CORRECTING BELIEFS TO INCREASE HEALTH INVESTMENTS: A FIELD EXPERIMENT AMONG DISADVANTAGED YOUTHS*

Bruno Crépon† Julie Pernaudet‡

October 28, 2019

Abstract

After showing that unmet health needs are prevalent among unemployed French youth, we assess two policies aimed at increasing their health investments. Some young people received personalized guidance on health insurance; others were additionally examined by a physician who referred them to adequate practitioners. While the first intervention has virtually no impact, the second reduces the perceived costs of healthcare and doubles the probability of consulting a psychologist. We use machine learning to test the belief revision channel and show that the latter impact can be more than twice larger for people who underestimate their mental health needs the most.

JEL Codes: I12, I13, J24 D83

Keywords: healthcare utilization, Bayesian learning, LASSO, heterogeneity

*This experiment would not have been possible without the support of the French Experimental Fund for the Youth (Fonds d'Expérimentation pour la Jeunesse). We also thank the Labex-ECODEC for its financial support and the Institutional Review Board of the Abdul Latif Jameel Poverty Action Lab (J-PAL) for its approval. The description of the project can be found under the AEA RCT identification number 0000761. The project has been co-implemented by Bruno Crépon, Simon Beck and Lucile Romanello at the Center for Research in Economics and Statistics (CREST); Didier Dugast, Virginie Kergoat and Joël Dutertre at the youth job center of Sénart; Pierre Chauvin and Sophie Lesieur at the French Institute of Health and Medical Research (INSERM). We are very grateful for the support of the five youth job centers which worked on this project, in particular to the physicians who took part in the experiment. We also thank the research assistants who worked on this project: François Dhont, Franck Gasser, Cécile Montastier, Elodie Pétorin, and Teldja Seniguer. We are grateful to participants of the 2016 Advances with Field Experiments conference, the 2016 North American Summer Meeting of the Econometric Society, the 2016 Annual Conference of the Royal Economic Society, to the 2015 J-PAL Field Experiments in Labor Economics and Social Policies Conference, the 2015 HCEO Summer School as well as seminar participants at the University of Chicago, CREST, Ecole Polytechnique, and Paris Dauphine University. We also thank Juanna Schrøter Joensen and Eric Maurin for their comments.

†CREST and J-PAL

‡University of Chicago, 5757 S University Ave, Chicago, IL 60637, jpernaudet@uchicago.edu

1 Introduction

Unemployment of low-skilled youths has been extensively studied, but one dimension of their human capital is often overlooked despite its potentially harmful effect on employability: their health capital. Like other dimensions of human capital, health is an input in which people from different socioeconomic backgrounds are likely to invest differently. If disadvantaged young people accumulate lower health capital, then poor health can be an additional barrier they face to enter the labor market (see García-Gómez et al. (2010), Cai (2010), Cai and Kalb (2006) or Prinz et al. (2018) for a review of the impact of health on labor market outcomes). In this paper, we study health investments behaviors of low-skilled unemployed youths in France and try to better understand the factors affecting their health decisions.

Health decisions have multifaceted determinants, ranging from financial access to healthcare to sociocultural factors such as the subjective representations of health and illness (e.g., which symptoms are normal or not), social norms about expected health behaviors, and trust in medical institutions, for example. A lot of attention has been devoted to the question of financial access to healthcare and the role of insurance (e.g. Baicker et al. (2013), Finkelstein et al. (2012), Kolstad and Kowalski (2012), Buchmueller et al. (2004)), but less is known on the role of non-financial determinants of health investments. In this paper, we examine both financial and non-financial barriers to healthcare utilization. In the French context where health coverage is universal and generous, the financial barrier should in theory be low. However, we show that disadvantaged youth tend to be unaware of their coverage and of the different insurance options (including free public insurance) they have, resulting in suboptimal coverage. On the non-financial side, we show that targeting people's beliefs about their health needs and their knowledge of the resources available to meet them can significantly raise health investments.

Our focus is on job seekers attending public employment agencies dedicated to unemployed people aged 16 to 25 (youth job centers) in France. It is well documented that people attending those job centers on average are very low-skilled, are low income, and are more exposed to stressful social or family circumstances. What is less documented is their state of health, that may also be a barrier to their socio-professional integration. In recent years, an increasing number of youth job centers have developed health services such as free consultations with a general practitioner

or a psychologist, partnerships with health centers or hospitals, and information campaigns. However, little is known about the impact of those initiatives on youth health investments and labor market outcomes. The experiment we ran aims at evaluating two possible interventions, one targeting exclusively health insurance and the other targeting both health insurance and perceptions of health needs and resources. In parallel, through the collection of unique data on individualized medical examinations, we also intend to better document the nature and extent of the different health issues faced by this population. The experiment was implemented between 2011 and 2013 in five French job centers, where young people were randomly assigned to one of the two interventions or to the control group.

In the first program, we provide young people with personalized information on their insurance benefits and on the procedures required to apply for free public insurance, if they are eligible for it.[1] Although the information provided is freely available, it may be hard to find or understand for disadvantaged youths, and that may lead them to make suboptimal decisions (Bhargava et al. (2017), Kling et al. (2012)). As lifting the financial barrier to healthcare is a prerequisite for observing increases in health investments, we maintain the insurance information component in the second program. Nonetheless, to address additional misperceptions young people may have of their health needs, the second program additionally encourages them to meet with a physician who makes a comprehensive diagnosis of their health and refer them to health services corresponding to their health issues, if any. A specificity of the intervention is that it does not target one specific health issue. Instead, it addresses the multiplicity of health problems disadvantaged youths may face, from sexually risky behavior or drug use to dental problems or depression. Another key aspect of this intervention is that the information provided is based on each person's specific situation and is thus more accurate than general information provided in a prevention campaign or a brochure for example (see Dupas and Miguel (2017) for a discussion of the impact of tailored information in health programs). Comparing the impacts of the two programs, our goal is to distinguish between barriers that are due to a misunderstanding of the health insurance system, and barriers that are due to biases in the subjective assessment of one's needs.

Underestimation of health needs might indeed be particularly prevalent among the deprived pop-

[1] In France, since 2000, the government provides on a means-tested basis a free health insurance that fully covers medical expenditures and exempts from advancing payment. According to the French Court of Audit (Cour des Comptes (2015)), approximately one-third of the eligible population does not benefit from it.

ulation we target. Subjective health assessment, at constant objective health status, has been found to be highly correlated with socioeconomic and cultural background (Etilé and Milcent (2006), Shmueli (2003), Mackenbach et al. (1996)). More generally, White (2015) suggests that representations and attitudes towards health, illness and pain, and subsequently health behaviors, can greatly differ between cultural groups. Studies in Australia and India find for instance that it is paradoxically among the subpopulations facing the biggest health issues that self-reported health tends to be the highest (Mathers and Douglas (1998), Murray and Chen (1992)). Sen (2002) puts forward a possible interpretation of this result in terms of reference point: taking an extreme case, someone surrounded by people in poor health, for instance, could tend to consider poor health as a normal state, and under-invest in his or her health as a consequence.

To assess the perception revision channel and hence better understand how the physician's information affects young people's health investments, we write a Bayesian learning model and test its predictions. A central prediction of the model is that the impact of the intervention is a function of the physician's referrals and people's own perceptions of their health status. The larger the gap between the physician's signal and the person's perception, the higher the impact. We measure health perceptions in the whole sample but the referrals are by construction observed for treated youths only. We thus use a LASSO regression on the various baseline measures of young people's socioeconomic environment and healthcare utilization to predict a counterfactual signal for the untreated sample. This way, we can test the prediction of the model by simply interacting the treatment variables with the predicted signal and with people's baseline perceptions.

The analysis gives rise to four main empirical results. Thanks to the medical diagnoses made by the physicians, we first find evidence of unmet health needs for a vast majority of young people; physicians directed them to another health practitioner in 82% of the cases. Physicians data also reveal that three major concerns in this population are risky sexual behaviors, for 31% of the examined young people, dental problems, in 30% of the cases, and poor mental health, with 22% of the examined young people directed to a psychologist and 17% diagnosed with depression. Second, we find that the program combining insurance information and physician recommendations strongly reduces the perceived cost of healthcare and increases health investments for some issues the physicians detected but not all. It has no impact on the probability of consulting a dentist and only a small impact on the likelihood to use contraception, but it leads to a large

increase in the probability of consulting a psychologist, from 9 to 17%. In addition, the recipients of the combined program tend to have a higher probability of undertaking job training programs, suggesting that the program could contribute to skills formation. Third, heterogeneity analyses show that the impact on the probability of consulting a psychologist varies with the extent to which young people lack awareness of their health needs. As Bayesian learning would predict, we find that the larger the gap between their initial perceptions of their mental health and the personalized signal they receive from the physician, the larger the impact of the intervention. Hence, impacts on the probability of consulting a psychologist are above 0.15 for more than 10 percent of our sample, versus 0.08 for the average treatment effect. Fourth, the limited program providing only information on health insurance does not lead to any health investment - the room of improvement of the level of coverage was low, and young people's understanding of the health insurance system is not affected.

The paper mainly relates to three strands of the literature. It contributes first to the literature investigating how subjective perceptions affect human capital decisions and how information may change those perceptions. In the field of education, the importance of the perceptions of costs and returns on schooling investments is documented in a variety of contexts (Wiswall and Zafar (2016), Attanasio and Kaufmann (2014), Delavande and Zafar (2019), Jensen (2010)). In the health field, the role of subjective perceptions and information has been considered mainly for the analysis of sexual health and behaviors. In Malawi, Paula et al. (2014) find for example that downward revision of the perceived probability of being HIV positive leads to a higher probability of having extramarital sex and/or multiple partners, while upwards revision reduces this probability. From previous waves of the same survey, Thornton (2008) finds a small positive effect of learning about HIV positive results on the number of condoms purchased, and no effect of learning about negative results. Those impacts do not significantly differ depending on prior beliefs about HIV status. On the contrary, Gong (2015), studying the impact of HIV testing in Kenya and Tanzania, find that only people for whom the information provided by the test contrasts with their prior beliefs about their HIV status change their behavior. More generally, predicting how people process information related to themselves is difficult because they may incur disutility from learning about unfavorable news and revise their beliefs in unexpected ways (Eil and Rao (2011)). In our analysis, we provide new empirical evidence of a significant be-

havioral response to a negative signal about mental health, which is consistent with Bayesian revision of beliefs.

A second strand of the literature our paper contributes to is the literature studying the determinants of the take-up of social programs. Low take-up of social programs is a broad policy question that concerns not only public health insurance but also unemployment insurance, college financial aid, disability programs, food stamps and housing benefits (see Currie (2006) or Remler and Glied (2003) for reviews). Three main reasons why people who are eligible for social programs do not use them have been stressed in the literature: social stigma associated with benefiting from the program (Dahl et al. (2014), Keane and Wolpin (2010), Schanzenbach (2009)), transaction costs induced by the administrative procedures (Finkelstein and Notowidigdo (2019), Deshpande and Li (2017), Schanzenbach (2009), Aizer (2007)) and unawareness of program eligibility (Finkelstein and Notowidigdo (2019), Armour (2018), Daponte et al. (1999)). The interventions we test in this paper target the two latter barriers in the case of public health programs in France. We provide evidence that informing young people on the existence of public insurance, its eligibility rules and application procedures tend to raise the proportion of people benefiting from it. We also find that additionally informing them about the possibility to benefit from free health care strongly reduces the proportion of people thinking that money is necessary to get care.

Finally, our paper comes within the scope of the recent literature applying machine learning techniques to policy evaluation problems (see Athey and Imbens (2017) or Mullainathan and Spiess (2017) for reviews). The two approaches have been creatively combined in many different ways. The LASSO (Least Absolute Shrinkage and Selection Operator) has been used in particular to estimate average causal effects in the absence of experimental variation (Belloni et al. (2017), Belloni et al. (2012)), but also to detect heterogeneous treatment effects in the presence of numerous covariates (Imai et al. (2013); see also Athey and Imbens (2016) and Athey and Wager (2018) who rely on regression trees and random forests, or Chernozhukov et al. (2018) who propose an inference method for different features of heterogeneity estimated by machine learning techniques). Here, we use a LASSO regression in the treatment group sample to detect the best predictors of the signal they receive from the physician, and then estimate the counterfactual signal people who were randomly assigned to the other groups would have received if they had seen the physician. This estimated signal allows us to assess the prediction of our

theoretical model of health investment decision and better understand the operating mechanism behind the effectiveness of information provision. To our knowledge, the closest paper in spirit is Oster (2018), who uses random forest to predict (missing) diabetes diagnoses and evaluate the impact of the diagnoses on food purchases.

The remainder of the article is organized as follows. Section 2 presents in detail the interventions and the context in which they were implemented, section 3 describes the data collected and the population studied, section 4 develops a simple model of health investment decisions and makes explicit the way each of the two programs is expected to affect decisions, section 5 presents the econometric framework, section 6 gives the results, section 7 tests the predictions of the model and section 8 concludes.

2 Interventions

2.1 Context

The experiment was implemented in five youth job centers, which are centers where counselors welcome 16-25 year-old unemployed people who encounter difficulties with social and professional integration. Youth job centers provide them with individualized career guidance in order to help them define a professional project, find training to acquire adequate skills and ultimately find a job. According to the national authority that coordinates the entire network of youth job centers, approximately 1.4 million people were registered in youth job centers in 2011 in France. At that time, the unemployment rate (in the ILO sense) among 15 to 24 year-olds was around 22% according to INSEE[2]. Young people attending these centers form a subpopulation sometimes referred to as "Not in Education, Employment or Training" (NEET): they left the education system with very low qualifications on average, they are unemployed and we show that a substantial proportion suffered from personal disruptions during childhood or adolescence (see section 3.3.1). One mission of the youth job centers is to help them acquire job searching skills (how to present a curriculum vitae, how to write a motivation letter, what type of jobs

[2]National Institute of Statistics and Economic Studies

their qualifications correspond to). However, as observed by an increasing number of youth job centers, an upstream barrier to social and professional integration for these young people is their poor health and low use of health care services.

In 2006, the French Ministry of Health, the Ministry of Labor and the head of youth job centers jointly wrote a charter that reveals the increasing health inequalities between disadvantaged unemployed youths and their better-off counterparts.[3] This charter formulates four main objectives for youth job centers: to increase access to health insurance, to increase access to health care, to provide information on prevention and to favor detection of health problems. However, few health data are available on this population; to our knowledge, only one survey from 2011 investigated those issues.[4] Based on a sample of more than 4 000 young people attending youth job centers or community health centers, this survey shows that these young people have 1.5 higher risk of suffering from ill-being than active youths holding a permanent job, 1.8 higher risks of declaring health problems and two times higher rate of attempting suicide. Labbe et al. (2007) compare 105 901 unemployed youths with 138 344 students or active youths in community health centers in France and find that unemployed youths have lower use of health care, especially visits to a general practitioner, dentist, or gynecologist, lower use of contraception, and poorer perceived health. These results also echo Currie et al. (2012), who find that young people with disadvantaged socioeconomic backgrounds tend to have worse health, especially mental health, compared to their more advantaged counterparts.

2.2 Programs

In this experiment, we test two possible interventions aimed at raising health investments among underprivileged youths, one tackling solely (perceived) financial barriers to healthcare, the other tackling both financial barriers and biases in the assessment of health needs.

In France, the health insurance system is generous, but it may be difficult for young people to navigate through the system. The rules are the following: people can benefit from public basic health insurance either if one member of their household contributes to social security, which is

[3] *Charte de la santé des jeunes en insertion sociale et professionnelle* [in French]
[4] Study "Bien être - mal être ? Mieux vous connaître", December 2011.

automatic for working people and for people receiving unemployment benefits, or, under means testing, they can benefit from free public health coverage (CMU[5]). These two public basic insurances reimburse 70% of healthcare expenditures (minus a lump sum of 1€) as long as care is provided by a physician (consultations with a dentist, an ophthalmologist, a gynecologist, for example, are reimbursed, but consultations with a psychologist are not, except if the consultation is provided in a hospital or in a community health center). For example, for a visit to a general practitioner for which the patient has to pay 23€, basic insurances will reimburse 15.10€ (0.7*23 - 1€). The remainder can be reimbursed by a complementary insurance. Under a given income threshold (720€ per month in 2015), people can benefit from free public complementary insurance (CMU-C[6]) that provides full medical coverage and cash advance exemption for most healthcare expenditures (people under CMU-C policy are charged 0€ when they consult a health practitioner). In 2010, more than four million people were covered by CMU-C. The objective of this policy is to reduce financial constraints in access to health care services for the poorest segments of society. Nonetheless, as mentioned earlier, there is evidence of low take-up (Cour des Comptes (2015)), resulting in a non-negligible share of people being incompletely covered (see section 3.3.3 for detailed statistics among our sample). For people whose income is above the 720€ threshold, they have to pay for a private complementary insurance of their choice to be reimbursed for the remainder. Each complementary insurance offers different levels of coverage, for different types of healthcare expenditures.

