Journal of Health Economics 27 (2008) 1418–1422

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

# To wait or to pay for medical treatment? Restraining ex-post moral hazard in health insurance

Stefan Felder

*Duisburg-Essen University, Department of Economics, 45117 Essen, Germany*

ARTICLE INFO

*Article history:*
Received 8 February 2008
Received in revised form 22 May 2008
Accepted 9 June 2008
Available online 18 June 2008

*JEL classification:*
D82
H21
I18

*Keywords:*
Waiting time
Coinsurance
Moral hazard
Optimal insurance contracts

ABSTRACT

We explore the hierarchy of two instruments, waiting time and coinsurance for medical treatment, for optimally solving the tradeoff between the economic gains from risk sharing and the losses from moral hazard. We show that the optimal waiting time is zero, given that the coinsurance rate is optimally set.



## 1. Introduction

Most health insurance plans do not use indemnity payments that depend on health status, but reimburse the costs of the care actually provided to consumers. The reason is that the health status of a person can only be imperfectly observed. When insurance reimburses health care expenses, it subsidizes the price of health care at the margin, giving rise to a dead-weight loss. An optimal insurance contract thus involves copayments by the consumers. Coinsurance solves the tradeoff between risk sharing and the incentives to consume increased medical care.

User charges apply in social health insurance systems in many countries. Some services, such as dental care, may be excluded from the social insurance package, and deductibles and copayment rates are used to restrain the demand for health services. Countries with a national health service, however, do not generally employ coinsurance to curb demand for health care. They instead use waiting time for this purpose.

There is a vast literature analysing waiting times and waiting lists in the provision of health care services (for an overview see Cullis et al., 2000). Lindsay and Feigenbaum (1984) develop a model in which waiting list queues function as a rationing device. Waiting time matters because the value of health care decays the longer the treatment is postponed after the diagnosis. Iversen (1993) outlines a supply side model in which waiting lists for inpatient care are the result of a political bargaining process over resources. Iversen (1997) and Olivella (2002) analyze the effect of a private sector on waiting in the national health care system. Another strand of literature explores the scope of using waiting as a method of rationing health care in

*E-mail address:* stefan.felder@uni-due.de.

0167-6296/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jhealeco.2008.06.002

order to meet distributional objectives. Hoel and Sæther (2003) show that queuing in the public health care system might entice high-income persons to opt out and buy health care privately while still participating in the financing of the public system. If sufficient weight is given to equity, welfare can be improved by using the self-selection mechanism. In the absence of distributional motives, however, it is never optimal to have a positive waiting time for health treatment.

Smith (2005) analyzes user charges for different treatments under an exogenous health care budget. For a given treatment, all consumers have the same benefit, but, as the consumers' wealth differs and marginal utility of wealth is decreasing, coinsurance reduces the demand for health services by discouraging poorer consumers from seeking treatment. In line with optimal taxation rules, the coinsurance rate for a treatment is inversely related to the price-elasticity of demand.

Gravelle and Siciliani (2008a) extend the Smith model by introducing heterogeneity with respect to the consumers' benefits from treatment and a public health care system which uses queuing to ration demand. Prioritization among treatments again follows an elasticity rule: the more elastic demand is with respect to the waiting time, the longer the optimal wait is. Adding user charges to the model, Gravelle and Siciliani demonstrate that Ramsey waits and prices apply simultaneously. While they do not check whether the optimum requires the use of both instruments, in a subsequent paper, they find a corner solution with respect to waiting time (see Gravelle and Siciliani, 2008b). If the disutility of waiting is higher for patients with large benefits of treatment, the optimal waiting time is zero.

In this paper we address the complementary use waiting time and coinsurance in a representative consumer model, which includes health and consumption as arguments of the utility function and a linear coinsurance scheme designed to restrict moral hazard. We add the investment aspect of health, determining the consumer's income in the process, and waiting time, which reduces income since an untreated ill person cannot work. We show that the optimal coinsurance rate is positive (due to moral hazard) and smaller than one (due to the consumer's risk aversion). For a given coinsurance rate, the optimal waiting time may or may not be positive. However, if the coinsurance rate is optimally set, it is not optimal to have a positive waiting time for treatment.

## 2. The model

We consider a one-period world in which the representative consumer's health status $h$ in each state of the world $s$ depends on a composite of health services, denoted by $m$, and a random variable $\theta > 0$ which represents exogenous shocks to the consumer's health status:

$$h_s = m_s - \theta_s. \tag{1}$$

The health status determines the consumer's income in the labour market. We assume that income $f$ depends on health according to:

$$f = f(h_s); \quad f' > 0; \quad f'' < 0; \quad f'(0) = \infty. \tag{2}$$

Waiting for treatment hinders the consumer from working and reduces her income by a factor $\tau$, the waiting time, with $0 \le \tau < 1$. Since $\theta$ is not observable to the insurer, the amount the insurer pays to the consumer in each state instead depends on her purchases of health services $m$. Assume that the insurer charges a constant coinsurance rate $\lambda$, with $0 \le \lambda \le 1$. Moreover, we restrict the consumer's total exposure to risk by assuming $\lambda + \tau \le 1$.

With an actuarially fair insurance premium, the consumer's consumption flow $c$ in state $s$ is restricted by the budget constraint:

$$c_s = (1-\tau)f(h_s) - \lambda m_s - (1-\lambda)\sum_s \pi_s m_s, \tag{3}$$

where units are chosen in such a way that all commodities sell at a price of unity and $(1-\lambda)\sum_s \pi_s m_s$ is the insurance premium with $\pi$ as state-specific probabilities ($\sum_s \pi_s = 1$). The insurance premium may also represent a lump-sum tax to finance health insurance in a public system where the state is the only insurer.

The consumer's utility depends on consumption $c$ and health status $h$. For a given value of $\theta$ the consumer's utility is given by

$$u(\theta_s) = u(c_s, m_s - \theta_s). \tag{4}$$

We assume

$$u^c, u^m > 0; \qquad u^{cc}, u^{mm} < 0; \qquad u^{cm} = 0; \qquad u^c(0, m-\theta), u^m(c, 0) = \infty. \tag{5}$$

The consumer faces a two-step optimization problem. After realization of $\theta$, she decides on a utility maximizing consumption $c$ and demand for health services $m$ with a given waiting time $\tau$ and coinsurance rate $\lambda$. Ex ante she solves for the optimal $\tau$ and $\lambda$ which maximize expected utility, taking into account her ex-post decision on $c$ and $m$.

The first-order condition for the ex-post maximum in state $s$ is

$$u^m(c_s^*, m_s^* - \theta) + u^c(c_s^*, m_s^* - \theta_s)f'(m_s^* - \theta_s)(1-\tau) = \lambda u^c(c_s^*, m_s^* - \theta_s). \tag{6}$$