

Effects of Dialogue Meetings - Pre-analysis plan Phase 1*

Matteo Alpino¹, Karen Evelyn Hauge¹, Andreas Kotsadam¹, and Simen Markussen¹

¹Ragnar Frisch Centre for Economic Research

February 12, 2018

Abstract

Sickness absence entails large individual and societal costs. Dialogue Meetings (DMs) where the absentee, the employer, and often the physician, discuss arrangements for full or partial work resumption have been in place in Norway since 2007. In collaboration with the Norwegian Labour and Welfare Administration (NAV) we are conducting a large scale randomized natural field experiment to investigate if and for whom DMs have effects. The experiment is still ongoing and is conducted in several phases to test for mechanisms iteratively. While the experiment is still ongoing, it is possible to analyze data on the individuals that have already been involved in the experiment so far (early sample). To do so we proceeded as follows: first, we randomly divided the early sample into an early training sample and an early test sample. NAV then merged the early training sample with outcome data and covariates and handed the data over to us. Based on the results in the early training sample, we here register some key decisions that will be followed in analyzing the rest of the data (the analysis sample, which comprises the late sample and the early test sample). The results from the early training sample are presented in a companion document and will be subject to a direct replication using the early test sample. Based on the findings in the early test sample we will decide whether and how to change the intervention. Such changes will be pre-registered at a later stage.

Keywords: RCT; Sickness absence; Employment; Causal tree and forest; Machine learning

*We would like to thank Pathric Hägglund and Per Johansson for valuable suggestions. Funding from the Norwegian Research Council (grant number 259512) is also acknowledged.

1 Introduction

We conduct a large scale randomized field experiment to investigate the effects of mandatory dialogue meetings (DMs). In such meetings, the absentee, the employer, and often the physician, discuss whether arrangements can be made at the workplace that make full or partial work resumption possible (e.g. to alter the number or the nature of tasks at work). A system of DMs have been in place in Norway since 2007 and it is explicitly intended to induce long-term absentees to fully or partly return to work. DMs constitute a combination of nudging, monitoring and facilitation of appropriate working conditions and tasks.

In order to test the causal effect of the DMs we team up with the Norwegian Labour and Welfare Administration (NAV) to conduct a large scale randomized field experiment. In particular, we randomize: 1) whether people receive letters calling them in to mandatory or voluntary meetings; 2) when in the sickness spell they receive the letter; 3) and when the meeting is scheduled (if mandatory). Calling in people to a mandatory meeting may have a notification effect in addition to the meeting itself being successful. The notification effect is similar to the "threat effect" in the unemployment insurance literature with reference to the effect of being summoned to an activation program. In the case of DMs, however, the notification effect is a mix of behavioral responses by the employee, the employer as well as by the physician. There may also be a reversed notification effect if people that expect to be called in to a meeting are instead told that a meeting is voluntary for them.

The experiment enables us to answer a wide set of policy relevant questions, such as: Can mandatory dialogue meetings reduce sickness absence and promote work resumption? Is it sufficient to nudge people with the threat of a meeting to make them come back to work faster? Are there adverse effects of calling people in to voluntary meetings? Randomized controlled trials are an excellent tool for recovering average treatment effects without bias. Underlying the average effect there may be substantial heterogeneity and for some

sub-populations the effects may be the opposite of the average effects. From a theoretical perspective, treatment effect heterogeneity may help us understand the mechanisms behind why a treatment works or not and facilitates the generalization of findings from an experiment to other settings. From a policy perspective, improved targeting is also important for cost-efficiency.

However, identifying treatment effect heterogeneity is more difficult than identifying average treatment effects. While it is important to establish which groups benefit or are hurt most by a policy, there are pitfalls of naively splitting the data to test for effects across subgroups. One issue is that multiple hypotheses are tested which is likely to lead to false positives and spurious correlations if not corrected for. Without a pre-specified analysis plan it is impossible to know how many non-reported tests were conducted by the researchers (Olken, 2015). Specifying all possible tests is, however, difficult, especially in a setting where there is an inherent uncertainty with respect to treatment heterogeneity.

Hence, there is a tradeoff between pre-specifying hypotheses and learning about heterogeneity from the data. A solution to the problem is replication of research findings (Coffman and Niederle, 2015; Coffman, Niederle, and Wilson, 2017) but often experiments are costly and difficult to replicate. Recently, economists have started to borrow methods from the machine learning literature to solve this issue. Machine learning methods are primarily aimed at predicting an outcome variable and use cross-validation to compare model predictions to actual outcomes in test samples not used to estimate the model (see e.g. Mullainathan and Spiess (2017) and Varian (2014) for recent reviews of this literature from the perspective of economics). Inspired by cross-validation, Athey and Imbens (2016) propose ways to estimate heterogeneous treatment effects relying on machine learning techniques and sample splitting: one sample is used to identify different subgroups based on all the available covariates, while the second sample is used to estimate treatment effects. Given that the second sample has not been used to select the subgroups, the subgroup structure is exogenous and standard

inference can be used. These methods, which Athey and Imbens (2016) and Wager and Athey (2017) label "honest", reduce the problem of multiple hypothesis testing by building replication in the process. A similar idea has already been suggested for randomized experiments as well whereby researchers are encouraged to only use a portion of the data for the analysis and replicate the results using the other portion (Anderson and Magruder, 2017; Fafchamps and Labonne, 2017). We will employ "honest" machine learning techniques to search for heterogeneous treatment effects and we also conduct a sample split approach on our early data to derive some early hypotheses and to guide our efforts in changing the experiment while it is still ongoing.

In particular, we want to test for mechanisms iteratively but in order to do so we must know what the results are. We therefore conduct an analysis on a set of individuals that were included in the experiment early on. To reduce the risks of finding false positives, yet still being able to learn from the data in an explorative fashion, we conduct an analysis on a random sample of the early sample (early training sample). The results of this analysis is presented in the companion document "Effects of Dialogue Meetings - results from the early training sample". Based on the results in this early training sample, we specify some key hypotheses. Some of these hypotheses will be replicated on the early test data (that is the remaining part of the early data). This replication is especially powerful as the data come from the same experiment and the same time period. The results from the replication will guide us in refining the experiment and test additional hypotheses. Had we conducted the analysis on the whole early data we would have been more worried about our results being false positives and had we specified the analysis before analyzing the early data we would have learnt less from the data and some hypotheses would have been left untested. The refinements and new hypotheses that arise from the early test data will be pre-specified at a later stage.

This plan also outlines decisions to guide us in the analysis of the the analysis sample,

which consists of the early test sample as well as all individuals that are included in the experiment after the date of the early sample (late sample). As the analysis sample is larger, some of the methods and tests are more suited to be tested on that sample, in particular the machine learning techniques which are very data intensive and the more fine grained heterogeneity (e.g. by offices, occupations, and diagnoses).