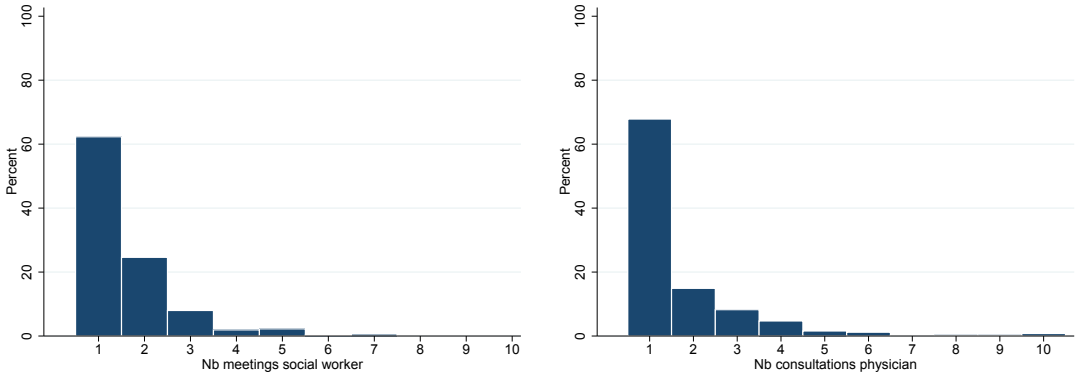
The objective of the first program is to inform young people about their health insurance options, in particular the CMU-C for eligible people, to help them understand the reimbursement rules and to reduce the perceived complexity associated with application procedures. It consists of encouraging young people to meet individually with a social worker present in the youth job center, whose role is to inform them on the most adequate health insurance given their economic situation and to explain the procedures required to obtain it. The social worker had access to the Social Security records and hence knew the exact insurance coverage of each young people. It consists only of information provision: the social worker did not complete the procedures in place of the young people. The functioning of the different types of medical coverage was

[5]Universal (basic) Medical Coverage
[6]Universal Complementary Medical Coverage

also explained to the young people in order to make sure they understood the benefits of the insurance. Young people could meet with the social worker several times during the year of the intervention if needed, but most of them (63.1%) met with the social worker only once, as shown in Figure 1.

In the second program, the objective is to provide personalized information on health needs and corresponding healthcare resources on top of the information about health insurance. It consists in encouraging young people to meet, in addition to the social worker, a physician (present at the youth job center as well) who is in charge of making a detailed evaluation of each person's health status (dental and visual acuity check-ups, BMI measures, psychological health assessment based on the Mini-International Neuropsychiatric Interview (Sheehan et al. (1997)), sexuality, tobacco, alcohol, and drug use assessments). Based on this evaluation, the physician informs young people on their specific health issues (if any) and refers them to the appropriate practitioner.[7] We keep the social worker program as a basis in this second program because physician's recommendations are useless if young people believe they cannot afford the consultations. Young people could meet with the physician several times during the year of the intervention if needed, but 67.8% met with the physician only once (Figure 1).

Figure 1: Programs intensity



Both programs are entrusted to experts (a social worker, a physician) and personalized (individual meetings) in order to make the information as accurate as possible. They have been designed

[7] The physician could not provide care or make appointments for young people, he was only informing them on which type of practitioner they should consult.

in accordance with researchers from the National Institute of Health and Medical Research (INSERM) and with the physicians of the five youth job centers. There was one physician and one social worker per site, who were present on average two days a week at the job center. The cost of such programs is mainly driven by salaries: having a physician working in the job center two days a week costs approximately 46 000€ per year and a social worker working in the job center two days a week costs approximately 20 000€ per year.[8] A back of the envelope calculation suggests that the first program (social worker only) would cost approximately 14€ per young person per year, assuming each young person meets the social worker only once and spends on average 30 minutes per appointment. Under these two maintained assumptions applied to both the social worker and the physician, the second program would cost approximately 47€ per young person per year.

People assigned to the control group kept seeing their counselor as any other young person in the job center and as people assigned to one of the programs, but they did not receive any encouragement to meet the social worker or the physician.

2.3 Encouragement protocol

The two treatments were only encouragements, they were not mandatory, and no one was excluded from taking them. Randomization was carried out separately in each youth job center at the individual level among the 1528 young people initially present in the experiment (see section 3.1). It resulted in 501 young people assigned to the social worker program, 532 assigned to the combined program involving both a social worker and a physician, and 495 assigned to the control group. The research assistant present in each job center was in charge of planning the appointments between young people and the physician and/or social worker for people in the treatment groups immediately after they answered the baseline questionnaire. Then the research assistant would call them the day before the meeting to remind them about the appointment. In case they did not come to the appointment, the research assistant was responsible for calling them and asking why they did not come and making a new appointment. This protocol was

[8] These numbers have been provided by the youth job center of Sénart who was in charge of the experiment budget for the five youth job centers involved.

repeated one more time if they did not come to the second appointment. In case young people were not reachable, the research assistant would call them up to five times in a period of two weeks, leaving a message each time on the answering machine. Table 1 presents the take-up rates we observe for each group, as well as the expected take-up rates in the case of perfect compliance (in parentheses below). It shows that the take-up rate is high for the social worker program, with 82.04% of the people assigned to this program who actually met the social worker. Take-up rate for the combined program involving both a social worker and a physician is lower, with 77.63% of the people assigned to this program who met the social worker and 67.29% who met the physician (63.34% met both). No people assigned to the control group met the physician, and only 0.61% of them met the social worker. Approximately 0.20% of the people assigned to the social worker program met the physician.

Table 1: Take-up rates

	SW group	SW+P group	C group
Met physician	0.20	67.29	0.00
	(0%)	(100%)	(0%)
Met social worker	82.04	77.63	0.61
	(100%)	(100%)	(0%)
N	501	532	495

Note: This table gives for each of the three groups the percentage of young people who met the physician and the social worker. In parentheses are the percentages that we would have expected in case of perfect compliance to random allocation.

In order to characterize the young people who took the treatment versus those who did not, we compare them along various baseline characteristics for each of the two treatment groups (Appendix I). People assigned to the combined program or to the limited program who took the

program are similar to those who did not take it according to socioeconomic characteristics and insurance situation. People who took the social worker program tend to have better perceived psychological health and to be more optimistic about their future health than people who did not take it. They are also more likely to have a regular doctor. People who took the combined program differ from those who did not only in their perception of how their body withstands illness.

In the endline survey, we asked treated young people a few questions about their opinion on the intervention they benefited from. These statistics should be interpreted with caution because of the face-to-face nature of the survey and the desirability bias it is likely to induce. However, among people who met the social worker, 78.8% declared the program was useful because they needed it, and 54.5% declared they needed it but did not want to go through the administrative procedures. Among this latter proportion of people, 64% declared the program convinced them to exercise their insurance rights. This latter statistic is in line with the increase in the proportion of CMU-C recipients we observe in both treatment groups (see section 6). Among people who met the physician, 76.8% declared that the program was useful because they needed it, and 39.1% declared they needed it but did not suspect it before the consultation(s). Among those who declared the program was useful because they needed it, 63.4% declared they would not have consulted a physician if they had not been encouraged to. These statistics, although they might be positively biased, tend to suggest that the program raised awareness about health needs for a large proportion of young people.

3 Data and population

3.1 Sampling of the Job Centers and young people

The selection of the five youth job centers which participated in the experiment resulted from the following procedure. First, we sent the 420 existing youth job centers in France an electronic letter presenting the project and asking them to answer a few questions on health services they had (if any). We received 280 questionnaires back, of which 99 youth job centers agreed to

volunteer to participate in the experiment - a high number that confirms that the project is relevant to the youth job centers. To be eligible, the condition was to have a physician working *inside* the job center.[9] Having a physician inside the youth job center minimizes the costs of the consultation, both in terms of transportation and in terms of time, which may be significant barriers if people have low expectations on the benefits of the consultations. We imposed this condition to maximize the take-up of the program. Among the 99 youth job centers which volunteered to participate in the experiment, only five had a physician working inside the job center. These five youth job centers are located in the Paris area (Clichy-sous-Bois and Sénart), in the north east (Reims), close to the center (Poitiers) and in the south of France (Toulouse) and reflect a certain diversity (cities of different sizes, different regional economic contexts). They are five youth job centers that are particularly involved in youth health issues, and as such, they are not representative of the 420 youth job centers. Nonetheless, there is no reason to think that young people attending these job centers have specific characteristics.

Inside these five job centers, the young people who participated in the experiment were selected using the following procedure. First, thanks to youth job centers recordings, we identified all the young people who came for the first time at least twice to the job center in a period of four months. This was motivated by the fact that we did not want to include people who would come only once to the center and never return, and who would have been very hard to convince to answer the survey and to see the social worker or physician. We thus identified 3 555 young people in the five youth job centers. The research assistants present in each center had to call them to present the survey (in a standardized way, see Appendix II) and ask them whether they would agree to participate in the experiment. At that stage, they excluded young persons who had an illness requiring immediate intervention and persons who did not speak French sufficiently well to understand the questionnaires. Approximately 7% of the 3 555 young people were excluded based on these two criteria. Among the remaining young people, 11% explicitly refused to participate in the experiment and 34% were impossible to reach (after five calls with messages left on the answering machine in a period of two weeks), which we interpreted as an implicit refusal. Finally, 5% of the young people were randomly excluded because the research

[9] Many youth job centers partner with physicians but they work outside the job center, typically in a hospital or in an association, and counselors refer young people to these places.

assistants were overwhelmed. The sample thus finally consisted of 1 528 young people, who signed a consent to state their willingness to participate in the experiment. As we excluded people with great difficulties in French and as people who were impossible to reach could not be included, our sample might miss some of the most marginalized young people.

Power calculations suggest this number of observations, for a level of statistical significance of 5% and power of 80%, would allow detecting effects as small as 0.175 standard deviation when we compare physician and control groups, 0.177 standard deviation when we compare social worker and control groups and 0.174 standard deviation when we compare physician and social worker groups. These minimum detectable effects are considered as a "small" effects (Cohen (1988)), which justified the continuation of the experiment.

3.2 Data collected

The first type of data we collected came from questionnaires the 1 528 young people answered face-to-face with a research assistant. We first surveyed them at the very beginning of the experiment, in 2011, before they knew which group they would be assigned to, and we again surveyed them one year later, after the interventions took place.[10] In these questionnaires, we ask young people about several aspects of their everyday life such as family, neighborhood, holidays, money issues, housing, sports, and cultural activities. We also ask them a set of questions aiming at capturing their level of social integration such as their number of friends, their feelings and experiences of discrimination, the support they have from close relatives, their level of confidence in various institutions (school, police, media, justice, government, health system, political parties). Finally we ask them a series of questions aimed at capturing their health status (perceived health and vulnerabilities based on the Health Perceptions Questionnaire (Ware Jr (1976)), mental health based on the Beck Hopelessness Scale (Beck et al. (1974))) as well as questions about their utilization of healthcare.[11]

A precise protocol was established in order to maximize response rates to the follow-up survey.

[10] A pilot study had been previously implemented to test the questionnaires.

[11] A complete list of the endline survey questions can be found in Appendix III. The endline questionnaire has been designed in partnership with INSERM researchers. It hence contains both questions aimed at capturing the impact of the intervention on the main domains of interest (medical coverage, health investments, health and labor market outcomes) and questions intended for complementary analyses by the INSERM team.

First, between the baseline and the endline surveys, research assistants had to establish a contact (a call or a short discussion in the job center) with each young person at least every three months in order to remind them about the experiment. Then, three weeks before the date when the young person was to answer the final questionnaire (one year after she answered the initial questionnaire), the research assistant had to contact her to make an appointment for answering the final questionnaire. If the young person was not reachable after four calls (on different days of the week and at different moments of the day) and had no appointment planned with a counselor at the job center in the forthcoming weeks, the research assistant tried other ways of reaching the person (the second phone number or the email she had to provide at the beginning of the experiment, and then online networks such as Facebook, LinkedIn, Badoo, etc.).

The second type of data that we have was collected by the physicians and the social workers only for the young people they met as part of the experiment. The physicians had to make an assessment of the state of health of each young person (see section 2.2) and to report the medical practitioners they recommended to them for a follow-up health consultation (general practitioner, dentist, optician, psychologist, gynecologist, dermatologist...) or the medical analyses they recommended to them (blood test, radiography). The social workers also had to report the medical coverage of each young person they met before implementing the program.

All the questionnaires and reporting materials have been conceived in accordance with INSERM researchers and physicians of the five youth job centers.

3.3 Descriptive analysis

In this section, to better understand the specificity of our sample, we first compare our baseline survey to another study that samples disadvantaged people living in the Paris area and asks the same series of questions on socioeconomic background. We show that young people in our sample tend to have a particularly deprived family environment and have experienced difficult life events. Second, we exploit the data collected by the physicians to identify the main health problems faced by young people in this population and document the importance of underconsumption of healthcare. Medical data reveal that dental and mental problems, risky sexual behavior, and to a lesser extent overweight are the most prevalent issues. Finally, we document the possible

hindrances that may prevent young people from using healthcare and show that perceived costs of healthcare and underestimation of health needs are important issues in this population.

3.3.1 Disadvantaged population

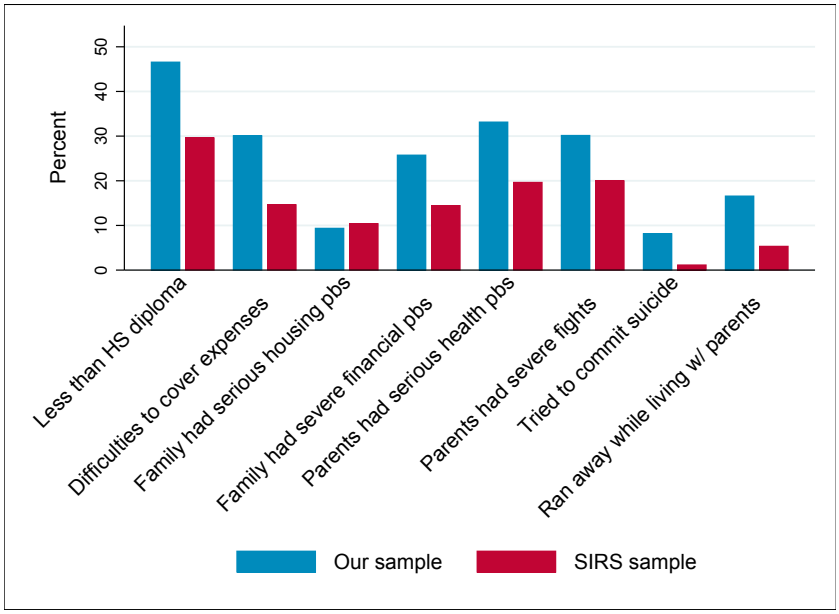
As already mentioned, young people attending youth job centers often are those who accumulate socioeconomic difficulties, such as a low level of education, financial difficulties and a poor family environment. To better assess their specificity, we compare our statistics to SIRS[12] 2010 survey, which was implemented by the INSERM researchers who participated in the design of our questionnaires as well. This survey is representative of people living in disadvantaged neighborhoods in the Paris area. We restrict the sample to people of the same age as our population, that is to say, people who were between 16 and 25 years of age in 2010 (208 observations). Figure 2 compares the answers to a series of questions on socioeconomic background that are common to our survey and SIRS survey. It indicates that in our baseline sample (1528 observations), 47% of the young people do not have a high school diploma, compared to 30% in SIRS (and 29% in the general French population of 20 to 24 year-old people in 2011 according to INSEE). They are also 33% more likely to declare they have difficulties reading or writing in French (not presented in the graph because not in SIRS survey). These low skills are likely to reduce their labor market prospects, but the literature suggests that low-educated people are also likely to have lower health investments (Cutler and Lleras-Muney (2010), or Lochner (2011) for a review) and to be less responsive to health information campaigns (Aizer and Stroud (2010)). Figure 2 also shows that 30% of them declare they have difficulties covering daily expenses (that is twice more than in SIRS), which also tends to be associated with lower use of healthcare and preventive care services (Devaux (2015), Van Doorslaer et al. (2006)).

A growing number of epidemiological studies highlight the importance of other determinants of healthcare use such as personal and social disruptions, which have rarely been considered in the literature. Bazin et al. (2005) show that people who faced serious disruptions during childhood or adolescence such as being placed under social services guardianship, having serious financial or housing difficulties, being convicted of a crime, or having made attempts to commit suicide

[12] Health, Inequalities and Social Isolation

tend to have low use of health care services. Here again, our sample of young people has been more exposed to such extreme events during childhood or adolescence than SIRS young people. Figure 2 shows that 9% of the young people in our sample declare their family had serious housing problems (close to SIRS), 26% stated their family had serious financial problems (compared to 15% in SIRS) , 33% declare their parents had serious health problems (20% in SIRS) and 30% recount their parents had severe fights (compared to 20% in SIRS). Finally, 8% declare they tried to commit suicide and 17% declare they ran away from their parents' home at least once, which is again much higher than for SIRS people. We can also add that for 46% of the young people in our sample, their parents are separated, divorced or never lived together; for 10% of them, either their father or mother have died, and 5% of them have no contact anymore with one or both of their parents or never got to know them (not presented in the graph). These figures suggest young people attending youth job centers tend to have an extremely precarious background and difficult life experiences, which is likely to negatively affect their health and in particular their mental health.

Figure 2: Baseline socioeconomic characteristics



3.3.2 Untreated health issues

As explained in section 2.2, the physicians made an evaluation of each young people health and based on this evaluation, told them which practitioner they should consult (if any). Unemployed young people are rarely surveyed, so collecting medical data and providing a detailed description of their health is the first step towards better understanding the difficulties they may face. Table 2 gives the main recommendations made by the physicians and compares them to young people's healthcare utilization in the year before the experiment. In this table, looking at the first column, the first striking finding is that the vast majority of young people needed healthcare, as suggested by the fact that almost 81,7% of the young people who met the physicians have been directed to at least one practitioner. More precisely, we see that the main recommendations were to consult a dentist, a general practitioner, to be tested for HIV, to consult a gynecologist, a psychologist, or an ophthalmologist.[13]

In the second column, we now compare these recommendations to young people's healthcare consumption in the year before they saw the physician in the experiment. We define underconsumption as not having consulted a given practitioner in the past year (or not having been tested for HIV) while the physician recommends it. The table shows that underconsumption rates are very high for HIV screening and for ophthalmologist and psychologist consultations; 74% of young people did not consult a psychologist in the past year while the physician in the experiment thinks it would be useful.[14] However, it is important to distinguish between unmet preventive health needs, that are not necessarily prejudicial to find a job on the short run, and unmet curative health needs. In other words, we need to distinguish between recommendations that correspond to a simple check-up and recommendations that correspond to an identified health issue that could be detrimental in the labor market. The next table presents the main health issues detected by the physicians and suggests that the ophthalmologist referrals, in particular, are primarily for check-up visits.

