative procedures are worth taking in view of the benefits obtained thanks to training. However, only 15% think that there are enough training programs corresponding to their needs and nearly 50% of our respondents considers that training supply is not close enough to their home. Here again, these percentages are approximately the same for the *PIC* population and non *PIC* population. Thus, the perceived quality of training program does not seem to be a major issue when considering joining a training program, contrary to jobseekers opinion about the availability of training supply which may lower the entry into training.

### **Barriers related to training returns**

As mentioned in the introduction, the perceived returns of education is an important determinant of human capital investment (Jensen 2010). For this reason, we pay special attention to beliefs jobseekers have about training returns: if they think that training programs do not lead to better labor market outcomes, they are not likely to enroll. To this end, we distinguish between refresher training, long and job specialized training, and the combination of both. We also distinguish between sector of activity, as these variables are likely to affect the actual training returns. Thus, we first look at beliefs jobs seekers have about the increase in employment access one can reach through training programs. To that end, we collect their opinion about the difference in percentage between the share of jobseekers finding a sustainable job within three months after a training program on the one hand, and after registering at *Pôle emploi* on the other hand (see questions 6 and 11 to 13 from our survey). This difference does not actually measure beliefs about training returns per se, as (i) we do not have jobseekers beliefs about the percentage of people finding a sustainable job *without a training program* (but instead *whether they completed or not a training program*), and (ii) people joining a training program and those who do not are not likely to be statistically identical. However, comparing then this difference with the empirical one calculated through historical data from *Pôle emploi* gives us a measure of jobseekers optimism towards training returns: if the difference between jobseekers beliefs and actual data is positive, they are optimistic about training returns, but if it is negative, they are rather pessimistic. Figure 3.9.5 shows the distribution of this double difference (*i.e* the distribution of the gap between jobseekers beliefs and actual data) when we consider long and job specialized training program (Figure 3.9.5 (a)), refresher training program (Figure 3.9.5 (b)) and for the combination of both (Figure 3.9.5 (c)). For each type of training we compare the distribution of the *PIC* population (in gray) and the whole sample (in green). Overall, the distribution of gaps is centered around 0 and is slightly right skewed, except for long and job specialized training programs, for which we observe a peak for gaps being slightly less than 0 (see Figure 3.9.5 (a)). Thus, training returns do not seem to be a major barrier preventing jobseekers from enrolling into training programs. We deepen our analysis exploring jobseekers beliefs about reemployment probability after registering at *Pôle emploi* and after a training program separately. Indeed, despite the fact that jobseekers do not seem to be particularly pessimistic about training returns, one would like to know to which extent this result is driven by their beliefs about reemployment probability after registering



at *Pôle emploi*, or reemployment probability after training. Table 3.8.4 shows for each sector of activity and each type of training jobseekers beliefs about the percentage of people finding a sustainable job after completing a training and after the registration at *Pôle emploi*, as well as the associated empirical values taken from *Pôle emploi* databases. Results suggest that jobseekers tend to overestimate the proportion of people finding a sustainable either after completing a training and after registering at *Pôle emploi* which is in line with the academic literature (see Spinnewijn (2015)). In the end, training returns do not seem to be an obstacle jobseekers face when considering training enrollment, which is consistent with the results we found above about the perceived quality of training programs.

### Barriers related to training funding

The last external barrier we look at relates to financial constraints. Training funding encompasses various dimensions such as registration fees (direct costs), indirect costs (transportation costs, etc.) but also the opportunity cost of pursuing a training course instead of working. The latest relates to training returns, which we explore above. Hence, we focus here on the first two dimensions, as both the amount of registrations fees and additional costs are likely to affect the decision to participate in a training program. To that end, we first look at the importance jobseekers attach to direct and indirect costs when considering training enrollment. Figure 3.9.6 indicates that for only 20% of jobseekers, indirect cost would prevent them from joining a training program. However, regarding registration fees, nearly two thirds of the whole sample say that they would not join a training program if they did not obtain a grant for it. This proportion is significantly lower for the *PIC* population (in gray on figure 3.9.6). Overall, these results suggest that contrary to expenses due to transportation or child caring, registrations fees matters a lot in the participation decision and may prevent jobseekers from enrolling.

To pursue our analysis, we then examine beliefs jobseekers have about the proportion of training programs whose registration fees are already covered through public funds (see question 14), as our results show that financial support is an important determinant of entry into training. As mentioned in subsection 3.2.2, regions and *Pôle emploi* buy each year training seats according to firms' recruitment intention. These training programs are listed in a catalogue host by *Pôle emploi*, where jobseekers can freely look at. If jobseekers underestimate the share of programs which are available for free, they should be less likely to apply for training, as the application procedure for individual grant is much longer, complicated, and uncertain. In addition, and the amount of registration fees is hard to sustain individually.<sup>15</sup> On the contrary, if they overestimate this proportion, it should not affect their likelihood to enroll. Figure 3.9.7 displays the distribution of beliefs jobseekers have about the percentage of subsidized programs. While 100% of training courses listed in the *Pôle emploi* catalogue are free for jobseekers (irrespective of the sector of activity), results reveal that all respondents (in green) as well as the *PIC* population (in gray) highly underestimate this proportion: 60% of respondents think that this proportion is less than 40%, and less than 10% think that this proportion is higher than 80%. In light of these results, we deepen our analysis, making the assumption that jobseekers may have misunderstood our question (question 14), and have considered the share of free training among whole the set of training that exists,

---

<sup>15</sup>Registration fees for programs targeted under the *PIC* typically range from 1000 to 6500 euros.



instead of among those listed in the *Pôle emploi* catalogue. For this reason, we use data from the website *defi-métiers*, which host the list of all training programs available in the region Ile-de-France. This list is intended to be exhaustive, and indicates which training are subsidized and which ones are not. We thus restricted our sample to jobseekers living in Ile-de-France, and find that more than half of jobseekers are still pessimistic about the share of funded training programs, at least in Ile-de-France (see figure 3.9.8). Among those, the difference between the estimated percentage and the actual values reach 24 percentage point on average. Finally, to confirm these findings, we checked in administrative data from *Pôle emploi* the proportion of jobseekers who completed a training program through an individual grant whereas a public grant was available for the same job.<sup>16</sup> We find that this percentage reaches 15%.<sup>17</sup> All these results tend to confirm the fact that jobseekers lack information about funding possibilities, which in line with our findings in the information module: nearly two thirds of jobseekers say that they lack information about funding opportunities and how to access them (see question 16).

Finally, results from question 27 indicate that approximately 60% of jobseekers believe that they cannot get financed for the direct costs of training. Given the above results, this belief is not rational, as jobseekers are highly pessimistic about the proportion of training places in their sector that are funded.

### Conclusion about external barriers

We end our analysis of external barriers by examining the relative importance of each external barrier (information, quality, returns and funding) preventing jobseekers from enrolling into a training program. Figure 3.9.9 shows that the most frequently cited constraint by all respondents (in green) as well as the *PIC* population (in gray) is the financial one: 30% of jobseekers point out financial constraints, while all other barriers get 12% or less. We have also seen throughout our survey results that jobseekers lack information about the registration steps and how to access basic information about training characteristics. Therefore delivering information about funding and registration procedure may lessen the importance of these external barriers and may favor training participation.

### 3.5.2. Internal barriers

In this subsection, we are interested in the internal barriers jobseekers face to when entering training. Contrary to external barriers, internal barriers are intrinsically linked to their personal traits. We are more specifically interested in self-efficacy, executive function, and self-esteem, and time preference, as their impact on the decision making process has been largely documented in the literature (see [Benabou and Tirole \(2002\)](#), [DellaVigna and Paserman \(2005a\)](#), or [Caliendo et al. \(2020\)](#) for instance). Figure 3.9.10 explores the relative importance of each of these barriers (see question 33), for both the the whole

<sup>16</sup>We performed this analysis considering training starting between the January 2019 and April 2019, as our data on training seats purchased by *Pôle emploi* only contained information within this interval.

<sup>17</sup>We matched training on both the *formacode* (which is a numerical code corresponding to a restricted set of jobs), and the region. Thus, we assume that training courses in a given formacode are very similar and are (imperfect) substitutes for one to another: if one is funded, people move to that one. Regarding the matching on the region, we are aware that regions can be very big, and if in a given region, a training grant is available in one city which is far from the job seeker's residence area, she will probably not choose this training. We thus run this analysis on very small regions separately (Corse and DOM-TOM), where transportation costs should be low in terms of money and time, and except for the region Ile-de-France, we find approximately the same percentage.

(in green) sample and the *PIC* population (in gray). We find that for each of these two populations, the proportion of jobseekers pointing out each barrier is approximately the same across self-efficacy, executive function, and time preference (self-esteem is the less cited issue). These results are consistent with jobseekers answers on depression (question 23), self-efficacy (question 24) and self-esteem (question 25): indeed, for each of these variables, we constructed an index ranging from 1 (very low) to 4 (very high). The detailed construction of these indices is shown in appendix .1. We find that both for the whole sample and the *PIC*, the index is above 3 for a large majority of people, which suggests that for at least these three variables, jobseekers have no more difficulties on one barrier than for the others (see figures .1.1 to .1.3 presented in appendix .1).

Overall, jobseekers do not report experiencing major issues related to depression, executive function and self-efficacy, and when they indicate experiencing external barriers, executive function and self-efficacy, self-esteem and are equally pointed out.

Finally, our results suggest that external and internal barriers are additive. Table 3.8.5 shows the distribution of jobseekers according to whether or not they report having external or internal barriers. We observe that among people who claim no external problem, almost everybody (90.81%) also claims no internal problem. However, of all people who do have an external problem, about half also claim an internal problem. Finally, almost nobody (9.19%) claims only internal problems. However, admitting to be held back purely by internal barriers may cause self-image concerns, which means that participants in this category may underreport more than others. Therefore, designing interventions hitting beliefs about the self can still be a way to improve training participation.

## 3.6. Outlook

The goal of this section is to propose an experimental design aimed at targeting the main external and internal beliefs jobseekers have to deal with when they want to join a training program. We first review the interventions we want to run and the hypotheses they rely on. We then detail our experimental design and the equations we will estimate. Finally, we end with a discussion about heterogeneity and power calculations.

### 3.6.1. The interventions

Results from our survey suggest that both the *PIC* and non *PIC* population underestimate the proportion of subsidized training available to them. In addition, they lack operational information to register into training programs. In the light of this diagnostic, we will run interventions aimed at modifying external beliefs about funding, and increasing self-efficacy. We choose to focus on this internal belief because beliefs are more malleable than preferences and skills (see Haushofer et al. (2020)), and because self-efficacy is equally cited by jobseekers reporting internal barriers. Therefore, interventions hitting this internal belief are more likely to have an effect than interventions targeting other internal beliefs such as self-control, executive function or self-esteem. Our final purpose is then to test the impact

of these interventions on the entry into training and to derive some policy recommendations to favor enrollment.

### 3.6.1.1. Treatments

As a first stage, a randomly selected sample of jobseekers will be invited to attend an online collective meeting with traditional, generic content about training. Job seekers will attend these meetings in groups of 20. At the end of the meeting, jobseekers will be introduced to a specific interactive app-based course taking place in the 6 weeks following the meeting, and whose content will depend on the random treatment assignment of the meeting:

- **External beliefs treatment:** jobseekers assigned to this group will be introduced to an app-based course on overcoming external barriers (specifically, how to obtain financing for training and complete administrative procedures). The course on external barriers will follow the same structure of weekly tasks and interchange with peers, but it will be focused on overcoming financial and administrative hurdles: how to find information about financed training, how to complete the administrative process, etc. The combined course will feature elements from both individual courses, comprising both tasks to build self-efficacy, and on overcoming financial and administrative hurdles.
- **Internal belief treatment:** jobseekers assigned to this treatment will be introduced to an app-based course on overcoming internal barriers (specifically, on building self-efficacy).
- **Combined treatment:** jobseekers assigned to this treatment group will be introduced to an app-based course on both external and internal barriers
- **Placebo group:** finally, jobseekers assigned to this group will be introduced to an app-based “placebo course” without psychologically active content (e.g., generic information on training).

As a second stage, jobseekers will then interact with the app in the following six weeks. They will receive specific tasks each week, depending on their treatment assignment.

The course on internal beliefs will largely follow the Generalized Self-Efficacy (GSE) curriculum developed by McKelway (2020): jobseekers will reflect on their own strengths and achievements, and how their strengths have led them to accomplish things in their life so far. The app will ask them questions each week, and jobseekers share and discuss their answers in an interactive forum with other jobseekers from their meeting unit. Over six weeks, the jobseekers build up the belief in their own capacity to achieve desirable outcomes and overcome obstacles (Bandura 1977).

The course on external beliefs will follow the same structure of weekly tasks and interchange with peers, but it will be focused on overcoming financial and administrative hurdles: how to find information about financed training, how to complete the administrative process, etc.

The combined course will feature elements from both individual courses, comprising both tasks to build self-efficacy, and on overcoming financial and administrative hurdles.

Finally, the placebo course will rely on generic training material provided by *Pôle emploi*. Interactive features (weekly tasks and exchange with peers) will focus on generic exchanges about their experiences with different training programs, without guidance on specific barriers.

### 3.6.1.2. Hypotheses

Our hypotheses are well described by our theory of change. We hypothesize that there will be several impacts associated with our interventions:

1. Knowledge jobseekers have about entry into training and about its financial aspects should improve for those receiving the external beliefs treatment or the combined treatment. Similarly, jobseekers' self-efficacy should be magnified due to the participation to the internal belief intervention or the combined intervention.
2. A larger share of jobseekers should plan to register into training programs.
3. A larger share of jobseekers should have taken concrete steps toward registration into training programs
4. Finally, a larger share of jobseekers should register into training programs.

Our main outcome variable will be the entry into training programs. However the theory of change behind the interventions implies additional steps. First, jobseekers are supposed to have improved their knowledge and skills along the dimensions related to the interventions. Second, jobseekers should have planned to register into training. Third, they should have taken concrete steps in order to join a training program, and last they should have enrolled. We will collect data that will enable us to have information along all these dimensions. To this end, we will run a survey shortly after the interventions have been completed. To avoid as much as possible issues related to low response rate, we will design a very short survey asking a small number of questions and will randomly assign participants in the experiment so that they answer just one of these questions. Our main outcome variable, *i.e.* entry into training, will be measured using administrative data from *Pôle emploi*. A nice feature of this variable is that it is available for all our sample, and over several years.

