

The discretionary nature of sick leave

In between *fit as a fiddle* and *halfway dead* there is a grey area in which workers may be able to work or may qualify for sick leave. Within this area there is room for subjective judgment. In many cases, two workers with the same health problem and the same job will evaluate their ability to work differently. We know from everyday life that health problems are tackled in many ways, by some with courage and by others with despair. From several studies we also know that physicians differ in how they evaluate health problems; suggesting different diagnoses and different treatments in response to the same symptoms (see e.g. Englund et al., 2000; Gorter et al., 2001). The subjectivity of health makes it difficult to agree upon a common set of criteria for eligibility to sick leave and other health related benefits.

Health and work ability is not only subjective, it is also difficult to measure. Not only will people differ in how they cope with real health problems, some may even prefer pretending to have a (more serious) health problem in order to qualify for sick-leave benefits. In many cases a diagnosis cannot be set on the basis of a simple test. The most common diagnoses for sick leave in Norway are related to pain and mental illness; inherently difficult to verify. Often physicians diagnose their patients based on a (very) short dialogue. The problem of asymmetric information creates an additional challenge for the welfare state. Policy makers must take both the difficulty of *subjectivity* and that of *asymmetric information* into account when designing policies for the common good.

Economists' advice: provide economic incentives

Recently, the traditional incentive perspective has been challenged by behavioral approaches that emphasize trust and reciprocity as factors that make welfare states work, despite incentive problems. Nevertheless, the common approach in economics is to solve the type of informational problems that is related to sick leave by ensuring that workers choose to self-report their health truthfully. Any sickness insurance scheme should satisfy an incentive compatibility constraint such that workers considered able to work do so voluntarily. Such incentives can take several forms. Perhaps the most obvious one is differences in pay between work and absence. Numerous contributions have studied the relationship between sickness absence and sickness benefits. Several of these studies focus on natural experiments arising from policy reforms and document, in a convincing manner, that more generous sickness benefits increase sick leave, and vice versa (see e.g. Henrekson and Persson, 2004; Johansson and Palme, 2002, 2005).

Incentives to work can also arise from fear of being fired. If workers believe that sick leave increases their probability of being laid off or fired, and they prefer to remain employed, this provides an incentive to work. How changes in job-security affect sick leave has been studied by several authors and the effects on sick leave are convincing. In one such study, conducted by Lindbeck et al. (2006), a Swedish policy reform is considered. This reform made the seniority principle ("last one in is the first one out") more flexible for employers ("one of the three last in is the first one out"), but only in small firms. In the smallest of these firms, where job security was reduced the most, sick leave fell by around 0.4-0.65 days per employee/year. In a study of sick leave among Italian bank employees Ichino and Riphan (2005) found that when newly employed workers' probation period is over and employment contracts become permanent, sick leave increases substantially.

A third way to create incentives to work is to make sick leave less pleasant. One such strategy is to require patients to visit a physician. Most countries require a medical certificate for long-term sick leave. From a Swedish randomized experiment it is documented that the additional monitoring and “hassle” related to such a visit reduces sick leave as workers tend to return to work the day before they are required to visit a physician (see Hesselius and Larsson, 2006). A Norwegian reform in 2004 shows that also *how* physicians certify sick leave can be affected by policies, and that such policies may be important for reducing sick leave. In the second essay I evaluate the Norwegian reform and find that it reduced sick leave by more than 20 percent. One way to interpret these results is that the reform made absence more unpleasant for workers who were able to work. Physicians were required to document the *medical necessity* of inactivity more thoroughly, and to promote part-time sick leave for workers able to work part-time. I believe the success of the Norwegian reform lies in that it changed how physicians treated patients whose health falls within the grey area.