[13] The category "another practitioner" includes X-ray, pediatrician, dermatologist, cardiologist, rheumatologist, psychiatrist, ENT specialist, surgery, anesthetist, phlebologist, pulmonologist, diabetologist, stomatologist, orthodontist, gastroenterologist, endocrinologist, neurologist, addictologist, allergist, physiotherapist, dietitian, and speech therapist. Each of these practitioners was recommended in less than 5% of the cases (and even less than 1% for most of them), which justified that we gathered them.

[14] It should be noted that it is possible for young people to have free and anonymous HIV screening in screening centers and free consultations with a psychologist in health centers or hospitals.

Table 2: Baseline physicians' recommendations and healthcare consumption

Physician recommendation	Share (%)	Underconsumption (%)	N
Directed to dentist	39.1	59.8	338
Directed to GP	33.3	27.7	336
Directed to HIV screening	32.9	87.9	277
Directed to gynecologist	28.5	61.2	172
Directed to psychologist	21.5	74.0	339
Directed to ophthalmologist	16.5	80.0	333
Directed to another practitioner	38.8	67.7	335
Directed to any practitioner	81.7	14.1	339

Note: Share gives the proportion of young people for whom the physician recommended a consultation. Underconsumption is for example not having consulted a psychologist in the year before meeting the physician while the physician recommends it. N gives the number of observations.

Table 3 presents the diagnoses made by the physicians. The table first shows that almost one out of three people engages in risky sexual behaviors, according to the physicians.[15] This assessment is based on a series of indirect questions that the physicians asked about the use of condoms and number of partners. This figure echoes the large proportion of young people directed to HIV screening and gynecologists in Table 2. Dental problems, defined as a decayed tooth or a tooth to extract, also concern almost one-third of the young people and could be detrimental in the labor market, as found in Glied and Neidell (2010). This proportion is smaller than the proportion of people directed to a dentist (Table 2), which suggests that the proportion of referrals contains check-up visits. Based on the Mini-International Neuropsychiatric Interview

[15] Only sexually active people are considered, which reduces the sample to 276 observations.

scale (see section 2.2), physicians detected a depression among 17.4% of the young people they met. Relying on a similar assessment tool, the Composite International Diagnostic Interview - Short Form, the 2017 *Health Barometer* national survey shows that the prevalence of depression is 11.7% among the 18-24 years old population (see Léon et al. (2018)).[16] Here again, 17.4% is smaller than the proportion of young people directed to a psychologist, but mental health issues can be multiple and encompass a much larger range of problems than just depression. Studies of the impact of mental health on labor market outcomes are rare because of measurement and endogeneity challenges, but the few papers trying to tackle those challenges tend to find a strong negative impact of poor mental health (see Prinz et al. (2018) for a review). Frijters et al. (2014), in particular, find that a one-standard-deviation decline in mental health reduces employment by 30 percentage points, using the death of a close friend as an instrument for mental health variability.

Table 3 also presents the weight distribution, which reveals that a non-negligible share of people is overweight or even obese according to their BMI. Nonetheless, when we compare our figures to the 2003 National Health Survey (the most recent available INSEE Health Survey) which sampled 40,900 persons who were representative of the overall French population, being overweight is 8 percentage points less prevalent in our sample than in the general population. Obesity rates are similar (12.6% in our sample versus 10% in the general population), and being underweight is more frequent in our sample (6.8% versus 4%).[17] Finally, we see that 15.1% have to renew their vaccinations, which is part of the share of young people directed to a general practitioner in Table 2. Only 8.2% have visual acuity lower than 8/10 (the average acuity in our sample, with or without glasses or contact lenses, is 9.5/10), suggesting that most of the recommendations to consult an ophthalmologist were for check-up visits.

[16] We can also refer to Leray et al. (2011) who analyze a representative national survey using the same tool as we use, the Mini-International Neuropsychiatric Interview scale, and find that the prevalence of depression is 11.0% on average among the 18-75 years old population between 1999 and 2003.

[17] It is interesting to note that studying the 2003 INSEE Health Survey, de Saint Pol (2009) finds that social inequalities are particularly important for dental health, obesity and HIV testing, which are also among the main health issues diagnosed by the physicians in our experiment.

Table 3: Baseline prevalence of health issues

Physician diagnosis	Prevalence (%)	N
Risky sex behavior	31.2	276
Dental problem	30.0	303
Depression	17.4	327
BMI < 18.5 (underweight)	6.8	325
25<= BMI < 30 (overweight)	19.7	325
BMI>30 (obesity)	12.6	325
Vaccines not up to date	15.1	325
Visual acuity <8/10	8.2	280

Note: Prevalence gives the proportion of young people diagnosed with a health issue by the physicians. N gives the number of observations.

3.3.3 Barriers to healthcare use

There are many factors that can lead people to underuse healthcare services. In this section, we first discuss the barriers reported by young people in the baseline survey, and then cross the survey data with the data collected by the social workers and physicians to document the gaps between young people's perceptions and their actual situation in terms of medical coverage and health issues.

In the baseline survey, we ask young people whether they forwent healthcare in the past year, and if yes, for what reason(s). Among the 1520 respondents, 28.2% declare that they forwent healthcare. Among those 28.2%, the main reasons young people reported are that they face financial barriers (37.8%) and that they prefer to wait for things to improve on their own (29.4%).

Other possible reasons such as the complexity of the healthcare system, the wait times, the fear of consulting a practitioner, the lack of trust or lack of time do not seem to be major barriers, young people selecting them in less than 10% of the cases. To further investigate the financial barrier, we review the medical coverage data collected by the social workers in collaboration with the social security administration. Those data show that 5.7% of the young people have no medical coverage at all, and 12.9% solely have basic coverage (see section 2.2), which means that only 18.6% of the young people in our sample have to pay when using healthcare. But if we now compare those statistics with young people's declared medical coverage in the baseline survey, we observe discrepancies showing that young people believe they are less covered than they really are.[18] Indeed, 7.9% declare they have no medical coverage at all, and 15.2% say they solely have basic coverage, which means that 23.1% of the young people believe they have to pay when using healthcare, compared to 18.6% according to the social security records. The baseline survey also reveals that 13.8% declare they do not know whether they are covered or not.

Another possible factor that could explain the prevalence of untreated health issues documented in the previous section is that young people may underestimate their health needs. Such cases would not be reflected in the question about healthcare forgoing that we previously discussed, as one must first be aware of one's needs in order to declare forgone care. To assess the extent to which young people underestimate their health needs, we compute the proportion of people who self-rate their health as good in a given dimension among people diagnosed by the physicians as having a health issue in this dimension. We observe young people's perceptions in the baseline survey for dental health, mental health and weight. On those three dimensions, we find that the proportion of people underestimating (according to the above definition) their dental issues is surprisingly high, with 62.2% of the people diagnosed with a decayed tooth or a tooth needing to be extracted self-rating their dental health as good or very good. A direct consequence is that they are very likely to underestimate their needs of dental care. Among people diagnosed with depression by the physicians, 20.0% self-rate their psychological health as good or very good, suggesting that they do not realize their mental health is poor. If we look among people who have been directed to a psychologist by the physician, which encompasses a broader range of

[18] We restrict the baseline sample to the subsample of young people for whom the social worker collected their medical coverage, i.e. 630 young people, for comparability.

possible mental health issues than just depression, 34.7% self-rate their psychological health as good or very good. Finally, among people identified as being overweight or obese, 15% self-rate their weight as normal or even too thin.[19]

In section 7, we explore the extent to which the impact of the combined program varies with the gaps between young people's perceptions of their health and the signal they receive from the physician. We show that in the case of mental health, the larger the gap, the bigger the treatment effect.

4 Theoretical framework

In this section, we write a simple model that links perception of health needs and cost of health investments to health behaviors. The goal of this model is to draw testable predictions that unravel possible mechanisms through which our interventions may operate on health behaviors. The first straightforward prediction of the model is that the lower the cost, the higher the probability of investing in one's health. A second prediction is that the larger the gap between the information provided by the physician and people's initial perceptions of their health, the larger the impact of the physician intervention on the probability to invest. In the empirical part, we test these predictions to assess the consistency of observed behaviors with such mechanisms (see section 7).

4.1 Setting

We denote individual i's state of health h_i, from which we define being in poor health as $\underline{h_i} := (h_i < 0)$. People judge their health based on the information I they have. Individual i's subjective probability of being in poor health is then denoted $Pr(\underline{h_i}|I_i)$. In the experiment, people who

[19] One may argue that a more fundamental determinant of healthcare underuse could be that young people simply do not value their health status, but several questions from the baseline survey indirectly capturing health valuation tend to discard this possible explanation. When asked about their attitude towards health, 56.5% of the young people indeed declare that they get easily anxious when something is wrong, and 30.4% declare that they worry about their health more than most other people do. 43.8% of young people also declare that they do something to improve their health on a regular basis (monitoring their diet or exercising, for instance), which suggest that they value being in good health.

meet the physician receive information I_i^1 on their health, whereas people who do not meet the physician evaluate their health status only based on their own perceptions of their health, I_i^0. Their respective subjective probabilities of being in poor health are then $Pr(\underline{h_i}|I_i^0, I_i^1)$ and $Pr(\underline{h_i}|I_i^0)$.

When in good health, or in poor health but cured, people have high utility level $\overline{U_i}$, whereas when they are in poor health and do not get cured, they have low utility level $\underline{U_i}$. The monetary cost of the cure is denoted c_i.[20]

Individuals will invest in their health if the level of utility they reach with the investment is higher than the expected utility they reach without; that is to say, they invest if:

$$\begin{aligned} \overline{U_i} - c_i &> \overline{U_i}(1 - Pr(\underline{h_i}|I_i)) + \underline{U_i} Pr(\underline{h_i}|I_i) \\ \Leftrightarrow Pr(\underline{h_i}|I_i) &> \frac{c_i}{\overline{U_i} - \underline{U_i}} \end{aligned} \tag{1}$$

In this decision rule, the limited program giving information on insurance benefits aims at reducing c_i. The combined program giving information on both insurance benefits and health status aims at both reducing c_i and improving $Pr(\underline{h_i}|I_i)$, from $Pr(\underline{h_i}|I_i^0)$ to $Pr(\underline{h_i}|I_i^0, I_i^1)$.

4.2 Predictions

An immediate prediction of the model is that when the costs of health investments c_i are reduced, people invest more. Another mechanism that this model describes and that is specific to the combined program is the revision of the subjective probability of being in poor health.

To further develop this revision process, we need to make distributional assumptions on people's perceptions of their health and on the signal they receive from the physician. We first assume that $h_i|I_i^0 \hookrightarrow \mathcal{N}(I_i^0, \sigma_0^2)$, which means that people have heterogeneous initial information on their health status, but homogeneous information precision (for simplicity). Then we assume

[20] Here we only consider the monetary costs of health investments, because we observe variations in monetary costs in the data. A more sophisticated model could decompose c_i into monetary and non-monetary components, which would include costs such as social stigma associated with certain health investments, or anxiety for example. Since we do not observe such costs and would subsequently be unable to test the predictions of such model, we restrict c_i to be purely monetary.

that $I_i^1|h_i \hookrightarrow \mathcal{N}(h_i, \sigma_1^2)$, which means that they consider the physician diagnosis to be unbiased but not fully accurate.

If we apply Bayes' rule, we obtain the following conditional distribution:

$$h_i|I_i^0, I_i^1 \hookrightarrow \mathcal{N}(\pi I_i^1 + (1-\pi)I_i^0, \widetilde{\sigma}^2)$$

with $\pi = \sigma_0^2/(\sigma_0^2 + \sigma_1^2)$ and $\widetilde{\sigma}^2 = \sigma_0^2\sigma_1^2/(\sigma_0^2 + \sigma_1^2)$ (see Appendix B)

It follows that the probability to perceive oneself in poor health for people who met the physician is:

$$Pr(\underline{h_i}|I_i^0, I_i^1) = \Phi(\pi(I_i^1 - I_i^0)/\widetilde{\sigma} + I_i^0/\widetilde{\sigma}) \tag{2}$$

whereas for people who did not meet the physician, it is simply:

$$Pr(\underline{h_i}|I_i^0) = \Phi(I_i^0/\sigma_0) \tag{3}$$

It can be shown that local linear approximations of $Pr(\underline{h_i}|I_i^0, I_i^1)$ and $Pr(\underline{h_i}|I_i^0)$ satisfy:

$$Pr(\underline{h_i}|I_i^0, I_i^1) - Pr(\underline{h_i}|I_i^0) \propto \alpha I_i^1 - \beta I_i^0 \tag{4}$$

where $\alpha > \beta > 0$ (see Appendix B)

Equation (4) suggests that the change in the perception of being in poor health increases as I_i^1, the signal provided by the physician, becomes higher (high meaning that the physician tells the young person she is in poor health), and the change is all the more important since the young person perceived herself as healthy (low I_i^0).

We define $Y_i^T(I_i^0, I_i^1)$ and $Y_i^C(I_i^0)$ as the decision to invest in one's health when treated and not treated, respectively: $Y_i^T(I_i^0, I_i^1) := Pr(\underline{h_i}|I_i^0, I_i^1) > c_i/(\overline{U}_i - \underline{U}_i)$ and $Y_i^C(I_i^0) := Pr(\underline{h_i}|I_i^0) > c_i/(\overline{U}_i - \underline{U}_i)$

The change in the perception of being in poor health (equation (4)) can easily be translated into

a change in the decision to invest in one's health, which is the ultimate outcome of interest (see Appendix B):

$$E(Y_i^T(I_i^0, I_i^1) - Y_i^C(I_i^0)|I_i^0, I_i^1) \propto \alpha I_i^1 - \beta I_i^0 \tag{5}$$

where $E(Y_i^T(I_i^0, I_i^1)|I_i^0, I_i^1)$ and $E(Y_i^C(I_i^0)|I_i^0, I_i^1)$ respectively denote the shares of young people who invest in their health in the combined program and in the control group, and the expectation refers to the distribution of $c_i/(\overline{U}_i - \underline{U}_i)$ in the population.

In the empirical part (section 7), we thus test whether the average difference in health investments between the combined treatment group and the control group is indeed higher as the gap between people's initial perceptions of their health and the signal they receive from the physician increases. To do so, we use the data collected by the physicians on their recommendations (the signal they give to young people) and the health perception measures we collected in the baseline survey.

5 Empirical framework

5.1 ITT Model

For all the outcomes presented in section 6, we estimate the following intent-to-treat model of being encouraged to meet the social worker or both the social worker and the physician:

$$Y_i = \alpha_0 + \alpha_1 Z_i^{SW+P} + \alpha_2 Z_i^{SW} + \epsilon_i \tag{6}$$

where Z_i^{SW+P} is a dummy variable indicating whether individual i was assigned to the social worker and physician meetings rather than the control group or alternative program, and Z_i^{SW} is a dummy variable indicating whether individual i was assigned to the social worker meeting rather than the control group or alternative program. Two robustness tables are provided in Appendix D: Table A.5 presents a nonlinear specification (logistic regression), as most of the outcome variables are binary, and Table A.6 presents the same linear probability model as above with control variables that have been selected based on Belloni et al. (2014). Results are very similar with these two alternative specifications.

All the results tables present both the usual unadjusted p-values (however obtained through a bootstrapping procedure) and the multiplicity-adjusted p-values that we obtain following the method described in List et al. (2019). As there are two interventions, when we consider a family involving K outcome variables, we subsequently have 2K null hypotheses in the family.

We present ITT estimates and not LATE estimates that would allow recovering the effect of meeting with the social worker or with the physician and the social worker rather than just the effect of being encouraged to meet them. We consider that the exclusion restriction that would be necessary is not fully plausible in this case: just mentioning health concerns could indeed directly affect people's perceptions and behaviors regarding health. Health is a sensitive topic, and young people might infer many things from being encouraged to meet the physician or the social worker, even if they decide not to participate. However, given that compliance rates are very high (see section 2.3), the differences identified by the ITT estimates are driven by the large proportion of people who are actually treated.

Finally, we assume that the outcomes of a given individual are not influenced by the treatment status of other individuals. There could be spillovers if people from the treatment groups share some information they got through the program with people from the control group. We cannot rule out the possibility of such spillovers since young people can interact when they come to the job center or outside of it if they know each other. Nonetheless, we believe interactions inside the job center are very limited since people do not come to the job center so often (in our sample, young people have on average 5.5 meetings with their counselor per year), and the probability that young people know each other outside the job center is quite low given that they come from various areas. Except for the youth job center of Toulouse, there is indeed only one job center for the whole city or even sometimes for a group of nearby cities (as is the case for Sénart and Clichy-sous-Bois) so they cover very vast territories. In case people in the control group have been affected, our estimates would most likely underestimate the true impacts.

5.2 Validity of the experimental design

The objective of the random assignment is to create three statistically identical groups so that any difference on average outcomes between the three groups can be attributed to the inter-

ventions we test. We check that this property holds at least for observed characteristics by comparing variables means in each treatment group to the control group means on a large set of baseline observed characteristics. Results are presented in Appendix A, Table A.1. We consider five families of baseline outcomes (socioeconomic characteristics, insurance, health, perceived vulnerabilities, and health investments). We implement List et al. (2019)'s procedure to account for multiple testing within each of these families. Only one significant difference appears among all the comparisons made: the proportion of youths in the social worker group who self-rate their physical health as bad is significantly larger compared to the control group mean.

