CEFOS REPORT 5

Risk Management and Health Care

Per-Johan Horgby

ISRN GU-CEFO-R--5--SE
ISSN 1104-327x

Contents

Foreword

In autumn 1994, Professor Lars Söderström at the Department of Economics, Göteborg University, initiated the project 'Risk Management and Health Care' at The Center for Public Sector Research. Unlike many other research projects in health economics, this project has a theoretical approach. The project analyzes health insurance arrangements in the light of recent developments in the area of finance economics and insurance theory. The present study consists of the main findings after a year of research and it is the fifth report published by CEFOS.

The report *Risk Management and Health Care* is also a licentiate dissertation in economics at the Department of Economics, Göteborg University.

Göteborg, November 1995

Prof. Lars Strömberg
Head of The Center for Public Sector Research

Author's foreword

During this study I have benefited from a number of persons. Without the inexhaustible support from Professor Lars Söderström, I would not have been able to publish this report. I am very thankful that he gave me the opportunity to work on this exiting theme and for his stimulating comments and good advises throughout the project.

A great part of the research took place at the Institut für Versicherungsbetriebslehre der Universität Hannover, Germany, and I am very grateful to Professor J.-Matthias Graf von der Schulenburg for the opportunity to visit his research institute and for his generous hospitality. In Hannover, I also had the fortune to meet Professor Franz Haslinger. I am indebted for Professor Haslinger's kindly support and that he agreed to be my co-author to one of the papers.

I would like to thank Dr. Lars-Göran Larsson, Mr. Klas Rikner, Mr. Matthias Lundbäck, Mrs. Anna Persson and others for helpful comments on seminars at the Center for Public Sector Research and at the Department of Economics, University of Gothenburg.

I have benefited greatly from discussions with professors and Ph.D.-students from the Department of Health and Society at the University of Linköping. Especially I would like to mention Professors Per Carlsson, John Christensen and Lennart Nordenfelt and Mrs. Åsa Persson, Mr. Jan Perselli and Mr. Anders Nordlund.

I have also received useful comments and advises from Professors Per-Olof Bjuggren, Jönköping International Business School; Terkel Christensen, Odense University; Louis Eeckhoudt, FUCAM, Mons; Frank A. Sloan, Duke University; Betrand Villeneuve, INSEE, Paris.

Financial support has been kindly provided by The Royal Swedish Academy of Sciences, The Swedish Institute and The Swedish Insurance Association.

I will give a special word of gratitude to my big family, especially to my mother and father whose love and thoughtfulness reach far over all borders. Last but not least, I should like to thank my wife, whose love and intellectual support have been an immense advantage for me. Annette, to get to know You is my greatest award in life.

PART ONE

Introduction to Risk Management and Health Care

1. Introduction

1.1 Background

Risk exists whenever the future is unknown. Regarding insurable risks we have a wide pattern of insurance technologies and institutions which successfully transform risky prospects to calculable consequences. Analogously, financial markets have experienced a tremendous evolution concerning transformation of financial risks. Insurance economics as well as financial theory have been productive in exploring the field called *risk management*. Risk management deals with efficient transformation of risks. A decision-maker who faces a risky prospect can choose among various risky assets to hedge against the risk considered. Insurance and financial tools of risk management are in this respect complements in the transformation process of a risk.

Since Arrow's seminal article, "Uncertainty and the Welfare Economics of Medical Care" in 1963, the economic study of medical care obtained a striking scientific demarcation. Arrow pointed out the prevalence of uncertainty in medical care and the relevance of risk-bearing. As illness to a considerable extent is an unpredictable phenomenon, the ability to shift the risk of illness to a health insurer is worth a price that many are willing to pay. If some risks are not covered, the market fails to achieve an optimal solution. The nonexistence of markets for bearing some risks reduces welfare in the society. Owing to the welfare loss, Arrow argued that the society will recognize the optimality gap and that social institutions will arise to bridge it. The medical market will therefore be characterized as regulated and intervened, where agents partly work under nonmarket behavior.

Uncertainty has now become one of the key-features of health economics. In this study, I will use Arrow's approach but also extend the notion of uncertainty in health care and especially investigate how health risks may be transformed in different ways. I shall argue that health risks can be transformed by exchange as well as by production. To reduce the financial risk of becoming ill, people exchange the risk with others through health insurance. But to reduce the risk of impairment as well as the risk of not becoming healthy again, individuals transform the health risk by investing in different forms of health care. Thus, a part of the health risk is transformed by production activities.

To discuss some aspects of health risk, I will use a risk management approach. No other areas of economics have investigated the concept of risk reduction as thoroughly as insurance and financial theory. It is therefore surprising that so few attempts have been made to enhance the development of health economics - where the concept of uncertainty plays such an important role - with basic tools of risk management.

Insurance economics and finance theory are well-established and still innovative. Resting upon seminal contributions in past decades, insurance economics and financial theory have become two major fields in economics. Today various terminologies and theorems of these fields can be used to enrich and to modulate research in health economics. This is especially urgent in the light of the need of reforms concerning financing and organization of health care.

1.2 Purpose and outline of the study

The purpose of this work is to investigate the role of uncertainty in health and health care with tools of risk management. I shall recognize the transformation process of health risks as the focal point of research and I shall introduce a theoretical framework about the economics of uncertainty in order to analyze problems associated with health risks. Thereafter, I shall present three papers concerning (i) health investment under uncertainty, (ii) adverse selection and high-cost protection, and (iii) social risk in health insurance.

The purpose has influenced the outline of the study. The work is divided into two parts. In part one, I shall give a survey over the theoretical framework to the research topics in the second part. The basic theory in the economics of uncertainty is presented (chapter 2). Owing to the importance of insurance in health care financing, I devote a section to insurance models in the classical expected utility framework and in the elegant framework of a mean-standard deviation approach. I also give a summary of the three papers (chapter 3).

In the second part, the three papers comprising the study of Risk Management and Health Care are presented. The first paper considers how people should allocate their health care resources according to portfolio theory (chapter 4). Contrary to traditional health demand models, I put the investment decision in the light of uncertainty considered here. The second paper begins with an empirical observation that the medical cost distribution is very skewed. Few people have very high costs (are bad risks) and the great part of the population utilizes very little or no medical service at all (are good risks). How this problem shall be taken care of is the main subject of the paper (chapter 5). The last essay penetrates an insurance-related problem, how social risk in health insurance is managed in two different organization forms. This paper brings statistical insurance models together with financial models of the insurance firm (chapter 6).

1.3 Methodology

In this study, I proceed from well-known theories in decision-making, insurance economics and finance. My main purpose is to discuss how these theories can be used to enrich research in health economics. Relying heavily on the economic theory of uncertainty, the study adopts an axiomatic approach. Individuals are assumed to maximize an objective function subject to certain constraints. Value is defined in a neoclassical context, in terms of ability and willingness to pay for a marginal unit of a commodity. In equilibrium, the market value of a commodity is equal to its marginal cost. All essential elements of the economics approach - stable preferences, rational choice and equilibria - apply, but I shall to some extent modify the neoclassical framework by introducing costly information, transaction costs and constraints on property.[1]

1.4 Transformation of health risks as a research issue

Since Grossman's (1972) seminal contribution to health economics, it is generally accepted that it is not medical care *per se* that consumers want, but rather health itself. Grossman´s health demand model is, however, based on the assumption of perfect information; individuals are fully informed about the inputs in the health production function. In Arrow's pioneering article about ten years earlier, economists' attention was drawn to the specific nature of the medical industry. Arrow especially emphasized the role of imperfect information and uncertainty. By combining Arrow's insights about uncertainty in medical care with Grossman's investment thinking on health, the economic study of health care adapts a risk management approach. In Figure 1.1, the risk management approach is illustrated as a function of the transformation process of health risks.

[1] Compare the extension of neoclassical economics in neo-institutional economics (Eggertson 1990).

Figure 1.1 Transformation process of health risks

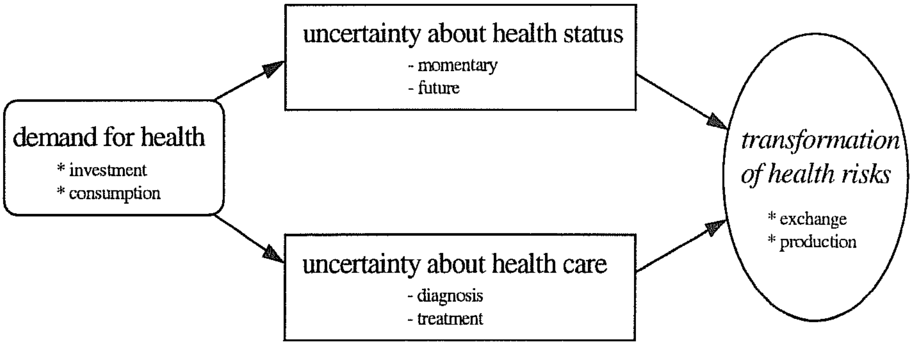


Figure 1.1 begins from the assumption that health is demanded as a consumption as well as an investment good. In respect to health and health care, many different kinds of uncertainty exist. Concerning health status, we have to deal with uncertainty about momentary as well as future health status. People do not always know their own present health status, which is a problem for correct decision-making. Uncertainty of future health status is a problem as far as incidence and severity of illness is concerned. In regard to uncertainty about health care, we have to consider the risk of bad quality of treatments and diagnoses as well as the risk of being exploited by physicians who possess superior information.[2]

To reduce uncertainty surrounding health and health care, health risks are transformed in two ways, through exchange and through production. The unpredictable incidence of illness together with high treatment costs speak in favor of insurance financing of health care. One dimension of the risk of becoming ill is reduced by exchanging the health risk with an insurer. A great part of the monetary risk associated with illness is thereby transformed via an exchange process.

Another dimension of health risks is transformed through health care production. To reduce the uncertainty about momentary health status, individuals demand information.[3] Medical diagnoses are demanded to

[2] Another important aspect of uncertainty of care is whether one has enough insurance coverage to assure qualified treatment in case of sickness. This is a question of justice, or more provocatively "right to health care", which I have discussed in Horgby (1993).

[3] An individual can believe that he is ill, but he is not, and he can be diseased without knowing it (cf. how HIV infection develops into AIDS). For accurate

reveal uncertainty about hidden diseases and to assure correct medical treatment. However, we do not exactly know the quality of diagnoses. Concerning treatments, uncertainty exists about which treatment methods are appropriate and what the consequences of the curative program will be. Physicians and other producers of health care work under conditions of uncertainty as well, and therefore *new* health risks appear in the production process. A patient who originally faces a health risk and who wants to transform the risk via medical care, has to calculate with a new therapeutic risk in the production process.

In conclusion, transformation of health risks goes through two processes, through exchange via insurance and through production via health care. This division of the risk transformation in two parts is also visible in health economics. We identify Grossman's approach as transformation by production. Likewise, Arrow's insurance approach deals with transformation by exchange.

These two approaches are not mutually exclusive, but have several points of interest in common. By setting the transformation process as the focal point of research, interdependencies between the two approaches are considered. This calls for a partly new theoretical framework as how to manage health risks in exchange and production. The present work is a step in that direction in what I have called a risk management approach to health care. The study gives a theoretical framework to the economics of uncertainty and it concentrates on some problems in the risk transformation process. It is my intention to enhance the development of health economics by focusing on some health risks and ways to handle them. At one hand I discuss new aspects of health investments by means of classical theories (chapter 4), and on the other hand, I use new approaches to well-known problems in health insurance financing (chapter 5 and 6).

---

investment and consumption, individuals demand information which can disclose the 'true' health status (Persson 1993).

2. Theoretical framework

To be able to analyze problems connected with the transformation process of health risks, we need tools of risk management. In this chapter, I will present a generalization of standard economic analysis to situations characterized by uncertainty. The presented theories of risk management will thereby serve as a theoretical framework for the three papers in part two of the study.

2.1 Decision-making under uncertainty

2.1.1 The Bernoulli principle and the von Neumann-Morgenstern utility function

The roots of modern decision theory go back to how Daniel Bernoulli (1738) treated the St. Petersburg Paradox, which was initiated by Daniel's cousin Nicholas Bernoulli. Daniel Bernoulli suggested that it is not expected gains which determine whether to participate in a game or not. Rather than using expected value, Daniel Bernoulli (Bernoulli henceforth) came up with the hypothesis that individuals possess a utility function, and that they will evaluate gambles on the basis of expected utility. By incorporating utility in the decision function, Bernoulli took the first step to what should be the standard model describing how people invest in uncertain prospects.

The basic decision model looks as follows. A decision maker faces a set of probabilistic states $(1,...,s...,S)$, for example, different states of health. Available to the decision maker is a set of actions $(1,...,x,...,X)$. If he chooses an action $x=1$ and state $s=1$ occurs, then the consequence c_{11} follows. For example, an individual has the opportunity to choose the action x = *vaccination against influenza.* If the state s = *influenza epidemic* occurs, then the consequence c = *health* follows. Hence, consequences, in the form of ultimate outcomes, follow from a chosen act *and* the realization of a particular state.

There is a crucial distinction between a preference function over consequences c and a derived preference ordering over actions x. Utility attaches direct to consequences and indirect to actions. Let us use the notation $u(c)$ to represent a person's utility function over the consequences and $U(x)$ for the utility function defined over actions. By assuming discrete states of the world and the possibility to assign to each state a probability belief (π_s)[1], the so called Bernoulli principle suggests that one should choose an action that maximizes the expected utility, that is

$$\max! \qquad U(x)=\sum_{s=1}^{S}\pi_s u_s(c), \qquad (s=1,\ldots,S) \qquad (1)$$

The utility of the action *vaccination against influenza* will, according to equation (1), be judged by the (dis-) utility of being vaccinated in other states than the state *influenza epidemic* as well. Adverse side-effects in other health states will thereby be considered.

Although Bernoulli justified his decision-principle rigorously, it was not until the work by von Neumann and Morgenstern (1947) that the expected-utility theory became generally accepted (Borch 1969, p. 54). The Bernoulli principle presupposes evaluation of utilities, which only consider expected values. Each action will in general imply a mixture of possible consequences. Therefore, in dealing with choices under uncertainty it is not evident how a ranking of consequences leads to an ordering of actions. Given some well-defined axioms of behavior,[2] von Neumann and Morgenstern showed that utility in some respect is "calculable". Hirshleifer and Riley (1992) mean that the great contribution of von Neumann and Morgenstern was to show that it is possible to construct a preference ordering function $u(c)$ whose joint use with the Bernoulli principle will lead to the correct ranking of actions.

It turns out that the Bernoulli principle is applicable if and only if the $u(c)$ function describes *cardinal* utilities to consequences. Only cardinal transformations of $u(c)$, that is positive linear, rather than positive monotonic transformations, will leave preference rankings unchanged.[3] The von Neumann-Morgenstern utility function enables us to handle utilities in the form of probabilities, and the Bernoulli principle becomes equivalent to the standard formula for compounding probabilities. This feature of the von Neumann-Morgenstern theorem has been very useful

---

[1] $\pi=(\pi_1,\pi_2,\ldots,\pi_S);\ \sum_{s=1}^{S}\pi_s=1$

[2] From four axioms about human behavior von Neumann and Morgenstern (1947, pp. 26-27, 617, 632) postulated a mathematical treatment of human preference ordering, which have been called *rational behavior*: (i) axiom of ordering; (ii) axiom of independence; (iii) Archimedes axiom; (iv) axiom of non-saturation.

[3] Machina (1987, p. 51) and Hirshleifer and Riley (1992, pp. 15-17).

in economic analysis of uncertainty, and the expected utility hypothesis has become the dominant principle in the field.

The expected utility hypothesis is the basic model in my research reports in part two of the study. The approach has, though, often been criticized. The linearity in probabilities is associated with the independence axiom. This axiom has frequently been challenged, but nevertheless, at the present time few contributions to the economics of uncertainty use non-linear models such as presented by, for example, Machina (1987).

2.1.2 *Two-moment decision model*

While the expected utility theorem assumes that an agent maximizes the expected value of a utility function defined of payoffs over consequences, the so-called mean-standard deviation approach focuses on the first two moments of random payoffs. The mean-standard deviation approach postulates that the probability distribution associated with any prospect $x=(c_1,...,c_s;\pi_1,...\pi_s)$, where c_s is the consequence in state s, and π_s is the probability of state s to occur, is effectively represented by the mean and standard deviation of consequences. The mean-standard deviation approach involves an assessment of likely returns associated with each alternative payoff and the associated risk. An individual is supposed to prefer a higher average and a lower variability of payoffs. The average payoff of consequences is measured by the mean $\mu(c)$ of any particular portfolio of assets, and the risk is measured as the standard deviation $\sigma(c)$ of payoffs. The preference function can therefore be represented by an indifference mapping on $\mu(c)$ and $\sigma(c)$ axes.

The mean-standard deviation approach has the great advantage of being a simple method of economic modeling. But in moving from expected utility values, where the probability distribution is fully defined in terms of consequences in each and every state, to a statistical summary of the mean or standard deviation of that distribution, it is obvious that some information has been lost. And henceforth, the $\mu-\sigma$-criterion and expected utility theorem is generally not consistent. In order to show coherence between the expected utility theorem and the mean-standard approach it is common that one has to assume a normal distribution or a quadratic utility function (thereby obtaining the unpleasant features of negative marginal utility and increasing absolute risk aversion).

It is true that some restrictions must be placed on either the agent's preferences or the set of random variables comprising the choice set. But neither quadratic utility nor normality is considered to be conditions which are acceptable to impose. Sinn (1983) and Meyer (1987) present another condition based on location and scale parameters of distributions.

Meyer and Sinn argue that the two-moment decision model is consistent with the expected utility model if the choice set is composed of random variables which differ from another only by location and scale parameters. The location and scale parameter condition states that the choice set should be constrained to selecting a distribution from a linear class. More formally, given two cumulative distribution functions $F_1(\cdot)$ and $F_2(\cdot)$, then the distributions differ only by location parameters α and β if $F_1(x) = F_2(\alpha + \beta x)$ with $\beta > 0$. Following the examination by Meyer (pp. 422-426), we can get some important characteristics of a mean-standard deviation approach.

Let us assume a choice set Y_i. All random variables in Y_i differ from another by location and scale parameters. Let X be the random variable obtained from one of the Y_i using the normalizing transformation $X = (Y_i - \mu_i)/\ \sigma_i$, where μ_i is the mean and σ_i is the standard deviation of Y_i. Rearranging the expression for X, we see that the random variable Y_i is equal in distribution to $\mu_i + \sigma_i X$. It is now possible to formulate the expected utility from the choice set Y_i for any individual with utility function $u(\ )$ as

$$U(Y_i) = \int_a^b u(\mu_i + \sigma_i x) dF(x) = V(\mu_i, \sigma_i), \tag{2}$$

where a and b bound the interval of X. Equation (2) defines the preference function of the mean and standard deviation approach. To analyze some qualities of $V(\mu_i, \sigma_i)$, its partial derivatives are examined.

$$V_\mu(\mu_i, \sigma_i) = \int_a^b u'(\mu_i + \sigma_i x) dF(x); \tag{3}$$

$$V_\sigma(\mu_i, \sigma_i) = \int_a^b u'(\mu_i + \sigma_i x) x dF(x) = \int_a^b \left(u''(\mu_i + \sigma_i x) \int_a^x x dF(x) \right) dx \tag{4}$$

From (3) and (4) Meyer shows that the objective function possesses the following properties:

$$V_\mu(\mu_i, \sigma_i) \geq 0 \qquad \forall\, \mu \text{ and } \sigma \geq 0, \text{ iff, } u'(\mu + \sigma x) \geq 0 \ \ \forall\, \mu + \sigma x, \tag{5}$$

$$V_\sigma(\mu_i, \sigma_i) \leq 0 \qquad \forall\, \mu \text{ and } \sigma \geq 0, \text{ iff, } u''(\mu + \sigma x) \leq 0 \ \ \forall\, \mu + \sigma x. \tag{6}$$

Let μ -values be mapped on the ordinate and σ -values on the abscissa. The property expressed in (5) implies that movements in the vertical direction in (μ,σ) space indicate higher utility. Likewise, (6) implies that movements in the horizontal direction indicate lower indifference curves. Properties (5) and (6) tell, thus, that higher mean value is preferred to lower, and that lower dispersion is preferred to higher.