2 The field experiment

The regular procedure of calling in to dialogue meetings differs across the country. In principle, all absentees should have a meeting before week 27 but in practice there are many exceptions to this rule. We work with 13 offices that are willing to change their procedures of targeting for the meetings. Most of these offices are in Oslo (the Norwegian capital), but three of them are outside of Oslo. We started the trial in September 2016 with one office and it has been expanding since then and is still ongoing.

The new procedure consists of several steps and absentees are randomly assigned to seven different treatments. Each week a caseworker at each office checks whom of the absentees are in their 8th week of the sickness spell and are employed. These individuals are assigned to different caseworkers. The assignment of individuals to caseworkers is not randomized and is different in different offices. In some of them, caseworkers are assigned equally many individuals and assignment is by birth dates. At other offices, some caseworkers are assigned more individuals. At yet other offices, some caseworkers are assigned absentees by industry or by employers. When the caseworkers have received their cases they log in to a secure internet page administered by us and write in the id number of the absentee.

Once the absent person has been added to the system, the caseworker is asked two questions: 1) "Based on the information you have, how many more weeks do you think the sickness absence will continue?". The alternatives are: less than 4 weeks, between 4 and 11 weeks, between 12 and 20 weeks, more than 20 weeks, or do not know. 2) "In your

opinion, how important is it that the absentee is called in to a meeting?”. The alternatives are: Very important, important, neither important nor unimportant, not that important, not important at all. The idea of including these questions is to investigate how caseworker discretion may improve the targeting of the system.

The absentees are then randomly assigned to different treatments, shown in Figure 1. The variations of the treatments are of different types. First of all, we vary whether the meeting is mandatory or not. In principle, everyone has a right to a meeting and NAV does not want to infringe on this right. The individuals not being called in to a mandatory meeting therefore receive an offer to have a voluntary meeting. That is, NAV sends them a letter explaining that they know that the person is on sickness absence but that they have decided not to call in to a meeting. The letter further explains that if the person herself, her doctor, or employer thinks there is a need for a meeting, she can contact the office and a meeting date will be arranged. They also send a similar letter to the employer of the absentee.

For the mandatory meetings, we vary the timing so that they can be held in week 13, 19, or 26 of the absence spell (that is 5, 11, or 18 weeks after the draw). By comparing people randomly assigned to having meetings at different times we can test whether earlier meetings are better than later meetings. The timing of the letter is also varied to test for notification effects and more broadly to test if there is a difference in the effects of early or late letters. The design is set up so that we have overlapping groups of letters and meetings with different combinations. In particular, we randomly assign the timing of letters for people with meetings in week 26 to be either in week 15 or in week 22. If there is a strong notification effect of receiving the mandatory meeting letter we should expect people getting earlier letters to have shorter sickness spells. We also vary the timing of the meeting week to be either in week 19 or in week 26 for people getting letters in week 15. The timing of letters for the voluntary meeting are also randomized to be sent out in week 9, 15, or 22. This

is useful in order to test for a type of reversed notification effect whereby people receiving the letter with a voluntary meeting may feel less monitored and perhaps stay on sick-leave longer.

Letter is sent	Offer of meeting	Mandatory meeting
Week 9		Meeting in week 13
Week 15		Meeting in week 19
Week 15		Meeting in week 26
Week 22		Meeting in week 26

Figure 1: Different treatments

The offices have different capacities and different wishes of how many meetings to have. We have as our baseline case equal probability of each of the eight cells (there are in fact only 7 groups but we make the voluntary meeting group for letter week 15 double as it is used as a comparison to two mandatory letters), i.e. a 12.5 percent chance of each cell. Some offices, however, wanted other combinations. In particular, many offices wanted a slower start with only late meetings to start with. Other offices need to change the number of meetings over time due to personnel shortages. So far, only one office wanted more than 50 percent of the meetings to be mandatory. In some rare cases, individual caseworkers have had personalized probabilities as well. For instance, one caseworker works with one large employer only and

this employer wants to have meetings for all absentees. In that case, only the timing of the mandatory meetings and letters are randomly assigned.

3 Samples

There are several different samples in the study and we show them in Figure 2. We have an early sample consisting of individuals that have already been involved in the experiment long enough to have meaningful outcome measures. In order to reliably test for early stage heterogeneous treatment effects, we randomly divide this sample further into an early training sample and an early test sample. We also have a late sample of people that become enrolled in the experiment at later stages. The final analysis sample will contain the late sample and the early test sample. In this section we describe how we divide the total sample into these sub-samples.

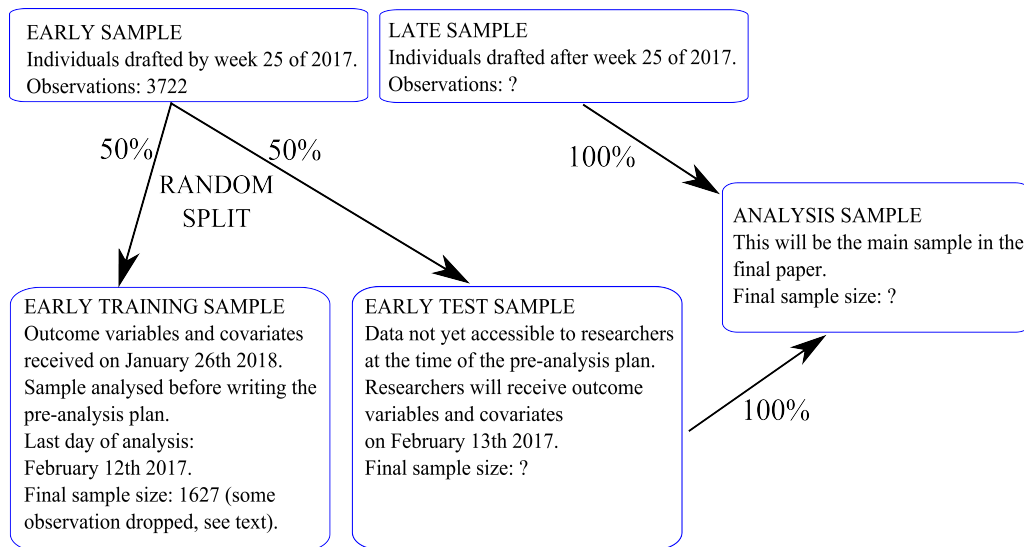


Figure 2: The different samples

On October 30th (week 44) 2017 we took out all the data from the secured internet page. This data only contains information about the treatment status and not any information about outcomes. At this date, 7,619 absentees had been entered into the system. To start the analysis, we wanted to extract a random sample of individuals that have had the opportunity

to have a meeting. In week 44 of 2017 we could thereby include everyone who had been drafted until 18 weeks before (as they are drafted in week 8 of their spell and have the meeting in week 26 of the spell). Hence, we extracted everyone drafted before week 26 of 2017. This amounts to 3,742 individuals and these individuals constitute our early sample. After some initial power calculations, suggesting that we would be able to detect medium-sized effects for the mandatory versus voluntary meeting using half of this sample (the minimum detectable effect was calculated to be 0.13 standard deviations), we decided to have NAV add outcomes and covariate data for a random half for the early sample. This half constitutes the early training sample and the other half of the early sample is called the early test sample.

