# Can compulsory dialogues nudge sick-listed workers back to work?

Simen Markussen

Ragnar Frisch Centre for Economic Research.

Knut Røed

Ragnar Frisch Centre for Economic Research and IZA.

Ragnhild C. Schreiner

Ragnar Frisch Centre for Economic Research and the University of Oslo.

# ABSTRACT

# Can compulsory dialogues nudge sick-listed workers back to work? *

We evaluate the impacts of a compulsory dialogue meeting for long-term sick-listed workers in Norway. The meeting is organised by the local social security administration after around six months of absence, and its purpose is to bring together the absentee, the employer, and the family physician to discuss whether arrangements can be made to facilitate partial or full work resumption. Our causal analysis is based on random-assignment-like geographical variation in the meeting propensity. We find that the meetings reduce absence duration considerably, both through a notification and an attendance effect. They also reduce the risk of premature labour market exit.

Corresponding author:
Ragnhild C. Schreiner
The Ragnar Frisch Centre for Economic Research
Gaustadalléen 21
0349 Oslo
Norway
E-mail: r.c.schreiner@frisch.uio.no.

---

# 1  Introduction

Programs insuring workers against income losses during sickness absence are of major importance in many welfare states. These programs are typically highly valued by workers as well as voters. Yet, researchers have found evidence of serious moral hazard problems: The higher and more long-lasting is the sick-pay, the higher is the rate of sickness absence (Johansson and Palme (1996); Henrekson and Persson (2004)). Higher absenteeism may in turn imply a greater risk that some absence spells develop into more serious long-term disabilities with low likelihood of successful return to regular employment. Policy makers therefore face a difficult trade-off, and are (or should be) interested in measures that can reduce the moral hazard problems associated with sick-leave insurance, such that the desired level of insurance can be provided at the lowest possible cost. This paper examines the effects of one such policy, namely a compulsory dialogue meeting (DM) for long-term sick-listed workers in Norway.

Like in most other industrialised countries, Norway's sick-leave insurance system entails moral hazard problems with respect to the behaviors of both employers and employees. The employee receives a 100% wage replacement (up to a relatively high ceiling; see details in next section) from the first day of absence and for up to one year. The cost the first 16 days of each absence spell is fully covered by the employer , after which the public purse takes over the bill. There is no experience rating. This insurance design (which with some variations is typical for European OECD countries (see OECD (2010)) may imply that employees exert less effort to avoid and escape from absenteeism than what is optimal from a social planner's point of view, whereas the employers exert too little effort to facilitate work resumption and prevent the employee from losing the foothold in the labour market. One way to reduce the moral hazard problems is to monitor the behavior of employers and employees. The main device used for this purpose is a requirement that absence spells exceeding three days (eight days in some firms) are certified by an authorised physician.[1] However, it is well known that physicians may be rather poor gatekeepers, not only because their health

---

[1]If the employer is part of the "Inclusive Workplace Agreement" (IA), workers can self-certify absence lasting up to eight days. This is the case for approximately 50% of Norwegian employees.

assessments largely are based on what the patients choose to tell them, but also because they have financial incentives to ensure that their patients are sufficiently satisfied to not find a new physician.(Markussen et al. (2013)) Thus, additional monitoring mechanisms may be required. Given that the financial cost of continuing a sick-pay spell a few extra days may be negligible for both the absentee and the employer, it is possible that even very small hurdles - e.g. the need for an unpleasant conversation between the two of them - are sufficient to repeatedly postpone work resumption. Procrastination in intertemporal effort choices is widespread, and a growing empirical literature indicates that many individuals discount the future in a hyperbolic fashion; i.e., with a bias toward the present (DellaVigna and Paserman (2005); Paserman (2008); Cockx et al. (2014)). This implies that activities for which future benefits must be weighed against immediate costs tend to be postponed repeatedly, even when the activities are optimal from a long-term perspective. Hyperbolic employers and employees will always be tempted to delay unpleasant work resumption efforts yet another few days. In such cases a small "nudge" may be all that is needed to speed things up.

The meeting evaluated in this paper may be viewed as a combined nudging and monitoring device. It is organised by the local social security administration (SSA) around six months into the sick-leave spell, and its main purpose is to bring together the absentee, the employer and the certifier of the sick-leave to discuss whether arrangements can be made at the workplace that make full or partial work resumption possible. Examples of such measures can be to alter the number or nature of tasks at work, to implement a home office, or to adjust work hours. An important feature of the Norwegian sick-pay insurance system is that it allows partial (graded) absence, implying, for example, that a sick-listed employee works at a 50% capacity (and receives the normal wage from the employer for this part), while collecting 50% sick-pay benefits. Previous evidence has indicated that promoting partial rather than full absence is a fruitful strategy toward reducing overall absenteeism (Markussen et al. (2012)), and the sick-pay legislation actually requires physicians to issue partial absence certificates for all absences exceeding eight weeks, unless there are strong medical grounds for maintaining 100% absence. Yet, the majority of absence spells are not graded, and the DM thus provides an arena where the social security administration can push for partial as well as full work resumption.

The questions we seek to answer in this paper are whether or not the DM has achieved its aim of speeding up the process of partial or full work resumption, and - if so - whether it has done so in a cost-efficient way. In addition, we are interested in identifying the mechanisms behind any DM effects. More concretely, we distinguish between "notification effects" and "attendance effects". The DM is typically summoned three weeks before it is supposed to take place through letters from the social security administration to the sick-listed employee, the employer, and (if deemed appropriate by the caseworker) the sick-leave certifier (the physician). If the sick-leave ends after the letters have been sent, but before the scheduled time of the meeting, the meeting will normally be canceled. The notification effect encompasses the responses triggered by the summons letters before any meeting actually takes place. It bears some resemblance to the notorious "threat effect" frequently encountered in the unemployment insurance literature with reference to the effect of being summoned to an activation program (Black et al. (2003); Rosholm and Svarer (2008); Geerdsen (2006); Graversen and Larsen (2013)). In our case, however, the notification effect not only refers to the behavior of the insured worker, but also to the behaviors of the employer and/or the physician. The attendance effect encompasses the impacts following from the meeting itself, e.g., in terms of an agreed strategy for work resumption.

The DM was introduced in Norway in 2007, but there is no data on meetings actually held until 2009. Our analysis therefore includes spells starting between January 2009 and December 2010, and it is based on complete administrative registers for all physician-certified absence spells in this period. To identify causal impacts of the DM, we exploit that even though the meeting is in principal compulsory, there has been ample scope for local social security administrations to make exemptions. This has resulted in a considerable geographical variation in both the overall use of DM's and in their precise timings within absence spells. We will argue that from the workers' and firms' points of view, this gives rise to a random-assignment-like variation in the duration-specific probabilities of being called to a DM, and, as we show below, this makes it possible not only to identify the effects of the meeting as such, but also to distinguish notification from attendance effects.

Our statistical analysis consists of two parts. The first part examines the effects of the

DM on sick-leave duration, based on a mixed proportional hazard rate framework. It aims at distinguishing notification from attendance effects by including in the return-to-work hazard one variable representing the probability of being called to a meeting and another representing the probability of having already (and recently) participated. In this part of the analysis, we either treat any form of work resumption (partial or full) as the outcome of interest, or we treat full work resumption as the outcome of interest while disregarding partial work resumption. The second part of the analysis examines the effects of DMs on long-term outcomes, such as future employment and earnings. In this part, we do not distinguish notification from attendance effects.

Our study relates to Johansson and Lindahl (2013), who examine the effects of an information meeting (IM) on the duration of absence spells of (largely) non-employed sick-pay claimants; i.e., persons who have already lost the job that originally made them eligible for sickness benefits. The IMs are organised by Swedish local SSA offices, and the purpose of the meeting is to inform absentees about the criteria for continued sickness benefits. The study is based on data where the timing of the call to the IM is randomised. The results from the evaluation of the IMs suggest a significant positive effect on the exit hazard from sick-leave. However, for these persons, the alternative to sick-pay will often be unemployment insurance, and the study indeed shows that the increase in outflow from sick-leave is partly met by an increase in the inflow to registered unemployment. By contrast, our own study focuses on employed sick-listed workers where the alternative to continued absence is partial or full work resumption.