### 3.6.2. Experimental design and randomization

Before assigning jobseekers to our treatments, we will select our population of interest according to several criteria: we will first sample jobseekers who never completed a training program since the beginning of their registration at *Pôle emploi*, as they are unlikely to be interested in taking up a vocational training program again. We will also sample jobseekers who agreed to receive news by emails and texts from *Pôle emploi*, as jobseekers will be invited to participate in collective meetings by email and text. Then, and as can be seen in Figure 3.9.11, jobseekers will be invited by groups of 50 to participate in an online meeting and will receive reminders from *Pôle emploi* to participate in these meetings. We will follow the same method as the one adopted in the experiment presented in chapter 1, aimed at estimating the impact of a website devoted to the provision of online advice to jobseekers. In

this experiment, we secured a registration rate on the job search website of 50% among jobseekers invited to attend a collective meeting. From this experience, we expect that inviting 50 jobseekers to collective meetings will lead to an average of 25 of them participating in the collective meeting. These jobseekers attending these collective meetings will form the sample of our experiment. The collective meeting group they will participate in will be jointly randomly assigned to each treatment. Given the experiment presented in chapter 1, we expect that 50% of jobseekers will participate in these interventions, which therefore corresponds to a take-up rate of 50%.

### 3.6.3. Main equation to consider

Our primary analysis will be based on the estimation of the following simple regression:

$$y = a + b_{internal}T_{internal} + b_{external}T_{external} + b_{both}T_{both} + u$$

where  $T_{internal}$  stands for the internal treatment variable,  $T_{external}$  stands for the external treatment variable and  $T_{both}$  stands for the combined treatment variable. The constant  $a$  corresponds to the intercept of the model and,  $u$  corresponds to the error term. Thanks to the randomization, the coefficients  $b_{internal}$ ,  $b_{external}$  and  $b_{both}$  associated with the treatment variables give the causal impact of our interventions. We will also consider, as frequently done, the introduction of covariates to improve the accuracy of estimates. To do that, we will use the rich datasets from *Pôle emploi* and implement the method proposed in [Belloni, Chernozhukov, and Hansen \(2014\)](#), which amounts to implement several Lasso steps to predict the treatment (in our case, there would be three Lasso for this step) and to predict the dependent variable. We will end up with the following regression:

$$y = a + b_{internal}T_{internal} + b_{external}T_{external} + b_{both}T_{both} + X(selected)^{\top}c + u$$

where  $X(selected)$  denotes the set of covariates selected following the methodology of [Belloni, Chernozhukov, and Hansen \(2014\)](#).

### 3.6.4. Heterogeneity

Heterogeneity of impacts is always significant to address; in our case, it is especially important as it enables us to inform the targeting of our interventions. Several aspects are worth discussing in this respect.

First, a key aspect of our paper consists of linking the heterogeneity of needs in the population with different interventions. Thanks to the survey, we have identified that some jobseekers face internal barriers while others encounter external barriers more, and we have designed several specific interventions in order to address those two types of problems. Consequently, it will be essential to show that these who are likely to face internal barriers are those for whom the internal-barrier intervention's impact is the largest, and the reverse for the external barrier. The interest of our analysis is to show there exists a mapping between different identified segments of the population and the intervention best-suited for that category

of jobseekers. To do so, we will conduct a short baseline questionnaire to be administered during the collective meeting to assess the strength of both internal and external barriers for each jobseeker included in the experiment. Then, we will examine the heterogeneity of impacts according to the measures of internal and external barriers' intensity as measured by the baseline survey. Formally, for a variable  $G \in \{0, 1\}$  to separate two sub-populations of jobseekers,<sup>18</sup> we will consider the regression:

$$y = a + b_{internal}T_{internal} + b_{external}T_{external} + b_{both}T_{both} \quad (3.1)$$

$$+ G(a_G + b_{G,internal}T_{internal} + b_{G,external}T_{external} + b_{G,both}T_{both}) + X(selected)^\top c + u.$$

Such a regression enables to measure the impacts of the three treatments for both categories of jobseekers,  $X = 0$  and  $X = 1$ , and thus to compare the effects for the two sub-populations and investigate potential heterogeneity of the effects.

A second important aspect relates to targeting and the propensity to enter training. Should such interventions be targeted at jobseekers who have a large propensity to enter training absent the intervention or at the other jobseekers? On the one hand, jobseekers with a large propensity to enter training might be more receptive to our interventions. On the other, they might merely not need the intervention to decide to enter training. To analyze the issue, we can estimate a slightly different version of equation (3.1), in which there is not just one heterogeneity group but two, corresponding to intermediate and large propensities to enter training.

### 3.6.5. Power calculations

In this subsection, we report the power calculations of our experiment. The Minimum Detectable Effect of the experiment (MDE), *i.e.*, the smallest effect associated with our treatments we will be statistically able to detect, is given by the following formula (see [Duflo et al. \(2007\)](#)):

$$MDE = (t_\alpha + t_{1-\pi}) \sqrt{\frac{2}{\Pi^2 K} \left( \rho + \frac{1-\rho}{n} \right) \sigma}$$

where  $K$  is the number of collective meetings organized in each treatment arm,  $\rho$  is the intra-cluster correlation of residuals, and  $n$  the number of individuals in each collective meeting.

To compute the MDE, we rely on assumptions about the value of the different parameters in the equation or direct measures. We need to compute the MDE for our final outcome in the theory of change, namely training enrollment in our experimental design. Indeed, we make the assumption that impacts are smaller and smaller throughout the theory of change. Therefore, our purpose is to have an experiment that enables us to detect impacts for the final step of the theory of change. Hence, we will be able to detect impacts on the other variables along the whole theory.

Computations from data available at *Pôle emploi* show that the propensity to enter training within a year is equal to 9.57% (from which we can easily compute the variance). We choose the parameter  $\rho$  to

<sup>18</sup>For instance, in our setting,  $G = 0$  (respectively 1) for jobseekers likely to face stronger external barriers than internal ones (resp. stronger internal barriers than external ones).

be either very small 0.01 or small 0.1. The number of individuals by collective meeting will be roughly 20 or 25 (we will document the rationale for such a small number) later. Last, based on the experiment on the online job search platform presented in chapter 1, we assume that the take-up  $\Pi$  is 50%.

Results are presented in figure 3.9.12. We first consider the case of a single test (MDE(20,1,x)). However, as there are three treatment arms in our experimental design, we also perform power computation accounting for multiple testing, with three simultaneous tests. When we perform tests at the 5% level controlling for the Family Wise Error Rate, we consider a threshold p-value that is not 5% anymore (as it was the case for one single treatment) 5/3% instead. On figure 3.9.12, those results appear as MDE(20,3,x). As can be seen, if the number of collective meetings in each arm is as large as 210, the MDE is 0.04. This is still large compared to the average in the control group (0.0957). It yields a total number of 4200 individuals per treatment arm. When reaching 410 collective meetings, the related MDE(20,1,x) decreases to approximately to 3%. These numbers are only slightly affected by the Bonferroni correction to account for multiple testing. However, as it is well-known, they are highly sensitive to the correlation parameter  $\rho$ .

### 3.7. Conclusion

This paper reviews the results of a survey we conducted on French jobseekers registered at *Pôle emploi*. The goal of this survey was to identify the obstacles they face regarding training enrollment. Distinguishing between “external” and “internal” beliefs, we find that jobseekers experience financial constraints as the main external barrier preventing them from joining a training program. In addition, our results suggest that jobseekers underestimate the proportion of subsidized programs available to them. We also find that half of jobseekers lack information about the registration process and training programs in general. Regarding internal beliefs, we find that obstacles related either to self-efficacy, self-control, self-esteem and executive function are equally mentioned among jobseekers reporting internal barriers in training enrollment. Finally, our last finding is that external and internal barriers seem additive: we find that of all people reporting external problem, about half also claim an internal problem while almost nobody claims only internal problems. In the last section of this paper, we use these results to inform the design of a randomized controlled trial. Drawing upon McKelway (2018), the interventions we design are based on the delivery of information through app-based courses, with interactive sessions involving groups of jobseekers. These courses aim at targeting either external or internal beliefs, or both of them simultaneously. This design will enable us to compare the impact of each intervention in absolute and relative terms.

### 3.8. Tables

Table 3.8.1: Probability to enroll into training 12 months after registering at *Pôle emploi*

	<b>Probability</b>
<b>PIC population</b>	<b>9.10%</b>
<b>Non PIC population</b>	<b>10.42%</b>
<b>Less than 26</b>	<b>11.13%</b>
<b>Above 26</b>	<b>8.78%</b>
<b>All categories</b>	<b>9.57%</b>

*Notes:* analysis conducted on 2,652,548 observations

Table 3.8.2: Distribution of strata within the sample

	Young	Not young	Total
Low or unskilled jobseekers	40%	20%	60%
Medium or high skilled jobseekers	20%	20%	40%
Total	60%	40%	100%



Table 3.8.3: Summary Statistics

	All		Respondents		Non respondents		Difference [5-3]
	Mean 1	SD 2	Mean 3	SD 4	Mean 5	SD 6	
Female (%)	50.46	50.00	59.29	49.13	49.60	50.00	9.69***
Age	29.94	11.76	34.87	13.57	29.46	11.45	5.41***
Married (%)	26.34	44.05	33.80	47.31	25.62	43.65	8.18***
Number of children	0.49	0.96	0.65	1.11	0.47	0.95	0.18***
Level of education below the high school degree (%)	60.00	48.99	50.84	50.00	60.88	48.80	-10.05***
Unemployment duration (in months)	19.84	24.67	22.57	27.59	19.58	24.35	2.99***
<i>Assistance track</i>							
Low (%)	18.57	38.89	19.27	39.45	18.51	38.83	0.77*
Moderate (%)	37.16	48.32	36.56	48.16	37.21	48.34	-0.65
Intensive (%)	23.83	42.61	21.44	41.05	24.07	42.75	-2.62***
<i>Sector of activity</i>							
Social fields (%)	16.06	36.71	18.03	38.45	15.86	36.53	2.17***
Agriculture, livestock and fisheries	2.60	15.90	1.79	13.26	2.68	16.14	-0.89***
Agrifood business (%)	1.92	13.72	1.15	10.68	2.00	13.99	-0.84***
Arts and crafts (%)	5.00	21.79	3.99	19.57	5.10	21.99	-1.11***
Audiovisual, communication (%)	2.63	16.02	3.10	17.33	2.59	15.88	0.51***
Building and construction (%)	6.92	25.38	5.18	22.16	7.09	25.67	-1.92***
Trade (%)	16.21	36.86	16.83	37.42	16.15	36.80	0.68
Electronics and energy (%)	2.44	15.42	1.96	13.86	2.49	15.57	-0.53***
Teaching, Sciences and Language (%)	2.19	14.65	2.85	16.63	2.13	14.44	0.72***
Management (%)	3.76	19.01	5.62	23.04	3.57	18.56	2.05***
Hotels, catering, tourism (%)	10.31	30.41	9.09	28.75	10.43	30.57	-1.34***
IT (%)	1.61	12.57	1.75	13.13	1.59	12.52	0.16
Handling (%)	6.72	25.04	5.47	22.73	6.85	25.26	-1.38***
Engineering and quality (%)	5.93	23.62	4.60	20.95	6.06	23.86	-1.46***
Health (%)	4.95	21.69	4.14	19.93	5.03	21.85	-0.88***
Secretariats and assistance (%)	7.25	25.94	11.22	31.56	6.86	25.28	4.36***
Transports (%)	3.51	18.40	3.22	17.65	3.54	18.47	-0.32
N =	100041		8842		91199		
			8.84%		91.16%		

*Notes:* This table presents summary statistics for socio-demographic variables of our sample of job seekers. Column 1 displays the mean of variables for the entire sample while column 2 shows the standard deviation. Column 3 et 4 on the one hand, and 5 and 6 on the other hand displays the same information, but respectively on the sub-sample of respondents and non respondents. The last column shows the average difference in means between respondents and non respondents, *i.e.* the coefficient we obtain by regressing the complete response dummy on the covariate of the row. Stars reflect the significance of this difference with robust standard errors. \*\*\* significance at the 1% level, \*\* significance at the 5% level, \* significance at the 10% level.

Table 3.8.4: Returns : comparison of average beliefs per sector with true probabilities

Sector	After registration		After pic training		After basic training		After combined training		% of job seekers
	True	Belief	True	Belief	True	Belief	True	Belief	
All sectors	6.28	23.06	16.86	33.8	8.5	30.45	15.1	35.6	100
1	4.26	22.35	16.1	33.61	8.5	29.54	13.24	35.67	.12
2	3.41	20.05	8.34	29.66	8.5	27.59	12.28	32.31	.03
3	8.11	22.53	18.89	31.53	8.5	31.07	20.83	34.84	.01
4	5.63	18.54	22.56	32.46	8.5	32.87	35.42	37.85	.05
5	6.4	18.6	12.97	34.04	8.5	32.21	14.81	38.84	.04
6	5.04	24.36	13.49	32.96	8.5	30.06	7.88	34.6	.05
7	7.21	22.63	17.91	31.79	8.5	29.68	13.95	33.72	.15
8	5.83	26.21	17.06	39.95	8.5	34.39	8.25	39.76	.01
9	7.61	21.04	10.48	32.99	8.5	28.89	6.56	37.47	.03
10	9.59	24.18	21.87	37.48	8.5	33.67	17.77	39.72	.03
11	7.83	23.95	17.9	33.21	8.5	29.92	16.8	34.88	.08
12	10	29.59	23.77	39.65	8.5	31.75	21.28	40.29	.03
13	4.03	23.42	8.69	31.49	8.5	27.97	5.15	30.95	.04
14	5.88	23.55	18.94	34.8	8.5	30.47	25	35.97	.06
15	5.8	27.2	16.36	37.6	8.5	32.56	9.46	36.84	.08
16	6.25	21.19	15.7	32.2	8.5	29.28	12.96	34.56	.13
17	6.99	26.21	25.12	38.39	8.5	33.24	23.81	37.47	.05

Notes: Analysis conducted on 12,948 observations.