Regardless of policies, incentives to work can also be created by employers taking a worker’s sick-leave history into account when negotiating wages and considering promotions. Contracts that provide incentives over time were studied theoretically by, among others, Lazear (1979, 1981), and Freeman (1977). In the third essay I study the individual cost of sick leave in terms of future earnings and employment prospects. Using variations in individual sick leave arising from differences in physicians’ leniency, I find that sick leave has substantial consequences for subsequent earnings and employment. This is a negative result, limiting the amount of social insurance welfare states can provide. In a generous sickness insurance scheme, like the Norwegian system with full wage replacement during sickness, employers’ hands are tied in the sense that they are less able to provide intra-temporal incentives to work. However, since future outcomes such as earnings and promotions can be made conditional on past behavior, firms can instead provide inter-temporal incentives.

Against economists' advice: sick leave in Norway

Most countries' sickness insurance schemes are in accordance with "economists' advice", as sickness benefits are lower than wages. Norway is an exception, offering sick-leave benefits equal to earnings for most workers, relatively strong employment protection legislation and a generous quota of self-reported sick-leave days. Perhaps unsurprisingly to most economists, Norwegian workers have higher absence rates than workers in most other countries. Every day, more than 7 percent of Norwegian workers are on sick leave – on average. Needless to say, if one could find a way to reduce sick leave without causing other problems, it would be beneficial to the economy. Currently, sick leave has been on the top of the Norwegian political agenda. The background for this is that there has been an increasing fraction of workers on sick leave since the early 1990s, and at the same time, the number of disability pension recipients is growing. A common explanation for the rise in absenteeism during the 1990s is changing business cycle conditions. It is well known that sick leave tends to be procyclical, but the reason why has been widely debated. Business cycle changes may affect incentives to work as excess supply of labor makes unemployment more frightening. Economic fluctuations will also – by definition – affect the composition the labor force since marginal workers, entering the labor force during booms and exiting during busts, are likely to be of poorer health than those permanently employed. However, most contributions conclude that such compositional changes hardly can explain the cyclicity of sick leave (see e.g. Arai and Skogman Thoursie, 2005; Askildsen et al., 2005; Nordberg and Røed, 2006; Markussen, 2007). The rise in absenteeism since the early 1990s is studied by Biørn et al. (2010). They find that the (intra persona) increase in sick leave is even stronger than reported

in official statistics because there has been a systematic sorting of workers with poor health *out* of the labor market.

There are few indications supporting that the rise in absenteeism can be explained by poorer general health. On the contrary, based on matched data on self-reported health complaints and actual absence behavior in 1996 and 2003, Ihlebaek et al. (2007) found poor concordance between changes in self-reported health complaints and changes in sickness absence at the individual level. During a period where sick leave increased by 65 percent, self reported health complaints remained constant. Objective health measures, such as expected life length, indicate instead that health is improving.

Despite little doubt that Norwegian sick leave could be reduced by a strengthening of economic incentives, the growth in sick leave over the past two decades, which is the rationale for the current debate, is not easily explained. The Norwegian sickness insurance scheme has been kept nearly unchanged since 1978. An explanation, linking the generous system and the rise in absenteeism, must thus include more than individual sick-leave decisions conditional on incentives. One such link is the theory of changing norms in welfare states (Lindbeck et al. 2003). In short, in addition to incentives, *social stigma* also matters for individual sick-leave decisions. Following a policy reform, some individuals alter their behavior, and the social stigma associated with sick leave is changed. This in turn makes more individuals alter their behavior. Social stigma is dependent on how others behave and creates a *social multiplier*. That the behavior of others affects individual absence decisions is convincingly documented by Lindbeck et al. (2009) and Hesselius et al. (2009). Changing norms may explain the rise in absenteeism in Norway, at least parts of it. How norms are transmitted over time is an unsolved question.¹

¹ See Bjørn et al. (2010) for an attempt to disentangle changes in absenteeism from *time*, *age*, and *cohort*.

The second commonly used explanation for the rise in absenteeism in Norway is related to *structural changes* in the labor market. Such changes may have many origins. *Skill biased technological change*,² *increased competition* in product and service markets arising from liberalization of international trade, *tight public budgets* (the *Baumol effect*), and relatively *high wage floors*, may all imply that the pressure to be productive has increased, and that the tolerance for less productive workers has diminished. Whether such structural changes can explain the rise in absenteeism is not much studied empirically, and far from settled.