Since not all the people of the baseline sample answered the endline survey that we use to estimate impacts, we check that attrition rates are comparable between the three groups and that the groups remain comparable in the endline sample. The overall response rate in the final survey is 71.9%. In the combined social worker and physician group, the response rate is 71.6%, compared to 72.7% in the limited social worker group and 71.7% in the control group. Pair-wise Chow tests suggest that these three rates are not statistically different. Nonetheless, it could be that although the rates are similar, respondents selected differently in each group. To get some insight on this possible source of bias, we redo the balancing checks for baseline characteristics on the sample of respondents only (see Appendix A, Table A.2). There is no significant difference, although it should be noted that smaller sample reduces our power to detect small differences.[21]

6 Results

6.1 Health insurance

The first question we want to answer is whether the interventions improved young people's coverage as well as their understanding of it. The intervention devoted to health insurance is the social worker program. However, we are also interested to know whether the program combining

[21] With 1 100 observations in the endline sample and imperfect compliance, we are able to detect effects as small as 0.237 standard deviation in the comparison between the physician and social worker versus the control group, 0.252 standard deviation in the comparison between social worker versus control group and 0.281 standard deviation in the comparison between the physician and social worker group versus the social worker group. These minimum detectable effects are still quite small.

information on insurance with information on health needs has a higher impact on the health insurance take-up compared to the program providing information on health insurance only, as it is likely to raise the perceived immediate utility of health insurance.

To do so, we consider the level and type of medical coverage young people report in the endline survey.[22] Results are presented in Table 4. In this table, each line corresponds to an outcome of interest. The columns give the results of the estimation of equation (6): the first column gives the mean in the control group, the second and third display the impact of the combined program and its p-value, and the fourth and fifth columns present the impact of the limited program and its p-value. The last one corresponds to the number of observation in the estimation sample. In columns "SW+P" and "SW", we show the estimated coefficients and their robust standard error below. The p-values columns give first the bootstrap per-comparison p-value (which is unadjusted for multiple testing) and then the multiplicity-adjusted p-value below in italics. As explained in section 5.1, we use the method suggested by List et al. (2019) to compute the adjusted p-values. The family of hypotheses corresponds to the list of variables presented in the table. Some variables have been removed from the family because they were collinear with another variable; in this case, the adjusted p-value is missing. As we test two interventions, the size of each family corresponds to two times the number of free variables.

To understand how medical coverage changes with the programs, we decompose it into different levels and types of coverage. The first decomposition indicates whether the coverage is complete or incomplete. Then we refine it, looking in the case of complete coverage at whether the complementary insurance is free and public (CMU-C) or private. In the case of incomplete coverage, we look at whether young people have only basic coverage or no coverage at all. Basic coverage is either the universal (basic) medical coverage provided to low-income people (CMU), or social security (SS), which is the basic medical coverage provided to taxpayers (see section 2.2 for detailed explanations). Finally, we look at whether people know their coverage or not (a possible answer to the question about medical coverage in the survey) and at their score in a quiz about health insurance functioning.

[22] It could be the case that the declared medical coverage in the endline survey does not correspond to the real medical coverage if young people do not know their coverage well. But we are interested in the coverage people *think* they have, since this is what they will take into consideration in their healthcare consumption decisions.

The table shows that overall, medical coverage is already high even without any intervention, leaving little room for improvement. Most young people in the control group have full coverage: only 4.3% have no health insurance at all, and 13.1% have partial coverage. However, 9.9% do not know their coverage. Looking now at columns "SW+P" and "SW", we indeed see that both interventions have little impacts. The main change is an increase in the share of young people having access to CMU-C. This is true for both interventions, with an increase of respectively 5.9 and 6.8 percentage points in the combined and limited interventions, with respect to an average of 19.3% in the control group. These impacts are significant at the 5% and 10% levels with standard p-values, but not with adjusted p-values.

The increase in the share of young people with CMU-C corresponds to two types of change. On the one hand, we observe a substitution between private and public complementary insurance. This is especially the case for the social worker intervention for which the increase of 6.8 percentage points is associated with a decrease of the share of private complementary insurance of 6.2 percentage points. In this case, there is thus no improvement on the level of medical coverage but only on its cost, CMU-C being free. On the other hand, we observe a decrease in the proportion of young people with no complementary insurance with the combined program. The share goes from 13.1% in the control group to 8.0% in the corresponding treatment group. In this case, the increase in the share of people with CMU-C corresponds to an improvement in coverage. Both cases suggest that some people were eligible for this free public insurance but did not ask for it. The difference in the impact of the two interventions could support the idea that providing additional information on health needs makes the utility of having complete coverage (basic *and* complementary coverage) more salient. Nonetheless, it should be noted that none of the impacts discussed above resist multiple testing adjustment.

The limited impact of the two interventions on the type of coverage is an interesting result. Dufour-Kippelen et al. (2006) emphasize that non-use of CMU-C mainly comes from a lack of knowledge about the degree of coverage of this insurance (which our intervention tackles) and from a reluctance to go through the associated administrative procedures. In our case, the social worker explains the procedures to follow, but it is the responsibility of young people to undertake the procedures. This could be an explanation for the weak impacts. Such interpretation is

consistent with Aizer (2007)'s findings that information and administrative costs are important barriers to enrollment in Medicaid. Similar results have been found on EITC benefits or financial aid programs for college enrollment (Bhargava and Manoli (2015), Bettinger et al. (2012)).

Table 4: Impact on Health insurance coverage (ITT estimates on full sample)

Variable	Mean C	SW+P	p-values	SW	p-values	N
Incomplete coverage	0.173	-0.038	0.155	-0.008	0.790	1093
		(0.027)	.	(0.028)	.	
-No insurance	0.043	0.013	0.413	0.004	0.799	1093
		(0.016)	*0.955*	(0.015)	*0.799*	
-(CMU or SS) and no comp	0.131	-0.051	0.025	-0.013	0.616	1093
		(0.023)	*0.180*	(0.025)	*0.946*	
Complete coverage	0.727	0.034	0.294	0.006	0.838	1093
		(0.032)	.	(0.033)	.	
-(CMU or SS) and CMU-C	0.193	0.059	0.053	0.068	0.023	1093
		(0.031)	*0.327*	(0.031)	*0.189*	
-(CMU or SS) and private	0.534	-0.025	0.508	-0.062	0.098	1093
		(0.037)	*0.936*	(0.037)	*0.485*	
Does not know coverage	0.099	0.004	0.855	0.002	0.919	1093
		(0.022)	.	(0.023)	.	
Health insurance quiz (0 to 5)	3.784	0.067	0.468	-0.032	0.727	1091
		(0.093)	*0.951*	(0.093)	*0.927*	

Continued on next page...

... table 4 continued

Variable	Mean C	SW+P	p-values	SW	p-values	N

Note: Mean C gives variable mean in the control group. SW+P gives the impact of physician plus social worker intervention versus control group (α_1 in the ITT model), and SW gives the impact of social worker only intervention versus control group (α_2 in the ITT model). Robust standard errors are in parentheses below. P-values gives per comparison (multiplicity-unadjusted) bootstrap p-values and multiplicity-adjusted p-values in italics below (except for colinear variables) corresponding to the left column estimated coefficients. For the p-values adjustment, we follow List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. N gives the number of observations in the regression.

6.2 Health investments

Second, we want to know whether and to what extent each program increased health investments. We focus on the dimensions physicians data revealed to be the most problematic (see descriptive analysis section 3.3.2, Table 3): dental health, mental health, sexual behavior, and to a lesser extent, weight issues. To measure investments in these dimensions, we use again young people's answers to a series of questions on their use of healthcare services in the year before the endline survey. More specifically, we focus on general practitioner, dentist, gynecologist and psychologist consultations, and add a last variable corresponding to a synthetic dummy variable indicating whether or not the young person met any other practitioner; this variable gathers 13 different practitioners.[23]. We also consider a second family of outcomes corresponding to declared health practices: having a regular doctor (which, although it does not respond to one of the specific health issues described above, suggests young people have a practitioner they can easily consult if they need healthcare), using contraception, and exercising or having a healthy diet.[24]

Results are presented in Table 5, which follows the same organization as the previous table

[23] It includes X-rays, pediatricians, dermatologists, cardiologists, rheumatologists, psychiatrists, ENT specialists, surgeons, anesthetists, phlebologists, pulmonologists, diabetologists, and ophthalmologists.

[24] The regular doctor was defined in the questionnaire as a doctor young people can easily consult if needed. It was made clear that the physician present at the job center as part of the experiment could not be considered as a regular doctor.

(the only difference is that here we have two sections, that correspond to the two families of hypotheses in the p-value adjustment procedure). The main result is that the combined program leads to a strong increase in the proportion of young people consulting a psychologist. While the proportion is 8.7% in the control group, with the encouragement to meet a social worker and a physician, it almost doubles, increasing by 7.8 percentage points. The impact is highly significant, with a per-comparison p-value of 0.1% and a multiplicity-adjusted p-value of 0.7%. This result is very robust (see Appendix D), and it is specific to the combined program. For the three other main recommendations made by the physicians (general practitioner, dentist, and gynecologist), we do not find the expected results: young people do not significantly increase their investments. It should nonetheless be noted that the proportion of young people consulting those three practitioners in the control group is already high so the room for improvement is more limited.

Concerning sexual behavior, we see a small increase in the proportion of young people using contraception with the combined program, which goes in the expected direction. This proportion goes from 75.3% in the control group to 80.5% for people who were encouraged to meet a physician and a social worker. Nonetheless, this impact is significant only at the 10% level before adjusting the p-value, and no longer significant at standard levels with the adjustment.[25] We do not find any impact on the probability of consulting a gynecologist, and we unfortunately do not observe whether young people got tested for HIV in the endline survey.

In the table, we focus on the binary variables indicating whether young people consulted the different recommended practitioners. If we look at the number of consultations (not presented here), we find no impact of the interventions.[26] The absence of impact of the combined program on the number of consultations with a psychologist, in particular, suggests that although the program is effective in initiating a consultation, it is unlikely to lead to sustained investments in psychological health. However, it should be noted that only 8.7% of the young people in

[25] Our measure of risky sexual behavior is quite limited; we only observe whether young people used contraception during their last intercourse and which contraception device they used. In the table, we present the aggregated variable. When we decompose it into the different possible contraception devices young people may have used, we do not find any impact on specific devices, in particular on the use of condoms, which is the only protection against STDs. It would have been difficult to ask more detailed questions due to the face-to-face nature of our survey, which is likely to induce desirability bias on sensitive issues such as sexual behavior.

[26] The only exception is the number of consultations with a gynecologist that decreases with both interventions, but further investigations show that this effect is driven by observations in the top 5% of the distribution of the number of consultations, the control group having a thicker tail.

the control group consult a psychologist, so our ability to detect an effect on the number of consultations is limited.

On other health practices, the probability of exercising or eating healthy is unchanged in spite of the few weight issues diagnosed by the physicians (see Table 3, section 3.3.2), but we do not observe whether physicians recommended that they exercise more or improve their diet. Finally, we find a small increase in the probability of having a regular doctor, which was expected as one objective of the combined program was to facilitate access to healthcare services and make young people more autonomous. However, the impact is no longer significantly different from zero when we adjust for multiple hypotheses testing.

Overall, mental health is the only dimension for which young people seem to underestimate their needs (section 3.3.3) and for which we find an increase in health investments. This is an important result as poor mental health has been found to be a barrier to finding a job (Prinz et al. (2018)). Investing in its improvement might thus be a first step towards reintegrating the labor market. Dental health investments are not so studied; thus our finding is hard to interpret. By contrast, the weak or null impacts we find on sexual behavior and sports or eating habits are not surprising. The literature indeed suggests that a meeting with a physician is quite unlikely to induce changes in these two types of behavior. Duflo et al. (2015) find for example that having teachers provide information on HIV in the classroom allows to increase students' knowledge on the disease but does not reduce childbearing rates or the risk of STD. On eating behavior, Prina and Royer (2014) find that providing parents with information on their child's BMI and on the risk of obesity has no impact on behaviors and weight.

Our theoretical model predicts that the impact should be larger as the gap between the signal sent by the physician and young people's initial perceptions of their health is itself larger. Estimations confirm these predictions in the case of mental health (see section 7).

Finally, the table shows that the second program providing only insurance information has no impact on health investments. This result is not so surprising since for the types of investments we consider – those that correspond to the main health issues diagnosed by the physicians – health insurance does not necessarily play a role. It does not enter in the decision to exercise or

eat healthy for example, and as mentioned earlier, young people can obtain free consultations with a physician for contraception and sexual issues at community health centers, where they can also obtain contraception devices for free. For the psychologist consultations, it should be noted that the youth job centers we worked with in the experiment have partnerships with health centers or hospitals that offer free consultations with a psychologist; so health insurance, in this case again, is not necessary.

Table 5: Impact on Health investments (ITT estimates on full sample)

Variable	Mean C	SW+P	p-values	SW	p-values	N
Healthcare consumption :						
Consulted GP	0.783	0.039	0.187	-0.021	0.497	1091
		(0.030)	*0.814*	(0.031)	*0.866*	
Consulted dentist	0.454	-0.040	0.268	0.016	0.652	1099
		(0.037)	*0.872*	(0.037)	*0.652*	
Consulted psychologist	0.087	0.078	0.001	0.017	0.444	1100
		(0.024)	*0.007*	(0.022)	*0.901*	
Consulted gynecologist	0.475	-0.051	0.314	-0.064	0.201	589
		(0.051)	*0.883*	(0.050)	*0.807*	
Consulted any other practitioner	0.437	-0.030	0.430	-0.023	0.526	1077
		(0.037)	*0.928*	(0.037)	*0.769*	
Health practices :						
Has regular doctor	0.810	0.050	0.065	0.021	0.459	1091
		(0.028)	*0.305*	(0.029)	*0.849*	
Contracep. used last intercourse	0.753	0.052	0.093	0.016	0.641	1068
		(0.031)	*0.364*	(0.032)	*0.871*	
Exercises and/or eats healthy	0.451	0.044	0.230	0.018	0.645	1078

Continued on next page...

... table 5 continued

Variable	Mean C	SW+P	p-values	SW	p-values	N
		(0.037)	*0.631*	(0.038)	*0.645*	

Note: Mean C gives variable mean in the control group. SW+P gives the impact of physician plus social worker intervention versus control group (α_1 in the ITT model), and SW gives the impact of social worker only intervention versus control group (α_2 in the ITT model). Robust standard errors are in parentheses below. P-values gives per comparison (multiplicity-unadjusted) bootstrap p-values and multiplicity-adjusted p-values in italics below (except for colinear variables) corresponding to the left column estimated coefficients. For the p-values adjustment, we follow List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. N gives the number of observations in the regression.

6.3 Perceptions of health system

In this section, we examine whether the beliefs young people hold about the health system changed with the interventions. A change in those beliefs could indeed be a mechanism through which the interventions affect people's investments. To do so, we consider three families of outcomes. First, we are interested in the perceived financial cost of healthcare, that we capture by asking people whether they think one needs money to get care. Second, we look at different types of transaction costs such as the difficulties young people may have to know which practitioner to consult, to find an available physician, or a good physician. The last family aims at capturing the level of trust people have in physicians and, more generally, in medicine, with three variables asking them whether they think that physicians know best what is good for a patient, that medicine has effective solutions for any health problem, and whether they find it difficult to confide in physicians.

The results suggest that trust and perceived transaction costs are not affected by the interventions. Perceived monetary costs, by contrast, significantly decrease with the combined program. Approximately 80.9% in the control group believe one needs money to get care. In the group

assigned to the combined intervention, this proportion shrinks to 72.2%. There is no impact with the social worker intervention, suggesting that the decrease we observe with the combined intervention comes from the information provided by the physician. As previously mentioned, the physicians told young people about possibilities to get free healthcare, in particular free consultations with a psychologist and free contraception, which can explain the strong decrease in the perceived costs of healthcare we observe in the combined program.

Table 6: Impact on Perceptions of health system (ITT estimates on full sample)

Variable	Mean C	SW+P	p-values	SW	p-values	N
Financial cost :						
You need money to get care	0.809	-0.087	0.006	-0.007	0.796	1092
		(0.031)	*0.011*	(0.030)	*0.796*	
Transaction costs :						
Difficult to know who to consult	0.691	-0.008	0.806	0.019	0.579	1086
		(0.035)	*0.806*	(0.034)	*0.913*	
Difficult to find available physician	0.531	-0.024	0.496	-0.024	0.520	1088
		(0.037)	*0.959*	(0.037)	*0.942*	
Difficult to find good physician	0.538	-0.018	0.608	0.029	0.414	1087
		(0.037)	*0.834*	(0.037)	*0.942*	
Trust :						
Physicians know what is good	0.801	0.001	0.973	-0.013	0.663	1095
		(0.030)	*0.973*	(0.030)	*0.872*	
Medicine has effective solutions	0.305	-0.046	0.176	-0.037	0.283	1091
		(0.033)	*0.589*	(0.034)	*0.723*	
Difficult to confide in physicians	0.392	0.031	0.376	0.058	0.123	1092
		(0.036)	*0.734*	(0.037)	*0.511*	

Continued on next page...

... table 6 continued

Variable	Mean C	SW+P	p-values	SW	p-values	N

Note: Mean C gives variable mean in the control group. SW+P gives the impact of physician plus social worker intervention versus control group (α_1 in the ITT model), and SW gives the impact of social worker only intervention versus control group (α_2 in the ITT model). Robust standard errors are in parentheses below. P-values gives per comparison (multiplicity-unadjusted) bootstrap p-values and multiplicity-adjusted p-values in italics below (except for colinear variables) corresponding to the left column estimated coefficients. For the p-values adjustment, we follow List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. N gives the number of observations in the regression.

