I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13582

The Role of Beliefs in Long Sickness Absence: Experimental Evidence from a Psychological Intervention

Gabriel Pons Rotger
Michael Rosholm

AUGUST 2020

IZA Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13582

The Role of Beliefs in Long Sickness Absence: Experimental Evidence from a Psychological Intervention

Gabriel Pons Rotger
VIVE

Michael Rosholm
Aarhus University and IZA

AUGUST 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

The Role of Beliefs in Long Sickness Absence: Experimental Evidence from a Psychological Intervention[1]

This paper makes use of the randomized allocation of workers on sick leave in Denmark into self-management support, to examine the role of beliefs about control for prolonged absenteeism due to illness. Our results demonstrate that the ability of the intervention to lead sick-listed workers toward resuming employment crucially depends on workers' control beliefs. The intervention increases the perception of control among control pessimists and substantially accelerates the decision to return to work. Furthermore, we identify a group of control-optimist workers for whom "learning" about control beliefs is self-defeating, and leads them toward reduced capacity in terms of return-to-work performance.

JEL Classification: J21, C93, D91

Keywords: sickness insurance, personality traits, randomized control trial, machine learning

Corresponding author:
Gabriel Pons Rotger
VIVE - The Danish Center for Social Science Research
Herluf Trolles Gade 11
1052 Copenhagen K
Denmark
E-mail: gpr@vive.dk

[1] This article is part of the project, "Evaluation of Lær at tackle job og sygdom." Rotger's participation is financed by a grant from the Danish Ministry of Employment. We would like to thank seminar participants in the Labor and Public Policy seminar at Aarhus University, the Rockwool Foundation, The Danish Center for Social Science Research, Workshop on Health and the Labour Market at the University of Nantes–IAE, for helpful comments. Preregistration protocol is ClinicalTrials.gov: NCT02136056. Neither the authors nor their employers have relevant or material financial interests that relate to the research described in this paper. The authors are responsible for all remaining errors and shortcomings.

I. Introduction

Long-term absence from work due to illness is often associated with lasting negative effects on subsequent employment, leading to economic and social deprivation and permanent disability (see e.g., Hansen 2000; Gjesdal and Bratberg 2003; Gjesdal et al. 2004; Kivimäki et al. 2004; Ichino and Moretti 2009; Markussen 2012). As the consequences of prolonged absence and its associated costs to society are considerable, understanding the roots of long sickness spells is crucial for designing interventions to help persons on the long-term sick list return to stable employment trajectories (Johansson and Palme 2002; Henderson, Glozier, and Elliott 2005; Heymann et al. 2010; Holm et al. 2017).

Numerous studies have demonstrated the importance of moral hazard in absence behavior (e.g., Dionne and St-Michel 1991; Johansson and Palme 1996, 2002, 2005; Henreksson and Persson 2004; Ichino and Riphahn 2005; Hesselius and Persson 2007; Puhani and Sonderhof 2010; Ziebarth and Karlsson 2010, 2014; Fevang, Markussen, and Røed 2014). However, the evidence also shows that, presumably due to the severity of the health impairment, the issue of moral hazard is less prevalent for prolonged sickness absence (Ziebarth 2013).

However, many long sickness spells start with common conditions like musculoskeletal and mental health disorders that, for whatever reasons, do not improve sufficiently over time (Henderson, Glozier, and Elliott 2005), and, apart from the influence of illness itself (e.g., Bültmann et al. 2006; Christensen et al. 2007; Dekkers-Sánchez et al. 2008), evidence is lacking for factors associated with long-term sick leave.

Our study contributes to the literature with experimental evidence regarding the role of individuals’ control beliefs in prolonged absence from work. Control beliefs such as locus of control and self-efficacy have long had a central place in social psychology as fundamental factors for motivation, effort (Rotter 1966; Bandura 1986, 1989), and self-control (Rosenbaum 1980). Locus of control refers to the extent to which individuals believe that what happens to them is related to their own behaviors or actions, or on the contrary to external factors such as luck (Rotter 1966). Self-efficacy captures the strength of individuals' beliefs in their own ability to undertake behavior or actions to improve their situation (Bandura 1986, 1989).

In economics, control beliefs are considered to be a key personality trait for labor market behavior (e.g., Coleman and DeLeire 2003; Borghans et al. 2008; Heckman, Stixrud, and Urzua 2006; Heckman and Kautz 2012; Cobb-Clark and Schurer 2013; Spinnewijn 2013, 2015; Caliendo, Cobb-Clark, and Uhlendorff 2015; Lekfuangfu et al. 2018), and growing empirical evidence suggests that beliefs about control may be of key importance in avoiding prolonged absence. Such evidence demonstrates that the locus of control is positively associated with emotional resilience against health shocks (Buddelmeyer and Powdthavee 2016), with healthy behavior (Chiteji 2010; Cobb-Clark, Kassenboehmer, and Schurer 2014), with labor supply after severe health shocks (Schurer 2017), and with a wide range of labor market outcomes (Gallo et al. 2003; Uhlendorff 2004; Ng, Sorensen, and Eby 2006; Heineck and Anger 2010; Cobb-Clark and Tan 2011; Ahn 2015; Caliendo, Cobb-Clark, and Uhlendorff 2015; McGee 2015; McGee and McGee 2016).

Sick-listed workers are likely to be more pessimistic regarding the controllability of their labor market outcomes than other groups in the labor force (Buddelmeyer and Powdthavee 2016; Schurer 2017; Marsaudon 2019). Furthermore, uncertainty about the consequences of illness on future health (Arrow 1963; Chandra, Handel, and Schwartzstein 2019) and the effectiveness of medical treatment (Frank 2004) may increase the difficulty for long-term sick-listed workers to choose recovery efforts wisely, which in turn may reduce belief in their own capacity for control. Workers with negative beliefs about control may therefore delay their return to work by not putting in adequate health investment (Babcock et al. 2012; Schurer 2017). Furthermore, workers with negative control beliefs may have self-discipline problems and procrastinate in their recovery efforts (Rosenbaum 1980; Bryan, Karlan, and Nelson 2010).

If such behavioral barriers are widespread among long-term sick-listed workers, then economic incentives will be less effective to reduce long absence spells, and control-enhancing interventions might be required to promote return to work. However, regardless of the substantial explanatory power of control beliefs for a wide range of labor market outcomes and health behavior, it is not clear that control-enhancing interventions will help. Evidence shows that locus of control is generally quite stable in working-age populations (Specht, Egloff, and Schmukle 2011, 2013; Cobb-Clark and Schurer 2012, 2013), although recent studies show that control beliefs may decrease among workers displaced from their jobs because of health shocks (Marsaudon 2019) or because their workplace closure (Preuss and Hennecke 2018).

As far as we know, there is no evidence that an intervention directly aimed at boosting control beliefs leads to better labor market outcomes (Almlund et al. 2011; Gutman and Schoon 2013; Cobb-Clark 2015). Gottschalk (2005) provides evidence that temporary subsidized employment of young welfare recipients leads to positive effects on both locus of control and earnings. Pedersen et al. (2015) directly test the effect of a control-enhancing course aimed at return to work among sick-listed workers at risk of mental health disorder—an intervention similar to the program studied in this article—and finds that the positive effects on locus of control were not reflected in increasing subsequent employment. As participants in the course investigated in Pedersen et al. (2015) were recruited shortly after the start of sick leave (4–8 weeks), the lack of responsiveness might reflect that negative perceptions of control are not widespread among short-term sick-listed workers. A review of 10 randomized controlled trials (Franek 2013) concludes that the Stanford

Chronic Disease Self-Management course (CDSMP), also similar to the intervention investigated in our study, led to clinically minimal improvements across a number of health status measures, healthy behaviors, and self-efficacy (Lorig et al. 1986, 1996, 1999, 2001). This report shows that CDSMP did not improve health care utilization, but the intervention was exclusively focused on short-term health outcomes. Moreover, results could be misleading given that they were based on performance of individuals who participated in the program, and did not include intent-to-treat analysis.

Consequently, two key questions remain:

- Does boosting belief in control among workers have a positive impact on employment?
- If so, which workers are the most likely to return to work?

Our article contributes to the empirical literature on behavioral economics with experimental evidence on the effect of a control-enhancing intervention, a self-management support course, on return to work for long-term sick-listed workers in Denmark. As in Malmendier, Tate, and Yan (2011), and in contrast to Buddelmeyer and Powdthavee (2016) and Schurer (2017), we explore variation in individuals' beliefs to identify their role in return to work. Different from Malmendier, Tate, and Yan (2011), we obtain such variation by means of a randomized intervention aimed at affecting individuals' beliefs.[2]

Our estimated mean impacts demonstrate that the intervention had no overall effect on employment rates. Moreover, the program did not contribute to transiting from sickness benefits to unemployment insurance benefits. However, motivated by the heterogeneity of individuals participating in the intervention with respect to their *ex ante* control beliefs, we examine the

[2] Malmendier, Tate, and Yan (2011) explore exogenous variation in managers' personal histories (shaped by the Great Depression and while serving in the military) to identify the role of overconfidence in their financial decision making.

heterogeneity of the employment effect using causal machine-learning techniques (Athey and Imbens 2016; Athey, and Wager 2018; Athey, Tibshirani, and Wager 2019). We provide strong evidence that the effectiveness of the intervention depends crucially on individual's control beliefs. For workers with negative pre-intervention control beliefs, self-management significantly increased the return-to-work rate 17 to 19 months after intervention assignment by 18.54 percent (a 110-percent increase compared to the control group). In contrast, for workers whose pre-intervention beliefs were more positive, self-management support *reduced* regular employment at 17 to 19 months after program assignment by 13.42 percent (a 29-percent reduction compared to the control group).

An exploration of plausible mechanisms reveals two main channels for the employment effect. First, the intervention has a positive impact on perceived personal control of illness among workers with negative control beliefs. This improvement in control beliefs persists for at least 6 months after program participation, strongly suggesting that reinforced control beliefs contribute to workers' return to work.

Furthermore, we identify a group of long-term sick-listed workers for whom "learning" about control beliefs is harmful, as it reduces their return-to-work rate. Concretely, the intervention initially leads control-optimistic workers to perceive their illness as more severe and to put more effort into coping with the psychological consequences of illness than if they had not participated in the program.

Finally, we investigate the influence of course peers on employment effects. Our results go against social learning theory and demonstrate that positive program impacts on workers with low control beliefs will hardly be affected by excluding peers with positive control beliefs from LTJS courses. Social learning during the course does not seem to be relevant, a result that could be explained by the fact that the health disorders affecting program participants are heterogeneous. The lack of peer influence suggests that disclosure of information via course participation has a direct impact on the control beliefs and the labor market behavior of sick-listed workers, leading less confident workers toward accelerating their return to work.

II. Experiment Design

A. Treatment as Usual: Paid Sick Leave in Denmark

In Denmark, employed workers, self-employed workers, and unemployed workers receiving unemployment insurance benefits are eligible for the sick-leave benefit. For wage earners, the benefit replaces 100 percent of wages up to DKK 4,135 per week (as of January 2015, €554/$637). Benefits are primarily funded by taxes. However, in the case of sick-listed employees, the employer covers benefits for the first 30 days of absence, and often also covers part of the difference between the full wage and the sickness benefit (depending on local agreements between trade unions and employers' associations). Employers can easily dismiss sick-listed workers, who will continue to receive sickness benefits for as long as they are eligible. During the period of our experiment, eligibility for sickness benefits was 26 weeks in total. At 26 weeks of absence, an assessment is made, and some workers are deemed eligible for a prolonged sickness benefit period. Individuals who cannot get their sickness benefits extended are eligible for the "job clarification" program, that is, an interdisciplinary assessment of work capacity. Individuals in this program receive a benefit lower than the sickness benefit, corresponding approximately to the level of social-welfare assistance.

In Denmark, the sickness benefit system is regulated by the central government and administered by municipalities. Municipal job centers monitor eligibility conditions of sickness benefit recipients and are obliged to schedule a first follow-up meeting within eight weeks after the first day of work absence. At follow-up, the case worker assesses the need for active intervention to maintain the attachment of the sick-listed worker to the labor market. The intervention toolbox consists of various types of vocational rehabilitation such as job counseling, wage-subsidized, on-the-job training, internships, and professional courses, as well as graded return to work (partial sick leave). For sickness periods lasting longer than eight weeks, subsequent follow-up must take place at four- or eight-week intervals, depending on the expected length of the sickness period.

B. The Intervention: Self-management Support

In late 2014, The Danish Agency for Labour Market and Recruitment in Denmark launched a randomized trial in 27 Danish municipalities. The aim was to examine whether long-term sick-listed workers, who participated in a self-management course called "Learn to Cope with Work and Illness" (in Danish: *Lær at Tackle Job og Sygdom*—LTJS), could achieve faster return-to-work and more stable employment trajectories than those receiving the "usual treatment" as presented previously.[3]

The target group for the intervention was comprised benefit recipients sick-listed from work with the following characteristics:

- Deemed unlikely to return to work quickly at the first follow-up meeting.
- Elapsed sickness benefit period of 6 to 19 weeks.
- Diagnosed with recognized chronic or long-term health problem.
- Wished to return to work.

[3] The intervention was pre-registered at ClinicalTrials.gov:r NCT02136056. Overall results of the intervention were first presented in a report in Danish (Rotger 2019).

- Fluent Danish speakers.
- No physical or psychological issues that would prevent standard course participation.

Participation in the LTJS intervention was voluntary and delivered in addition to the “usual treatment.” Recruitment for the intervention was performed by caseworkers at job centers, typically during the first or second follow-up meeting.

The LTJS program is an adaptation of the Stanford CDSMP (Lorig et al. 1986, 1996, 1999, 2001) given to all-cause long-term sick-listed workers in Denmark. The program was implemented as a group-based course. Each course was attended by an average of nine participants of mixed age, gender, and diagnoses. The courses were taught by two trained volunteers, who acted as facilitators. At least one of the facilitators previously experienced a long-term sickness absence. The facilitators received four days of training based on a teaching manual. Rather than prescribing specific behavior changes, the leaders assist participants in making their own disease management choices to reach self-selected goals. The program consisted of six weekly 2.5-hour sessions. The topics covered during the 15-hour course included information and techniques to manage fatigue, stress, depression, insomnia, pain, anger, shortness of breath, itching, memory problems, and involuntary urination; use of cognitive symptom management techniques; exercise; pacing; nutrition; use of medication; communication with others including health professionals, case workers, employers, and family; cooperation problems with health care services and problem solving.[4]

The course incorporates strategies suggested by Bandura to enhance individuals’ confidence in their own ability to carry out actions necessary to deal with sickness absence and speed up return

[4] The content of the course is described in more detail in The Danish Committee for Health Education (2014).

to work. These strategies include weekly action planning and feedback, reinterpretation of symptoms, group problem solving, individual decision making, and facilitators acting as role models (Bandura 1986, 1989).

The effectiveness of the LTJS program is based on three premises. First, long-term sick-listed workers experience difficulties in managing their health problems. Second, the provision of self-management support enhances their perceived control of sickness absence. Third, workers with positive control beliefs display better behavioral responses to illness and return to work than those with nonexistent or lower levels of positive beliefs.

C. Evaluation Design

The evaluation was designed as a stratified randomized trial implemented between October 2014 and February 2016. As soon as a recruited person had given consent to participate and completed a baseline survey, the Department of Psychology and Behavioral Sciences at Aarhus University randomly allocated the person to the LTJS intervention or the control group. Randomization was conducted separately for each team (i.e., for each team there is an associated control group) with a probability of 65/35 of being randomized into the treatment group. The stratified randomization schedule ensured that we could make comparisons between treatment and control groups within municipal job center and calendar time windows.

Once a sufficiently large number of participants necessary for assembling a team were recruited, the intervention began. This implied that some individuals recruited for a given team had to wait for a considerable period of time from randomization to the course start date. This situation and its consequences for estimated results are analyzed in detail by Rotger (2019).

The LTJS evaluation project initially recruited 657 individuals. We excluded 36 in the treatment group and 17 controls due to municipality drop-out, and another 8 in the treatment group from a municipality that did not allocate any individuals to the control group (i.e., violated the protocol).

Next, we excluded 7 in the treatment group and 6 controls who did not answer the baseline survey and therefore did not participate in the experiment. One person in the treatment group and 2 controls were removed who had withdrawn their consent to participate. Finally, we excluded 7 in the treatment group and 5 controls for whom we could not match data in the administrative registers with their baseline survey information. The analysis sample comprised data on individuals from 27 municipalities and 43 blocks (43 treatment teams and 43 control groups). Seventeen municipalities contributed two teams each, 18 municipalities contributed one team, and two municipalities merged participants into a single team.[5] The resulting sample is comprised of 568 individuals (T = 371, C = 197).[6]

The described sample was used for investigating the main effect of LTJS on return-to-work and sickness absenteeism. Our investigation of heterogeneous effects uses a subsample of 511 observations with non-missing information on all pre-treatment covariates. We also defined a smaller subsample of 400 observations to investigate group effects on several mediating outcomes measured immediately after the course, and six months after the course.

Compliance was remarkably high: seven persons in the treatment group (2 percent) indicated during surveys that they had not participated at all in the intervention. On average, treated individuals participated in 4.7 out of 6 sessions.

III. Data

A. Data Sources

To study the effect of the LTJS intervention, we merged a very rich survey on sickness absence collected by the Unit for Psycho-Oncology and Health Psychology from the University of Aarhus

[5] See Figure A1 in the Online Appendix. Table A1 in the Online Appendix shows the dates for allocation of teams, and number of control and treatment observations for each team.
[6] See Figure A2 in the Online Appendix.

with the Ministry of Employment's longitudinal register database on public transfer incomes and employment (the DREAM register).[7] The DREAM register contains weekly information about income transfers from public programs as well as employment. The DREAM register is known to be of high reliability for measuring sick-leave spells (Stapelfeldt et al. 2012) as well as employment, as it is based on administrative register data for sickness benefit payments and compulsory income reporting by employers to the tax authorities. In the analyses, we use a period of 2 years before randomization to construct labor market and sickness absence histories, and we follow all individuals for approximately 1.5 years after the program allocation week (precisely, we followed them for at least 81 weeks). Thus, for each person, we have a weekly time series covering 3.5 years. Because individuals recruited by the participating municipalities entered the program at different time points between October 2014 and February 2016, everyone's time series on regular employment and sickness absence is located at an individual-specific calendar time interval.