In taking out the early training sample we wanted to have as similar individuals as possible in the training sample and in the test sample used for replication. We first stratified the data based on caseworker and week and randomly assigned half of the observations to each sample. In cases of odd numbers of absentees we stratified the rest by office and week. The few remaining ones were randomly assigned within weeks only.

Individuals drafted after week 25 of 2017 up until we change or stop the experiment constitute our late sample. The final analysis sample will consist of all individuals in the late sample and the individuals in the early test sample.

4 Data and coding of main variables

Decision rules for dropping observations: For the early training sample we asked NAV to add data on outcomes as well as background characteristics. We received this data on January 26th 2018. Some individuals out of the 1871 had to be dropped since they were drawn after their sickness spell had ended. We also drop one individual with an error in the sickness absence date. In total, we have 1627 observations with 48 percent assigned to mandatory meetings and 52 percent assigned to the voluntary meetings. The same decision

rules for dropping observations will be used in the other samples.

Primary and secondary outcome variables: We choose our primary outcome variable to be the total days of sickness absence (*Total*). This variable counts the total number of days of sickness absence between the draw and the date of data extraction. We have two secondary outcome variables that we expect will give us similar results but that will not be analyzed as extensively nor by themselves be seen as confirmatory. These variables are *Days*, which is a measure of days within the current sick leave spell and *Graded days*, which is the grade adjusted total number of days. People may be on sick leave 100 percent, which is the most common, but they may also be partly on sick leave. *Graded* is a measure of full-day equivalents for the total sick days. If a person is on half time sick-leave for a period, for instance, this measure counts each day as a half day during this period.

The health related part of the Norwegian social insurance program consist mainly of three programs. For employees getting sick the first program is sickness leave (SL), as is mainly studied in this experiment. The worker can be on SL for up to one year, with full wage replacement (up to a threshold of 6 base amounts). If the worker is unable to return to work after 1 year, and her work ability is reduced by at least 50 percent, she can start on temporary disability insurance (TDI). On TDI the replacement rate is 66 percent and one can normally receive TDI for a period up to 4 years. During TDI one is supposed to participate in training and rehabilitation programs and the goal is to return to employment. If this does not succeed one can apply for permanent disability insurance (DI). For the early sample we will not investigate effects on TDI as most individuals have not been on leave for a long enough time to qualify for this. We will add this variable as a secondary outcome variable in later analyzes, however.

Based on the days within the spell we will dummy code whether an individual has returned before week 13, 19, 22, and 26. As described in Section 5, these variables will be used to test the different types of threat effects.

Treatment variables: We will use whether individuals are assigned to receiving a letter for a mandatory meeting (Treatment) or voluntary meeting (Control) as our main treatment variable of interest. We also dummy code each of the seven individual treatment variables and use them in some analyzes (see Section 5).

Other variables: The baseline background variables we use as controls and for heterogeneity are female, birth year, days before (the absence history measured as the total number of days on sickness absence since 2015 up until the date of the draw), grade of sickness absence (at the time of the draw), and number of employees at the workplace (at the time of the draw). In the balance tests these will either be coded as a dummy (female) or as continuous variables. Alternative codings for the heterogeneity analysis are described in Section 5.

We have the diagnoses of all the individuals in the data classified according to the second edition of the International Classification of Primary Care (ICPC-2). This classification uses a letter for each broad type of diagnosis such as a bodily part (e.g. F=Eye) or type (e.g. P=Psychological), and a number within each letter. A distinction is made between symptoms/complaints and established diagnoses within each letter. We create a dummy variable that equals one for individuals that are classified as having symptoms and zero otherwise.

We also create two variables based on the caseworkers predictions. These are *CW predicted meeting important* which equals 1 if the caseworkers predicted the meeting to be "important" or "very important" and *CW predicted long spell* which equals one if the caseworker predicted a spell of over 20 weeks. For the early training sample the summary statistics are presented in Table 1. We will create the same table for the analysis sample.

Balance tests: To test for balance we will regress our main treatment variable on the other variables described above both individually and together, while controlling for block fixed effects (a combination of week and caseworker). We will judge whether the randomization worked by conducting an F-test of whether the control variables jointly predict treatment

Table 1: Descriptive statistics.

	Total		Treated		Control	
	Mean	SD	Mean	SD	Mean	SD
<i>Dependent variables</i>						
Total days	123.80	(89.1)	121.37	(90.3)	126.44	(87.8)
Days (within spell)	103.24	(88.7)	101.76	(89.3)	104.85	(88.2)
Graded days	90.56	(77.9)	87.55	(77.7)	93.83	(78.1)
<i>Returning before...</i>						
13 weeks	0.55	(0.5)	0.56	(0.5)	0.53	(0.5)
19 weeks	0.65	(0.5)	0.66	(0.5)	0.65	(0.5)
22 weeks	0.69	(0.5)	0.70	(0.5)	0.69	(0.5)
26 weeks	0.74	(0.4)	0.75	(0.4)	0.74	(0.4)
<i>Baseline control variables</i>						
Female	0.64	(0.5)	0.64	(0.5)	0.64	(0.5)
Birth year	1974.92	(11.8)	1974.95	(12.0)	1974.88	(11.6)
Days before	116.85	(80.1)	115.13	(75.5)	118.72	(84.8)
Grade	72.79	(27.1)	71.92	(27.0)	73.72	(27.3)
nr employees	503.16	(1501)	529.38	(1657)	474.69	(1311)
<i>Other main heterogeneity variables</i>						
Symptoms	0.41	(0.5)	0.40	(0.5)	0.42	(0.5)
<i>Caseworker predictions</i>						
CW predicted meeting important	0.62	(0.5)	0.61	(0.5)	0.63	(0.5)
CW predicted long spell	0.16	(0.4)	0.17	(0.4)	0.16	(0.4)
<i>N</i>	1627		847		780	

Notes: The sample consists of the early training sample.

status. For the early training sample the corresponding table is seen in Table 2. We create a similar balance table for the more fine grained definitions of the treatment and as there are 21 different tests involved in comparing them to each other we will not judge the success of the randomization based on any single F-test in the table. A version of this table for the early training sample is provided in Table 3. We will create the same tables for the analysis sample.