This table shows for each sector of activity and each type of training jobseekers beliefs about the percentage of people finding a sustainable job after completing a training and after the registration at *Pôle emploi*, as well as the associated empirical values taken from *Pôle emploi* databases.

Table 3.8.5: Distribution of jobseekers across barriers

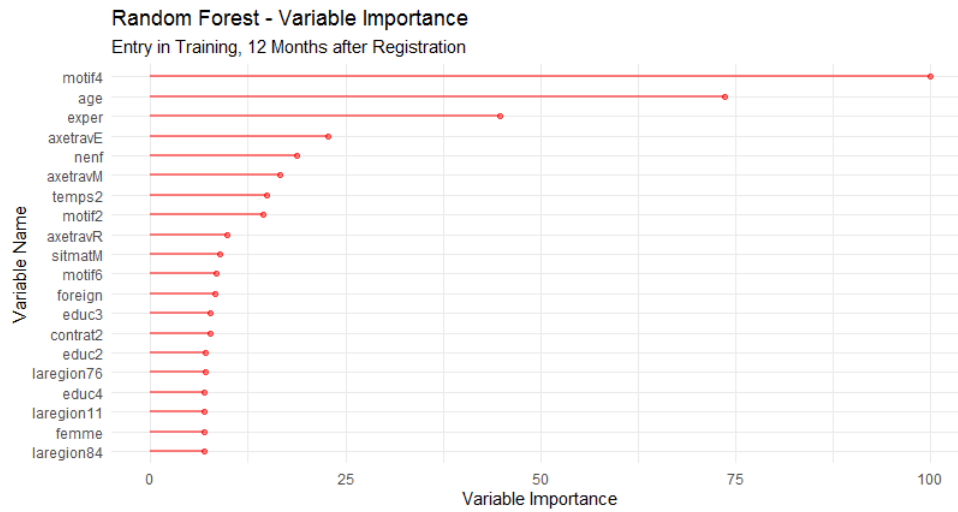
	Internal barriers	No internal barriers	Total
External barriers	52.25%	47.75%	42.09%
No external barriers	9.19%	90.81%	57.91%
Total	34.12%	65.88%	100%

Notes: analysis conducted on 8,842 observations.

Reading : the proportion of job seekers reporting no external barriers is 57.91%. The proportion of job seekers reporting no internal barriers is 34.12%. Among jobseekers reporting no external barriers, 9.19% report internal barriers.

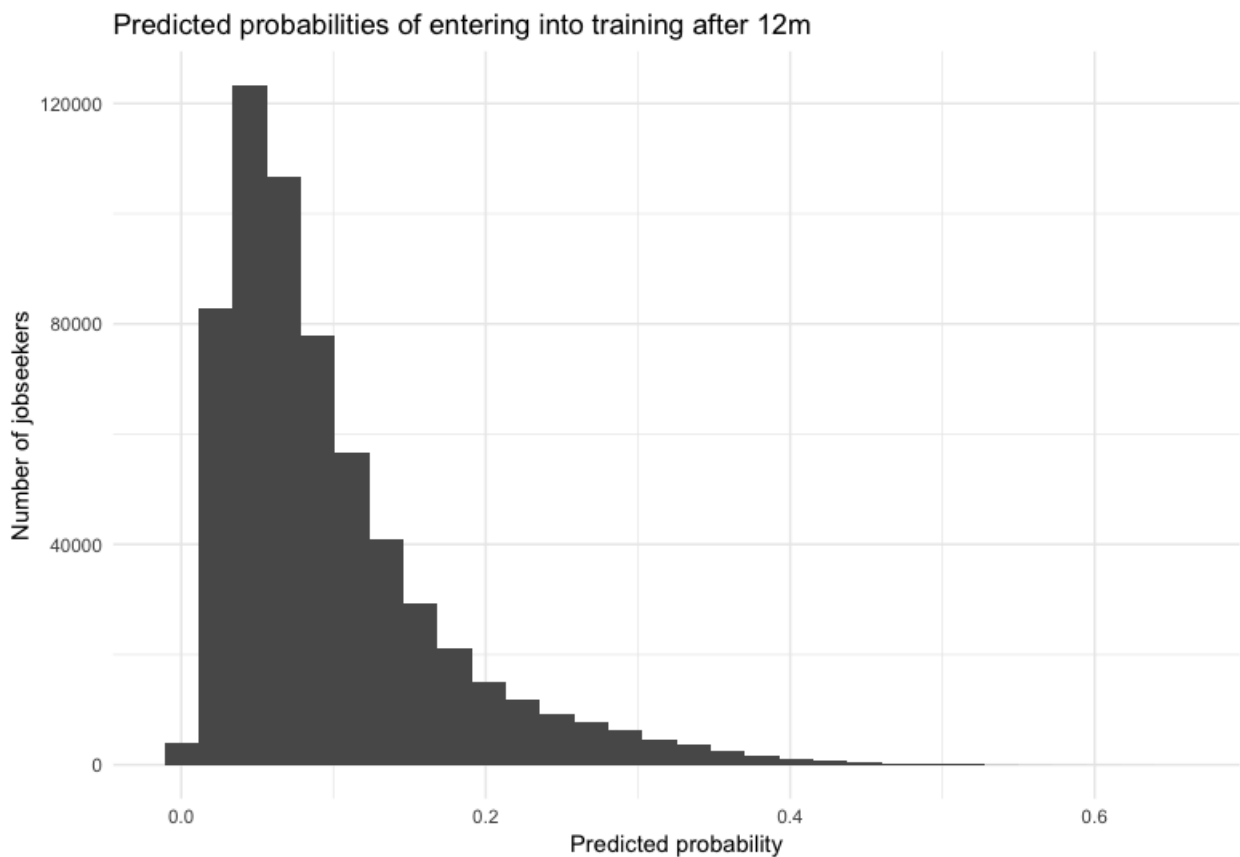
### 3.9. Figures

Figure 3.9.1: Random forest - variable importance



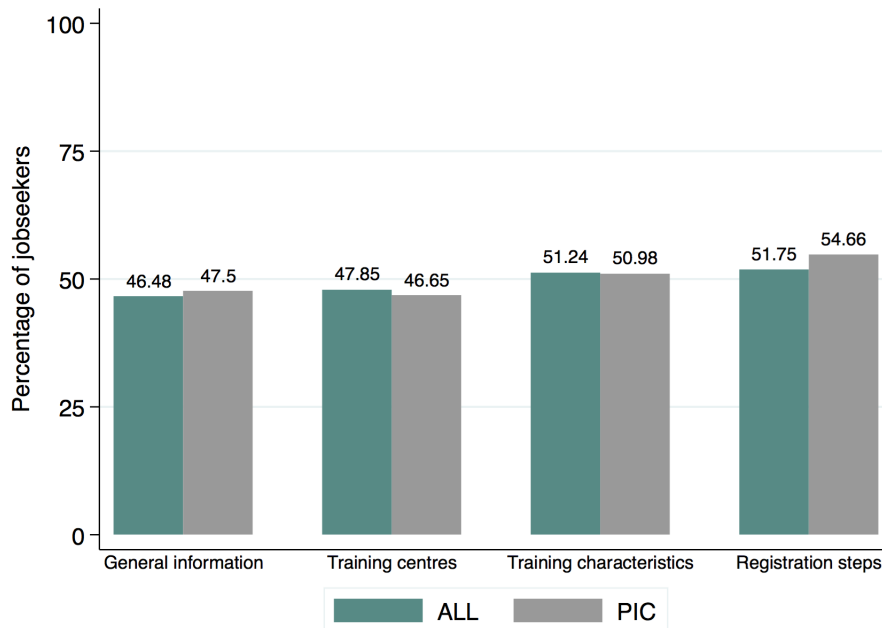
Notes: analysis conducted on 2,652,548 observations.

Figure 3.9.2: Predicted probability of training enrollment



Notes: analysis conducted on 2,652,548 observations.

Figure 3.9.3: Informational constraints

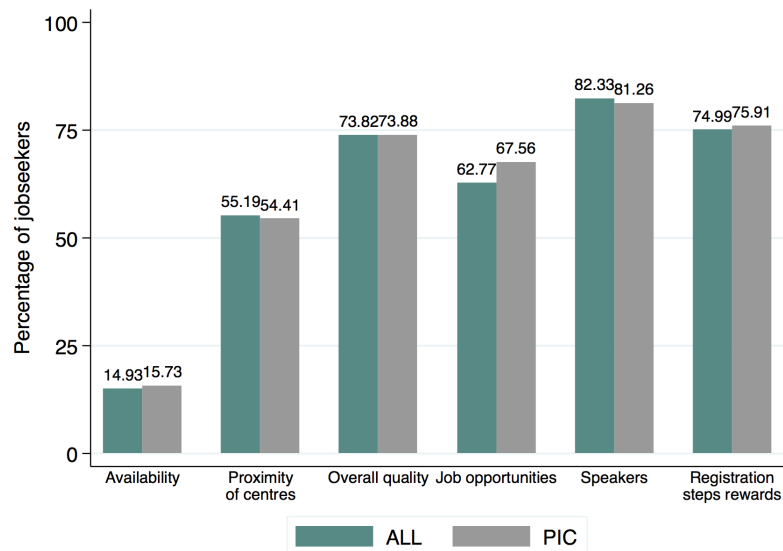


Notes: analysis conducted on 11,034 observations.

This figure shows the percentage of jobseekers reporting being well informed about respectively training programs in general, training centers, training characteristics, and the registration procedure to enroll.

The variable on general information is built from question 15, and is equal to 1 if the jobseeker answers one of the two first items (“Oui, tout à fait”, or “Suffisamment”), 0 otherwise. The other dummy variables are built on statements from question 16, and are equal to 1 if the jobseeker indicates that she mostly or strongly agrees, 0 otherwise.

Figure 3.9.4: Beliefs about quality



Notes: analysis conducted 10,636 on observations.

This figure shows the percentage of jobseekers having a good opinion about respectively the availability of training programs, the proximity of training centers, the overall quality of training programs, their reemployment opportunities they lead to, and the quality of speakers. The last variable corresponds to the proportion of jobseekers considering that the registration steps to join a training program are worth taking.

The first dummy variable on the amount of available training programs is built from question 17, and is equal to 1 if the jobseeker answers the first item (“Oui, il y en a assez”), 0 otherwise. The other dummy variables are built on statements from question 18, and are equal to 1 if the jobseeker indicates that she mostly or strongly agrees, 0 otherwise.

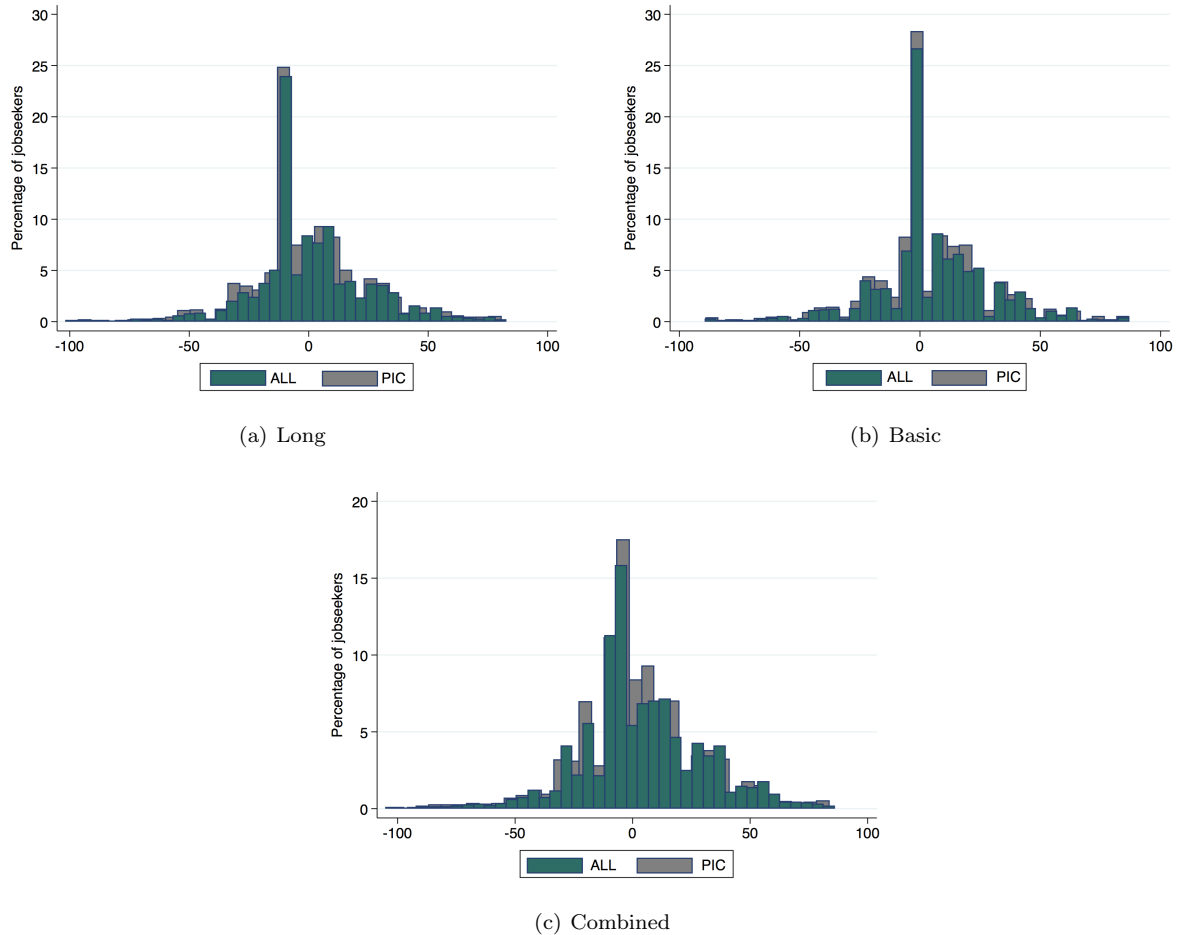


Figure 3.9.5: Training returns optimism

*Notes:* analysis conducted on 12,948 observations.

This figure shows the distribution of the difference between jobseekers beliefs and actual data about training returns.

