Healer or gatekeeper?

Physicians’ role conflict when symptoms are non-verifiable

Benedicte Carlsen,1 Jo Thori Lind2 and Karine Nyborg34

Abstract

Although physicians are often expected to be gatekeepers to health insurance benefits such as paid sick leave, research indicates a substantial reluctance to reject patient requests for sickness certificates. We show that private information on the patient’s part creates a conflict between the healer and gatekeeper roles: if a patient reports subjective symptoms indicating a need for sick leave, the physician is unable to tell if the patient is truly sick or a shirker. We show that even if most physicians prefer to be good gatekeepers, all of them may trust their patients in Nash equilibrium. These ideas are illustrated using results from focus group interviews with Norwegian primary care physicians.

Keywords: Sicklisting; subjective diagnoses; asymmetric information; focus group interviews.

JEL codes: D11, D21, H42, I11, I18

1 Department of Health Promotion and Development, University of Bergen, P.O.Box 7807, N-5020 Bergen, Norway. E-mail: benedicte.carlsen@uib.no.
2 Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N-0317 Oslo, Norway (corresponding author). E-mail: j.t.lind@econ.uio.no.
3 Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N-0317 Oslo, Norway. E-mail: karine.nyborg@econ.uio.no.
4 This project was originally part of the Viabile Welfare State project at the Ragnar Frisch Centre for Economic Research, funded by the Research Council of Norway through its Programme for Welfare Research. We are grateful to colleagues and conference/seminar participants for helpful comments to an earlier version.
1. Introduction

Primary care physicians serve two roles: they are healers and gatekeepers. While their gatekeeper role may seem particularly important in welfare states, where physicians are often expected to prevent misuse of publicly financed benefits such as paid sick leave (Mykletun et al., 2014), they also participate in restricting access to benefits of private health insurance schemes.

In the medical profession, it is widely acknowledged that there exists a conflict between the healer and gatekeeper roles, and that physicians struggle with this conflict (Alexanderson et al. 2009, Carlsen & Norheim 2005; Pearson 2000; Arnesen & Fredriksen 1995; Ayres 1996; Angell 1993). For an economist, this might appear puzzling. If benefit eligibility is conditional on medical need, wouldn’t eligibility simply follow from the physician’s assessment of the diagnosis and associated recommended medical treatment?

Private information, however, complicates this picture. While only the physician can be expected to have medical expertise and access to e.g. results of blood tests and x-rays, only the patient knows her subjectively experienced symptoms such as pain, nausea or dizziness. To arrive at the correct diagnosis, thus, one may need information from both parties.

In the present paper, we show that patients’ private information on subjective symptoms gives rise to a conflict between the healer and gatekeeper roles. In particular, we study physicians’ decision of whether to grant sickness certificates qualifying patients for paid sick leave. Patients aiming to get an unjustified paid leave can potentially exploit the physician’s inability to verify their reported subjective symptoms. Even if physicians are fully aware of this, they may be unable to distinguish strategically misreporting shirkers from the truly sick. We show that in Nash equilibrium, physicians will act as if they believe every patient is reporting truthfully, unless the physician has strong preferences for being a good gatekeeper. We illustrate these conclusions using results from focus group interviews with Norwegian physicians.

Previous research on information asymmetries in the patient – physician relationship has studied physicians’ superior ability to judge patients’ need of treatment (e.g. Dranove 1988; Rochaix 1989; de Jaegher and Jegers 2000), and the information advantages of a potential patient - physician collusion vis-á-vis the insurer (Chalkley and Malcolmson 1998, Choné and Ma 2010, Alger and Ma 2003). The role of patients’ private information vis-á-vis their physician has, to our knowledge, hardly been studied within health economics.

5 Similar dilemmas may arise when prescribing medication such as opioids.
One underlying premise for our analysis is that physicians, rather than being purely selfish in the narrow sense, care about doing a good job. This is not new: In the health economics literature, assumptions of physician altruism towards patients have long been common (e.g. Arrow 1963, Farley 1986, de Jaegher and Jegers 2000, Alger and Ma 2003, Ma 2007). One reason for this is presumably that the pure self-interest model is hard to reconcile with empirically observed physician behavior (Godager et al., 2009). In the current context, however, it is important to note that physicians’ gatekeeper behavior is influenced not only by possible conflicts between their own narrow self-interest and the public interest, but also by conflicts between different aspects of the public interest, such as healing versus gatekeeping. Thus, it is not at the outset obvious whether strong social preferences makes a physician more or less strict compared to a purely self-interested one.

In the medical profession, the general consensus seems to be that when conflicts arise, a physician’s primary responsibility lies with the patient, not the insurer. The World Medical Association’s Code of Medical Ethics asks physicians to pledge that “the health of my patient will be my first consideration” (WMA 2006); the Charter of Medical Professionalism (Medical Professionalism Project, 2002) states, as its first fundamental principle, the primacy of patient welfare: “Market forces, social pressures, and administrative exigencies must not compromise this principle” (op.cit, p. 244).

Below, we first present a stylized formal model where we assume that every physician is perfectly competent: if given the full diagnostic information, she can costlessly determine the patient’s correct diagnosis. We thus abstract from incompetence, effort, and diagnostic uncertainty as sources of variation in sick-listing practices. We also disregard motives such as the physician’s preference for the patient’s social approval, assuming that physicians are motivated by income, by a desire to be a good healer, and possibly also a desire to be a good gatekeeper. We find that even for such socially responsible and perfectly competent physicians, there is an inherent conflict between healing and gatekeeping, which they will plausibly resolve by giving up on gatekeeping.