Table 2: Test of balance.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Treated	Treated	Treated	Treated	Treated	Treated	Treated	Treated	Treated
Female	0.0049 (0.027)								0.0013 (0.028)
Birth year		0.00053 (0.0011)							0.00040 (0.0011)
Days before			-0.000057 (0.00016)						-0.000051 (0.00016)
Grade				-0.00013 (0.00047)					-0.00010 (0.00049)
nr employees					-0.00000055 (0.0000077)				-0.00000073 (0.0000078)
Symptoms						0.016 (0.026)			0.014 (0.027)
CW predicted meeting important							-0.00060 (0.029)		0.00034 (0.030)
CW predicted long spell								-0.012 (0.036)	-0.0087 (0.037)
Share of individuals treated	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52	0.52
No. of observations	1627	1627	1627	1627	1627	1627	1627	1627	1627
R-squared	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13	0.13
Controls	Block	Block	Block	Block	Block	Block	Block	Block	Block

Notes: All regressions control for block fixed effects (a combination of week of draw and caseworker). Robust SE in parentheses. The F-value from an F-test of the significance of all variables in column 9 is 0.02 ($p=0.8985$).

Table 3: Balance tests for all seven treatments

Variable	F1						F2						F3						F4						F5						F6						F7						F8						F9						F10						F11						F12						F13						F14						F15						F16						F17						F18						F19						F20						F21						F22						F23						F24						F25						F26						F27						F28						F29						F30						F31						F32						F33						F34						F35						F36						F37						F38						F39						F40						F41						F42						F43						F44						F45						F46						F47						F48						F49						F50						F51						F52						F53						F54						F55						F56						F57						F58						F59						F60						F61						F62						F63						F64						F65						F66						F67						F68						F69						F70						F71						F72						F73						F74						F75						F76						F77						F78						F79						F80						F81						F82						F83						F84						F85						F86						F87						F88						F89						F90						F91						F92						F93						F94						F95						F96						F97						F98						F99						F100						F101						F102						F103						F104						F105						F106						F107						F108						F109						F110						F111						F112						F113						F114						F115						F116						F117						F118						F119						F120						F121						F122						F123						F124						F125						F126						F127						F128						F129						F130						F131						F132						F133						F134						F135						F136						F137						F138						F139						F140						F141						F142						F143						F144						F145						F146						F147						F148						F149						F150						F151						F152						F153						F154						F155						F156						F157						F158						F159						F160						F161						F162						F163						F164						F165						F166						F167						F168						F169						F170						F171						F172						F173						F174						F175						F176						F177						F178						F179						F180						F181						F182						F183						F184						F185						F186						F187						F188						F189						F190						F191						F192						F193						F194						F195						F196						F197						F198						F199						F200						F201						F202						F203						F204						F205						F206						F207						F208						F209						F210						F211						F212						F213						F214						F215						F216						F217						F218						F219						F220						F221						F222						F223						F224						F225						F226						F227						F228						F229						F230						F231						F232						F233						F234						F235						F236						F237						F238						F239						F240						F241						F242						F243						F244						F245						F246						F247						F248						F249						F250						F251						F252						F253						F254						F255						F256						F257						F258						F259						F260						F261						F262						F263						F264						F265						F266						F267						F268						F269						F270						F271						F272						F273						F274						F275						F276						F277						F278						F279						F280						F281						F282						F283						F284						F285						F286						F287						F288						F289						F290						F291						F292						F293						F294						F295						F296						F297						F298						F299						F300						F301						F302						F303						F304						F305						F306						F307						F308						F309						F310						F311						F312						F313						F314						F315						F316						F317						F318						F319						F320						F321						F322						F323						F324						F325						F326						F327						F328						F329						F330						F331						F332						F333						F334						F335						F336						F337						F338						F339						F340						F341						F342						F343						F344						F345						F346						F347						F348						F349						F350						F351						F352						F353						F354						F355						F356						F357						F358						F359						F360						F361						F362						F363						F364						F365						F366						F367						F368						F369						F370						F371						F372						F373						F374						F375						F376						F377						F378						F379						F380						F381						F382						F383						F384						F385						F386						F387						F388						F389						F390						F391						F392						F393						F394						F395						F396						F397						F398						F399						F400						F401						F402						F403						F404						F405						F406						F407						F408						F409						F410						F411						F412						F413						F414						F415						F416						F417						F418						F419						F420						F421						F422						F423						F424						F425						F426						F427						F428						F429						F430						F431						F432						F433						F434						F435						F436						F437						F438						F439						F440						F441						F442						F443						F444						F445						F446						F447						F448						F449						F450						F451						F452						F453						F454						F455						F456						F457						F458						F459						F460						F461						F462						F463						F464						F465						F466						F467						F468						F469						F470						F471						F472						F473						F474						F475						F476						F477						F478						F479						F480						F481						F482						F483						F484						F485						F486						F487						F488						F489						F490						F491						F492						F493						F494						F495						F496						F497						F498						F499						F500						F501						F502						F503						F504						F505						F506						F507						F508						F509						F510						F511						F512						F513						F514						F515						F516						F517						F518						F519						F520						F521						F522						F523						F524						F525						F526						F527						F528						F529						F530						F531						F532						F533						F534						F535						F536						F537						F538						F539						F540						F541						F542						F543						F544						F545						F546						F547						F548						F549						F550						F551						F552						F553						F554						F555						F556						F557						F558						F559						F560						F561						F562						F563						F564						F565						F566						F567						F568						F569						F570						F571						F572						F573						F574						F575						F576						F577						F578						F579						F580						F581						F582						F583						F584						F585						F586						F587						F588						F589						F590						F591						F592						F593						F594						F595						F596						F597						F598						F599						F600						F601						F602						F603						F604						F605						F606						F607						F608						F609						F610						F611						F612						F613						F614						F615						F616						F617						F618						F619						F620						F621						F622						F623						F624						F625						F626						F627						F628						F629						F630						F631						F632						F633						F634						F635						F636						F637						F638						F639						F640						F641						F642						F643						F644						F645						F646						F647						F648						F649						F650						F651						F652						F653						F654						F655						F656						F657						F658						F659						F660						F661						F662						F663						F664						F665						F666						F667						F668						F669						F670						F671						F672						F673						F674						F675						F676						F677						F678						F679						F680						F681						F682						F683						F684						F685						F686						F687						F688						F689						F690						F691						F692						F693						F694						F695						F696						F697						F698						F699						F700						F701						F702						F703						F704						F705						F706						F707						F708						F709						F710						F711						F712						F713						F714						F715						F716						F717						F718						F719						F720						F721						F722						F723						F724						F725						F726						F727						F728						F729						F730						F731						F732						F733						F734						F735						F736						F737						F738						F739						F740						F741						F742						F743						F744						F745						F746						F747						F748						F749						F750						F751						F752						F753						F754						F755						F756						F757						F758						F759						F760						F761						F762						F763						F764						F765						F766						F767						F768						F769						F770						F771						F772						F773						F774						F775						F776						F777						F778						F779						F780						F781						F782						F783						F784						F785						F786						F787						F788						F789						F790						F791						F792						F793						F794						F795						F796						F797						F798						F799						F800						F801						F802						F803						F804						F805						F806						F807						F808						F809						F810						F811						F812						F813						F814						F815						F816						F817						F818						F819						F820						F821						F822						F823						F824						F825						F826						F827						F828						F829						F830						F831						F832						F833						F834						F835						F836						F837						F838						F839						F840						F841						F842						F843						F844						F845						F846						F847						F848						F849						F850						F851						F852						F853						F854						F855						F856						F857						F858						F859						F860						F861						F862						F863						F864						F865						F866						F867						F868						F869						F870						F871						F872						F873						F874						F875						F876						F877						F878						F879						F880						F881						F882						F883						F884						F885						F886						F887						F888						F889						F890						F891						F892						F893						F894						F895						F896						F897						F898						F899						F900						F901						F902						F903						F904						F905						F906						F907						F908						F909						F910						F911						F912						F913						F914						F915						F916						F917						F918						F919						F920						F921						F922						F923						F924						F925						F926						F927						F928						F929						F930						F931						F932						F933						F934						F935						F936						F937						F938						F939						F940						F941						F942						F943						F944						F945						F946						F947						F948						F949						F950						F951						F952						F953						F954						F955						F956						F957						F958						F959						F960						F961						F962						F963						F964						F965						F966						F967						F968						F969						F970						F971						F972						F973						F974						F975						F976						F977						F978						F979						F980						F981						F982						F983						F984						F985						F986						F987						F988						F989						F990						F991						F992						F993						F994						F995						F996						F997						F998						F999						F1000						F1001						F1002						F1003						F1004						F1005						F1006						F1007						F1008						F1009						F1010						F1011						F1012						F1013						F1014						F1015						F1016						F1017						F1018						F1019						F1020						F1021						F1022						F1023						F1024						F1025						F1026						F1027						F1028						F1029						F1030						F1031						F1032						F1033						F1034						F1035						F1036						F1037					
----------	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	-----	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--	-------	--	--	--	--	--