Figure (a) displays this difference for long and job specialized training programs and is computed as follows :

$$[\text{belief}(\text{reemployment probability after completing a training program}) - \text{belief}(\text{reemployment probability with or without having completed a training program})] - [\text{data}(\text{reemployment probability after completing a training program}) - \text{data}(\text{reemployment probability with or without having completed a training program})]$$

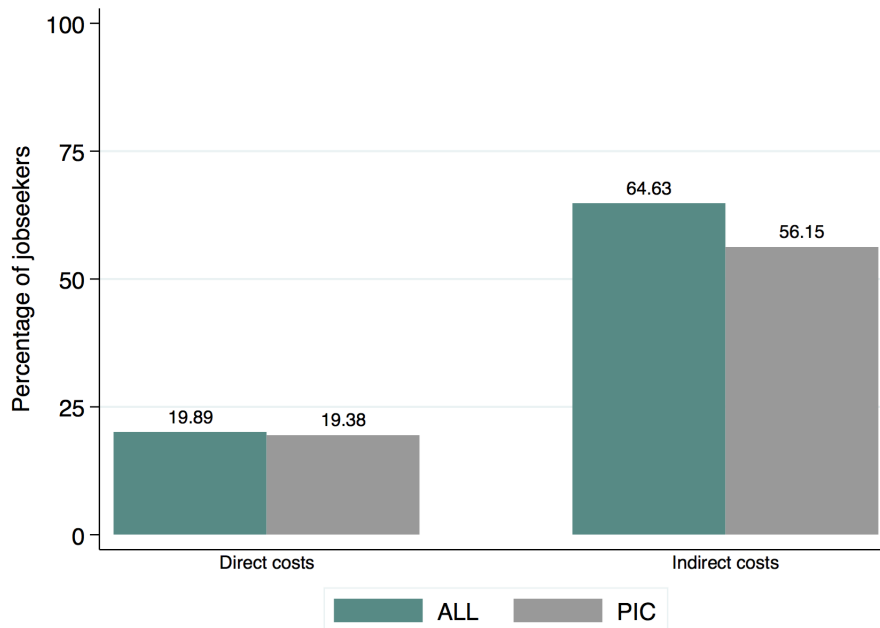
Actual data about reemployment probability come from *Pôle emploi* databases. Data about beliefs come from jobseekers answers to our survey : beliefs about reemployment probability after long and job specialized training programs come from question 11. Beliefs about reemployment probability with or without training come from question 6.

Figure (b) displays this difference for refresher training programs. Beliefs about reemployment probability after these training programs come from question 12. Beliefs about reemployment probability with or without training come from question 6.

Finally, figure (c) displays this difference for both the combination of refresher training programs and long and job specialized training programs. Beliefs about reemployment probability after these training programs come from question 13. Beliefs about reemployment probability with or without training come from question 6.

In our survey, we asked jobseekers to estimate reemployment probabilities within percentage intervals. Therefore, to be able to compare their beliefs with actual values from *Pôle emploi* databases, we took for each belief variable the median of the interval.

Figure 3.9.6: Financial constraints

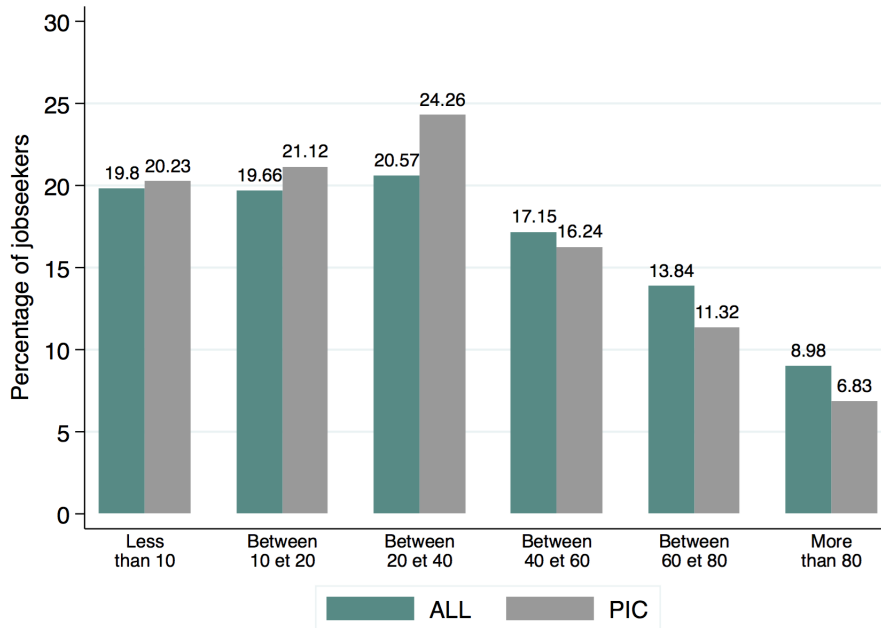


Notes: analysis conducted on 8,931 observations.

This figure shows the percentage of jobseekers for which direct costs or indirect costs would prevent them from enrolling into a training program.

The first dummy variable on direct costs is built from question 29, and is equal to 1 if the jobseeker answers the first item (“Indispensable, sans quoi je ne la suivrai pas”), 0 otherwise. The second dummy variable on indirect cost is built from question 30, and is equal to 1 if the jobseeker answers the second item (“Non, même en ne payant pas les frais d’inscription, je ne peux pas financer les frais annexes (coûts de transport, restauration, etc.”)).

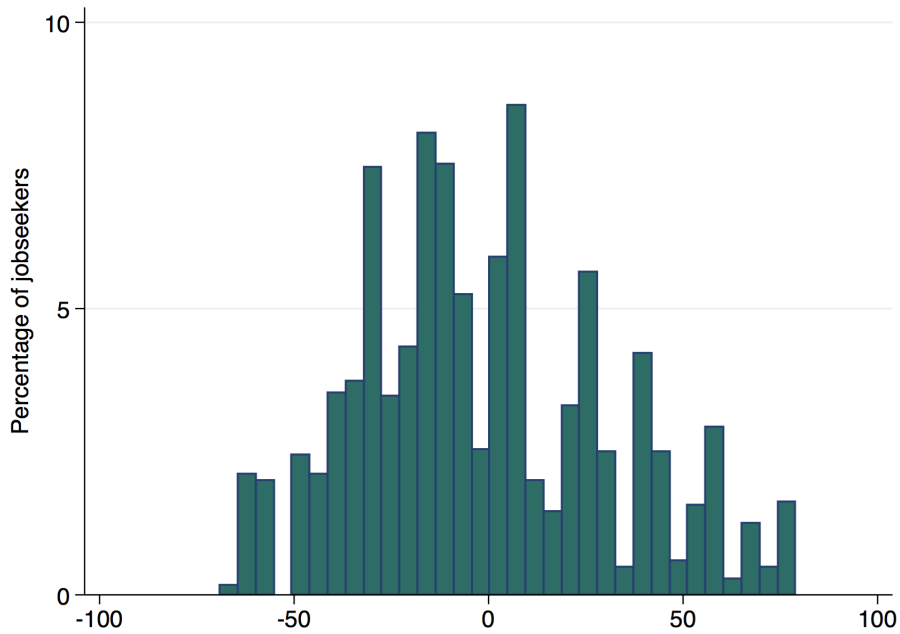
Figure 3.9.7: Beliefs about the share of subsidized training in the *Pôle emploi* catalogue



Notes: analysis conducted on 12,063 observations.

This figure shows the distribution of jobseekers beliefs about the percentage of subsidized programs. This variable comes from question 14 in our survey.

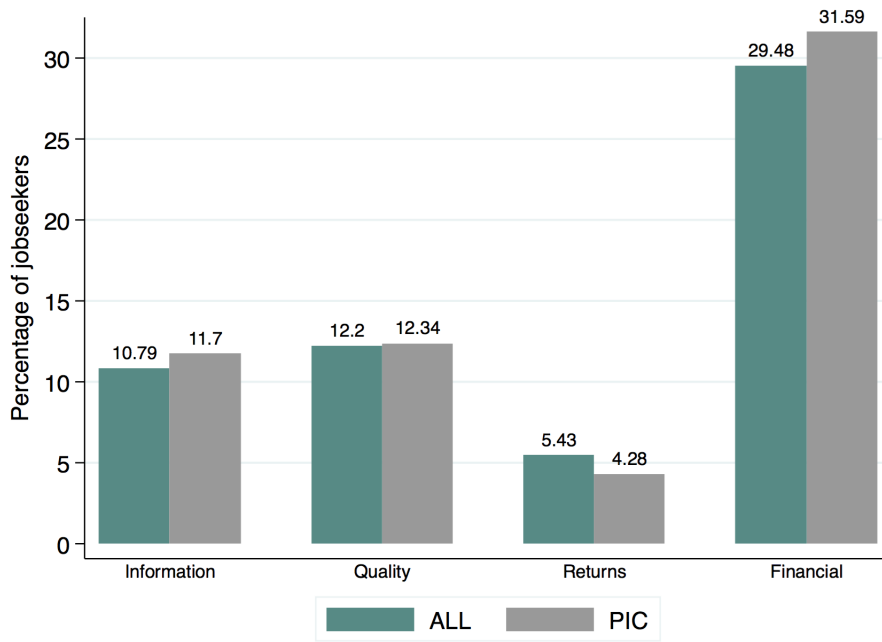
Figure 3.9.8: Difference between the estimated proportion of subsidized training programs and the actual share in the region Ile-de-France



Notes: analysis conducted on 1,847 observations.



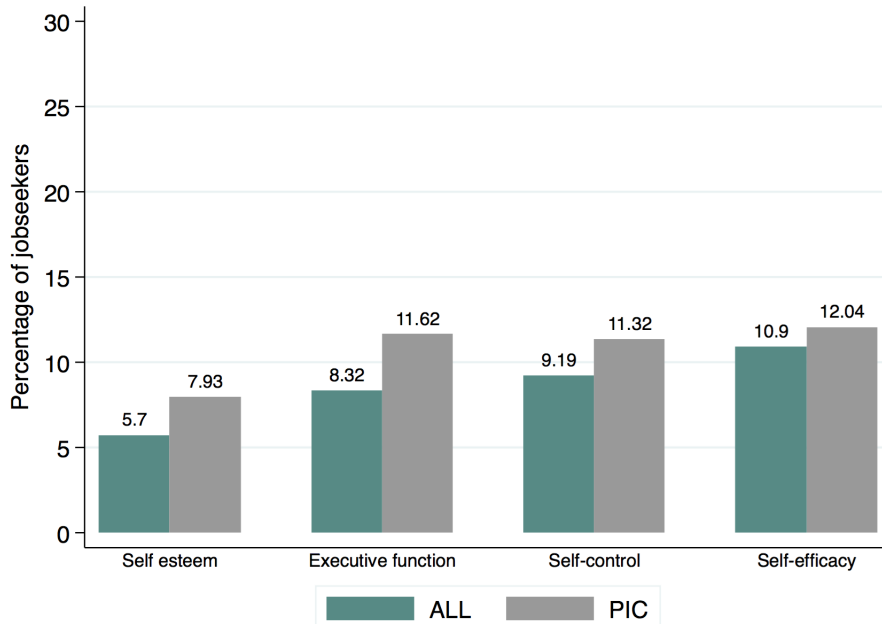
Figure 3.9.9: Main external barrier face by jobseekers



Notes: analysis conducted on 8,842 observations.

The distribution of the main external barrier faced by jobseekers comes from question 34.

Figure 3.9.10: Main internal barrier faced by jobseekers



Notes: analysis conducted on 8,842 observations.

The distribution of the main internal barrier faced by jobseekers comes from question 33.

Figure 3.9.11: Design of the experiment

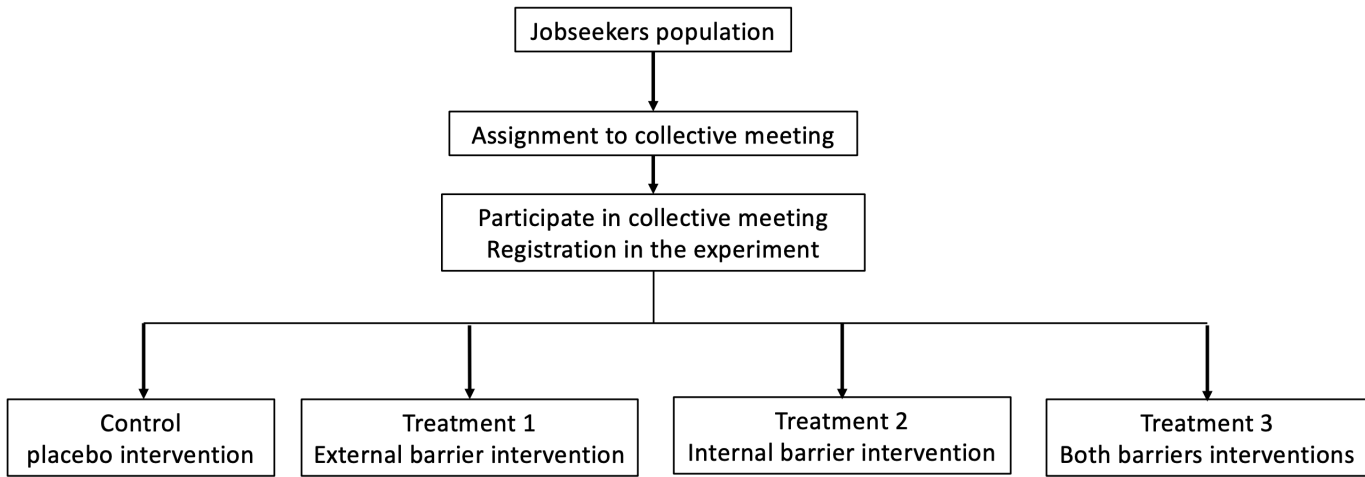
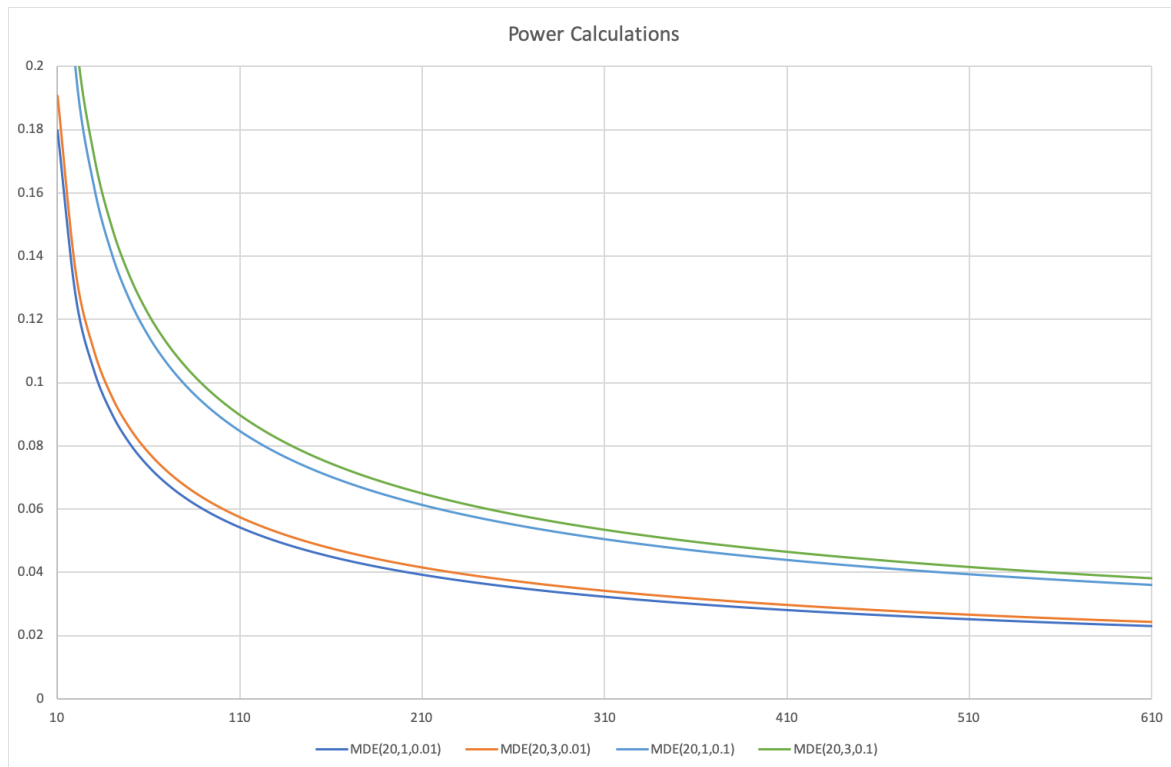