Carlsen’s (2008) focus group interviews with Norwegian primary care physicians illustrate our findings nicely, providing interesting examples of how primary care physicians tackle the gatekeeper – healer dilemma in their everyday work. Some of her findings, including previously unpublished data, will thus be presented below. While the data is not new, the issues discussed by the physicians in the study are concerned with very general aspects of the perceived healer – gatekeeper conflict. The data is based on a small set of participants and cannot be viewed as representative, but provides

---

6 More generally, work motivation can be influenced by concerns such as self-image (Benabou and Tirole 2002, 2003, 2006; Akerlof and Kranton 2005; Dur and Glazer 2008; Brekke and Nyborg 2008, 2010), vocation (Heyes, 2004; Besley and Ghatak, 2005), or a strong preference for public goods or public sector work per se (Francois 2007, Deffgaauw and Dur 2008).
a glimpse into how these primary care physicians reason when faced with patients’ private information. While the physicians were familiar with the role conflict, they mostly tackled it by trusting their patients. This is in line with several other Scandinavian surveys indicating that primary care physicians are reluctant to reject patients’ requests for sickness certificates (Nilsen et al. 2015; Wahlström and Alexanderson 2004; Carlsen and Norheim 2003; Gulbrandsen, Førde and Aasland 2002; Englund, Tibblin and Svardsson 2000; Larsen, Førde and Tellnes 1994).7

2. The model

2.1 The environment
Consider a society with a mass N of workers. A share s of these workers (0 < s < 1) are sick, while the remaining are healthy. Let health status be binary and observable only by the worker herself: either one is and feels sick, or one is and feels healthy.8 While patients observe their own subjectively experienced symptoms, they are generally unable to treat themselves if sick: they lack medical expertise, and are unable to perform and interpret the results of medical tests (such as x-rays, blood test etc.).

Let each sick worker j be characterized by a set of symptoms \( \hat{\sigma}_j \in \Sigma \) where \( \hat{\sigma}_j \) can consist of j’s subjective symptoms, verifiable symptoms, or both. Subjective symptoms are observable only by the patient, for example pain, nausea, dizziness, blurred vision, hallucinations or fatigue. We denote the set of subjective symptoms by \( \Sigma^s \). Verifiable symptoms can be observed by a physician, e.g. blood test results and interpretations of ultrasound or x-ray images. We disregard symptoms observable by neither patient nor physician, as these would not matter substantially to the argument.9

Let \( D(\hat{\sigma}_j) \) denote the diagnosis corresponding to the set of symptoms \( \hat{\sigma}_j \). In general, knowledge of both subjective and verifiable symptoms are required to arrive at the correct diagnosis. However, for what we will call subjective diagnoses, the correct treatment is based on the patient’s subjectively experienced symptoms only. ME, schizophrenia and migraine might be examples of this.

7 Alexanderson et al. (2009) report that 9 percent of Swedish physicians say “no” at least once a week to a patient who wants sick leave, 76 percent do so at least once a year, while 15 percent never do so. While these numbers may appear to indicate a stricter practice than the body of research quoted above, note that the middle category is very broad: saying “no” once a year is not very often.

8 We disregard preventive health care; the only medical reason to see a doctor is that one feels unwell.

9 “Symptom” includes negative observations indicating that nothing is wrong. We assume that the distribution, possible values and possible combinations in \( \Sigma \) are complex enough that we can disregard the possibility of an uninformed person correctly guessing a worker j’s symptoms.
Let the share workers who are truly sick, but with subjective diagnoses, $s^S < s$, be exogenously given, and let $s$ and $s^S$ be common knowledge.

Workers do not possess general medical knowledge. Hence, they have to report a set of symptoms $\sigma_j \in \Sigma$ to the physician who then determines a diagnosis. If a worker reports a $\sigma_j$ with verifiable symptoms she does not experience, this is immediately revealed by the physician. She can, however, report any set of subjective symptoms. Although workers do not have medical training, we assume that every worker knows the symptoms of at least one subjective diagnosis requiring sick leave.

Hence, if a shirker strategically misreports her subjective symptoms, pretending to suffer from some subjective diagnosis requiring sick leave, the physician is unable to distinguish the shirker from a truly sick patient.

There is a mass $\Pi < N$ of physicians. Workers decide whether or not to see a physician, and, if they do, which physician to see and which symptoms to report. Workers only visit physicians once. They report their subjectively experienced symptoms (pain, dizziness etc.) to the physician before the physician performs tests and reports the results of these tests to the patient. The physician then decides which treatment (if any) to prescribe, including whether or not to sicklist.

For some diagnoses, absence from work is required for healing. Workers with these diagnoses are entitled to sick leave. We disregard any other reason for sicklisting (from the physician’s point of view) than healing. Sick leave is not required for all diagnoses, and presence at work can also be part of the treatment.

Several simplifying assumptions allow us to focus on the effects of patients’ private information, disregarding other factors that may complicate physicians’ sicklisting decisions. First, physicians are assumed to be perfectly competent: if given correct information on a worker’s full set of symptoms, every physician can effortlessly arrive at the patient’s true diagnosis and associated correct medical treatment. Some symptoms can readily be verified by the physician (e.g., bruises, arrhythmia). Other symptoms are experienced by the patient but cannot be verified by the physician, implying that the physician has to choose whether to believe a patient or not. We denote the action of trusting $T$ and the action of mistrusting $M$. As outlined below, the physician’s behavior depends on her preferences, hence her type, and the equilibrium behavior of patients.

Second, we disregard uncertainty and ambiguity, both in the relationship between symptoms, diagnoses and treatments, and in the healing process itself: a sick worker is assumed to heal – feel well – if and only if she receives the medically correct treatment for her diagnosis.
Sick workers’ objective is to get well: that is, to maximize their probability of getting the correct treatment.\(^\text{10}\)

### 2.2 Worker behavior

A sick worker maximizes her probability of getting the correct treatment for her diagnosis. She chooses a doctor and presents a set of symptoms \(\sigma_j\) given her true set of symptoms \(\hat{\sigma}_j\).

Healthy workers have no medical reason to see a doctor. Those who do not see a physician do not present any diagnosis. We denote this by the action \(\sigma_j = \emptyset\). Some healthy workers may still see a doctor in order to get an unjustified sickness certificate. Below, the phrase *shirker* denotes a healthy worker who sees a physician with the aim of getting an unjustified sickness certificate, and who chooses which physician to see and which symptoms to report to maximize her chances of succeeding in this aim.

Shirking, however, may be costly. Seeing the doctor takes time, effort and money. There may be psychological costs of lying; workers may worry about social sanctions from their peers, reduced career prospects if absent from work, less than full insurance coverage; they may also care about social and/or intrinsic benefits of work. Such costs can differ between workers. Thus, the easier it is to get an unjustified sickness certificate, the more workers may find it worthwhile to seek one.