Our main finding is that the DM has a positive and substantial effect on the hazard to partial as well as full work resumption. Both notification and attendance effects contribute to earlier work resumption, and based on a simulation exercise, we find that the two effects are of roughly the same quantitative importance. Together, the notification and attendance effects estimated in this paper imply that for each realised DM, the duration until partial or full work resumption is reduced by approximately 20 calendar days (including weekends). Given the low cost of arranging the meetings, we also find that the meetings are highly cost effective, in the sense that the value of the additional work they generate - as measured by

earnings (which are again equal to the saved insurance payouts) is much larger than the meeting costs. Moreover, when estimating whether the meeting has an effect on long-term outcomes such as employment status, income and benefit dependency around two years after the DM, the results consistently point towards favorable long-term effects. The long-term effects are small in magnitude. However, this is as expected since the DM is a relatively modest policy intervention. Overall, we conclude that an activation policy in terms of a mandatory meeting some time into sick-leave spells is an efficient policy measure to reduce long-term absence and potential disability inflow.

# 2 Data and institutional Setting

Conditional on receiving a certificate from a physician, all absentees are entitled to a 100% wage replacement ratio for up to one year. The replacement ratio is 100% up to a ceiling of six times the base amount in the Norwegian pension system. The base amount is adjusted every year, and was equal to 85,245 NOK in 2013. Using the average exchange rate for 2013, this corresponds to approximately 14,500 USD.[2] A major challenge with the Norwegian system is the lack of economic incentives for absentees to return to work during the first year on sick-leave benefits, as well as the weak incentives for the employer to promote work resumption. The employer pays for the first 16 days of the absence spell whereas the social security administration (SSA) covers the benefit payments thereafter. This means that the social security system undermines the employer's economic incentives to exert effort in helping long-term absentees back to work. As a matter of fact, there is empirical evidence that employers sometimes discourage long-term absentees from returning to work due to the risk of the absentee starting a new spell shortly thereafter (Fevang et al. (2014)). Moreover, if the absence spell outlast the one-year sick-pay period, a separate employment protection for absent workers no longer applies, implying that the employer can legally fire the worker without having any responsibility for subsequent social insurance payments.

---

[2]All monetary amounts in this paper are inflated to 2013 value, based on the adjustment factor used in the Norwegian pension system. The translations to US dollars are based on the average exchange rate applying in 2013, such that 1 USD=5.875 NOK.

Several measures, the DM included, have been implemented to counteract these incentives. Their purpose is both to facilitate the return to work and to monitor that the absentees reveal their actual need for absence. For employed workers, the employer is obliged to arrange a meeting with the employee during the first seven weeks of the spell to agree on a plan for work resumption, if necessary involving changes in tasks or work hours. No later than 26 weeks into the spell, the SSA is supposed to organise a DM. The meeting was introduced in 2007 to induce long-term absentees to fully or partly return to work, as opposed to continuing on a path to disability insurance dependency. The DM is intended to provide a setting where the employer and the absentee can discuss possible measures to be made at the workplace to make work resumption possible. The certifier of the sick-leave should attend the meeting if the caseworker at the SSA office considers it to be appropriate. Even though the local SSA offices are required by the law to arrange a DM within 26 weeks of the absence spell, there is significant variation between counties in the frequency and timing of the DM. There are 19 counties in Norway with an average of 24 SSA offices within each county. The county administration face binding budgets and this can result in considerable differences in the extent of use of different labour market programs, including the DM. Failing to organise a DM can be justified as an exemption, something the law actually gives room for. The law states that an exemption is allowed if "such a meeting is assumed to be clearly unnecessary" (Folketrygdloven §8-7,8 a).[3] As it turns out, this "exemption option" has been interpreted and exploited very differently in different parts of the country, which - viewed from the sick-listed workers' point of view - constitutes a variation in the likelihood of being exposed to a meeting which is as good as randomly assigned.

Figure 1 illustrates the variation across counties in DM exemption practices, as reflected in the share of all long-term absence spells (longer than five months) started between January 2009 and December 2010 for which a realised DM is recorded. The share varies from around 13-18% in Østfold and Oslo (the capital) to more than 30% in Vestfold and Telemark.[4]

---

[3]The most common reasons for making exemptions are if the absentee is severely ill and/or admitted to a health institution, if the absentee is expected to return to work shortly after the time of the DM, or is part of an active labour market program that is likely to lead to an upcoming return to work.

[4]A list of DM use by county is provided in appendix A.2.

Figure 1: Geographical variation in the use of DMs
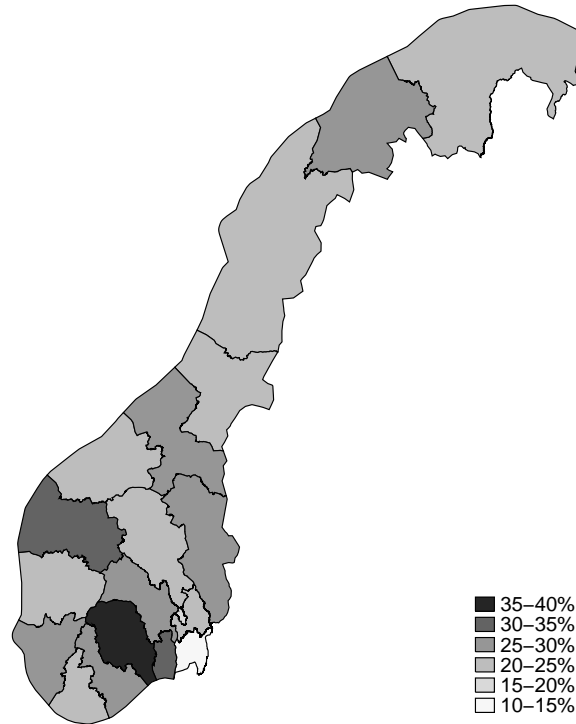


Notes: Map of Norway, where the 19 counties are divided into six groups by the share of long-term spells (longer than five months) started between January 2009 and December 2010 for which a realised DM is recorded. A darker shade illustrates a more intensive use of DMs.

The main data used in the empirical analysis consists of all physician-certified sick-leave spells in Norway that were started between January 2009 and December 2010. As pointed out above, The DM was implemented already in 2007; however, we do not have data on the use of the meeting before 2009. This also implies that in order to find a time period without DMs to use as a suitable control period, we must go back to spells that started before the summer of 2006, and thus were unlikely to be affected by the introduction of DMs in 2007. Moreover, to avoid interference from another reform in July 2004 (which changed the regulations regarding sick-leave certification), we use spells that started between July 2004 and June 2006 to establish a control group of pre-DM spells. Inclusion in the data is further conditional on full sick-pay eligibility, meaning that the absentee cannot have received any sickness benefits the previous six months. Finally, we require that the absentee is between the age of 18 and 66, are registered in the Employee Register (Arbeidstakerregisteret), and have an annual income exceeding an amount corresponding to approximately 14,500 USD.

All absentees are followed from the beginning of the spell until a transition to partial or full work resumption. A spell is right censored if no transition takes place within the twelve month period of sickness benefits.

Table 1 shows descriptive statistics on the data used in the empirical analysis. We estimate two alternative hazard rate models with different definitions of the work resumption event and corresponding analysis populations. Since both models are estimated on data for both the pre-DM and DM periods, we have in total four data sets used for estimation. Columns I and II show descriptive statistics for all sick-leave spells that were classified as full-time (100%) absence from the beginning, for the periods with and without DMs respectively. These spells will be used to examine DM-effects on any degree of work resumption - partial or full. Most of the absence spells in Norway (95%) are indeed full-time at the time of entry, and, as explained above, an important aim of the dialogue meetings is to encourage (at least) a gradual return to work for those who have little or no contact with their employer during the spell. Columns III and IV show the corresponding descriptive statistics for all spells - including those that start out with a graded absence certificate, and, hence, where the claimants partially work already from the start of the absence spells. These spells will be used to examine DM-effects on the transition to full work resumption (and thus exit from sick-leave) only. In each of the four data sets, there are around two million absence spells - experienced by around one million persons. For the spells starting with full-time absence, the average duration until partial or full work resumption takes place is 28-29 days, and around 3-4% of the claimants have not resumed any kind of work after six months (the typical timing of a DM). For all spells (including those that start out with partial work), the average duration until full work resumption is close to 40 days, and around 6% of the claimants are still on sick-pay after six months.