6.4 Health

Third, we investigate whether the changes we observe in health investments, in particular, the higher propensity to consult a psychologist, are associated with improved health status. To do so, we use three five-level self-rated health questions from the endline survey on general, physical and psychological health that we dichotomized (variables take value 1 when young people rate their health as good or very good, and 0 if they rate it as average, bad or very bad), a dummy variable indicating whether young people reported having a chronic health problem, and another dummy variable capturing moderate to severe hopelessness according to the Beck scale (Beck et al. (1974)).

The expected effect of the interventions on the three self-rated health questions is ambiguous. On the one hand, as they measure the perceptions young people have of their health (and not their objective health status), and as the combined program aims at correcting these perceptions, if young people initially tended to overestimate their health status as we have seen earlier (see section 3.3.3), then the program could lead to a deterioration of self-rated health. On the other hand, as the proportion of people consulting a psychologist increased, if the consultations led to improved mental health, then young people could also have better self-rated health, in particular,

psychological health. If the two mechanisms are operating and compensates, the effect is expected to be null.

Results are presented in Table 7. They show that the combined program did not affect perceived health on any of the three dimensions, and did not affect the probability to report a chronic health problem nor the probability to be detected with critical Beck scores (scores corresponding to moderate to severe hopelessness are considered as indicative of suicide intentions). Although those null impacts may be disappointing, it should be reminded that we observe those variables in the endline survey simultaneously as we measure whether young people consulted a psychologist in the past year, so we do not know how much time there is between their first consultation with a psychologist and the report of their health status. At most, if young people consulted the psychologist right after they met the physician in the experiment, the time passed is one year. This duration might be insufficient to observe large improvements in mental health.

Looking at the limited program, we find no effect on health measures except a little deterioration of self-rated psychological health, which is robust to p-value adjustment and alternative specifications (see Appendix D). This effect was unexpected and is hard to interpret. Nonetheless, when we look at the impact of the interventions on health-related information seeking (not presented here), we see that the limited intervention led people to seek more information (while the combined intervention led them to seek less). We also know from the baseline survey that people in our sample tend to be very worried about their health (87.8% think one can be seriously ill without noticing it, 56.5% declare they get easily anxious when something is wrong, 19.7% declare they have been so ill in the past that they thought they would die). It could be the case that the limited intervention, by reminding the importance of medical coverage, raised concerns about health that translated into lower self-rated psychological health, as observed in Table 7. It should also be noted that this negative impact is not paralleled by a significant impact on our second measure of mental health, the Beck Hopelessness scale.

Table 7: Impact on Health status (ITT estimates on full sample)

Variable	Mean C	SW+P	p-values	SW	p-values	N
Good general self-rated health	0.839	-0.018	0.489	-0.034	0.235	1097
		(0.028)	*0.978*	(0.029)	*0.838*	
Good physical self-rated health	0.833	-0.041	0.148	-0.015	0.619	1098
		(0.029)	*0.684*	(0.028)	*0.991*	
Good psychological self-rated health	0.807	-0.011	0.728	-0.082	0.010	1096
		(0.029)	*0.921*	(0.031)	*0.079*	
Has a chronic health pb	0.184	0.013	0.676	-0.013	0.637	1095
		(0.029)	*0.966*	(0.029)	*0.981*	
Moderate to severe hopelessness	0.127	-0.003	0.907	-0.027	0.306	877
		(0.028)	*0.907*	(0.026)	*0.901*	

Note: Mean C gives variable mean in the control group. SW+P gives the impact of physician plus social worker intervention versus control group (α_1 in the ITT model), and SW gives the impact of social worker only intervention versus control group (α_2 in the ITT model). Robust standard errors are in parentheses below. P-values gives per comparison (multiplicity-unadjusted) bootstrap p-values and multiplicity-adjusted p-values in italics below (except for colinear variables) corresponding to the left column estimated coefficients. For the p-values adjustment, we follow List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. N gives the number of observations in the regression.

6.5 Labor market and social integration

Although the first objective of the interventions is to increase health investments, we are also interested to see if they allow improving labor market and social integration. As explained earlier, the motivation for implementing the experiment in youth job centers was that poor health could be a barrier for unemployed young people in finding a job. This is particularly true for poor

mental health, which, according to psychologists working in youth job centers, tends to take the form of low self-confidence and self-esteem, difficulties in making career decisions and building social ties.

Results are presented in Table 8, which is organized as previous tables. We consider two families of outcomes. First, we use direct measures of labor market integration: human capital investments variables such as participation in training programs or workshops, and a binary variable indicating whether young people worked in the previous year.[27] Second, we consider a family of outcomes capturing social integration such as whether young people are involved in associations, whether they have a high number of friends ("more than 10 friends" is the highest category in our survey), and whether they have someone they can count on if they need help.

The table shows that there is no significant impact of the interventions on labor market outcomes, considering the multiplicity-adjusted p-values. However, we can note that there is a positive impact of the combined program on the probability of participating in training courses. The proportion of young people who participated in at least one training course in the previous year is 42% in the control group, versus 49.8% with the combined program (significant at the 5% level according to the per comparison p-value). Other studies similarly investigated the link between access to healthcare and work and reached inconclusive findings as well. Kaestner et al. (2017) and Baicker et al. (2014), for example, find no employment improvement with Medicaid expansion in spite of insurance coverage improvements. In regard to social integration, the second part of the table shows that the combined program doubles the probability of having more than 10 friends, and tends to increase the probability of being involved in associations, although the estimate is no longer significantly different from zero when we adjust for multiple hypothesis testing.

[27] Youth job centers regularly organize "workshops" in the job center on specific topics such as writing a motivation letter, a CV, they may also consist in mock job interviews for example.

Table 8: Impact on Labor market and social integration (ITT estimates on full sample)

Variable	Mean C	SW+P	p-values	SW	p-values	N
Work and training :						
Did training course(s)	0.420	0.078	0.036	0.053	0.150	1092
		(0.037)	*0.184*	(0.037)	*0.527*	
Participated in workshop(s)	0.220	0.041	0.202	0.032	0.316	1082
		(0.032)	*0.555*	(0.032)	*0.655*	
Worked	0.751	0.026	0.400	0.006	0.871	1091
		(0.031)	*0.597*	(0.032)	*0.871*	
Social integration :						
Involved in associations	0.158	0.051	0.078	0.013	0.667	1095
		(0.029)	*0.311*	(0.028)	*0.952*	
Has more than 10 friends	0.040	0.042	0.017	0.007	0.640	1094
		(0.018)	*0.089*	(0.015)	*0.981*	
Has someone (s)he can count on	0.946	-0.002	0.932	-0.001	0.931	1098
		(0.017)	*0.932*	(0.017)	*0.995*	

Note: Mean C gives variable mean in the control group. SW+P gives the impact of physician plus social worker intervention versus control group (α_1 in the ITT model), and SW gives the impact of social worker only intervention versus control group (α_2 in the ITT model). Robust standard errors are in parentheses below. P-values gives per comparison (multiplicity-unadjusted) bootstrap p-values and multiplicity-adjusted p-values in italics below (except for colinear variables) corresponding to the left column estimated coefficients. For the p-values adjustment, we follow List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. N gives the number of observations in the regression.

7 Mechanisms

The model we develop in section 4 emphasizes the role of people's perceptions of their health in their decision to invest in it - perceptions which can be very biased, as shown in section 3.3.3. In the model, if the *perceived* probability of being in poor health is higher than a given threshold, then the individual invests (equation (1)). In this section, we assess whether the information provided by the physician leads young people to revise their subjective probability of being in poor health. Equation (5) (section 4.2) indeed shows that the impact of the combined intervention is predicted to be heterogeneous and to depend positively on the signal provided by the physician and negatively on individual's prior health perceptions.

To test those predictions, we would ideally like to estimate the following model:

$$\begin{aligned} Y_i = a \quad &+ \quad \alpha Z_i^{SW+P}.I_i^1 - \beta Z_i^{SW+P}.I_i^0 + bZ_i^{SW+P} \\ &+ \quad \alpha' Z_i^{SW}.I_i^1 - \beta' Z_i^{SW}.I_i^0 + b' Z_i^{SW} + cI_i^1 + dI_i^0 + \epsilon_i \end{aligned} \tag{7}$$

where variables Z^{SW+P} and Z^{SW} are the combined and limited program assignment variables, and I_i^1 and I_i^0 are the signal provided by the physician and the individual's prior health perceptions, respectively. Notice that we introduce $-\beta$ and $-\beta'$ (instead of the standard $+\beta$ and $+\beta'$) for consistency with the theoretical model. In equation (7), we are interested in coefficients α and β, and we expect, in line with the predictions of the model, that $\alpha > \beta > 0$. In contrast, we do not expect similar pattern for α' and β' since the limited program does not provide information I_i^1.

However, we cannot directly estimate equation (7). The first problem is that by construction, we observe physicians' signal, I_i^1, only for young people who met them. Second, we face data limitations for the measure of young people's perceptions, I_i^0. For mental and dental health, we ask young people in the baseline survey to give a rating on a four or five-level Likert scale from "very bad" to "very good" (that we dichotomize, with value 1 corresponding to "good" and "very good").[28] Since those perception variables are measured up to a multiplicative or an additive

[28] For sexual health, which is another important concern in our population according to physicians' diagnoses, we do not have any measure of young people's perceptions.

constant, we cannot assess the complete prediction $\alpha > \beta > 0$, but only $\alpha > 0$ and $\beta > 0$. This implies that we can only examine the role of perceptions biases and not perceptions precision (which would require to test $\alpha = \beta$ versus $\alpha > \beta$).

To address the first problem, we use a machine learning algorithm to estimate a prediction of the signal provided by the physician using the sample of young people who were examined by the physician and our large set of baseline data. We extrapolate this prediction to the whole sample and take it as a measure of the counterfactual signal young people would receive if they were examined by a physician (signal that is revealed only for treated people, and remains latent for others). More precisely, the left-hand side quantity that we want to predict is the probability of being recommended to consult a psychologist in the case of mental health and the probability of being recommended to consult a dentist in the case of dental health.[29] Our prediction is based on the LASSO algorithm. The algorithm selects the right-hand side variables among the set of baseline variables, which contains in particular various measurements of the family environment, young people's social ties, possible traumatic events that happened during childhood or adolescence, that are likely to be strong determinants of health status. We use cross-validation and leave-one-out techniques to avoid overfitting. More detailed explanations of the procedure we implement are provided in Appendix C. The second problem cannot be addressed, so we only test the hypotheses that $\alpha > 0$ and $\beta > 0$ for mental and dental health.

Instead of estimating equation (7), we thus estimate the following equation:

$$\begin{aligned} Y_i = a \quad &+ \quad \alpha Z_i^{SW+P}.\widehat{I}_i^1 - \beta Z_i^{SW+P}.I_i^{0^m} + bZ_i^{SW+P} \\ &+ \quad \alpha' Z_i^{SW}.\widehat{I}_i^1 - \beta' Z_i^{SW}.I_i^{0^m} + b' Z_i^{SW} + c\widehat{I}_i^1 + dI_i^{0^m} + \epsilon_i \end{aligned} \tag{8}$$

whose difference with equation (7) is that we take the prediction $\widehat{I}_i^1$ instead of the (partially) observed measure I_i^1, and where $I_i^{0^m}$ stands for our measure of health perceptions.

The results are presented in Table 9, where the first four columns respectively display estimates

[29] Although we also observe the recommendation to consult a dietitian and weight issues were among the important concerns reported by the physicians (see Table 3), we could not accurately predict the probability of being recommended to consult a dietitian as only 3.5% of the treated young people have been directed to a dietitian. So we excluded it from the heterogeneity analysis.

of coefficients α, β, α' and β', and the two lines give the two investment decisions (Y_i) we are interested in: consulting a psychologist and consulting a dentist. We see that for the probability of consulting a psychologist, the estimated coefficients are consistent with the theoretical predictions: $\hat{\alpha} > 0$ and $\hat{\beta} > 0$.[30] Those results are in line with the interpretation that young people revise their perceptions of their health needs in a Bayesian fashion after receiving the recommendation from the physician. Unsurprisingly, the heterogeneous impacts of the social worker program are not statistically different from zero. Concerning the probability of consulting a dentist, none of the estimated coefficients is significant, suggesting that the perception revision mechanism described in our model does not hold in this case. This could explain why we do not find any impact in the general analysis (Table 5, section 6). However, the absence of significant heterogeneity might also be due to the poor performance of the algorithm procedures to identify meaningful covariates to predict dental health issues (see Appendix C).

Table 9: Heterogeneity analysis

Variable	α	$-\beta$	α'	$-\beta'$	N
Consulted psychologist	0.517	-0.322	0.058	-0.108	1097
	(0.193)	(0.145)	(0.189)	(0.135)	
	0.007	*0.027*	*0.760*	*0.425*	
Consulted dentist	-0.847	0.715	0.769	-0.166	1065
	(3.207)	(1.964)	(3.341)	(2.055)	
	0.792	*0.716*	*0.818*	*0.936*	

Continued on next page...

[30] We also get that $\hat{\alpha} > \hat{\beta}$, consistent with the model prediction, but as explained before, we cannot fully trust this result due to the measurement issues concerning health perceptions I^{0^m}.

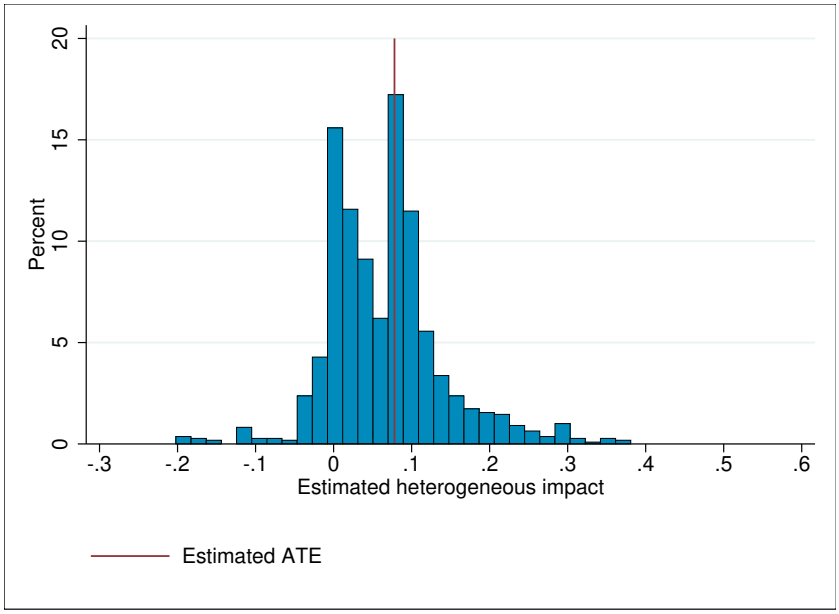
... table 9 continued

Variable	α	β	α'	β'	N

Note: α is the coefficient of the physician plus social worker intervention interacted with the (predicted) signal provided by the physician. β is the coefficient of the physician plus social worker intervention interacted with the baseline health perceptions young people have. α' and β' are the same coefficients for the social worker only intervention. Robust standard errors are provided below the coefficients in parentheses, and (multiplicity-unadjusted) per comparison p-values are provided in italics. N gives the number of observations in the regression.

Those findings are important in two respects. They confirm that our main empirical result is consistent with a model in which young people *learn* from the personalized information they receive in the intervention. In the case of mental health, where symptoms of deterioration might be hard to self-diagnose, the results are indeed in line with the prediction that young people revised their beliefs about their health needs. Another implication of our analysis is that the impact of providing such personalized information is highly heterogeneous. It depends on each individual's true health status and how strongly his or her perceptions are misaligned with it. Figure 3 below displays the distribution of the estimated impact, $\widehat{\alpha}.\widehat{I}_i^1 - \widehat{\beta}.I_i^{0^m} + \widehat{b}$, together with the estimated average treatment effect, that ignores the impact's heterogeneity. It shows that the estimated individual impact has a large variability, ranging from -0.2 to more than +0.4. The average treatment effect, $\widehat{\alpha_1}$ (equation (6)), is represented by the red vertical line, and corresponds to approximately 0.078 (see Table 5). The distribution reveals that while the impact of the intervention on the probability to consult a psychologist is limited for a large share of the population, it is substantial for others: the 75th percentile of the distribution is approximately 0.097, and the 90th percentile approximately 0.154. On the other hand, it is null or even negative for a 16.4% of our sample.

Figure 3: Impact distribution



In the same spirit as Chernozhukov et al. (2018), to learn more about the sources of heterogeneity in the effect of the intervention on mental health investments, we compare the average socioeconomic characteristics of young people in the first and last quartiles of the impact distribution. Table 10 shows that the variables associated with a higher impact tend to correspond to more precarious family and financial situations. The 25% of young people who benefits the most from the intervention are significantly more likely to be single (for 76.3% of them, versus 54.4% of the young people who are among the 25% who benefits the least from the intervention), to have divorced parents, to have a very low level of education, to have been taken charge of by the Department of Children and Family Services (DCFS), and to have experienced severe financial problems by the age of 18. Unsurprisingly, variables reflecting situations of extreme emotional and psychological distress are also more strongly correlated with the probability of being among the quartile of people for whom the intervention had the strongest impact on the probability to consult a psychologist. The proportion of young people who tried to commit suicide is indeed more than six times higher in the last quartile of the impact distribution than in the first, and the proportion of young people who ran away while they were living with their parents is also significantly higher in the last quartile.

As clearly shown in figure 3, although the impact can be large for some youths, it is also very small for many of them. If the objective is to address mental health issues among unemployed youths, it seems to be inefficient to have them all meet with the physician. The previous comparisons reveal that there are some factors that are good predictors of which young people are the most likely to highly benefit from the intervention and that the job centers could use as a screening tool to better target the intervention (e.g., asking young people about their marital status, their level of education, DCFS placement, the financial situation of their family). A simple policy implication could be to systematically direct those targeted people to the physician to be examined and, if needed, referred to a psychologist.