Let \(q \in [0,1]\) be the (endogenous) success probability for a shirker. That is, \(q\) is the probability that a healthy worker seeing a physician gets a sickness certificate, given an optimal choice of physician and symptoms to present. When physician types cannot be observed, \(q\) is the probability that a randomly chosen physician is of a type offering sick-leaves (playing \(T\) in the terminology introduced below).

Denote the pecuniary, social and moral costs of obtaining an unwarranted sick leave for individual \(j\) by \(k_j\). If the utility of being on unwarranted sick leave (compared to working) is \(B\), then a given healthy worker would sicklist if \(k_j < qB\). Denote by \(q_j = \frac{k_j}{B}\) her cut off probability (i.e. she will attempt to sicklist whenever \(q \geq q_j\)), and let \(F\) be the cumulative distribution of \(q_j\) in the population. Assume that \(F(0) = 0\) and that \(F\) is the same among sick and healthy individuals. A worker that would not shirk even if access to sick leave was certain is said to have \(q_j > 1\). Then for a given \(q\), the share of workers attempting to go on unwarranted sick leave is

---

\(^{10}\) This would be justified if, for example, nothing is enjoyable when feeling sick, making the worker willing to sacrifice any other concern in order to heal.
$$l = (1 - s)[1 - F(q)].$$

It follows that $0 \leq l < (1 - s)$. A share $c = (1 - s)[1 - F(1)]$ of workers face sufficiently large costs of shirking that they never seek an unjustified sickness certificate regardless of the probability of getting one. We have $1 - s > c > 0$.

### 2.3 Physician behavior

The physician choice we are primarily interested in is whether or not to issue sickness certificates.\(^{11}\)

Assume that physicians care about their net income, about being a good healer, and possibly also about being good gatekeeper (focusing on the patient’s private information, we disregard variation in physician effort). Let each physician’s net income per patient (monetary payment minus costs, including effort) be exogenously fixed at $a$. Let $\tau_i$ denote the physician’s strategy for granting or not granting sickness certificates (chosen from the set of all feasible rules for whom to grant a sickness certificate). The number of patients coming to see physician $i$, $P(\tau_i)$, the number of patients $i$ heals, $H(\tau_i)$, and the number of unjustified sickness certificates she grants, $G(\tau_i)$, may all depend on her sicklisting strategy.\(^{12}\) Physician $i$ maximizes her expected utility, given by

$$U_i(\tau_i) = aP(\tau_i) + hH(\tau_i) - g_iG(\tau_i).$$

Here, $h > 0$ measures the strength of the preference to be a good healer, while $g_i \geq 0$ reflects physician $i$’s preference for being a good gatekeeper. Linear separability is assumed for simplicity. Let $g_i$ have a cumulative distribution function $G$ on the proper interval $[0, \bar{g}]$ with $G'(0) > 0$ and $G' (\bar{g}) > 0$.\(^{13}\) Then the type of physician, which we initially take to be private knowledge, is given by $g_i$.

For a given set of patients with known symptoms, there would be a straightforward optimal sicklisting strategy: to establish the correct diagnosis, granting a sickness certificate if and only if sick leave is part of the medical treatment for this diagnosis. This rule would maximize the physician’s success as a healer $H(\tau_i)$, while minimizing the number of unjustified sick leaves at $G(\tau_i) = 0$ (recall

---

\(^{11}\) Note that if sick leave were the only element of medical treatment, anyone would have a fair chance of guessing a patient’s correct treatment (sick leave or not) without knowing the patient’s symptoms. We assume, however, that such correct guesses are sufficiently unlikely to be disregarded. Our implicit assumption is that establishing the correct diagnosis is a prerequisite for healing because medical treatments may involve a multitude of additional required elements not made explicit in the formal model (e.g., medication, physiotherapy, exercise, radiation, surgery). Including such elements explicitly would complicate the formal model without adding substantial insights.

\(^{12}\) The populations of both workers and physicians are modelled as continua. However, as they have the same cardinality, we get natural numbers when we divide the size of a set of the mass of workers on the size of a set of physicians.

\(^{13}\) Eq. (2) implies that doctors’ altruism towards patients, possibly also towards the insurer, is of the “impure” kind (Andreoni 1990): physicians care about their own contributions to patients’ health and to the prevention of shirking. A “purely” altruistic doctor would care about the total level of healing and gatekeeping in society, not her own contributions in particular. Pure altruism, however, yields implausible empirical predictions (Bergstrom et al. (1986); for a discussion, see Nyborg and Rege 2003).
that the payment per patient is fixed). To determine the correct diagnosis, however, private information from the patient may be required, and some patients, the shirkers, may choose to misrepresent their subjectively experienced symptoms. In addition, the physicians’ sicklisting strategy might affect the number of patients.

In theory a physician can choose an action for any specific combination of patient and reported set of symptoms \( \sigma_j \). To simplify, we focus on the following sicklisting practices for physician \( i \):

**Trust**, \( \tau_i = T \): If sick leave is part of the medically correct treatment for the diagnosis \( D(\sigma_j) \), \( i \) issues a sickness certificate. In every other case, \( i \) does not issue a sickness certificate.

**Mistrust**, \( \tau_i = M \): If sick leave is part of the medically correct treatment for \( D(\sigma_j) \), and \( D(\sigma_j) \) is not a subjective diagnosis, \( i \) issues a sickness certificate. In every other case, \( i \) does not issue a sickness certificate.

That is, trusting physicians issue sickness certificates whenever this is justified on medical grounds, based on the patient’s reported symptoms and the verified observations. Mistrusting physicians do not issue sickness certificates unless the need for sick leave can be verified, granting no sick leaves based on subjective symptoms alone.\(^{14}\)

When a share \( l \) of workers wrongfully attempt to pass as sick, a physician trusts patients if \( g_i \leq h \frac{s^x}{1-s-c} \).

Below, we show that even if most physicians do have a preference for gatekeeping, every one of them may still be trusting in Nash equilibrium. This holds whether or not patients can observe doctors’ sicklisting strategies.

### 3. Trusting and shirking with unobservable physician strategies

Our first result, formalized below as Proposition 1, is concerned with the case where patients cannot observe physicians’ sicklisting strategies. Thus, patients cannot choose their doctor based on the doctor’s sicklisting strategy; similarly, doctors cannot use their sicklisting strategies to attract patients.