Figure 3.9.12: Power calculations



## .1. Appendix - Construction of psychological variables analyzed in section 3.5.2 and results

To construct psychological our indexes, we relied both on the psychological and behavioural literature:

- Our measure of the depression comes from the behavioral activation for depression scale (BADS) - Short-Form, developed by [Manos et al. \(2011\)](#) and adapted in French by [Wagener et al. \(2014\)](#). Question 23 uses items 1, 5, 6, 7 from this BADS scale.
- Our measure of Generalized Self-Efficacy comes from Generalized Self-Efficacy scale, developed by [Schwarzer \(1995\)](#) and adapted in French by Michelle Dumont, Ralf Schwarzer and Matthias Jerusalem in 2000. Question 24 uses items 1, 3, 8 from French GSE scale.
- Our measure of self-esteem comes from the Rosenberg Self-Esteem scale, developed by [Rosenberg \(1965\)](#), and adapted in French by [Vallieres and Vallerand \(1990\)](#). Question 25 uses items 1, 7, 9, 10 from this scale.

To create the psychological scales associated with depression, self-efficacy, self-esteem, we first numerized answers for each corresponding question :

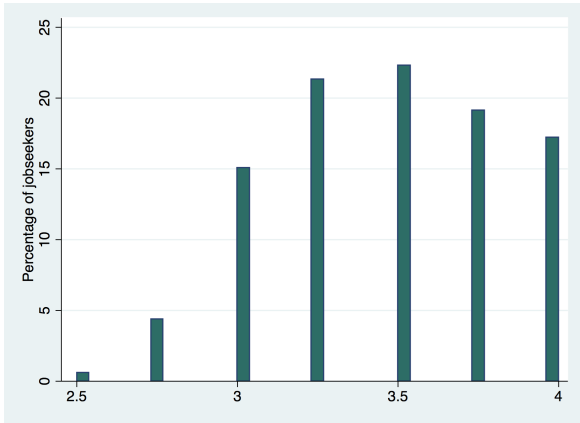
- 1 corresponds to the answer “Pas du tout vrai”;
- 2 corresponds to the answer “A peine vrai”;
- 3 corresponds to the answer “Plutôt vrai”;
- corresponds to the answer to “Totalement vrai”.

However, when the item statement is negative, and as recommended by the literature, we invert the numbers. 4 becomes 1, 3 becomes 2, 2 becomes 3, 1 becomes 4. Finally, we end by averaging the scores for all items within each question in order to obtain the scale<sup>19</sup>.

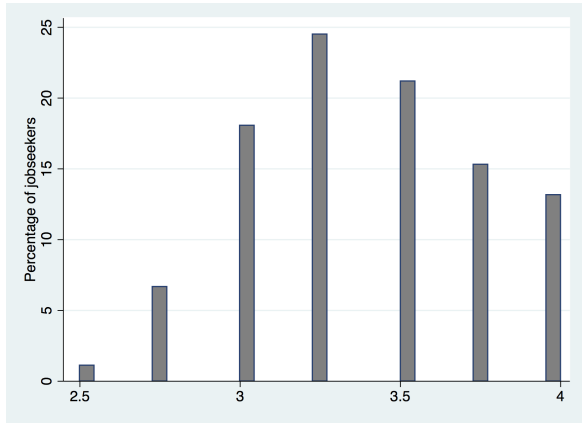
Figures .1.1 to .1.3 shows the distribution of depression, self-efficacy, and self-esteem for the whole sample (in green) and the *PIC* population (in gray).

---

<sup>19</sup>In the literature, summing the scores for all items within each question is more common. However, since we have selected only a few items from the original scales and do not have the whole scale, we average the scores for all items within each question.

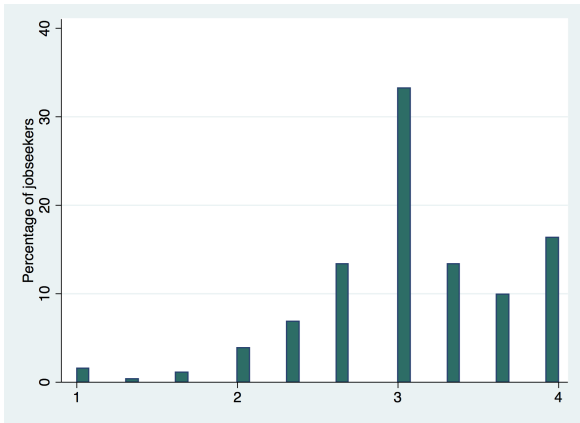


(a) Whole sample

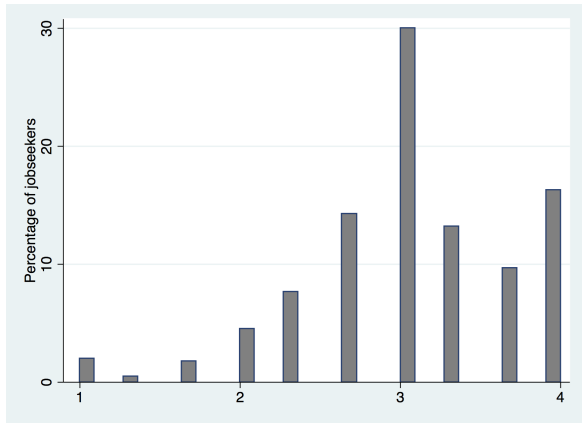


(b) *PIC* population

Figure .1.1: Depression index

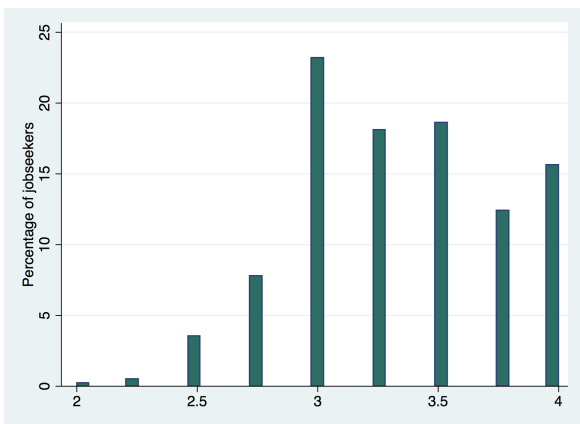


(a) Whole sample

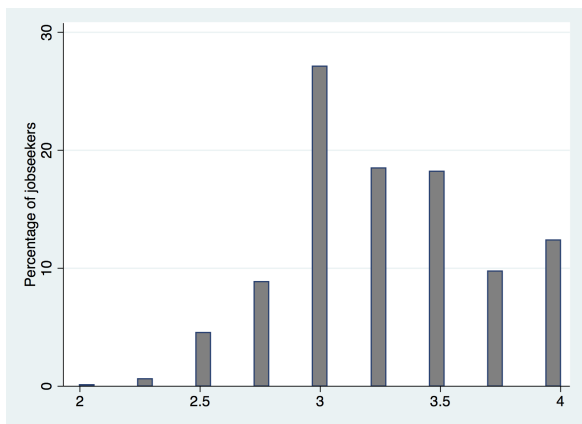


(b) *PIC* population

Figure .1.2: Self-efficacy index



(a) Whole sample



(b) *PIC* population

Figure .1.3: Self-esteem index

## .2. Appendix - Survey questionnaire

## Introduction

Cette enquête dure en moyenne 10 minutes.

Vos réponses sont anonymisées et ne seront jamais communiquées à votre conseiller.


Vos réponses sont très utiles ! Elles permettent d'améliorer les services d'accompagnement.

**\* 1. Selon vous, quel est l'obstacle principal auquel vous faites face dans votre recherche d'emploi ? Indiquez l'obstacle le plus important.**

- Le manque d'offres d'emploi dans les métiers qui m'intéressent
- Les offres d'emploi sont des emplois précaires (CDD, Intérim, auto-entrepreneuriat)
- L'accompagnement dont je bénéficie est insuffisant
- Il me manque des qualifications ou des expériences
- Je ne sais pas comment me comporter sur le marché du travail
- Mon âge
- Il n'y a pas d'obstacle
- Autre (veuillez préciser)

**\* 2. Quel est l'objectif de votre recherche d'emploi ?**


- Je cherche un emploi dans un métier que j'ai déjà occupé
- Je cherche un emploi dans un nouveau métier
- Je souhaite faire le point et réfléchir à de nouvelles pistes professionnelles qui me plairaient davantage
- Je veux créer ou reprendre une entreprise
- Je ne sais pas bien vers quel métier m'orienter
- Je ne recherche pas d'emploi (j'ai déjà retrouvé un emploi, je m'occupe des enfants ou personnes âgées de ma famille, je suis retraité, etc.)



\* 3. Quel type de contrat êtes-vous prêt à accepter ?

(Plusieurs réponses possibles)

- Un CDI ou un CDD d'au moins 6 mois
- Un CDD de moins de 6 mois
- Une mission d'intérim
- Un autre type de contrat

  
\* 4. Indiquez dans la liste déroulante ci-dessous le secteur qui correspond aux métiers qui vous intéressent le plus et dans lequel vous cherchez un emploi (*une seule réponse possible et exigée*).

Si vous n'êtes pas en recherche d'emploi, indiquez le dernier secteur dans lequel vous avez travaillé.

\* 5. Précisez le métier :



**Cette question s'intéresse à votre estimation du retour à un emploi durable. Un emploi durable correspond à un contrat de six mois minimum (CDI ou CDD de six mois minimum).**

\* 6. Selon vous, sur 100 personnes qui s'inscrivent à Pôle emploi et qui recherchent un emploi dans le même secteur que vous, combien ont retrouvé un emploi durable trois mois après leur inscription ?

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment tout le monde (plus de 80)

## Attentes vis à vis des formations et de leur intérêt

Dans le cadre du Plan d'Investissement dans les Compétences, Pôle emploi et ses partenaires proposent des **formations professionnelles** qui visent à faire monter en qualification les demandeurs d'emploi.

Il s'agit principalement de **formations de deux mois minimum pour acquérir les compétences nécessaires à un métier spécifique** (par exemple : chaudronnier, fraiseur, commis de cuisine, etc.).

**Il peut également s'agir de formations de remise à niveau des connaissances fondamentales** (maîtrise de la langue écrite et orale, compétence de calculs) qui sont des **prérequis aux formations à un métier spécifique.**

\* 7. Avez-vous déjà suivi une formation de deux mois minimum en tant que demandeur d'emploi (pour n'importe quel métier / secteur) ?

- Oui, une seule
- Oui, plusieurs
- Non, jamais

La suite du questionnaire porte uniquement sur ces types de formations professionnelles, c'est-à-dire les formations d'une durée de deux mois minimum, pour apprendre un métier spécifique, parmi ceux qui vous intéressent dans le secteur précisé en question 4.

Vous pouvez modifier votre réponse aux questions 4 et 5 à tout moment en utilisant le bouton "Préc." situé en bas de la page.

\* 8. Selon vous, pour vous aider à trouver un travail dans le secteur indiqué en question 4, suivre une formation professionnelle de ce type est :

- Indispensable
- Très utile mais pas indispensable
- Peut-être utile
- Inutile

\* 9. Quel genre de formation serait le plus adapté dans le cadre de votre recherche d'emploi ?

*Si vous envisagez de combiner plusieurs formations, merci d'indiquer les formations concernées (jusqu'à deux formations possibles)*

- Une formation pour me perfectionner dans un métier que je connais déjà
- Une formation pour acquérir une certification pour valider des compétences que je maîtrise déjà
- Une formation pour apprendre un nouveau métier
- Une formation pour revoir mes connaissances fondamentales (langue, écriture, calcul)
- Une formation aux outils et méthodes de recherche d'emploi (écriture de CVs, recherche internet, préparation aux entretiens, etc.)
- Aucune de ces propositions car une formation, quel que soit son objectif, ne me serait pas utile

**\_\_\_\_\_**

\* 10. Considérez les formations professionnelles de deux mois minimum dans le secteur indiqué en question 4. Selon vous, sur 100 personnes inscrites au départ à la formation, combien finissent et valident la formation ?