Table 1: Descriptive statistics

| | Full-time sick-leave spells | | All sick-leave spells | |
| --- | --- | --- | --- | --- |
| | DM period | Pre-DM period | DM period | Pre-DM period |
| | I | II | III | IV |
| Fraction of females (% ) | 57.3 | 56.3 | 58.1 | 57.0 |
| Mean age at entry to sick-leave | 40.0 | 39.9 | 40.1 | 40.0 |
| Level of achieved education | | | | |
| High school (%) | 44.8 | 48.1 | 44.4 | 47.6 |
| University/college or higher (%) | 29.1 | 27.0 | 30.1 | 27.8 |
| Immigrant background (%) | 20.8 | 15.7 | 20.5 | 15.5 |
| Mean income in Base Amounts (14,500 USD) | 5.0 [2.2] | 5.0 [2.1] | 5.0 [2.2] | 5.0 [2.1] |
| Mean duration until partial or full work resumption (days) | 29.5 [62.0] | 28.2 [59.1] | N/A | N/A |
| Mean duration until full work resumption (days) | N/A | N/A | 39.6 [78.0] | 37.4 [75.0] |
| Fraction with a DM (%) | 1.3 | N/A | 2.4 | N/A |
| Mean time before DM (in days) | 182 | N/A | 184 | N/A |
| Fraction returning to work (%) | 98.5 | 98.5 | 97.9 | 97.9 |
| Number of spells | 2,014,576 | 1,880,895 | 2,115,424 | 1,966,940 |
| Number of individuals | 1,060,256 | 1,009,139 | 1,098,709 | 1,042,308 |
| *Long-term spells > 26 weeks* | | | | |
| Fraction with dialogue meeting (%) | 33.6 | N/A | 36.5 | N/A |
| Fraction returning to work (%) | 65.6 | 64.4 | 70.5 | 68.0 |
| Number of spells | 74,245 | 61,732 | 132,975 | 112,025 |

[Standard deviations in brackets]

Notes: Column I shows all sick-leave spells started between January 2009 and December 2010. Column II shows all sick-leave spells started between July 2004 and June 2006. Column III (IV) is a sub-sample of column I (II) and shows all full-time (100% sick-leave from work) spells started within this period.
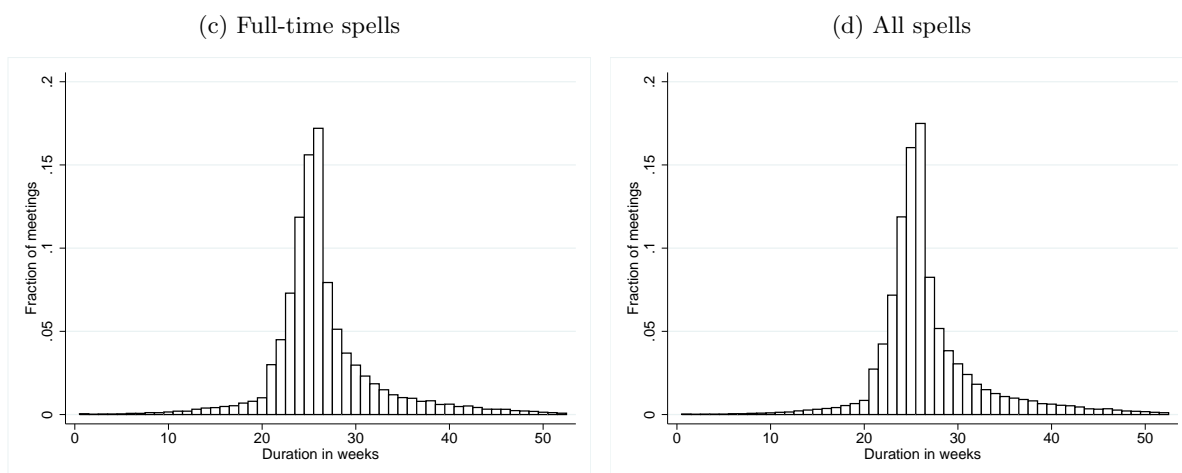
Panel (a) [(b)] of Figure 2 shows the distribution of spell durations for the sample of full-time [all] spells in the pre-DM and DM periods respectively. The distributions for the pre-DM and DM periods are very similar (so similar that it is actually difficult to spot any differences at all), and regardless of spell type, the vast majority of the spells last just a couple of weeks or less. Since our analysis focuses on a meeting scheduled to take place around 26 weeks into the absence spells, we are not really interested in the many short-term spells per se. However, to avoid sample selection problems, we include them in the data. This allows us to appropriately model the selection process that determines which spells eventually reaches the DM-relevant durations. The figure shows a spike in the frequency of spells with a duration of 51-52 weeks. This is probably a result of the sickness benefit payments coming to an end after 52 weeks. Panel (c) [(d)] of Figure 2 show the distributions of the spell durations at which the DM is held for the sample of full-time [all] spells. This figure reveals significant variation in spell duration at the time of the meeting, which can be explained both by individual needs and differences in local priorities. The figure further shows that very few meetings are arranged before week 20.

Figure 2: Timing of the DM and weekly transition rates

i) Distributions of spell duration

(a) Full-time spells

(b) All spells



ii) Timing of the DM
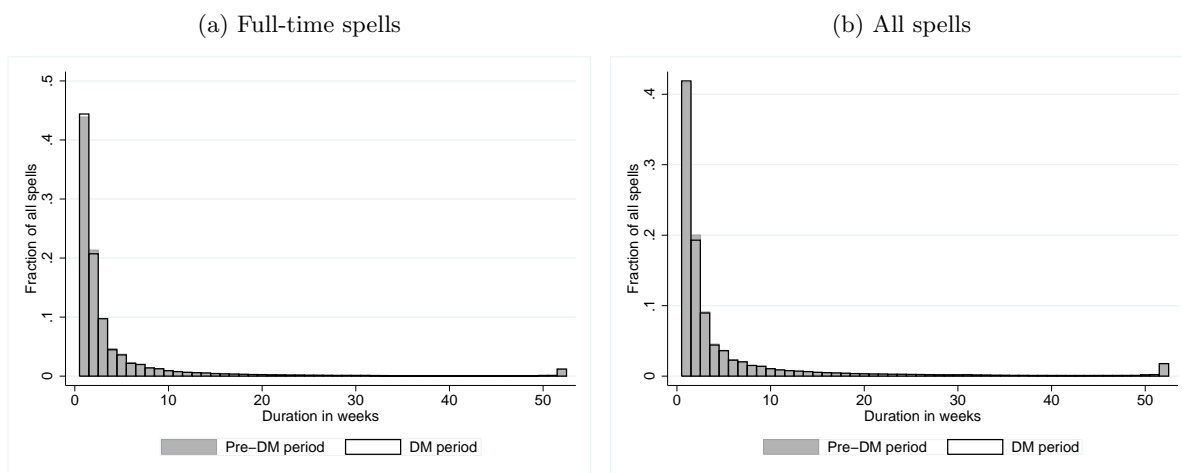
(c) Full-time spells

(d) All spells



Notes: Panel (a) [(b)] of Figure 2 shows the distribution of spell durations for the sample of full-time [all] spells in the pre-DM and DM periods respectively. In both panels, the gray areas are the distributions in the pre-DM period and the areas outlined in black are the distributions in the period with DMs. Panel (c) [(d)] shows the distribution of spell durations at which the DM is held for the sample of full time [all] spells. See Table 1 for details on the four different samples.

# 3 Empirical analysis

In this section, we set up and estimate statistical models designed to identify the causal effects of DMs on work resumption, as well as on subsequent employment, earnings, and social insurance dependency. The basic idea of our identification strategy is to use the

random-assignment-like variation in absentees' exposure to DMs generated by differences in DM-intensities across counties to identify the causal effects of interest. Hence, the first step of our empirical approach is to compute county-specific DM intensities for each duration-week. By exploiting the county-variation in the use of DMs, we obtain causal effect estimates identified on the basis of persons who potentially are treated differently in different counties. To the extent that the true causal effects are heterogeneous, this implies that our estimates will be representative for this "marginal" population. This is also the population that primarily will be affected by attempts to scale up or down the use of DMs. Estimates based on observed county-variation is thus arguably policy relevant. The main aim of our empirical analysis is to examine the extent to which DMs influence the propensity to partially or fully resume work, and further to decompose any estimated impacts into notification and attendance effects. Intuitively, it appears possible to separate notification from attendance effects on the basis of their timing within spells. Effects occurring prior to the typical timing of the meeting reflect notification, while effects occurring afterwards reflect attendance. However, if there really is a notification effect, this may itself imply that the population actually attending the meeting is selected. This particular selection problem remains even if the calls to the meetings are completely randomly assigned. Hence, to correctly distinguish the two effects, we need to model explicitly how any shifts in hazard rates following from notification changes the population at risk of being exposed to an attendance effect also. We do that by means of a hazard rate model, allowing for observed as well as unobserved heterogeneity. While we can exploit data on the exact timing of work resumption to disentangle impacts of notification and attendance within the framework of a duration model, we are not able to offer a similar disentanglement for the longer term impacts of DMs. Hence, when we examine the influence of DMs on later employment, earnings and social insurance dependency (two years after the start of the absence spell), we limit ourselves to a more reduced form analysis, where we seek to identify the overall impacts of the counties' choice of DM intensity.