In this case, there is always at least some trust and some shirking in the Bayesian Nash equilibrium. In fact, we show that unless \( \tilde{g} > h \frac{s^x}{(1-s-c)} \) all physicians trust in Bayesian Nash equilibrium. That is, the higher the prevalence of subjective diagnoses requiring sick leave, and the higher the share of

\(^{14}\) Still another candidate strategy would be to randomize between Trust and Mistrust. Since the use of mixed strategies changes very little in the analysis below, we mostly ignore this option.
workers who will never shirk, the stronger the maximal gatekeeping preference must be for any physicians at all to choose Mistrust.\textsuperscript{15}

Note that this result arises in the absence of economic or social incentives to please one’s patients, and in spite of the assumption that establishing the correct diagnosis is costless. The gatekeeping problem is caused by patients’ private information – and physicians’ resulting inability to distinguish sick from shirkers – combined with physicians’ preference to be good healers: by mistrusting, physicians deprive some of their truly sick patients the opportunity to heal, namely those suffering from subjective diagnoses requiring sick leave.

In the game, each patient $j$ chooses whether to not see a doctor ($\sigma_j = \emptyset$), or to see a doctor and present symptoms $\sigma_j \in \Sigma$. Sick workers would reporting truthfully $\sigma_j = \hat{\sigma}_j$, whereas healthy workers trying to pass for sick misreport $\sigma_j$.

Shirking is not only restricted by physicians’ choices, but also by the (implicit) decisions of workers themselves. Even if all physicians trust, thus leaving the gate open, shirking is still limited to those with $q_j < 1$.

Let $\Sigma^s$ denote the symptoms of a subjective diagnosis requiring sick leave that $j$ is familiar with. Let $q^* > 0$ be a threshold such that if $q_j \geq q^*$, worker $j$ prefers to be a shirker; if $q_j < q^*$, the worker prefers not to be a shirker.\textsuperscript{16} As the patient meets a physician whose strictness he doesn’t know and the physician meets a patient whose true health state she doesn’t know, we have to look for Bayesian Nash Equilibria. We then have the following:

**Proposition 1.** There is a cut off probability $q^*$ and cut off weight $g^*$ such that when a worker and a physician are matched without knowing each other’s identity, there is a Bayesian Nash Equilibrium where

1. Every sick worker $j$ plays $\sigma_j = \hat{\sigma}_j$
2. Every healthy worker $j$ with $q_j < q^*$ plays $\sigma_j \in \Sigma^s$
3. Every healthy worker $j$ with $q_j \geq q^*$ plays $\sigma_j = \emptyset$
4. Every physician $i$ with gatekeeping preferences $g_i \leq g^*$ plays $\tau_i = T$
5. Every physician $i$ with gatekeeping preferences $g_i > g^*$ plays $\tau_i = M$

In equilibrium, $q^* = G \left( \frac{h \cdot s^s}{(1-s-c)} \right)$ and $g^* = \frac{h \cdot s^s}{(1-s) [1 - F(q^*)]}$.

\textsuperscript{15} There is no other strategy which is strictly preferred to Trust and Mistrust.

\textsuperscript{16} That is, for a share $c$ of all workers, $q_j \geq 1$. 

9
The Bayesian Nash equilibrium can be understood as follows: Sick workers always report their symptoms truthfully, aiming to get the correct medical treatment. Shirkers misreport their symptoms, pretending to suffer from a subjective diagnosis requiring sick leave, hoping that their doctor is trusting. Physicians thus cannot distinguish shirkers from the truly sick with subjective diagnoses requiring sick leave. The probability that a patient reporting such symptoms is actually sick, depends on the aggregate prevalence of such diagnoses, $s^s$, and the aggregate share of shirkers, $l$. Physicians prefer trust or mistrust depending on these probabilities and the physicians' healing and gatekeeping preferences. Their decisions determine the share of trusting physicians $t$, which in turn determines how many healthy workers become shirkers (i.e., how many workers decide to see a doctor aiming to get an unjustified sickness certificate).

This is illustrated in Figure 1. Let $l(q)$ be the best response curve of healthy workers, in terms of the share of shirkers $l$ as a function of the share of trusting physicians $q$. We know that $l(0) = F(0) = 0$, since no shirker bothers to see a doctor if there is no chance of success; moreover, $l(1) = F(1) = 1 - s - c$.

Further, let $q(l)$ be the best response curve of physicians, in terms of the share of trusting physicians $t$ as a function of the share of shirkers $l$. We know that $q(0) = 1$: a physician trusts if $hs^s \geq g_i l$, and $hs^s \geq 0$ always holds. The largest possible value of $l$ is $(1 - s - c)$. Whether $t(1 - s - c)$ is smaller than or equal to 1, depends on whether $\tilde{g}_i > h \frac{s^s}{(1 - s - c)}$, as shown in the two panels of Figure 1.17

---

17 The steepness and curvature of the two reaction curves, and the exact location of the point $t(1 - s - c)$, depend on the properties of the distributions of $q_j$ and $g_i$. 

10
Panel a: Best response curves when gatekeeping preferences are strong
\((g^* > h \frac{s^s}{1-s-c})\)

Panel b: Best response curves when gatekeeping preferences are weak
\((g^* \leq h \frac{s^s}{1-s-c})\)

Figure 1. Physicians’ and healthy workers’ reaction curves
- Workers’ reaction curve, \(l(q)\)
- Physicians’ reaction curve, \(q(l)\)
If the insurer finds that there is too much shirking in equilibrium, and would thus like physicians to be mistrusting, there is a straightforward way to achieve that goal: the insurer can simply make patients with subjective diagnoses ineligible for sick leave. In terms of consequences, this is equivalent to making all physicians mistrust their patients, while moving the gatekeeping versus healing dilemma from the physician to the insurer. For such a change to be welfare improving, however, the welfare gain to society caused by $qLN$ fewer individuals taking unjustified sick leave must exceed the welfare loss caused by $qs^SN$ fewer sick patients being healed. To judge whether this is the case, the levels of $s^s$ and $c$ would obviously matter, in addition to factors such as individual benefits of healing and the cost of funds.