To guarantee risk-aversion, the utility function has to be concave. To establish concavity of the objective function $V(\mu_i,\sigma_i)$, we must examine its second derivatives.

$$V_{\mu\mu}(\mu_i,\sigma_i) = \int_a^b u''(\mu_i+\sigma_i x)dF(x); \tag{7}$$

$$V_{\sigma\sigma}(\mu_i,\sigma_i) = \int_a^b u''(\mu_i+\sigma_i x)x^2 dF(x); \tag{8}$$

$$V_{\mu\sigma}(\mu_i,\sigma_i) = \int_a^b u''(\mu_i+\sigma_i x)x dF(x). \tag{9}$$

If $V_{\mu\mu}(\mu_i,\sigma_i)$ and $V_{\sigma\sigma}(\mu_i,\sigma_i)$ are nonpositive and if $V_{\mu\mu}V_{\sigma\sigma}-V_{\mu\sigma}^2$ is nonnegative, concavity is established. This is assured for all μ and all $\sigma \geq 0$ if and only if $u''(\mu_i+\sigma_i x) \leq 0$ for all $\mu_i+\sigma_i x$.[4]

Now we see that the mean-standard deviation approach does not have to rely on normality assumptions or quadratic utility functions. The location and scale parameter condition is sufficient to ensure consistency between the expected utility hypothesis and a two-moment decision model. Meyer has thereby succeeded in showing the coherence between the (μ,σ) criterion and the expected utility criterion. More generally, the following theorem can be stated:

> *Meyer's theorem:*
> If all distributions in the choice set differ only by location and scale parameter, that is they belong to the same linear class, then a moment-based rankings of random variables coincides with the expected utility hypothesis.

Most useful is that the location and scale parameter condition places no restriction on the functional form of the distribution describing any particular random variable. The tools provided by mathematical statistics can therefore be used even for skewed distributions, which are very common in health insurance applications. To my knowledge, few contributions have been made by using a two-moment approach to in-

[4] For proofs, see Meyer (1987, pp. 424-425, pp. 428-429)

surance economics. Sinn (1983, pp. 295-334) gives an general introduction to the theory of insurance demand and Schlesinger (1994) explains moral hazard and adverse selection in a (μ,σ) space. In chapter 4 and 5, I shall use a two-moment approach to discuss health investment under uncertainty and adverse selection in health insurance, respectively. In chapter 5, I put Schlesinger's examination of adverse selection a bit further, by analyzing effects of a governmental intervention on the insurance market.

2.2 Optimal risk-bearing under uncertainty

2.2.1 Contingent Claims Market

Walras (1874) was the first economist who constructed a formal model in which prices and quantities were simultaneously and interdependently determined. However, Walras, like many others after him, only considered optimal allocation of resources under conditions of certainty. It was first through the significant contributions of Arrow (1953) and Debreu (1959) that an extension of the equilibrium theory to conditions of uncertainty was considered in a fruitful way.

Arrow drew attention to the fact that goods can be distinguished by any characteristic that agents care about. If agents value goods differently under different circumstances, then goods can be distinguished by the state of nature in which they will be provided. Inability to act and perform daily activities make the use of, for example, sport facilities less valuable. In case of illness or impairment we rather demand curative care than athletic activities. An individual would like to transfer claims contingent upon states of nature that ensure him sports when he is healthy and medical care at time of incapacity. By combining the general equilibrium approach initiated by Walras with the expected utility theorem from von Neumann-Morgenstern, Arrow succeeded in explaining how a market can handle this kind of trade under uncertainty.

Arrow's model is characterized by preferences, endowments and markets. Agents have preferences over consumption streams and there are markets available for trading consumption at different states of the world. On markets securities are traded in form of contingent claims. A contingent claim c_s pays off if and only if the corresponding state $s = 1,...,S$ obtains. Let $u(c_{is})$ denote the i:th individual's utility as a function of claims c_s in state s. The preference ordering is described by quasi-concave von Neumann-Morgenstern utility functions ($u_i' > 0; u_i'' \leq 0$) and each individual i in the economy is assumed to maximize his expected utility over contingent claims. Due to an atomized market, individuals are price-takers and the claims can be exchanged in accordance with the market prices $p = p_1, \ldots, p_s$. Hence, given some endowment position $\bar{c}_{is}$, each individual stands for the task of maximizing expected utility (10) subject to the budget constraint (11) that he cannot invest more than his available income allows.

$$\text{Max!} \qquad U_i = \sum_{s=1}^{S} \pi_s u(c_{is}) \tag{10}$$

$$\text{s.t.} \qquad \sum_{s=1}^{S} p_s c_{is} = \sum_{s=1}^{S} p_s \bar{c}_{is} \qquad (s = 1, \ldots, S,; i = 1, \ldots, N) \tag{11}$$

Given that an inner solution exists, Arrow shows one of the cornerstones in the economics of uncertainty which Hirshleifer and Riley (1992) call "the fundamental theorem of risk-bearing":

$$\frac{\pi_s u'(c_{is})}{p_s} = \phi \tag{12}$$

The meaning of equation (12) is that at the individual's risk-bearing optimum the expected marginal utility per monetary unit will be equal in every state. Transfers in form of commodities or cash may be partial or complete. Under assumption of perfect markets and state independent utility, an individual will smooth out income streams between different states until he enjoys the same net income in every state of nature. Moreover, in accordance with the concept of Pareto-optimality to expected utilities, Arrow proves the important result that perfect competition on $s \cdot c$ markets for contingent claims is sufficient to achieve Pareto-optimality of *ex ante* as well as *ex post* sense (Dréze 1987, p. 127).

By these fundamental insights, Arrow has enriched economics with a powerful theoretical framework to analyze markets which are characterized by uncertainty. State-contingent markets may be markets in incomes, assets or goods and the analysis may be conducted using the standard techniques described above. In section 2.3, we shall more closely consider how Arrow's model can be extended to health insurance, and as we shall see, an insurance contract is a very good example of a contingent claim.

2.2.2 *Risk-bearing optimum and* $\mu - \sigma$ *preferences*

As we have discussed in 2.1.2, the $\mu - \sigma$ approach is perfectly suitable for decisions under uncertainty if all attainable distributions belong to the same linear class. Let us here consider optimal portfolio management in more detail before we in 2.3.2 use the $\mu - \sigma$ approach to extend insurance economics.

The portfolio theory is a general approach to investment decisions. The works of Markowitz (1959) and Tobin (1965) concentrate on two fundamental characteristics of investments: return and risk (measured as variability in outcome). The central theme in Markowitz' and Tobin's works was to provide an understanding of the process of spreading risk and although the theory is often expressed in terms of investment in shares, the concepts are of much wider relevance. Basically any kind of asset will do if the return distribution can be couched by the properties of mean and standard deviation. In chapter 4, we will see an application of the portfolio theory to health investments. Investment in health goes through more or less healthy activities, such as prevention, self-care or professional medical care. Since an individual has limited control over the effects of health investments, he must assess the activities on a probabilistic ground. Consequences of activities are identified in terms of the mean and variance of healthy time. By using portfolio theory, an 'health investor' can receive normative guidelines for efficient investment strategies. However, for convenience and illustration of the basics in portfolio theory, let us here consider an asset whose return is income.

In the portfolio theory it is assumed that the investor can assess at least the first two moments of probability distributions associated with alternative investment portfolios. In accordance with this assumption, for any asset a $(a = 1,\dots,n)$, let μ_a represent the mean of the income yield $\tilde{z}_a$ per unit of a held. Similarly, let σ_a represent the standard deviation of $\tilde{z}_a$ and σ_{ab} the covariance between assets a and b. If γ_a is the budget share[5] of asset a in a portfolio of n investments, then the expected return, μ, from a portfolio of the n investments, is given by

$$\mu = \sum_{a=1}^{n} \gamma_a \mu_a , \tag{13}$$

[5] If a portfolio consists of q_a units of assets a = 1,...,n and each asset has a market price $p = p_a, \dots, p_n$, then the individual's budget constraint can be written in terms of his asset holdings as: $\sum_a p_a q_a = \sum_a p_a \bar{q}_a = W$, where W is wealth devoted to investment. Each budget share then takes the form of $\gamma_a = q_a p_a / W$.

and the standard deviation in portfolio returns, σ, is given by

$$\sigma = \left(\sum_a \sum_b \gamma_a \sigma_{ab} \gamma_b \right)^{1/2} = \left(\sum_a \sum_b \sigma_a \gamma_a \rho_{ab} \gamma_b \sigma_b \right)^{1/2}, \qquad (14)$$

where ρ_{ab} is the correlation coefficient between the distributions of $\tilde{z}_a$ and $\tilde{z}_b$. Since γ_a represent proportions, we also have the following relationship:

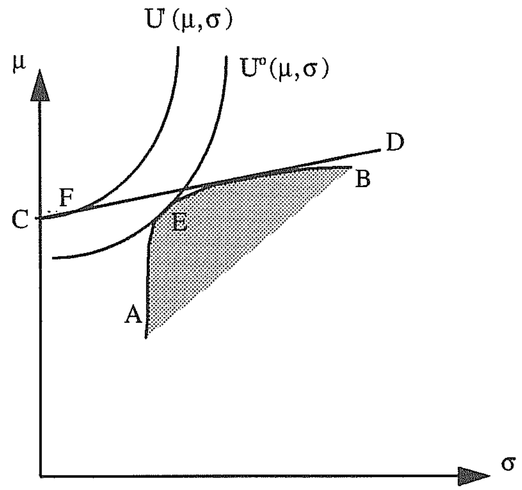
$$\sum_{a=1}^{n} \gamma_a = 1. \qquad (15)$$

The rule is to diversify assets in a portfolio which give maximum expected return, that is maximize $U(\mu,\sigma)$ subject to the budget constraint that the sum of portfolio shares equals unity (Markowitz 1952).[6] In Figure 2.1, this portfolio choice is described graphically.

The assumption of risk aversion, implies that indifference curves will be upward-sloping. The set of feasible portfolio sets will be the shaded area. Every investment plan available may be depicted by a point in the shaded area. As long as there is less than perfect positive covariance between assets, portfolios will have probability returns superior to single asset holdings. Those portfolios that are optimally diversified lie on the efficient boundary *AB*. The introduction of a riskless asset, whose return is represented by point *C*, enlarges the opportunity set to include the boundary *CD*.

[6] Investors are assumed to prefer higher expected value, i.e. $U_\mu > 0$, and given the level of μ, investors are assumed to choose the investment with lowest risk, i.e. $U_\sigma < 0$.

Figure 2.1



By maximizing expected utility over $\mu - \sigma$ subject to the efficient boundary, we can now determine the optimal risk-bearing of an individual. The investor will choose the portfolio-mix which places him on the indifference curve representing the highest level of utility. In absence of a riskless asset, the individual's risk-bearing optimum would be at point *E* in Figure 2.1. For this investor, no other strategy will dominate the portfolio obtained at point *E*.

If we also take account of the possibility of lending and borrowing money at some risk-free rate of interest, the efficient boundary is extended to *CD*. If an individual invest in the market portfolio, which is the tangent of *CD* and *AB*, and lend or borrow at the risk-free rate, he can achieve any point along *CD*. If he lends money by, for example, investing in Treasury bills, and place the rest in the market portfolio, he can obtain a better strategy than *E* at point *F*. Hence, in presence of a riskless asset, the efficient boundary is enlarged to *CD* and the optimum is at point *F*.

By introducing the riskless asset, we have introduced the main properties of the Capital Asset Pricing Model (CAPM) (Sharpe 1964, Mossin 1966). While the pure portfolio theory is strongly normative, the CAPM describes how capital assets are priced in a competitive market. The CAPM says that there is a linear relationship between the expected returns on securities and their covariance with the market portfolio. Usually this relationship is expressed in terms of beta-values, a scaled measure obtained by dividing a security's covariance by the variance of the market portfolio. The CAPM doctrine declares that investors will only be rewarded for bearing systematic risk but not for taking unsystematic risk. The key-implication of this doctrine is that all efficient portfolios will be equivalent to investment in the market portfolio plus, possibly, lending or borrowing (Sharpe 1991). In other words, all individuals will invest in the market portfolio and lend or borrowing along the line *CD* in Figure 2.1.

Although the evolution of new theories in finance is rapid, the CAPM still constitutes "the centerpiece of modern finance theory" (Hirshleifer and Riley 1992, p. 79). The CAPM rests on many simplifying assumptions, but it is easy to understand and to extend to other fields, such as insurance pricing. In chapter 6, we will use the CAPM approach to discuss the problem of social risk in health insurance. By using insights from the CAPM, we will discuss the possibilities for a health insurer to reduce social risk and lessen the safety loading required.

2.3 The economics of insurance

> "To my knowledge, there are no other major institutions in which the shift of risks through a market appears in such an explicit form as in insurance and common stocks." (Arrow 1970, p. 136)

For the purpose of becoming acquainted with risk management technologies we have been analyzing decision making under uncertainty and diversification of risks. Considering the polar case of the existence of a perfect capital market, there is no demand for specific insurance contracts. On a perfect capital market "individuals can sell shares in any claim (including human capital claims) and effectively eliminate insurable risks through diversification" (Mayers and Smith 1983), p. 305). A necessary condition for a specific insurance demand is that diversification is costly. More specifically, the costs of diversification must exceed the costs of hedging the risk with insurance. In order to compete with the capital market, insurance institutions must operate with lower transaction costs. The emergence of large insurance companies points towards their success in this respect.

The beginning of modern insurance economics goes back to the early 1960s. Almost at the same time Arrow (1963, 1965) and Borch (1961, 1962) published several important articles in the same framework as Arrow's analysis about contingent claims. Actually, the simplest way to issue a security that pays a claim in a certain state, for example, in case of illness, is to issue an insurance contract. Since the beginning of the 1960s, the economic analysis of insurance activity has become a wide research field with many exciting applications. In this section, I shall introduce some mainstream topics in insurance economics related to the outline in sections 2.1 and 2.2.

2.3.1 *Insurance models in an expected utility framework*

A large number of fundamental results in insurance economics have been derived from the linear expected utility model. Dionne and Harrington (1992, p. 4) claim that "the classical expected utility model remains the most useful approach for applications in insurance". The generalization of the expected utility model to a non-expected utility framework is an important contribution, primarily to treat small-probability events in an appropriate way. In this direction some important research contributions have recently been made.[7] But the non-expected utility analysis does not seem to imply a scientific ´revolution´ in insurance economics, "[it] just [gives] some added insights on the relationship between the assumptions of risk and insurance theorems and their conclusions" (Machina 1995, p. 37).

Let us consider an expected utility insurance model in more detail. In the expected utility theory, "consumers are presumed to form probability judgments about the likelihood of losses and to value their consequences, and then to choose the coverage that maximizes probability-weighted value" (Camerer and Kunreuther 1993, p. 7). Usually one describes a scenario, where an individual with gross income Y faces an insurable loss L, that can occur with the probability π. Let L denote the financial costs of getting ill and π the probability that illness occurs. Without any insurance an individual will meet the risky prospect of becoming Y when he is healthy and $Y-L$ when he is ill, $(Y, Y-L; 1-\pi, \pi)$. If the individual is risk-averse, we know that he is willing to pay a premium P and get reimbursed by indemnity $I(L)$ when the financial loss occurs. The premium is calculated as expected medical costs plus some multiplicative loading, $P=(1+\lambda)\pi I(L)$, where λ is a loading factor due to, for example, transaction costs or safety funding.

The individual can specify the desired amount of coverage $I(L)$ in the range of $0 \leq I(L) \leq L$. With coverage $I(L)$, the premium P is proportional to $I(L)$. Denoting the premium *rate* for each unit insurance coverage that the company charges for p, the premium can also be expressed as P =

[7] See Machina (1995) and Schlee (1995).

$pI(L)$. In other words, to reduce the risk exposure, an applicant has to pay p per unit insurance coverage. We see that the premium rate p is equivalent with the marginal cost for insurance protection.[8] Now we can formulate the individual insurance decision as to maximizing expected utility (16) subject to the market prices for risk reducing (17):

$$\text{max!} \qquad EU = \pi u(y_s) + (1-\pi)u(y_h), \tag{16}$$

$$\text{s.t.} \qquad py_s + (1-p)y_h = p(Y-L) + (1-p)Y, \tag{17}$$

where (y_s, y_h) denotes net income in the state of sickness (s) and health (h), respectively.[9] Solving this maximization problem, one obtains the following first-order conditions for a Pareto optimum:

$$\frac{u'(y_s)}{u'(y_h)} = \frac{p(1-\pi)}{(1-p)\pi}. \tag{18}$$

Equation (18) establishes the result that marginal rate of substitution between income in the state of loss and in the state of no loss will be equal, if and only if, the premium rate equals the probability rate of loss-occurrence. Hence, when $p > \pi$, marginal utility of wealth must be greater when one is sick, meaning that optimal insurance coverage is only partial, that is $I(L) < L$.

From these basic assumptions about maximizing expected utility subject to a budget constraint given market prices, a great number of insurance models have been produced.[10] However, almost every problem considered has a certain structure of asymmetric information in common. Since Coase (1937), it is known that a market with perfect information and no transaction cost will always allocate resources optimally. Given uncertainty and imperfect information, knowledge and information become strategic market possibilities that have values in themselves. The presence of informational asymmetry causes imper-

[8] Thus, the premium rate is equivalent to $\partial P/\partial I = (1+\lambda)\pi = p$. In case of no loading ($\lambda = 0$) the premium rate equals the probability value and the insurance premium is called actuarially fair.

[9] $y_h = Y - P;\ y_s = Y - P - L + I$

[10] A complete list of research topics would be very long. In the literature, the following papers appear frequent concerning insurance demand: Mossin (1968), Schlesinger (1981); insurance and portfolio choice: Mayers and Smith (1983), Briys (1988); state dependent utility: Cook and Graham (1977), Karni (1985); moral hazard: Pauly (1974), Shavell (1979); adverse selection: Pauly (1974), Rotschild and Stiglitz (1976), Crocker and Snow (1986); insurance pricing: Fairley (1979), Doherty and Garven (1986).

fections in comparison to a "frictionless" world. In the economics of uncertainty in general, and in insurance economics in particular, problems connected with asymmetric information can be divided into two parts: moral hazard due to hidden actions, and adverse selection due to hidden knowledge. The asymmetric information problem is of course valid for insurers as well as for policyholders. But, interestingly, the major contributions have been written on the possibility of opportunistic behavior by the policyholders. Only an assumption about competitive markets can justify this shortcoming.