## 3.1 Measuring county-specific meeting intensity profiles

To compute county-and-duration-specific DM intensities relevant for each absence spell s, we use all absence spells exceeding 16 week duration and set up a transition rate model with participation at a DM as the endogenous event (shorter spells are dropped for the reason that a DM almost never occurs during the first 16 weeks). The model consists of one observation per absentee per week of the absence spell until the absentee is no longer under risk of attending a meeting. This means that an additional observation is included for each week the absentee has not attended a meeting or terminated the spell. The spell is right censored when the absentee returns to work, or after 12 months.

Our strategy is similar to that of Markussen and Røed (2014), who study the effects of a variety of vocational rehabilitation programs by exploiting local variation in the use of the different programs as a source of randomness in the probability of being assigned to the different programs. When constructing the measures of DM intensities, the contribution of a person's own spell must be removed from the indicator to avoid a mechanical correlation between the indicator and anything unobserved about the spell. To simplify these calculations, we follow Markussen and Røed (2014), and use a linear transition rate model to construct the treatment intensities. Let $DM_{sjd}$ be the event of attending a meeting for an absentee with spell $s$ who has been under risk of meeting attendance for $d > 16$ weeks in county $j$:

$$DM_{sjd} = \boldsymbol{\delta_d} + \boldsymbol{x'_{st}}\boldsymbol{\theta} + u_{sjd}, \tag{1}$$

where $\boldsymbol{\delta_d}$ is a vector of weekly duration dummy parameters and $\boldsymbol{x_{st}}$ is a vector of individual observable control variables including age, sex, medical diagnosis, earnings level, industry of employment, education (level and type), contracted work-hours, local labour market tightness, and calendar time. To avoid invalid functional form restrictions, all the controls are included in the form of relatively large numbers of indicator variables; see Appendix A.1 for details. The measure of treatment intensity faced by the individual with spell $s$, still at meeting risk at duration $d$ in county $j$ is constructed by calculating the mean residual within

county for each duration, while removing the contribution of spell $s$:

$$\phi_{jd,-s} = \frac{1}{N_{jd} - 1} \Big[ \sum_{k \epsilon N_{jd}} \hat{u}_{kjd} - \hat{u}_{sjd} \Big], \tag{2}$$

where $N_{jd}$ is the number of spells still at meeting risk at duration $d$ in county $j$, and $\hat{u}_{sjd}$ is the residual from regression Equation (1). As we are aiming at disentangling notification and attendance effects of the DMs, we are interested in both the probability of attending a DM in each given duration week and in the probability of already having attended a meeting. Both variables are expressed in terms of the county and duration-specific treatment intensities, which are scaled in order to make the attendance probabilities vary in the range zero-one. For spell s, we hence express the probability of being called to a DM in week $d$ conditional on not having been called before (the weekly hazard) as follows:

$$h_{sjd} = \begin{cases} 0 & \text{for } d = 0, 1, ..., 16 \\ \hat{\phi}_{jd,-s} + \hat{\delta}_d + c & \text{otherwise,} \end{cases} \tag{3}$$

where $\hat{\delta}_d$ is the estimated coefficient on element $d$ of the vector of duration parameters from Equation (1), and c is a constant such that $c + \hat{\delta}_{d=26} = D\bar{M}_{d=26}$, where $D\bar{M}_{d=26}$ is the observed mean DM transition rate for all spells in Norway at duration week 26.

From this expression, we can derive the duration-specific probability of already having attended a meeting, $S_{sjd}$ as

$$S_{sjd} = S_{sj,d-1} + (1 - S_{sj,d-1}) h_{sjd}, \tag{4}$$

with $S_{sj,d=0} = 0$. The second part of this expression is the unconditional probability of attending a DM for each duration week which we denote $g_{sjd}$:

$$g_{sjd} = (1 - S_{sj,d-1}) h_{sjd}. \tag{5}$$

Figure 3 illustrates the spatial variation in the two predicted attendance probabilities (attending this week ($g_{sjd}$) and already having attended ($S_{sjd}$)) over spell duration, by

comparing two counties in each end of the meeting intensity distribution with the national average. The filled circles in panel (a) and (b) are the duration-specific means of the attendance probabilities in Telemark, the county that tends to use DMs the most, while the hollow circles are the duration-specific means in Østfold, the county that tends to use the DMs the least. Finally, the triangles represent the duration-specific national means. As can be seen from the figures, the attendance probabilities vary considerably both over spell duration and over counties. In line with national regulations, there is a clear attendance spike around week 26, yet a number of meetings are held some time before or after week 26 also. Conditional on sick-pay exhaustion (52 week duration), the probability of having attended a DM at some stage during the spell varies from around 18-50% over counties, with a national mean of approximately 35%.

Figure 3: DM attendance probabilities

| (a) Probability of attendance this week | (b) Probability of already having attended |
|---|---|
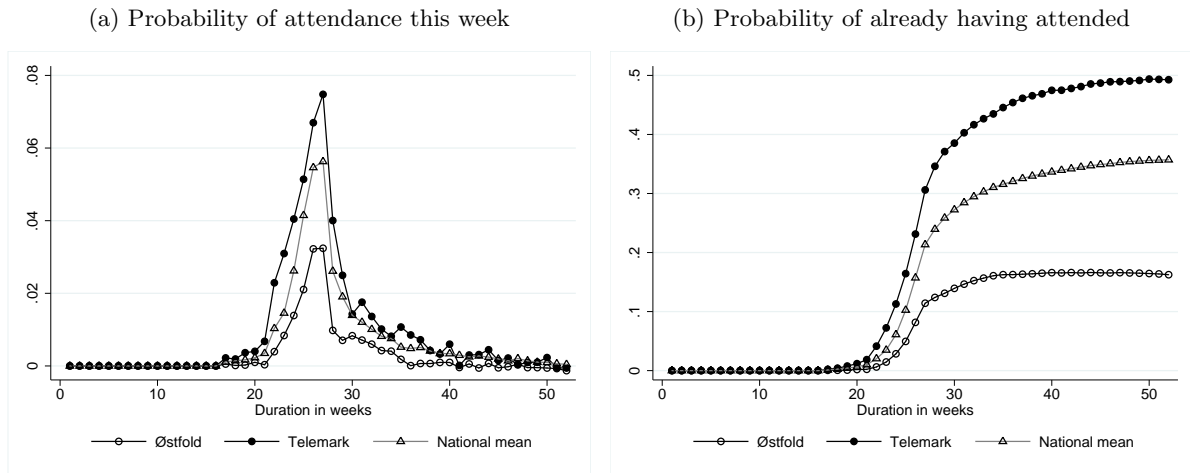


Notes: Panel (a) illustrates the variation in the predicted probability of attending a DM over spell durations for the county in Norway that tends to use the DM the least (Østfold; illustrated by hollow circles), the county that tends to use the DM the most (Telemark; illustrated by filled circles) and the country as a whole (illustrated by triangles). Panel (b) illustrates the variation in the predicted probability of already having attended a DM, also for Østfold, Telemark and the whole country.

## 3.2 Effects on the work resumption propensity

We model the probability of resuming work - partially or fully - by means of a mixed proportional continuous time hazard rate model (MPH) with piecewise constant duration effects estimated separately for each week. Given the suspected selectivity of DM

participation - both due to selectivity in the calls to the meetings and due to the selectivity resulting from the possible effects of the calls - we do not exploit data on actual DM participation at all in this subsection. Instead, our key explanatory variables are going to be proxy-variables representing the probability of being notified about a forthcoming DM and the probability of already (and recently) having participated. Although it is clear from the previous discussion that we intend to derive these probabilities on the basis of the county-by-duration specific DM intensities, it is far from obvious exactly how this should be done. Since DM notifications are unobserved, we need to make assumptions regarding their timing relative to the actually held meetings, and also regarding the duration of any notification effects. In addition, we need to make assumptions about the duration profile of any attendance effects. Since none of these assumptions can be completely based on prior knowledge, a central element of our modeling strategy will be to test out a number of (reasonable) alternative specifications, and then to choose the best one based on an information criterion. For the moment, we take an agnostic view on this, and include in the hazard rate the unspecified variables $N_{sjd}$ and $A_{sjd}$ as representing notification and attendance, respectively, together with controls for individual (and spell) heterogeneity, calendar time, spell duration, and county. Let $\boldsymbol{x}_{st}'$ be a vector of observable characteristics, $\boldsymbol{\gamma_j}$ be a vector of county fixed effects, $\boldsymbol{\lambda_t}$ denote calendar month fixed effects (monthly dummy parameters), $\boldsymbol{\rho_d}$ denote weekly spell duration parameters and $\boldsymbol{v_i}$ be a measure of individual-specific unobserved heterogeneity. The work resumption hazard rate can then be expressed as

$$\theta_{sijtd} = exp(\boldsymbol{x}_{st}'\boldsymbol{\alpha} + \boldsymbol{\lambda_t} + \boldsymbol{\rho_d} + \boldsymbol{\gamma_j} + \beta^1 A_{sjd} + \beta^2 N_{sjd} + \boldsymbol{v_i}). \qquad (6)$$

The vector of observed characteristics $\boldsymbol{x_{st}}$ is exactly the same as the one we used to compute the county-specific DM-indices in Equation (1); see the Appendix for details. It may be noted that while observed individual explanatory variables are spell-specific and time-varying, the individual unobserved covariate, $\boldsymbol{v_i}$, is person-specific (hence the i-subscript) and time-invariant. This implies that we exploit the existence of repeat spells to disentangle the impacts of duration dependence and unobserved heterogeneity; see, e.g. Van den Berg (2001). Person-specific unobserved heterogeneity can in our case be justified by the fact that we have controlled for the spell-specific medical diagnoses; hence the role of

the unobserved covariate is primarily to capture more deep-rooted individual characteristics, such as motivation, work ethics and general health status.