Table 10: Average characteristics of the first and last quartiles

Variable	First quartile	Last quartile	Difference	N
Single	0.544	0.763	0.219	802
	(0.499)	(0.426)	*0.000*	
Divorced parents	0.399	0.462	0.063	803
	(0.490)	(0.499)	*0.071*	
Less than high school diploma	0.415	0.513	0.098	802
	(0.493)	(0.500)	*0.005*	
Has been placed to DCFS	0.059	0.127	0.068	799
	(0.237)	(0.333)	*0.001*	
Family had serious housing pbs	0.090	0.116	0.026	803
	(0.286)	(0.320)	*0.226*	
Family had severe financial pbs	0.267	0.347	0.079	801
	(0.443)	(0.477)	*0.015*	
Tried to commit suicide	0.029	0.184	0.156	801
	(0.167)	(0.388)	*0.000*	
Ran away while living w/ parents	0.147	0.218	0.072	803

Continued on next page...

... table 10 continued

Variable	First quartile	Last quartile	Difference	N
	(0.354)	(0.414)	*0.008*	

Note: The first and second columns give the mean of the variable for the first and last quartiles, with the standard deviation in parentheses below. The third column gives the difference between the two, with the p-value in italics below. N gives the number of observations.

8 Conclusion

In this paper, we approach the problem of youth unemployment with a new perspective, focusing on health capital as a possible determinant of employability. We target unemployed youths coming from disadvantaged backgrounds in France and investigate the factors that induce or deter health investments.

Through the collection of unique medical data, we first find that these young people indeed face a number of health issues that are left untreated. The most prevalent issues diagnosed by the physicians are mental and dental health issues, as well as risky sexual behaviors. We consider interventions designed to address different possible barriers to health investments. A specificity of the interventions is that the information on health needs is provided by a physician after examining each young person; and the information on health insurance is provided by a social worker after investigating their eligibility and insurance options. In both cases, it is a highly personalized information.

Relying on a randomized control trial, we find that informing unemployed youths about health insurance slightly increases the take-up of free public health insurance, but is not sufficient to increase health investments. This suggests that enhancing the visibility of public health insurance and simplifying the procedures could improve young people's coverage, but more is needed to change their health behaviors. Results from the second intervention show that additionally

providing young people with information on their health needs and the corresponding services they can use additionally doubles their probability of consulting a psychologist, decreases the perceived cost of healthcare, and to some extent, induces safer sexual behaviors. It nonetheless does not affect the probability of consulting a dentist, in spite of the physicians' recommendations. We also find suggestive evidence that the intervention leads young people to raise their investments in job training, and might improve social integration.

To further document the mechanism by which the provided information prompts young people to raise their mental health investments, we specify a Bayesian model of perceptions revision and derive predictions that we empirically assess. Using a LASSO regression and our rich survey and medical data, we predict the probability of being directed to a psychologist for each young person in our sample. We then use our self-rated health survey data to show that the impact of the intervention on the probability of consulting a psychologist increases with the gap between physicians' diagnoses and young people's perceptions of their psychological health. This result is consistent with Bayesian learning. On the methodological side, this result suggests that it is important when studying how effective information programs are to go beyond average impacts and analyze how the impacts vary with people's prior level of information.

One conclusion of our heterogeneity analysis is that impacts of information policies on health behaviors can be large. The simple Bayesian model we develop shows that the accuracy of the information provided is crucial. Our experiment's originality is to provide information based on individual assessments made by experts. The provided information is thus likely to be very accurate. Our findings suggest that targeted to people who need it the most, such information policies can substantially affect health investments, while being relatively inexpensive and easy to implement.

References

Aizer, A. (2007). Public health insurance, program take-up, and child health. *The Review of Economics and Statistics 89*(3), 400–415.

Aizer, A. and L. Stroud (2010). Education, knowledge and the evolution of disparities in health. Technical report, National Bureau of Economic Research.

Armour, P. (2018). The role of information in disability insurance application: An analysis of the social security statement phase-in. *American Economic Journal: Economic Policy 10*(3), 1–41.

Athey, S. and G. Imbens (2016). Recursive partitioning for heterogeneous causal effects. *Proceedings of the National Academy of Sciences 113*(27), 7353–7360.

Athey, S. and G. W. Imbens (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives 31*(2), 3–32.

Athey, S. and S. Wager (2018). Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association 113*(523), 1228–1242.

Attanasio, O. P. and K. M. Kaufmann (2014). Education choices and returns to schooling: Mothers' and youths' subjective expectations and their role by gender. *Journal of Development Economics 109*, 203–216.

Baicker, K., A. Finkelstein, J. Song, and S. Taubman (2014). The impact of Medicaid on labor market activity and program participation: Evidence from the Oregon health insurance experiment. *American Economic Review: Papers & Proceedings 104*(5), 322–328.

Baicker, K., S. L. Taubman, H. L. Allen, M. Bernstein, J. H. Gruber, J. P. Newhouse, E. C. Schneider, B. J. Wright, A. M. Zaslavsky, and A. N. Finkelstein (2013). The Oregon experiment - Effects of Medicaid on clinical outcomes. *New England Journal of Medicine 368*(18), 1713–1722.

Bazin, F., I. Parizot, and P. Chauvin (2005). Original approach to the individual characteristics associated with forgone healthcare. *The European Journal of Public Health 15*(4), 361–367.

Beck, A. T., A. Weissman, D. Lester, and L. Trexler (1974). The measurement of pessimism: the Hopelessness scale. *Journal of Consulting and Clinical Psychology 42*(6), 861.

Belloni, A., D. Chen, V. Chernozhukov, and C. Hansen (2012). Sparse models and methods for optimal instruments with an application to eminent domain. *Econometrica 80*(6), 2369–2429.

Belloni, A., V. Chernozhukov, I. Fernández-Val, and C. Hansen (2017). Program evaluation and causal inference with high-dimensional data. *Econometrica 85*(1), 233–298.

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.

Bettinger, E. P., B. T. Long, P. Oreopoulos, and L. Sanbonmatsu (2012). The role of application assistance and information in college decisions: Results from the H&R block FAFSA experiment. *The Quarterly Journal of Economics 127*(3), 1205–1242.

Bhargava, S., G. Loewenstein, and J. Sydnor (2017). Choose to lose: Health plan choices from a menu with dominated option. *The Quarterly Journal of Economics 132*(3), 1319–1372.

Bhargava, S. and D. Manoli (2015). Psychological frictions and the incomplete take-up of social benefits: Evidence from an IRS field experiment. *The American Economic Review 105*(11), 3489–3529.

Buchmueller, T. C., A. Couffinhal, M. Grignon, and M. Perronnin (2004). Access to physician services: Does supplemental insurance matter? evidence from France. *Health Economics 13*(7), 669–687.

Cai, L. (2010). The relationship between health and labour force participation: Evidence from a panel data simultaneous equation model. *Labour Economics 17*(1), 77–90.

Cai, L. and G. Kalb (2006). Health status and labour force participation: Evidence from Australia. *Health Economics 15*(3), 241–261.

Chernozhukov, V., M. Demirer, E. Duflo, and I. Fernandez-Val (2018). Generic machine learning inference on heterogenous treatment effects in randomized experiments. Technical report, National Bureau of Economic Research.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Routledge.

Cour des Comptes (2015). Communication à la commission des finances du Sénat, article 58, alinéa 2 de la Loi organique du 1er août 2001. Technical report.

Currie, C. et al. (2012). *Social determinants of health and well-being among young people*. World Health Organization Regional Office for Europe Copenhagen, Denmark.

Currie, J. (2006). The take-up of social benefits. In D. C. A.J. Auerbach and J. Quigley (Eds.), *Public Policy and the Income Distribution*, pp. 80–148. New York: Russell Sage.

Cutler, D. M. and A. Lleras-Muney (2010). Understanding differences in health behaviors by education. *Journal of Health Economics 29*(1), 1–28.

Dahl, G. B., A. R. Kostøl, and M. Mogstad (2014). Family welfare cultures. *The Quarterly Journal of Economics 129*(4), 1711–1752.

Daponte, B. O., S. Sanders, and L. Taylor (1999). Why do low-income households not use food stamps? Evidence from an experiment. *Journal of Human resources 34*(3), 612–628.

de Saint Pol, T. (2009). Les inégalités sociales face à la santé en France. *Regards croisés sur l'économie* (1), 65–70.

Delavande, A. and B. Zafar (2019). University choice: the role of expected earnings, non-pecuniary outcomes, and financial constraints. *Journal of Political Economy* (forthcoming).

Deshpande, M. and Y. Li (2017). Who is screened out? application costs and the targeting of disability programs. Technical report, National Bureau of Economic Research.

Devaux, M. (2015). Income-related inequalities and inequities in health care services utilisation in 18 selected OECD countries. *The European Journal of Health Economics 16*(1), 21–33.

Duflo, E., P. Dupas, and M. Kremer (2015). Education, HIV, and early fertility: Experimental evidence from Kenya. *The American Economic Review 105*(9), 2757–2797.

Dufour-Kippelen, S., A. Legal, and J. Wittwer (2006). Comprendre les causes du non recours à la CMUC. *Economie sociale et droit*.

Dupas, P. and E. Miguel (2017). Impacts and determinants of health levels in low-income countries. In *Handbook of economic field experiments*, Volume 2, pp. 3–93. Elsevier.

Eil, D. and J. M. Rao (2011). The good news-bad news effect: asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics 3*(2), 114–38.

Etilé, F. and C. Milcent (2006). Income-related reporting heterogeneity in self-assessed health: evidence from France. *Health economics 15*(9), 965–981.

Finkelstein, A. and M. J. Notowidigdo (2019). Take-up and targeting: Experimental evidence from SNAP. *The Quarterly Journal of Economics* (forthcoming).

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker, et al. (2012). The Oregon health insurance experiment: Evidence from the first year. *The Quaterly Journal of Economics 127*(3), 1057–1106.

Frijters, P., D. W. Johnston, and M. A. Shields (2014). The effect of mental health on employment: evidence from australian panel data. *Health economics 23*(9), 1058–1071.

García-Gómez, P., A. M. Jones, and N. Rice (2010). Health effects on labour market exits and entries. *Labour Economics 17*(1), 62–76.

Glied, S. and M. Neidell (2010). The economic value of teeth. *Journal of Human Resources 45*(2), 468–496.

Gong, E. (2015). HIV testing and risky sexual behaviour. *The Economic Journal 125*(582), 32–60.

Imai, K., M. Ratkovic, et al. (2013). Estimating treatment effect heterogeneity in randomized program evaluation. *The Annals of Applied Statistics 7*(1), 443–470.

Jensen, R. (2010). The (perceived) returns to education and the demand for schooling. *The Quarterly Journal of Economics 125*(2), 515–548.

Kaestner, R., B. Garrett, J. Chen, A. Gangopadhyaya, and C. Fleming (2017). Effects of ACA Medicaid expansions on health insurance coverage and labor supply. *Journal of Policy Analysis and Management 36*(3), 608–642.

Keane, M. P. and K. I. Wolpin (2010). The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women. *International Economic Review 51*(3), 851–892.

Kling, J. R., S. Mullainathan, E. Shafir, L. Vermeulen, and M. Wrobel (2012). Comparison friction: Experimental evidence from Medicare drug plans. *The Quarterly Journal of Economics 127*(1), 199–235.

Kolstad, J. T. and A. E. Kowalski (2012). The impact of health care reform on hospital and preventive care: evidence from massachusetts. *Journal of Public Economics 96*(11-12), 909–929.

Labbe, E., J.-J. Moulin, C. Sass, C. Chatain, R. Guéguen, and L. Gerbaud (2007). Etat de santé, comportements et environnement social de 105 901 jeunes en insertion professionnelle. *Pratiques et organisation des soins 38*(1), 43–53.

Leray, E., A. Camara, D. Drapier, F. Riou, N. Bougeant, A. Pelissolo, K. Lloyd, V. Bellamy, J. Roelandt, and B. Millet (2011). Prevalence, characteristics and comorbidities of anxiety disorders in France: results from the "Mental Health in General Population" survey (MHGP). *European Psychiatry 26*(6), 339–345.

List, J. A., A. M. Shaikh, and Y. Xu (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics* (forthcoming).

Lochner, L. (2011). Nonproduction benefits of education: Crime, health, and good citizenship. In *Handbook of the Economics of Education*, Volume 4, pp. 183–282. Elsevier.

Léon, C., C. Chan Chee, and E. du Roscoät (2018). Depression in France among the 18-75 year-olds: Results from the 2017 Health Barometer. *BEH 637*.

Mackenbach, J. P., C. Looman, and J. Van der Meer (1996). Differences in the misreporting of chronic conditions, by level of education: the effect on inequalities in prevalence rates. *American Journal of Public Health 86*(5), 706–711.

Mathers, C. and R. Douglas (1998). Measuring progress in population health and well-being. *Measuring progress: Is life getting better*, 125–155.

Mullainathan, S. and J. Spiess (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Murray, C. J. and L. C. Chen (1992). Understanding morbidity change. *The Population and Development Review*, 481–503.

Oster, E. (2018). Diabetes and diet: purchasing behavior change in response to health information. *American Economic Journal: Applied Economics 10*(4), 308–48.

Paula, Á. D., G. Shapira, and P. E. Todd (2014). How beliefs about HIV status affect risky behaviors: Evidence from Malawi. *Journal of Applied Econometrics 29*(6), 944–964.

Prina, S. and H. Royer (2014). The importance of parental knowledge: Evidence from weight report cards in Mexico. *Journal of Health Economics 37*, 232 – 247.

Prinz, D., M. Chernew, D. Cutler, and A. Frakt (2018). Health and economic activity over the lifecycle: Literature review. Technical report, National Bureau of Economic Research.

Remler, D. K. and S. A. Glied (2003). What other programs can teach us: Increasing participation in health insurance programs. *American Journal of Public Health 93*(1), 67–74.

Schanzenbach, D. (2009). *Experimental estimates of the barriers to Food Stamp enrollment.* Institute for Research on Poverty, University of Wisconsin-Madison.

Sen, A. (2002). Perception versus observation: self reported morbidity has severe limitations and can be extremely misleading. *British Medical Journal 324*(7342), 860.

Sheehan, D., Y. Lecrubier, K. H. Sheehan, J. Janavs, E. Weiller, A. Keskiner, J. Schinka, E. Knapp, M. Sheehan, and G. Dunbar (1997). The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability. *European Psychiatry 12*(5), 232–241.

Shmueli, A. (2003). Socio-economic and demographic variation in health and in its measures: the issue of reporting heterogeneity. *Social science & Medicine 57*(1), 125–134.

Thornton, R. L. (2008). The demand for, and impact of, learning HIV status. *The American Economic Review 98*(5), 1829.

Van Doorslaer, E., C. Masseria, and X. Koolman (2006). Inequalities in access to medical care by income in developed countries. *Canadian Medical Association Journal 174*(2), 177–183.

Ware Jr, J. E. (1976). Scales for measuring general health perceptions. *Health Services Research 11*(4), 396.

White, R. G. (2015). Health and illness: mental representations in different cultures. *In: Wright, J. D. (ed.) International Encyclopedia of the Social and Behavioral Sciences.*, 567–570.

Wiswall, M. and B. Zafar (2016). Human capital investments and expectations about career and family. Technical report, National Bureau of Economic Research.

Appendix

A Balancing checks

Table A.1: Balancing checks for the baseline sample

Variable	Mean C	SW+P-C	SW-C	N
Socioeconomic characteristics :				
Age (in years)	20.984	0.094	0.110	1523
		0.963	*0.986*	
Female	0.535	-0.029	-0.002	1527
		0.987	*0.931*	
Less than high school diploma	0.496	-0.055	-0.033	1522
		0.748	*0.978*	
Has difficulties in French	0.356	-0.036	-0.043	1525
		0.968	*0.898*	
Has been placed to DCFS	0.089	-0.014	-0.013	1518
		0.975	*0.975*	
Family had serious housing pbs	0.103	-0.016	-0.011	1525
		0.987	*0.963*	
Family had severe financial pbs	0.245	0.003	0.035	1522
		0.999	*0.972*	
Difficulties in covering expenses	0.292	0.024	0.003	1509
		0.981	*0.995*	
Already had a job	1.508	0.060	0.061	1525
		0.964	*0.966*	

Continued on next page...

... table A.1 continued

Variable	Mean C	SW+P-C	SW-C	N
Insurance :				
No insurance	0.071	0.022	0.043	1518
		0.708	*0.136*	
(CMU or SS) and no comp	0.209	-0.038	-0.059	1518
		0.548	*0.129*	
(CMU or SS) and CMU-C	0.163	-0.003	0.011	1518
		0.905	*0.979*	
(CMU or SS) and private	0.407	0.014	0.017	1518
		0.958	*0.984*	
Does not know coverage	0.150	0.005	-0.012	1518
		!0.814	*!0.587*	
You need money to get care	0.799	-0.046	0.005	1523
		0.420	*0.977*	
Health :				
Bad general self-rated health	0.184	0.007	0.024	1526
		0.950	*0.882*	
Bad physical self-rated health	0.168	0.038	0.065	1523
		0.636	*0.090*	
Bad psychological self-rated health	0.275	0.037	0.039	1521
		0.733	*0.706*	
Has a chronic health pb	0.194	-0.013	-0.000	1526
		0.980	*0.990*	
Moderate to severe hopelessness	0.151	-0.012	0.016	1150
		0.950	*0.973*	

Continued on next page...