Three surveys were collected; the first immediately before program allocation (baseline LTJS questionnaire), the second one week after completion of the course (first follow-up LTJS questionnaire), and the third, 26 weeks following course end (second follow-up LTJS questionnaire).[8] The surveys contain a wide range of measures, described below, on cause of absence, health status, health behavior, motivation, behavioral self-management, control beliefs, and expectations.

[7] DREAM is a Danish acronym for The Register-Based Evaluation of Marginalization Scope (in Danish: *Den Registerbaserede Evaluering Af Marginaliseringsomfang*). The main outcomes used in the analysis are employment and temporary disability, including sickness benefits (codes 890–899), rehabilitation benefits (codes 760–767), and job clarification programs (codes 870–878).

[8] The survey information was not made accessible to job centers.

B. Variables

The primary outcome used in this study is the proportion of time spent employed 69 to 81 weeks (17–19 months) after participants were placed in treatment and control groups. We consider effects on regular employment for the initial period (1–13 weeks after randomization), and for the whole period of 81 weeks (corresponding to 1–19 months) as secondary outcomes. These choices of outcomes are made to assess the initial impact of the intervention including the waiting time and any lock-in effects, and the total impact of the intervention. The primary outcome measures the longer-term employment, where we try to circumvent the problem of waiting times between randomization and start of the intervention (any anticipation effects will likely taper off after the intervention) and lock-in effects.

Since some may return to unemployment, we also consider as secondary outcomes the number of weeks spent receiving sickness benefits at 69–81 weeks, 1–13 weeks, and 1–81 weeks after participants enter intervention groups.

Our survey includes two measures of control beliefs, our primary mediator of the LTJS. Self-efficacy and locus-of-control scales can both refer to specific outcomes and behaviors (Amtmann et al. 2012; Wallston, Strudler Wallston, and DeVellis 1978) or can be used to measure individuals' perception of control across different contexts (Ajzen 2002). Generalized measures of control beliefs like locus of control, preferred in the empirical economic literature (Cobb-Clark 2015), may imperfectly reflect individuals' perceived controllability of absence behavior, and therefore we favor two sickness absence–specific measures of control beliefs. First, we use the six items from the short form of the University of Washington Self-efficacy Scale (UWSES; Amtmann et al. 2012), which we call "Self-efficacy in Sickness Absence." This instrument was translated into Danish by researchers at the Unit for Psycho-Oncology and Health Psychology, Aarhus University. The scale measures individual's sickness self-efficacy, which refers to people's belief that they can succeed at keeping their illness from interfering with diverse situations and tasks in the future.

The second measure of control beliefs is the third item of the Brief Illness Perception questionnaire (BIPQ), developed by Broadbent et al. (2006), which measures individuals' expected personal control of illness. This BIPQ item, which we call "Perceived Personal Control of Illness," is strongly associated with self-efficacy and is most likely, in the context of sickness absence, to be reflecting internal locus of control beliefs regarding illness and return to work (Broadbent et al. 2006).

Our survey includes one measure of a secondary mediator of the LTJS, cognitive-behavioral self-management (hereinafter referred to as behavioral self-management), the chronic disease self-management questionnaire (CDSM) (6-item questionnaire from Lorig et al. 1996). This scale measures in which ways respondents cope with the psychological aspects of illness, such as by mentally distancing themselves from discomfort.

Our survey information includes several baseline beliefs or items covering individuals' expectations regarding re-employment. These items refer to individuals' perceived certainty to return to work for the number of hours per week as before sickness onset, expected duration of sickness absence, expected job tasks, and expected working-time after re-employment.

Finally, we capture individuals' motivation for work with the Treatment Self-Regulation Questionnaire (Levesque et al. 2007). This scale measures the respondent's autonomous motivation (6 items), introjected regulation (2 items), external regulation (4 items), and amotivation (3 items).

To keep the number of control variables and potential mediators to a manageable level, most of these surveys are summarized in standard scales. The English versions of the scales used to measure mediating outcomes and their construction of final scores are included in the Online Appendix.

The covariate set includes 57 continuous, categorical, and binary indicators mostly taken from the baseline survey. Perceived illness is measured via the BIPQ. We use the Symptom Checklist Scale (12 items) from the Common Mental Disorders Questionnaire (CMDQ) (Christensen et al. 2005) to measure the extent to which the individual is troubled by diverse symptoms, and we also rely on the Illness Worries Scale (7 items) from the CMDQ. We measure mental disorders with the Anxiety Scale (4 items) and the Depression Scale (6 items) obtained from CMDQ (Christensen et al. 2005). Next, perceived physical impairment is measured via a 9-item scale obtained from the SF36 Scale (Bjorner et al. 1998). We measure the consequences of sickness and chronic illness in terms of self-rated pain, fatigue, and shortness of breath. Finally, the survey includes a measure of health-related quality of life (Noerholm et al. 2004) and the WHO-5 questionnaire on well-being (Bech et al. 2003).

We capture *health-related behavior* with different questions. We measure intensity of exercise and physical activity, alcohol consumption (4-item scale from the CMDQ), purchase and consumption of medicine, intensity of contacts with general practitioner (GP), specialists, emergency room episodes, and hospitalization. We acknowledge that the database does not include two important indicators for individuals' lifestyle choices, namely, diet and smoking. However, the available indicators are likely to reflect significant rehabilitation efforts and changes in health-related behavior.

The survey and administrative data includes a wide range of measures describing the socio-economic characteristics of sick-listed workers regarding age, gender, education, marital status, children, employment, cause of sickness, sickness history, multiple disorders, working capacity, and motivation for work.

C. Participants

Table 1 presents the sample means of selected characteristics of program participants by treatment and control group of the three analysis subgroups. Summary descriptive statistics are presented for the entire sample (N = 568, T = 371, C = 197), the sample used to investigate heterogeneity of post-program effects (N = 511, T = 336, C = 175), and the sample used to estimate group effects on mediating outcomes (N = 400, T = 267, C = 133). The table shows that women comprise 75 percent of the sample, average age is about 47 years, and approximately 50 percent of the sample has higher education. Most individuals claim sickness benefits connected to regular employment or self-employment (approximately 70 percent) rather than unemployment. Sixty percent of illnesses and symptoms are caused by work conditions. Approximately half the sample claims sickness benefits because of mental disorders (stress, anxiety, depression, etc.). The fraction of time in regular employment during two years immediately prior to treatment assignment is approximately 50 percent. The fraction of time with sickness benefits during the same period is 25 percent. Finally, the average size of control group "teams" is 5 individuals, and the average size of treatment teams is 9. Averages of selected covariates are remarkably similar across samples.[9]

[Table 1]

[9] The sample used to estimate group effects on mediating outcomes (referred to as Mechanism Analysis Sample) excludes three blocks (23, 33, and 43) due to lack of survey information on variables needed for measuring these outcomes. The exclusion of these blocks slightly reduces the proportion of participants with higher education, but does not alter the means of other covariates.

Panel A in Figure 1 shows the distribution of expected return to work (baseline beliefs) measured immediately before program assignment.[10] The baseline beliefs are heterogeneous, but with a higher concentration of program participants at the bottom of the distribution, as expected from workers with long-term health disorders. Panel B in Figure 1 shows the distribution of self-efficacy in sickness absence (control beliefs) measured immediately before treatment assignment. The figure shows that the recruited group of individuals is characterized by quite different levels of self-efficacy despite their prolonged sickness absence, recognized chronic health problems, and low re-employment prospects. The distribution of self-efficacy in sickness absence is quite stable across types of health disorders, workers' previous status in the labor market, workers' education and across severity of health disorder (see Figures A3 to A6 in Online Appendix).

[Figure 1]

To test for balance in the characteristics of the control and treatment samples, we use two tests. First, we run regressions of all pre-treatment covariates $\mathbb{X}_i$ (binary indicators, categorical variables, and continuous variables) collected immediately before assignment to the treatment indicator (W_i) and a vector of 43 randomized-block fixed effects (B_i):

(1) $\mathbb{X}_{ki} = \delta_k W_i + B_i' \beta_k + u_{ki}$

We test for the significance of δ_k in model (1). Table A2 in the Online Appendix reports estimated coefficients δ_k and the p-values for the null hypothesis of individual characteristic $\mathbb{X}_{ki}$. Balance tests are presented for the samples used to estimate mean impacts, conditional effects, and group effects on mediating outcomes.

[10] See the Online Appendix for the definition of baseline beliefs.

Columns 1 to 3 in Table A2 report estimated coefficients δ_k and p-values for the main sample, for the sample used in the heterogeneity analysis, and for the sample used in the analysis of mechanisms. We detect statistically significant coefficients at the 5-percent level for only two characteristics: physical impairment and expected long illness. Column 2 in the same table presents similar results for the heterogeneity analysis sample. In this case, we reject the null hypothesis of variable balance for three characteristics: physical impairment, expected long illness, and perceived medical treatment effectiveness. Column 3 reports coefficients and p-values for the mechanism analysis sample. In this case, there is only one characteristic, "expected long illness," which has a p-value smaller than 0.05.

Second, we implement a LASSO model selection approach to identify eventual covariates among our set of covariates $\mathbb{X}_i$ that potentially could predict our treatment assignment indicator W_i (Belloni, Chernozhukov, and Hansen 2014). The LASSO algorithm does not select covariates for any of the three samples.

IV. Results

This section presents estimates of intention-to-treat (ITT) effects and heterogeneity for self-management support on labor supply, and mediating outcomes. Section IV.A begins by evaluating the overall effects of the psychological intervention on employment, sickness absence, and control beliefs. Section IV.B provides evidence on heterogeneity of the post-program effects of self-management on employment based on machine learning analysis. Section IV.C discusses the plausibility of alternative mechanisms by examining the effects of self-management support on mediating outcomes. Finally, Section IV.C investigates the role of course peer effects on the effects of the program on subgroups of workers with low- and high-level control beliefs.

A. Mean Impact

Despite the vast resources invested in active labor market policies, growing evidence shows that interventions such as rehabilitation programs and graded return to work tend to have a more limited impact on long-term sick-listed workers than on workers on shorter sick leave (e.g., Frölich et al. 2004; Høgelund et al. 2012; Andrén and Svensson 2012; Kools and Koning 2018; Rehwald, Rosholm, and Rouland 2019). Crucial knowledge about the contribution of behavioral interventions such as control-enhancing course support is missing, and this section investigates the overall impact of the self-management course on return to work for long-term sick-listed workers.

Let Y_i be the observed outcome of individual i, $Y_i(1)$ be the potential outcome of individual i under LTJS assignment, and $Y_i(0)$ the potential outcome without LTJS assignment. The ITT effect, the effect of being assigned to LTJS for individual i is $Y_i(1) - Y_i(0)$, and the average effect for long-term sick-listed workers of being assigned to LTJS is given by $\tau = \mathbb{E}[Y_i(1) - Y_i(0)]$. We estimate the ITT effect τ with the following regression:

(2) $Y_i = \tau W_i + B_i'\beta + X_i'\delta + \varepsilon_i$

where W_i indicates treatment assignment, and X_i is a set of pre-intervention covariates. The vector B_i ensures consistency of the estimated effect in the presence of different probabilities of being treated across randomization blocks, and the inclusion of variables X_i improves the precision of our estimates. Selected covariates are highly correlated with the outcome.[11] Robust standard errors are clustered at the block level.

[11] The set of covariates was selected for each outcome using the LASSO method (Belloni, Chernozhukov, and Hansen 2014).

Table 2 (Panel A, columns 1–3) reports post-program effects (weeks 69–81) on employment and on sickness absence. In this period, the average fraction of time in regular employment in the control group is 31 percent, while the average sickness fraction is 29 percent. The LTJS program does not have a significant impact on return to work in regular employment, nor does it lead to a reduction in sickness absence among long-term sick-listed workers. These results are robust to the inclusion of strong predictors of employment, and slightly different but statistically insignificant when we control in addition for the few unbalanced characteristics.

Table 2 (Panel B, columns 1–3) presents initial effects (weeks 1–13) on regular employment and sickness absence. These are most likely dominated by *ex-ante* waiting time effects (before the intervention begins) and to some extent also by lock-in effects. The average fraction of time with sickness benefits during the first quarter after randomization in the control group is 67 percent, and the average employment fraction in the control group is only 5 percent. In this context, we find that the LTJS intervention increases sickness absence by 4.92 percentage points (7.3-percent increase with respect to the control group). This increase in sickness absence in the intervention group does not lead to a negative employment effect, suggesting that the program initially reduces transitions from sickness benefits to unemployment insurance benefits.

Finally, Panel C in Table 2 presents ITT estimates of program impacts throughout the entire period (weeks 1–81). These effects capture waiting-time, lock-in, and post-program effects on employment. The estimated effect on regular employment is highly insignificant, and the estimated effects on sickness absence is also not significant.

Table 3 presents ITT estimates of program impacts on the main mediating outcome, control beliefs. We present results for four measures of control beliefs. Self-efficacy and perceived personal control of illness are measured at the first and second follow-ups.[12]

Panel A in Table 3 shows small impacts on control beliefs measured immediately after termination of the course. The immediate effect on self-efficacy is roughly a statistically insignificant 5-percent increase compared to the control group. We find slightly stronger evidence perception of illness control. Results show increases between 14 and 16 percent higher than the control group. Panel B of Table 3 shows that the program does not enhance control beliefs at the second follow-up.

Mean impacts shown in this section seem to indicate that even though long-term sick-listed people are exposed to an intervention aimed at improving their control beliefs, neither beliefs nor outcomes are strongly affected.

[Table 3]

B. Heterogeneity of Impact with Control Beliefs

The selection process for the LTJS course was based on the assumption that *all* sick-listed workers with long-term health problems faced behavioral barriers in managing their recovery. However, as shown in Section III.C, individuals recruited for the LTJS reported quite heterogeneous levels of control beliefs. Therefore, we examine whether such heterogeneity led to heterogeneous employment effects obscured by focusing on mean impacts.

The pre-analysis plan establishes workers' control beliefs as primarily mediating outcomes but does not consider the analysis of subgroups with different control beliefs. For example, one might investigate subgroups with different UWSES scores. However, we lack clear guidelines on which

[12] See the Online Appendix for the definition of control beliefs.

subgroups are the most relevant. Furthermore, heterogeneity of impacts may also be driven by characteristics such as baseline beliefs, preferences, or illness severity, which are correlated with control beliefs.

Choosing sample *ex-post* splits opens the possibility of detecting spurious heterogeneity due to searching over many possible subgroups. Therefore we apply machine learning to discover which subgroups are most relevant for program heterogeneity, specifically the causal forest (CF) method recently developed in Wager and Athey (2018) and Athey, Tibshirani, and Wager (2019).[13]

We look at effect heterogeneity as a prediction problem, and assess the relative contribution of control beliefs to the conditional average treatment effect (CATE), that is, the effect of LTJS assignment conditional on all observable characteristics of participants before program assignment ($\mathbb{X}_i$), as follows:

(3) $\mathrm{CATE}(x) = \mathbb{E}[Y_i(1) - Y_i(0) | \mathbb{X}_i = x]$

We first test for the presence of effect heterogeneity of the form

(4) $\mathrm{CATE}(x) = \tau + (x_i - \bar{x})'\gamma$

where $\bar{x}$ is the vector of sample means of $\mathbb{X}$, and the vector of coefficients γ includes all differences in groups effects $\gamma_k = \mathbb{E}[Y_i(1) - Y_i(0) | X_{ki} = 1] - \mathbb{E}[Y_i(1) - Y_i(0) | X_{ki} = 0]$. Model (4) can be estimated in a standard regression framework by interacting the treatment variable W_i with all demeaned binary indicators $Z_{ki} = X_{ki} - \bar{X}_k$[14]:

[13] For other economic applications of CF, see Davis and Heller (2017), Bertrand et al. (2017), or O'Neill and Weeks (2019).

[14] We consider 59 binary indicators constructed with available 57 pre-treatment covariates and three age-group indicators. The binary indicators are defined in such a way that they split the sample into two groups of roughly similar size. For example, the indicator associated to perception of illness control takes the value of 1 if the PI3 item score is at least 0.511, and zero otherwise. For the full set of indicators, see Table 4.

(5) $Y_i = Z_i'\delta + \tau W_i + W_i Z_i'\gamma + B_i'\beta + \varepsilon_i$

We estimate the coefficients γ_k more parsimoniously with the modified covariate approach (Tian et al. 2104):

(6) $Y_i = \tau W_i^* + \frac{1}{2} W_i^* Z_i'\gamma + B_i'\beta + \varepsilon_i$

where $W_i^* = 2W_i - 1$. The estimation of CATE with model (6) has the additional advantage over model (5) in that it does not impose linearity of the conditional average outcome of the control group.

Column (1) in Table 4 presents the estimated differences in 59 group effects with model (6) for the primary outcome; the fraction of time spent in employment 69 to 81 weeks after randomization. This analysis reveals the presence of heterogeneous impacts of the intervention on employment rates. We reject the null for all zero group effect differences for employment (F-statistic = 552, *p*-value = 0.00). Furthermore, we reject the null for zero effect differences among groups with different levels of control beliefs (F-statistic = 2.35, *p*-value = 0.03). We highlight two findings from Table 4. First, the pre-treatment level of self-efficacy in physical discomfort strongly predicts differences in the employment effect of LTJS. The sign of the estimated coefficient indicates that the LTJS program has a 33-percentage-point larger impact on the employment of individuals with low self-efficacy in physical discomfort (individuals who reported no or only a little confidence that they can keep the physical discomfort related to their health condition or disability from interfering with the things they want to do) compared to individuals with higher self-efficacy in physical discomfort (individuals who reported average, much, or total confidence in the controllability of their physical discomfort).