One possibly surprising feature of the equilibrium is that physician behavior does not depend on the share of the workers attempting to shirk. One could imagine physicians behaving differently in a setting where a tiny fraction of sick leaves were granted to shirkers and in one where the majority of sick leaves were granted to shirkers. To some extent, this is an artefact of the linearity of physicians’ utility functions. If for instance the share of sick leaves granted to shirkers entered, the cut off $q^*$ would depend on the composition of workers. However, there would still be a single cut off for all physicians and the main intuition would go through.

4. Trusting and shirking with observable physician strategies

Consider now the situation where patients can, for some reason or the other, observe which physicians are mistrusting and trusting, respectively. Proposition 2 demonstrates that in this case, physicians’ gatekeeping preferences must be extremely strong to make them mistrust their patients.

The reason is that, since trusting physicians are poorer gatekeepers and better healers, every patient – whether shirker or sick – prefers a trusting physician. This includes the sick with verifiable diagnoses, and the sick with subjective diagnoses not requiring sick leave: although these patients would have been treated correctly by mistrusting physicians, they do not know, before seeing the physician, what kind of diagnosis they have. If there is no limit to the number of patients a physician can accept, every patient will see a trusting physician.

The mistrusting physician must hence not only be willing to forego the lost healing opportunities for patients truly sick with subjective diagnoses requiring sick leave, as in Proposition 1 above. She must also forego all income from patients: from shirkers, the sick with subjective diagnoses requiring sick leave, and all other sick workers. In addition, she must forego the pleasure of being a good healer – for any sick patient at all, since no patient comes to her office.
Even at this substantial cost, the physician’s mistrust will not reduce the overall shirking level in society: Healthy workers who prefer shirking will simply see someone else. Since doctors who mistrust in Nash equilibrium (if any) get no patients, the share of trusting physicians $t$ is irrelevant to the share of shirkers $l$ as well as for the number of healed patients.

When patients can observe physicians’ sicklisting strategies, the share of shirkers equals $1 - s - c$, regardless of the maximal strength $\bar{g}_i$ of physicians’ gatekeeping preferences. Shirking is limited only by $c$, the share of workers who are healthy but prefer not to shirk. The gate is open to shirkers, but not every healthy worker is a shirker.

*Game structure:*

First, nature draws a type $g_i$ for every physician. The type is common knowledge. Hence her trust strategy $\tau_i$ it is also common knowledge. Assume that there is no binding limit to a physician’s number of patients $P(\tau_i)$.

Second, each worker chooses whether to see a physician, and if so which physician to visit, and which symptoms to present.

Physicians do not know the true health status of patients, but there is no strategic uncertainty and signaling, so we look for a Bayesian Nash Equilibrium.

**Proposition 2:** In the game where physicians’ type $g_i$ is common knowledge, there is a Bayesian Nash Equilibrium with a cut off probability $q^{**}$ and a cut off weight $g^{**}$ where:

i. Every sick worker $j$ chooses a physician randomly among physicians with $g_i \leq g^{**}$, reporting $\sigma_j = \hat{\sigma}_j$.

ii. Every healthy worker $j$ for whom $q_j \leq 1$ chooses a physician randomly among physicians with $g_i \leq g^{**}$, and plays $\sigma_j \in \Sigma^s$;

iii. Every healthy worker $j$ for whom $q_j > 1$ plays $\sigma_j = \emptyset$.

iv. Every physician $i$ for whom $g_i \leq g^{**}$ plays $\tau_i = T$; any physician $i$ for whom $g_i > g^{**}$ plays $\tau_i = M$.

In equilibrium, $g^{**} = \frac{s(a+h)}{1-s-c} + a$.

**Proof:** See Appendix 1.
Again, the insurer can reduce shirking by restricting the right to paid sick leave to patients with verifiable diagnoses. This would reduce the mass of shirkers by $N(1 - s - c)$; it would also reduce the mass of healed patients by $Ns^5$. As before, this would improve social welfare only if preventing a mass of $N (1 - s - c)$ unjustified sick leaves matters more for social welfare than the healing of the mass of $Ns^5$ sick workers.

The above assumes that there is no binding limit to how many patients the trusting physicians can accept. If there were such limits, even mistrusting physicians would get some patients. Their gatekeeping preferences would thus not need to be as strong as claimed above to make them prefer mistrust. Nevertheless, as long as doctors’ capacity constraint (assuming it is equal for all doctors) exceeds $N/\Pi$, mistrusting doctors would still have strictly fewer patients than the trusting ones. The only patients seeing them would be the truly sick not accepted by a trusting doctor.

5. Discussion: the model

The main result of our formal analysis is that unless physicians’ gatekeeping preferences are strong, possibly implausibly strong, the gate is essentially open. This does not necessarily mean that shirking is widespread. It does mean, however, that the final decision on whether or not to get an unjustified sick leave lies with the patient, not, as is commonly assumed, with the physician: when faced with a conflict between the healer and gatekeeper roles, physicians may be expected to give priority to the former. Thus, in our model, the physician’s main function is to be the patient’s expert medical advisor, not her keeper.

Before proceeding to the focus group interview illustrations, let us mention a few caveats. First, if there is some imperfect signal indicating whether a patient is lying – e.g., the patient is avoiding eye contact – physicians could use the following kind of strategy: “Among patients with subjective diagnoses requiring sick leave, trust only those who do not avoid eye contact.” While this would reduce the conflict between healing and gatekeeping, it would not eliminate it: as long as the signal is imperfect, the differentiation between sick and shirkers is still to some extent arbitrary. Physicians using this strategy would thus still be poorer healers than those trusting all patients. Consequently, substantial parts of our analysis would follow through as before, although with somewhat lowered stakes.

Second, the only legitimate reason for sicklisting considered above is that rest (absence from work) is required for healing. In practice, patients may be sicklisted because of limited functionality: the

---

18 Mykletun et al. (2014).
worker may simply not be capable of performing her usual tasks, even if being present at work would not have impaired her healing prospects. Empirical evidence indicates that many physicians consider functionality judgements a particularly troublesome aspect of sicklisting (Alexanderson et al. 2009). Such judgement requires, in addition to medical knowledge and information on symptoms, knowledge of the patient’s working conditions and alternative job tasks. Physicians typically do not have first-hand access to the latter kind of information. Hence this type of sicklisting decisions are even more complex than those discussed here, as they involve three parties, all of whom may have private information and diverging interests – the patient, the physician, and the employer.