The general debate on sick leave is sometimes confusing. By some, sick leave is equalized with shirking. By others, sick leave is not a matter of choice but of necessity. In the first essay I suggest a more nuanced way of thinking about individual sick-leave decisions. In the proposed theory model, sick leave is a decision made by the individual *conditional* on health and incentives. In the grey area, workers must decide how sick one must be to choose sick leave. Incentives change individuals' threshold for claiming to be sick. The essay's innovation is to allow work and sick leave today to affect health tomorrow by assuming that work is healthy for healthy workers, but may delay recovery for sick workers. Health is thus an endogenous variable and the essay discusses potential trade-offs between health and sick leave as the policy makers' object is to maximize welfare, not minimize sick leave.

Our understanding of sick leave is fairly limited. When studying individual behavior by applying statistical models to micro data, one is often stunned by this realization. The degree of variation actually explained by our models is quite small – at least in my experience. In a joint work with Røed, Røgeberg and Gaure we decompose the variation in individual sick leave into a number of explanatory factors (as many as we could), including all sorts of family characteristics, individual characteristics, workplace effects and physician

² See e.g. Acemoglu (2002).

effects. Still, unexplained individual heterogeneity accounts for nearly all the variation (Markussen et al. 2009). Sick-leave behavior is strongly related to personal and often unobserved factors, and one should keep this in mind when research is presented to the public.

This dissertation has by no means been written with an ambition of covering all aspects of sick leave. As probably most researchers experience, the more work put into a field of study, the larger the field is perceived. Common for all three essays is that the action described occurs in the grey area, where absence to some extent is a matter of choice. The first essay discusses decisions in this area explicitly in a theoretical model. The second essay studies the effect of policies towards physicians set to navigate within this area when certifying sick leave. The third essay studies how variations in physicians' judgment with regard to sick-leave certification have effects on workers' subsequent earnings and employment.

Essay 1: Health and sick leave - policy makers' trade-offs

The design of optimal social insurance policies often involves difficult trade-offs. This paper discusses one such trade-off, namely that between health and sick leave. The basic idea is that work is good for healthy workers, while it may delay recovery for sick workers. This implies that if absence is "too costly", workers might go to work even when they should have stayed at home. If absence is "too inexpensive", they might stay home even when they should have gone to work. The main novelty of this paper is that it sets out a theory designed to integrate this simple idea into a more standard labor supply model, where workers trade off absence and leisure against the economic loss involved.

The model presented allows for several policy variables to affect sick leave. These are sick-leave benefits, unemployment benefits and job-security. However, key in the model is the suggested health dynamics. These are an attempt to incorporate what I believe to be most

physicians' advice to employees if they are ill: rest if you are sick, otherwise go to work. This is implemented in the model through a law of motion for health which is such that when you are ill, working will deteriorate your health, whilst working is good you when you are not ill. The model is solved by finding an agent's decision rule which specifies for which levels of health he chooses work or sick leave, conditional on the policy variables. By simulating this model one obtains predictions for health and sick leave. Some of these predictions are as expected. Low costs of absence imply higher sick leave. (Very) high costs of absence imply poorer health, as workers never really take time to recover from their illness. As costs are reduced, health improves, but in a non-linear fashion such that health improvements are large to begin with and then peter out. Interestingly, high costs of absence may also lead to high levels of sick leave - because workers with poor health "bounce back and forth" between work and sick leave. Hence, certain parameterizations of the model involve a U-shaped relationship between sick leave and work incentives.

In the final section, an application of the model is presented. The model is calibrated and used to investigate the predicted effects on health, sick leave and expected utility from changes in sick-leave benefits. The results of the model are as follows: Having no sick-leave benefits is not optimal because the increases in sick leave from more generous sick-leave benefits are small compared to the gains in health. Nor is full wage replacement during sickness optimal. By reducing sick-leave benefits somewhat, sick leave falls substantially while health is kept almost unchanged. Conditional on existing policies regarding job-security and unemployment benefits, the predicted optimal amount of wage replacement during sickness is around 70 percent.