Table 4 also shows that the LTJS program has significantly larger employment effects for individuals with low physical impairment (estimated difference in subgroups effect of –26 percent), individuals who apply behavioral self-management techniques to a larger extent (estimated coefficient of 23 percent), individuals with high levels of autonomy motivation (estimated difference in group effects of 18 percent), and individuals who feel they understand their illness better (estimated difference in group effects of 16 percent).

The estimated differences in group effects shown in Table 4 will properly reveal underlying heterogeneity if program effects depend in a linear way on few covariates. If that is the case, then the reported results strongly suggest that the effectiveness of LTJS crucially depends on a single individual characteristic: self-efficacy in physical discomfort. However, if more complex patterns drive effect heterogeneity, regression results could be biased and thus misleading.

We now estimate CATE with the causal forest (CF) method (Wager and Athey 2018; Athey, Tibshirani, and Wager 2019).[15] CF analysis allows us to uncover underlying heterogeneity in causal effects based on minimum assumptions on the role of a large number of covariates available in our LTJS baseline questionnaire. An important feature of CF analysis is that, in contrast to quantile treatment-effect analysis, it can detect heterogeneous effects even for treatments that result in changes in individuals' rank in the distribution of outcomes.[16]

The CF is comprised of thousands of recursive partitions of the sample (causal trees) that, given a randomly chosen subset of covariates, maximize the heterogeneity of LTJS effect (Athey

[15] Other methods for nonparametric estimation of heterogeneous effects include nearest neighbor and series estimation (Crump et al. 2008; Lee 2009; and Wilke et al. 2012). We chose CF over these methods because CF is asymptotically Gaussian (Wager and Athey 2018), has good finite sample properties (Knaus et al. 2018), and is better suited to datasets that include large numbers of covariates.

[16] The self-management course will not preserve individual rank in the outcome distribution if the course has a much higher impact on individuals with negative control beliefs (low employment without intervention) than on individuals with positive control beliefs (high employment without intervention).

and Imbens 2016).[17] A desirable feature of CF is that one part of the sample is used for constructing causal trees, and the second part is used for estimating group effects for specified causal trees. Sample splitting allows estimates of group effects as if partitions of the population were exogenously given (Athey and Imbens 2016). CF effect estimates are consistent with true effects and are asymptotically normally distributed (Wage and Athey 2018).

We set the number of causal trees to 25,000, and tune the parameters by cross-validation.[18] The algorithm used to predict effects with CF analysis is summarized in the Online Appendix.[19] Column (2) in Table 4 reports the importance of each binary indicator for the CATE on employment. The information in this table summarizes how often each variable is chosen by the CF algorithm. As all covariates are binary indicators, we can directly assess the importance of the variable with the proportion of splits and their depth in the tree (O'Neill and Weeks 2019). CF results on the key determinant of effect heterogeneity are consistent with the results obtained with linear regression models (column 1, Table 4). Self-efficacy in physical discomfort is the most important driver of employment effect heterogeneity with a much higher weight in the CF than other variables (CF weight = 0.15). In addition, the CF analysis reveals that a variable detected as important by our regression analysis, "behavioral self-management," is the second

[17] For an example of a causal tree chosen by our CF analysis of employment effects, see Figure A6 in the Online Appendix.
[18] We apply the R package grf. Parameters tuned by cross-validation are fraction of the data used to build each tree; number of variables tried for each split; minimum number of observations in each tree leaf; fraction of data used to build the tree used for determining splits; the parameter that controls the maximum imbalance of a split; and the parameter that controls how harshly imbalanced splits are penalized.
[19] For an example of a causal tree specified by our CF analysis of post-program employment, see Figure A6 in the Online Appendix.

most important determinant for employment effect (CF weight = 0.09). The remaining variables, including those detected with model (6), are less relevant for the employment effect.

[Table 4]

Figure 2 shows the distribution of the effects on employment 69 to 81 weeks after program assignment predicted with CF. The distribution of employment effects is bimodal with the two modes located around the first and third quartile of the distribution. Given that the two most important characteristics for the employment effects, self-efficacy in physical discomfort and behavioral self-management are quite independent from one another (correlation = 0.05), these pronounced peaks strongly suggest the presence of two different groups in terms of program efficacy: one group of control-pessimistic workers for whom the intervention is beneficial, and another group of optimistic workers for whom the course results in lower performance in terms of employment. Additional variation in employment effects for these groups most likely arise because of small differences in their levels of controls beliefs and behavioral self-management activity.

[Figure 2]

We now compare the group effect for worker splits according to their predicted employment effect with the following model:

(7) $Y_i = G_i \cdot \{\tau_+ W_i + B_i' \beta_+\} + (1 - G_i) \cdot \{\tau_- W_i + B_i' \beta_-\} + \varepsilon_i$

where G_i is a dummy variable representing a splitting rule (e.g., low-level control beliefs), τ_+ is the ITT effect of self-management support on the targeted group ($G_i = 1$), and τ_- is the ITT effect of the course on the residual group ($G_i = 0$). In Table 5, we compare estimates of post-program effects on employment, first for individuals with high CF expected effects (CF predicted employment effect above the median), and then for individuals with low control beliefs (low levels of self-efficacy in physical discomfort).

Row 1 in Table 5 shows that for the group of workers with high expected effects, the program increases regular employment in weeks 69–81 by 14.02 percentage points (71.4-percent increase compared to control group). Row 2 in Table 5 shows the opposite estimates for the other half of the sample with predicted employment effect below the median. The LTJS intervention has a negative but statistically insignificant effect on employment for this group.

The implementation of targeting rules based on CF results require that employment services collect information on a wide range of covariates and compute predicted program efficacy at the individual level with CF. Given that heterogeneity of LTJS effects are mainly driven by individuals' self-efficacy in physical discomfort, a much simpler recruitment method is to target sick-listed workers according to their control beliefs. Panel B in Table 5 show that the differences in estimated effects for groups of workers with low and high control beliefs (measured in terms of self-efficacy in physical discomfort) are consistent with the results obtained in our regression analysis of heterogeneity effects (see Table 4). Row (3) in Table 5 shows that for workers who reported no or a little confidence in controlling physical discomfort, the intervention increased employment in weeks 69–81 by 18.54 percentage points (110-percent increase with respect to the control group).

Row (4) of Table 5 shows that for workers who reported average, much, or total confidence in keeping physical discomfort under control, the LTJS course reduced regular employment by 13.42 percentage points (29 percent reduction compared to control group).

[Table 5]

C. Mechanisms

A major question in the literature is whether workers' control beliefs can be affected in a way that leads to lasting impacts on their employment. There is simply no evidence on such mechanism (Almlund et al. 2011; Gutman and Schoon 2013; Cobb-Clark 2015), although there are several channels through which they could unfold. We examine this issue in this section. Workers with low control beliefs can experience prolonged absences because their beliefs are low *ex ante* (Buddelmeyer and Powdthavee 2016; Schurer 2017), or they can experience a reduction in control beliefs during their absence (Preuss and Hennecke 2018; Marsaudon 2019). The intervention program may increase employment of sick-listed workers with such pessimistic control beliefs by enhancing control beliefs. More confidence in the controllability of sickness absence may increase individuals' health investment (Schurer 2017), improving their health status and increasing their motivation to return to work (Bandura 1986), and thereby further improving employment outcomes in the longer term. Second, the enhancement of control beliefs may directly reduce the perceived consequences of illness on long-term health outcomes (Buddelmeyer and Powdthavee 2016), advancing return to work of persons who would otherwise remain on sick leave.

Table 6 reports standardized percentage bias for a select group of pre-intervention characteristics for the partition of long-term sick-listed workers that maximize the heterogeneity in post-program employment effects. The table presents between-group differences with highest versus lowest predicted effects on employment in weeks 69 to 81 (columns 1 and 2), and the groups with low versus high self-efficacy in physical discomfort (column 3). Results across partitions clearly show huge differences in control beliefs between individuals who benefit the least and the most in terms of post-program return to employment. It is quite clear that individuals who benefit most from self-management are individuals who had negative control beliefs before program participation. There is a huge difference in “self-efficacy in physical discomfort” among workers at the bottom of the distribution of predicted employment effects and those at the top of the distribution, consistent with this variable being the key driver of heterogeneous employment effects. Furthermore, individuals with the most positive employment effects are also more pessimistic regarding their re-employment prospects. Similarly, self-efficacy (perceived control) in unexpected events, social interaction, at the center of one’s life, and new illness issues is much lower for program participants in the top quartile than the group in the bottom quartile.

[Table 6]

Tables 7 and 8 present estimates of group effects on a range of outcomes potentially mediating employment effects, first for control beliefs, the primary mechanism, and then secondary mediating outcomes (behavioral self-management, expected return to work, perception of illness, and motivation for work). We measure these mediating outcomes both at the first and second follow-ups to detect eventual sustained changes in outcomes.

In Panel A of Table 7, the first two rows compare program effects on control beliefs for individuals in the top part of predicted impacts on employment during weeks 69 to 81 (labeled “high expected effect”), with impacts for individuals in the bottom part (50-percent group) of predicted employment effects (designated “low expected effect”). The first row shows that the top 50 percent group experience important increases in their control beliefs measured immediately after participation in the program. The program increases self-efficacy in sickness absence by 1.94 points (19.5-percent increase compared to control group) and has a substantial positive effect on perception of personal control of illness, which increases by 1.48 points (58.3-percent increase compared to control group). Row (2) in Table 7 shows that control beliefs of the bottom 50-percent group is not affected by the program. These results are consistent with the positive role of improved control beliefs for employment. In addition, despite the negative impact of LTJS on employment of the bottom 50-percent group, exposure to the intervention does not reduce control beliefs among people less vulnerable in terms of their perception of controllability of sickness absence and illness.

These results confirm that the recruitment of program participants mainly based on the length of their sickness spell and on expected long-term health problems failed in the sense that it assigned a group to the LTJS—the bottom 50 percent—for which the program did not enhance control beliefs.

Rows 3 and 4 in Panel A of Table 7 reveal striking similarities in terms of program impacts on control beliefs, when a much simpler targeting rule is applied to program participants. For the group of workers with low self-efficacy in physical discomfort, the program increases their self-efficacy in sickness absence by 1.73 points (17.5-percent increase relative to control group) and increases their perception of personal control of illness by 1.43 points (58.6-percent increase compared to control group). Row 4 shows again that the intervention does not change control beliefs among workers with pre-intervention high levels of self-efficacy in discomfort.

Panel B in Table 7 reports the estimated impacts of LTJS on control beliefs six months after participation in the program. Interestingly, though the estimates of mediating outcomes are more strongly affected by attrition than the estimated effects measured at the first follow-up, the results suggest that the LTJS program enhances individuals' perception of personal control of illness at least for six months after the program. The program increased perceived control of illness among workers with low self-efficacy in physical discomfort by 1.02 points (31.5 percent increase relative to control group), and increased the same measure of control beliefs by 1.4 points (39.6-percent increase compared to control group).

Overall Table 7 confirms that immediately after the program, the program benefits the more control-pessimistic group of long-term, sick-listed workers by enhancing their control beliefs, without altering the control beliefs of control-optimistic workers. Moreover, enhancement of control beliefs persists for at least six months after the intervention, suggesting that LTJS course might have longer-term impacts on control beliefs. However, as control-optimistic workers exposed to the psychological intervention do not change their control beliefs, results reported in Table 7 leave unexplained the source of negative employment effects among control-optimistic workers.

[Table 7]

Table 8 reports effects of the program via four alternative mediating outcomes. As the positive effects on control beliefs can be the artifact of correlated effects, we investigate the impact of the program on three outcomes strongly correlated with control beliefs (expected return to work, perception of illness, and motivation for work). We use motivation for work as a proxy for the individual's employment preferences.

In addition, as behavioral self-management activity is one of the main topics addressed by the program, and the employment effects decrease with low pre-intervention levels of behavioral self-management, we examine whether the negative employment effects of LTJS on control-optimistic workers are associated with changes in psychological coping strategies.

Panel A in Table 8 displays our estimates of first follow-up impacts on secondary mediating outcomes, and Panel B reports estimated effects measured 26 weeks after program completion. Results clearly show that the increase in control beliefs taking place post-program cannot be attributed to improvements in workers' baseline beliefs (measured in terms of expected return to work or in terms of the perceived impact of illness). First, the LTJS course increased control beliefs of vulnerable workers without matched changes in re-employment prospects. Second, the results show that the LTJS program increased individuals' motivation for work, but such increases are modest (roughly 10-percent increase relative to control group) compared with effects on control beliefs and are not statistically significant.

Overall, Table 8 confirms that the intervention's strong positive impact on individuals' perception of personal illness control cannot be attributed to improvements in baseline beliefs or preferences for work proxied by motivation for work.

We now turn our attention to program impacts on secondary mediating outcomes for control-optimistic workers. Panel A in Table 8 indicates that for groups with low expected effect and high control beliefs, the intervention leads them to perceive the consequences of their illness as more severe than if they had not participated in the course. Rows 2 and 4 in (column 3 show that LTJS program increases perceived illness by 5.16 points (12.4-percent increase with respect to control group) and by 6.3 points (15.1-percent increase relative to control group), respectively.

These results imply that the intervention leads control-optimistic workers toward a behavior consistent with a deterioration of perceived health. Self-management support increases psychological coping activities among control-optimistic workers. The program increases behavioral self-management by 0.47 points among the low expected effect group (26.8-percent increase compared to control group); and by 0.56 points in the high control beliefs group (31.0-percent increase with respect to control group). Results reported in Panel B show that such more intensive psychological coping activities eventually die out.

Overall, the results presented in Table 8 for less vulnerable sick-listed workers indicate that the intervention deteriorates their perception of health and leads these workers to intensify psychological coping. Despite such increased effort, the negative employment effects for this group suggest that the control-enhancing program leads control-optimistic workers away from their correct beliefs, and such biased beliefs reduce their participation in the labor market.[20]

[Table 8]

It seems crucial in our particular context to understand how individuals with heterogeneous beliefs interacting with each other during course participation (i.e., peer effects) affect return to work. A major mechanism of the program may operate through exposure to different control beliefs. Such a phenomenon would be consistent with both positive effects for control-pessimistic workers and *vice versa*. The empirical findings open the possibility for improving program effectiveness by targeting it in a more restrictive way. Analysis of heterogeneity of program impacts points to the possibility of targeting individuals with low control beliefs. However, if peer effects are responsible for the observed heterogeneity effect, targeting would "throw the baby out with the bath water."

In other words, is social learning (from control-optimistic participants) necessary for enhancing control beliefs among control-pessimistic workers? Social learning refers to the idea that what others do provides relevant information for one's own choices, particularly when people are uncertain about the consequences of their health investment (Bikhchandani, Hirshleifer, and Welch 1992). We address this issue with an analysis of the role of course participants with low expectations for post-program employment effects.

[20] See Heidhues, Kőszegi, and Strack (2018) for a theoretical analysis and other examples for self-defeating learning.

As program assignment is randomized and does not depend on course peers, group composition is also random. We therefore identify the influences of heterogeneous shares of course peers by means of the following model:

(8) $Y_i = G_i\{\tau_+ W_i + \vartheta_+ W_i P_i + \theta_+ P_i + B_i'\beta_+\} + (1 - G_i)\{\tau_- W_i + \vartheta_- W_i P_i + \theta_- P_i + B_i'\beta_-\} + \varepsilon_i$

where P_i is the share of course peers with low expected employment effect. We exclude from the analysis sample three randomized blocks (8, 34, and 42) with fewer than five individuals in the treatment. Table 9 reports peer effects for two different specifications. Column 1 in Table 9 reports results for subgroups of workers according to their expected effect, where P_i indicates the share of peers with expected employment effect below the median. Results show that a higher presence of "bad" peers has a positive but statistically insignificant impact on the group of individuals with a high expected employment effect. As the peer effects are estimated quite imprecisely, we cannot discard that excluding individuals with low predicted program effects leads to lower employment effects among vulnerable sick-listed workers. The bottom row of column 2 demonstrates that the presence "bad" peers has the opposite impact on the employment of the group of workers with low expected effect. In this case, the higher presence of "bad" peers lead to less vulnerable towards lower performance in terms of employment. The estimated coefficient for the peer effect is –0.27 and is statistically significant at the 0.05 level.

Column 2 in Table 9 examines the simple splitting rule based on self-efficacy in physical discomfort. The estimated marginal effect of an increase in the share of "bad" peers on control-pessimistic workers is statistically insignificant and of modest size. These results suggest that targeting program participation to workers with low control beliefs is most likely to have a more moderate impact on the effectiveness of the program than doing so based on CF predicted effects.

V. Conclusion

Using a psychological intervention in the form of a self-management support course for Danish workers, this paper systematically examines the role of control beliefs in prolonged sickness absence. We present first experimental evidence that a control-enhancing intervention leads to better labor market outcomes (see Almlund et al. 2011; Gutman and Schoon 2013; Cobb-Clark 2015), and that control beliefs constitute an important factor for the return to work of sick-listed workers.

While the intervention has zero overall effect on return to work and control beliefs, subgroup analyses show that average outcomes hide substantial heterogeneity. Our regression and machine learning analyses, relying on a randomized trial as well as the collection of very rich data on sickness absence, detect that the effectiveness of the control-enhancing course strongly depends on the individual's self-efficacy in physical discomfort. Namely, for such control-pessimistic workers, we find large positive effects on return to work, while for more control-optimistic workers we find significant negative effects on return to work.