Third, in our model, those who sicklist the most are the best healers. This particular conclusion depends crucially on the assumptions of 1) perfectly competent physicians, and 2) effortless establishment of the correct diagnosis (given that information is available). In practice, physicians’ competence and effort levels vary – and matter. If lack of competence and/or effort make physicians sicklist too much for patients’ own good, this may well reverse the positive correlation between healing and sicklisting. Since we have disregarded differences in competence and effort, our model is silent on the possible consequences of such differences on sicklisting.

In our view, the most interesting finding from the theoretical model is the following: physicians face an intrinsic conflict between healing and gatekeeping – even if they are perfectly competent, do not care about patients’ social approval, are not afraid to lose patients, and are willing to exert the efforts required to establish patients’ diagnoses. While social and/or economic incentives may certainly amplify the problem, as in our observable physician strategy case with patient competition, the role conflict is caused by private information on the patient’s part, not by economic or social incentives as such. Thus, it cannot be fully resolved simply through changing physicians’ incentives, nor by improvements in their medical skills.

6. Physicians’ strategies and views: Focus group interviews

As mentioned in the introduction, previous research has confirmed that physicians struggle with the conflict between their healer and gatekeeper roles (Nilsen et al. 2011; Alexanderson et al. 2009; Carlsen & Norheim 2005; Arnesen & Fredriksen 1995). Focus group interviews with Norwegian primary care physicians conducted some years ago by one of the authors (Carlsen, 2008) provide interesting examples of physicians’ strategies to handle this conflict in practice. We summarize her most relevant findings below, adding also some previously unpublished material.

The reported views may not be representative of Norwegian physicians or physicians in general, and are not suited for rigorous econometric analysis. It may still be of interest to explore whether the
strategies physicians report to be using are broadly consistent with our theoretical story – and if not, what kind of alternative stories their views may be suggesting (for a discussion of use of qualitative data, see e.g. Huberman and Miles 2002).

Details of the methodological design are provided in Appendix 2. The interviews were conducted in groups, and were concerned with the physicians’ experiences with and attitudes to sickness certification. Five focus group interviews with a total of 28 physicians were conducted. Although the sample is small, not randomized, and not necessarily representative (limitations that are normal for such studies), it is worth noting that the findings are largely in line with other quantitative and qualitative studies internationally (Englund and Svartsudd 2000; Wahlström and Alexanderson 2004; Freeman et al. 1999; Hussey et al. 2004, Winde et al. 2012).

Group interviews are often thought to be inadequate for surveying the range of different viewpoints within a group, while being better suited to expose generally accepted norms and attitudes (Kitzinger 1995). The latter was precisely the aim here. While group discussion may be dominated by ‘politically correct’ statements, this is not necessarily problematic if the aim is to depicture common norms, i.e. the professional culture of a group. Below, we restrict ourselves to report findings which represent “consensus views”, in the sense that many interviewees expressed similar views, while no-one explicitly opposed these views.

The participating physicians consistently differentiated between short and long sickness spells, defining short sick leaves as lasting up to two or three weeks. Usually, these spells were characterized as “unproblematic cases”, mostly related to bio-medically observable symptoms. Longer spells or repeated spells were frequently perceived as more problematic, and were also more often characterized by diffuse symptoms (in our terminology, including a greater proportion of subjective diagnoses).

The physicians agreed that for shorter periods, they generally issued sickness certificates without questioning the patient’s request at all. Such decisions were often made quickly, as most patients had easily verifiable symptoms of common, undramatic ailments. Several participants said that for all practical purposes, patients asking for certification for a single, shorter period of sickness absence

---

19 The Norwegian National Insurance Scheme offers full wage compensation during sick leave (except for the highest wages) for up to one year. While a sickness certificate issued by a physician is required for long-term absences, workers are allowed to self-declare sickness for absences of no more than a few days; the number of days depends on the type of employer (NAV, 2017).

20 The observable profile of the study sample was fairly similar to the population of Norwegian primary care physicians, see Appendix 2.
grant themselves the certification. None of the physicians could come up with examples of having denied patients shorter sick leaves, as this extract from one of the focus group discussions illustrates:

*Doctor A*: You rarely question the patient’s request for a sickness certification. It is hard to steer that process.

*Doctor B*: For many conditions, the patient in principle grants himself or herself sick leave; we just sign.

*Doctor C*: It is like B says: at the outset, they often sicklist themselves; I rarely deny a sick leave when they ask for it, at least.

The physicians felt that most people preferred working, and would therefore not attempt deceiving the physician, even though a minority might be trying to stay away from work for reasons other than strictly medical ones.

*Doctor D*: It is my opinion, and my point of departure, that people do want to work; hence I do not have the feeling that it is important to identify people who misuse the social security system.

When it comes to longer sickness spells, the physicians often found their role to be more complicated. The participants said they occasionally attempt to negotiate with the patient and make an effort to convince the patient of trying, sometimes little by little, to return to full time work; but even in these cases the physicians admitted to give in to the patient’s request more often than not.

*Doctor E*: At the outset, they sicklist themselves, I rarely deny a sickness certificate when they ask for it. But I do play a role in getting them back. Then I negotiate and wring and wriggle and play a much more important role, I think.

*Doctor F*: We try to make them join us in a setup where they can function, where they feel that this will work out. Then you give and take a little: “Now you try out this, and then we’ll talk in a week or two, and perhaps then you can manage to work even a bit more.” If we just say straight no, the whole situation gets locked.

In the interviews, some of the physicians claimed that they simply do not believe that patients attempt lying to their physician; firstly, because they believed few would want to shirk, and secondly, because they thought it would be difficult and uncomfortable for the patient to betray the reciprocal trust characterizing the doctor – patient relationship. However, as discussions in the groups evolved, several of the participants reflected further on this issue or were opposed by colleagues in the group, and then acknowledged that it is logical that they are probably, sometimes, deceived by patients.
Doctor G: If you arrive at the wrong diagnosis, then your conclusion about whether sick leave is required will of course be wrong too.