An important point to note from this specification is that since we have included county and duration fixed effects in the hazard rate, systematic differences in work resumption rates between counties or over spell duration will not contribute to identification of the causal parameters of interest. The two DM effects are instead identified by shifts in the hazard rate caused by the county-specific meeting intensities at durations corresponding to the typical timing of notification and recent participation, respectively, i.e., by the variation in duration-profiles across counties and their correlations with the DM intensity variables. The model is estimated by maximum likelihood. To do this in practice, each spell must be divided into a number of "spell parts", such that each part is characterised by covariate constancy. This implies that any change in an explanatory variable (e.g., because a new duration week is beginning) triggers a new spell part. Let the subscript $sp$ denote spell part $p$ of spell $s$, let $l_{sp}$ denote the length of that spell part (measured in days) and let $y_{sp}$ be an indicator variable taking the value 1 if the spell part ended with a transition back to work and zero otherwise. The contribution of individual $i$ to the likelihood function is then written:

$$L_i(\boldsymbol{v_i}) = \prod_{s \in S_i} \prod_{p \in s} [\theta_{sp,ijtd}(v_i)]^{y_{sp}} exp\left(-l_{sp}[\theta_{sp,ijtd}(\boldsymbol{v_i})]\right). \tag{7}$$

The vector of unobserved individual heterogeneity, $\boldsymbol{v_i}$ is approximated with an *a priori* unknown discrete probability distribution. The probability distribution is estimated non-parametrically by adding support points in the distribution until the model is saturated (Heckman and Singer (1984)). The preferred model is then chosen on the basis of the Akaike Information Criterion (AIC). See Gaure et al. (2007) for details on the estimation algorithm. Let $Q$ be the number of support points in the distribution of unobserved heterogeneity $v_m$ with associated probability $q_m$, $m = 1, 2, ..., Q$. The sample likelihood function can be written:

$$L = \prod_{i=1}^{N} \sum_{m=1}^{Q} q_m \prod_{s \in S_i} L_i(v_m), \qquad \sum_{m=1}^{Q} q_m = 1. \tag{8}$$

In order to carry out this estimation strategy, we first need to specify the two key variables $N_{sjd}$ and $A_{sjd}$. The notification variable $N_{sjd}$ should ideally represent the county-by-duration specific probability of receiving a call to a DM. While we have a good indicator of the probability of weekly participation, we have no data on meeting calls. We have been informed by the Norwegian social security administration that a commonly used practice is to summon the DM three weeks prior to its planned date. A natural assumption to make is thus that the county-by-duration specific probability of receiving a meeting call is approximately proportional to the corresponding probability of participation three weeks later.[5] The timing of notification is likely to vary somewhat across spells, however, and in order to assess this further, we specify a more general measure by leading the attendance probabilities derived in Equation (5) by $z_N$ weeks:

$$N_{sjd} = g_{sj(d+z_N)} \tag{9}$$

This specification of the notification variable implies that the work resumption hazard during a particular week is modeled as a function of the county-specific participation rate applying $z_N$ weeks later. The specification further involves an assumption of the notification effect lasting for exactly one week. This assumption is made since the data do not allow us to distinguish between a one-week notification effect and a more dispersed effect pattern over the weeks close in time to the actual week of notification. Hence, the coefficient attached to $N_{sjd}$ needs to be interpreted with some care. It is also of some interest to examine specifications of $N_{sjd}$, based on alternative choices of $z_N$. For example, specifying the lead as two instead of three weeks may be preferable if either the call tends to be submitted somewhat closer to the actual meeting, or if the main effect comes during the week after the notification was received. Similar arguments can be made regarding the specification of the attendance variable. We will define the attendance variable $A_{sjd}$ such that it measures the county-by-duration specific probability of recently having participated at a DM. In this case, we need to operationalise the concept of "recently"; i.e., the period of time following a DM for which we allow for a causal effect. Assuming that the attendance effect lasts for $z_A$ weeks, we specify the attendance variable as the probability of having participated at a

---

[5]Note that the relationship will not generally be proportional, as the ratio of the call-propensity to the realised meeting propensity depends on the level of the involved baseline hazard rates, on the county-specific probability of participation, as well as on the size of the notification effect.

DM some time during the last $z_A$ weeks, which based on the notation in Equation (4) can be expressed as

$$A_{sjd} = S_{sjd} - S_{sj(d-z_A)} \tag{10}$$

The specification of a constant effect lasting for a given number of weeks after meeting attendance is of course somewhat arbitrary, and it should be interpreted as a simplification rather than as a substantive assumption. In practice, we expect there to be a dynamic effect pattern, possibly with effects that taper off with time. The probability of having participated during the last $z_A$ weeks is of course also highly correlated with the probability of having participated during, say, the last $z_A + 1$ weeks, and in practice it is impossible to disentangle the precise nature of a dynamic effect pattern from the distribution of participation probabilities. Hence, the coefficient attached to $A_{sjd}$ also needs to be interpreted with some care. Note that the variation in both $N_{sjd}$ and $A_{sjd}$ is basically county-by-duration specific; the only reason why we include an s-subscript is that they vary slightly across individuals due to the leave-out-mean strategy described in Equation (2); i.e., that the intensities assumed to affect spell s are computed on the basis of all other spells in the same county and at the same duration.

To choose appropriate values of $z_N$ and $z_A$, we estimate the hazard rate model under a number of alternative $z_N$-$z_A$-combinations, and chose the combination that maximises the likelihood function. In order to save computational resources, this initial model-competition is carried out without incorporating unobserved heterogeneity into the likelihood function (the inclusion of unobserved heterogeneity does in practice not appear to influence the ranking of models). The chosen model is then re-estimated with full nonparametric specification of unobserved heterogeneity based on the likelihood function in Equation (8). The results from our model competition exercise for the model based on full-time absence spells, with partial or full work resumption defined as the outcome variable are presented in Figure 4. With respect to the number of leads $z_N$ on the notification variable, a clear "winner" emerges, namely $z_N = 3$. This gives the highest likelihood regardless of how the attendance variable is specified. As $z_N = 3$ corresponds to our prior knowledge regarding the timing of the DM call (the call is indeed supposed to be sent out three weeks prior to

the meeting), we find this reassuring. With respect to the duration of the attendance effect $z_A$, the results are less clear-cut (which is no surprise given the extremely high correlation we expect to find between measures based on small variations in $z_A$). Yet, for all choices of $z_N$, the likelihood is maximised for $z_A = 12$. Hence, our preferred model is based on the assumption that the notification of meetings occurs three weeks prior to their planed implementation and that the attendance effect lasts for 12 weeks. It should be noted, however, that our results would have been very similar had we selected other specifications in the neighborhood of the chosen one.[6]