We show that the most plausible mechanism of the positive employment effect is an increase in individuals' control beliefs. Furthermore, we show that the psychological intervention leads optimistic workers to perceive their illness as more severe than if they had not participated in the program. The analysis of peer effects suggests that targeting control-pessimistic workers still leads to positive effects for them. Peer effects results also suggest that negative effects on control-optimistic workers may be reduced if they were grouped with more pessimistic workers. This final point warrants further investigation. For now, the results clearly support only further targeting of the intervention on control-pessimistic workers.

References

Ahn, T. 2015. "Locus of control and job turnover." *Economic Inquiry*, 53(2), 1350–1365.

Ajzen, I. 2002. "Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior." *Journal of Applied Social Psychology*, 32(4), 665–683.

Almlund, M., Duckworth, A. L., Heckman, J., and T. Kautz. 2011. "Personality psychology and economics." In *Handbook of the Economics of Education* (Vol. 4, pp. 1–181). Elsevier.

Amtmann, D., Bamer, A. M., Cook, K. F., Askew, R. L., Noonan, V. K., and J. A. Brockway. 2012. "University of Washington self-efficacy scale: A new self-efficacy scale for people with disabilities." *Archives of Physical Medicine and Rehabilitation*, 93(10), 1757–1765.

Andrén D., and M. Svensson. 2012. "Part-Time Sick Leave as a Treatment Method for Individuals with Musculoskeletal Disorders." *Journal of Occupational Rehabilitation*, 22(3): 418–426.

Arrow, K. 1963. "Uncertainty and the Welfare Economics of Medical Care." *The American Economic Review*, 53(5): 141–149.

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

——. 2017. "The Econometrics of Randomized Experiment." In *Handbook of Economic Field Experiments* (Vol. 1, pp. 73–140). North-Holland.

Athey, S., Tibshirani, J., and S. Wager. 2019. "Generalized random forests." *The Annals of Statistics*, 47(2), 1148–1178.

Babcock, L., Congdon, W. J., Katz, L. F., and S. Mullainathan. 2012. "Notes on behavioral economics and labor market policy." *IZA Journal of Labor Policy*, 1(1), 2.

Bandura, A. 1977. “Self-efficacy: toward a unifying theory of behavioral change.” *Psychological Review*, *84*(2), 191.

——. 1986. “The explanatory and predictive scope of self-efficacy theory.” *Journal of Social and Clinical Psychology*, 4(3), 359–373.

——. 1989. “Regulation of cognitive processes through perceived self-efficacy.” *Developmental psychology*, 25(5), 729.

Benjamin, D. J., and J. M. Shapiro. 2005. “Does Cognitive Ability Reduce Psychological Bias?” *Retrieved May*, 16, 2007.

Belloni, A., Chernozhukov, V., and C. Hansen. 2014. “Inference on treatment effects after selection amongst high-dimensional controls.” *Review of Economic Studies*, 81(2), 608–650.

Bertrand, M., Crépon, B., Marguerie, A., and P. Premand. 2017. “Contemporaneous and Post-Program Impacts of a Public Works Program: Evidence from Côte d'Ivoire.” World Bank.

Bikhchandani, S., Hirshleifer, D., and I. Welch. 1992. “A theory of fads, fashion, custom, and cultural change as informational cascades." *Journal of Political Economy*, 100(5), 992–1026.

Bjorner, J. B., Damsgaard, M. T., Watt, T., and M. Groenvold. 1998. “Tests of data quality, scaling assumptions, and reliability of the Danish SF-36.” *Journal of Clinical Epidemiology*, 51(11), 1001–1011.

Borghans, L., Duckworth, A. L., Heckman, J. J., and B. Ter Weel. 2008. “The economics and psychology of personality traits.” *Journal of Human Resources*, 43(4), 972–1059.

Broadbent, E., Petrie, K. J., Main, J., and J. Weinman. 2006. “The brief illness perception questionnaire.” *Journal of Psychosomatic Research*, *60*(6), 631–637.

Bryan, G., Karlan, D., and S. Nelson. 2010. "Commitment devices." *Annu. Rev. Econ.*, 2(1), 671–698.

Buddelmeyer, H., and N. Powdthavee. 2016. "Can having internal locus of control insure against negative shocks? Psychological evidence from panel data." *Journal of Economic Behavior & Organization*, *122*, 88–109.

Bültmann, U., Rugulies, R., Lund, T., Christensen, K. B., Labriola, M., and H. Burr. 2006. "Depressive symptoms and the risk of long-term sickness absence." *Social Psychiatry and Psychiatric Epidemiology*, *41*(11), 875–880.

Caliendo, M., Cobb-Clark, D. A., and A. Uhlendorff. 2015. "Locus of control and job search strategies." *Review of Economics and Statistics*, *97*(1), 88–103.

Carneiro, P., Lee, S., and D. Wilhelm. 2016. "Optimal data collection for randomized control trials." *arXiv preprint arXiv:1603.03675*.

Chandra, A., Handel, B., and J. Schwartzstein. 2019. "Behavioral economics and health-care markets." In *Handbook of Behavioral Economics: Applications and Foundations* 1 (Vol. 2, pp. 459–502). North-Holland.

Chiteji, N. 2010. "Time preference, noncognitive skills and well being across the life course: do noncognitive skills encourage healthy behavior?." *American Economic Review*, 100(2), 200–204.

Christensen, K. S., Fink, P., Toft, T., Frostholm, L., Ørnbøl, E., and F. Olesen. 2005. "A brief case-finding questionnaire for common mental disorders: the CMDQ." *Family Practice*, *22*(4), 448–457.

Christensen, K. B., Lund, T., Labriola, M., Bültmann, U., and E. Villadsen. 2007. “The impact of health behaviour on long term sickness absence: results from DWECS/DREAM.” *Industrial health*, *45*(2), 348–351.

Carneiro, P., Lee, S. S., and D. Wilhelm. 2019. “Optimal Data Collection for Randomized Control Trials” (No. CWP21/19) Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Cobb-Clark, D. A. 2015. “Locus of control and the labor market.” *IZA Journal of Labor Economics*, 4(1), 3.

Cobb-Clark, D. A., Kassenboehmer, S. C., and S. Schurer. 2014. “Healthy habits: The connection between diet, exercise, and locus of control.” *Journal of Economic Behavior & Organization*, 98, 1–28.

Cobb-Clark, D. A., Kassenboehmer, S. C., and M. Sinning. G. 2016. “Locus of control and savings.” *Journal of Banking & Finance*, 73, 113–130.

Cobb-Clark D. A., and S. Schurer. 2012. “The stability of big-five personality traits." *Economic Letters* 115(1):11–15.

——. 2013. “Two economists’ musings on the stability of locus of control.” *The Economic Journal*, 123(570), F358–F400.

Cobb-Clark, D. A., and M. Tan. 2011. “Noncognitive skills, occupational attainment, and relative wages.” *Labour Economics*, 18(1), 1–13.

Coleman, M., and T. DeLeire. 2003. “An economic model of locus of control and the human capital investment decision.” *Journal of Human Resources*, 38(3), 701–721.

Crump, R. K., Hotz, V. J., Imbens, G. W., and O. A. Mitnik. 2008. “Nonparametric tests for treatment effect heterogeneity.” *The Review of Economics and Statistics*, 90(3), 389–405.

Davis, J., and S. B. Heller. 2017. “Using causal forests to predict treatment heterogeneity: An application to summer jobs.” *American Economic Review*, 107(5), 546–50.

Dekkers-Sánchez, P. M., Hoving, J. L., Sluiter, J. K., and M. H. Frings-Dresen. 2008. “Factors associated with long-term sick leave in sick-listed employees: a systematic review.” *Occupational and environmental medicine*, 65(3), 153–157.

Dionne, G., and P. St-Michel. 1991. “Workers' compensation and moral hazard.” *The Review of Economics and Statistics*, 236–244.

Falk, A., and F. Zimmermann. 2018. “Information processing and commitment." *The Economic Journal*, 128(613), 1983–2002.

Fevang, E., Markussen, S., and K. Røed. 2014. ”The sick pay trap.” *Journal of Labor Economics*, 32(2), 305–336.

Franek, J. 2013. “Self-management support interventions for persons with chronic disease: an evidence-based analysis.” *Ontario Health technology Assessment Series*, 13(9), 1.

Frank, R. G. 2004. “Behavioral economics and health economics." No. w10881. National Bureau of Economic Research.

Frölich M., Heshmati A., and M. Lechner. 2004. “A Microeconometric Evaluation of Rehabilitation of Long-term Sickness in Sweden." *Journal of Applied Econometrics*, 19(3): 375–396.

Gallo, W. T., Endrass, J., Bradley, E. H., Hell, D., and S. V. Kasl. 2003. “The influence of internal control on the employment status of German workers.” *Schmollers Jahrbuch*, *123*(1), 71–81.

Gjesdal, S., and Bratberg, E. 2003. "Diagnosis and duration of sickness absence as predictors for disability pension: results from a three-year, multi-register based and prospective study." *Scandinavian Journal of Public Health*, *31*(4), 246–254.

Gjesdal, S., Ringdal, P. R., Haug, K., and J. G. Mæland. 2004. "Predictors of disability pension in long-term sickness absence: results from a population-based and prospective study in Norway 1994–1999." *The European Journal of Public Health*, *14*(4), 398–405.

Gottschalk, P. 2005. "Can work alter welfare recipients' beliefs?." *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, 24(3), 485–498.

Gutman, L. M., and I. Schoon. 2013. "The impact of non-cognitive skills on outcomes for young people." *Education Endowment Foundation*, 59(22.2), 2019.

Hansen, J. 2000. "The effect of work absence on wages and wage gaps in Sweden."
Journal of Population Economics, 13, 45–55.

Heckman, J. J., and T. Kautz. 2012. "Hard evidence on soft skills." *Labour economics*, 19(4), 451–464.

Heckman, J. J., Stixrud, J., and S. Urzua. 2006. "The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior." *Journal of Labor economics*, 24(3), 411–482.

Heidhues, P., Kőszegi, B., and P. Strack. 2018. "Unrealistic expectations and misguided learning." *Econometrica*, 86(4), 1159–1214.

Heineck, G., and S. Anger. 2010. "The returns to cognitive abilities and personality traits in Germany." *Labour Economics*, *17*(3), 535–546.

Henderson, M., Glozier, N., and K. H. Elliott. 2005. “Long term sickness absence: Is caused by common conditions and needs managing.” *BMJ: British Medical Journal*, *330*(7495), 802.

Henrekson, M., and M. Persson. 2004. “The effects on sick leave of changes in the sickness insurance system.” *Journal of Labor Economics*, 22(1), 87–113.

Hesselius, P., and M. Persson. 2007. “Incentive and spill-over effects of supplementary sickness compensation." IFAU-Institute for Evaluation of Labour Market and Education Policy Working Paper No. 2007:16.

Heymann, J., Rho, H. J., Schmitt, J., and A. Earle. 2010. “Ensuring a healthy and productive workforce: comparing the generosity of paid sick day and sick leave policies in 22 countries.” *International Journal of Health Services*, *40*(1), 1–22.

Holm, A., Høgelund, J., Gørtz, M., Rasmussen, K. S., and H. S. B. Houlberg. 2017. “Employment effects of active labor market programs for sick-listed workers.” *Journal of Health Economics*, 52, 33–44.

Huijs, J. J., Koppes, L. L., Taris, T. W., and R. W. Blonk. 2012. “Differences in predictors of return to work among long-term sick-listed employees with different self-reported reasons for sick leave.” *Journal of Occupational Rehabilitation*, 22(3), 301–311.

Høgelund J., Holm A., and L. Eplov. 2012. “The Effect of Part-time Sick Leave for Employees with Mental Disorders." *Journal of Mental Health Policy and Economics*, 15(4): 157–170.

Ichino, A., and E. Moretti. 2009. “Biological gender differences, absenteeism, and the earnings gap.” *American Economic Journal: Applied Economics*, 1(1), 183–218.

Ichino, A., and R. T. Riphahn. 2005. “The effect of employment protection on worker effort: Absenteeism during and after probation.” *Journal of the European Economic Association*, 3(1), 120–143.

Kivimäki, M., Forma, P., Wikström, J., Halmeenmäki, T., Pentti, J., Elovainio, M., and J. Vahtera. 2004. “Sickness absence as a risk marker of future disability pension: the 10-town study." *Journal of Epidemiology & Community Health*, 58(8), 710–711.

Knaus, M. C., Lechner, M., and A. Strittmatter. 2018. “Machine Learning Estimation of Heterogeneous Causal Effects: Empirical Monte Carlo Evidence”. No. 13402 CEPR Discussion Paper.

Kools, L., and P. Koning. 2019. “Graded return-to-work as a stepping stone to full work resumption." *Journal of Health Economics*, 65, 189–209.

Johansson, P., and M. Palme. 1996. “Do economic incentives affect work absence? Empirical evidence using Swedish micro data.” *Journal of Public Economics*, 59(2), 195–218.

——. 2002. “Assessing the effect of public policy on worker absenteeism.” *Journal of Human Resources*, 381–409.

——. 2005. “Moral hazard and sickness insurance.” *Journal of Public Economics*, 89(9–10), 1879–1890.

Lee, M.-J. 2009. “Nonparametric Tests for Distributional Treatment Effect for Randomly Censored Responses.” *Journal of the Royal Statistical Society, Series* B, 71, 243–264.

Lekfuangfu, W. N., Powdthavee, N., Warrinnier, N., and F. Cornaglia. 2018. “Locus of control and its intergenerational implications for early childhood skill formation.” *The Economic Journal*, 128(608), 298–329.

Levesque, C. S., Williams, G. C., Elliot, D., Pickering, M. A., Bodenhamer, B., and P. J. Finley. 2007. “Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors.” *Health education research*, 22(5), 691–702.

Lorig K., Seleznick M., Lubeck D., Ung E., Chastain R. L., and H. R. Holman. 1989. "The beneficial outcomes of the arthritis self-management course are not adequately explained by behavior change." *Arthritis Rheum. Jan*; 32(1):91–5

Lorig K., Stewart A., Ritter P., González V., Laurent D., and J. Lynch. 1996 *Outcome Measures for Health Education and other Health Care Interventions*. Thousand Oaks CA: Sage Publications, 24–38.

Lorig, K. R., Sobel, D. S., Stewart, A. L., Brown Jr, B. W., Bandura, A., Ritter, P., ...and H. R. Holman. 1999. "Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization: a randomized trial." *Medical Care,* 5–14.

Lorig, K. R., Ritter, P., Stewart, A. L., Sobel, D. S., Brown Jr, B. W., Bandura, A., ... and H. R. Holman. 2001. "Chronic disease self-management program: 2-year health status and health care utilization outcomes." *Medical Care*, 1217–1223.

Malmendier, U., Tate, G., and J. Yan. 2011. "Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies." *The Journal of Finance*, 66(5), 1687–1733.

Markussen, S. 2012. "The individual cost of sick leave." *Journal of Population Economics*, 25(4), 1287–1306.

Marsaudon, A. 2019. "Do Health Shocks Modify Personality Traits? Evidence from Locus of Control." ffhalshs–01976868f.

McGee, A. D. 2015. "How the perception of control influences unemployed job search." *ILR Review*, *68*(1), 184–211.

McGee, A., and P. McGee. 2016. “Search, effort, and locus of control.” *Journal of Economic Behavior & Organization*, *126*, 89–101.

Ng, T. W., Sorensen, K. L., and L. T. Eby. 2006. “Locus of control at work: a meta-analysis.” *Journal of Organizational Behavior*: *The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 27(8), 1057–1087.

Noerholm, V., Groenvold, M., Watt, T., Bjorner, J. B., Rasmussen, N. A., and P. Bech. 2004. “Quality of life in the Danish general population–normative data and validity of WHOQOL-BREF using Rasch and item response theory models." *Quality of Life Research*, 13(2), 531–540.

Nordberg, M., and K. Røed. 2009. “Economic Incentives, Business Cycles, and Long-Term Sickness Absence.” *Industrial Relations: A Journal of Economy and Society*, 48(2), 203–230.

Offerhaus, J. 2013. “The type to train? Impacts of personality characteristics on further training participation. Impacts of Personality Characteristics on Further Training Participation.” (January 1, 2013). SOEPpaper, (531).

O'Neill, E., and M. Weeks. 2018. “Causal Tree Estimation of Heterogeneous Household Response to Time-Of-Use Electricity Pricing Schemes.2 arXiv preprint arXiv:1810.09179.

Pedersen, P., Søgaard, H. J., Labriola, M., Nohr, E. A., and C. Jensen. 2015. “Effectiveness of psychoeducation in reducing sickness absence and improving mental health in individuals at risk of having a mental disorder: a randomised controlled trial.” *BMC public health*, 15(1), 763.

Preuss, M., and J. Hennecke. 2018. “Biased by success and failure: How unemployment shapes locus of control.” *Labour Economics*, 53, 63–74.

Puhani, P. A., and K. Sonderhof. 2010. “The effects of a sick pay reform on absence and on health-related outcomes.” *Journal of Health Economics*, *29*(2), 285–302.

Rehwald, K., Rosholm, M., and B. Rouland. 2018. “Labour market effects of activating sick-listed workers.” *Labour Economics*, 53, 15–32.

Rosenbaum, M. 1980. “A schedule for assessing self-control behaviors: Preliminary findings.” *Behavior Therapy*, 11(1), 109–121.

Rotger, G. P. 2019. ”Evaluering af Lær at tackle job og sygdom: Et randomiseret kontrolleret studie af beskæftigelseseffekten.” VIVE rapport

Rotter, J. B. 1966. “Generalized expectancies for internal versus external control of reinforcement.” *Psychological monographs: General and applied*, *80*(1), 1.

Schurer, S. 2017. “Bouncing back from health shocks: Locus of control and labor supply.” *Journal of Economic Behavior & Organization*, *133*, 1–20.

Smith, F. J. 1977. “Work attitudes as predictors of attendance on a specific day.” *Journal of Applied Psychology*, *62*(1), 16.