Doctor H: Yes, obviously. I just expose my naïveté by saying that I don’t discover much cheating.

Nevertheless, they also claimed that it is very difficult to reveal a lie indicating a subjective diagnosis:

Doctor I: It can be hard with musculoskeletal problems that you cannot really pin down. The patient may for example claim to feel pain, and you perform thorough tests, but cannot find anything, except that the patient says he or she is in pain. That is difficult. To what extent should you say that you think this is only psychological?

Additionally they felt inhibited by a sense of duty to trust the patient and center the diagnosis on the patient’s subjective experience; thus some said they had “decided to believe” the patients.

Doctor J: You cannot relate to a world in which everyone is a potential cheater – that you are standing there holding back a whole nation of shirkers who really just want to be at home. Then you could not work. A horrible thought!

Or, it could simply be too socially distressing for the physician to relate to patients while thinking of them as liars:

Doctor K: [...] otherwise you have to sit and think that this person in front of me is lying, right to my face. And we cannot sit like that all day, thinking that way.

The physicians also felt hindered from revealing shirkers by their professional responsibility as the patient’s regular primary care physician. This responsibility allegedly makes them averse to taking health risks on the patient’s behalf, and incites primary care physicians to practice defensively.

Doctor L: Then there are those times when I’ve been really suspicious, but then I’m afraid that my own prejudices play a role. Am I certain about this? And if I think this guy is a real shirker, there’s a good chance that I’m actually wrong; that this person seems unsympathetic to me does not necessarily mean that he does not qualify for a sick leave. For it has happened, sometimes, that I’ve been a bit harsh with the patient, and then they end up in a hospital policlinic and it turns out that “good heavens, man, this person has meniscrupture, I have to do an arthroscopy […]”. And I think: Oh [...], I’ve been rude to that patient.

Doctor M: I remember two rather young patients, 30 to 40 years, with substantial back pains, where I was skeptical to one of them. She has ended up being operated by specialists twice. Such experiences strengthen my conviction that I have to consider my own prejudices.
In summary, the physicians experienced the gatekeeper role regarding sickness certification as problematic, largely because they had reason to believe that a minority is granted sick leave without real medical reasons, while at the same time feeling that they had no choice but to trust their patients. Hence shorter sickness certifications were issued practically without questions asked. Longer sickness leaves were more often negotiated, and the physician would attempt to probe further for possibilities for the patient to return to work; nevertheless, these sick leaves were also usually granted in the end. Hence, the physicians in the study regretted that they were, to a large extent, unable to function as gatekeepers for sickness certification.

*Doctor N: I find this a difficult task. One feels a squeeze between what society wants and what the patient wants. Usually I feel that I work more for the patient than for society.*

7. Concluding remarks

Several studies have found that physicians struggle with the conflict between their healer and gatekeeper roles, and that as a result, gatekeeping tends to be lax (e.g., Nilsen et al. 2015, Alexanderson et al. 2009, Carlsen & Norheim 2005; Arnesen & Fredriksen 1995). If insurers wish to address this problem, they first need to understand the root of the perceived dilemma.

We show that a healer – gatekeeper role conflict arises due to the patient’s private information about her own symptoms. This information asymmetry makes the physician unable to distinguish shirkers from a subgroup of truly sick patients. Given this, keeping the gate closed to shirkers unavoidably involves closing it to some of the truly sick as well.

Our theoretical analysis demonstrates that if patients *cannot* observe physicians’ sicklisting strategies, some physicians may mistrust their patients in the Bayesian Nash equilibrium, provided that these physicians have relatively strong preferences for gatekeeping. If patients *can* observe physicians’ strategies, however, the healer – gatekeeper conflict is considerably amplified. In this case, both shirkers and the truly sick shy away from physicians who do not trust, leaving the latter with no or few patients. In this case, implausibly strong gatekeeping preferences are required to make physicians mistrust patients’ reported symptoms in the Bayesian Nash equilibrium.

Physicians’ trust can be exploited by rational and cynical shirkers who seek unjustified sickness certificates. However, since this dilemma is caused by asymmetric information, not mere sloppiness or incompetence by physicians, it cannot be removed simply by changing physicians’ incentives.

Given that professional norms among physicians stress the duty to place the patient’s interests first (WMA 2006), it seems likely that many physicians will give priority to healing over gatekeeping when
conflicts arise. If the insurer wishes to reverse this priority, this can be achieved in a very straightforward manner, simply by making the sick with non-verifiable diagnoses ineligible for sick leave benefits (which may impair their healing prospects and involve substantial distributional effects). Expecting physicians to take responsibility for such a decision, however, seems unlikely to succeed.

Among physicians participating in the focus group study reported above, the healing – gatekeeping conflict was a familiar one. Of those interviewed, none could remember to have denied a short-term sick leave at the patient’s request. While physicians reported that they sometimes, in cases of long-term sick leaves, try to persuade patients that it would be better to return to work, a leave would usually be granted anyway if they were unable to convince the patient about this. Thus, they seemed to agree that the choice of whether to be sicklisted lies, essentially, with the patient herself, while the physician primarily serves as the patient’s medical advisor. These findings are in line with previous empirical work (e.g. Carlsen and Norheim 2005; Larsen et al 1994).

If the gate is indeed open, the interesting puzzle may not be why sick leave rates are so high in welfare states with generous sick leave coverage, but rather why they are so low. In Norway, where the interviews reported above were conducted, the National Insurance Scheme is exceptionally generous, offering workers a 100% salary coverage during sick leave (although restricted for very high salaries) for up to one year. However, in addition to the possible social and intrinsic benefits of working (Brekke and Nyborg 2010) and psychological costs of lying (Gneezy et al. 2013), empirical research indicates substantial long-term pecuniary losses associated with sick leaves, even if there is formally full coverage (Markussen 2012).

The bottom line of our study is that primary care physicians’ main role seems to be that of the patient’s expert advisor, not her keeper. As advisors, they may certainly restrict sick leaves by suggesting to the patient, whenever relevant, that her healing prospects are best served by returning to work. Nevertheless, given the conflicts between gatekeeping and healing pointed out above, the expectation that physicians can fulfill both roles without substantial conflict seems unfounded.