Essay 2: How physicians can reduce sick leave – evidence from a natural experiment

On July 1st 2004 Norwegian authorities launched an apparently minor reform, changing the regulations for physicians' sick-leave certification while leaving economic incentives for workers, workplaces and physicians unchanged. At the same time, sick leave fell by around 20 percent, constituting a 1.34 percentage point increase in labor supply. This is a lot compared to any sickness insurance reform, and if a causal link can be established between the reform and the sudden drop in sick leave, the reform should be relevant for researchers as well as policy makers.

The Norwegian reform was built on the notion that inactivity (not working) is unsuitable as treatment for many kinds of illnesses, such as back pain or mental illness. Therefore, this reform was “activity oriented” in the sense that primary care physicians were instructed specifically that workers should remain active unless their medical status clearly made working not recommendable or impossible. To promote *activity as treatment*, physicians were instructed to encourage the use of part-time sick leave for workers with health problems and, if the worker had not performed any work-related activity after eight weeks of sick leave, to file a report to the Social Security Administration stating why – on medical grounds – inactivity is necessary.

I find that the reform caused a reduction in sick-leave spell length and frequency. In total, the reform reduced total sick leave by as much as 23-24 percent. Perhaps most importantly, return probabilities from full-time sick leave improved during the first 8 weeks of each spell, leading to a persistent reduction in total sick leave of more than 10 percent. This suggests that the extended documentation requirements imposed on physicians really had a substantial impact. One possible explanation why this measure had such an impact is that it

put weight behind physicians' arguments. Tired of patients demanding sick leave – reducing them to “social benefit providers” – physicians could now point to the reform and their hands being tied by the extended documentation routines, and encourage a return to work.

There was a substantial heterogeneity between physicians in how they responded to the reform. Interestingly, physicians who had not been allocated their desired number of patients were substantially less responsive to the reform. This is relevant in policy terms, as it suggests that the Norwegian market based system for patient allocation has certain undesired side-effects. The physicians' gatekeeping role may come into conflict with their business interests. This *externality* of the market based system should be weighed against any benefits from it, such as improved service, freedom of choice and reduced waiting periods.

Essay 3: The individual cost of sick leave

Despite Norwegian workers enjoy full wage replacement during sickness; sick leave is not for free. Two years later, workers with sick leave earn less and are more seldom employed than others. This paper estimates the individual cost of sick leave using Norwegian register data covering all workers, their primary care physicians and sick-leave spells certified by these physicians. To isolate causal effects of sick leave from other factors, an instrumental variable strategy is applied. The practice style, or leniency, of each primary care physician is used as instrumental variable for individual sick leave.

I find that sick leave has a substantial impact on earnings two years later. On average, a one percentage point increase in an individual's sick-leave rate reduces this person's earnings two years later by 1.2 percent. Taking into account the fact that earnings are fairly persistent over time, these effects make it relevant to ask whether sickness insurance really is insurance – or a loan against future earnings. A one-percentage point increase in sick leave also reduces the probability of being employed full-time two years later by 0.5 percentage

points. Consequently, the effects on earnings are partly caused by reduced employment.

Physicians' leniency, used as instrumental variable, is estimated in conjunction with a rich set of observables for each worker (earnings, education, age, family situation, county of residence, etc.) including a workplace dummy, and is taken from Markussen et al. (2009). Physicians' leniency, if well measured, should be a suitable instrumental variable for sick leave, capturing the variation in absence propensity due to differences in leniency among doctors. A series of robustness tests supports this, testing for confounding factors related to geography, endogenous matching of physicians and patients and the problem of *reflection* that arises because the leniency indicator in essence is a group mean in which the individual worker contributes.

The estimated effects differ among different workers. Men's earnings are much more affected by sick leave than what is the case for women. The findings also indicate that wage effects are strongest at the top of the income distribution (or education distribution) while employment effects are strongest at the bottom.