Moral hazard

Moral hazard has been subject of intense discussions in insurance theory. The concept expresses an insurance-induced change in the policyholders behavior and it can be manifested in two different forms: (i) moral hazard *ex ante*: the individual changes his preventive efforts and therefore the probabilities of becoming ill; (ii) moral hazard *ex post*: the individual chooses more expensive treatments because the insurance covers all or parts of his health care expenditures.

It is obvious that most of us do not consciously seek to bring about illness. But the fact that we have insurance causes us to take more chances, to take too little preventive efforts, to claim more services, etc. than we would if we had no insurance. The crucial point is to which extension the insurer can observe the level of precautionary taken by the policyholder. If the insurer cannot observe the extent of action, the premium increases progressively with coverage rate as a rule. As the insurers on a competitive market cannot issue actuarial premiums under these circumstances, we know from the first order conditions in equation (18) that the policyholder chooses less than full coverage. Thus, the existence of moral hazard makes out an extra cost on the premiums, compared to the situation with no insurance.

Much empirical efforts have been invested as to how health insurance change the demand for health care. The price elasticity estimates from these studies have all the expected sign, though span a wide range from *-0.2* to *-2.0*.[11] Accordingly, by forcing consumers to share the marginal

---

[11] Studies by e.g. Phelps and Newhouse (1974), Newhouse et al. (1980), Schulenburg (1987) have suggested that the price of insurance is indeed an important determinant of demand for health services. In contrast to these studies, the RAND study used experimental policies with a coinsurance rate and a stop-loss feature (Manning et al. 1987). In a randomized experiment, they investigated how different forms of coinsurance rates, ranging from free care to care with 95 percent coinsurance, affected the level of average medical expenditures. Going from 95 percent coinsurance to free care, the average family's medical costs increased by nearly 45 percent. Even the hospitalization rates were responsive to the changes in coinsurance

cost of care, coinsurance and deductibles are effective means of combating moral hazard.[12]

Adverse selection

It is quite likely that the potential policyholders know more about their expected health expenditures than do the insurance companies. Under such circumstances, health insurers cannot differentiate between applicants. Individuals with high health risks will try to buy policies that are accounted for individuals with low health risks. If they succeed, expected health care costs will exceed the premiums paid in. As the insurers are aware of this, they know that low risk contracts will make losses. As a result, the insurers can only issue contracts accounted for health care consumption of the high risks. Because the premium is higher than the expected health expenditures for low risks, all potential beneficiaries who expect an expenditure level below the high risks' consumption will choose to self-insure instead. High risks will, thus, drive out low risks from the market, leading to an adverse selection of high risks in an insurance pool.

Adverse selection is a real problem for markets involving health insurance. It expresses "the tendency of high risks to be more likely to buy insurance or to buy larger amounts than low risks" (Cummins et al. 1983). The inefficiency that arises due to adverse selection is that low risk individuals cannot obtain as much insurance coverage as they wish. In addition, since low risks face a favorable premium price, high risk individuals will insure against risks that they would not otherwise insure against.

The adverse selection hypothesis has been tested in a number of empirical studies concerning health insurance. The findings are neither homogeneous nor completely consistent with the adverse selection theories. Nevertheless, empirical evidences seem to support the thesis that low risks cannot obtain as much insurance coverage as they wish due to adverse selection.[13]

---

rates. Hospitalization increased by about 33 percent due to a reduction of the coinsurance rate from 95 percent to zero. The RAND study found that the price elasticity of medical care was approximately *-0.2*.

[12] In this context, it is worth noting, that the spread of health insurance cannot be said to be responsible for the cost explosion in the health care sector. As Manning et al. (p. 269) and Weisbrod (1991) have pointed out, the driving force of health care expenditures is the development of new medical technologies in *combination* with third-party-financing.

[13] According to Price and Mays (1985), adverse selection has raised the premiums for high risks and lowered the premium for low risks, relative to premiums in absence of any selection. Likewise, Marquis and Phelps (1987) and Marquis (1992) found that conditions for adverse selection seem to exist in a health insurance mar-

In order to reduce the adverse selection problem, insurers have developed mechanisms such as risk categorization[14], experience rating[15] and partial insurance coverage[16]. One interesting result is that the introduction of compulsory insurance covering just a part of the costs and charging uniform premiums can potentially lead to a Pareto improvement (Eckstein et al. (1985), Strassl (1988)). In this case the insurance market will be characterized as a mixed system, with a incomplete cumpolsory insurance scheme and a private market for supplementary insurance.

By introducing a compulsory insurance scheme, high and low risk individuals are pooled together. The mandatory insurance subsidizes, thus, the premiums for high risks. This makes out an utility loss for the low risks, who have to pay a higher mandatory premium than their expected costs. However, the subsidization increases the expected utility for high risks. The willingness for high risk individuals to buy supplementary insurance calculated for low risks is diminished. In this way, the self-selection constraint, that low risk individuals can only buy low risk contracts, will be eased. Now low risks can buy supplementary insurance coverage at actuarial terms, and the end allocation might be Pareto superior over a completely private insurance market.

Although an unanimous picture of the effects of the compulsory insurance is lacking, the prerequisite for a Pareto improvement is that

---

ket. Also Brown (1992) and Brown and Doerpinghaus (1993) conclude their empirical studies with findings in favor for adverse selection in the health insurance market. By using all available risk factors in the premium setting, Van de Ven and van Vliet (1995) argue that a premium can be risk-adjusted to the health risk considered. In an empirical study, they conclude that there need not be any substantial unavoidable consumer information surplus, and thus no "essential adverse selection". Nevertheless, as long as premiums are not risk-adjusted, for any reason, the consumer information advantages may result in adverse selection.

[14] Insurers can classify risks using many variables, such as the easily verified variables of age and sex. Crocker and Snow (1986) showed that imperfect categorization always enhances efficiency if information is costless (as for age and sex), but if classification is costly, the efficiency gains were ambiguous.

[15] In models concerning experience rating, insurers use information related to past experience of the insureds in order to motivate them to reveal their true risk. See for example Dionne (1983), Cooper and Hayes (1987).

[16] In the literature about partial insurance coverage, two types of insurance policies can be observed. Pauly (1974) introduced so-called "price-only" policies, that means insurers charge a uniform premium rate per unit of coverage to all applicants. The other type is the "price-quantity" policies, introduced by Rothschild and Stiglitz (1976), where insurers are assumed to offer a menu of policies with different prices and quantities so that different risks choose different insurance policies. The "price-quantity" approach is, for example, used in comparative statics about the usefulness of a governmental intervention (cf. Horgby 1995).

compulsory insurance should be incomplete. For a health insurance this means that only a part of the medical care expenses should be covered, with the remainder left to individually contracted complementary insurance. Interestingly, few attempts have been made to try to explain what 'partial' means in insurance terms. In the health insurance literature, authors often only make reference to some notion of need to limit the social side of health insurance.[17] In chapter 5, we will see another approach to limit the compulsory health insurance schemes, resting upon the notion of the state as superior risk-bearer for losses which occur with low frequency and at high cost. This so-called high-cost protection reduces not only adverse selection but increases also the allocation efficiency on the market.

Social risk

We have now discussed the main problems in insurance markets which can be derived from informational asymmetries. Several other reasons can be given why the market does not allocate risks optimally.[18] Hirshleifer and Riley (1992) have shown that because of the fact that total wealth of society will be lower following systematic losses, full insurance cannot be obtained on every potential loss. This kind of risk is termed social risk.

Social risk is not only a result of small numbers where the law of large numbers cannot work fully. If the risks in an insurance pool are on average positively correlated, then even with large numbers the variance of the average loss does not approach zero. Owing to social risk the totals of premiums paid in could not always match the total indemnities payable, and an insurer has to consider the risk of bankruptcy.

The risk of insolvency is in most countries taken care of by regulation. The rationale for solvency regulation is that consumers are not in the position to monitor the risk of insurer default. Regarding health insurance the regulators have forced the insurers to organize business around one single health insurance line, such as Health Maintenance Organization in the USA, Gesetzliche Krankenversicherung in Germany or Family Doctors in Sweden. However, in chapter 6, we will see that the problem of social risk is larger in such a line of business. This is because of lim-

---

[17] See further chapter 5.4.

[18] Schlesinger and Doherty (1985) list several other reasons than mentioned here for markets to be incomplete: general market risk, transaction costs of insurance, search costs for insurance, nonmarketable assets. These market imperfections make out another sort of risk, the so-called background risk (Eeckhoudt and Kimball 1992) in insurance decision making. Briys and Viala (1995) consider how a background risk affect the Pareto optimal design of an insurance policy.

ited diversification possibilities, compared to a health insurer who spreads the risks with other risks than health risks and financial devices.

2.3.2 *Insurance models in a two-moment framework*

In order to utilize the ease and elegance of a two-moment decision model, I will here apply the mean-standard deviation approach to insurance economics. First we change the fixed accident loss L to a loss distribution $\tilde{L}$. The net income will then take the following form:

$$y = Y - \tilde{L} + I(\tilde{L}) - P\left[I(\tilde{L})\right], \tag{19}$$

where Y is the gross income, I is the insurance coverage and P is the premium. The insurance coverage, as well as the premium, is now a function of the loss distribution. Let us consider an insurance with proportional coverage rate α, that is an insurance with a coinsurance arrangement. In the presence of coinsurance, the premium will also be proportionally adjusted to the coverage rate and is calculated as the expected accident costs, $P = \alpha E(\tilde{L})$. By assuming a competitive market without any administrative costs, we neglect the impact of loading costs. We can now express the expected income with insurance protection as:

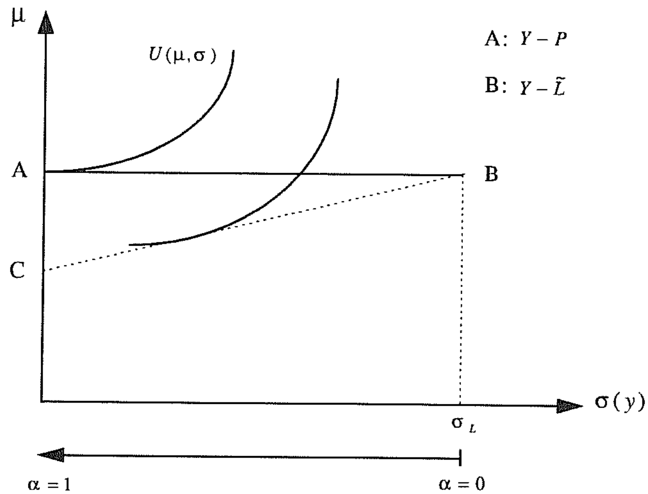
$$E(y) = \mu = Y - E(\tilde{L}) + \alpha E(\tilde{L}) - \alpha E\left[I(\tilde{L})\right], \tag{20}$$

with the standard deviation

$$\sigma(y) = (1-\alpha)\sigma(\tilde{L}). \tag{21}$$

From equation (20), one can see that for each coverage rate α the mean income μ will only change by a constant and a proportional factor. The local and scale parameter conditions are thereby fulfilled and we do not have to rely on quadratic utility functions or normal assumptions (Schlesinger 1994, p. 120). By solving for α in (21), equation (20) can be used to express the mean income as a function of expected losses and standard deviation of income. It is thus possible to analyze different combinations of insurance coverage in a mean-standard deviation space (see Figure 2.2)

Figure 2.2



Suppose an applicant has the opportunity of choosing the degree of coverage α. He can choose different levels of coverage at different prices along the budget constraint connecting points *A* and *B* in Figure 2.2.[19] If he does not buy any insurance ($\alpha = 0$), then the end-of-period wealth distribution is represented by point *B*. Every higher degree of coverage will lead to a lower dispersion of wealth. Assuming risk aversion, an individual will buy some insurance coverage mapped on the line *AB*. Excluding any kind of loading, the individual's optimum is characterized by a full coverage (point *A*), and hence no dispersion of wealth. This is the same solution as in the framework of expected utility in the previous section. In case of some loading factor, the budget constraint would have a positive slope, such as *CB*, and the optimal allocation would be a policy with a coinsurance arrangement.

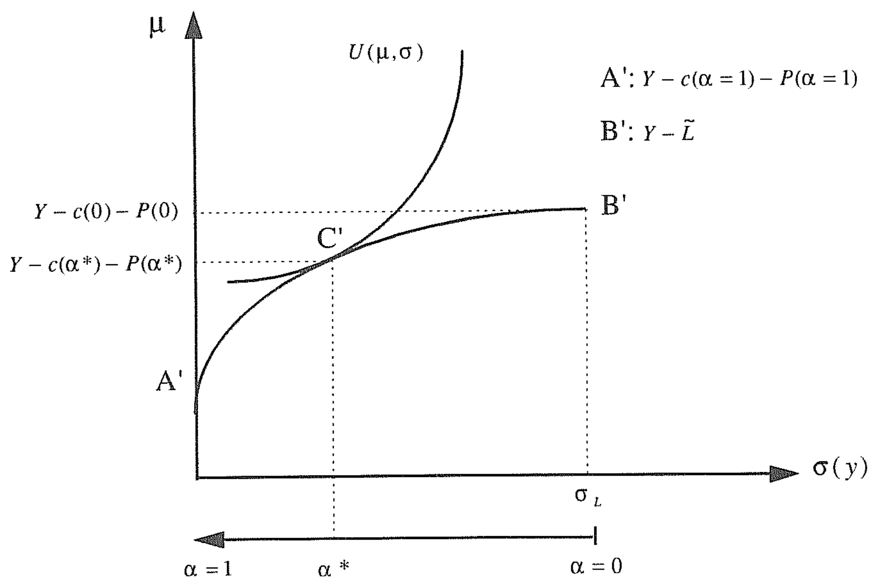
If an individual can influence his own probability loss, the problem of hidden action emerges. We make two plausible assumptions. First we assume that prevention is costly, and second, that the optimal level of prevention decreases with coverage rate.[20] The insurer knows that the more insurance coverage the policyholder purchases the more likely he is to take less precautionary activities. As the prevention intensity de-

[19] This illustrates, thus, a "price-quantity" policy.

[20] It is assumed that increases in prevention reduce the likelihood of the loss to happen, i.e. $\pi'(c) \leq 0$, where c denotes the costs of prevention. The impact of prevention is though diminishing, $\pi''(c) \geq 0$.

creases with coverage, the expected costs will increase. This means that the premium price must change with the coverage rate. More specifically, the marginal cost of a comprehensive insurance protection will be greater than the marginal cost of a lesser insurance coverage, $(p(\alpha_0) < p(\alpha_1); \alpha_0 < \alpha_1)$. Graphically this implies that the budget line cannot be a straight line as in Figure 2.2, but has to be concave, see Figure 2.3.

Figure 2.3



As we now consider the case of costly prevention, we have to rewrite the budget line (20) as

$$E(y) = \mu = Y - c(\alpha) - E(\tilde{L}) + \alpha E(\tilde{L}) - \alpha E\left[I(\tilde{L})\right], \tag{22}$$

where $c(\alpha)$ denotes the costs of prevention as a function of insurance coverage. The logic behind the concave budget line is straightforward. A more comprehensive insurance coverage leads to lower expected accident costs which implies lesser prevention taken. Therefore, the premium price must be higher per insurance unit as the insurance coverage increases. As a result, the premium expenses for full coverage will exceed the expected losses. And as we knew previously, under such circumstances a rational insurance buyer will not demand complete risk elimination, but rather bear some risk of his own in terms of coinsurance. Hence, optimal allocation of risk-reducing investment is characterized by an incomplete insurance when policyholders can take preventive care.[21] This optimum is shown at point C' in Figure 2.3. Owing to this theoretical analysis of moral hazard, we can conclude that an insurance changes incentive to take prevention. The changed behavior makes out a new cost for the insurer, which implies that the insurer cannot issue actuarial policies.

Even the problem of adverse selection is set in another perspective by a mean-standard deviation approach. As a major part of chapter 5 in part two of the study is devoted to this presentation, I refer the reader to that section.

[21] Shavell (1979) emphasizes that the cost of care has a major impact on the optimal solution. If the cost of taking care is too high, then a complete insurance protection is optimal, since the cost situation makes care priceless. And if the cost of taking care tends to zero, the level of coverage approaches full coverage, since the risk exposure will be decreased at very low costs. In between these polar cases partial rather than full coverage is optimal.

3. Summary

The purpose of the study is to investigate the role of uncertainty in health and health care with tools of risk management. This has been done in two separate parts.

In part one, I give an introduction to Risk Management and Health Care. I consider the transformation process of health risks as the focal point of research. Moreover, I present a theoretical framework of the economic study of uncertainty. Especially, a new approach to the $\mu-\sigma$-criterion is introduced and how it can be used in insurance economics.

In part two, three papers comprising the theoretical study of Risk Management and Health Care are presented. Each paper consider an aspect of the transformation process of risk; the first paper how an individual should invest in health under uncertainty; the second paper how adverse selection in health insurance can be reduced by a governmental high-cost protection; and the third paper how social risk in health insurance can be managed efficiently.

The present work has enlarged the economic study of health care in two ways; at one hand by focusing on new aspects of health investment by means of classical theories (paper I); and on the other hand by using new approaches to well-know problems (paper II and III).

3.1 Paper I (Chapter 4)

This report examines health investments under uncertainty. The traditional approach to dealing with health investments relies heavily on concepts from the consumer theory. A common feature of the production models of health is that they neglect uncertainty about the input variables. They thereby overlook the importance of uncertainty in health investment decisions.

In the traditional insurance literature, the proper context for a discussion of prevention and cure has been accepted as the medical one. The consequence of such acceptance of medical care is the classification and legitimization of causes on a biological rather than social, cultural and psychological basis. In dealing with moral hazard, the literature normally considers just one single variable, as "level of care" for moral hazard *ex ante* or "amount of medical service" concerning moral hazard *ex post*. Multi-activity decisions and other treatments than medical care are therefore excluded from the health insurance analysis.

In this paper a mixed approach is considered. From the production models of health, the investment-strategy attitude towards health is applied, and from the insurance models, the concept of uncertainty is applied to the input variables. It is assumed that individuals can undertake many activities which influence the probability of being healthy. Activities are graded according to their return *and* to their risk. People are assumed to be risk-averse maximizers of healthy time and they construct a portfolio of activities which yields the highest return of healthy time for a given risk. The concept of diversification in the portfolio theory is used. In order to reduce risk associated with activities, consumers should hold diversified portfolios of competing activities. Owing to the power of diversification, almost all efficient health portfolios represent multi-activity behaviors.

In contrast to traditional production models of health, this model considers uncertainty of input variables and contrary to traditional health insurance models, prevention and cure are subject to multi-activities decisions and characterized in terms of bundles of activities.

3.2 Paper II (Chapter 5)

This research report considers adverse selection in a mixed health insurance system, with many private health insurers and a single governmental high-cost insurer. The purpose of the paper is to theoretically discuss the gains of state interventions in the health insurance market in regard to adverse selection. Many seminal contributions to the adverse selection problem are based on simple loss processes and insurance contracts. The policyholders typically meet the same uncertain loss and differ only in their loss probabilities. By doing so, the authors overlook the importance of distributional risk in insurance pricing.