Notes: The value displayed for t-tests are the difference in the means across the groups. The value displayed for F-tests are the F-statistic. All regressions control for block fixed effects (a combination of week of draw and caseworker). ***, **, and * indicate significance at the 1, 5, and 10 percent critical level.

5 Empirical strategies

5.1 Regression analysis

We will start by estimating the effects of mandatory versus voluntary meetings on the total number of days the absentee is absent after the draw. In doing so, we estimate versions of the following linear regression model:

$$Total_i = \beta DM_i + \theta Block_{cw} + \lambda X_i + \varepsilon_i \text{ (eq. 1)},$$

where DM_i is a dummy that equals one for being summoned to a mandatory dialogue meeting for individual i and zero otherwise, $Block_{cw}$ are fixed effects for the strata (a combination of caseworker (c) and week (w) fixed effects). These blocking variables are included in all regressions as randomization takes place within these strata. In particular, the assignment of cases to caseworkers is not always random, the treatment share and intensity varies over time, and the types of absentees are likely to change over seasons. X_i are baseline controls that can be added to possibly increase power. The baseline controls were chosen by testing different specifications in the early training data. These are: a dummy variable for female, dummy variables for age in 5 year groups, the number of days of sickness absence since 2015 and the number of employees at the workplace of the absentee as a continuous variables, and dummies for graded sickness absence in categories of 10. We will present results with and without the baseline controls. We use robust standard errors in all estimations.

The standard errors do not need to be clustered at any level as the randomization is at the individual level (see Abadie et al. (2017)).

The same specification is also run for the secondary outcome variables, i.e. the number of days within the sickness spell that an individual is absent, $Days_i$, and the grade adjusted total number of days, $Graded\ days_i$. We will later on create a dummy variable for temporary disability insurance (TDI) and run a linear probability model with TDI as outcome. We also test for different treatment effects by replacing DM_i with dummies for different mandatory meeting weeks.

We will also conduct tests where we run regressions for people with letters in the same week but with different content. That is, we will restrict the sample to people getting letters in week 9 and test whether the letter is one for a mandatory meeting in week 13 or for a voluntary meeting. The same type of analysis will be done for individuals receiving letters in week 22. For individuals with letters in week 15 we will run three different regressions: i) Excluding those with a mandatory meeting in week 19; ii) Excluding those with a mandatory meeting in week 26; and iii) Including both groups with mandatory meetings but with separate dummy variables.

We will also include all seven treatments in the same regression. To test for the independent effect of receiving voluntary letters and the timing effects of mandatory letters we will run regressions separately for people with voluntary and mandatory letters. If there is an independent reversed notification effect we expect that people getting voluntary letters early will have longer absence than people getting voluntary letters later.

To explore heterogeneity we will first interact the treatment variable(s) with the baseline control variables. For ease of interpretation we dummy code the Grade variable into 1 for people with any amount of graded sickness absence and zero for those with 100 percent sick leave. We call this variable Graded. We retain the dummy coding for female and transform the continuous variables so that they are in standardized form. We will further

split the sample by gender and into symptoms or diagnoses (i.e. symptoms=0) and we will run separate regressions for each main diagnosis (as defined by the letter in the ICPC-2 or a dummy for Multiple diagnoses in cases where the individual has several diagnoses). The sample will be split by caseworker predictions (long spell and importance of meeting), by offices as well as by occupations. We also apply machine learning techniques to search for heterogenous treatment effects more systematically (see Section 5.3).

5.1.1 Identification of threat effects

To identify the different types of threat effects we will use the dummy variables for having returned before week X (13, 19, 22) and regress this on the different sets of dummy variables for the treatments. To start with, we will run the following linear probability model:

$$\textit{Return before week 13}_i = \beta \textit{DM 13 letter 9}_i + \gamma \textit{Letter9}_i + \theta \textit{Block}_{cw} + \lambda X_i + \varepsilon_i \text{ (eq. 2)},$$

where *Return before week 13_i* is a dummy for returning before week 13, *DM 13 letter 9_i* equals 1 if the individuals are randomly assigned to a mandatory meeting in week 13 with a notification letter in week 9, and *Letter9_i* is a dummy for being randomly assigned to a notification of a voluntary meeting in week 9. In this specification, the ones receiving the early letters are compared to everyone not having received the letter yet. This allows for a direct test of the threat effect as well as the reversed threat effect. More specifically, if β is positive, there is a threat effect whereby calling people in to a mandatory meeting induces them to end their spells faster without a meeting having taken place. Similarly, if γ is negative, there is a reversed threat effect whereby telling people that a meeting is voluntary for them reduces the probability that they end their sickness spell in the following weeks.