References

- Acemoglu, D., 2002: Technical change, inequality, and the labor market, *Journal of Economic Literature*, Vol. 40, No. 1, 7-72.
- Arai, M., Skogman Thoursie, P., 2005: Incentives and selection in cyclical absenteeism, *Labor Economics*, 12, 269-280.
- Askildsen, J. E., Bratberg, E., Anti Nilsen, Ø., 2005: Unemployment, labor force composition and sickness absence: a panel data study, *Health Economics*, 14, 1087-1101.
- Biørn, E., Gaure, S., Markussen, S., Røed, K., 2010: The rise in absenteeism, working paper (in progress).

- Englund, L., Tibblin, G., Svärdsudd, K., 2000: Variations in sick-listing practice among male and female physicians of different specialties based on case vignettes, *Scandinavian Journal of Primary Health Care*, 18, 48-52.
- Freeman, S., 1977: Wage trends as performance displays productive potential: a model and application to academic early retirement, *The Bell Journal of Economics*, Vol. 8, No. 2, 419-443.
- Gorter, K., de Poel, S., de Melker, R., Kuyvenhoven, M., 2001: Variation in diagnosis and management of common foot problems by GPs, *Family Practice*, Vol. 18, No. 6, 569-573.
- Henrekson, M., Persson, M., 2004: The Effects on Sick Leave of Changes in the Sickness Insurance System, *Journal of Labor economics*, Vol. 22, 87-114.
- Hesselius, P., Johansson, P., Nilsson, P., 2009: Sick of your colleagues's absence? *Journal of the European Economic Association*, Vol. 7, (2-3), 1-12.
- Hesselius, P., Johansson, P., Larsson, L., 2006: Hur påverkar kravet på läkarintyg sjukfrånvaron? Erfarenheter från ett socialt experiment, *Ekonomisk Debatt*, Vol. 2, 5-16.
- Ichino, A., Riphan, R.T., 2005: The Effect of Employment Protection on Worker Effort: Absenteeism During and After Probation, *Journal of the European Economic Association*, Vol. 3, 120-143.
- Ihlebaek, C., Brage, S., Eriksen, H. R. 2007: Health Complaints and Sickness Absence in Norway, 1996-2003. *Occupational Medicine*, 57, 43-49.
- Johansson, P., Palme, M., 2005: Moral hazard and sickness insurance, *Journal of Public Economics*, 1879-1890.
- Johansson, P., Palme, M., 2002: Assessing the effects of a compulsory sickness insurance on worker absenteeism, *Journal of Human Resources*, Vol. 37, No. 2, 381-409.
- Lazear, E. P., 1981: Agency, earnings profiles, productivity, and hours restrictions, *American Economic Review*, Vol. 71, No. 4, 606-620.

- Lazear, E. P., 1979: Why is there mandatory retirement? *The Journal of Political Economy*, Vol. 87, No. 6, 1261-1284.
- Lindbeck, A., Palme, M., Persson, M., 2009: Social interaction and sickness absence, IFN Working Paper No 725, Stockholm.
- Lindbeck, A., Palme, M., Persson, M., 2006: Job security and work absence: evidence from a natural experiment, Seminar paper no. 743, Institute for international economic studies, Stockholm University.
- Lindbeck, A., Nyberg, S., Weibull, J. W., 2003: Social norms and welfare state dynamics, *Journal of the European Economic Association*, 1, 533-542.
- Markussen, S., Røed, K., Røgeberg, O., Gaure, S., 2009: The anatomy of absenteeism, IZA Dp. 4240.
- Markussen, S., 2007: Økonomisk sykefraværsforskning: hva vet vi og hvor bør vi gå?, *Søkelys på arbeidslivet*, Nr. 1, Årgang 24, 63-81.
- Nordberg, M., Røed, K., 2009: Economic incentives, business cycles and long-term sickness absence, *Industrial relations*, Vol. 48, No. 2, 203-230.