References


Appendix 1

Proof of Proposition 2

i) By assumption, any sick worker $j$ maximizes the probability of getting the correct treatment for her diagnosis. Given physicians’ strategies, $j$ cannot do better than choosing a doctor with $g_i \leq g^*$, and hence whose $\tau_i = T$ and reporting truthfully ($\sigma_j = \hat{\sigma}$): For all diagnoses except the subjective requiring sick leave, any doctor will provide the correct treatment; if $j$’s diagnosis is a subjective one requiring sick leave, $j$ will only be treated by trusting doctors. Unless $\hat{\sigma} \in \Sigma^s$, $j$ does not know whether her diagnosis is a subjective one requiring sick leave. Hence if she chooses a doctor for whom $\tau_i = T$, she is sure to be treated correctly even if her diagnosis turns out to be a subjective one requiring sick leave. If $\hat{\sigma} \in \Sigma^s$, $j$ knows for sure that she will only be treated by a trusting doctor. Thus, all sick workers strictly prefer a doctor for whom $\tau_i = T$. By the same reasons as in Proposition 1 i), she will report truthfully ($\sigma_j = \hat{\sigma}$).

ii) Every worker for whom $q_j < 1$ is a potential shirker. Given physicians’ strategies, a shirker $j$ cannot get a sickness certificate from a doctor $i$ whose $\tau_i = M$; she can get one from a doctor $i$ whose $\tau_i = T$, but only by reporting $\sigma_j \in \Sigma^s$. Hence every shirker strictly prefers a doctor whose $g_i \leq g^*$. Due to the assumption that there are no binding limits to $P(\tau_i)$, trusting physicians share all patients between themselves. Hence by choosing a doctor $i$ whose $\tau_i = T$, and reporting $\sigma_j \in \Sigma^s$, the success probability of a shirker is 1.

iii) Every worker for whom $q_j \geq 1$ is a non-shirker. Although she may acquire a sickness certificate, her costs of doing so exceeds the benefits. Hence she plays $\sigma_j = \emptyset$.

iv) Using eq. (2), we know that a physician prefers trust to mistrust if

$$aP(T) + hH(T) - g_iG(T) \geq aP(M) + hH(M) - g_iG(M).$$

First, given patients’ strategies, $P(M) = 0$, since all patients prefer trusting physicians; $P(T) = \frac{N(s+1)}{\epsilon \Pi}$, since trusting physicians share all patients between them, and patients cannot distinguish between trusting doctors, thus choosing randomly among them. Second, a trusting doctor heals all her sick patients, hence $H(T) = \frac{Ns}{\epsilon \Pi}$; mistrusting doctors do not get any patients, and thus cannot heal anyone, so $H(M) = 0$. Third, a trusting doctor grants sick leave to all liars seeing her: $G(T) = \frac{ln}{\epsilon \Pi}$. A mistrusting doctor grants no unjustified sick leaves, so $G(M) = 0$. Inserting this into the inequality above and rearranging, gives that $i$ prefers trust whenever $s(a + h) + l(a - g_i) \geq 0$, or
\( \left( \frac{s(a+h)}{l} + a \right) \geq g_i \). As all healthy workers with \( q_j < 1 \) get a sickness certificate, we get \( l = (1 - s - c) \). Hence physicians with \( g_i \leq \left( \frac{s(a+h)}{1-s-c} + a \right) \) play \( \tau_i = T \) while physicians with \( g_i > \left( \frac{s(a+h)}{1-s-c} + a \right) \) play \( \tau_i = M. \) ■
Appendix 2

Empirical method and procedures

During the fall of 2007, 97 invitations to participate in focus group interviews were sent to a random sample of coordinators for physicians’ educational groups from the Norwegian counties of Hordaland and Oslo. Groups that responded by contacting us were included in the study consecutively. We continued interviewing groups until a point where no substantial new findings appeared; a point of data saturation (Glaser and Strauss 1967). In total we conducted five group interviews with a total of 28 physicians participating. The profile of the study sample was fairly similar to the population of Norwegian primary care physicians (see Table 1). All group participants knew each other from monthly meetings in the educational groups.

Table 1: Characteristics of the participating physicians and all primary care physicians in Norway

<table>
<thead>
<tr>
<th></th>
<th>Interview sample</th>
<th>Population of primary care physicians in Norway*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physicians</td>
<td>28</td>
<td>3891</td>
</tr>
<tr>
<td>Percentage male</td>
<td>61%</td>
<td>68%</td>
</tr>
<tr>
<td>Mean age</td>
<td>45</td>
<td>48</td>
</tr>
<tr>
<td>Mean no of listed patients</td>
<td>1172</td>
<td>1196</td>
</tr>
<tr>
<td>Percentage with vacancies on list</td>
<td>38%</td>
<td>47%</td>
</tr>
<tr>
<td>Percentage specialists</td>
<td>46%</td>
<td>56%</td>
</tr>
</tbody>
</table>


One of the authors21 conducted the interviews following a semi-structured interview guide.22 The researcher attempted to promote free discussion around the main topics of the interview guide. These topics were based on earlier Scandinavian studies of primary care physicians’ gatekeeper role in sickness certification (Carlsen and Norheim 2005, Englund, Tibblin, and Svardsudd 2000, Gulbrandsen, Førde and Aasland 2002, Wahlström and Alexanderson 2004). Interviews lasted between 60 and 90 minutes and were recorded and transcribed word for word by a research

---

21 This author is trained as an anthropologist and experienced in conducting group interviews.
22 The project was approved by the Privacy Ombudsman against the privacy and license requirements of the Data Inspectorate in relation to the Personal Data Act and Health Register Act. Participants signed an informed consent form containing information about their right to anonymity and right to withdraw from the study at any point. The participants also filled in a brief questionnaire about themselves and their medical practice.
assistant. The transcripts were then analyzed using thematic content analysis (Green and Thorogood 2004) to extract shared concerns and fundamental premises.

The interview guide included the following themes:

- Experience with the gatekeeper role
- Perceptions of the regulations regarding sick leave and the concept of illness
- Trust and mistrust in the patient-physician relationship
- Assessing patients’ diagnoses and work capacity
- Examples of difficult decisions
- Suggestions for improvements of the system