The medical cost distribution per capita consumption is skewed and extremely heavy-tailed. The skewed medical cost distribution illuminates the heterogeneity and variability of health risks. This brings about two difficulties. First, how can the market discriminate between policyholders with different risks; and second, how can the insurer set accurate premium prices if an average cost varies?

In the paper, the gains of a mixed system is shown, where the state insures the high-costs in the long tail of the cost distribution, leaving the low-cost range to the private health insurance market. The point is that if the state insures for high medical costs, the rest of the risk distribution gets closer in shape to a normal distribution. The "normalization" transforms the originally heterogeneous risks to more homogeneous ones, thereby reducing the adverse selection effects. Moreover, public coverage for severe illness interacts with private coverage for less severe illness as the high-cost protection reduces the variance of an average health risk in the private market.

Important findings indicate that high-cost protection reduces adverse selection and improves market efficiency. The market for health insurance will be characterized as a mixed system, with a governmental high-cost insurer and many private low-cost insurers.

In the paper a two-moment approach based on location and scale parameter conditions is used. An additional contribution of the paper is to show the usefulness of a two-moment approach to insurance problems.

3.3 Paper III (Chapter 6)

In this paper we consider health risks that are correlated between individuals. A frequent example is the annulally returning influenza, striking large parts of the population. A non-frequent example is radioactive downfall following an accident in a nuclear power plant. The individual health risks are in these cases not perfectly offsetting and an insurance pool cannot eliminate all risk through diversification. The part of an individual risk exposure that cannot be diversified within an insurance system is called social risk.

The presence of social risk causes distortion in the health insurance market. The purpose of this paper is to discuss the consequences of social risk in health insurance and how it is administrated by two different organizations, a mutual insurance pool and by an insurance entrepreneur.

A mutual pool for health insurance offers diversification among similar health risks. When policyholders systematically become sick at the same time, the solvency of the mutual pool is put under pressure. To hedge against insolvency, a buffer fund per policy will be charged. Hence, owing to social risk, health insurance premiums will become loaded and an insured will choose an incomplete coverage.

An insurance entrepreneur working according to risk management technologies is characterized by looking for optimal diversification opportunities. Instead of diversification in a pool consisting only of health risks (economies of scale), the entrepreneur uses also the capital market as a diversification pool (economies of scope). The analysis suggests that a competitive market for health insruance would benefit from economies of scope in addition to scale production.

PART TWO

Three Papers in Risk Management and Health Care

4. Individual Risk Management and Health Investments*

4.1 Introduction

The traditional approach to dealing with health investments relies heavily on concepts from consumer theory.[1] The focus is on health as a consumption as well as an investment good. In equilibrium, the marginal utility of holding a unit of health must be equal to its marginal user-cost consisting of interest and depreciation. A common feature of production models of health is, however, that they neglect uncertainty surrounding the input variables.[2] They thereby overlook the importance of uncertainty in health investment decisions.

In the traditional insurance literature, the proper context for a discussion of prevention and cure has been accepted as the medical one.[3] The consequence of such acceptance of medical care is the classification and legitimization of causes on a biological rather than social, cultural and psychological basis. In dealing with moral hazard, the literature normally considers just one single variable, as "level of care" for moral hazard *ex ante* or "amount of medical service" concerning moral hazard *ex post*. Multi-activity decisions and other treatments than medical care are, therefore, excluded from the health insurance analysis.

In this paper a mixed approach is considered. From the production models of health, I am influenced by the investment-strategy attitude towards health, how an individual consciously chooses to perform health-related activities. And from the insurance models, I adopt the

---

* Paper presented at the third European Conference on Health Economics, Stockholm, Sweden, 20-22 August, 1995. Original title: "Individual Risk Management and Health Investments - A Portfolio Approach".

[1] See, for example, Grossman (1972), Muurinen (1982), Wagstaff (1986, 1993).

[2] Exemptions are Dardoni and Wagstaff (1990) and Zweifel (1992). Dardoni and Wagstaff follow Arrow's view about uncertainty surrounding medical care. Other input variables in the health function than medical care are, however, not examined. Zweifel argues that the production function for health is dependent on the state of health prevailing during the decision period. In addition, Zweifel claims that production of health amounts to a modification of transition probabilities which are essentially governed by chance. This paper shares his opinion.

[3] Concerning health insurance, see, for example, Arrow (1963), Nordquist and Wu (1976), Breyer and Zweifel (1992). Concerning general problems of moral hazard, seminal contributions have been made by Pauly (1974) and Shavell (1979).

concept of uncertainty and apply it to input variables. This approach extends, thus, investment models with uncertainty and insurance models with activities other than only medical care.

4.2 The portfolio approach to health behavior

In many of the everyday decisions which we make, we expose ourselves to the possibility of a uncertain outcome. Such decisions can be characterized by the presence of choice and risk. In the context of health investments, the investor is faced with a wide range of more or less healthy activities. All these activities determine the health behavior of an individual. Moreover, it lies in the nature of activities that one does not know how they affect the probability of being healthy. People imagine that running is good for their health, but they are not sure about how much and under what circumstances. It is also said that some alcohol can prevent strokes, but there are many doubts about this. Living in a world with incomplete information and knowledge, we still act and invest our time and money in different kinds of activities. It is trivial but worth noting that we usually do not perform only one activity. We have an intuition that it is good to do several activities, an intuition about spreading our risks. This cluster of activities makes up the health behavior of an individual. And to analyze health behavior, I will use a risk management approach by extending the portfolio theory (Markowitz 1959) to health investment decisions.[4]

4.2.1 Towards a portfolio model of health

Activities can be performed for different reasons. It is obvious that an activity can please more than one need, as well as resulting in adverse side-effects. Investments in health not only add increments to the stock of health, but produce satisfaction and enjoyments in life. The well-being aspect of activities should not be neglected. Sports are for some people the essence of life, while others suffer just thinking about it. Decisive to performing sports are not only the biological health effects, but also the pleasure and well-being of doing it.

Focusing first on the return of an activity (a), it is natural to think of gains in term of healthy time. As in traditional health production models, an activity yields utility in the current period as a consumption

[4] A recent interesting application of the portfolio theory is given by Iannaccone (1995), who uses the portfolio theory to explain religious behavior.

good and effects the health stock at the end of the period as an investment good. Health has a consumption feature because good health makes people feel better and health is also desired as an investment good because it increases the number of healthy days available to work and to earn income.

If an individual wishes to increase his probability of being healthy from one period to the next, he must undertake some health investments. It is the time profile of health that individuals select and it is achieved by investments in activities. Maximizing healthy time goes by maximizing the probability of being healthy (π_a) in the following period.[5] Contrary to the traditional health production models, I will emphasize the uncertainty of health investments.[6] That means that an individual has a limited control over his health status. As return on investments is uncertain, one has to assess activities on a probabilistic ground at the beginning of each period. The uncertainty of the input variables is expressed by the distribution function ($\tilde{\pi}_a$). The underlying assumption is that experienced knowledge exists about the health effects of activities. This knowledge is not only a function of own experiences, but also of the social context people live within.

For complicated illnesses, we seek professional advices.[7] Doctors make diagnoses and express themselves in fuzzy probabilistic terms, such as "the probability to be perfectly cured after a hip fracture is about 80 percent after a year". For medical care investments we can express the consequences of the health investments as a function of mean and variance of healthy time. In short, we have an intuition about how an activity may affect our health but some uncertainty exists which is expressed in terms of deviation from our experienced knowledge.

The increased probability of being healthy yielded by activity *a* will be noted c_a and is accounted as $c_a = \pi_a - \pi$, where π is the probability without taking any activity and π_a is the probability after taken activity

[5] Normally in the literature of health economics the "probability of becoming ill" is considered instead of the "probability of being healthy". It is, however, convenient to illustrate positive rather than negative values in a $\mu - \sigma$-space, and I therefore choose to use "probability of being healthy".

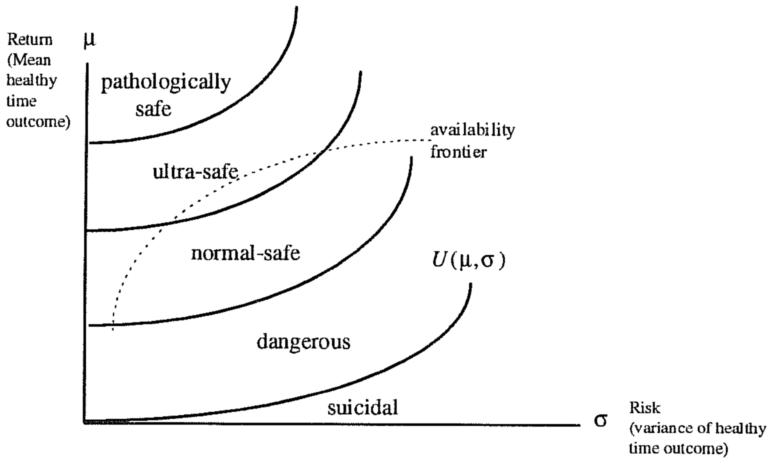
[6] Grossman (1972, p 19-21) is aware of the necessity to bring in uncertainty in his model. He therefore suggests to change the character of the depreciation rate from a deterministic to a probabilistic one. In a critique on Grossman, Dowie (1975) means, however, that the "...prime uncertainty faced by health producers is the impact of their investment inputs on their health stock" (p 634). A probabilistic depreciation rate, which Grossman suggests, will not lead to much better understanding of uncertainty in health production.

[7] The importance of self-care shall, however, not be neglected. Dean et al. (1983) estimated that the proportion of illnesses cared for by lay people without help from professionals is 66 to 75 percent.

a.[8] A unit of activity a has a positive impact on the probability of being healthy in the following period if the marginal probability is greater than zero, that is $\partial\pi/\partial a > 0$.[9]

In an early attempt to construct a health portfolio, Dowie (1975) divides activities in five categories: suicidal, dangerous, normal-safe, ultra-safe and pathologically safe. See Figure 4.1.

Figure 4.1[10]



In figure 4.1, each activity group is plotted according to their mean and standard deviation of probability of being healthy in next period. The important insight is that activities can be graded according to their return *and* their risk. Healthy time available after performing activity a is thus uncertain ($\tilde{c}_a = \tilde{\pi}_a - \pi$), and the standard deviation of healthy time is called "risk" in the model. This risk-feature of health investments has to a large extent been neglected in the health production literature.

---

[8] The change of probability measures the likelihood of increasing the probability of being healthy. If an activity has a positive influence on health, then $c_a > 0$, and $c_a \leq 0$ otherwise.

[9] How the functional relationship between a and π_a appears cannot be stated generally. Usually activities are good to some degree and for investments above a certain limit the marginal yield is negative. If $\pi_a = \pi(a)$ and γ is equal optimal investment, then $\pi'(a) > 0$ for $a < \gamma$, and $\pi'(a) < 0$ for $a > \gamma$. Some activities, however, do not have a positive influence even in small amounts, such as the extreme activity of suicide.

[10] From Dowie (1975, p 629).

Dowie claims that there is a high demand for high-return, low risk activities, as it is shown by indifference curves in the north-west direction in Figure 4.1. It is, however, technologically difficult to meet this demand at any relevant price. Therefore an availability boundary exists - equivalent to a budget restriction - which reflects both technological possibilities and economic restrictions. This gives rise to

> *Conclusion 1*: People demand health by means of investments in activities. The outcome of these activities can be characterized in terms of the mean and variance of healthy time. This represents labelling activities in an indifference map.

4.2.2 *Diversification of activities*

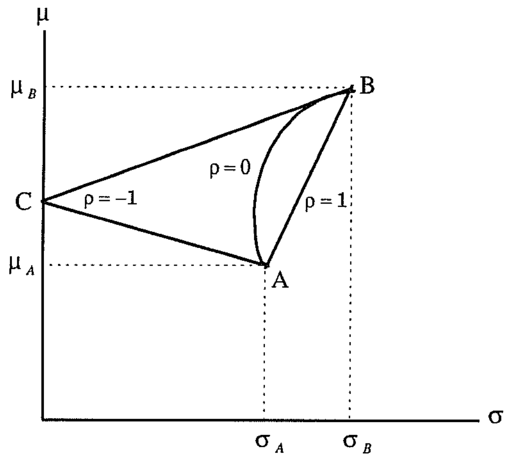
Performing a mixture of activities tends to reduce portfolio standard deviation. If an individual engages in a number of activities, he can diversify away some of the uncertainty surrounding the health effect of the input variables. Decisive to perform an activity is how the activity varies with other activities. One cannot forecast the health effect from one single activity. What counts is our health portfolio, our lifestyle. And the power of diversification is a function of the size and sign of the correlation coefficient between the activity return distributions. Before we continue with a portfolio composition, let us more closely consider how diversification works in health investments.

Let points *A* and *B* in Figure 4.2 below represent the characteristics of two activities *a* and *b*. Consider first the limiting case of a perfect positive correlation ($\rho = 1$). Here the combinations associated with mixtures of activities *a* and *b* would be proportionate to the relative budget shares.[11] That means that the portfolio outcome would lie along the straight line *AB*. Let us reflect on the health effects of vitamin C. We know that vitamin C is effective to prevent colds. To ensure the right doses of vitamin C, one can, for example, eat vitamin C capsules or eat oranges. The effect of the two activities is quite the same, that is they are highly correlated. In this respect, vitamin C capsules and oranges are complements.

---

[11] For a proof, see appendix equation (A4).

Figure 4.2



If activities are uncorrelated ($\rho = 0$), the attainable portfolio outcome would fall on the line represented by the bowed curve connecting *AB* in Figure 4.2. As the middle bowed curve has been drawn, the slope is negative in the region of point *A*. Apparently, a portfolio must exist with a mixture of *a* and *b* having higher return and lower standard deviation than the single-activity portfolio of activity *a*.[12] This case could be illustrated by a serious disease for which the medical profession does not yet have a suitable treatment program, such as medication against AIDS. In absence of efficient medicine, well-proven therapies developed for similar disease pattern are used instead. From time to time new risky medical technologies will be tested. A patient usually has to give his consent to *one* of the available treatment programs. According to portfolio theory such a choice is inefficient. Provided it is possible to combine parts of safe and risky technologies, a better outcome will be received if a mixed strategy is chosen. For instance, select a new risky technology as the main treatment program and hedge the risk by experiences from a well-proven technology.

Considering the limiting case of perfect negative correlation, the risk-reducing effect of diversification is so great that the curve *AB* breaks into two lines meeting at point *C*. With an optimal combination of activity *a* and *b* the standard deviation of the portfolio converges to zero.[13] The portfolio at point *C* yields, thus, a riskfree return by combining an

[12] For a proof, see appendix equation (A5).

[13] For a proof, see appendix.

appropriate mixture of the two activities. A general example is of course the optimal trade-off between prevention and cure. The appeal of prevention is straightforward. It is better not to suffer from a disease than to have it and try to repair the damage afterwards. Illustrative examples are vaccinations against smallpox or measles.[14]

Summing up the discussion about diversification, we agree about

> *Conclusion* 2: In order to reduce the risk associated with activities, consumers will seek to hold diversified portfolios of competing activities. Optimal health investments are thereby best seen as bundles of activities which aim to prevent illnesses and/or to restore states of health.

4.3 A portfolio model of health investments

4.3.1 A health portfolio consisting of two activities

We can now examine how an individual should compose a health portfolio of different activities to construct an efficient lifestyle. Let μ_a represent the mean change of probability $\tilde{c}_a$ per unit of activity a, and let σ_a denote the standard deviation of $\tilde{c}_a$.[15] Hence, the return from a single-activity portfolio consisting of q_a units of activity a amounts to $\mu = q_a \cdot \mu_a$.

Let us now consider the simplest case with only two activities. The first activity considered is medical care (a_1). Medical care is assumed to be a safe activity. Even though medical technologies are set under scientific observations, the exactness of their promises are uncertain to some degree. It is therefore possible to combine medical care with another activities to reduce overall risk exposure. Consider any of those activities

---

[14] Another highly effective investment is to use a condom to prevent sexually transmitted diseases. People with risky sexual behavior can strongly reduce the probability of developing the human immunodeficiency virus (HIV) to AIDS if they practice safe sex. One of the keystones of AIDS prevention is therefore to subsidize testing for HIV and to inform about methods for safe sex. Hence, diversifying highly dangerous sexual behavior with safe sex can reduce the private risk of engaging in "the market for sexual trades".Philipson and Posner (1995, p 446) argues "...that the market for sexual trades has many features similar to other economic markets and that the problem of AIDS can be viewed as a problem of market participants' uncertainty concerning the quality of the services exchanged." Thus, even here a risk management approach is appropriate.

[15] $\mu_a = E(\tilde{c}_a)$ and $\sigma_a = \left[E(\tilde{c}_a - \mu_a)^2\right]^{1/2}$.

available, say activity 2, for which riskiness is expressed as $\sigma_2 > \sigma_1 > 0$. The outcome - measured as the total probability change of the portfolio investments - is now a random variable given by:

$$\tilde{c} = q_1 \cdot \tilde{c}_1 + q_2 \cdot \tilde{c}_2 . \tag{1}$$

Let us assume that the structure of the portfolio is fixed for a period of time. First at the end of the period new information is at hand and the agent can reconsider his investments and choose new investment strategies. If the individual's disposable income for health investments is W and each unit of activity costs P_a, then the budget constraint can be written as:

$$P_1 \cdot q_1 + P_2 \cdot q_2 = W . \tag{2}$$

For any activity a, we can define its rate of return. If $\tilde{R}_a = \tilde{c}_a / P_a$, then one dollar invested in activity a today will yield a change of the probability of being healthy of $\tilde{R}_a$ in the next period. Using the budget identity and the notations for return on medical care (activity 1), we can rewrite expression (1) as:

$$\tilde{c} = (W - P_2 \cdot q_2) \cdot \tilde{R}_1 + q_2 \cdot \tilde{c}_2 . \tag{3}$$

Taking the expectation of (3), the parameters for the probability distribution of the portfolio become:

$$E(\tilde{c}) = \mu = W \cdot \tilde{R}_1 + \left(\mu_2 - \tilde{R}_1 \cdot P_2\right) \cdot q_2 , \tag{4}$$

$$\sigma = \left(\sigma_1^2 \cdot q_1^2 + 2 q_1 \cdot q_2 \sigma_{12} + q_2^2 \cdot \sigma_2^2\right)^{1/2} . \tag{5}$$

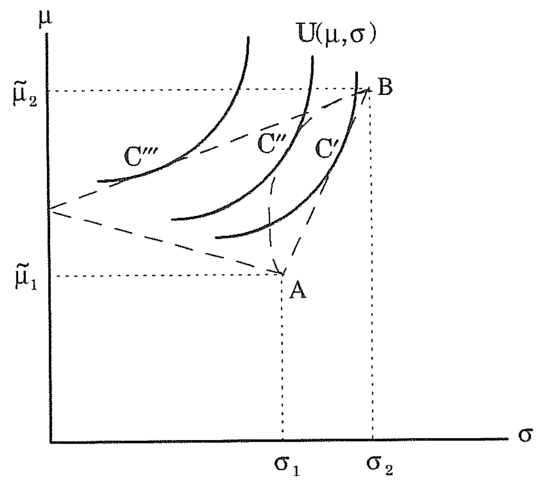
where μ equals the expected return of the portfolio investments and σ is equal to the standard deviation of the portfolio.[16] The valuation of the parameters in equations (4) and (5) rests upon professional as well as layman's wisdom. As medical care is traded at the health care market, diagnoses and treatments are sold to the consumers. The information advantage of physicians vis-à-vis consumers put consumers in an unfavorable position. Consumers have to make decisions on the information they receive from physicians. They have therefore limited possibilities to check the validity of the information by their own experiences.