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment tout le monde (plus de 80)

**Les questions suivantes s'intéressent à votre estimation du retour à un emploi durable trois mois après avoir fini et validé une formation. Un emploi durable correspond à un contrat de six mois minimum (CDI ou CDD de six mois minimum).**

\* 11. Considérez des demandeurs d'emploi qui ont fini et validé **une formation professionnelle** de deux mois minimum dans le secteur indiqué en question 4. Selon vous, sur 100 personnes, combien ont retrouvé un emploi durable trois mois après la fin de la formation ?

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment tout le monde (plus de 80)

\* 12. Considérez des demandeurs d'emploi qui ont fini et validé **une formation pour revoir leurs connaissances fondamentales** (langue, écriture, calcul). Selon vous, sur 100 personnes, combien ont retrouvé un emploi durable trois mois après la fin de la formation ?

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment tout le monde (plus de 80)

\* 13. Considérez des demandeurs d'emploi qui ont fini et validé **une formation pour revoir leurs connaissances fondamentales** (langue, écriture, calcul) **puis une formation professionnelle** de deux mois minimum dans le secteur indiqué en question 4. Selon vous, sur 100 personnes, combien ont retrouvé un emploi durable trois mois après la fin de ces deux formations ?

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment tout le monde (plus de 80)

**\* 14. Considérez les formations professionnelles de deux mois minimum dans le secteur indiqué en question 4. Selon vous, sur 100 places proposées pour ces formations dans le catalogue de formations de Pôle emploi, combien sont gratuites pour les demandeurs d'emploi (car déjà payées par Pôle emploi ou d'autres partenaires) ?**

- Très peu (moins de 10)
- Peu (entre 10 et 20)
- Une bonne part mais moins de la moitié (entre 20 et 40)
- A peu près la moitié (entre 40 et 60)
- Beaucoup (entre 60 et 80)
- Quasiment toutes les places (plus de 80)

### Information à propos des formations

Cette section s'intéresse aux informations dont vous disposez au sujet des formations professionnelles, quel que soit le secteur qui vous intéresse.

\* 15. Avez-vous le sentiment d'être bien informé sur l'offre des formations professionnelles disponibles et sur les procédures d'inscription ?

- Oui, tout à fait
- Suffisamment
- Insuffisamment
- Non, absolument pas

\* 16. Cette question s'intéresse plus précisément à votre accès à l'information au sujet des formations professionnelles de deux mois minimum dans les métiers que vous recherchez. Pour chacune des propositions suivantes, indiquez si vous êtes d'accord.

Tout à fait en désaccord    Plutôt en désaccord    Plutôt en accord    Tout à fait d'accord

Je connais les différents types de formation (certifiante, non certifiante, etc.)

Je connais les principaux organismes délivrant des formations et comment les contacter

Je sais distinguer les formations gratuites pour les demandeurs d'emploi et les formations qui nécessitent une participation financière des demandeurs d'emploi

Je sais rechercher les formations gratuites pour les demandeurs d'emploi dans le catalogue des formations de Pôle emploi

Je sais où rechercher des formations spécifiques et leurs caractéristiques (intervenants, nombre d'heures, stage éventuel, prérequis, entretiens/test de sélection, etc.) pour identifier celles qui m'intéressent

J'ai bien en tête les différentes étapes et démarches nécessaires pour m'inscrire en formation

Pour les frais d'inscription, je sais si des financements sont accessibles et connais leurs conditions

Je sais où rechercher des informations pratiques sur la disponibilité, les dates et les lieux des formations qui m'intéressent



**\* 17. Comment vous informez-vous sur les formations ?**

**Indiquez vos principales sources d'information (jusqu'à trois réponses possibles)**

- Avec mon conseiller (Pôle emploi, Mission locale, Cap emploi, etc.)
- Avec mon entourage (famille, amis, connaissances)
- Sur les sites internet de Pôle emploi
- Sur d'autres sites internet
- Directement auprès des organismes de formation
- Par les ateliers proposés par Pôle emploi
- Par d'autres moyens (associations, réseau professionnel, etc.)
- Je ne sais pas vraiment où trouver de l'information
- Je ne cherche pas vraiment d'information

## Offre et qualité des formations

Cette section s'intéresse à vos opinions à propos de l'offre, de la qualité, de l'adéquation à vos besoins, et de l'intérêt d'une formation professionnelle de deux mois minimum pour votre recherche d'emploi.

\* 18. Y a-t-il des offres de formation adaptées à votre projet professionnel proches de chez vous ?

- Oui, il y en a assez
- Oui, mais il y en a peu
- Non, il n'y en a pas
- Je ne sais pas

\* 19. Cette question s'intéresse à votre sentiment et votre jugement sur l'offre des formations. Nous comprenons parfaitement que vous puissiez ne pas être familier du contenu ou de la qualité. Dans ce cas, répondez en fonction de ce que vous imaginez des formations professionnelles de deux mois minimum dans les métiers que vous recherchez.

	Tout à fait en désaccord	Plutôt en désaccord	Plutôt d'accord	Tout à fait d'accord
Les formations proposées sont adaptées à vos besoins	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les formations proposées sont proches de chez vous	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les formations proposées sont de bonne qualité	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les formations proposées vous permettront de retrouver un emploi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les formations proposées vous permettront de réussir dans votre nouvel emploi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les bénéfices obtenus de la formation valent la peine d'effectuer les démarches nécessaires à l'inscription	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Le contenu des formations est suffisamment pratique pour acquérir des compétences qui seront utiles sur le marché du travail	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Le contenu des formations est trop théorique pour être utile sur le marché du travail	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Etant donné vos compétences, vous parviendrez à suivre et valider ces formations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Le niveau théorique de ces formations est trop élevé pour vous	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les intervenants sont compétents	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Les intervenants respectent les chômeurs et s'intéressent à leur situation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

**\* 20. Pour vous, est-ce important d'obtenir un diplôme ou une certification professionnelle à l'issue de la formation ?**

- Oui, le diplôme ou la certification est la principale raison de suivre la formation
- Je valorise autant le diplôme ou la certification que les compétences apprises
- Non, les compétences apprises sont plus importantes

## Capacités à s'inscrire et suivre une formation

Cette section s'intéresse à ce que vous pensez de vos compétences et capacités.

\* 21. Pensez aux **métiers** que vous visez actuellement dans votre recherche d'emploi. En considérant les connaissances et compétences que vous avez actuellement (sans formation additionnelle), indiquez si vous êtes d'accord avec les propositions suivantes.

	Tout à fait en désaccord	Plutôt en désaccord	Plutôt d'accord	Tout à fait d'accord
J'ai les diplômes et certificats requis pour trouver un emploi dans ces métiers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les compétences techniques requises pour réussir dans mon futur emploi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les aptitudes sociales nécessaires pour être à l'aise dans mes relations avec mes collègues et ma hiérarchie professionnelle	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les qualités personnelles (conscientieux, fiable, organisé) pour réussir dans mon futur emploi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 22. Pensez maintenant aux **formations professionnelles de deux mois minimum** dans les métiers que vous recherchez.

Pour chacune des propositions, indiquez si vous êtes d'accord.

	Tout à fait en désaccord	Plutôt en désaccord	Plutôt d'accord	Tout à fait d'accord
J'ai les diplômes et certificats requis pour entrer dans cette formation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les compétences techniques requises pour réussir dans cette formation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les aptitudes sociales nécessaires pour être à l'aise avec les encadrants, professeurs, et autres demandeurs d'emploi au cours de cette formation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai les qualités personnelles (conscientieux, fiable, organisé) pour réussir cette formation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Etre au chômage peut être une expérience difficile au cours de laquelle il est parfois dur de savoir où on en est, par où commencer, comment s'organiser et accomplir les différentes étapes nécessaires pour aller de l'avant. En particulier, il peut être difficile d'identifier et déterminer les formations professionnelles qui seraient bénéfiques, effectuer les démarches d'inscription, trouver des offres d'emploi et y candidater. Cela peut également être une période stressante et parfois angoissante où confiance et motivation sont mises à l'épreuve.

\* 23. Pensez à ce que vous avez fait la semaine dernière, y compris aujourd'hui.

Pour chacune des propositions suivantes, indiquez à quel point l'énoncé est vrai pour vous.

	Pas du tout vrai	A peine vrai	Moyennement vrai	Totalement vrai
J'ai été une personne active et j'ai accompli les objectifs que je m'étais fixés	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Il y a eu beaucoup de choses que je devais faire que je n'ai pas faites	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
La plupart des choses que j'ai faites avaient pour objectif d'échapper à quelque chose de désagréable ou à l'éviter	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai passé beaucoup de temps à ressasser mes problèmes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 24. Les propositions suivantes s'intéressent plus généralement à votre expérience de demandeur d'emploi et à votre activité de recherche d'emploi ou de formation.

Pour chacune des propositions suivantes, indiquez si vous êtes d'accord.

	Tout à fait en désaccord	Plutôt en désaccord	Plutôt d'accord	Tout à fait d'accord
Je peux toujours arriver à résoudre mes difficultés si j'essaie assez fort	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
C'est facile pour moi de maintenir mon attention sur mes objectifs et accomplir mes buts	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Lorsque que je suis confronté à un problème, je peux habituellement trouver plusieurs solutions	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 25. Les propositions suivantes concernent votre perception de vous-même dans cette période de recherche d'emploi ou de formation.

Pour chacune d'elles, indiquez si vous êtes d'accord.

	Tout à fait en désaccord	Plutôt en désaccord	Plutôt d'accord	Tout à fait d'accord
Je pense que je suis une personne de valeur, au moins égale à n'importe qui d'autre	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Dans l'ensemble, je suis satisfait de moi	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Parfois, je me sens vraiment inutile	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Il m'arrive de penser que je suis un bon à rien	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

\* 26. Quelles qualités essentielles attribuez-vous aux personnes qui réussissent professionnellement ?

(Jusqu'à trois réponses possibles)

- Le talent, l'intelligence
- La persévérance
- Etre travailleur
- La chance
- L'appartenance à un milieu social privilégié
- Avoir un bon réseau personnel ou professionnel
- L'apparence physique
- La capacité à bien travailler avec ses collègues et ses supérieurs
- Autre (veuillez préciser)

## Accès aux formations - sélection et financement

Cette section s'intéresse à vos chances d'être admis dans une formation et d'obtenir un financement.

\* 27. Considérez les formations professionnelles de deux mois minimum dans les métiers que vous recherchez. Si vous faisiez toutes les démarches nécessaires pour y candidater, pensez-vous être sélectionné par l'organisme délivrant la formation ?

- Oui, j'ai de bonnes chances
- Peu probable, car il n'y a pas assez de places
- Peu probable, car je n'ai pas les prérequis nécessaires
- Peu probable, car je ne réussirai pas les tests d'entrée en formation
- Peu probable, pour d'autres raisons

\* 28. Si vous faisiez toutes les démarches nécessaires pour solliciter un financement, pensez-vous pouvoir en obtenir un ?

- Oui, j'ai de bonnes chances
- Peu probable, car peu de financements sont accordés dans les métiers que je recherche pour ce type de formation
- Peu probable, car Pôle emploi considère mon projet de formation inadapté
- Peu probable, pour d'autres raisons

\* 29. Est-ce qu'obtenir un financement pour pouvoir suivre une formation est :

- Indispensable, sans quoi je ne la suivrai pas
- Important mais pas indispensable, je la suivrai quand même
- Secondaire, avec ou sans financement je la suivrai

\* 30. Si vous aviez un financement couvrant les frais d'inscription, pourriez-vous vous permettre de suivre une formation ?

- Oui, tout à fait
- Non, même en ne payant pas les frais d'inscription, je ne peux pas financer les frais annexes (coûts de transport, restauration, etc.)
- Non, qu'elle soit financée ou non, je ne peux pas suivre une formation. Je dois retrouver un travail rapidement pour avoir davantage de revenus



## Plans et intentions

Cette section s'intéresse à vos projets en matière de formation. Pensez aux formations professionnelles de deux mois minimum dans les métiers que vous recherchez.

\* 31. Envisagez-vous de vous inscrire en formation prochainement ?

- Oui, je suis en train de faire les démarches
- Oui, dans les prochaines semaines
- Pas immédiatement mais dans les six prochains mois
- Non, je n'ai pas de projet défini sur les six prochains mois
- Non, je n'ai aucune intention de suivre une formation

\* 32. Au cours de la semaine dernière, combien de temps environ avez-vous consacré à la recherche de formation ?

- Plusieurs journées complètes
- Quelques demi-journées
- Quelques heures à l'occasion
- Je n'y ai pas passé de temps

## Obstacles à l'entrée en formation


\* 33. Quel obstacle majeur de motivation ou d'organisation personnelle, s'il y en a un, vous empêche de vous inscrire en formation ?

- Il n'y a pas d'obstacle majeur
- Je n'ai pas la motivation car j'ai le sentiment d'être inutile
- Je suis débordé par toutes les démarches et je ne parviens pas à m'organiser
- J'ai prévu et j'aimerais m'inscrire en formation mais je ne parviens pas à me forcer à le faire, je repousse sans arrêt
- Même si je faisais toutes les démarches, je pense que je n'aurais aucun contrôle sur mon accès en formation, donc à quoi bon

\* 34. Quel obstacle pratique majeur, s'il y en a un, vous empêche de vous inscrire en formation ?