Figure 4: Choice of model
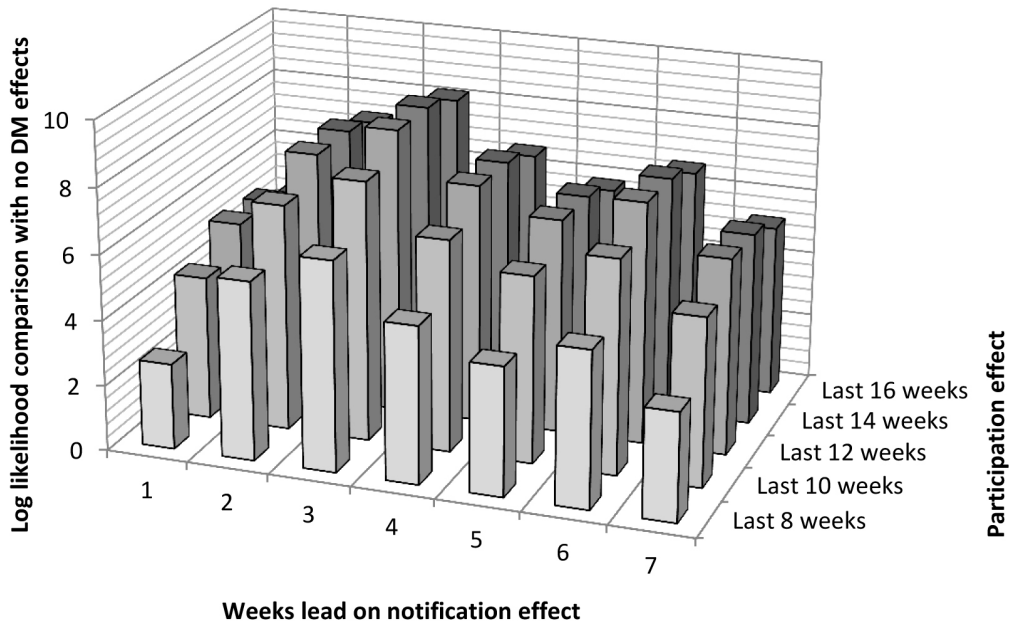


Notes: Sample of full-time spells with partial or full work resumption as outcome variable. The height of the columns represent the log likelihood for each of the 35 estimated models when subtracting the log likelihood from a model where both DM intensities $A_{sjd}$ and $N_{sjd}$ are excluded.

The main results based on the chosen model are presented in Table 2. Column I and II show the results for the models with partial or full work resumption as outcome variable for the pre-DM and DM periods respectively. Similarly, column III and IV show the results for the models with only full work resumption as outcome variable. Estimates for the pre-DM period

---

[6]The results of the model competition are very similar for the model estimated on all spells and full work resumption defined as the outcome, although it is in this case more difficult to identify a clear "winner". For this model $z_N = 2$ is marginally better than $z_N = 3$, and $z_A = 12$, $z_A = 14$, and $z_A = 16$ give almost exactly the same likelihood. To make the results directly comparable, we use $z_N = 3$ and $z_A = 12$ for both models/datasets.

may be interpreted as a sort of placebo analysis. Given the ambiguities regarding the precise interpretation of the coefficient estimates, we will use them primarily for statistical inference regarding the existence and sign of notification and attendance effects. Our discussion of magnitudes and substantive significance will instead be based on simulation exercises where we compare results of imposing the meeting intensity profiles actually observed in different counties. Our results indicate statistically significant positive effects of both notification and attendance. The effects are considerably larger when the outcome is measured as an indicator for any degree of work resumption - partial or full - rather than an indicator for full work resumption only. Estimating on the sample of pre-DM spells instead (placebo analysis) gives much smaller and (in all cases) statistically insignificant coefficient estimates, supporting the causal interpretation of the estimates obtained based on spells from the period with DMs.

Table 2: Results

| | Partial of full work resumption | | Full work resumption | |
| | I | II | III | IV |
| | DM period | Pre-DM period (Placebo) | DM period | Pre-DM period (Placebo) |
|---|---|---|---|---|
| Notification | 1.691*** [0.600] | 0.360 [0.658] | 1.216*** [0.447] | 0.230 [0.537] |
| Attendance last 12 weeks | 0.328*** [0.112] | 0.174 [0.125] | 0.169** [0.080] | 0.004 [0.100] |
| # Masspoints | 12 | 10 | 10 | 11 |
| Number of spells | 2,014,576 | 1,880,895 | 2,115,424 | 1,966,940 |

\* 0.1 \*\* 0.05, \*\*\* 0.01. [Standard errors in brackets]

Notes: # Masspoints refers to the number of masspoints in the distribution of unobserved heterogeneity. For a list of included control variables, see Appendix A.1.

To shed some light on the magnitudes of the estimated effects, we perform a simulation exercise based on the estimated models where we simulate the duration of all absence spells under alternative assumptions regarding the two key variables $N_{sjd}$ and $A_{sjd}$ and their effects on the transition rates. More specifically, to ensure that we do not extrapolate outside the range of variation actually observed in the data, we compare simulated work resumption outcomes when we impose three alternative sets of DM profiles on $N_{sjd}$ and $A_{sjd}$:

i Mixed DM intensity: The profiles actually estimated for each county.

ii High DM intensity: The profiles estimated for the county that uses DMs the most (Telemark).

iii Low DM intensity: The profiles estimated for the county that uses DMs the least (Østfold).

To disentangle the quantitative impacts of notification and attendance, we repeat each simulation with only one of the estimated effects "turned on" at the time. The results are presented in Table 3. The upper [lower] part of the table shows results from the model including full-time [all] spells. Column I and II report summary statistics, while column III reports the effect on spell duration per meeting held implied by the simulation of the models. Looking first at the results for the model with partial or full work resumption defined as the outcome of interest, we note that the different DM profiles have very small impacts on overall absence duration. For example, substituting the high for the low DM profile reduces the average spell duration by a mere 0.24 days. However, this "small" effect primarily reflects that the DM profiles are relevant only for the tiny fraction of spells that lasts at least 5-6 months. Hence, even the high DM intensity profile implies that less than 2% of the absentees ever attend a DM. Looking at the impact on average spell duration per realised meeting instead, we find that each extra DM reduces the duration until partial or full work resumption by as much as 19-20 days. Hence, for the target population of long-term absentees, the DM profile appears to be of considerable importance. Repeating the simulation exercises with only one of the estimated notification and attendance effects turned on at the time (the other one set to zero), we find that the two effects are of similar importance for average duration. The attendance effect is slightly larger than the notification effect. Moving on to the model with full work resumption as the outcome of interest, the simulated impacts are considerably smaller, yet far from negligible. Each extra DM is estimated to reduce the duration until full work resumption by around 10-11 days. And for this outcome, the notification effect is slightly larger than the attendance effect.

Table 3: Simulated impacts on duration until work resumption of alternative DM profiles

| | I | II | III |
|---|---|---|---|
| | Average spell duration (days) | Fraction attending a DM (%) | Implied effect on spell duration per extra meeting (days) |
| **Partial or full work resumption** | | | |
| (N= 2,014,576) | | | |
| Assumed DM profile: | | | |
| Mixed (as estimated for each county) | 29.16 | 1.35 | |
| High (as estimated for Telemark) | 29.04 | 1.93 | |
| Low (as estimated for Østfold) | 29.28 | 0.69 | |
| Substituting high for mixed DM intensity | -0.12 | +0.58pp | -20.3 |
| With notification effect only | | | -9.1 |
| With attendance effect only | | | -12.1 |
| Substituting high for low DM intensity | -0.24 | +1.24pp | -19.0 |
| With notification effect only | | | -8.4 |
| With attendance effect only | | | -10.6 |
| **Full work resumption** | | | |
| (N= 2,115,424) | | | |
| Assumed DM profile: | | | |
| | 39.18 | 2.42 | |
| Mixed (as estimated for each county) | | | |
| High (as estimated for Telemark) | 39.07 | 3.47 | |
| Low (as estimated for Østfold) | 39.29 | 1.22 | |
| Substituting high for mixed DM intensity | -0.11 | +1.05pp | -10.8 |
| With notification effect only | | | -5.6 |
| With attendance effect only | | | -4.8 |
| Substituting high for low DM intensity | -0.22 | +2.25 | -10.1 |
| With notification effect only | | | -5.3 |
| With attendance effect only | | | -4.5 |

Notes: The upper [lower] part of the table shows results from the model including full-time [all] spells. Column I and II show summary statistics when imposing the three alternative assumptions on the DM profile, while column III is the effect on spell duration per meeting held implied by the simulation of the models.

### 3.3 Effects on long-term outcomes

To learn more about the effectiveness of arranging DMs, we estimate the effect of the meeting intensity on a number of long-term outcomes; i.e., employment, earnings, and social insurance dependency, where earnings and employment are measured in the second year after the year of entry into sick-pay, and disability benefit receipt is measured exactly 24 months after the month of entry into sick-pay. We do this by comparing how long-term and short-term absence spells are differentially affected by the overall county-specific DM intensity, using the pre-DM period to control for any spurious correlation between the DM-intensity measure and the county-specific performance of long-term versus short-term sick-listed workers. Hence, in this subsection, we use the pre-DM spells directly as "controls" in a difference-in-differences setting, rather than as inputs to a separate placebo analysis.