Specht J., Egloff B., and S.C. Schmukle. 2011. “Stability and change of personality across the life course: the impact of age and major life events on mean-level and rank-order stability of the big five." *J Pers Soc Psychol* 101(4):862–882, doi:10.1037/a0024950

——. 2013. “Everything under control? The effects of age, gender, and education on trajectories of perceived control in a nationally representative German sample." *Dev Psychol* 49(2):353–364, doi:10.1037/a0028243

Spinnewijn, J. 2013. “Insurance and perceptions: how to screen optimists and pessimists.” *The Economic Journal*, 123(569), 606–633.

Spinnewijn, J. 2015. “Unemployed but optimistic: Optimal insurance design with biased beliefs.” *Journal of the European Economic Association*, *13*(1), 130‒167.

Stapelfeldt, C. M., Jensen, C., Andersen, N. T., Fleten, N., and C.V. Nielsen. 2012. “Validation of sick leave measures: self-reported sick leave and sickness benefit data from a Danish national register compared to multiple workplace-registered sick leave spells in a Danish municipality." *BMC Public Health*, 12(1), 661.

The Danish Committee for Health Education 2014. *Lær at tackle job og sygdom*.

Tian, L., Alizadeh, A. A., Gentles, A. J., and R. Tibshirani. 2014. “A simple method for estimating interactions between a treatment and a large number of covariates.” *Journal of the American Statistical Association*, 109(508), 1517‒1532.

Tipton, R. M., and E. L. Worthington 1984. “The measurement of generalized self-efficacy: a study of construct validity.” *Journal of Personality Assessment*.

Uhlendorff, A. 2004: “Der Einfluss von Persönlichkeitseigenschaften und sozialen Ressourcen auf die Arbeitslosigkeitsdauer,” *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 56, 279–303.

Waddell, G., Aylward, M., and P. Sawney. 2002. *Back pain, incapacity for work and social security benefits: an international literature review and analysis*. RSM Press.

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

Wallston, K. A., Strudler Wallston, B., and R. DeVellis. 1978. “Development of the multidimensional health locus of control (MHLC) scales." *Health Education Monographs*, 6(1), 160‒170.

Willke, R. J., Zheng, Z., Subedi, P., Althin, R., and C. D. Mullins. 2012, “From Concepts, Theory, and Evidence of Heterogeneity of Treatment Effects to Methodological Approaches: A Primer,” *BMC Medical Research Methodology*, 12, 185.

Ziebarth, N. R. 2013. “Long-term absenteeism and moral hazard—Evidence from a natural experiment.” *Labour Economics*, *24*, 277–292.

Ziebarth, N. R., and M. Karlsson. 2010. “A natural experiment on sick pay cuts, sickness absence, and labor costs.” *Journal of Public Economics*, *94*(11–12), 1108–1122.

——. 2014. “The effects of expanding the generosity of the statutory sickness insurance system.” *Journal of Applied Econometrics*, 29(2), 208–230.

Tables and Figures

Table 1

Means of Selected Individual Characteristics

Variable	Mean Impact Analysis Sample		Heterogeneity Analysis Sample		Mechanism Analysis Sample	
	C	T	C	T	C	T
Female (%)	75	77	75	77	78	77
Age (years)	47	47	48	47	48	47
Higher education (%)	49	50	47	51	44	41
Sickness transition from employment (%)	70	71	70	71	71	72
Sickness caused by work conditions (%)	60	60	59	59	57	60
Sickness caused by mental disorder (%)	51	46	50	45	49	46
Employment (%)	49	50	48	50	48	51
Sickness absence (%)	24	24	25	24	25	25
Baseline beliefs (0–20 scale)	4	4	4	4	5	4
Motivation for work (0–90 scale)	36	36	37	36	37	37
Perception of illness (0–80 scale)	32	32	31	31	31	32
Perceived personal control of illness (0–10 scale)	3	3	3	3	3	3
Self-efficacy in sickness absence (0–24 scale)	11	11	11	11	11	11
Self-efficacy in physical discomfort (0–4 scale)	2	2	2	2	2	2
Self-efficacy in unexpected events (0–4 scale)	2	2	2	2	2	2
Self-efficacy in social interaction (0–4 scale)	2	2	2	2	2	2
Self-efficacy at center of one’s life (0–4 scale)	2	2	2	2	2	2
Self-efficacy in frustration (0–4 scale)	2	2	2	2	2	2
Self-efficacy in new illness issues (0–4 scale)	2	2	2	2	2	2
Behavioral self-management (0–30 scale)	9	8	9	8	9	8
Group size	5	9	5	9	5	9
Observations	197	371	175	336	133	267

Source: DREAM, LTJS evaluation dataset, baseline LTJS questionnaire, and authors’ calculations. C = control; T = treatment.

Table 2

Effect of Self-management Support on Employment and Sickness Absence

	(1)	(2)	(3)
Panel A: Post-program Effect			
Sickness absence	0.0455	0.0445	0.0226
	(0.0435)	(0.0423)	(0.0415)
Control group mean	0.2948	0.2948	0.2912
Employment	–0.0037	–0.0022	0.0254
	(0.0293)	(0.0261)	(0.0235)
Control group mean	0.3116	0.3116	0.3132
Panel B: Initial Effect			
Sickness absence	0.0602*	0.0421*	0.0492*
	(0.0253)	(0.0208)	(0.0227)
Control group mean	0.6751	0.6738	0.6738
Employment	–0.0109	–0.0105	–0.0067
	(0.0126)	(0.0114)	(0.0114)
Control group mean	0.0480	0.0480	0.0483
Panel C: Total Effect			
Sickness absence	0.0616#	0.0608#	0.0374
	(0.0327)	(0.0318)	(0.0313)
Control group mean	0.4270	0.4270	0.4242
Employment	–0.0247	–0.0234	–0.0072
	(0.0247)	(0.0204)	(0.0195)
Control group mean	0.2341	0.2341	0.2353
Strata fixed effects and strong predictors	no	yes	yes
Unbalanced variables	no	no	yes
Observations	568	568	567

Notes: Robust standard errors clustered at block level in parentheses. Column 1 presents results from model (2) with B_i. Column 2 presents results from model (2) with B_i and X_i. The indicator *At least 60% of employment for two years before assignment* is included in the model of all outcomes, with the exception of model for initial effect on sickness absence, which includes the indicator *Expect at least 1 to 3 months before return to work*. Column 3 reports the estimated effects from model (2) with B_i, X_i, and the inclusion of physical impairment, expected long illness, and perceived effect of medical treatment. # $p < 0.100$; * $p < 0.050$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

Table 3

Effect of Self-management Support on Control Beliefs

	(1)	(2)	(3)
Panel A: First Follow-up			
Self-efficacy in sickness absence	0.3938 (0.5849)	0.6914 (0.4194)	0.8186# (0.4055)
Control group mean	18.0779	18.0779	18.0779
Observations	447	447	447
Perceived personal control of illness	0.4578# (0.2514)	0.4606# (0.2491)	0.5013* (0.2363)
Control group mean	3.1922	3.1922	3.1922
Observations	447	447	447
Panel B: Second Follow-up			
Self-efficacy in sickness absence	–0.9469 (0.6010)	–0.5974 (0.4875)	–0.3899 (0.4682)
Control group mean	19.3611	19.3611	19.3611
Observations	416	416	416
Perceived personal control of illness	0.2962 (0.2676)	0.2713 (0.2717)	0.3433 (0.2691)
Control group mean	3.6452	3.6452	3.6452
Observations	398	398	398
Strata fixed effects and strong predictors	No	Yes	Yes
Unbalanced variables	No	No	Yes

Notes: Robust standard errors clustered at block level are in parentheses. Column 1 presents results from model (2) with B_i. Column 2 presents results from model (2) with B_i and X_i. The covariates measured before program assignment—self-efficacy in sickness absence and perception of personal control of illness—are included in the models for self-efficacy and perception of illness control, respectively. Column 3 reports estimated effects from model (2) with B_i, X_i, and the inclusion of physical impairment, expected long illness, and perceived effect of medical treatment. # $p < 0.100$; * $p < 0.050$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, baseline LTJS questionnaire, first follow-up LTJS questionnaire, and second follow-up LTJS questionnaire.

Table 4

Heterogeneity of the Effect of Self-management Support on Employment

Sample Split	Difference in group effects (1)	Weight in Causal Forest (2)
Sickness transition from employment	–0.09	0.01
Sickness caused by work conditions	0.04	0.01
Higher education	0.11	0.04
Sickness from private sector job	–0.11	0.01
Sickness from public sector job	0.04	0.01
Sickness caused by work conditions	0.01	0.01
Sickness caused by mental disorder	–0.06	0.01
Sickness caused by musculoskeletal	–0.02	0.01
Female	0.00	0.01
Age 20–39	0.08	0.02
Age 40–54	–0.05	0.02
Age 55–72	0.00	0.00
Cohabitation with partner	–0.06	0.02
Cohabitation with children	0.05	0.01
Previous long absence	0.08	0.05
Previous sickness absence	0.02	0.03
Multiple serious illnesses	0.15	0.00
Pain	–0.03	0.01
Fatigue	–0.04	0.01
Dyspnea	0.07	0.03
Sleep quality	–0.00	0.01
Physical activity	–0.10	0.02
Prescribed medicine	0.06	0.03
Contact with GP	0.02	0.01
Contact with psychologist	0.10	0.01
Contact with physiotherapist	0.12	0.01
Contact with emergency department	0.03	0.00
Somatic disorder	0.03	0.01
Illness worries	0.15	0.02
Anxiety	–0.13	0.00
Depression	0.09	0.01

(*continued*)

Table 4 (*continued*)

Sample Split	Difference in group effects (1)	Weight in Causal Forest (2)
Physical impairment	–0.26*	0.01
Perception of consequences	–0.03	0.01
Expected long illness	–0.12	0.01
Perceived personal control of illness	–0.06	0.01
Perceived medical treatment effectiveness	0.08	0.01
Perception of symptoms	–0.00	0.01
Illness concern	–0.03	0.01
Illness understanding	0.16*	0.03
Emotional response	–0.02	0.01
Autonomy motivation	0.18*	0.01
Introjected regulation	0.01	0.01
External regulation	0.04	0.02
Amotivation	0.08	0.04
Expected full return to work	0.08	0.00
Expected long sickness	–0.11	0.02
Expected similar job tasks	–0.02	0.02
Expected working time	–0.04	0.01
Work capacity	0.11	0.01
Life quality	–0.10	0.02
Health-related life quality	0.01	0.00
Well-being	0.11	0.00
Self-efficacy in physical discomfort	–0.33**	0.15
Self-efficacy in unexpected events	0.13	0.01
Self-efficacy in social interaction	–0.07	0.01
Self-efficacy at center of one's life	–0.00	0.01
Self-efficacy in frustration	–0.02	0.02
Self-efficacy in new illness issues	–0.07	0.04
Behavioral self-management	0.23*	0.09

Notes: Column 1 reports estimated difference in group effects with model (6) for regular employment during weeks 69–81. Column 2 reports variable importance corresponding to the causal forest predicted effect on regular employment in weeks 69–81. * $p < 0.050$; ** $p < 0.010$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

Table 5

Group Effects of Self-management Support on Employment by Targeting Rule

Group	Employment (1)
Panel A: Causal Forest Expected Effect	
High expected effect	0.1402* (0.0534)
Control group mean	0.1735
Low expected effect	–0.0564 (0.0586)
Control group mean	0.4633
Panel B: Control Beliefs	
Low control beliefs	0.1854*** (0.0468)
Control group mean	0.1678
High control beliefs	–0.1342* (0.0638)
Control group mean	0.4562
Observations	511

Notes: Columns 1 to 6 report group effects τ_+ and τ_- estimated with model (7) for targeting rules: CF predicted employment effect above the median (designated high CF expected effect and low CF expected effect); low self-efficacy in physical discomfort (labeled low control beliefs and high control beliefs). Robust standard errors clustered at block level appear in parentheses. * $p < 0.050$; *** $p < 0.001$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

Table 6

Standardized Percentage Bias in Selected Characteristics by Targeting Rule

	Group with expected effect above median versus group with expected effect below median (1)	Group with expected effect above 66th percentile versus group with expected effect below 33rd percentile (2)	Group with low self-efficacy in physical discomfort versus group with high self-efficacy in physical discomfort (3)
Age	–16.6	–25.2	–17.1
Female	–7.6	4.7	–1.8
Higher education	37.9	43.3	31.2
Sickness transition from employment	–22.7	–34.5	–10.1
Sickness work related	–26.8	–29.6	–18.1
Sickness caused by mental disorder	–53.1	–67.9	–56.4
Employment	–43.7	–46.8	–19.5
Sickness absence	32.3	40.1	13.9
Baseline beliefs	–71.1	–112.4	–52.1
Motivation for work	–8.4	–19.2	–11.8
Perception of illness	65.1	94	52.2
Perceived personal control of illness	–42.3	–55.4	–52.4
Self-efficacy in sickness absence	–121.8	–169.2	–138.6
Self-efficacy in physical discomfort	–176.7	–300.1	–327.4
Self-efficacy in unexpected events	–84.6	–126.5	–99.9
Self-efficacy in social interaction	–69.5	–92.3	–69.1
Self-efficacy at center of one's life	–75	–99.1	–76.7
Self-efficacy in frustration	–71.7	–93.9	–68
Self-efficacy in new issues	–85.2	–112.5	–87.3
Behavioral self-management	24.3	37.8	–9.6

Notes: Columns 1 to 3 report the standardized percentage bias for selected characteristics and for different targeting rules. Column 2 reports the standardized difference of means for individuals with CF predicted employment effect above the 66th percentile compared to individuals with CF predicted employment effect below the 33rd percentile.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and Baseline LTJS questionnaire.

Table 7

Group Effects of Self-management Support on Control Beliefs by Targeting Rule

Group	Self-efficacy in Sickness Absence (1)	Perceived Personal Control of Illness (2)
Panel A: First Follow-up		
High expected effect	1.9442* (0.8349)	1.4832*** (0.3963)
Low expected effect	0.7772 (1.1543)	–0.1813 (0.5402)
Low control beliefs	1.7270# (0.9484)	1.4257** (0.4060)
High control beliefs	0.1991 (1.1396)	–0.3343 (0.4343)
Observations	408	415
Panel B: Second Follow-up		
High expected effect	0.9771 (1.2989)	0.5917 (0.4796)
Low expected effect	–0.8938 (1.1075)	0.7301 (0.4702)
Low control beliefs	0.4205 (1.3784)	1.0214* (0.4830)
High control beliefs	–0.6604 (1.3415)	0.3382 (0.4812)
Observations	376	377

Notes: Columns (1) and (2) report group effects τ_+ and τ_- for outcomes self-efficacy in sickness absence, and perception of personal control of illness, respectively. Panel A reports estimated effects on these outcomes measured at the first follow-up, and Panel B reports estimated effects measured at the second follow-up. The group effects are estimated with model (7) for following targeting rules: CF predicted employment effect above the median and low self-efficacy in physical discomfort. Robust standard errors clustered at block level appear in parentheses. # $p < 0.100$; * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$.

Sources: Authors' calculations based on DREAM, LTJS evaluation dataset, baseline LTJS questionnaire, first follow-up LTJS questionnaire, and second follow-up LTJS questionnaire.

Table 8

Group Effects of Self-management Support on Secondary Mediated Outcomes by Targeting Rule

Group	Behavioral Self-management (1)	Expected Return to Work (2)	Perception of Illness (3)	Motivation for Work (4)
Panel A: First Follow-up				
High expected effect	0.1484 (0.2247)	1.0812 (0.9891)	–0.1450 (5.7505)	5.5390 (6.0738)
Low expected effect	0.4687** (0.1575)	–0.5628 (0.8111)	5.1568# (2.7762)	3.7844 (2.7594)
Low control beliefs	0.0709 (0.2366)	0.8811 (1.0779)	0.1033 (5.7338)	6.8088 (5.7219)
High control beliefs	0.5573*** (0.1282)	–0.7798 (0.8799)	6.3003* (2.9774)	3.7932 (2.8870)
Observations	408	400	415	405
Panel B: Second Follow-up				
High expected effect	0.0815 (0.2476)	0.7519 (1.2954)	1.4359 (5.4558)	5.5558 (6.4005)
Low expected effect	0.1697 (0.2075)	–0.7672 (0.8793)	3.3135 (2.3117)	1.2371 (3.1772)
Low control beliefs	0.0804 (0.2433)	0.9592 (1.2451)	1.3493 (5.4081)	5.9014 (6.1321)
High control beliefs	0.1959 (0.2142)	–0.7322 (1.1729)	3.4814 (2.9391)	2.1323 (3.3409)
Observations	373	367	377	371

Notes: Columns 1 to 4 report group effects τ_+ and τ_- for outcomes behavioral self-management, expected return to work, perception of illness, and motivation for work, respectively. Panel A reports estimated effects on these outcomes measured at the first follow-up, and Panel B reports estimated effects measured at the second follow-up. The group effects are estimated with model (7) for targeting rules: CF predicted employment effect above median and low self-efficacy in physical discomfort. Robust standard errors clustered at block level are in parentheses. # $p < 0.100$; * $p < 0.050$; ** $p < 0.010$; *** $p < 0.001$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, baseline LTJS questionnaire, first follow-up LTJS questionnaire, and second follow-up LTJS questionnaire.