- Il n'y a pas d'obstacle majeur
- Il n'existe pas assez d'informations sur l'offre des formations et les modalités de candidature
- Il n'y a pas de formation de qualité, adaptée à mes besoins, assez proche de chez moi
- Même une formation de qualité ne me sera d'aucune aide pour retrouver un emploi
- Je ne peux pas me le permettre financièrement

**Cliquez sur le bouton "Suiv." pour terminer et valider l'enquête !  
Vous trouverez ensuite quelques conseils et astuces utiles sur la  
recherche d'emploi !**

	
--	--

# References

- Abdulkadiroğlu, A., P. A. Pathak, and C. R. Walters (2018, January). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics* 10(1), 175–206.
- Abel, M., R. Burger, E. Carranza, and P. Piraino (2019). Bridging the intention-behavior gap? the effect of plan-making prompts on job search and employment. *American Economic Journal: Applied Economics* 11(2), 284–301.
- Adomavicius, G. and A. Tuzhilin (2005, 07). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on* 17, 734–749.
- Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018). Learning about job search: A field experiment with job seekers in germany. *Journal of Public Economics* 164, 33–49.
- Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018). Learning about job search: A field experiment with job seekers in Germany. *Journal of Public Economics* 164(C), 33–49.
- Andersson, F., H. J. Holzer, J. I. Lane, D. Rosenblum, and J. A. Smith (2016). Does Federally-Funded Job Training Work? Non-experimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms.
- Annexe au Projet de Loi de Finances, F. P. (2018).
- Athey, S., J. Tibshirani, S. Wager, et al. (2019). Generalized random forests. *The Annals of Statistics* 47(2), 1148–1178.
- Autor, D. (2009, 01). Studies of labor market intermediation : Introduction to "studies of labor market intermediation". *Studies of Labor Market Intermediation*.
- Babcock, L., W. J. Congdon, L. F. Katz, and S. Mullainathan (2012). Notes on behavioral economics and labor market policy. *IZA Journal of Labor Policy* 1(1), 2.
- Barnow, B. S. and J. Smith (2015, October). Employment and training programs. (21659).
- Barr, A. and S. Turner (2017). A letter and encouragement: Does information increase post-secondary enrollment of ui recipients? (23374).
- Beaman, L. and J. Magruder (2012). Who gets the job referral? evidence from a social networks experiment. *American Economic Review* 102(7), 3574–93.

- Behaghel, L., B. Crépon, and M. Gurgand (2014). Private and public provision of counseling to job seekers: Evidence from a large controlled experiment. *American economic journal: applied economics* 6(4), 142–74.
- Belloni, A., V. Chernozhukov, and C. Hansen (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies* 81(2), 608–650.
- Belot, M., P. Kircher, and P. Muller (2019). Providing advice to jobseekers at low cost: An experimental study on online advice. *The review of economic studies* 86(4), 1411–1447.
- Benabou, R. and J. Tirole (2002, 02). Self-confidence and personal motivation. *The Quarterly Journal of Economics* 117, 871–915.
- Berger, E. M., H. Hermes, G. Koenig, F. Schmidt, and D. Schunk (2019). Self-regulation training and job search behavior: A natural field experiment within an active labor market program. *Working Paper*.
- Berger, M. C., D. Black, and J. A. Smith (2001). Evaluating profiling as a means of allocating government services. In M. Lechner and F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies*, Heidelberg, pp. 59–84. Physica-Verlag HD.
- Bertrand, M., E. Shafir, and S. Mullainathan (2004). A behavioral economics view of poverty.
- Bleemer, Z. and B. Zafar (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics* 157(C), 184–211.
- Blundell, R., M. C. Dias, C. Meghir, and J. V. Reenen (2004, June). Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association* 2(4), 569–606.
- Breiman, L. (2001). Random forests. *Machine learning* 45(1), 5–32.
- Caliendo, M., D. Cobb-Clark, and A. Uhlendorff (2015). Locus of control and job search strategies. *The Review of Economics and Statistics* 97(1), 88–103.
- Caliendo, M., D. A. Cobb-Clark, C. Obst, H. Seitz, and A. Uhlendorff (2020). Locus of control and investment in training. *Journal of Human Resources*.
- Card, D., J. Kluge, and A. Weber (2017, 10). What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association* 16(3), 894–931.
- Carrell, S. and B. Sacerdote (2013). Why do college going interventions work? (19031).
- Carrell, S. and B. Sacerdote (2017, July). Why do college-going interventions work? *American Economic Journal: Applied Economics* 9(3), 124–51.
- Castleman, B. and L. C. Page (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior Organization* 115(C), 144–160.

- Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research.
- Conlon, J. (2018). Major malfunction: A field experiment correcting undergraduates' beliefs about salaries.
- Cour des Comptes, L. f. d. d. d. (2018).
- Dasgupta, U., L. Gangadharan, P. Maitra, S. Mani, and S. Subramanian (2012). Choosing to Be Trained: Evidence from a Field Experiment. *SSRN Electronic Journal*.
- Dellavigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47(2), 315–372.
- DellaVigna, S., J. Heining, J. F. Schmieder, and S. Trenkle (2020, April). Evidence on job search models from a survey of unemployed workers in Germany. (27037).
- DellaVigna, S. and M. D. Paserman (2005a). Job search and impatience. *Journal of Labor Economics* 23(3), 527–588.
- DellaVigna, S. and M. D. Paserman (2005b). Job search and impatience. *Journal of Labor Economics* 23(3), 527–588.
- Dinerstein, M., L. Einav, J. Levin, and N. Sundaresan (2018, July). Consumer price search and platform design in internet commerce. *American Economic Review* 108(7), 1820–59.
- Dizon-Ross, R. (2019, August). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review* 109(8), 2728–65.
- Duflo, E., R. Glennerster, and M. Kremer (2006, December). Using randomization in development economics research: A toolkit. Working Paper 333, National Bureau of Economic Research.
- Duflo, E., R. Glennerster, and M. Kremer (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics* 4, 3895–3962.
- Dynarski, S. and J. Scott-Clayton (2008). Complexity and targeting in federal student aid: A quantitative analysis.
- Falk, A., D. Huffman, and U. Sunde (2006a). Self-confidence and search.
- Falk, A., D. B. Huffman, and U. Sunde (2006b, December). Self-Confidence and Search. IZA Discussion Papers 2525, Institute of Labor Economics (IZA).
- Finkelstein, A. and M. Notowidigdo (2018). Take-up and targeting: Experimental evidence from SNAP. (11558).
- Gee, L. K. (2019). The more you know: Information effects on job application rates in a large field experiment. *Management Science* 65(5), 2077–2094.
- Ghosal, S., S. Jana, A. Mani, S. Mitra, S. Roy, et al. (forthcoming). Sex workers, self-image and stigma: Evidence from Kolkata brothels. *Review of Economics and Statistics*.

- Greenberg, D. H., C. Michalopoulos, and P. K. Robins (2003). A meta-analysis of government-sponsored training programs. *ILR Review* 57(1), 31–53.
- Hanna, R., S. Mullainathan, and J. Schwartzstein (2014, 06). Learning Through Noticing: Theory and Evidence from a Field Experiment \*. *The Quarterly Journal of Economics* 129(3), 1311–1353.
- Haushofer, J., A. John, and K. Orkin (2020). Can simple psychological interventions increase preventive health investment? *NBER Paper*.
- Herz, B. and T. van Rens (2019, 05). Accounting for Mismatch Unemployment. *Journal of the European Economic Association*. jvz018.
- Horton, J. J. (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics* 35(2), 345–385.
- Insee (2018). Emploi, chômage, revenus du travail. *Insee Références*.
- Jacobson, L. and J. Davis (2017). The Relative Returns to Workforce Investment Act-Supported Training in Florida by Field, Gender, and Education and Ways to Improve Trainees’ Choices. *Journal of Labor Economics* 35(S1), 337–375.
- Jensen, R. (2010, 05). The (Perceived) Returns to Education and the Demand for Schooling\*. *The Quarterly Journal of Economics* 125(2), 515–548.
- Krueger, A. B., A. Mueller, S. J. Davis, and A. Şahin (2011). Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high frequency longitudinal data [with comments and discussion]. *Brookings Papers on Economic Activity*, 1–81.
- Krueger, A. B. and A. I. Mueller (2012). Time use, emotional well-being, and unemployment: Evidence from longitudinal data. *American Economic Review* 102(3), 594–99.
- Krug, G. and G. Stephan (2013). Is the contracting-out of intensive placement services more effective than provision by the pes? evidence from a randomized field experiment. (7403).
- Kuhn, P. and H. Mansour (2014a). Is internet job search still ineffective? *The Economic Journal* 124(581), 1213–1233.
- Kuhn, P. and H. Mansour (2014b). Is internet job search still ineffective? *The Economic Journal* 124(581), 1213–1233.
- Kuhn, P. and M. Skuterud (2004a). Internet job search and unemployment durations. *American Economic Review* 94(1), 218–232.
- Kuhn, P. and M. Skuterud (2004b). Internet job search and unemployment durations. *American Economic Review* 94(1), 218–232.
- Künzel, S. R., J. S. Sekhon, P. J. Bickel, and B. Yu (2017). Meta-learners for estimating heterogeneous treatment effects using machine learning. *arXiv preprint arXiv:1706.03461*.
- Manos, R. C., J. Kanter, and W. Luo (2011). The behavioral activation for depression scale-short form: development and validation. *Behavior therapy* 42 4, 726–39.

- Marinescu, I. and R. Rathelot (2018, July). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics* 10(3), 42–70.
- McCall, B., J. Smith, and C. Wunsch (2016, June). Government-Sponsored Vocational Education for Adults. 5, 479–652.
- McGee, A. D. (2015). How the perception of control influences unemployed job search. *ILR Review* 68(1), 184–211.
- McKee-Ryan, F., Z. Song, C. R. Wanberg, and A. J. Kinicki (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology* 90(1), 53.
- McKelway, M. (2018). Women’s self-efficacy and women’s employment: Experimental evidence from india. *Working Paper*.
- Meyer, C. J. (2018). In search of a better life: Self-control in the ethiopian labor market.
- Mortensen, D. T. (1970). A theory of wage and employment dynamics. *Microeconomic Foundations of Employment and Inflation Theory*.
- Papageorgiou, T. (2014, 02). Learning Your Comparative Advantages. *The Review of Economic Studies* 81(3), 1263–1295.
- Paserman, M. D. (2008, 08). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation. *The Economic Journal* 118(531), 1418–1452.
- Patterson, C., A. Şahin, G. Topa, and G. Violante (2016, May). Working hard in the wrong place: A mismatch-based explanation to the uk productivity puzzle.
- Resnick, P. and H. R. Varian (1997). Recommender systems. *Commun. ACM* 40, 56–58.
- Ridley, M., G. Rao, F. Schilbach, and V. Patel (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Working paper*.
- Rosenberg, M. (1965). Society and the adolescent self-image. *Princeton University Press*.
- Rubin, D. (1974a). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.
- Rubin, D. B. (1974b). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66(5), 688.
- Schilbach, F., H. Schofield, and S. Mullainathan (2016). The psychological lives of the poor. *American Economic Review Papers and Proceedings* 106(5), 435–40.
- Schwarzer, R., . J. M. (1995). Generalized self-efficacy scale. in j. weinman, s. wright, m. johnston, measures in health psychology: A user’s portfolio. causal and control beliefs (pp. 35- 37). *Windsor, England: NFER-NELSON*.
- Skandalis, D. (2019). Breaking news: The role of information in job search and matching. *Working paper*.
- Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, Vol. 13, No 1, pp. 130-167.



- Stigler, G. (1962). Information in the labor market. *Journal of Political Economy*.
- Vallieres, E. F. and R. J. Vallerand (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi de rosenberg. *International Journal of Psychology* 25(2), 305–316.
- Varian, H. R. (2010, May). Computer mediated transactions. *American Economic Review* 100(2), 1–10.
- Wagener, A., M. Van der Linden, and S. Blairy (2014, 10). Psychometric properties of the french translation of the behavioral activation for depression scale - short form (bads-sf) in non-clinical adults. *Comprehensive Psychiatry* 56.
- Wager, S. and S. Athey (2017). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* (just-accepted).
- Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* 67(2), 301–320.
- Şahin, A., J. Song, G. Topa, and G. L. Violante (2014, November). Mismatch unemployment. *American Economic Review* 104(11), 3529–64.

# References

- Abdulkadiroğlu, A., P. A. Pathak, and C. R. Walters (2018, January). Free to choose: Can school choice reduce student achievement? *American Economic Journal: Applied Economics* 10(1), 175–206.
- Abel, M., R. Burger, E. Carranza, and P. Piraino (2019). Bridging the intention-behavior gap? the effect of plan-making prompts on job search and employment. *American Economic Journal: Applied Economics* 11(2), 284–301.
- Adomavicius, G. and A. Tuzhilin (2005, 07). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on* 17, 734–749.
- Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018). Learning about job search: A field experiment with job seekers in germany. *Journal of Public Economics* 164, 33–49.
- Altmann, S., A. Falk, S. Jäger, and F. Zimmermann (2018). Learning about job search: A field experiment with job seekers in Germany. *Journal of Public Economics* 164(C), 33–49.
- Andersson, F., H. J. Holzer, J. I. Lane, D. Rosenblum, and J. A. Smith (2016). Does Federally-Funded Job Training Work? Non-experimental Estimates of WIA Training Impacts Using Longitudinal Data on Workers and Firms.
- Annexe au Projet de Loi de Finances, F. P. (2018).
- Athey, S., J. Tibshirani, S. Wager, et al. (2019). Generalized random forests. *The Annals of Statistics* 47(2), 1148–1178.
- Autor, D. (2009, 01). Studies of labor market intermediation : Introduction to "studies of labor market intermediation". *Studies of Labor Market Intermediation*.
- Babcock, L., W. J. Congdon, L. F. Katz, and S. Mullainathan (2012). Notes on behavioral economics and labor market policy. *IZA Journal of Labor Policy* 1(1), 2.
- Barnow, B. S. and J. Smith (2015, October). Employment and training programs. (21659).
- Barr, A. and S. Turner (2017). A letter and encouragement: Does information increase post-secondary enrollment of ui recipients? (23374).
- Beaman, L. and J. Magruder (2012). Who gets the job referral? evidence from a social networks experiment. *American Economic Review* 102(7), 3574–93.