[16] $\sigma_{12} = E(\tilde{c}_1 - \mu_1) \cdot (\tilde{c}_2 - \mu_2)$ is the covariance of $\tilde{c}_1$ and $\tilde{c}_2$.

From a health investor's point of view, the size of the parameters $\tilde{R}_1$ and σ_1 are exogenously determined. Contrary to medical care, activity 2 is subjectively valued.

Without considering the covariation of activities in health investments, gains of diversification will be lost. A person who only has confidence in the medical profession, will hold a single-activity portfolio consisting of medical care exclusively. In Figure 4.3, this position is denoted A and the return takes the value $\tilde{\mu}_1$. A risk-loving person would on the other hand not invest in medical care at all, but rather rely on the more risky activity which has a higher expected return per monetary unit. This strategy is point B in Figure 4.3.

Figure 4.3



In daily life none of these extreme positions seem familiar. We engage in activities whose consequences we can not foresee and we rely on medical doctors and their diagnoses. According to the portfolio theory this is exactly what we shall do. As we saw in section 4.2.2, diversification reduces variability in healthy time and we should therefore combine activities in a portfolio along the budget line AB.

Let us assume a utility function that is a function of the mean and standard deviation of healthy time: $U = U(\mu,\sigma)$. To ensure quasi concavity of $U(\mu,\sigma)$, we assume that the choice of activities is constrained to selecting a distribution from a linear class (Meyer 1987).[17] An individuals maximizes healthy time under the constraint that he cannot invest in more activities than the budget allows for. In optimum the marginal value of a risk reducing activity must equal the marginal cost of its performance.

The optimal solution will strongly be affected by the covariance term in (5). In Figure 4.3, we can see three optimal solutions, depending on the size and sign of the covariance term. The tangential points lie at higher indifference curves the more uncorrelated, or even better, the more negative correlated the activities are, that is $C''' > C'' > C'$. To maximize healthy time, an investor should search for activities that are slightly positively or highly negatively correlated with each other. The imperative from the portfolio theory on health investments is straightforward. A rational health investor should diversify and construct a portfolio of activities which yields the highest return of healthy time for a given risk! Let us summarize these findings in

> *Conclusion 3*: As perception and reception of activities vary among people, also healthy lifestyles are different. In combining medical care with a risky activity, a rational agent constructs a health portfolio which yields the highest return of healthy time for a given risk.

4.3.2 *A health portfolio with many activities*

Generalizing the two-activity choice to any number of activities, we get a multi-activity portfolio. As before, let μ_a represent the mean change of probability $\tilde{c}_a$ per unit of activity a. Similarly, let σ_a denote the standard deviation of $\tilde{c}_a$ and σ_{ab} the covariance of $\tilde{c}_a$ and $\tilde{c}_b$. If the individual invests his time and money in a portfolio consisting of q_a units each of activities $a = 1,\ldots,A$, the return of his "lifestyle" is related to the activity return parameters via:

[17] $U_\mu \geq 0$ and $U_\sigma \leq 0$ for all μ and all $\sigma \geq 0$, and $U_{\mu\mu} \leq 0$, $U_{\sigma\sigma} \leq 0$, and $U_{\mu\mu}U_{\sigma\sigma} - U^2_{\mu\sigma} \geq 0$ for all μ and all $\sigma \geq 0$. The proofs for quasi concavity can be found in Meyer (1987, pp. 424-425).

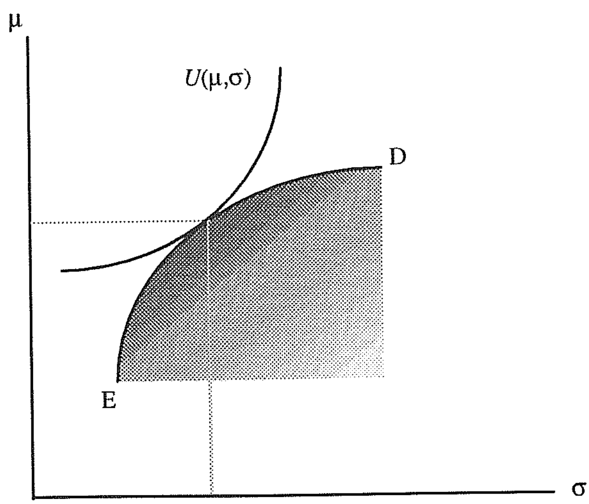
$$\mu = \sum_{a=1}^{A} q_a \mu_a \tag{6}$$

$$\sigma = \left(\sum_{a=1}^{A} \sum_{b=1}^{B} q_a \sigma_{ab} q_b \right)^{1/2} = \left(\sum_{a=1}^{A} \sum_{b=1}^{B} \sigma_a q_a \rho_{ab} \sigma_b q_b \right)^{1/2}, \tag{7}$$

where ρ_{ab} is the correlation coefficient between the distributions of activities $\tilde{c}_a$ and $\tilde{c}_b$.

In Figure 4.4 below, the multi-activity choice is shown graphically. The figure shows how the choice is enlarged when an investor has a larger selection of activities. By mixing activities in different proportions, he can reduce the risk and obtain a more favorable outcome. The expected returns and standard deviations that can be generated by some sets of risky activities are the opportunity sets in the shaded area. As long as there is less than perfect positive covariance among the activities many portfolios will have probability returns superior to a single activity. Efficient risk management is characterized by those diversified portfolios that dominate all other portfolios and single activity performances. Those sets that are optimally diversified lie on the efficient boundary *DE*. Owing to the power of diversification, almost all portfolios on the boundary *DE* would most likely represent multi-activity lifestyles.

Figure 4.4



Since an investor wishes to increase expected healthy time and reduce standard deviation, he will be interested only in lifestyles that lie along *DE*. These activity sets are efficient lifestyles. Whether individuals want to choose minimum risk portfolios, lifestyles close to E, or some other riskier portfolio in the neighborhood to D, depend on individuals' attitude towards risk.

Recent studies have pointed out the importance of lifestyles when health investments are considered. With data from the 1985 Health Interview Survey, Kenkel (1995) points out the importance of the chosen lifestyle for the health status. Excessive weight, cigarette smoking, heavy drinking, excessive or insufficient sleep and stress are harmful inputs in the investigated health production functions. Beneficial inputs are, for example, exercise and moderate alcohol consumption. In another study about demand for cigarettes, Hu et al. (1995) conclude that smoking behavior occurs in the context of an overall health behavior pattern. Individuals who smoke exhibit a higher level of other unhealthy behaviors. Certain high-risk behaviors, such as acute and chronic drinking and physical inactivity, tend to cluster with cigarette smoking. Hu et al. try to explain this observation as a consequence from the covariance of smoking with other unhealthy behaviors, such as low educational skill, stress, and preferences for risk-taking.

In this empirical context a risk management approach gives normative guidelines. The important insight from the portfolio approach on health investments is that diversifying activities tends to reduce the variability of healthy time. To reach the efficient boundary we should diversify our investments in a multi-activity portfolio. Treatment, prevention and health promotion ought to be characterized in terms of bundles of activities and not as single activities. However, in order to explain some behaviors, we do not have to assume risk-loving preferences as Hu et al. do. The portfolio theory may disclose "irrational" behaviors as "rational". I pointed out the two-fold meaning of health, as an investment good as well as a consumption good. The well-being effect of, for example, smoking may induce an investor to continue smoking, even if the biological effects of smoking are dangerous! These findings may be summarized in

> *Conclusion 4*: Efficient lifestyles are those diversified portfolios that dominate all other single- and multi-activity behaviors. Owing to the power of diversification, almost all efficient lifestyles represent multi-activity behaviors. Decisive to undertake an activity is how the activity covaries with other activities.

4.4 Concluding remarks

Contrary to traditional production models of health which describe how people *do* invest in health, the portfolio theory is strongly normative. A risk management approach raises imperatives how people *should* invest in health. I think that the portfolio theory can shed some light on particular problems which have been neglected in the traditional framework of consumer theory. Moreover, the portfolio theory on health investments raises new insights in rational behavior. Some problems must, however, be admitted.

Diversification presupposes knowledge about covariance between activities. The question arises if it is possible to find out how, for example, jogging is correlated to eating behavior or the correlation between medical care and psychotherapy. Data sets on cancer and smoking habits can surely be incorporated in a portfolio framework. But for more qualitative variables, such as health promoting and health education, adequate proxies must be found. This task is not easy. Nevertheless, the question Markowitz asked in his pioneering article from 1952 concerning financial securities - "Perhaps there are ways, by combining statistical techniques and the judgment of experts, to form reasonable probability beliefs (μ_i, σ_{ij})"[18] - may also find its answer concerning health behavior as it has received in the vast literature on financial theory.

The portfolio theory is a new way of thinking about health behavior. If the underlying condition holds, that activity outcomes can be graded according to their mean and variance of healthy time, then a portfolio approach gives guidance for an optimal health behavior. The portfolio thinking is valid even if we cannot measure qualitative variables, because prior investments have given us and our environment (e.g. the family) experiences about how different kinds of activities affect our health.

However, as the portfolio approach at the present time is highly prescriptive, let me finally conclude the analysis on health behavior in some normative statements that highlight the portfolio-thinking approach to health investments:

[18] Markowitz (1952, p 82).

Conclusion 5:

i) By a risk management approach to health investments, prevention and cure are characterized in terms of bundles of activities.
ii) To reduce the risk exposure of health investments, people should diversify their activities in a health portfolio,
iii) this means that they should construct a lifestyle which yields the highest return of healthy time for a given risk.
iv) As perception and reception of activities vary among people, healthy lifestyles are also different.
v) Health policies, like health education, health regulation and health promotion, should rather be comprehensive and focus on the whole individual, not on a single disease.

Appendix

Diversification of two risky activities[19]

Let μ_a represent the mean change of probability $\tilde{c}_a$ per unit of activity a, and let σ_a denote the standard deviation of $\tilde{c}_a$. Assume two arbitrary risky activities, a and b, with the characteristics of $(\mu_a < \mu_b)$, $(0 < \sigma_a < \sigma_b)$ and $(\rho_{ab} = \sigma_{ab}/\sigma_a \cdot \sigma_b)$. If the individual's disposable income for health investments is W and if each unit of activity costs P_a, then the budget constraint can be written as:

$$P_1 \cdot q_1 + P_2 \cdot q_2 = W. \tag{A1}$$

In Figure 4.2, the single-activity portfolios consisting of activity a and b are represented by points A and B, respectively. If the respective budget shares are $\alpha = q_a \cdot P_a/W$ and $(1-\alpha) = q_b \cdot P_b/W$ and the mean yields on each single-activity portfolio are $\mu_A = W \cdot \mu_a/P_a$ and $\mu_B = W \cdot \mu_b/P_b$, then the portfolio return is:

$$\mu = \alpha \cdot \mu_A + (1-\alpha) \cdot \mu_B, \tag{A2}$$

and the standard deviation of the two-activity portfolio becomes:

[19] Diversification of two healthy activities is a straightforward extension of diversification of financial assets (see for example Sharpe (1964)).

$$\sigma = \sqrt{(\alpha\sigma_A)^2 + ((1-\alpha)\sigma_B)^2 + 2\rho\alpha(1-\alpha)\sigma_A\sigma_B} \,. \qquad \text{(A3)}$$

If perfect positive correlation ($\rho = 1$) exists, the expression under the square root sign in (A3) can be factored:

$$\sigma = \sqrt{[\alpha\sigma_A + (1-\alpha)\sigma_B]^2} = \alpha\sigma_A + (1-\alpha)\sigma_B \qquad \text{(A4)}$$

Hence, the portfolio outcome will be linearly related to the proportions invested. Graphically this result implies that the combination will lie along the straight line *AB*. A budget share for activity *a* of $\alpha = 0.5$ results in a $\mu - \sigma$-combination in the middle of line *AB*.

If activities are uncorrelated ($\rho = 0$), the attainable portfolio outcome will fall on the line represented by the bowed curve connecting *AB* in Figure 4.2. As the middle curve has been drawn, the slope $d\mu/d\sigma$ is negative in the region of point *A*. Apparently, a portfolio must exist with $\alpha > 0$ having a higher return and a lower standard deviation than the single-activity portfolio of activity *a*. To prove this, we consider the case when α increases. The slope connecting *A* and *B* can be written:

$$\frac{d\mu}{d\sigma} = \frac{d\mu/d\alpha}{d\sigma/d\alpha} = \frac{(\mu_A - \mu_B)\cdot\sigma}{\left[\alpha\sigma_A^2 + (\alpha - 1)\sigma_B^2 + (1-2\alpha)\sigma_A\rho\sigma_B\right]} \,. \qquad \text{(A5)}$$

At point *A* $\alpha = 0$. If $\rho \leq 0$, then the denominator will be negative at point *A*, the implication being that the efficient boundary does not include point *A* itself.

In case of perfect negative correlation, the risk-reducing effect of diversification is so great that the curve *AB* breaks into two lines meeting at point C. As $\rho = -1$, the standard deviation of the portfolio becomes equal to $\sigma = \alpha\sigma_A - (1-\alpha)\sigma_B$. Setting $\alpha = \sigma_B/(\sigma_A + \sigma_B)$, the standard deviation of the portfolio converges to zero. The portfolio at point C in Figure 4.2 yields, thus, a riskfree return by combining an appropriate mixture of the two activities.

5. A proposal for a reorganization of health insurance in presence of asymmetric information*

5.1 Introduction

Health insurance raises a number of interesting problems, practically as well as theoretically. Most economists interpret the practical problems, at least in part, due to moral hazard and adverse selection effects. Politically there are disputes about whether governmental interventions would lead to a more efficient allocation or to increased fairness. Those economists who are in favor of governmental interventions are usually not very clear about how the interventions should appear. However, many of them argue for a mixed system in which the government should cover basic health expenses whereas health care above a certain coverage should either be taken care of privately, on an out-of-pocket basis or by private health insurance. Contrary to this, this paper argues for a mixed system in which the government covers for large expenditures, whereas all expenses up to a certain limit are to be taken care of privately.

The seminal contributions to the adverse selection literature, such as Akerlof (1970) and Rothschild and Stiglitz (1976) are based on simple loss processes and insurance contracts. The policyholders typically meet the same uncertain loss and differ only in their loss probabilities. By doing so, they overlook the importance of distributional risk in insurance pricing.[1] Most risk distributions are skewed and extremely heavy-tailed. The distribution of medical costs is no exception.

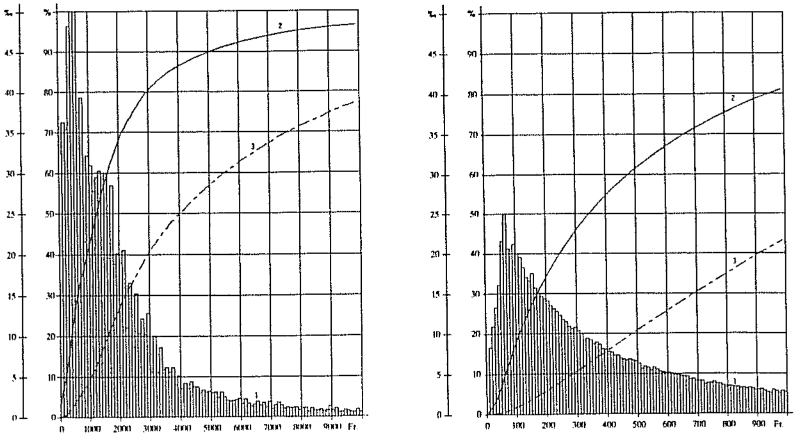
Data on health care spending from the United States indicates that about 10 per cent of eligibles account for over 70 per cent of total spending (Riley et al. 1986). As can be seen from the Swiss data in Figure 1, the main part of the medical cost utilization occurs in a narrow range in the left tail of the costs distributions. The greatest part of the population consumes limited medical services or nothing at all. The skewed

* The paper was co-authored by Professor Franz Haslinger, University of Hannover and it was presented at the 22nd Seminar of the European Group of Risk and Insurance Economists, Geneva, Switzerland, September 18-20, 1995.

[1] See the discussion in Berger and Cummins (1992, p. 273)

medical cost distribution illuminates the heterogeneity and variability of health risks. This brings about two difficulties: First, how can the market discriminate between policyholders with different risks and second, how can the insurer set accurate premium prices if an average cost varies?

Figure 5.1 **Medical cost distribution per capita, stationary and ambulatory care, Switzerland 1983**[2]



stationary care ambulatory care

This paper focuses on the possibilities for the state to intervene in the health insurance market to reduce the negative effects of adverse selection and at the same time reduce the variance in the insurance pool. The point is that if the state insures for high medical costs, the rest of the risk distribution gets closer in shape to a normal distribution. The "normalization" transforms the originally heterogeneous risks to more homogeneous ones, thereby reducing the adverse selection effects. Additionally, public coverage for severe illness interacts with private coverage for less severe illness as high-cost protection reduces the variance of an average health risk on the private market. Important findings indicate that high-cost protection reduces safety loading and improves market efficiency. The market for health insurance will be characterized as a mixed system, with a governmental high-cost insurer and many private low-cost insurers.

[2] Schulenburg (1987, p. 32)

5.2 A model for a health insurance market

5.2.1 Model specification

Let us consider a person who is exposed to health risks such as diseases or injuries. We neglect the intangible effects of the health risks and consider only the financial consequences.[3] The financial outcomes are a function of medical utilization and vary in a period of time as well in incidence as in intensity. Let y^0 denote the individual income without any insurance and let X denote the medical cost distribution, where X is limited to an upper level, i.e. $X \in \{0, X^1, ..., X^m\}$. An individual ($i$) faces a prospect of becoming ill according to a certain probability distribution $\pi_i \in \{\pi_i^0, ..., \pi_i^m\}$, and without any insurance the individual has to cover medical expenses out of his own pocket. In case of illness, the available income amounts to $Y = y^0 - X$, and to $Y = y^0$, otherwise. Since a risk-averse individual would rather prefer to pay a fixed premium and receive a certain income than face the stochastic prospect of paying the medical bill on an out-of-pocket basis, he demands a health insurance. The available income with an insurance protection can be written as

$$Y = y^0 - P - X + I, \tag{1}$$

where P is the premium and I the insurance coverage. The insurance coverage varies proportionally with the cost distribution; $I = \alpha X$, where $0 \leq \alpha \leq 1$, and the premium is based on the expected medical costs. Considering a discrete distribution the expectation of X is defined as

$$E(X) = \sum_{j=0}^{m} \pi_i^j x^j. \tag{2}$$

Because of the irregularity of payments there is likely to be a cost of capital tied up. Each insurance policy includes therefore a buffer fund, a so-called *safety loading*. The safety loading is assumed to reflect the distribution risk of the skewed medical cost distribution. Even if the average medical cost in an insurance pool normally does not vary to any appreciable extent, distortions in the form of unforeseeably costly

[3] Schulenburg (1987) notices that if the intangible and financial effects are highly positively correlated, a risk-adverse individual will demand excessive insurance coverage. If the financial and intangible risk could be separated in time, the correlation would be reduced and also the insurance demand. Zweifel (1992) shows how this is possible with a bonus system.

treatments can cause rather large deviations. To treat an AIDS-patient is extremely expensive and even a few policyholders who are HIV-positive can put the solvency of the insurer heavily under pressure.[4]

Borch (1985) also considers a multiperiod decision-making process. A decision-maker who maximizes future cash flows knows that after a bankruptcy all future cash flows are forfeited. He will therefore internalize the costs of bankruptcy in the premium setting. Hence, the premiums will not only be calculated as expected loss but also reflect the contribution of policies to the risk of ruin. In summary, for some reasons mentioned above an insurer is in need of a safety buffer fund. The premium formula is therefore set as:

$$P = (1+\lambda)\alpha E(X), \tag{3}$$

where λ equals the safety loading per policy. We assume a perfectly competitive market without administration costs and no social risk[5]. The safety loading is the only loading factor considered.