Table 9

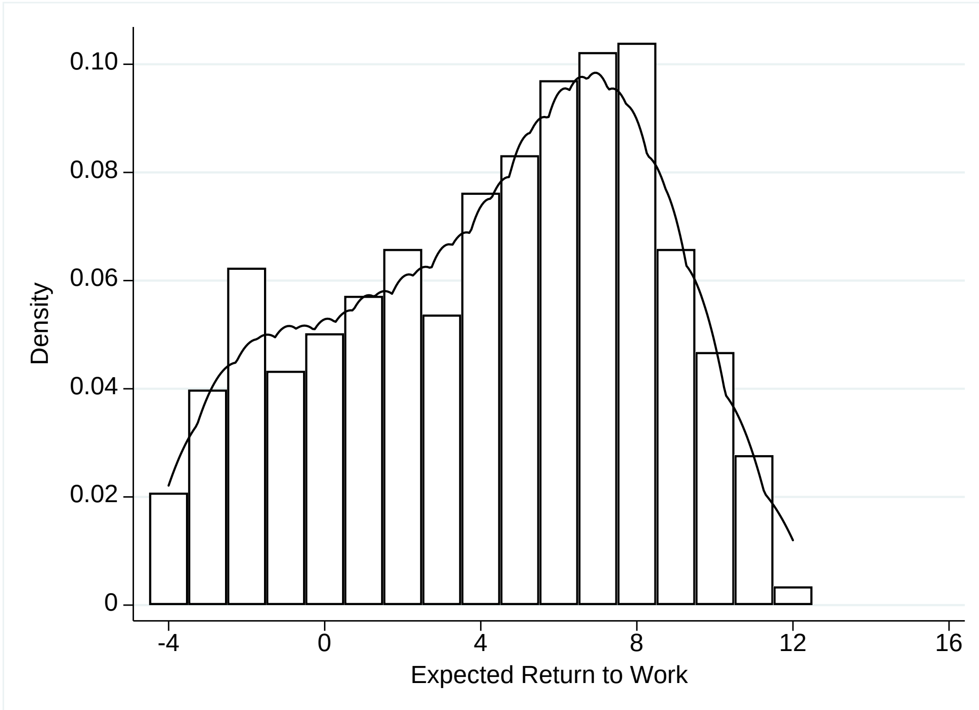
Peer Effects on Employment by Targeting Rule

Group	Causal Forest (1)	Control Beliefs (2)
High expected effect	0.00671 (0.0898)	0.1230 (0.0899)
× More than 50% peers have low expected effect	0.2038 (0.1515)	0.1527 (0.1507)
Low expected effect	0.0248 (0.0782)	–0.0619 (0.0911)
× More than 50% peers have low expected effect	–0.2699* (0.1278)	–0.2609# (0.1471)
Observations	493	493

Notes: Columns 1 and 2 report high expected effects (τ_+), low expected effects (τ_-), and their interaction with a dummy indicating whether the proportion of course peers with low expected effect is greater than 0.5 (coefficients ϑ_+ and ϑ_- in model (8)). # $p < 0.100$; * $p < 0.050$.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire. Observations from blocks 8, 34, and 42 are excluded as course participants numbered less than 5.

Panel A: Baseline Beliefs



Panel B: Control Beliefs

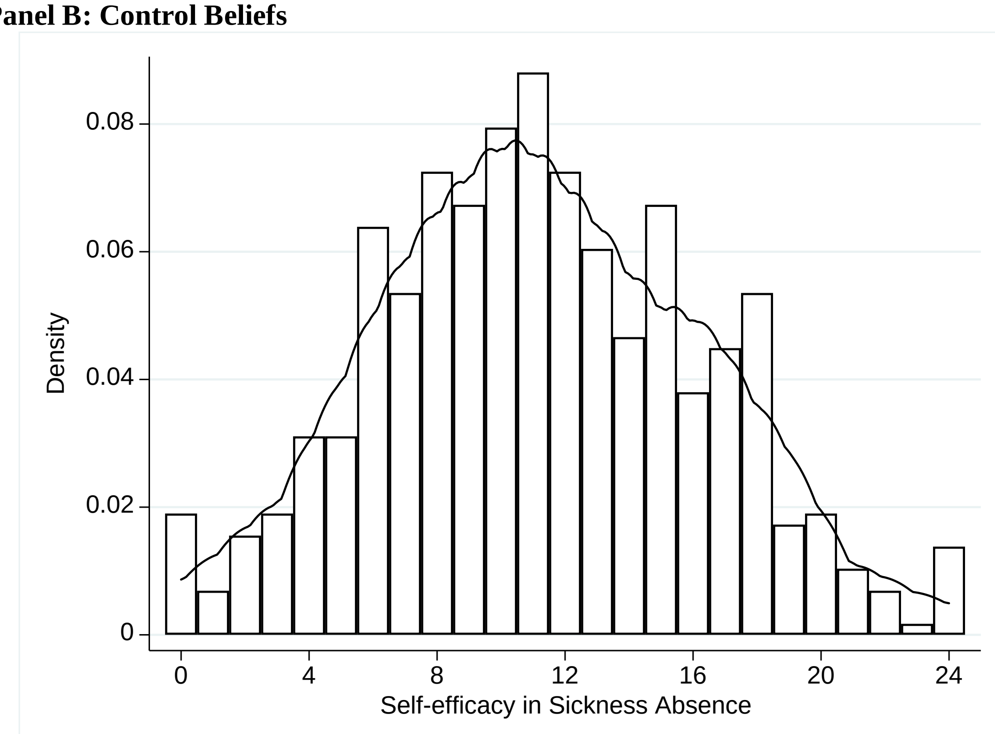


Figure 1

Distribution of Pre-intervention Beliefs

Source: Authors' calculations based on LTJS evaluation dataset and baseline LTJS questionnaire.

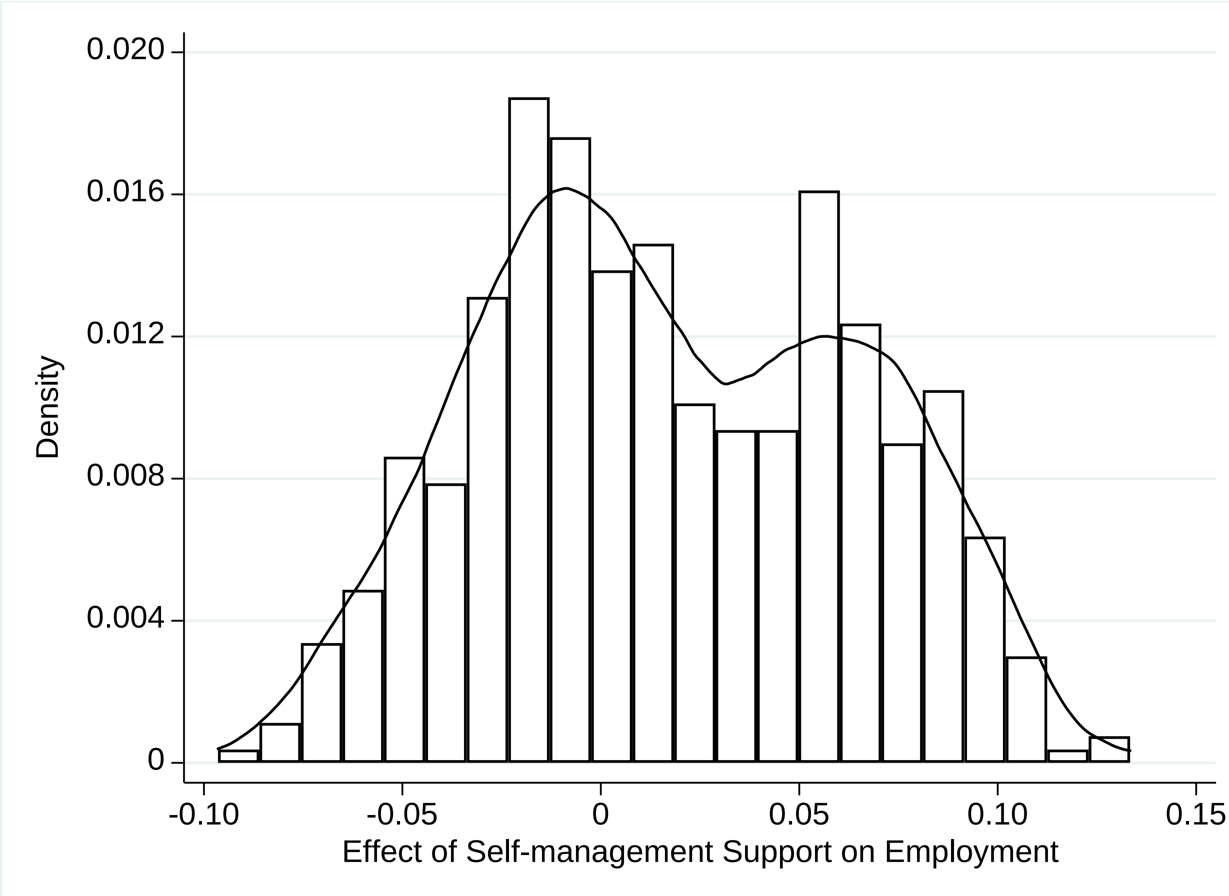


Figure 2

Distribution of Expected Effects of Self-management Support on Employment

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

Online Appendix

A. Supplementary Tables and Figures

Table A1

Selection Period and Size of Treatment and Control Samples for 43 Randomized Blocks

Block	Allocation Start	Allocation End	Treated	Control
1	September 2015	February 2016	9	4
2	February 2015	March 2015	7	4
3	March 2015	May 2015	7	4
4	January 2015	March 2015	13	6
5	April 2015	April 2015	7	5
6	November 2014	December 2014	13	5
7	February 2015	April 2015	9	4
8	September 2015	October 2015	5	3
9	April 2015	May 2015	12	7
10	October 2015	October 2015	7	4
11	July 2015	October 2015	9	5
12	January 2015	April 2015	11	3
13	April 2015	April 2015	9	6
14	August 2015	September 2015	8	3
15	July 2015	August 2015	12	8
16	August 2015	October 2015	9	6
17	February 2015	March 2015	11	5
18	March 2015	July 2015	6	4
19	January 2015	February 2015	5	4
20	April 2015	May 2015	7	2
21	May 2015	May 2015	7	4
22	June 2015	August 2015	9	4
23	November 2015	November 2015	16	9
24	December 2014	March 2015	5	5
25	February 2015	March 2015	7	4
26	March 2015	May 2015	9	5

(*continued*)

Table A1 (*continued*)

Block	Allocation Start	Allocation End	Treated	Control
27	February 2015	February 2015	9	7
28	April 2015	April 2015	7	4
29	February 2015	March 2015	9	4
30	January 2015	March 2015	13	7
31	March 2015	May 2015	14	9
32	October 2014	November 2014	11	6
33	December 2014	February 2015	9	1
34	February 2015	March 2015	4	3
35	April 2015	May 2015	6	3
36	February 2015	Maj 2015	11	4
37	June 2015	December 2015	9	7
38	January 2015	January 2015	7	4
39	February 2015	March 2015	6	3
40	February 2015	May 2015	7	4
41	July 2015	September 2015	7	4
42	May 2015	July 2015	5	1
43	July 2015	July 2015	8	3

Sources: LTJS evaluation dataset.

Table A2

Balance Tests. Coefficient and p-Values

Variable	Mean Impact Analysis Sample (1)	Heterogeneity Analysis Sample (2)	Mechanism Analysis Sample (3)
Sickness transition from employment	0.04 (0.50)	0.05 (0.48)	0.03 (0.62)
Work-related sickness	0.03 (0.50)	0.04 (0.48)	–0.02 (0.77)
Higher education	–0.01 (0.78)	–0.02 (0.72)	–0.00 (0.95)
Sickness transition from private sector job	0.05 (0.30)	0.05 (0.23)	0.05 (0.29)
Sickness transition from public sector job	0.04 (0.53)	0.05 (0.44)	0.01 (0.93)
Sickness caused by work conditions	0.04 (0.54)	0.04 (0.57)	0.06 (0.45)
Sickness caused by mental disorder	–0.03 (0.62)	–0.02 (0.73)	–0.01 (0.88)
Sickness caused by musculoskeletal	0.01 (0.76)	0.01 (0.92)	–0.02 (0.75)
Female	0.08 (0.21)	0.08 (0.26)	0.02 (0.80)
Age	3.10 (0.36)	2.80 (0.40)	0.38 (0.82)
Cohabitation with partner	0.08 (0.17)	0.06 (0.33)	0.09 (0.20)
Cohabitation with children	0.03 (0.63)	0.02 (0.67)	0.02 (0.71)
Previous long sickness absence	0.10 (0.07)	0.10 (0.07)	0.09 (0.10)
Previous sickness absence	0.02 (0.40)	0.02 (0.43)	0.01 (0.53)
Previous regular employment	0.03 (0.42)	0.05 (0.27)	0.05 (0.19)
Multiple serious illnesses	0.04 (0.35)	0.02 (0.66)	–0.01 (0.78)
Pain	0.55 (0.12)	0.41 (0.28)	0.32 (0.44)
Fatigue	0.13 (0.76)	0.23 (0.62)	0.07 (0.86)
Dyspnea	0.10 (0.73)	0.17 (0.62)	0.02 (0.96)
Sleep quality	0.40 (0.18)	0.32 (0.27)	0.21 (0.41)

(*continued*)

Table A2 (*continued*)

Variable	Mean Impact Analysis Sample (1)	Heterogeneity Analysis Sample (2)	Mechanism Analysis Sample (3)
Physical activity	0.03	0.08	0.16
	(0.91)	(0.75)	(0.43)
Prescribed medicine	0.02	0.00	–0.06
	(0.77)	(0.99)	(0.27)
Contact with GP	0.30	0.30	0.35
	(0.29)	(0.33)	(0.24)
Contact with psychologist	–0.25	–0.10	–0.13
	(0.29)	(0.68)	(0.59)
Contact with physiotherapist	0.08	0.07	0.22
	(0.85)	(0.86)	(0.62)
Contact with emergency department	0.04	0.04	0.06
	(0.35)	(0.38)	(0.29)
Somatic disorder	0.78	0.76	–0.12
	(0.53)	(0.58)	(0.91)
Illness worries	0.54	0.44	0.07
	(0.53)	(0.62)	(0.93)
Anxiety	0.59	0.72	0.47
	(0.20)	(0.15)	(0.27)
Depression	1.20	1.47	1.02
	(0.09)	(0.05)	(0.08)
Physical impairment	6.24	5.70	2.42
	(0.02)	(0.04)	(0.30)
Perceived serious consequences	0.62	0.74	0.41
	(0.20)	(0.16)	(0.28)
Expected long illness	1.01	0.94	0.64
	(0.03)	(0.04)	(0.10)
Perceived personal control of illness	0.11	0.15	0.16
	(0.68)	(0.61)	(0.50)
Perceived medical treatment effectiveness	0.73	0.78	0.58
	(0.05)	(0.04)	(0.08)
Perceived symptoms	0.47	0.57	0.24
	(0.30)	(0.25)	(0.54)
Illness concern	0.61	0.72	0.58
	(0.20)	(0.15)	(0.14)
Illness understanding	0.15	0.25	0.28
	(0.70)	(0.52)	(0.46)
Emotional response	0.66	0.80	0.42
	(0.18)	(0.14)	(0.33)
Autonomy motivation	1.39	1.25	0.80
	(0.54)	(0.58)	(0.60)
Introjected regulation	0.79	0.80	0.90
	(0.14)	(0.16)	(0.14)

(*continued*)

Table A2 (*continued*)

Variable	Mean Impact Analysis Sample (1)	Heterogeneity Analysis Sample (2)	Mechanism Analysis Sample (3)
External regulation	0.80 (0.32)	0.76 (0.35)	0.78 (0.33)
Amotivation	0.65 (0.16)	0.81 (0.10)	0.64 (0.16)
Expected full return to work	0.08 (0.70)	0.05 (0.84)	–0.05 (0.80)
Expected long sickness absence	0.38 (0.13)	0.43 (0.10)	0.38 (0.05)
Expected similar job tasks	0.17 (0.39)	0.14 (0.50)	0.08 (0.62)
Expected working time	0.11 (0.69)	0.11 (0.69)	0.08 (0.70)
Work capacity	–0.71 (0.85)	–0.91 (0.81)	–3.06 (0.34)
Life quality	0.15 (0.51)	0.16 (0.45)	0.09 (0.58)
Health-related life quality	0.14 (0.44)	0.17 (0.35)	0.11 (0.41)
Well-being	0.17 (0.83)	0.14 (0.86)	0.17 (0.76)
Self-efficacy in physical discomfort	0.02 (0.92)	0.03 (0.91)	–0.04 (0.77)
Self-efficacy in unexpected events	0.10 (0.63)	0.12 (0.57)	0.05 (0.73)
Self-efficacy in Social Interaction	0.10 (0.69)	0.12 (0.63)	0.07 (0.69)
Self-efficacy at center of one's life	0.11 (0.67)	0.12 (0.64)	0.00 (0.99)
Self-efficacy in frustration	0.06 (0.79)	0.09 (0.72)	0.04 (0.79)
Self-efficacy in new illness issues	0.12 (0.58)	0.16 (0.46)	0.15 (0.30)
Behavioral self-management	0.22 (0.80)	0.12 (0.89)	–0.21 (0.75)
Observations	568	511	400

Notes: P-values in parentheses. Column 1 presents balance tests obtained from model (1) for the sample used to estimated mean impacts. Column 2 presents balance tests obtained from model (1) for the sample used to estimate heterogeneity impacts. Column 3 presents balance tests obtained from model (2) for the sample used to estimate effects on mediating outcomes.