The starting point of the analysis is the same datasets as we used in previous subsection; see Table 1. However, since we are not attempting to disentangle notification and attendance effects in this analysis, and therefore do not need to worry about modeling the sorting into attendance generated by notification, we drop the large number of very short spells; i.e., those lasting less than three weeks. These spells are clearly not affected by DMs, and since the diagnosis composition is very different from the longer spells, they may also constitute a poor comparison group to the spells that are affected. The remaining spells are then classified as either "short-term" or "long-term", such that the spells in the first category are not under any significant risk of attending a DM, while the spells in latter category face a risk of DM attendance. "Short-term" is defined as having a duration between 3 and 15 weeks, and "long-term" is defined as having a duration of at least 19 weeks.[7] Spells with a duration between 15 and 19 weeks are dropped from the sample to amplify the differences in DM risks between the spells in the two categories. Table 6 in appendix A.3 shows descriptive statistics on short and long-term spells in the pre-DM and DM periods for both the sample of full-time spells and the sample of all spells. The composition of absentees on observable individual characteristics is similar in the two periods, both for short and long-term spells.

---

[7]The results are not sensitive to small variations in these definitions.

To evaluate the long-term effects, we model the different outcome variables of interest, all denoted $y_{sj}$, as linear functions of the probability of the county-specific probability of attending a meeting during the course of a full 52 week sick-pay period, i.e., $S_{sj,52}$. The outcomes are labour earnings, social insurance claims, and employment, all measured in the second calendar year after the year of entry into sick-leave, and an indicator for disability insurance claims, measured 24 months after the month of entry into sick-leave. We estimate the following triple difference model:

$$y_{sj} = \phi + \boldsymbol{x'_{s,t}}\boldsymbol{\pi} + \boldsymbol{\xi_j} + \boldsymbol{\eta_t} + \tau_1 Long_d + \tau_2 Long_d \times Post_t + \tau_3 Long_d \times S_{sj,d=52}$$
$$+ \tau_4 Post_t \times S_{sj,d=52} + \tau_5 Post_t \times Long_d \times S_{sj,d=52} + v_i, \quad (11)$$

where $Long_d$ is an indicator for the spell lasting at least 19 weeks and $Post_t$ is an indicator for the spell occurring in the DM-period. Further, $\boldsymbol{x'_{it}}$ is the same vector of covariates as in Equation (6), $\boldsymbol{\xi_j}$ is a vector of county fixed effects and $\boldsymbol{\eta_t}$ is a vector of year and (seasonal) month fixed effects (measured at the time the spell was started). We have normalised the coefficients such that $\tau_5$ captures the causal effect of an increase from the lowest to the highest observed intensity (i.e., a change in $S_{sj,52}$ from the Østfold to the Telemark average levels).

The results are presented in Table 4. The estimated impacts on longer term labour market outcomes are generally small. This is as expected, given that the evaluated policy intervention is also very small (after all, it's just a meeting). Yet, they all point in the same direction and a number of them are statistically different from zero. Hence, it appears that more intensive use of DMs raises long-term absentees' employment and earnings prospects and reduces their risk of continued disability insurance claims. For example, in the model with all absence spells included, we find that an increase in the DM intensity corresponding to the difference between the lowest and the highest observed intensity leads to an increase of around 1.5 percentage points in the probability that a long-term absentee has a labour income exceeding around 29,000 USD in the second year after the year of sick-pay entry. Furthermore, such an increase raises expected earnings in that year by around 1450 USD, and it reduces the probability of claiming temporary disability benefits 24 months after

sick-pay entry by approximately 2.3 percentage points.

Table 4: Effects of county-specific DM intensity on outcomes for long-term absentees measured two years after sick-pay entry

|  | Full or partial work resumption (Full-time absence spells sample) | Full work resumption (All absence spells sample) |
|---|---|---|
| Income $\geq 2\times$ Base Amounts (29,020 USD) | 0.015* [0.008] | 0.013*** [0.005] |
| Income $\geq 3\times$ Base Amounts (43,530 USD) | 0.012 [0.007] | 0.009* [0.005] |
| Income in Base Amounts (14,510 USD) | 0.100* [0.057] | 0.099*** [0.035] |
| On temp. disability | -0.023*** [0.007] | -0.014*** [0.005] |
| Number of spells | 905,265 | 1,112,104 |

* 0.1 ** 0.05, *** 0.01. [Standard errors clustered at the county level in brackets]

Notes: The table shows the estimates of $\tau_5$ in Equation (11), with $S_{sj,d=52}$ normalised such that its range equals the difference between the two counties that use DMs the most and the least. The dataset is described in Table 6 in Appendix A.3. The income-related outcomes are measured during the second calendar year after the year of sick-leave entry. The outcome "On temporary disability" is a dummy variable indicating whether a person receives temporary disability benefits exactly 24 months after sick-leave entry. Permanent disability benefits are not included in this outcome due to missing data. For a list of included control variables, see Appendix A.1.

# 4    Cost efficiency

In order to assess cost efficiency, we compare the estimated DM gains with the associated costs. To obtain a measure of the benefits of the DM, we multiply the number of sick-leave days reduced per meeting by the average daily salary from the data set on all sick-leave spells (presented in Table 1). The average daily salary of the sick-listed workers in 2009 and 2010 is calculated to 1,580 NOK (269 USD). Our most conservative estimate of the number of saved sick-leave days is based on the model with only full work resumption counted as exit from sick-leave, and in this model we found that each realised DM reduces sick-leave duration by approximately ten days. Taking into account that this number also includes weekends, we estimate that each DM saves at least seven working days, or $1,580 \times 7 = 11,060$ NOK (1,885 USD). Since the replacement rate for most of the workers is 100%, the saved social insurance payouts will be of similar magnitude. Our cost estimates are based on information from

27

the Norwegian social security administration (SSA). The SSA caseworker spends around three hours preparing and organizing a meeting. With an annual employee cost of 650,000 NOK and 1,700 work hours per year, this corresponds to 1,150 NOK per meeting. The physician is on average compensated with 1,800 NOK per meeting. The employer is not compensated for DM attendance; however, we assume a cost of 1,800 NOK for the time of the employer. In total, this gives a cost of 4,750 NOK (809 USD). According to these crude calculations, the estimate of the benefits is more than twice as large as the estimate of the costs of arranging a meeting. And, notably, on the benefit side, this cost-benefit analysis includes only the added full-time work days. Based on the model for partial or full work resumption, we estimated that an extra seven working days can be added if we also count partial work resumption (since the most typical grade for partial absence spells is 50%, this constitute approximately 3-4 full-time days). And on top of that, we have indications that DMs raises subsequent employment and slightly reduces disability insurance dependency. Taken together, our results clearly substantiate that arranging compulsory dialogue meetings is a cost efficient way to reduce long-term absence.

# 5   Conclusion

This paper examines the effect of a policy aimed at inducing long-term sick-listed workers in Norway to full or partial work resumption. The policy is a dialogue meeting (DM) organised by the local social security administration around six months into sick-leave spells. To identify the causal effects of the DM, we exploit that there is considerable geographical variation in both the overall use of DMs, and in the timing of DMs within absence spells. We argue that this gives rise to random-assignment-like variation in the duration-specific probabilities of being called to a DM, which makes it possible not only to identify the overall effects of the meeting, but also to distinguish notification from attendance effects. Similar to findings in the literature on active labour market policies for unemployed workers, we find that the mere notification of the meeting has a considerable impact on the work resumption hazard. However, in contrast to the existing literature on labour market programs, we also find large favorable effects of actual participation. In total, our results imply that each

realised DM yields a reduction of around ten days in time until full work resumption, and a reduction of around 20 days in time until any form of work resumption (partial or full). The notification and attendance effects are of similar magnitude. Based on some simple cost-benefit calculations, we show that economic gains derived from earlier work resumption by far exceed the costs of arranging the meetings.

The expressed aim of the DM policy is to encourage contact between the employee and the workplace, and thus to smooth the process of returning to regular duties. It has been hoped that this will not only contribute to lower absenteeism, but also to reduce the risk of long-term absence becoming a first step toward labour market exclusion. The findings reported in this paper indicate that the DMs are successful in promoting partial work resumption, and to some extent also in strengthening future labour market attachment. High DM intensity appears to raise employment and reduce social insurance dependency two years after sick-leave entry.

In our empirical analysis, we are not able to shed light on the relative importance of DM effects operating through the behaviors of the absentee, the physician and the employer. The meeting may be viewed as a monitoring device, whereby the social insurer can assess all three parties' efforts - or lack thereof - to ensure a successful return to work. The absentee may realise that the health situation no longer meets the requirements for continued full-time sick-pay, and thus take steps to resume work before this is established. The physician may respond to a call to a DM by not renewing a poorly justified absence certificate. Also the employer might respond by reconsidering the opportunities for modified work, as required by the law, even when it entails some additional costs. Although we cannot distinguish the contribution of the absentee from that of the physician and employer, our conclusion is nevertheless the same; arranging mandatory dialogue meetings some time into an absence spell is an efficient policy measure to reduce long-term sickness absence and curb inflow to disability programs.