5.2.2 *A two-moment health insurance model*

In the insurance literature the expected utility model has frequently been used to determine the optimal coverage rate.[6] And concerning adverse selection, authors have used the same approach to explain why a market with asymmetric information fails to supply optimal insurance.[7] Instead of using an expected utility model, we use a two-moment decision

---

[4] Focusing on a hospital level, cost of treatments can vary for different reasons. Siegel et al. (1992) say that the mainly reasons for variations in treatment intensities are due to economics of scale and scope of the providers, level of personal skills and age of capital equipment.

[5] One ground for incomplete diversification is what Hirshleifer and Riley (1992) call *social risk*. "Social risk comes about whenever private risks are not perfectly offsetting" (p. 129) and "...provides two reasons why insurance prices may not be fair or actuarial: (i) if the number of risks in the insurance pool is small, the Law of Large Numbers cannot work very fully; (ii) if the separate risks are on average positively correlated, then even with large numbers the variance of the per capita return does tend to diminish but does not approach zero" (p. 131). Concerning the problem with dependent risks, Horgby and Söderström (1995) suggest a financial market approach, diversifying insurance contracts at the financial market instead of in an insurance pool.

[6] See for example Borch (1961, 1962), Arrow (1963), Mossin (1968), Smith (1968), Shavell (1979).

[7] Important contributions have been made by Pauly (1974), Rothschild and Stiglitz (1976), Wilson (1977), Dionne (1983), Crocker and Snow (1986).

model to analyze a health insurance market with imperfect information. The mean-standard deviation approach does not lead to any inconsistencies if the choice set is composed of random variables, and the variables only differ from one another by location and scale parameters.[8] An agent's expected utility ranking of the elements of the choice set could therefore be represented by a mean-standard ordering.[9] An individual's expected income with insurance protection amounts to

$$E(Y) = \mu = y^0 - P - E(X) + I = y^0 - (1+\alpha\lambda)E(X), \tag{4}$$

and the standard deviation of income is

$$\sigma(Y) = (1-\alpha)\sigma(X), \tag{5}$$

where $\sigma(X)$ is the square root of the second moment of $[X - E(X)]$. The utility of an individual in the two-moment decision model is a function of the expectation and the standard deviation of income. The optimization problem is to

$$\text{max!} \quad EU[Y,\sigma(Y)] = V(\mu,\sigma), \tag{6}$$

$$\text{s.t.} \quad E(Y) = \mu = y^0 - E(X) - \lambda\left(\frac{\sigma(X) - \sigma(Y)}{\sigma(X)}\right)E(X).^{10} \tag{7}$$

The mean income in our context is now a function of the expected medical costs and the standard deviation of income and medical costs, respectively. It is therefore possible to analyze different combinations of insurance coverage in a mean-standard deviation space (see Figure 5.2 below).

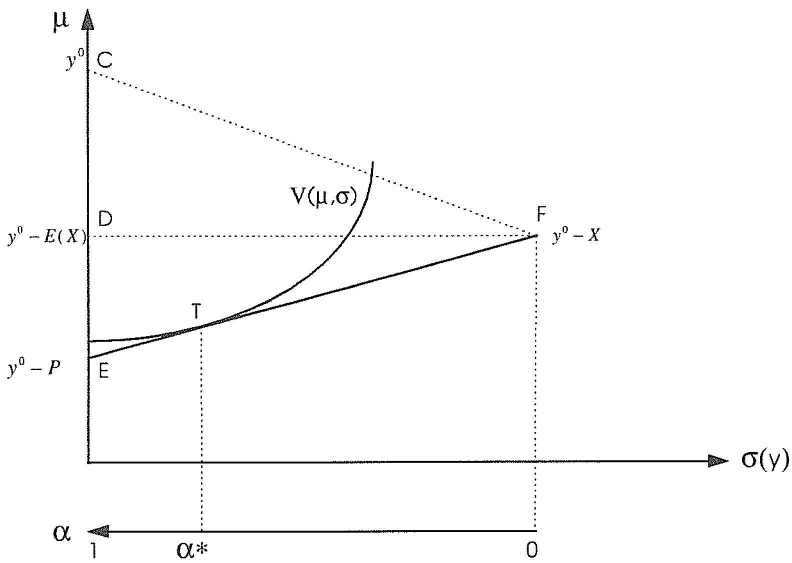
---

[8] The location and scale parameter condition is further explained in the Appendix.

[9] Sinn (1980) and Meyer (1987) prove the consistency of a mean-standard approach with expected utility theory, and Schlesinger (1994) gives some insurance applications.

[10] The budget constraint in (7) is compounded by equations (4) and (5).

Figure 5.2 The mean-standard deviation model of health insurance



The slope of the budget constraint (ray *EF*) in Figure 5.2 is positive as long as the insurer claims a higher price per insurance unit than its actuarial value, i.e. as long as $\lambda > 0$. Increasing the coverage rate α, a move to the left implies therefore not only a decreasing standard deviation but also a decreasing final expected income distribution. Without any insurance the individual faces the prospect of $(y^0 - X)$, which is illustrated as point *F*. Along ray *EF* he can choose among combinations of insurance coverage to corresponding premium prices. At point *E*, the individual is fully insured $(\alpha = 1)$, pays a premium $P = (1+\lambda)E(X)$ and receives a certain state independent income of $y = y^0 - P$. On all other points along *EF*, the individual does not pay the entire expected value, but rather $\alpha E(X)$, and bears some risk of his own as is shown by the corresponding standard deviation of income. The optimal final income distribution is received if (6) is maximized subject to (7). Analytically the solution is

$$\frac{V'_{\mu}}{V'_{\sigma}} = \frac{\sigma(X)}{\lambda E(X)}, \tag{8}$$

which is illustrated as the tangential point *T*. The optimal choice in the presence of a safety loading is thus characterized by an incomplete insurance coverage $(\alpha^* < 1)$. Consistent with the classical findings by

Arrow (1963) and Mossin (1968), a rational agent will not fully insure if the premium price is not actuarially fair. Only if the premium price is actuarially fair will a rational agent demand a complete insurance. In that case, the budget constraint would coincide with the line *DF* and the agent would fully insure at point *D*.[11]

5.3 Heterogeneous risks on the health insurance market

Now we consider a large group of individuals who are identical in all respects, except for the probability of becoming ill. For simplicity, we suppose that there are two kinds of individuals, one group represents good risks - A-individuals - and the other represents bad risks - B-individuals. Hence, the probability distribution of utilization of medical treatments is greater for the B-group than for the A-group, i.e. $\pi_A < \pi_B$. Accordingly, both the expected medical costs, $E(X_A) < E(X_B)$, as well as the standard deviation, $\sigma_{X_A} < \sigma_{X_B}$, are greater for the B-group. To determine whether an allocation is in equilibrium, the Rothschild and Stiglitz (1976) equilibrium conditions are used.[12]

First, the case is analyzed when information about individual health risks is known to all agents in the market (public information) and second, a more realistic scenario is taken into consideration, when only the policyholders know their own health risks, and the health profile cannot be verified but for very high diagnostic expenditure (private information).

---

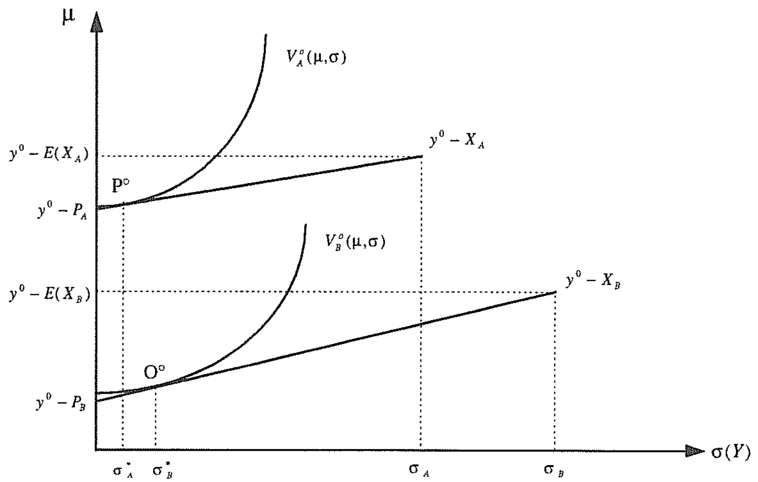
[11] Schlesinger (1994) uses a new theory of risk aversion developed by Segal and Spivak (1990) who make a difference between risk aversion of the first and second order. In case of risk aversion of first order, the indifference curves have a positive slope at $\sigma = 0$. This means that even if a rational actor faces a loaded premium, the optimal choice can be a complete insurance coverage $(\alpha = 1)$. If ambiguity regarding the probability of a particular event occurring and/or uncertainty about the magnitude of the resulting loss exist, then newly results claim that it is possible to charge higher premium prices than its expected value (Kunreuther et al. 1995). These extensions are not further considered here.

[12] "Equilibrium in a competitive insurance market is a set of contracts such that, when customers choose contracts to maximize expected utility, (i) no contract in the equilibrium set makes negative expected profits; and (ii) there is no contract outside the equilibrium set that, if offered, will make a nonnegative profit." (Rothschild and Stiglitz 1976, p. 633)

5.3.1 *Heterogeneous health risks and public information*

With public information, each health profile can easily be verified. As the objective is to maximize each individual's utility subject to his budget constraint, the market will be divided into two distinct segments: one for good risks and one for bad risks. Each individual is assumed to maximize the objective function in (7) subject to the budget constraint in (8). But, as $E(X_A) < E(X_B)$ the premium is greater for the B-group and the expected mean-income will be lesser for B-individuals than A-individuals.

Figure 5.3 A heterogeneous health insurance market with public information



Let us consider this case with help of Figure 5.3. Without any insurance the individuals in each group face the prospect of becoming $(y^0 - X_A)$ and $(y^0 - X_B)$, respectively. If possible, B-individuals would like to buy the same contracts as the A-individuals. However, as the equilibrium conditions state that an equilibrium contract cannot make negative profits and the insurer can easily disclose the health profile, A-contracts cannot be sold to the B-individuals. Hence, A-individuals can only buy policies along their budget constraint $(y^0 - X_A) \rightarrow (y^0 - P_A)$, and B-individuals are restricted to choose policies along $(y^0 - X_B) \rightarrow (y^0 - P_B)$. Note that the budget line for B-individuals is steeper than the budget line for A-individuals. Intuitively this can be explained from the greater risk to insure B-individuals. As can be seen from equation (7) the impact of a given safety loading increases with the variance of an insurance contract.

The optimal insurance allocation is reached at the tangential point O^o for B-individuals and correspondingly at P^o for A-individuals. This allocation is consistent with the equilibrium conditions and is Pareto optimal, since it is not possible to find another allocation that is better for one group without making the other group of policyholders worse off. The optimal allocation with public information is thus characterized by a separating equilibrium where the bad risks pay a greater premium rate than the good risks.

5.3.2 *Heterogeneous health risks and private information*

We now assume that the health insurance market is characterized by private information, i.e. individual health risks are only known by the individuals themselves and cannot be revealed to another except at very high diagnostic costs.[13] To keep the track clear we assume that it is impossible for the insurers to discriminate at all. There is, however, knowledge about group health costs at an aggregated level.[14] Each insurance contract must therefore be supplied to the entire market, to both A- and B-individuals. Following the logic of Rothschild and Stiglitz the insurers offer a number of policies (1,2,...n) with different prices and quantities, so that individuals with different risks have the opportunity to choose different insurance policies.[15]

Even though the insurers are not able to discriminate the agents according to their health risks, a binding constraint on a competitive market is that in equilibrium only insurance policies along the budget

[13] This assumption is rather sharp, since often people do not know their own health risks. It is, however, plausible that each individual has superior knowledge about, i.e. genetic inheritance.

[14] We assume that the size of the parameters π_A and π_B are known to all, including the state.

[15] During a price and quantity competition, it is conceivable that insurance policies with different prices will exist in equilibrium. Those who want a more comprehensive insurance coverage may be willing to pay a higher price than those who are satisfied with a superficial coverage. This extension of offering many policies on the market was not necessary in the previous section, since the insurers had perfect information about expected health costs and could perfectly discriminate between A- and B-individuals. In contrast to price-quantity policies Pauly (1974) only discusses price-policies. As Pauly assumes that the insurers cannot observe the total amount of coverage purchased by a client, the insurers charge a uniform premium rate per unit of coverage to all policyholders.

lines will be issued. As noted in equation (3), the premium will equal expected medical costs plus safety loading. If a policy exists with a premium rate which does not fulfill equation (3), the insurer will make a profit/loss and the free entrance/exit mechanism will force the market to action. Equation (3) is, however, only a necessary condition. If, for example, an insurer issues optimal insurance for A-individuals (policies at point P^o) not only the target group will buy it, but also the B-individuals who have a higher expected consumption of health services. Such a contract will make a loss even though equation (3) is fulfilled and since the insurers cannot discriminate between the two groups, such a contract will not be supplied at the market.

Because of their bad risks the B-individuals will dominate the market outcome. They will maximize their utility at point O^o in Figure 5.3. If the insurer issues only P_B^*, it is possible that no one from the A-group will buy any insurance, since A-individuals have lower expected costs than the B-group. In that case the asymmetric information between policy-holders and insurers has destroyed the health insurance market for A-individuals. As Akerlof (1970) pointed out, the bad risks drive out the good risks from the market. Another outcome is, however, possible. Before we search for a second-best equilibrium, we change the independent variable $\sigma(Y)$ in Figure 5.3 to the coverage rate α. As the standard deviation of income varies proportionally to the coverage rate, $\sigma(Y)=(1-\alpha)\sigma(X)$, the $\mu-\sigma$-model can also be illustrated in a $\mu-\alpha$ space. Accordingly, the budget constraint takes the form of

$$\mu = y^0 - (1+\lambda)\alpha E(X) - E(X) + \alpha E(X) = y^0 - E(X) - \lambda\alpha E(X) \tag{10}$$

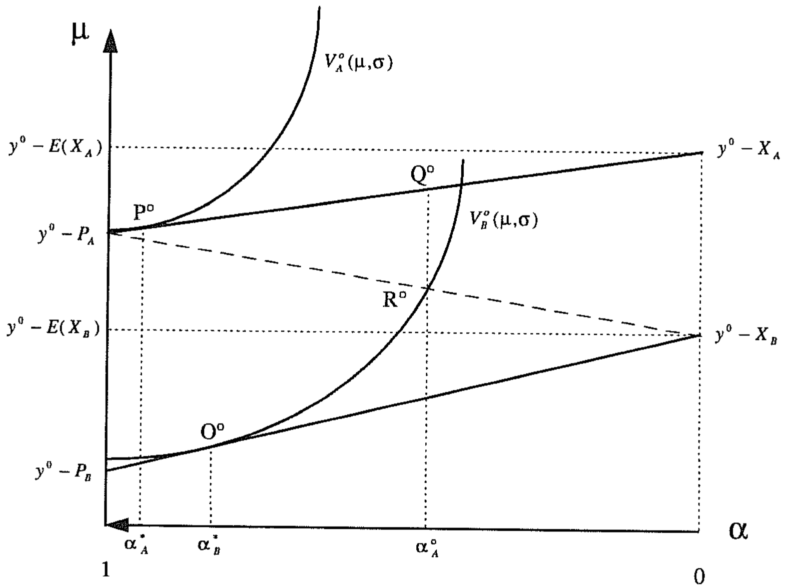
and graphically the relationship can be seen in Figure 5.4 below.

Because the market is characterized by a price and quantity competition, the insurers can issue insurance contracts at different premium prices, $(P_i^1, P_i^2, ..., P_i^n)$, resulting from different levels of coverage, $(\alpha_i^1, \alpha_i^2, ..., \alpha_i^n)$. An equilibrium is possible only if different contracts comply with the self-selection constraint that only high-risk B-individuals choose only high-risk contracts and only low-risk A-individuals choose only low-risk contracts. We have already mentioned that the B-group will dominate the market through choosing an optimal insurance coverage at point O^o. The contracts open to the A-individuals are constrained by the need to make B-individuals remain with coverage rate α_B^*. A sufficient condition for a market equilibrium is fulfilled when the utility for the B-individuals is the same regardless of which policy they choose, i.e.

$$V(\mu_B,\sigma_B|\alpha_B^*,P_B^*) \geq V(\mu_B,\sigma_B|\alpha_A,P_A). \tag{11}$$

A feasible equilibrium is where the indifference curve of the B-individuals cuts the long broken line $(y^0 - X_B) \rightarrow (y^0 - P_A)$ at point R^o in Figure 5.4.

Figure 5.4 A health insurance market with private information ($\mu - \alpha$ space)



At the intersection at point R^o the bad risks are indifferent between the optimal allocation O^o at price P_B^* and the allocation R^o at a lower price per insurance unit but with lesser coverage.[16] We assume that B-individuals buy the optimal policy P_B^*. Now A-individuals can buy policies with the coverage rate α_A^o and therefore reach an allocation of insurance contracts which is denoted Q^o.

At large, the allocation (O^o,Q^o) does not differ from the separating equilibrium by Rothschild and Stiglitz. As at equilibrium in the Rothschild and Stiglitz-model the premium rate will still be lower for the good risks, but they will not enjoy an optimal insurance protection

16 $V(\mu_B^*,\sigma_B^*|\alpha_B^*,P_B^*) \geq V(\mu_B,\sigma_B|\alpha_A^0,P_A^0)$

($\alpha_A^o < \alpha_A^*$). And the bad risks face a higher premium rate than the good risks but reach on the other hand their optimal coverage rate ($\alpha_B^* > \alpha_A^o$).

In summary, if the insurer could discriminate the policyholders according to their health risks, the market would clear at the Pareto optimal equilibrium in O^o and P^o in Figure 5.4. If information about health status is private, the insurer cannot discriminate the policyholders and the market clears at O^o and Q^o. In a private health insurance market with private information there will be a separating equilibrium, and the outcome is not Pareto-optimal.