- Behaghel, L., B. Crépon, and M. Gurgand (2014). Private and public provision of counseling to job seekers: Evidence from a large controlled experiment. *American economic journal: applied economics* 6(4), 142–74.
- Belloni, A., V. Chernozhukov, and C. Hansen (2014). Inference on treatment effects after selection among high-dimensional controls. *The Review of Economic Studies* 81(2), 608–650.
- Belot, M., P. Kircher, and P. Muller (2019). Providing advice to jobseekers at low cost: An experimental study on online advice. *The review of economic studies* 86(4), 1411–1447.
- Benabou, R. and J. Tirole (2002, 02). Self-confidence and personal motivation. *The Quarterly Journal of Economics* 117, 871–915.
- Berger, E. M., H. Hermes, G. Koenig, F. Schmidt, and D. Schunk (2019). Self-regulation training and job search behavior: A natural field experiment within an active labor market program. *Working Paper*.
- Berger, M. C., D. Black, and J. A. Smith (2001). Evaluating profiling as a means of allocating government services. In M. Lechner and F. Pfeiffer (Eds.), *Econometric Evaluation of Labour Market Policies*, Heidelberg, pp. 59–84. Physica-Verlag HD.
- Bertrand, M., E. Shafir, and S. Mullainathan (2004). A behavioral economics view of poverty.
- Bleemer, Z. and B. Zafar (2018). Intended college attendance: Evidence from an experiment on college returns and costs. *Journal of Public Economics* 157(C), 184–211.
- Blundell, R., M. C. Dias, C. Meghir, and J. V. Reenen (2004, June). Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association* 2(4), 569–606.
- Breiman, L. (2001). Random forests. *Machine learning* 45(1), 5–32.
- Caliendo, M., D. Cobb-Clark, and A. Uhlendorff (2015). Locus of control and job search strategies. *The Review of Economics and Statistics* 97(1), 88–103.
- Caliendo, M., D. A. Cobb-Clark, C. Obst, H. Seitz, and A. Uhlendorff (2020). Locus of control and investment in training. *Journal of Human Resources*.
- Card, D., J. Kluge, and A. Weber (2017, 10). What Works? A Meta Analysis of Recent Active Labor Market Program Evaluations. *Journal of the European Economic Association* 16(3), 894–931.
- Carrell, S. and B. Sacerdote (2013). Why do college going interventions work? (19031).
- Carrell, S. and B. Sacerdote (2017, July). Why do college-going interventions work? *American Economic Journal: Applied Economics* 9(3), 124–51.
- Castleman, B. and L. C. Page (2015). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior Organization* 115(C), 144–160.

- Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018). Generic machine learning inference on heterogeneous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research.
- Conlon, J. (2018). Major malfunction: A field experiment correcting undergraduates' beliefs about salaries.
- Cour des Comptes, L. f. d. d. d. (2018).
- Dasgupta, U., L. Gangadharan, P. Maitra, S. Mani, and S. Subramanian (2012). Choosing to Be Trained: Evidence from a Field Experiment. *SSRN Electronic Journal*.
- Dellavigna, S. (2009). Psychology and economics: Evidence from the field. *Journal of Economic Literature* 47(2), 315–372.
- DellaVigna, S., J. Heining, J. F. Schmieder, and S. Trenkle (2020, April). Evidence on job search models from a survey of unemployed workers in Germany. (27037).
- DellaVigna, S. and M. D. Paserman (2005a). Job search and impatience. *Journal of Labor Economics* 23(3), 527–588.
- DellaVigna, S. and M. D. Paserman (2005b). Job search and impatience. *Journal of Labor Economics* 23(3), 527–588.
- Dinerstein, M., L. Einav, J. Levin, and N. Sundaresan (2018, July). Consumer price search and platform design in internet commerce. *American Economic Review* 108(7), 1820–59.
- Dizon-Ross, R. (2019, August). Parents' beliefs about their children's academic ability: Implications for educational investments. *American Economic Review* 109(8), 2728–65.
- Duflo, E., R. Glennerster, and M. Kremer (2006, December). Using randomization in development economics research: A toolkit. Working Paper 333, National Bureau of Economic Research.
- Duflo, E., R. Glennerster, and M. Kremer (2007). Using randomization in development economics research: A toolkit. *Handbook of development economics* 4, 3895–3962.
- Dynarski, S. and J. Scott-Clayton (2008). Complexity and targeting in federal student aid: A quantitative analysis.
- Falk, A., D. Huffman, and U. Sunde (2006a). Self-confidence and search.
- Falk, A., D. B. Huffman, and U. Sunde (2006b, December). Self-Confidence and Search. IZA Discussion Papers 2525, Institute of Labor Economics (IZA).
- Finkelstein, A. and M. Notowidigdo (2018). Take-up and targeting: Experimental evidence from SNAP. (11558).
- Gee, L. K. (2019). The more you know: Information effects on job application rates in a large field experiment. *Management Science* 65(5), 2077–2094.
- Ghosal, S., S. Jana, A. Mani, S. Mitra, S. Roy, et al. (forthcoming). Sex workers, self-image and stigma: Evidence from Kolkata brothels. *Review of Economics and Statistics*.

- Greenberg, D. H., C. Michalopoulos, and P. K. Robins (2003). A meta-analysis of government-sponsored training programs. *ILR Review* 57(1), 31–53.
- Hanna, R., S. Mullainathan, and J. Schwartzstein (2014, 06). Learning Through Noticing: Theory and Evidence from a Field Experiment \*. *The Quarterly Journal of Economics* 129(3), 1311–1353.
- Haushofer, J., A. John, and K. Orkin (2020). Can simple psychological interventions increase preventive health investment? *NBER Paper*.
- Herz, B. and T. van Rens (2019, 05). Accounting for Mismatch Unemployment. *Journal of the European Economic Association*. jvz018.
- Horton, J. J. (2017). The effects of algorithmic labor market recommendations: Evidence from a field experiment. *Journal of Labor Economics* 35(2), 345–385.
- Insee (2018). Emploi, chômage, revenus du travail. *Insee Références*.
- Jacobson, L. and J. Davis (2017). The Relative Returns to Workforce Investment Act-Supported Training in Florida by Field, Gender, and Education and Ways to Improve Trainees’ Choices. *Journal of Labor Economics* 35(S1), 337–375.
- Jensen, R. (2010, 05). The (Perceived) Returns to Education and the Demand for Schooling\*. *The Quarterly Journal of Economics* 125(2), 515–548.
- Krueger, A. B., A. Mueller, S. J. Davis, and A. Şahin (2011). Job search, emotional well-being, and job finding in a period of mass unemployment: Evidence from high frequency longitudinal data [with comments and discussion]. *Brookings Papers on Economic Activity*, 1–81.
- Krueger, A. B. and A. I. Mueller (2012). Time use, emotional well-being, and unemployment: Evidence from longitudinal data. *American Economic Review* 102(3), 594–99.
- Krug, G. and G. Stephan (2013). Is the contracting-out of intensive placement services more effective than provision by the pes? evidence from a randomized field experiment. (7403).
- Kuhn, P. and H. Mansour (2014a). Is internet job search still ineffective? *The Economic Journal* 124(581), 1213–1233.
- Kuhn, P. and H. Mansour (2014b). Is internet job search still ineffective? *The Economic Journal* 124(581), 1213–1233.
- Kuhn, P. and M. Skuterud (2004a). Internet job search and unemployment durations. *American Economic Review* 94(1), 218–232.
- Kuhn, P. and M. Skuterud (2004b). Internet job search and unemployment durations. *American Economic Review* 94(1), 218–232.
- Künzel, S. R., J. S. Sekhon, P. J. Bickel, and B. Yu (2017). Meta-learners for estimating heterogeneous treatment effects using machine learning. *arXiv preprint arXiv:1706.03461*.
- Manos, R. C., J. Kanter, and W. Luo (2011). The behavioral activation for depression scale-short form: development and validation. *Behavior therapy* 42 4, 726–39.

- Marinescu, I. and R. Rathelot (2018, July). Mismatch unemployment and the geography of job search. *American Economic Journal: Macroeconomics* 10(3), 42–70.
- McCall, B., J. Smith, and C. Wunsch (2016, June). Government-Sponsored Vocational Education for Adults. 5, 479–652.
- McGee, A. D. (2015). How the perception of control influences unemployed job search. *ILR Review* 68(1), 184–211.
- McKee-Ryan, F., Z. Song, C. R. Wanberg, and A. J. Kinicki (2005). Psychological and physical well-being during unemployment: a meta-analytic study. *Journal of applied psychology* 90(1), 53.
- McKelway, M. (2018). Women’s self-efficacy and women’s employment: Experimental evidence from india. *Working Paper*.
- Meyer, C. J. (2018). In search of a better life: Self-control in the ethiopian labor market.
- Mortensen, D. T. (1970). A theory of wage and employment dynamics. *Microeconomic Foundations of Employment and Inflation Theory*.
- Papageorgiou, T. (2014, 02). Learning Your Comparative Advantages. *The Review of Economic Studies* 81(3), 1263–1295.
- Paserman, M. D. (2008, 08). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation. *The Economic Journal* 118(531), 1418–1452.
- Patterson, C., A. Şahin, G. Topa, and G. Violante (2016, May). Working hard in the wrong place: A mismatch-based explanation to the uk productivity puzzle.
- Resnick, P. and H. R. Varian (1997). Recommender systems. *Commun. ACM* 40, 56–58.
- Ridley, M., G. Rao, F. Schilbach, and V. Patel (2020). Poverty, depression, and anxiety: Causal evidence and mechanisms. *Working paper*.
- Rosenberg, M. (1965). Society and the adolescent self-image. *Princeton University Press*..
- Rubin, D. (1974a). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.
- Rubin, D. B. (1974b). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology* 66(5), 688.
- Schilbach, F., H. Schofield, and S. Mullainathan (2016). The psychological lives of the poor. *American Economic Review Papers and Proceedings* 106(5), 435–40.
- Schwarzer, R., . J. M. (1995). Generalized self-efficacy scale. in j. weinman, s. wright, m. johnston, measures in health psychology: A user’s portfolio. causal and control beliefs (pp. 35- 37). *Windsor, England: NFER-NELSON*.
- Skandalis, D. (2019). Breaking news: The role of information in job search and matching. *Working paper*.
- Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, Vol. 13, No 1, pp. 130-167.

- Stigler, G. (1962). Information in the labor market. *Journal of Political Economy*.
- Vallieres, E. F. and R. J. Vallerand (1990). Traduction et validation canadienne-française de l'échelle de l'estime de soi de rosenberg. *International Journal of Psychology* 25(2), 305–316.
- Varian, H. R. (2010, May). Computer mediated transactions. *American Economic Review* 100(2), 1–10.
- Wagener, A., M. Van der Linden, and S. Blairy (2014, 10). Psychometric properties of the french translation of the behavioral activation for depression scale - short form (bads-sf) in non-clinical adults. *Comprehensive Psychiatry* 56.
- Wager, S. and S. Athey (2017). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association* (just-accepted).
- Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* 67(2), 301–320.
- Şahin, A., J. Song, G. Topa, and G. L. Violante (2014, November). Mismatch unemployment. *American Economic Review* 104(11), 3529–64.

**Titre :** Essai sur le rôle de l'information sur le comportement de recherche et la demande pour la formation professionnelle

**Mots clés :** Recherche d'emploi, information, économie comportementale, évaluation de politiques publiques

**Résumé :** Cette thèse explore l'impact de l'information sur le comportement de recherche d'emploi et la demande pour les formations professionnelles. Le premier chapitre évalue l'impact du site internet Bob Emploi, qui vise à délivrer de l'information aux demandeurs d'emploi à propos du marché du travail. Les résultats indiquent qu'il n'y a aucun impact sur l'effort de recherche des demandeurs d'emploi et le périmètre géographique et sectoriel de la recherche. Cependant, les demandeurs d'emploi ayant recours à Bob Emploi mobilisent davantage leur réseau personnel ainsi que les services publics de l'emploi. Enfin, il n'y a aucun effet sur le bien-être et sur le retour à l'emploi. Le deuxième chapitre examine le rôle de l'information sur l'entrée en formation professionnelle. Les résultats indiquent que la réception d'un email avec un message mettant l'accent sur les opportunités de retour à l'emploi après la formation fait plus que doubler la probabilité que les demandeurs d'emploi rappellent le centre de formation. Cependant, les taux d'appel sont faibles en valeur absolue (moins de 1%) et il n'y a aucun impact sur l'inscription en formation. Nos résultats suggèrent que l'impact détecté sur les appels est davantage dû à l'augmen-

tation de l'importance accordée aux informations sur la formation plutôt qu'à la mise à jour des croyances des demandeurs d'emploi. Enfin, le troisième chapitre étudie également la demande pour la formation professionnelle, mais prend en compte les contraintes comportementales. Distinguant les croyances "externes" (sur le monde) et les croyances "internes" (sur soi-même), les résultats montrent que les demandeurs d'emploi subissent des contraintes financières les empêchant de rejoindre un programme de formation, et qu'ils sous-estiment la proportion de formations disponibles qui sont financées. Les obstacles internes liés à l'auto-efficacité, aux préférences intertemporelles, à l'estime de soi et à la capacité d'organisation sont mentionnés à part égale par les demandeurs d'emploi indiquant avoir des obstacles internes à l'inscription en formation. À partir de ce diagnostic, la dernière partie est consacrée au design d'un essai randomisé contrôlé, avec des interventions reposant sur la transmission d'informations via des cours en ligne, et des sessions interactives par groupes de demandeurs d'emploi. Ces cours visent à cibler les croyances externes, internes, ou les deux simultanément.

**Title :** Essays on the role of information on job search behavior and the demand for vocational training

**Keywords :** Job search, information, behavioural economics, policy evaluation

**Abstract :** My dissertation examines the impact of information on job search behavior and demand for vocational training. The first chapter evaluates the impact of the website Bob Emploi, which aims at delivering news to jobseekers about the labour market. Results indicate that there is no impact on jobseekers' search effort and search scope. However, job seekers using the website are more likely to rely on personal networks and to use resources provided by public employment services. Finally, there is no effect on self-reported well-being and on employment. The second chapter focuses on the role of information about training on the enrollment rate. Results indicate that receiving an email with a message emphasizing training returns in terms of employment more than doubles the likelihood that job seekers call back the training center. However, callback rates are low in absolute value and there is no impact on enrollment. Our results suggest that the effects on callbacks are driven by increasing salience of basic information about training

rather than by belief updating. Finally, the third chapter focuses on the demand for vocational training as well, but takes into account behavioural constraints. Distinguishing between "external" beliefs (about the world) and "internal" beliefs (about the self), results show that jobseekers experience financial constraints preventing them from joining a training program, and that they underestimate the proportion of subsidized programs available to them. Obstacles related either to self-efficacy, self control, self esteem and executive function are equally mentioned among jobseekers reporting internal barriers in training enrollment. In light of this diagnosis, the last part is dedicated to the design for a randomized controlled trial, with interventions relying on the delivery of information through app-based courses, and interactive sessions involving groups of jobseekers. These courses aim at targeting either external or internal beliefs, or both of them simultaneously.