... table A.1 continued

Variable	Mean C	SW+P-C	SW-C	N
Perceived vulnerabilities :				
Get sick more easily than other people	0.221	-0.023	-0.018	1523
		0.928	*0.955*	
Body withstands illness very well	0.807	-0.010	-0.038	1513
		0.970	*0.847*	
Usually catch viruses	0.383	-0.037	-0.044	1521
		0.927	*0.835*	
Easily anxious when something wrong	0.545	0.031	0.029	1524
		0.935	*0.931*	
Expect better health	0.509	0.008	0.056	1485
		0.957	*0.660*	
Has already been seriously ill	0.249	0.002	0.017	1520
		0.923	*0.953*	
Has been so ill thought die	0.178	0.029	0.025	1525
		0.910	*0.953*	
Health investments :				
Consulted GP	0.791	-0.035	-0.019	1512
		0.853	*0.946*	
Consulted dentist	0.470	-0.052	-0.042	1526
		0.726	*0.831*	
Consulted psychologist	0.087	0.007	0.003	1528
		0.959	*0.867*	
Consulted gynecologist	0.441	-0.039	-0.062	793
		0.962	*0.792*	
Consulted another practitioner	0.374	-0.024	-0.047	1509

Continued on next page...

... table A.1 continued

Variable	Mean C	SW+P-C	SW-C	N
		0.960	*0.787*	
Has regular doctor	0.737	-0.051	-0.037	1526
		0.596	*0.796*	
Exercises and/or eats healthy	0.391	0.017	-0.006	1485
		0.971	*0.971*	

Note: SW+P-C gives the difference between the combined social worker and physician group and the control group, SW-C gives the difference between the social worker program and the control group. P-values in italics below the differences have been adjusted for multiple testing following List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. For variables that have been removed from the family because of colinearity (signaled by !), the per comparison bootstrap p-values is provided. N gives the number of observations in the regression.

Table A.2: Balancing checks for the endline sample

Variable	Mean C	SW+P-C	SW-C	N
Socioeconomic characteristics :				
Age (in years)	21.062	0.107	0.076	1095
		0.999	*1.000*	
Female	0.561	-0.055	-0.000	1099
		0.867	*0.998*	
Less than high school diploma	0.452	-0.042	-0.020	1094
		0.982	*0.999*	

Continued on next page...

... table A.2 continued

Variable	Mean C	SW+P-C	SW-C	N
Has difficulties in French	0.336	-0.017	-0.006	1097
		0.999	*1.000*	
Has been placed to DCFS	0.074	-0.013	0.001	1092
		0.999	*1.000*	
Family had serious housing pbs	0.093	-0.001	-0.013	1099
		1.000	*0.999*	
Family had severe financial pbs	0.256	-0.004	0.029	1099
		1.000	*0.998*	
Difficulties in covering expenses	0.305	0.006	-0.029	1086
		1.000	*0.998*	
Already had a job	1.494	0.046	0.055	1099
		0.999	*0.997*	
Insurance :				
No insurance	0.065	0.022	0.047	1094
		0.867	*0.214*	
(CMU or SS) and no comp	0.184	-0.022	-0.025	1094
		0.935	*0.925*	
(CMU or SS) and CMU-C	0.153	0.006	0.012	1094
		0.965	*0.954*	
(CMU or SS) and private	0.431	0.028	-0.008	1094
		0.918	*0.845*	
Does not know coverage	0.167	-0.035	-0.027	1094
		!0.197	*!0.316*	
You need money to get care	0.807	-0.055	-0.014	1096
		0.461	*0.980*	

Continued on next page...

... table A.2 continued

Variable	Mean C	SW+P-C	SW-C	N
Health :				
Bad general self-rated health	0.197	-0.002	0.012	1099
		0.943	*0.969*	
Bad physical self-rated health	0.164	0.047	0.061	1098
		0.548	*0.304*	
Bad psychological self-rated health	0.274	0.045	0.050	1097
		0.718	*0.670*	
Has a chronic health pb	0.223	-0.015	-0.011	1099
		0.973	*0.922*	
Moderate to severe hopelessness	0.158	-0.031	0.018	831
		0.865	*0.978*	
Perceived vulnerabilities :				
Get sick more easily than other people	0.229	-0.018	-0.033	1096
		0.983	*0.935*	
Body withstands illness very well	0.793	0.019	-0.014	1089
		0.986	*0.985*	
Usually catch viruses	0.392	-0.033	-0.051	1095
		0.964	*0.775*	
Easily anxious when something wrong	0.514	0.083	0.070	1097
		0.267	*0.517*	
Expect better health	0.504	-0.002	0.073	1068
		0.971	*0.505*	
Has already been seriously ill	0.251	0.012	0.012	1097
		0.916	*0.972*	

Continued on next page...

... table A.2 continued

Variable	Mean C	SW+P-C	SW-C	N
Has been so ill thought die	0.186	0.037	0.025	1098
		0.871	*0.972*	
Health investments :				
Consulted GP	0.810	-0.040	-0.022	1092
		0.876	*0.945*	
Consulted dentist	0.479	-0.038	-0.050	1100
		0.939	*0.886*	
Consulted psychologist	0.076	0.021	0.023	1100
		0.921	*0.947*	
Consulted gynecologist	0.437	-0.021	-0.045	589
		0.975	*0.921*	
Consulted another practitioner	0.388	-0.009	-0.039	1088
		0.961	*0.934*	
Has regular doctor	0.761	-0.047	-0.052	1099
		0.826	*0.776*	
Exercises and/or eats healthy	0.393	0.017	0.001	1072
		0.981	*0.971*	

Note: SW+P-C gives the difference between the combined social worker and physician group and the control group, SW-C gives the difference between the social worker program and the control group. P-values in italics below the differences have been adjusted for multiple testing following List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. For variables that have been removed from the family because of colinearity (signaled by !), the per comparison bootstrap p-values is provided. N gives the number of observations in the regression.

B Proofs for theoretical framework

Proof that $\mathbf{h_i|I_i^0, I_i^1 \hookrightarrow \mathcal{N}(\pi I_i^1 + (1-\pi)I_i^0, \widetilde{\sigma}^2)}$

Step 1:

Assuming young people's own perceptions, I^0, and physcian's signal, I^1, are independent conditional on true health status, $l(I^0, I^1|h) = l(I^0|h)l(I^1|h)$, the conditional distribution $l(h|I^0, I^1)$ can be written as a function of $l(h|I^0)$ and $l(I^1|h)$:

$$l(h|I^0, I^1) = C(I^0, I^1)l(h|I^0)l(I^1|h) \tag{9}$$

where $C(I^0, I^1)$ is a normalization function.

To see that, we apply Bayes' rule:

$$l(h|I^0, I^1) = \frac{l(I^0, I^1|h)l(h)}{l(I^0, I^1)} \tag{10}$$

Then using the conditional independence assumption, we can rewrite (10) as:

$$l(h|I^0, I^1) = \frac{l(I^0|h)l(I^1|h)l(h)}{l(I^0, I^1)} = \frac{l(h|I^0)l(I^0)l(I^1|h)}{l(I^0, I^1)}$$

where we have again applied the Bayes' rule to $l(I^0|h)$. Equation (9) thus holds with $C(I^0, I^1) = l(I^0)/l(I^0, I^1)$.

Step 2:

Specifying $l(h|I^0)$ and $l(I^1|h)$ as normal distributions, $h|I^0 \rightsquigarrow \mathcal{N}(I^0, \sigma_0^2)$ and $I^1|h \rightsquigarrow \mathcal{N}(h, \sigma_1^2)$, we get that the perceived health conditional on initial perceptions and physician's signal is also distributed as a normal variable with moments:

$$h|I^0, I^1 \rightsquigarrow \mathcal{N}((I^0\sigma_1^2 + I^1\sigma_0^2)/(\sigma_0^2 + \sigma_1^2), \sigma_0^2\sigma_1^2/(\sigma_0^2 + \sigma_1^2))$$

From the normality assumptions, we indeed get that $l(h|I^0) \propto exp(-0.5(h-I^0)^2/\sigma_0^2)$ and $l(I^1|h) \propto exp(-0.5(h-I^1)^2/\sigma_0^2)$. Using the result obtained in step 1, the posterior $l(h|I^0, I^1)$ satisfies:

$$l(h|I^0, I^1) \propto exp(-0.5(h^2(1/\sigma_0^2 + 1/\sigma_1^2) - 2h(I^0/\sigma_0^2 + I^1/\sigma_1^2)))$$

from which we get that $h|I^0, I^1$ has a normal distribution $\mathcal{N}(\widetilde{m}, \widetilde{\sigma}^2)$, where $\widetilde{m}$ and $\widetilde{\sigma}^2$ relate to the previous expression through $1/\widetilde{\sigma}^2 = 1/\sigma_0^2 + 1/\sigma_1^2$ and $\widetilde{m}/\widetilde{\sigma}^2 = (I^0/\sigma_0^2 + I^1/\sigma_1^2)$.

Proof that $\mathbf{Pr(\underline{h_i}|I_i^0, I_i^1) - Pr(\underline{h_i}|I_i^0) \propto \alpha I_i^1 - \beta I_i^0}$ with $\alpha > \beta > 0$

To get the approximation, we develop the function Φ around 0 in equations (2) and (3) (section 4.2), as we are mainly interested in youth perceptions and physicians' signal around the threshold 0 ($h_i > 0$ or $h_i < 0$). We then get:

$$Pr(\underline{h_i}|I_i^0, I_i^1) - Pr(\underline{h_i}|I_i^0) \propto \pi \frac{I_i^1 - I_i^0}{\widetilde{\sigma}} + \frac{I_i^0}{\widetilde{\sigma}} - \frac{I_i^0}{\sigma_0} = \frac{\pi}{\widetilde{\sigma}} I_i^1 + \left(\frac{1-\pi}{\widetilde{\sigma}} - \frac{1}{\sigma_0}\right) I_i^0$$

Hence in equation (4) (section 4.2), $\alpha = \pi/\widetilde{\sigma}$ and $\beta = -((1-\pi)/\widetilde{\sigma} - 1/\sigma_0)$.
Obviously $\alpha > 0$, and $\alpha > \beta$ since $\alpha - \beta = 1/\widetilde{\sigma} - 1/\sigma_0$, which is clearly positive as $\widetilde{\sigma}^2 = \sigma_0^2\sigma_1^2/(\sigma_0^2 + \sigma_1^2) < \sigma_0^2$. To get $\beta > 0$, notice that given the expression of β, $\beta > 0$ is equivalent to $(1-\pi)\sigma_0^2 < \widetilde{\sigma}_i\sigma_0$. Given the expressions of π and $\widetilde{\sigma}^2$, we have $1 - \pi = \sigma_1^2/(\sigma_0^2 + \sigma_1^2)$, and $(1-\pi)\sigma_0^2 = \widetilde{\sigma}^2$. Then $\beta > 0$ is equivalent to $\widetilde{\sigma}^2 < \widetilde{\sigma}\sigma_0$, which is granted since $\widetilde{\sigma}^2 < \sigma_0^2$.

Proof that $\mathbf{E(Y_i^T(I_i^0, I_i^1) - Y_i^C(I_i^0)|I_i^1, I_i^0) \propto \alpha I_i^1 - \beta I_i^0}$

We can easily translate the change in the perception of being in poor health in a change in the decision to invest in one's health, assuming that, in the youth population, $c_i/(\overline{U_i} - \underline{U_i})$ is independent from health information, I_i^1 and I_i^0.
Then the treatment effect conditional on the extended information set I_i^1, I_i^0 writes as:

$$E(Y_i^T(I_i^0, I_i^1) - Y_i^C(I_i^0)|I_i^1, I_i^0) = S(Pr(\underline{h_i}|I_i^0, I_i^1)) - S(Pr(\underline{h_i}|I_i^0)) \propto \alpha I_i^1 - \beta I_i^0$$

where S is the survival function of $c_i/(\overline{U_i} - \underline{U_i})$. From a linear expansion of the survival function around $Pr(\underline{h_i}|I_i^0)$, we get that $S(Pr(\underline{h_i}|I_i^0, I_i^1)) - S(Pr(\underline{h_i}|I_i^0)) \propto Pr(\underline{h_i}|I_i^0, I_i^1) - Pr(\underline{h_i}|I_i^0)$.

Notice that the impact of providing information about health needs on young people's decision to invest in their health is heterogeneous. Nonetheless, when we study the *average* impact of information programs, we ignore this heterogeneity. If the initial health perceptions, I_i^0's, are large on average, the final impact can be rather weak. It can even be negative if the initial health perceptions are on average well above the average of the signals sent by the physician, I_i^1's.

C Signal estimation

In this section, we first explain the LASSO procedure we implement to select the predictors used to estimate the physicians' signals. Then we give the list of selected variables for the psychologist recommendation prediction and the dentist recommendation prediction.

C.1 LASSO procedure

We use a LASSO procedure to select relevant variables to predict the physician signal because we cannot use all the baseline variables, as this would lead to having more variables than observations. The advantage of the LASSO (compared to the older ridge regression method in particular) is to shrink the coefficients of irrelevant variables to exactly zero and hence operate a selection of a subset of relevant variables.

We run the procedure as follows:

1. We start by randomly splitting the sample of young people who met the physicians into two subsamples, and we choose a grid of possible values in $[0, +\infty[$ for the penalty parameter λ that shrinks the regression coefficients (when $\lambda = 0$, we are back to usual OLS, there is no shrinkage)

2. We fit the regression for each value of λ in the first subsample (train sample) and compute

the corresponding Mean Squared Error in the second subsample (test sample)

3. We select the λ that gives the lowest MSE, λ^*, and compute the vector of coefficients β that minimize the penalized Residual Sum of Squares in the test subsample, where the penalty parameter is λ^*. Estimates of β are presented in Appendix C.2 for all the variables whose coefficient was not shrunk to zero, which correspond to the selected variables

Then we take the selected variables and estimate a standard logit regression of each of the two recommendations (consultation with a psychologist or a dentist) on the selected variables for the sample of young people who met with the physicians. With the estimated coefficients, we compute the predicted signals for young people who did not meet the physicians (everybody except the estimation sample). To compute the prediction for the sample of young people who met with the physicians (the previous estimation sample), we estimate one logit regression per young person in this sample, each time removing him from the estimation (leave-one-out). In line with the LASSO procedure that splits the sample into a train and a test subsample, the leave-one-out procedure is a cross-validation procedure aimed at avoiding overfitting.

C.2 Variables selected to predict the physician's signal

In this section, we present the lists of variables selected by the LASSO procedure for the psychologist recommendation and for the dentist recommendation. We start with the full set of 386 baseline variables. For the psychologist recommendation, the LASSO selects eight variables: the binary variables indicating whether young people declared they forwent laboratory analyses, whether they think their body withstands illness very well, whether their Beck score corresponds to severe hopelessness (Beck et al. (1974)), whether they consulted a psychologist in the year before the baseline survey, whether they forwent medicines, whether they tried to commit suicide at least once between 13 and 18 years old, whether they rate their psychological health as bad, and their level of satisfaction with their love life.[31] This set of variables gives a pseudo R2 for

[31] The seed for the random initial split (step 1 of the algorithm, see C.1) was set to 123456789. We randomly changed this seed to check the sensitivity of the selection of predictors to the initial split, but the same set of variables tend to be selected, and the final results are similar. This inference issue has been studied in Chernozhukov et al. (2018), who develop a methodology to fully account for split dependence. In their case, the machine learning is used to directly measure individual-specific treatment effects. Our case is different since the

the prediction of the signal of approximately 0.23. For the dentist recommendation, the LASSO only selects the variable capturing how young people rate their dental health. The pseudo R2 for the prediction of the signal of approximately 0.06.

Table A.3: Variables selected by the LASSO procedure for the psychologist recommendation

Variable	Coefficient
Forewent laboratory analyses	0.020
Body withstands illness very well	-0.084
Severe hopelessness Beck	0.083
Consulted psychologist	0.151
Forewent medicines	0.050
Tried to commit suicide 13-18 yo	0.010
Bad psychological self-rated health	0.709
Level of satisfaction love life	-0.054

Table A.4: Variables selected by the LASSO procedure for the dentist recommendation

Variable	Coefficient
Perceived dental health	0.129

LASSO is used to extrapolate one of the two explanatory variables (the physician's referrals) that the theoretical model predict to determine heterogeneity.

D Robustness analysis

In this section, we present results for the general analysis based on slightly different specifications: logit specification instead of linear probability model, and incorporation of control variables in equation (6). For the choice of the control variables to add, since we have many variables in the baseline survey and no strong theoretical prior on which variables are the best predictors of the outcome, we rely on the method described in Belloni et al. (2014). The method consists in selecting the variables that best predict each of the two treatments with two separate LASSO regressions and then selecting the variables that best predict the outcome with a LASSO regression as well. In the final regression (equation (6)), we include the union of the three sets of selected variables as controls. In the first two LASSO regressions, no variable is selected, which confirms that the randomization was successful. In the LASSO regression of the outcome variables, different covariates are selected for each outcome, but often, the algorithm selects the baseline variable corresponding to the outcome variable (e.g. self-rated general health at baseline is selected among the best predictors of self-rated general health at endline).

Table A.5: Robustness to nonlinear specification (logit, no control variables)

Variable	SW+P	SW	N
Insurance coverage :			
Incomplete coverage	-0.039	-0.008	1093
	(0.027)	(0.026)	
-No insurance	0.013	0.004	1093
	(0.016)	(0.017)	
-(CMU or SS) and no comp	-0.053	-0.011	1093

Continued on next page...