The same specification can also be run with the dependent variable being *Return before week 19_i* and including dummies for *DM 19 letter 15_i*, *DM 26 letter 15_i*, and *Letter15_i*. The coefficient for the early letter with meeting in week 13 will then not be interpretable as a threat effect as people will have started to have meetings at this point. Rather, the

coefficient can be interpreted as a combination of the effect of being assigned to a meeting (threat effect) and a short run early meeting effect. We also run the same regression with the outcome variable being return before week 22. The excluded group still consists of people not having received any letter but as meetings have started also for the *DM 19 letter 15_i* group, only the *DM 26 letter 15_i* coefficient can be interpreted as a threat effect. The coefficient for *DM 19 letter 15* can now be interpreted as a combination of a letter effect and a short run effect of having a meeting in week 19. The coefficient for *DM 13 letter 9* can now be interpreted as a total effect of early letter and a medium term effect of meetings (as 9 weeks have passed since the intended dates of the meetings). We will investigate heterogeneity in threat effects by interacting the treatment variables with baseline covariates.

5.2 Duration analysis

We will conduct duration analyzes to investigate transition rates in the weeks the letters are received and in the weeks leading up to the meeting. These analyzes will allow us to test even finer predictions about the different types of letters. These methods will also shed light on the separate effects of the conducted meetings and the effects of the letters.

In the duration analysis we put more structure on the timing of the relationship between being called in to a meeting and returning to work. For this purpose, we start by restructuring the dataset. For each week in which the absentee is under risk for returning to work (ending the sickness spell) the spell is given an additional line in the dataset. The outcome variable is then whether or not the absentee actually returns to work this week. The advantage of this approach is that it enables us to test more specific behavioral hypotheses, such as *when* workers eventually respond to a letter (e.g. same week, the week after etc.), and *when* the meetings actually impact return to work, if at all, e.g. before the meeting (threat effect), same week or after the meeting. All models are estimated using Cox proportional hazards models (stcox in STATA). This is a partial likelihood model, implying that we do not estimate the underlying hazard rate or duration dependence. We thus do not impose

any parametric assumptions on the shape of the hazard function over the duration. The models are estimated using the same baseline covariates (and the block fixed effects) as in the linear models above and we also add dummies for the seven treatment groups.

As these models use the exact timing of the letters and the meetings there will be measurement errors whenever the meetings and letters are not held/sent at the times they are supposed to be. To the extent that the errors in sending of letters or timing of meetings are uncorrelated with the treatment received this will bias our coefficients towards zero. While we have no way of checking when the letters are sent, we can and will investigate when the meetings are held.¹

We estimate 5 different models:

1) First we investigate whether receiving a letter impacts immediate return to work. To do so we construct a time varying variable which start out being equal to 0. One week before the letter arrives the variable is changed to 1. The same week as the letter arrives it equals 2, and the week after it takes the value 3. Then, for two weeks after and onwards it is again equal to 0. To separate between letters actually summoning for a meeting (the treatment group), and letters stating that a meeting is voluntary (the control group) we interact the variable with a dummy for the treatment group.

2) We move on to compare early and late letters. We now ask whether the timing of the letter has any importance for its effect. We do so by constructing another time-varying variable taking the value of 1 the week the letter arrives and 0 otherwise. This variable is then interacted with a dummy for the week in the spell the letter is arriving (9, 15 or 22) and treatment.

3) In the third model, we separate between three types of letters and investigate their effect in the week the letter arrives: those for a voluntarily meeting, those saying that there

¹In the accompanying draft we present some descriptive statistics on meeting dates for our different treatments.

is a meeting in four weeks from now and those saying there is a meeting 11 weeks from now.

4) We then set up a model in which we can estimate effects of letters as well as meetings. The model contains a dummy variable indicating that the individual has received a letter. This variable is turned on the week the letter is sent, and never turned off again. It also includes a dummy indicating that you receive a letter this week. This variable takes the value 1 in the letter-week, and zero otherwise. The model also includes an interaction between mandatory meeting (DM) and having received a letter as well as an interaction between DM and being in the letter week. The model further includes a set of dummy variables for being 1, 2, or 3 weeks before the meeting, a dummy for the meeting week and a dummy for all the weeks after the meeting. Analyzing the independent effects of the meetings is tricky as this involves mediation analysis. The main problem is that the only ones having meetings constitute a selected sample of individuals not induced to end their spell by receiving the letter. The variables for having received the letter and its interaction with DM are included to mitigate the potential selection problem arising from the fact that there is a positive/negative effect on return to work from receiving a treatment/control letter.

5) Finally we want to investigate whether being called in to meetings have the largest effect early or late in the spell. We construct a variable which takes the value 1 the week before and same week as the meeting is held, and zero otherwise. This variable is then interacted with the timing of meeting (i.e. week 13, 19, or 26).

We will also estimate these models for different sub-groups.

5.3 Uncovering heterogeneity using machine learning techniques

We also use machine learning techniques to automate the search for heterogeneous treatment effects. There are many different types of machine learning algorithms and we have decided to use classification and regression trees (R package *causalTree*, (Athey and Imbens, 2016)); and random forests (R package *grf*, (Wager and Athey, 2017)). As this field is moving rapidly, however, it is possible that there will be other techniques that are relevant

for us once we start analyzing the data.

Both techniques start by splitting the sample in two partitions. We will split our analysis sample (i.e. early test sample + late sample) in two random subsets of equal size. We stratify the random draw using groups in which individuals have constant treatment probabilities and we set the seed=1987. We name the two partitions $S1$ (for the machine learning training sample) and $S2$ (for the machine learning test sample).

Our main analysis will use *Total* as outcome variable, and *DM* as treatment. We will also test different treatments, and in particular heterogeneity in the threat and reversed threat effects, and whenever this includes different samples we will draw $S1$ and $S2$ again (e.g. if we investigate heterogeneity in the effects for those treated with early versus late letters in the group receiving mandatory meetings). We will include as covariates all the variables mentioned in the paragraph *Other variables* in section 4, and also a dummy for each separate diagnosis. Continuous covariates are standardized.

One of the advantages of machine learning techniques is that they are able to handle a very large number of covariates, even more covariates than the number of observations. For that reason we will also include in the set of covariates any other observables that we will be able to collect (for example by merging our data with register data) and that we deem potentially interesting. All analyses will, however, also include the covariate set explicitly specified here.