The findings reported in this paper may justify a more intensive use of dialogue meetings in Norway (at least up to the levels observed in the counties with the most intensive use so

far) and possibly to introduce similar arrangements elsewhere. One question that naturally arises in this context is whether the dialogue meetings should be organised earlier than six months into the sick-pay spells. We obviously do not have any sound data-based foundation for evaluating such a policy. However, provided that the effects of DMs on work resumption hazards would have been the same for meetings held somewhat earlier, a simple simulation exercise indicate that arranging the meetings eight weeks earlier would have increased the number of actually held meetings by around 30-35%, but increased the effect per realised meeting - in terms of reduced spell duration - by around 50%. It is possible, though, that a change in the timing of the meeting also changes its effects on the hazard rates. On the other hand, the exact timing of the meeting might not change its efficiency if the success of the meetings lies in their nudging of absentees and employers into breaking a pattern of procrastination of action.

# References

Black, D. A., J. A. Smith, M. C. Berger, and B. J. Noel (2003). Is the threat of reemployment services more effective than the services themselves? Evidence from random assignment in the UI system. *American Economic Review 93 (4)*, 1313–1327.

Cockx, B., C. Ghirelli, and B. Van der Linden (2014). Is it socially efficient to impose job search requirements on unemployed benefit claimants with hyperbolic preferences? *Journal of Public Economics 113*, 80–95.

DellaVigna, S. and M. D. Paserman (2005). Job search and impatience. *Journal of Labor Economics 23*(3), pp. 527–588.

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

Gaure, S., K. Roed, and T. Zhang (2007). Time and causality: A Monte Carlo assessment of the timing-of-events approach. *Journal of Econometrics 141*(2), 1159–1195.

Geerdsen, L. P. (2006). Is there a Threat Effect of Labour Market Programmes? A Study of ALMP in the Danish UI System. *The Economic Journal 116*(513), 738–750.

Graversen, B. and B. Larsen (2013). Is there a threat effect of mandatory activation programmes for the long-term unemployed? *Empirical Economics 44*(2), 1031–1051.

Heckman, J. and B. Singer (1984). A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data. *Econometrica 52*(2), pp. 271–320.

Henrekson, M. and M. Persson (2004). The Effects on Sick Leave of Changes in the Sickness Insurance System. *Journal of Labor Economics 22*(1), 87–114.

Johansson, P. and E. Lindahl (2013). Can sickness absence be affected by information meetings? Evidence from a social experiment. *Empirical Economics 44*(3), 1673–1695.

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.

Markussen, S., A. Mykletun, and K. Røed (2012). The case for presenteeism - Evidence from Norway's sickness insurance program. *Journal of Public Economics 96*(11–12), 959 – 972.

Markussen, S. and K. Røed (2014). The impacts of vocational rehabilitation. *Labour Economics 31* (0), 1 – 13.

Markussen, S., K. Røed, and O. Røgeberg (2013). The changing of the guards: Can family doctors contain worker absenteeism? *Journal of Health Economics 32* (6), 1230 – 1239.

OECD (2010). Sickness, Disability and Work: Breaking the Barriers. Technical report, OECD, Paris.

Paserman, M. D. (2008). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation. *The Economic Journal 118* (531), 1418–1452.

Rosholm, M. and M. Svarer (2008). The Threat Effect of Active Labour Market Programmes. *Scandinavian Journal of Economics 110* (2), 385–401.

Van den Berg, G. J. (2001). Duration models: Specification, Identification and Multiple Durations. In J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5 of *Handbook of Econometrics*, Chapter 55, pp. 3381–3460. Elsevier.

# 6 Appendix

## A.1. Included control variables

The following conditioning variables are included in all all models in section 3.

Spell duration (one category for each weekly duration; 53 categories).

Calender time (one category for each calendar month; 36 categories).

County (19 categories).

Age (one dummy for each age from 18 to 66).

Native born (2 categories).

Diagnosis group. (31 categories based on average spell duration measured in five-days intervals up to a duration of 150 days, and then one category for diagnosis with an average spell durations exceeding 150 days. These are computed on the basis of sick-pay claimants who entered in 2008 and who are therefore not included in the analysis in this paper.)

Level of education (10 categories).

Field of education (10 categories).

Income group (16 categories).

Sector/occupation (84 categories).[8]

Local unemployment rate (linear variable).

---

[8]There has been a change in the sector coding over the time period of study. This means that the sector codes differ for the data from the pre-DM and the DM periods. In the model on long-term outcomes, the sector coding is hence different for spells started in the two different time periods.

## A.2. Variation in the use of DMs

Table 5: Variation in the use of dialogue meetings

| Norwegian counties | Share of spells with DM attendance |
|---|---|
| Østfold | 13.3% |
| Akershus | 21.2% |
| Oslo | 18.1% |
| Hedmark | 28.2% |
| Oppland | 22.9% |
| Buskerud | 25.6% |
| Vestfold | 30.7% |
| Telemark | 35.4% |
| Aust-Agder | 27.4% |
| Vest-Agder | 24.4% |
| Rogaland | 28.7% |
| Hordaland | 22.1% |
| Sogn og Fjordane | 30.1% |
| Møre og Romsdal | 20.4% |
| Sør-Trøndelag | 29.7% |
| Nord-Trøndelag | 24.6% |
| Nordland | 20.8% |
| Troms | 25.9% |
| Finnmark | 24.6% |
| Mean | 26.1% |

Notes: The share of the sickness spells started between January 2009 and December 2010 with a duration exceeding five months where the absentee attended a DM.

## A.3. Descriptive statistics - Long-term effects

Table 6: Descriptive statistics Long-term outcomes

| | Full-time sick-leave spells | | | | All sick-leave spells | | | |
|---|---|---|---|---|---|---|---|---|
| | DM period | | Pre-DM period | | DM period | | Pre-DM period | |
| | Short-term | Long-term | Short-term | Long-term | Short-term | Long-term | Short-term | Long-term |
| | I | II | III | IV | V | VI | VII | VIII |
| Fraction of females (% ) | 56.7 | 59.5 | 56.8 | 60.1 | 60.6 | 60.8 | 59.9 | 61.1 |
| Mean age at entry to sickness absence | 41.9 | 44.3 | 41.8 | 44.4 | 41.9 | 44.4 | 41.8 | 44.6 |
| Mean income in Base Amounts (14,500 USD) | 5.1 [2.3] | 4.9 [2.0] | 4.9 [2.2] | 4.8 [2.0] | 4.9 [3.0] | 5.0 [2.0] | 5.0 [2.2] | 4.9 [2.0] |
| Mean duration (weeks) | 6.9 [3.0] | 39.8 [12.3] | 6.9 [2.9] | 39.9 [12.3] | 6.9 [3.0] | 39.4 [12.4] | 6.9 [2.9] | 39.6 [12.4] |
| *Long-term outcomes* | | | | | | | | |
| Income $\geq$ 2$\times$ Base Amounts (29,020 USD) (%) | 85.7 | 66.2 | 86.7 | 65.8 | 87.1 | 68.5 | 87.7 | 67.9 |
| Income $\geq$ 3$\times$ Base Amounts (43,530 USD) (%) | 76.9 | 56.8 | 65.8 | 55.9 | 78.9 | 58.8 | 79.2 | 57.7 |
| Income in Base Amounts (14,500 USD) | 4.8 [3.0] | 3.7 [3.4] | 4.8 [2.8] | 3.5 [3.0] | 4.9 [3.0] | 3.8 [3.4] | 4.9 [2.7] | 3.6 [3.0] |
| On temp. disability (%) | 3.9 | 21.0 | 3.8 | 19.5 | 3.7 | 19.9 | 3.7 | 18.7 |
| N | 317,783 | 151,911 | 305,915 | 129,656 | 410,959 | 174,205 | 380,733 | 146,207 |

[Standard deviations in brackets]

Notes: Columns I-II show descriptives on the sample consisting of all full-time sick-leave spells started in the period with DMs (January 2009-December 2010) and columns III-IV show descriptives on the sample consisting of all full-time sick-leave spells started in the pre-DM period (July 2004-June 2006). Columns V-VI and VII-VIII shows similar descriptives on all spells (included graded) started in the DM period and the pre-DM period respectively. Odd numbered columns show of short-term spells, i.e. spells with a duration between 3 and 15 weeks. Even numbered columns show of long-term spells, i.e. spells with a duration of at least 19 weeks. For the long-term outcomes, income-related outcomes are measured during the second year after the year of sick-leave entry. The outcome "On temporary disability" is a dummy variables indicating whether a person receives temporary disability benefits in the second year after the year of sick-leave entry.