5.4 High-cost insurance protection

The heterogeneity of risks together with asymmetric information creates adverse selection. To reduce the adverse selection as described above, many authors proclaim compulsory insurance.[17] Although a unanimous picture of the effects of the compulsory insurance is lacking, the authors agree that compulsory protection should not be complete but only partial.

> "If individuals cannot be discriminated according to their individual health risks, at most a separating equilibrium may exist on a private insurance market provided that the share of ´good risks´ is not too large. In this case a compulsory insurance *covering just a part of the costs* and charging uniform contribution leads to a Pareto-improvement" (Breyer and Zweifel 1995, p. 155, our italics).

Interestingly, Breyer and Zweifel, as most other authors, do not give any further details about what "partial" means in insurance terms, although partial insurance is a key result for a Pareto improvement. Later on Breyer and Zweifel also analyze moral hazard and state that

> "...a mandated unique insurance coverage cannot be welfare-maximizing. It seems to be a better idea to let the lawmakers determine a *minimal extent of insurance coverage* while leaving it up to the individual to extend his coverage by buying complementary health insurance" (p. 206, our italics).

We can see that compulsory insurance *covering just a part of the costs* corresponds to a *minimal extent of insurance coverage.* The minimal insurance coverage which Breyer and Zweifel suggest seems to be based on a notion of need: only medical treatments which belong to a category

[17] See for example Dahlby (1981), Eckstein et al (1985), Eisen (1986), Strassl (1988).

of basic human needs shall be mandatory.[18] "Luxury" medical services can be privately insured for. But what is luxury and what is basic need? Schulenburg gives a hint what normally is meant:

> "...a basic compulsory insurance, i.e. a restriction of the social health insurance on the necessary and social political imperative" (Schulenburg 1994, p. 441).[19]

If the terminology "necessary" by Schulenburg has a meaning at all, there has to be a way to define "necessary" treatments and to determine the amount of care for a typical illness. Then one has to categorize different kinds of treatments and give priority according a system of need. Even though it is possible to agree about a definition about necessary health care - nevertheless the normative statements from the WHO (1946, 1985) are good tries - it will not do to implement an efficient allocation mechanism. The problem is that there are no objective standards about how much care a typical case claims, especially not in an environment of rapid change of medical technology. And we can therefore not restrict the health care supply on the notion of need.[20] It is almost certain that we end up in a system that takes care of everything due to the program of "social political imperative" health care.[21]

From a theoretical standpoint the economic objective is clear: a mandatory partial insurance can be Pareto sanctioned, but no one seems to be capable to formulate what 'partial' really means. We suggest a slightly different way to limit the compulsory insurance scheme and the arguments are purely economical.

We argue that compulsory insurance should only cover the most expensive health care services, acting as a high-cost protection. Private insurance is thus supplemented with a limited state insurance scheme which only provides coverage for the worst health states. On the health insurance market there will be a governmental high-cost insurer and many private low-cost insurers. The notion of need is not present in the argumentation and we do not try to classify different health services from a theory of justice. This does not mean, however, that the high-cost

---

[18] For a definition and an examination of the term *need* in health care see Liss (1993).

[19] Original quotation: "...einer Grundsicherung, d.h. einer Beschränkung der sozialen Krankenversicherung auf das Notwendige und als sozialpolitisch unabweisbare".

[20] Compare the discussion by Hayek (1960, pp. 297-300) on the objectivity-criterion in health care.

[21] In the initial model by the Dekker Commission in Holland the compulsory insurance package should cover some 85 % of health care spending. After some political consultation the compulsory part has increased to about 95 %.

protection is not compatible with widely-accepted moral principles of biomedical ethics. On the contrary, we believe that this approach is well-suited to combine the difficult task of economic efficiency and social justice.

As in the Besley (1988) article on public disaster insurance, we do not permit any comparative advantages for the public insurer in preventing increased expenditures on health care. Hence, given a certain coverage rate, health expenditures are assumed to be at the same level regardless if the health insurance market is supplied by a public or a private insurer.[22] However, as the state has larger properties and shares the risk among all its citizens, it has superior diversification possibilities. We use the Arrow-Lind (1970) approach and assume that the state should evaluate public projects by their certainty equivalent, this means we neglect the safety loading on premium pricing for high-cost policies. The state has thus a comparative advantage vis-à-vis the private sector regarding safety and is therefore a superior risk-bearer for great risks.[23]

5.4.1 *Efficiency gains through high-cost protection*

With a compulsory high-cost insurance scheme, the most severe risks are removed from the market and are taken care of by the state. It implies that the private market only insures for minor and frequent health services. Obviously, the high-cost protection has two advantages.

[22] Most arguments for a high-cost protection rest upon the presence of moral hazard. Besley (1988) shows how a public high-cost protection reduces the purchase of private insurance. This means, according to Besley, "...that the marginal reduction in the insurance premium outweighs the increase in the lump-sum tax required to finance the government insurance" (p. 150). The additional effect from the reduced premium rate in the private sector is that moral hazard is also reduced. "Hence, the insurance premium in the private sector is reduced still further, and the insured is better off" (p. 150). The welfare-improving effects of a mixed system which Besley formulates has been criticized by Selden (1993). According to Selden, Besley´s high-cost insurance only replicates the market equilibrium, and in that case the private insurance decision is unaffected. But, in insurance systems with fewer possibilities to observe the behavior of the policyholders, the incentives for over consumption are at hand. "Higher co-insurance or deductibles (i.e. high-cost protection), and stiffer medical tests and controls than today are probably the only feasible way to deal with moral hazard and cheating in these systems" (Lindbeck 1994, p. 388). Lindbeck as well as Besley and Selden neglect, however, the distributional effects of a high-cost protection. And that is the contribution of this paper.

[23] If we did consider inter-temporal insuring, also the private sector could escape the solvency problem by smoothing out the cash-flows by using the capital market. But as the state has lower risk of insolvency than private actors, it can always become better credit terms.

First, the private market only meets risks with a *high frequency* which simplifies accurate risk calculations. The limit for privately insured medical costs is set at the border where the variance for an average medical treatment in the private sector is at its lowest. With a high-cost scheme the medical cost distributions in the private sector tend to normal distribution. If the number of risks in the insurance pool is large enough, the law of large numbers can work effectively, and the mean medical cost could be regarded as almost constant. It follows that the insurers would not need as much safety loading as before the transition and the insurance premiums would become practically actuarial.

Second, since severe risks are removed from the market, the high-cost protection "transforms" the original heterogeneous risk distribution into a more *homogeneous risk distribution.* In spite of personal characteristics, we are all at times in need of medical care. On most of these occasions the utilization is limited and the differences between distinct groups are not so large. If we focus on Figure 1, we believe that the differences lie rather in the long tail than in the great bulk of minor health services.[24]

5.4.2 *A mixed health insurance system*

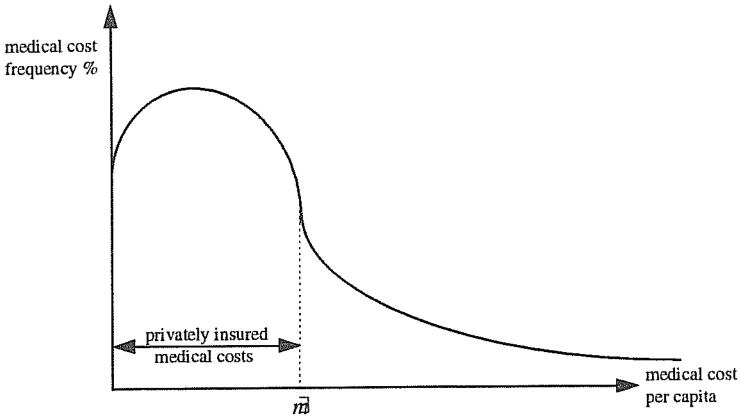
The medical cost distribution goes from health state $j=0$, which is healthiness and requires zero medical costs, to $j=m$, which is the maximum medical care expenditure that can be insured for. With a governmental high-cost scheme the state takes over financial responsibility above the ceiling, $\overline{m}$. The limit for $\overline{m}$ is determined from the shape of the medical cost distribution and it is set at the border where the variance in the private part is at its lowest.[25] In Figure 5.5, it is

[24] Concerning a closely-related theme to health insurance - sickness benefit - the pattern is the same. In Sweden, for example, long-term absence due to illness for practical nurses is longer than for nurses. Social workers are more often at home than other desk-workers such as administrators. The overall frequency of long-term absence due to illness is, however, very small. Only one-fourth of the total absence due to illness is categorized as long-term, but the costs for long-term cases add up to over eighty percent of the total absence costs. Thus, the differences in sickness benefits between groups can mainly be assigned to the differences in high-cost utilization. (Cf. Bjurulf et al. (1990))

[25] This approach differs from other high cost models, as Söderström (1993) or (Söderström et al. 1994), where it is argued that high cost protection should cover the costs when medical costs exceed a certain percentage of available income. The argument for the income independent ceiling, $\overline{m}$, rests solely on efficiency grounds, while the Söderström models also include equity arguments. It should be noted, however, that under the circumstances stipulated here (health risk is exogenous) a high-cost scheme is consistent with a widely-accepted moral principle that no one

shown how individuals bear the entire responsibility for their health care financing up to $\overline{m}$. After $\overline{m}$, the government insures for severe and costly treatments.

Figure 5.5 Split medical cost distribution for the private health insurance market



The medical cost distribution X_i^j is thus divided in two parts: one private part ($j = 0,1,\ldots,\overline{m}$) and one high-cost public part ($j = \overline{m}+1,\ldots,m$). In equation (2), the expectation of X was defined and now we can define the following identity:

$$E(X_i^j) \equiv E(X_i^{pr}) + E(X_i^{pub}), \tag{12}$$

where $E(X_i^{pr}) = \sum_{j=0}^{\overline{m}} \pi_i^j x^j$ and $E(X_i^{pub}) = \sum_{j=\overline{m}+1}^{m} \pi_i^j x^j$.

Each individual is responsible to finance the private part. Either he buys a health insurance on the private market or pays the medical bills on an out-of-pocket basis as they come. This decision depends on the attitude towards risk. Let us assume that the risk aversion for smaller financial hazards is large enough to induce an insurance demand.

According to the law of large numbers, the sample mean will be arbitrarily close to the distributional mean as the number of policy-holders approaches infinity. This does not mean, however, that an

should be punished for circumstances that are outside his or her control. Genetic inheritance is thus no grounds for complete discrimination. For a closer review of an egalitarian theory of justice, see, for example, Rawls (1971). Concerning justice in biomedical ethics, see, for example, Beuchamp and Childress (1989, pp. 256-307).

insurer can operate without any safety loading. A buffer fund is always required. But the buffer fund *per policy*, i.e. the safety loading here considered, can be arbitrarily close to zero. To understand the subject better, let us investigate the central limit theorem (cf. Cummins 1991).

Central Limit Theorem: If $X_1, X_2, ...$ is a random sample from a probability distribution with mean μ and variance σ^2, then the random variable Y with the probability distribution $F_N(y)$ approaches the standard normal distribution N(y) as the number of policyholders (N) approaches infinity:

$$Y = \frac{\sum_{i=1}^{N} X_i - N\mu}{\sigma\sqrt{N}} = \sqrt{N}\left(\frac{\overline{X} - \mu}{\sigma}\right) \tag{13}$$

From equation (13) it is possible to account an acceptable probability of ruin for an insurance pool. Assume that the limiting result implied by the central limit theorem applies exactly so we can write:

$$\Pr\left[\left|\frac{\sum_{i=1}^{N} X_i - N\mu}{\sigma\sqrt{N}}\right| < y_\alpha\right] = \Pr\left[\left|\sum_{i=1}^{N} X_i - N\mu\right| < y_\alpha \sigma\sqrt{N}\right] = 1 - \alpha\,, \tag{14}$$

where y_α is the standard normal deviate factor such that $N(y_\alpha) = 1 - \alpha$. The formula says, that an insurer must have on hand a total buffer fund of $y_\alpha \sigma\sqrt{N}$ to be assured with probability $1-\alpha$ that all claims will be paid. This implies a safety loading per policy in the height of: $y_\alpha \sigma / \sqrt{N}$. With the worst losses covered by the state, the standard deviation as well as the insurable amount for an average medical cost that the market faces, will decrease.[26] As the variability is diminished - but not the numbers of policyholders - the required safety loading for private contracts is also reduced. An important feature of a high-cost protection is that the policies issued on the private market do not require as much safety loading per policy to be as safe as before the transition. To be distinct, we believe that safety loading in the mixed system with many competitive low-cost insurers almost goes to zero, i.e. $\lambda \approx 0$.[27] Omitting

---

26 $E(X_i^{pr}) < E(X_i)$ and $\sigma_{X_i^{pr}} < \sigma_{X_i}$.

27 However, as Cummins (1991, p. 265) says, the central limit theorem "...is of limited practical relevance in working with insurance loss distributions because the distributions usually encountered in practice approach normality so slowly that they

the safety loading, the allocative efficiency is increased and the gross premium is calculated as

$$P_i^{pr} = E(X_i^{pr}). \tag{15}$$

The high-cost protection is financed by a proportional income tax (z) under the restriction that the budget must break even. If θ is the share of the total population (N) that belongs to group A, we can define the following relationship between compulsory insurance-costs and taxation income:

$$N(\theta E(X_A^{pub}) + (1-\theta)E(X_B^{pub})) = \sum_{j=\overline{m}+1}^{m} x_i^j = y^0 zN \tag{16}$$

Equation (16) says that the sum of the medical costs above the ceiling must equal the total revenue from the income tax. The net income will then be $y^0(1-z)$ for both groups. Since $\sigma_{X_A} < \sigma_{X_B}$, the B-group has a greater part of their medical costs in the high-cost range than the A-group. The B-individuals will therefore enjoy more benefits from high-cost protection than the A-individuals. Thus, the income tax redistributes wealth from A to B. If the new endowment position is compared to the old one, the following relationship must hold:

$$\left[y^0 - X_A\right] > \left[y^0(1-z) - X_A^{pr}\right] > \left[y^0(1-z) - X_B^{pr}\right] > \left[y^0 - X_B\right]. \tag{17}$$

The relationship in (17) tells us that taxation smoothes out the differences in expected income between the two risk groups. Before any insurance purchases have been realized the A-individuals are worse off and B-individuals better off than before taxation. To find out the market equilibrium under governmental intervention, we maximize the expected utility subject to the high-cost budget constraint:

---

remain significantly skewed even in large insurance pools." Cummins argument is not so critical under a high-cost protection scheme, since the distribution for the privately insured health expenditures are more normal distributed than the total loss distribution. Nevertheless, also the private loss distribution is skewed. But, let us for reasons of simplicity and clearness assume that an average cost in the insurance pool with a high-cost protection will be approximately constant. We can therefore neglect the safety loading, although we are aware that an insurer under a high-cost scheme has to take out some safety loading as well. But as the safety loading will be much smaller in a mixed system in comparison with a completely private market system, this simplification will not be misleading.

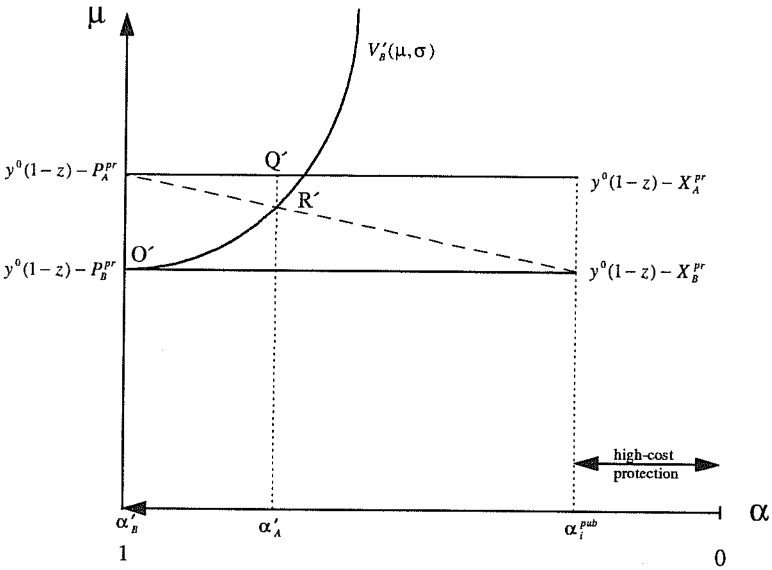
$$\max! \quad V(\mu_i, \sigma_i(\alpha)) \tag{18}$$

$$\text{s.t.} \quad \mu_i = y^0(1-z) - \left[E(X_i^{pub}) + E(X_i^{pr})\right] + \left[E(X_i^{pub}) + \alpha E(X_i^{pr})\right] - \alpha E(X_i^{pr})$$

$$\mu_i = y^0(1-z) - E(X_i^{pr}), \tag{19}$$

where the standard deviation of income now is a function of the standard deviation of the private medical cost distribution: $\sigma_i = (1-\alpha)\sigma(X_i^{pr})$. As the premium is actuarially fair, the value of the expected wealth will not change with α. A risk averse applicant will therefore avoid the whole risk by buying an insurance contract with full coverage, $\alpha = 1$. Accordingly, an individual faces the sure net income in the height of: $y^0(1-z) - P_i^{pr}$, where $P_i^{pr} = E(X_i^{pr})$. The graphical solution can be seen in Figure 5.6.

Figure 5.6 Mixed system: Private insurance market with a governmental high-cost protection



On the insurance market only contracts for minor health care costs are traded. On the completely private market the A- and the B-group had the endowment positions $\left[y^0 - X_A\right]$ and $\left[y^0 - X_B\right]$, respectively. In the mixed system, taxation reduces the endowed income for each group with the same size, $y^0(1-z)$. On an aggregated level, taxation is completely offset by the lesser insurable risk. On an individual level, however, taxation is only partly compensated by lesser expected loss for the A-group and overcompensated for the B-group, as can be seen from the shifts to the new endowment positions $\left[y^0(1-z)-X_A^{pr}\right]$ and $\left[y^0(1-z)-X_B^{pr}\right]$.

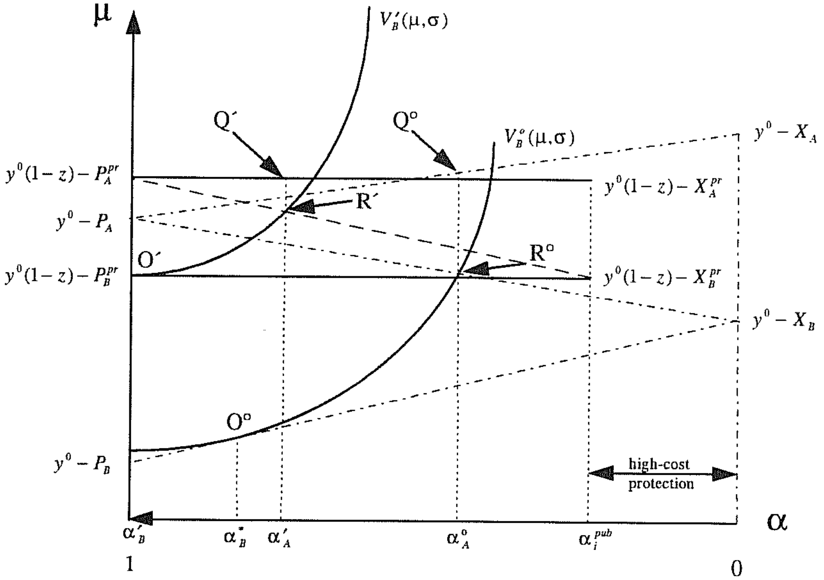
As the public insurance scheme has taken care of the most severe risks, the variation of an average insurance policy on the private market as well as the probability of ruin is diminished. The policies must not be loaded by any safety loading and as we have neglected administrative costs, the premiums are practically actuarial. Graphically this means that the budget lines become horizontal in the limit.