5.3.1 Honest causal tree

We will estimate conditional average treatment effects (CATEs) using the honest causal tree by Athey and Imbens (2016), with the R package `causalTree`. We will proceed as follows:

1) We grow the tree using only $S1$ with the R function `causalTree` with options `split.Rule = "CT"`, `split.Honest = TRUE`, `cv.option = "CT"`, `cv.Honest = TRUE`. We take into account that we have different treatment probabilities by passing to the routine the ex-ante

probability of being treated in different strata. We want to grow a tree with a manageable number of leaves, so we set the minimum number of treated and control observations in each leaf that is split (`minsize`) to be approximately equal to 5% of the sample size in $S1$. We select the number of cross validations equal to 100 (`xval`).²

2) We prune the tree (cut back leaves) by selecting the complexity parameter corresponding to the minimum normalized cross validation error (`xerror`). Note that each leaf of the pruned tree is defined by the same values of given covariates (for example women born after 1978.) The routine calculates the treatment effects in each terminal leaf. We name this tree dishonest tree.

3) To make the procedure what Athey and Imbens (2016) call honest we use one sample for growing the tree and another sample for testing the predictions. That is, we classify each observation in $S2$ as belonging to one of the terminal leaves defined by the tree grown at point 2). Using only $S2$, we estimate the average treatment effect in each leaf, which is called the conditional average treatment effect (CATE, where conditional refers to conditional on being in a given leaf). We name this tree honest tree.

4) Using only $S2$, we regress (OLS) the outcome on dummies for each terminal leaf, their interactions with the treatment and block fixed effects (without constant nor treatment included). We use robust standard errors. This regression recovers the same CATEs as in the previous point, but with standard errors.

5) We consider evidence of treatment heterogeneity if we find statistically significant treatment effects in one or more leaves in $S2$, and if in these leaves the CATE has the same sign in both the honest and in the dishonest tree.

²As recommended by the authors we also use the option `split.Bucket = TRUE` with parameters `bucketNum = minsize*3`, `bucketMax = 20`.

5.3.2 Honest causal forest

We will estimate conditional average treatment effects (CATEs) using an honest causal forest by (Wager and Athey, 2017) with the package `grf`. We proceed as follows:

1) Using the R function `causal_forest` we grow the forest (ensemble of trees) using only $S1$. Each tree uses only a random fraction (50%) of $S1$ (`sample.fraction = 1/2`). This sub-sample will be further cut in half at random to implement honest estimation (grow tree in one half, and estimate CATE in the other half). The estimated CATEs for each leaf and each tree are stored. We set the following parameters:

- Number of trees: at least 25000 (and if feasible 50000).
- Number of covariates used in each tree: half the total number of available covariates.
- Penalty parameter `lambda=0`, and `downweight.penalty=FALSE`
- Number of trees for each subsample (to construct confidence intervals): 50 if computation time allows (based on simulation it should take around 30 minutes). Otherwise we reduce it until feasible (minimum is 2).
- `precompute.nuisance = FALSE`
- `min.node.size = 4`
- `alpha = 0.05`
- `seed = 1987`
- `sample.fraction = 1/2`

2) Using the R function `predict` we predict the CATEs in $S2$ based on the estimates from point 1). We report a histogram, summary statistics and a variance plot of these CATEs.

We also report the fraction of observations that have a CATE significantly different from 0 at the 10% level.

3) To assess whether the forest detected true heterogeneity, we create a dummy for observations in *S2* that have a positive CATE, and one for those that have a negative one (obtained in point 2)). We regress the outcome on these dummies, their interactions with the treatment, and block fixed effects. If the treatment heterogeneity is not spurious, this regression will estimate a positive treatment effect in the group of observations with positive CATEs, and negative treatment effects for the other group. We will also test if the two are significantly different from each other.

4) We will do the same as in the previous point but splitting observations in *S2* at the median of CATE instead of at CATE=0, as well as in quartiles of CATE.

5) If evidence from the previous two points suggests that the treatment heterogeneity is not spurious, we will characterize observations with positive and negative CATEs and in different quartiles of CATEs inspecting summary statistics of their covariates in these different groups.

6 Main results in the early training sample

In the companion document "Effects of Dialogue Meetings - results from the early training sample" we present results from the specifications described here as well as some additional analyzes.

The main results from the early training sample are that the effect of mandatory as opposed to voluntary meetings seems to be negative but the effect is not statistically significant. There seems to be substantial effect heterogeneity with respect to the caseworkers' predictions of long meetings whereby those that are predicted to be away for a long period have a much larger effect of being called in to a mandatory meeting. Interestingly, we find less evidence of heterogeneity when using the caseworkers predictions for the effects of the

meetings.

The duration analysis suggests that there is a threat effect such that individuals are more likely to end the sick leave spell in the week they get the letter and in the week before the meeting. There also seems to be a reversed threat effect whereby individuals called in to a voluntary meeting are less likely to end their sick-leave spell in the week they get the letter. We also observe this reversed threat effect when running a regression of returning before week 13 on receiving a letter about a voluntary meeting in week 9.

The causal tree analysis suggests that there is heterogeneity with respect to the caseworker predictions of long spells but in particular for individuals with a higher share of full-time absence (as opposed to part time absentees). The results from the causal forest are inconclusive at this point but we expect this method to be more powerful once we receive more data.

7 Hypotheses to test in the test sample

When we get outcome data and covariates for the early test sample we will run all the analyzes for this sample that we ran in creating the companion document "Effects of Dialogue Meetings - results from the early training sample". The machine learning algorithms will use the early training sample as $S1$ and the early test sample as $S2$.

We will then compare the results and draw more reliable conclusions about the different effects. The conclusions we draw will be based on the full picture of the results but if we find the same results as those described in Section 6 with respect to effect heterogeneity by caseworker predictions of long spells, threat effects and reversed threat effects, we will suggest some policy changes.

The proposed changes in policy will be made in close cooperation with NAV and they have the final word on any change. We expect two relatively easily implementable changes to the design already. If the result of a reversed threat effect of sending out letters of voluntary

meetings hold up, a suggested policy change is to send these types of letters out later (if they have to be sent at all). If the results that the meetings have larger effects for people with long predicted spells hold up, these individuals can be targeted to have a higher probability of mandatory meetings.

Irrespective of the results in the early test sample we are also having discussions about other types of policy changes. We are discussing changes to the letters to see if the nudging aspect can be improved. We are also discussing ways to test different types of meetings by conducting some targeted experiments. For instance, do meetings work better when the physician attends or not? We are currently surveying the caseworkers and conducting interviews with physicians in order to derive more focused hypotheses here.

If a change is implemented we will register an additional plan detailing how that change will be tested and we will document how the change affects our analysis sample.

8 Hypotheses to test in the analysis sample

When we decide that the current phase of the experiment is over we will have NAV add data on outcomes and covariates for the full analysis sample. On that sample we will conduct all the analyzes specified in this plan. Other analyzes not specified here nor in later registrations will be seen as more explorative. As there are many hypotheses and tests in this pre-registration we here outline what we view as our main hypotheses.