... table A.5 continued

Variable	SW+P	SW	N
	(0.024)	(0.022)	
Complete coverage	0.034	0.006	1093
	(0.033)	(0.032)	
-(CMU or SS) and CMU-C	0.061	0.070	1093
	(0.032)	(0.032)	
-(CMU or SS) and private	-0.025	-0.061	1093
	(0.037)	(0.037)	
Does not know coverage	0.004	0.002	1093
	(0.022)	(0.023)	
Healthcare consumption :			
Consulted GP	0.041	-0.020	1091
	(0.031)	(0.030)	
Consulted dentist	-0.041	0.016	1099
	(0.037)	(0.037)	
Consulted psychologist	0.076	0.021	1100
	(0.024)	(0.027)	
Consulted gynecologist	-0.051	-0.064	589
	(0.050)	(0.049)	
Consulted any other practitioner	-0.030	-0.023	1077
	(0.037)	(0.037)	
Health practices :			
Has regular doctor	0.050	0.020	1091
	(0.028)	(0.027)	
Contracep. used last intercourse	0.052	0.015	1068
	(0.031)	(0.031)	
Exercises and/or eats healthy	0.043	0.018	1078

Continued on next page...

... table A.5 continued

Variable	SW+P	SW	N
	(0.037)	(0.038)	
Financial cost :			
You need money to get care	-0.084	-0.008	1092
	(0.030)	(0.033)	
Transaction costs :			
Difficult to know who to consult	-0.008	0.020	1086
	(0.034)	(0.035)	
Difficult to find available physician	-0.024	-0.024	1088
	(0.037)	(0.037)	
Difficult to find good physician	-0.018	0.029	1087
	(0.037)	(0.037)	
Trust :			
Physicians know what is good	0.001	-0.013	1095
	(0.030)	(0.030)	
Medicine has effective solutions	-0.045	-0.036	1091
	(0.033)	(0.033)	
Difficult to confide in physicians	0.031	0.058	1092
	(0.037)	(0.037)	
Health status :			
Good general self-rated health	-0.019	-0.034	1097
	(0.029)	(0.029)	
Good physical self-rated health	-0.041	-0.015	1098
	(0.029)	(0.030)	
Good psychological self-rated health	-0.011	-0.080	1096
	(0.032)	(0.031)	
Has a chronic health pb	0.012	-0.014	1095

Continued on next page...

... table A.5 continued

Variable	SW+P	SW	N
	(0.028)	(0.029)	
Moderate to severe hopelessness	-0.003	-0.028	877
	(0.026)	(0.027)	
Work and training :			
Did training course(s)	0.078	0.054	1092
	(0.037)	(0.037)	
Participated in workshop(s)	0.042	0.033	1082
	(0.032)	(0.033)	
Worked	0.027	0.005	1091
	(0.032)	(0.032)	
Social integration :			
Involved in associations	0.050	0.013	1095
	(0.028)	(0.030)	
Has more than 10 friends	0.041	0.009	1094
	(0.018)	(0.020)	
Has someone (s)he can count on	-0.002	-0.001	1098
	(0.017)	(0.017)	

Note: SW+P gives the impact of physician plus social worker intervention versus control group, SW gives the impact of social worker intervention versus control group. They correspond to average marginal effects. Standard errors are in parentheses below. N gives the number of observations in the regression.

Table A.6: Robustness to control variables (based on Belloni et al. 2014)

Variable	SW+P	SW	N
Insurance coverage :			
Incomplete coverage	-0.039	-0.011	1093
	(0.027)	(0.027)	
-No insurance	0.015	0.004	1093
	(0.016)	(0.015)	
-(CMU or SS) and no comp	-0.053	-0.016	1092
	(0.023)	(0.024)	
Complete coverage	0.025	0.005	1088
	(0.032)	(0.032)	
-(CMU or SS) and CMU-C	0.056	0.060	1088
	(0.027)	(0.029)	
-(CMU or SS) and private	-0.033	-0.056	1088
	(0.032)	(0.032)	
Does not know coverage	0.008	0.005	1093
	(0.022)	(0.022)	
Health insurance quiz (0 to 5)	0.033	-0.050	1091
	(0.091)	(0.092)	
Healthcare consumption :			
Consulted GP	0.045	-0.021	1083
	(0.028)	(0.030)	
Consulted dentist	-0.035	0.030	1098
	(0.036)	(0.036)	
Consulted psychologist	0.080	0.017	1100

Continued on next page...

... table A.6 continued

Variable	SW+P	SW	N
	(0.024)	(0.022)	
Consulted gynecologist	-0.044	-0.049	582
	(0.048)	(0.048)	
Consulted any other practitioner	-0.032	-0.004	1065
	(0.036)	(0.036)	
Health practices :			
Has regular doctor	0.057	0.034	1090
	(0.026)	(0.027)	
Contracep. used last intercourse	0.056	0.016	1068
	(0.031)	(0.032)	
Exercises and/or eats healthy	0.036	0.021	1052
	(0.034)	(0.035)	
Financial cost :			
You need money to get care	-0.065	0.001	1089
	(0.029)	(0.028)	
Transaction costs :			
Difficult to know who to consult	-0.009	0.019	1086
	(0.035)	(0.034)	
Difficult to find available physician	-0.024	-0.024	1088
	(0.037)	(0.038)	
Difficult to find good physician	-0.016	0.030	1085
	(0.037)	(0.037)	
Trust :			
Physicians know what is good	0.001	-0.012	1095
	(0.029)	(0.030)	
Medicine has effective solutions	-0.040	-0.035	1091

Continued on next page...

... table A.6 continued

Variable	SW+P	SW	N
	(0.033)	(0.034)	
Difficult to confide in physicians	0.031	0.058	1092
	(0.036)	(0.037)	
Health status :			
Good general self-rated health	-0.020	-0.028	1086
	(0.026)	(0.025)	
Good physical self-rated health	-0.044	-0.012	1097
	(0.027)	(0.026)	
Good psychological self-rated health	0.006	-0.066	1093
	(0.028)	(0.030)	
Has a chronic health pb	0.016	-0.010	1092
	(0.026)	(0.026)	
Moderate to severe hopelessness	-0.023	-0.044	705
	(0.029)	(0.029)	
Work and training :			
Did training course(s)	0.075	0.052	1092
	(0.037)	(0.037)	
Participated in workshop(s)	0.051	0.035	1082
	(0.031)	(0.031)	
Worked	0.027	0.012	1081
	(0.030)	(0.031)	
Social integration :			
Involved in associations	0.044	0.015	1062
	(0.028)	(0.027)	
Has more than 10 friends	0.043	0.007	1094
	(0.018)	(0.015)	

Continued on next page...

... table A.6 continued

Variable	SW+P	SW	N
Has someone (s)he can count on	-0.002	-0.002	1098
	(0.017)	(0.017)	

Note: SW+P gives the impact of physician plus social worker intervention versus control group, SW gives the impact of social worker intervention versus control group. Robust standard errors are in parentheses below. N gives the number of observations in the regression.

Supplementary material

I Determinants of the participation to the programs

Table B.1: Determinants of the program take-up for each program

Variable	Diff. SW+P	N	Diff. SW	N
Socioeconomic characteristics :				
Age (in years)	0.179	530	-0.156	499
	0.744		*0.964*	
Female	0.051	531	0.054	501
	0.834		*0.892*	
Less than high school diploma	-0.061	529	-0.059	501
	0.721		*0.940*	
Has difficulties in French	0.042	530	0.016	501
	0.832		*0.990*	
Has been placed to DCFS	-0.043	530	-0.033	496
	0.537		*0.926*	
Family had serious housing pbs	0.023	530	-0.011	500
	0.808		*0.943*	
Family had severe financial pbs	-0.015	528	-0.056	500
	0.904		*0.922*	
Difficulties in covering expenses	-0.003	526	-0.126	496
	0.932		*0.238*	
Already had a job	-0.140	532	-0.024	501
	0.402		*0.803*	

Continued on next page...

... table B.1 continued

Variable	Diff. SW+P	N	Diff. SW	N
Insurance :				
No insurance	0.033	526	-0.037	500
	0.441		*0.722*	
(CMU or SS) and no comp	0.018	526	0.007	500
	0.597		*0.868*	
(CMU or SS) and CMU-C	-0.049	526	-0.045	500
	0.454		*0.844*	
(CMU or SS) and private	0.045	526	0.043	500
	0.501		*0.688*	
Does not know coverage	-0.047	526	0.033	500
	!0.160		*!0.377*	
You need money to get care	0.088	530	0.046	500
	0.120		*0.808*	
Health :				
Bad general self-rated health	0.034	530	-0.058	501
	0.524		*0.523*	
Bad physical self-rated health	0.042	529	-0.081	501
	0.548		*0.346*	
Bad psychological self-rated health	0.087	528	-0.142	499
	0.137		*0.048*	
Has a chronic health pb	0.019	530	-0.048	501
	0.583		*0.532*	
Moderate to severe hopelessness	0.051	389	-0.018	396
	0.433		*0.731*	

Continued on next page...

... table B.1 continued

Variable	Diff. SW+P	N	Diff. SW	N
Perceived vulnerabilities :				
Get sick more easily than other people	0.062	530	-0.011	499
	0.305		*0.972*	
Body withstands illness very well	-0.126	525	0.115	497
	0.000		*0.181*	
Usually catch viruses	0.069	529	0.048	501
	0.371		*0.844*	
Easily anxious when something wrong	0.018	530	-0.005	500
	0.700		*0.929*	
Expect better health	0.061	516	0.183	490
	0.334		*0.012*	
Has already been seriously ill	0.080	528	0.051	499
	0.183		*0.829*	
Has been so ill thought die	0.053	530	0.018	501
	0.348		*0.969*	
Health investments :				
Consulted GP	0.031	526	-0.028	497
	0.956		*0.988*	
Consulted dentist	0.029	531	0.006	501
	0.973		*0.994*	
Consulted psychologist	0.035	532	0.001	501
	0.726		*0.968*	
Consulted gynecologist	-0.026	266	0.018	264
	0.681		*0.999*	
Consulted another practitioner	-0.028	525	0.011	492

Continued on next page...

... table B.1 continued

Variable	Diff. SW+P	N	Diff. SW	N
	0.903		*0.995*	
Has regular doctor	0.024	530	0.163	501
	0.827		*0.064*	
Exercises and/or eats healthy	0.027	515	-0.113	489
	0.949		*0.306*	

Note: Diff. SW+P gives the difference between people who met the physician and social worker and people who did not among people assigned to physician and social worker intervention. Diff. SW gives the difference between people who met the social worker and people who did not among people assigned to social worker intervention only. In italics below are the p-values, they are adjusted for multiple testing following List, Shaikh and Xu (2019). Families correspond to the list of variables in each section of the table. For variables that have been removed from the family because of colinearity (signaled by !), the per comparison bootstrap p-values is provided. N gives the number of observations in the regression.

II Presentation of the experiment

Your job center is participating in a study on young people health. Health is important, but we know that a lot of young people have difficulties getting healthcare. Yet, to find an internship or a job, you need to feel good. This study is important because it will allow us to better understand what young people's needs are, but it is also important because it could help to reduce difficulties young people in France face in their access to healthcare services.

Would you be willing to spend some time asnwering a questionnaire? It would be very valuable for us. Of course, your participation to this study is not compulsory, but the more young people that participate, the more reliable the study will be.

In this study, a randomly determined share of the participants will be invited to meet healthcare

professionals working in the job center and will possibly have a medical follow-up. Whatever the result of the lottery, you will have the same follow-up with your counselor in the job center. This study is fully anonymous: names of participants will be erased from any data transmitted to the researchers in charge of the study.

III List of questions of the endline survey

This table presents the list of questions asked in the endline survey and indicates the ones we use for the impact analysis.

Table B.2: List of endline survey variables

Variable	Used
Socioeconomic situation :	
Persons you currently live with?	
Their employment status?	
Your current financial resources?	
Any work project? What project?	x
Perceived chances to get a job?	
Worked past year? When? Full or part time? Number of hours worked?	x
Training courses? How many? How many days?	x
Workshops? How many? How many days?	x
Involved in associations? How many?	x
Provide daily help to someone in need for free? Time it takes?	
Number of friends? How often do you have contacts with them?	x
Have someone you can count on if you need it?	x
Feel alone?	

Continued on next page...

... table B.2 continued

Variable	Used
Are there sick or disabled persons in your household?	
Do you trust school? Police? Health system? Media? Justice? Government?	
Political parties?	
Health :	
General state of health?	x
Physical state of health?	x
Psychological state of health?	x
Any chronic health problems?	
Concerned with a close relative's health?	
Beck Hopelessness Questionnaire	x
Think physicians know best what is good for a patient?	x
Think you need money to have good health?	
Think you need money to get healthcare?	x
Think illness and recovery depends on God or Fate?	
Think medicine has effective solutions for any health problem?	x
Think information provided by physicians is hard to understand?	
Think recommendations made by physicians are hard to apply?	
Think you can be seriously sick without noticing it?	
Think it is difficult to find good physicians?	x
Think it is difficult to find a physician who have available slots?	x
Think it is difficult to know who to consult?	x
Think it is difficult to confide in physicians?	x
Think you get sick more easily than other people?	
Think your body withstands illness very well?	
Think when there is a virus you catch it?	

Continued on next page...

... table B.2 continued

Variable	Used
Think you worry easily when something is wrong?	
Think you have a very healthy life ?	
Think in the future you will have better health than people you know?	
Think you have already been seriously sick?	
Think you have already been so sick that you thought you would die?	
Think you care more about your health than most people?	
Do you do something regularly to have better health? What?	x
Have a regular doctor to consult in priority if health problem? If not, why?	x
Do you trust him?	
Consulted a physician in a hospital apart from emergency past year? Why?	
Have been to the emergency room the past year? How many times? Why?	
Consulted a GP past year? How many times?	x
Consulted a company doctor past year? How many times?	x
Consulted a dentist past year? How many times?	x
Consulted a psychologist past year? How many times?	x
Consulted a pharmacist past year? How many times?	x
Consulted an alternative medicine doctor past year? How many times?	x
Consulted a radiologist past year? How many times?	x
Consulted a gynecologist past year? How many times?	x
Consulted an ophthalmologist past year? How many times?	x
Consulted a pediatrician past year? How many times?	x
Consulted a dermatologist past year? How many times?	x
Consulted a cardiologist past year? How many times?	x
Consulted a rheumatologist past year? How many times?	x
Consulted a psychiatrist past year? How many times?	x
Consulted an ENT specialist past year? How many times?	x

Continued on next page...

... table B.2 continued

Variable	Used
Consulted a surgeon past year? How many times?	x
Consulted an anesthetist past year? How many times?	x
Consulted a vein specialist past year? How many times?	x
Consulted a pulmonologist past year? How many times?	x
Consulted a diabetologist past year? How many times?	x
Consulted another specialist past year? How many times?	x
Have a close relative working in the health sector?	
Women only: Followed up on gynecological health? If not, why?	
Forwent healthcare past year? Why? Which type of care? Were they important?	
Would like to have a doctor close to your home?	
Would like to have emergency service close to your home?	
Would like to have access to free health care?	
Would like to know a pharmacist close to your home?	
Would like to know a pharmacist who gives you medicines without a prescription?	
Would like to be in touch with a health worker you fully trust?	
Anything detrimental to your health in your current life?	
Have basic health coverage? Standard or CMU?	x
Have a complementary health coverage? Private or CMU-C?	x
CMU-C only: CMU-C helped to get better healthcare?	
Followed for a long-term health problem? Since when?	
Followed for disability? Receive disability benefit?	
Contraception used for the last intercourse? How did you get it?	x
Satisfied with - your diet? - your health? - your love life? - your family relationships?	
- your neighborhood? - your level of self-confidence? - your life in general?	
Information and knowledge :	

Continued on next page...

... table B.2 continued

Variable	Used
Interested in information on health and medicine?	
Feel informed on alcohol? Tobacco? Cannabis? Contraception? STIs? HIV?	
Air pollution? Water pollution? Diet? Vaccines? Cancer? Depression?	
Where do you look for health information?	
In the past year, have you looked for information on - pregnancy? - contraception?	
- health centers, hospitals or physicians? - travelers health? - children health?	
- sexuality? - depression? - tobacco, drug or alcohol? - nutrition? - a specific illness?	x
- a specific medicine? - insurance system procedures? - blood or organ donation?	
- health associations?	
Think the French health system works well?	
Think the French health system allows to treat everybody?	
Think the French health system is badly managed?	
Think the French health system allows people to choose one's physician?	
Think the French health system has good physicians?	
Think the French health system cost too much?	
Insurance system quiz?	x
Morning after pill quiz?	
Have your insurance card on you?	
Know the insurance center address?	
Have you been there in the past year?	
Satisfaction about the programs (only for treated people) :	
Met the physician from the experiment past year? How many times? If not, why?	x
Think it was nice because you usually never meet with physicians?	
Think it was useless because you did not need it?	x
Think you needed it but did not suspect it?	x

Continued on next page...

... table B.2 continued

Variable	Used
Think you would have consulted a physician anyway?	x
Think you did not understand why the physicians asked certain questions?	
Think you learned health-related things?	
Think confiding to a physician was good?	
Think it improved your health?	
Think consultation was too long?	
Think you would have liked to consult the physician more times?	
Think it convinced you to see doctors more regularly?	
Think you were anxious before the consultation?	
Think you were anxious after the consultation?	
Met the social worker from the experiment past year? How many times? If not, why?	x
Think it was nice because you usually never meet with social workers?	
Think it was useless because you did not need it?	x
Think you needed it but did not do the procedures?	x
Think you would have seen a social worker anyway?	
Think you did not understand why the social worker asked certain questions?	
Think you learned things?	
Think it improved your situation?	
Think meeting was too long?	
Think it discouraged you because procedures look too complex?	
Think you did not dare to ask the social worker certain questions?	
Think it convinced you to exercise your insurance rights?	x
Think you were anxious after the meeting?	

Acronyms

CMU Universal (basic) Medical Coverage

CMU-C Universal Complementary Medical Coverage

INSEE National Institute of Statistics and Economic Studies

SIRS Health, Inequalities and Social Isolation