Source: Authors' calculations based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

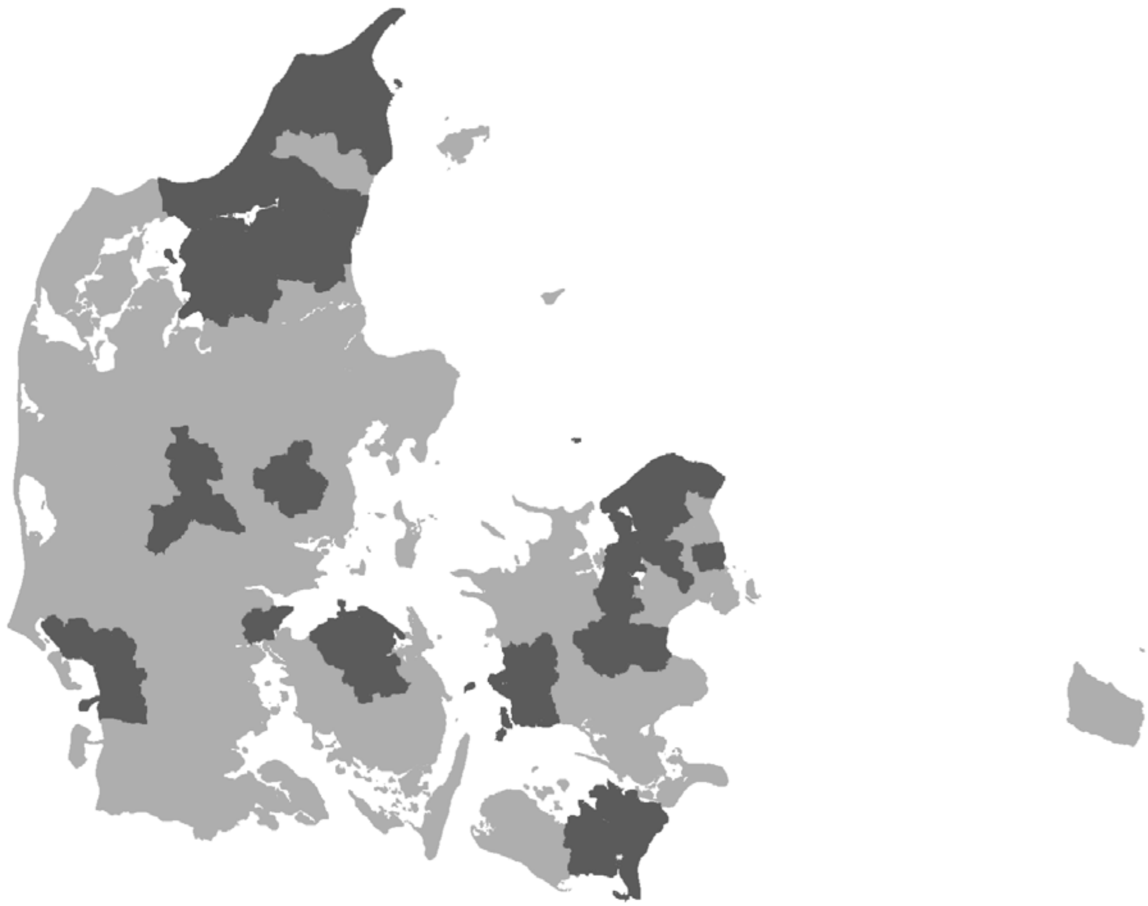


Figure A1

Participating Municipalities (dark greey)

Sources: LTJS evaluation dataset.

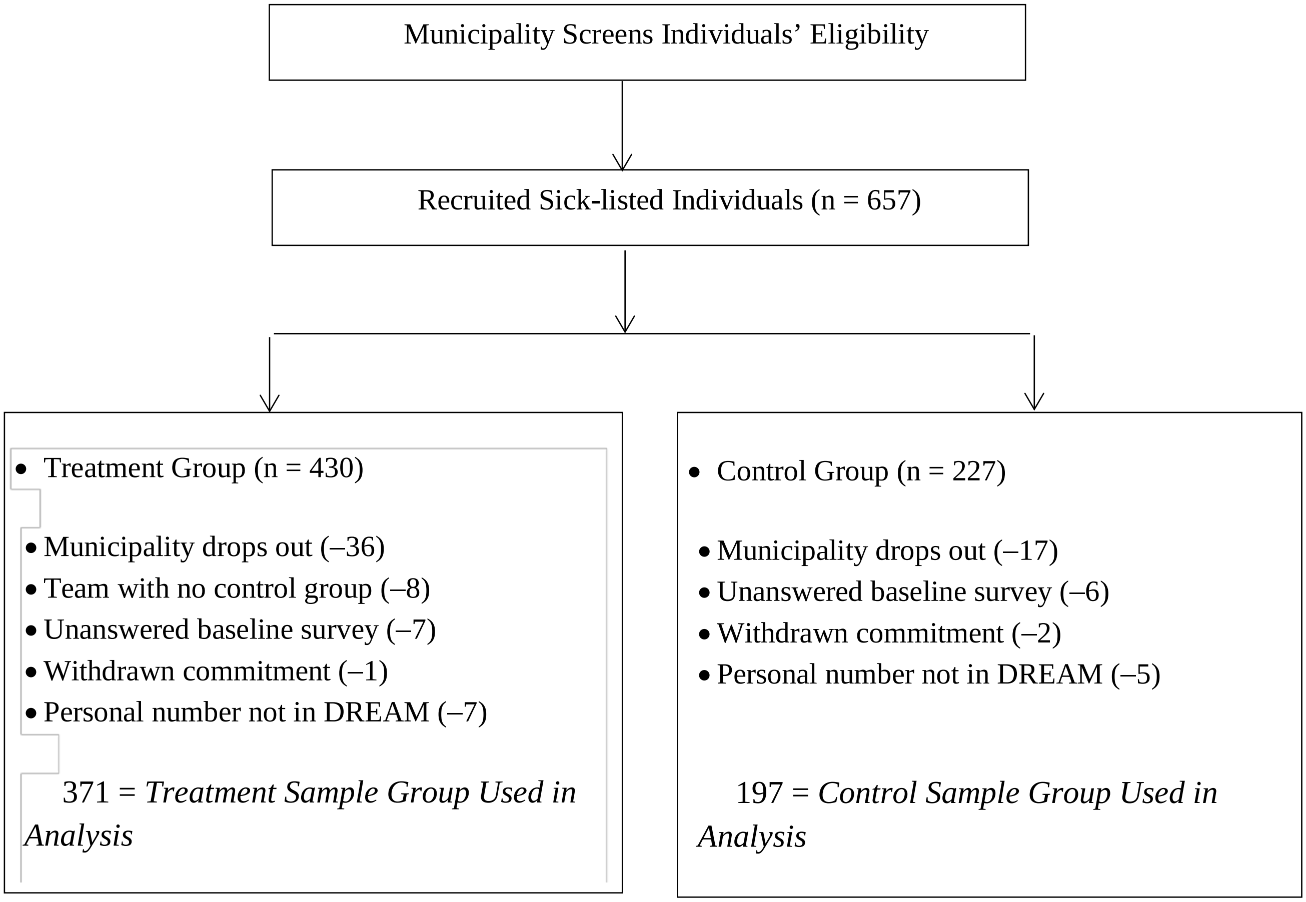
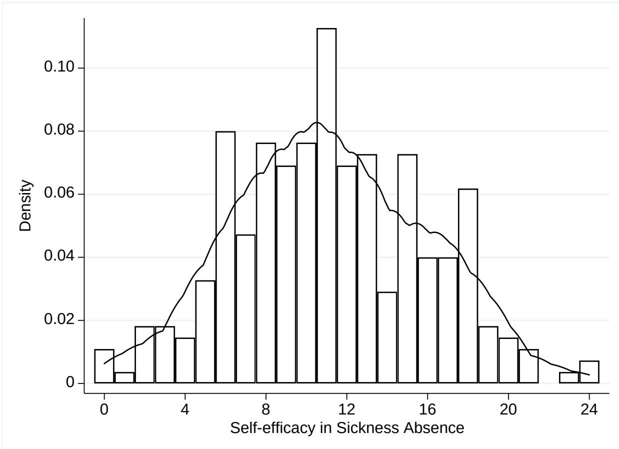


Figure A2

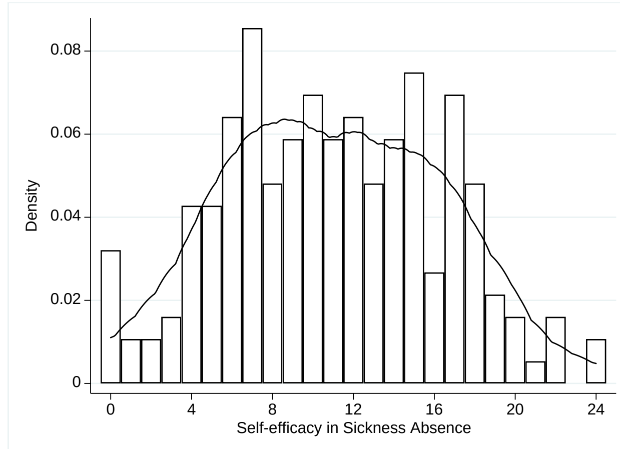
Sample Selection

Sources: LTJS evaluation dataset.

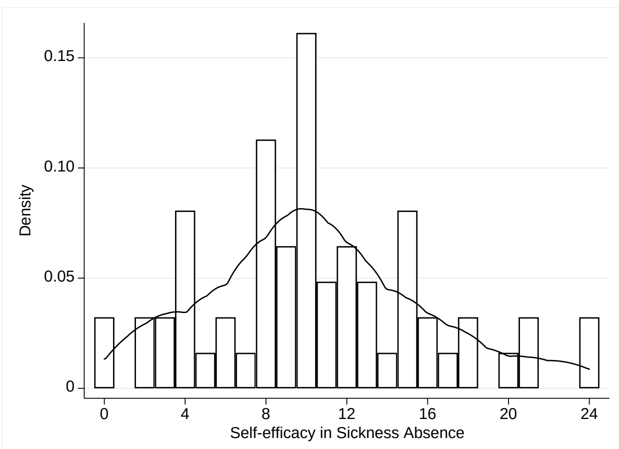
Panel A: Mental Health Disorder



Panel B: Musculoskeletal Disorder



Panel C: Other Physical Cause of Sickness



Panel D: Severe Injury, Cancer, or Cardiovascular

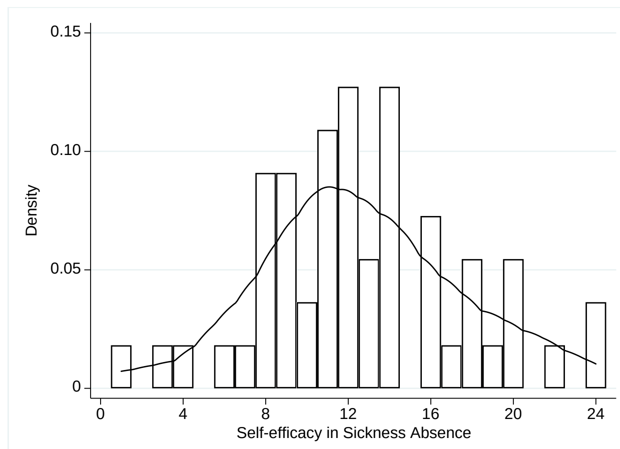
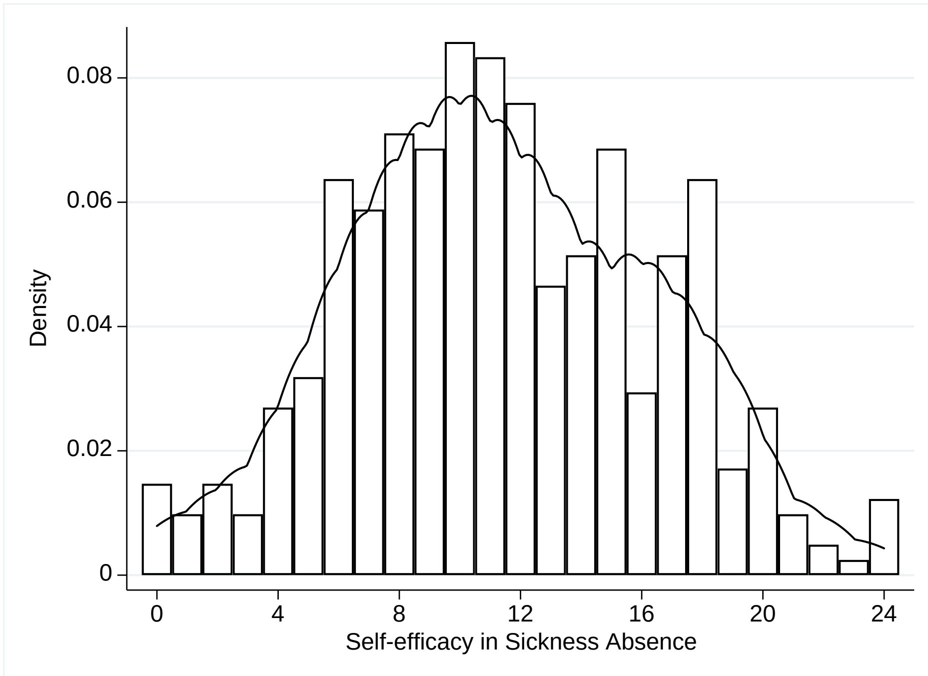


Figure A3

Distribution of Pre-Intervention Control Beliefs By Health Disorder

Sources: LTJS evaluation dataset, and baseline LTJS questionnaire.

Panel A: Sickness from Regular Employment or Self-employment



Panel B: Sickness from Unemployment or Subsidized Employment

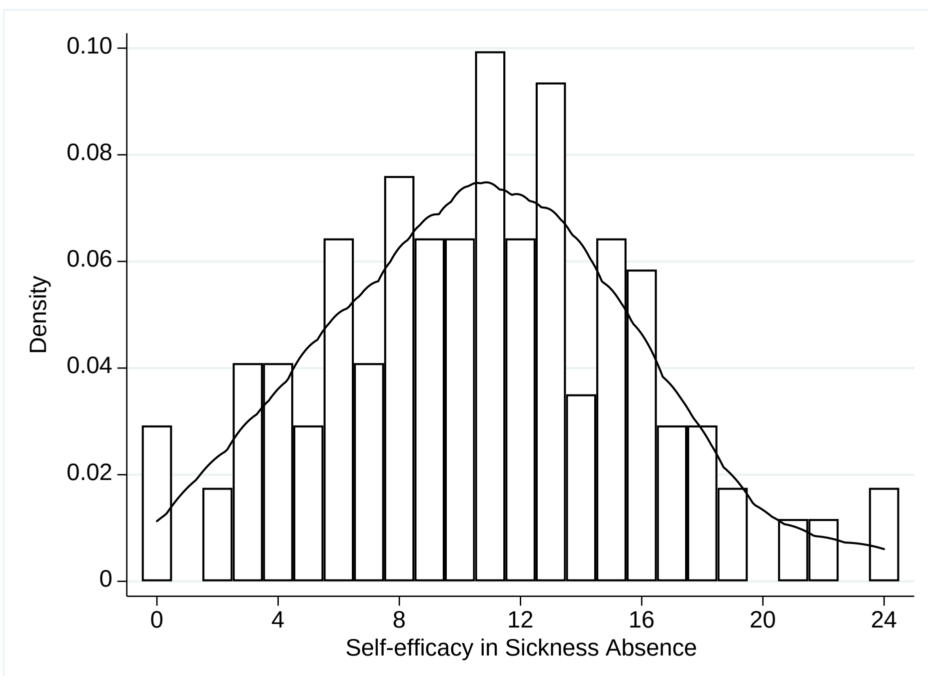
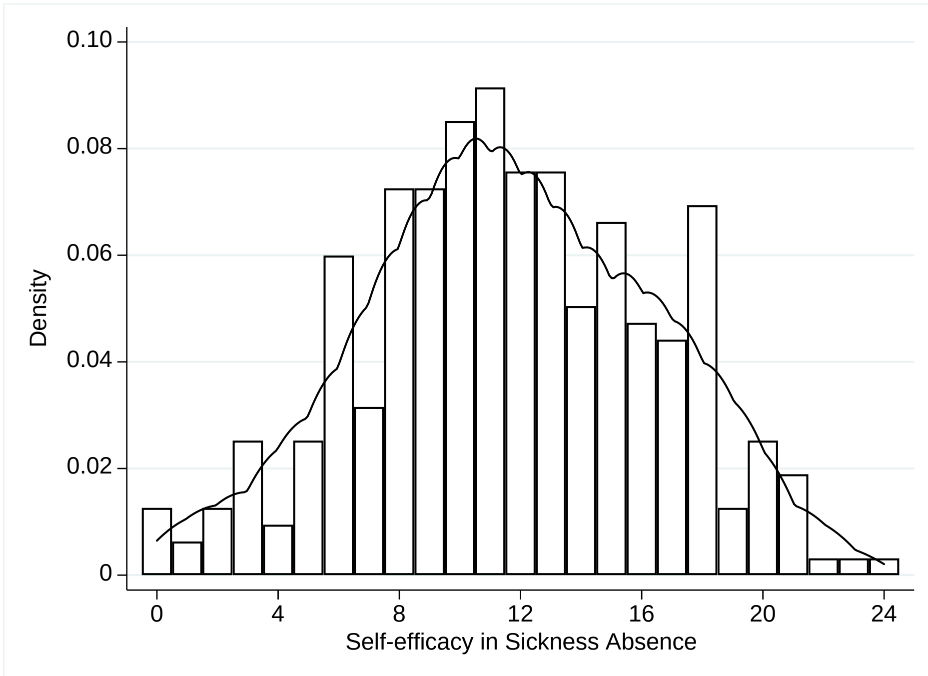


Figure A4

Distribution of Pre-Intervention Control Beliefs By Socio-economic Status

Source: LTJS evaluation dataset, and baseline LTJS questionnaire.