Hypothesis 1: Mandatory meetings will cause a shorter total absence spell than voluntary meetings. Support of this hypothesis would be a statistically significant negative coefficient for DM_i in the OLS regression of $Total$, controlling for the specified baseline covariates and block fixed effects (equation 1).

Hypothesis 2: There is a threat effect of letters calling in to mandatory meetings. This can be tested in different ways. As we saw in the duration analysis individuals randomly assigned to mandatory meetings may be more likely to end their spell in both the letter

week and in the week before meeting. We will, however, use equation 2 and view a statistically significant positive coefficient for *DM 13 letter 9_i* as supporting the hypothesis.

Hypothesis 3: There is a reversed threat effect of voluntary letters. Also this hypothesis can be tested in different ways and we saw indications in the duration analysis of a reversed threat effect by people being less likely to end their spell in the week the letter arrives if the letter called in to a voluntary meeting. Again, however, we will use equation 2 and view a statistically significant negative coefficient for *Letter9_i* as supporting the hypothesis.

The remaining hypotheses in this plan can be seen as secondary hypotheses. With respect to the heterogeneity in treatment effects we expect that the machine learning techniques will give a clear and consistent answer. These techniques are in some sense self validating as they are based on sample splits. We will nonetheless always include the covariates specified here.

9 Power analysis

We have calculated minimum detectable effects (MDE) under different assumptions. On February 9 2018 we have a total of 11,196 individuals entered into the system. We have already used 1,871 for the early training sample. A lower bound for the analysis number of observations in the analysis sample is therefore 9,325 before dropping any individual. In the early training sample we had to drop 13 percent of the observations since they had returned from their sick-leave spell already at the date of the draft. If we assume a 15 percent drop rate in the analysis sample and round downwards we will have a sample of 7,900.

The minimum detectable effect for a sample of 7,900 with a significance level of 0.05 is 0.063. That is, we would be able to find effects as small as 0.063 standard deviations of the dependent variable. As seen in Table 1 the standard deviation for *Total* is 90 days. Hence, we would be able to identify effects as small as 5.67 total days. We further know that the R-squared in a regression including *DM* and the baseline covariates is 0.22. The corresponding MDE if we take this into account is 0.056 (which corresponds to 5.04 total

days).

As we are testing multiple hypotheses we will also adjust the p-values. We have already anticipated the adjustment and therefore we have restricted the number of main hypotheses to test to be only three. To account for having three different hypotheses we follow the recommendations of Fink, McConnell, and Vollmer (2014) and use a method developed by Benjamini and Hochberg (1995) and Benjamini and Yekutieli (2001) to minimize the false non-discovery rate (see also Almeida (2012) and Finseraas and Kotsadam (2017) for pre-analysis plans with the same decision rules for correction of p-values). The main advantage of the method is that it is limiting the risk of false discoveries while only adjusting the critical values based on other true hypotheses. The false discovery rate method developed by Benjamini and Hochberg (1995) implies that the m p-values of the i hypotheses are ordered from low to high and that the critical value of the p-value is then $p(i) = \alpha \cdot i/m$. To take a concrete example, with 3 hypotheses and a significance level (α) of 0.05, the critical p-value would be 0.017 for the one with the lowest p-value ($0.05 \cdot 1/3$, which is the same as a Bonferroni correction). For the second and third hypotheses, the critical p-values would be 0.033 ($0.05 \cdot 2/3$) and 0.05 ($0.05 \cdot 3/3$). With a p-value of 0.017 our calculated MDE is 0.074 (0.064 accounting for an R-squared on 0.22). We expect the sample size to be larger but we conclude that even with the lowest bound of the sample size, the experiment is well powered to detect small effects.

10 Archive

The pre-analysis plan is archived before the early test sample is matched with any outcome variables. We archive it at the registry for randomized controlled trials in economics held by The American Economic Association: <https://www.socialscienceregistry.org/> on February 12 2018. We will receive the matched early test sample on February 13. The experiment is still running and it is unclear when the analysis sample will be finished. We

will test the results from the early training sample in the early test sample and we may file a separate analysis plan with more registered decisions taken after this point.

References

- Abadie, Alberto et al. (2017). *When Should You Adjust Standard Errors for Clustering?* Tech. rep. National Bureau of Economic Research.
- Almeida, Rita et al. (2012). *The Impact of Vocational Training for the Unemployed in Turkey: Pre-Analysis Plan*. Pre-analysis plan posted at povertyactionlab.org.
- Anderson, Michael L and Jeremy Magruder (2017). *Split-sample strategies for avoiding false discoveries*. Tech. rep. National Bureau of Economic Research.
- Athey, Susan and Guido Imbens (2016). “Recursive partitioning for heterogeneous causal effects.” In: *Proceedings of the National Academy of Sciences* 113.27, pp. 7353–7360.
- Benjamini, Yoav and Yosef Hochberg (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing.” In: *Journal of the royal statistical society. Series B (Methodological)*, pp. 289–300.
- Benjamini, Yoav and Daniel Yekutieli (2001). “The control of the false discovery rate in multiple testing under dependency.” In: *Annals of statistics*, pp. 1165–1188.
- Coffman, Lucas C and Muriel Niederle (2015). “Pre-analysis plans have limited upside, especially where replications are feasible.” In: *Journal of Economic Perspectives* 29.3, pp. 81–98.
- Coffman, Lucas C, Muriel Niederle, and Alistair J Wilson (2017). “A Proposal to Organize and Promote Replications.” In: *American Economic Review* 107.5, pp. 41–45.
- Fafchamps, Marcel and Julien Labonne (2017). “Using Split Samples to Improve Inference on Causal Effects.” In: *Political Analysis* 25.4, pp. 465–482.
- Fink, Günther, Margaret McConnell, and Sebastian Vollmer (2014). “Testing for heterogeneous treatment effects in experimental data: false discovery risks and correction procedures.” In: *Journal of Development Effectiveness* 6.1, pp. 44–57.

- Finseraas, Henning and Andreas Kotsadam (2017). “Does personal contact with ethnic minorities affect anti-immigrant sentiments? Evidence from a field experiment.” In: *European Journal of Political Research* 56.3, pp. 703–722.
- Mullainathan, Sendhil and Jann Spiess (2017). “Machine learning: an applied econometric approach.” In: *Journal of Economic Perspectives* 31.2, pp. 87–106.
- Olken, Benjamin A (2015). “Promises and perils of pre-analysis plans.” In: *Journal of Economic Perspectives* 29.3, pp. 61–80.
- Varian, Hal R (2014). “Big data: New tricks for econometrics.” In: *Journal of Economic Perspectives* 28.2, pp. 3–28.
- Wager, Stefan and Susan Athey (2017). “Estimation and inference of heterogeneous treatment effects using random forests.” In: *Journal of the American Statistical Association* Forthcoming.