Panel A: No Higher Education



Panel B: Higher Education

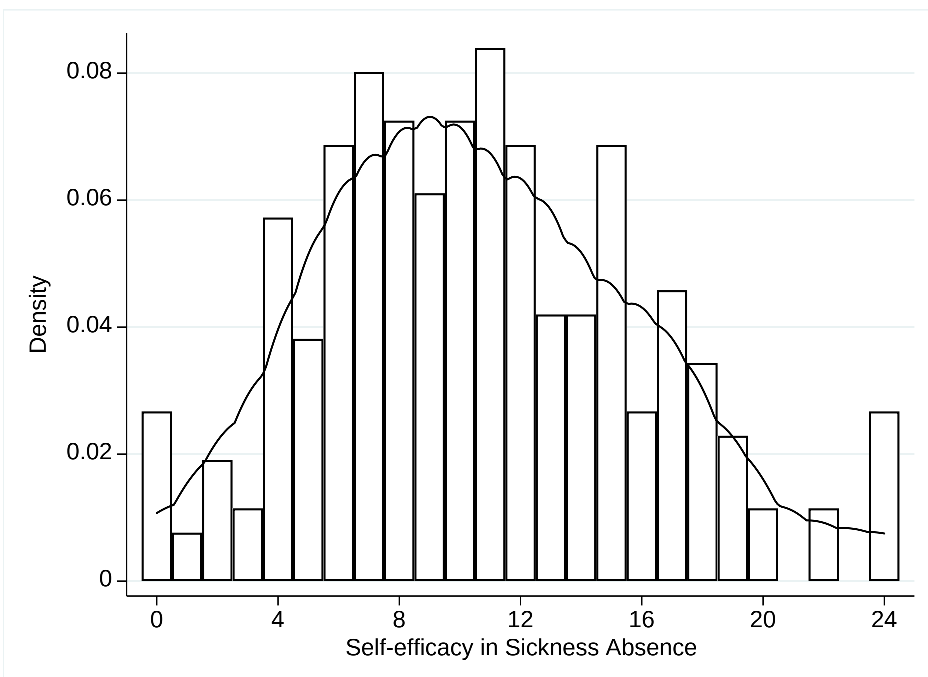
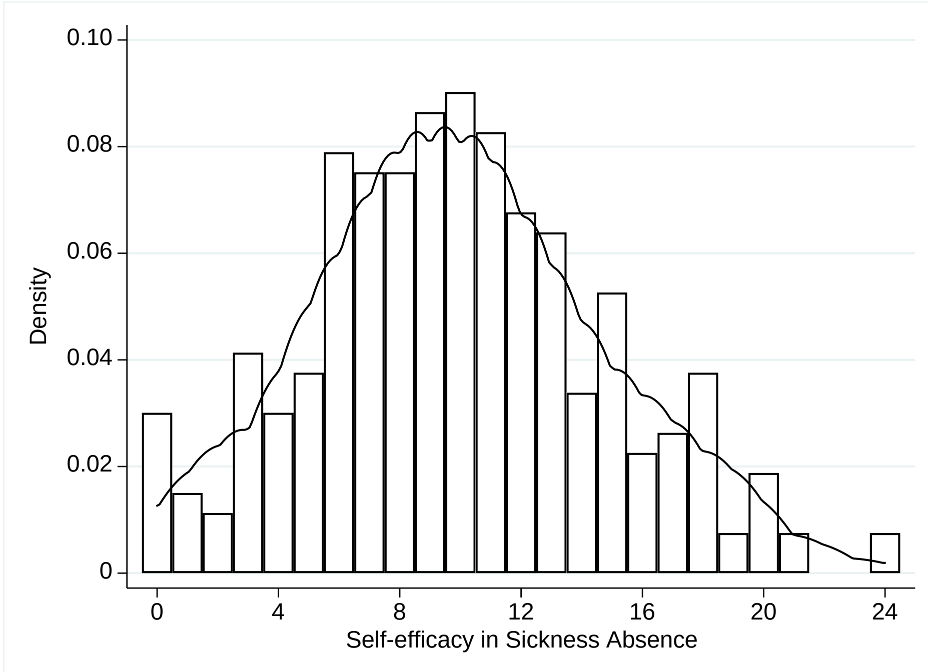


Figure A5

Distribution of Pre-Intervention Control Beliefs By Individual's Education

Source: LTJS evaluation dataset, and baseline LTJS questionnaire.

Panel A: Work Capacity under 40%



Panel B: Work Capacity at least 40%

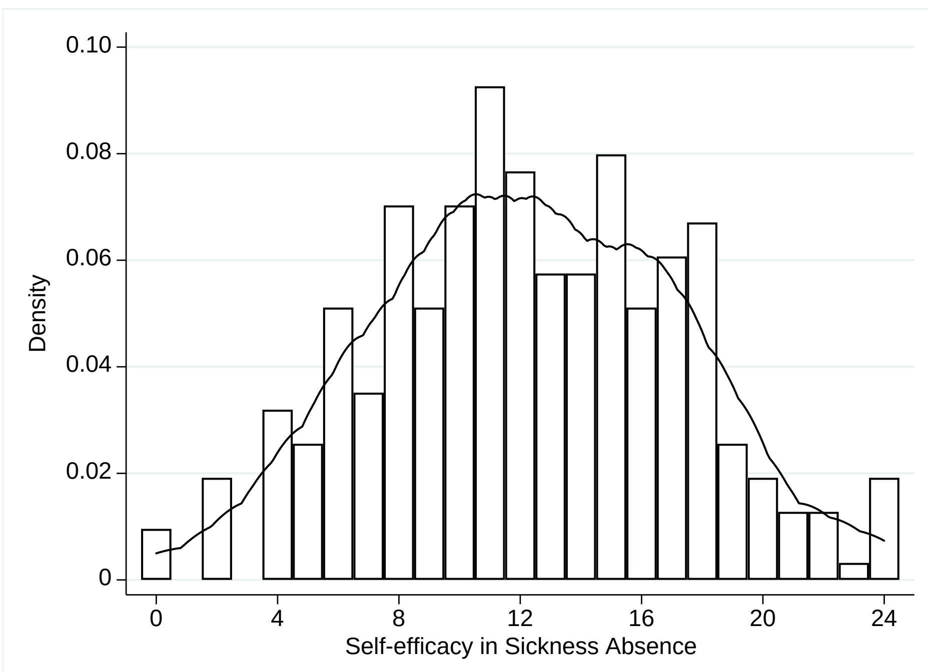
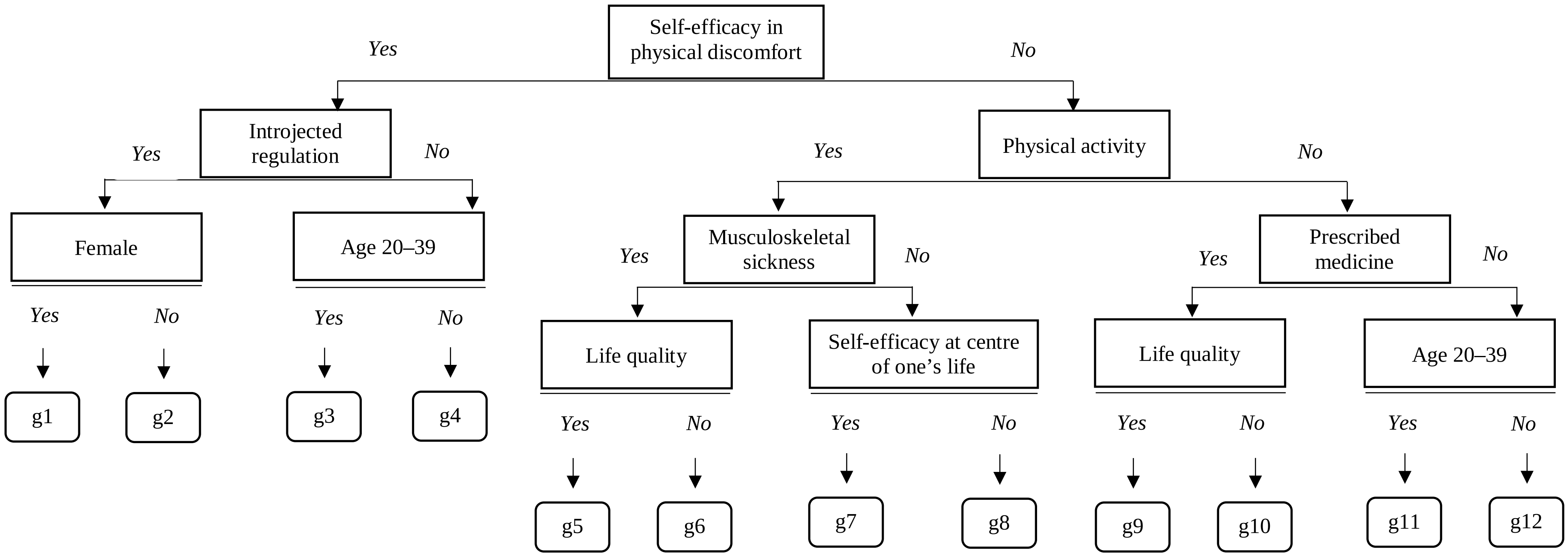


Figure A6

Distribution of Pre-Intervention Control Beliefs By Severity of Health Shock

Source: LTJS evaluation dataset, and baseline LTJS questionnaire.



g1: Female workers with self-efficacy in physical discomfort and introjected regulation (effect=–0.0182)
g2: Male workers with self-efficacy in physical discomfort and introjected regulation (effect=–0.0167)
g3: 20–30-year-old workers with self-efficacy in physical discomfort and no introjected regulation (effect=–0.0217)
g4: 31–39-year-old workers with self-efficacy in physical discomfort and no introjected regulation (effect=–0.0228)
g5: Workers with no self-efficacy in physical discomfort, sickness caused by musculoskeletal disorder, physically active and good quality of life (effect=0.0561)
g6: Workers with no self-efficacy in physical discomfort, sickness caused by musculoskeletal disorder, physically active and bad quality of life (effect=0.0637)
g7: Workers with no self-efficacy in physical discomfort, no musculoskeletal sickness physically active and self-efficacy in centre of one’s life (effect=0.0440)
g8 Workers with no self-efficacy in physical discomfort, not musculoskeletal sickness, physically active and no self-efficacy in centre of one’s life (effect=0.0513)
g9: Workers with no self-efficacy in physical discomfort not physically active, that receive prescribed medicine and have good quality of life (effect=0.0487)
g10: Workers with no self-efficacy in physical discomfort not physically active, that receive prescribed medicine and have bad quality of life (effect=0.0623)
g11: 20–30-year-old workers not physically active, with no self-efficacy in physical discomfort, who do not receive prescribed medicine (effect=0.0397)
g12: 31–39-year-old workers with no self-efficacy in physical discomfort, not physically active, who do not receive prescribed medicine (effect=0.0301)

Figure A7

Example of Causal Tree from the Causal Forest of the Effect of Self-Management Support on Employment

Sources: Authors’ calculations of group employment effects (in parenthesis) based on DREAM, LTJS evaluation dataset, and baseline LTJS questionnaire.

B. Control Beliefs

Individual's control beliefs are measured via two variables. We use the sum of scores from the University of Washington Self-efficacy Scale (Amtmann et al. 2012), and the third item of the Brief Illness Perception Questionnaire, which measures individual's perception of personal control of illness. The self-efficacy scale focuses on the individual's ability to control various key aspects of the consequences of illness on behavior, whereas the second measure proxies for the individual's locus of control regarding illness outcome.

How confident are you that:

SE1. You can keep the physical discomfort related to your health condition or disability from interfering with the things you want to do.

SE2. You can keep your health condition or disability from interfering with your ability to deal with unexpected events.

SE3. You can keep your health condition or disability from interfering with your ability to interact socially.

SE4. You can keep your health condition or disability from being the center of your life.

SE5. You can bounce back from frustration, discouragement, or disappointment that your health condition or disability may cause you.

SE6. You can figure out effective solutions to issues that come up related to your health condition or disability.

0: Not at all; 1: A little; 2: Average; 3: Mostly; 4: Totally

C. Baseline Beliefs

Baseline beliefs are measured via an expected return-to-work scale and a perception of illness scale. We first measure the expected return to work as $RW1 - RW2 + RW3 + RW4$ where RWj items read:

RW1. How confident do you feel that you can return to work completely? (Same number of hours as before your sick leave. Reply even if you have already returned.)

0: Not at all; 1: A little; 2: Average; 3: Very; 4: Totally

RW2. How long do you think it will take before you can return to work? (Same number of hours as before your sick leave. Reply even if you have already returned.)

0: Less than 1 month; 1: Between 1–3 months; 2: Between 3–6 months; 3: Between 6–12 months; 4: More than 1 year

RW3. What tasks do you envision returning to?

0: Cannot return at all; 1: Have to find a whole new career path; 2: Have to switch to other tasks; 3: Returning to most of my past duties; 4: Returning to all of the same tasks as before

RW4. How many hours per week do you expect to work if you return?

0: Under 15 hours; 1: 15–20 hours; 2: 20–25 hours; 3: 25–30 hours; 4: Full time, or more

The perception of illness is measured as PI = $PI1 + PI2 + PI3 + PI4 + PI5 + PI6 + PI7 + PI8$, where the items are obtained from the short-form of the illness perception questionnaire (Broadbent et al. 2006):

PI1. How much does your illness affect your life?

0: No effect at all; 1 …., 10: Severely affects my life

PI2. How long do you think your illness will continue?

0: A very short time; …., 10: Forever

PI3. How much control do you feel you have over your illness?

0: Absolutely no control; …., 10: Very high level of control

PI4. How much do you think your treatment can help your illness?

0: Not at all; …., 10: Extremely helpful

PI5. How much do you experience symptoms from your illness?

0: No symptoms at all; ....; 10: Many severe symptoms

PI6. How concerned are you about your illness?

0: Not at all concerned; …., 10: Extremely concerned

PI7. How well do you feel you understand your illness?

0: Don't understand at all; …., 10: Understand very clearly

PI8. How much does your illness affect you emotionally? (For example, does it make you angry, scared, upset, or depressed?)

0: Not at all affected emotionally; …., 10: Extremely affected emotionally

D. Behavioral Self-management

Behavioral self-management of illness is measured as the sum of scores of *BSMj* items from the chronic disease self-management questionnaire (Lorig et al. 1996):

BSM1. Try to distance yourself from the discomfort and pretend that it is not part of your body.

BSM2. Think of it not as discomfort, but as another feeling, such as a sensation of warmth or numbness.

BSM3. Try to direct your thoughts away from the discomfort, such as by singing or low brain gymnastics

BSM4. Perform muscle relaxation.

BSM5. Use visualization of positive feelings or experiences, such as visualizing that you are somewhere else.

BSM6. Think positive thoughts.

0: Never; 1: Almost never; 2: Sometimes; 3: Pretty often; 4: Very often; 5: Always

E. Motivation for Work

Motivation for work is measured as $\sum_j Mj + \sum_j IRj + \sum_j ERj - \sum_j AMj$, where items are obtained from Treatment Self-Regulation Questionnaire (Levesque et al. 2007).

The reason I try to return to work is:

M1. Because working is very important for me to be as healthy as possible

M2. Because I personally believe that work is the best for my well-being.

M3. Because I feel I want to take responsibility for my health.

M4. Because working involves important decisions that I really want to make for myself.

M5. Because I have considered it carefully and believe that work is of great importance to several significant parts of my life.

M6. Because working is related to my life goals.

IR1. Because I would feel guilty or ashamed if I did not try to return.

IR2. Because I would feel like a bad person if I did not try to return.

ER1. Because I feel that others are pushing to make me return.

ER2. Because others would be annoyed and angry with me if I did not try to return.

ER3. Because I want to prove to others that I can work.

ER4. Because I want the praise and recognition of others.

AM1. I am not thinking about returning to work at all.

AM2. I actually don't know why I'm trying.

AM3. Because it is easier to do what I am told than to think about it.

0: Completely wrong; ...; 3: About right; ...; 6: Completely true

F. Causal Forest Algorithm

(1) Draw a subsample without replacement from the heterogeneity analysis sample ($N_b = f \cdot N$).

(2) Randomly split the sample used to build tree b in the training sample ($N_{bt} = t \cdot N_b$) and the estimation sample ($N_{be} = (1 - t) \cdot N_b$).

(3) Use the training sample to specify the causal tree:

(3.1) Calculate the expected variance of employment effect in the training sample (V).[21]

(3.2) Randomly select v pre-treatment indicators from the 59-dimensional set X_i.[22]

(3.3) For each indicator X_{ki} form a candidate sample split by placing the observations of the training sample with $X_{ki} = 0$ in a left child leaf and the remaining observations of the training sample with $X_{ki} = 1$ in the right child leaf.

[21] See Athey and Imbens (2016) for the objective function used to measure goodness of fit.

[22] Step (3.2) improves prediction by ensuring that the same highly correlated indicator does not appear in all the trees.

(3.4) If there are at least h treatment and h control observations in the left and right child leaves, calculate the variance of treatment effects across both leaves while penalizing for overfitting (V').

(3.5) Implement a sample split based on the indicator with highest variance of treatment effect ($\max V'$) as long as $\max V' > V$.

(3.6) After the sample split, repeat the process 3.1 to 3.5 in each child leaf, until it is not possible to make more sample splits to the child leaf.

(4) The causal tree is fully grown when all leaves are terminal leaves.

(5) Assign all the observations of the estimation sample (N_{be}) into subgroups determined by the tree specified with the training sample.

(6) Calculate $\hat{\theta}_{l,b} = \sum_{i \in l, W=1} Y_{il} / N_{l_{be}}^{(1)} - \sum_{i \in l, W=0} Y_{il} / N_{l_{be}}^{(0)}$, where $N_{l_{be}}^{(1)}$ is the number of treated observations in the terminal leaf l of the estimation sample for bootstrap b.[23]

(7) Assign $\hat{\tau}_{l,b} = \hat{\theta}_{l,b}$ to all observations of the sample N that, given their full set of covariates X_i, correspond to subgroup l.

(8) Iterate (1) to (7) for $B = 25{,}000$ bootstrap samples.

(9) The individualized conditional average treatment effect for observation i is given by $\hat{\tau}_{CF,i}(x) = B^{-1} \sum_{b=1}^{B} \hat{\tau}_{l,b}$.

[23] Step (6) corrects for overfitting by using a different sample to estimate the leaf effects than the sample used to specify the model.