As earlier, the information is private and the insurers cannot discriminate the individuals. The bad risks will therefore dominate the market outcome. In the same way as in 5.3.2, B-individuals optimize their insurance protection at point O', which now means that B-individuals can enjoy a complete health insurance coverage since the premiums are not loaded. And the self-selection constraint implies that B-individuals must be indifferent between O' and the contract scheme offered to A-individuals. A possible equilibrium is where the utility-curve of the B-individuals (V_B') cuts the long dashed line at point R'. At R', insurance policies can be bought at a unit price that is lower than the unit price for an optimal B-policy, but the corresponding coverage rate is smaller, i.e. $(\alpha_A', P_A') < (\alpha_B', P_B')$. If (α_A', P_A') is offered to A-individuals, B-individuals have no incentive to change their purchasing behavior, and therefore the A-individuals can reach the point Q', Hence, the allocation (O', Q') is an equilibrium.

5.4.3 *Is governmental intervention Pareto sanctioned?*

The important question now arises whether the transition from a completely private solution to a mixed system is Pareto sanctioned. The equilibrium position in the mixed system, (O', Q'), has to be compared with the equilibrium position in the completely private market (O^o, Q^o). The mixed system is Pareto-superior over the completely private market, if the equilibrium (O', Q') dominates over (O^o, Q^o). Since the tax-distribution already has put the B-individuals in a better position than before the transition, the question is whether A-individuals experience a greater utility in a mixed system than on the completely private market, i.e. if $V_A(Q') > V_A(Q^o)$. Two forces determine whether this is the case, the degree of risk aversion and the efficiency gains.

Figure 5.7 Comparison of a mixed system with a complete private health insurance market



First, the subsidization from A- to B-individuals has increased the expected utility for the B-individuals. Accordingly, the self-selection constraint has been eased, which is indicated by a shift of the B-individuals´ utility curve in Figure 5.7, $V'_B > V^o_B$. It has become easier for the good risks to reach a high coverage rate. If risk aversion is large enough, the utility loss due to lost wealth for the A-individuals through taxation can be offset by a utility gain through higher insurance coverage.

Second, with a high-cost protection the insurable amount as well as the variance will decrease. Not only the irregularities of payments will be reduced, but also the insurers´ probability of ruin. The premium will therefore only reflect the expected costs and rather small safety arrangements, ($\lambda \approx 0$). This leads to a more efficient resource allocation which is indicated by the horizontal high-cost budget lines in Figure 5.7. And since the premiums are almost actuarial, the insurance coverage will be larger for both groups. In comparison to other partial insurance schemes, the success of the mixed system depends mainly on which effect the high-cost protection has on the allocation on the private market.

The mixed system is, thus, Pareto superior if the utility gains of the increased insurance coverage due to relaxing the self-selection constraint and more efficient market allocation is greater than the utility loss due to taxation and redistribution, or formally expressed as

$$V_A(Q') = V(\mu'_A, \sigma'_A | \alpha'_A, P'_A) > V(\mu^o_A, \sigma^o_A | \alpha^o_A, P^o_A) = V_A(Q^o). \qquad (20)$$

In summary, health states in the interval $[\overline{m}+1, m]$ are covered by compulsory insurance. The cost of doing so is the expected value of medical expenditures in this interval at full coverage rate $(\alpha = 1)$. The expenses for private insurance are reduced by this amount so that the tax used to finance the public insurance is offset by a reduction in the private insurance premium. In addition, we claim that the high-cost protection is welfare-improving. More specifically, compulsory high-cost insurance creates two advantages: (i) it reduces the adverse selection by smoothing out the expected loss-distributions and (ii) it increases the allocative efficiency through reducing the variability of an insurance unit on the private market for less severe illness.

5.5 Concluding remarks

It has been argued in this paper, that a mixed system of health insurance in which the government covers high-costs risks and voluntary private insurance covering low-cost risks, could Pareto dominate to a purely private insurance system in presence of adverse selection. This result has been derived, as is very common, in a static one-period framework. Given the problem addressed in this contribution, this approach is limited and what is wanted for is an extension towards a more-period model. There are, however, some reasons to defend the one period approach, at least in part. Given a finite number of policyholders, it is well known, that the working of the law of large numbers need not lead to a complete reduction of the variance even if there is insurance supplied by a monopolist. Thus we expect a positive loading. Moreover, competition would lead to higher premiums, since loading must be higher due to smaller numbers of insurance contracts held by each insurer. Thus we expect the convergence of a natural monopoly and the existence of monopoly rents - even if market competition is allowed for. In any case, managers of insurance companies (monopolies or competitive companies) are under pressure to pay dividends to shareholders instead of building up sustainable buffer funds to secure for the rare events of catastrophic expenses. Expecting this each period leads us to belief, that the one period model is not as bad a theoretical guide for policy as it may look like.

Appendix

Location and scale parameter condition

To show the consistency between a two-moment decision model and the classical expected utility model, we base the proofs of a survey on the theme made by Meyer (1987). The location and scale parameter condition is defined as follows:

> Two cumulative distribution function $G_A(\cdot)$ and $G_B(\cdot)$ are said to differ only by location parameters α and β if $G_A(x) = G_B(\alpha + \beta x)$ with $\beta > 0$.

Meyer says that in order for the two-moment decision model to be consistent with the expected utility model, the choice set is to be composed of random variables which differ from one another only by location and scale parameters. This means, according to Rothschild and Stiglitz (1970), that changes in distributions "...amount only to a change in the centering of the distribution and a uniform shrinking or stretching of the distribution - equivalent to a change in units" (p 241). The location and scale condition places, however, no restrictions on the functional form of the cumulative density function. Any distribution will do, as long as all alternative distributions to which the original distribution is to be compared are similarly distributed in the location and scale parameter sense. Quadratic utility functions or normally distributed pay-offs are thus no binding constraints.

If $X_i \in \{0, X_i^1, ..., X_i^m\}$, i = A,B, is a set of medical cost distributions for two groups of individuals, A and B, then the location and scale condition is fulfilled when the random variables X_i differ from one another only by location and scale parameters. As the disposable income is a direct function of medical expenses, $Y_i = y^0 - X_i$, the distributional properties for the medical cost distribution are also valid for the income distributions $Y_i \in \{y^0, Y_i^1, ..., Y_i^m\}$.

Let Z be the random variable obtained from one of the Y_i using the normalizing transformation $Z = (Y_i - \mu_i)/\sigma_i$, where μ_i is the mean and σ_i is the standard deviation of Y_i. The random variable Y_i is therefore equal in distribution to $\mu_i + \sigma_i Z$. Hence, the expected utility from Y_i for individuals in groups A and B with utility function u can be written as

$$EU(Y_i) = \sum_{j=0}^{m} \pi_i^j u(\mu_i + \sigma_i z) \equiv V(\mu_i, \sigma_i). \tag{A1}$$

The objective function in (A1) possesses the following properties:

$$V_\mu \geq 0 \qquad \forall\ \mu \text{ and } \sigma \geq 0 \text{ iff } u'(\mu + \sigma z) \geq 0\ \forall\ \mu + \sigma z$$

$$V_\sigma \leq 0 \qquad \forall\ \mu \text{ and } \sigma \geq 0 \text{ iff } u''(\mu + \sigma z) \leq 0\ \forall\ \mu + \sigma z$$

The slope of an indifference curve is thus

$$\frac{d\mu}{d\sigma}\Big|_{V(\mu,\sigma)} = -\frac{V_\mu(\mu,\sigma)}{V_\sigma(\mu,\sigma)} = S(\mu,\sigma), \tag{A2}$$

and the indifference curve possesses the following property:

$$\frac{d\mu}{d\sigma}\Big|_{V(\mu,\sigma)} \geq 0 \quad \forall\ \mu \text{ and } \sigma \geq 0 \text{ if } u'(\mu + \sigma z) \geq 0 \text{ and } u''(\mu + \sigma z) \leq 0\ \forall\ \mu + \sigma z.$$

To find out how the slope of an indifference curve changes as one moves in the vertical (μ) direction, we differentiate (A2) subject to the mean income:

$$\frac{\partial S(\mu,\sigma)}{\partial \mu} \leq (=, \geq)\ 0 \tag{A3}$$

if, and only if, $u(\mu + \sigma z)$ displays decreasing (constant, increasing) absolute risk aversion $\forall \mu + \sigma z$, where the Arrow-Pratt measure of absolute risk aversion: $R_A(\mu + \sigma z) = -u''(\mu + \sigma z)/u'(\mu + \sigma z)$, is used. Considering proportional rather than absolute changes in an individual's income levels, we analyze the income vector as it moves along a ray:

$$\frac{\partial S(t\mu, t\sigma)}{\partial t} \geq (=, \leq)\ 0 \tag{A4}$$

if, and only if, $u(\mu + \sigma z)$ displays increasing (constant, decreasing) relative risk aversion $\forall\ \mu + \sigma z$, where the relative risk aversion is measured according to Arrow-Pratt as:

$$R_A(\mu + \sigma z) = -(\mu + \sigma z)\, u''(\mu + \sigma z)/u'(\mu + \sigma z). \tag{A5}$$

The location and scale conditions allow for the ordinary assumption of decreasing absolute risk aversion and increasing relative risk aversion in a $\mu - \sigma$ space.

Proofs of the properties in (A3) and (A4) can be found in Meyer (1987, pp. 428-429).

6. Social risk in health insurance*

6.1 Introduction

In this paper we consider health risks that are correlated between individuals. A frequent example is the annually returning influenza, striking large parts of the population. A non-frequent example is radioactive downfall following an accident in a nuclear power plant, similar to Chernobyl in 1986. Individual risks are in these cases not perfectly offsetting and an insurance pool cannot eliminate all risk through diversification. Risk exposure that cannot be diversified within an insurance system is called *social risk*.

The purpose of this paper is to discuss consequences of social risks in health insurance and how it is administrated by two different organizations, a mutual insurance pool and an insurance entrepreneur. The mutual pool is assumed to be specialized in health risks. The main characteristic of an insurance entrepreneur is that he works according to a risk management concept, not only managing health risks, but all forms of insurance contracts and securities. Our discussion can hopefully put the debate on regulations of health insurance markets into perspective.

6.2 Health insurance in a mutual insurance pool

6.2.1 *A small mutual insurance pool*

Let us first consider the case of two individuals (*A* and *B*) with the same gross income (*Y*) and also the same probability of getting ill ($\pi^i = \pi\ ; i = A, B$). During a particular period of time they can be either healthy or ill. We assume that there is public information concerning health risks and that the financial cost of illness is *L*. The endowment of

* The paper was co-authored by Professor Lars Söderström, University of Gothenburg, and it was presented at the 15th Nordic Health Economists' Study Group Meeting, Reykjavik, Iceland, August, 18-19, 1994. Original title: "Social Risk in Health Insurance - Comparison between a mutual insurance pool and an insurance entrepreneur".

individual i in a new period is stochastic, equal to Y in a healthy state and Y-L in the presence of illness. Due to risk-aversion, an individual wishes to transfer money from the healthy state to the state of illness by means of health insurance.

If, initially, there is no insurance market, A and B can join in a mutual insurance pool for the purpose of sharing health costs. Assume that this pool is totally dependent on premiums currently paid by the policyholders and that it delivers no other service than risk sharing. Furthermore, assume that the pool which is owned by its members operates without a profit motive and that there are no administration costs.

The mutual pool is exposed to four different states of the world ($s=1,...,4$), to wit, both individuals are healthy; A is healthy and B is ill; A is ill and B is healthy; both are ill. If we assume that health risks are independent, the probability of illness and the corresponding health care costs for the various states are as follows:

individual A \ individual B *(probability; illness costs)*	*healthy* $(1-\pi;0)$	*ill* $(\pi;L)$
healthy $(1-\pi;0)$	$(1-\pi)^2;0$	$(1-\pi)\pi;L$
ill $(\pi;L)$	$\pi(1-\pi);L$	$\pi^2;L+L$

How may such a mutual insurance pool be organized? Let us first assume that the pool does not charge any premium but instead charges policyholders, *ex post*, whatever is needed to match losses as they come. If both individuals are healthy, there will be no charge at all and each individual will get his gross income Y. If one is healthy and the other is ill, a transfer from the healthy to the ill individual takes place and both will get the disposable income $Y-L/2$. But, if both are ill at the same time, no transfer is possible and each individual will have to bear the whole financial cost due to illness, $Y-L$. Hence, in this case there is no insurance available. For a small mutual insurance pool there is no solution to this problem. Complete insurance is not feasible.

A large insurance pool has a different situation. By taking advantage of the Law of Large Numbers the pool will be able to offer a wider insurance coverage.

6.2.2 *A large mutual insurance pool*

Let us now consider a pool with many members. As before, individuals are assumed to have the same gross income (Y) and the same probability of becoming ill ($\pi^i = \pi$), but here the number of policyholders is large ($i = 1,...,N$). *Ex ante*, an individual pays a fixed premium (P) for an insurance coverage and in case of illness he gets an indemnity (I) for health care costs. An individual's disposable income will be $y = Y - P$ when healthy and $y = Y - P - L + I$ when he is ill. Let $u(y)$ denote the individual's utility as a function of his disposable income (y), and assume that he is risk averse, ($u' > 0; u'' \leq 0$).

Policyholders are allowed to choose their own insurance coverage. By setting a price (p) per unit of compensation,[1] it is up to each policyholder to choose appropriate coverage. The insurance policy he chooses may cover the entire loss or parts of it, $I = \alpha L$, where $0 \leq \alpha \leq 1$. Each policyholder maximizes expected utility given his budget constraint.

Max! $$U = \pi u(y_s) + (1-\pi)u(y_h) \qquad (1)$$

s.t. $$p y_s + (1-p) y_h = p(Y - L) + (1-p)Y \qquad (s = \text{ill}, h = \text{healthy}) \quad (2)$$

Given that an inner solution exists, the optimal risk allocation implies that the ratio of expected marginal utility per monetary unit is equal in every state of nature:

$$\frac{u'(y_s)}{u'(y_h)} = \frac{p(1-\pi)}{(1-p)\pi}. \qquad (3)$$

A necessary condition for full reimbursement of health care costs ($I = L$) is that $p = \pi$, i.e. that the premium per unit of compensation equals the probability of becoming ill. This is the case of an actuarially fair insurance. In this case, people demand complete insurance.

For the mutual insurer to be able to provide actuarially fair insurance, the average risk of the entire pool must be regarded as approximately constant. Then no uncertainty concerning the amount of future claims exists and the insurer can issue policies equal to the expected value of becoming ill.

Consider now the case when the average loss in the pool varies from one state of nature to another. To stay solvent, the pool must have a

---

[1] $p = P/I(L)$

buffer fund. For the purpose of building this fund the pool adds a loading factor (λ) to the premiums asked from policyholders: $P(I) = (1+\lambda)\pi I$. Given that premiums are paid one period ahead, each policyholder has to pay the premium

$$P(I) = \frac{(1+\lambda)\pi I}{1+r_f}, \tag{4}$$

where r_f is a risk free rate of interest. Unless λ is equal r_f, this premium is not actuarially fair and from (3) we see that the optimal coverage implies a higher marginal utility of income in the state of illness than in the state of health. Hence, the optimal insurance coverage will not be complete, $I < L$.

According to the Law of Large Numbers the variation in claims decreases as more individuals participate in the pool. The safety loading will accordingly become smaller, but it will not vanish. There is a limit to what diversification can do.

If we expand the number of members in the pool ($i = 1,\ldots,\mathrm{N}$) the per capita medical cost amounts to $\theta = \frac{1}{N}\sum_{i=1}^{N} L^i$. We assume that individual losses (L^i) are identically distributed, has the same mean (μ) and variance (σ^2), and also that the correlation between all medical costs (ρ) is constant. The variance of the per capita cost will then be[2]

$$\sigma_\theta^2 = E\{(\theta - \mu)^2\} = \sigma^2\left\{\frac{1+\rho(N-1)}{N}\right\}. \tag{5}$$

When health risks are independent ($\rho = 0$), private risks offset each other completely if the number of policyholders becomes large enough. In this case the variance of the medical cost per capita goes to zero as the number of pool members becomes very large.[3]

When health risks are correlated ($\rho > 0$), this result does not hold. In the limit, as the number of members in the pool approaches infinity, the variance of the per capita loss will not go to zero but $\rho\sigma^2$, that is the covariance between the medical costs for any pair of individuals.[4] No matter how large the insurance pool is, as long as average health risks are positively correlated, the variance of per capita losses will remain

---

[2] See Markowitz (1959, pp. 109) or Hirshleifer and Riley (1992, p. 130).

[3] $\lim_{N\to\infty} \sigma_\theta^2 = \lim_{N\to\infty}(\sigma^2/N) = 0$.

[4] In the limit $\rho\sigma^2 = E\{(L^i - \mu)(L^j - \mu)\} = \sigma_{L^i L^j}$.

positive. Hence, social risk is not only due to small numbers. If average health risks are correlated, the social risk persists even with large numbers.[5]

In conclusion, we do not expect that a mutual insurance pool will offer its services on an actuarially fair basis. One reason is that the Law of Large Numbers does not apply. A small mutual insurance pool is unable to provide insurance that covers all possible cases. Another reason is that average health risks are positively correlated. In both cases, insurance premiums are inflated with a loading factor.

6.2.3 *Social risk and premium setting in a mutual insurance pool*

Due to social risks a mutual pool cannot eliminate the risk of insolvency even though the pool becomes very large. By adding a loading factor, however, the risk of insolvency diminish. Let us here consider how this loading can be calculated statistically.[6]

Assume that medical costs follow a normal distribution. Under this assumption, it is straightforward to calculate a confidence interval within which the insurance pool can manage its payment obligations,

$$\Pr\left(\left| \frac{\tilde{\theta} - \mu}{\sigma_\theta} \right| < Z_\eta \right) = \Pr\left(\left| \tilde{\theta} - \mu \right| < Z_\eta \sigma_\theta \right) = 1 - \eta \, , \tag{6}$$

[5] Diversification works over time as well. Provided that the premiums paid exceed expected claims the pool will successively acquire a fund to cover extreme cases. As the fund grows, premiums may be lowered. After the first period ($t = 1$), the pool will have a premium fund, $S_1 = P_1 - X_1$, where X_1 denotes claims actually paid in the first period. Repeating this procedure over several periods, the cost of illness will be diversified over time, reducing safety loadings in the long run. For example, an influenza in period 1 does not influence illness in period 2. Other events may not be diversified completely over time, for example, diseases following radioactive downfall. However, to illustrate social risk, it is enough to consider diversification among policyholders in one period. In reality, these two diversification systems work side by side.

[6] Regarding the possibility of insolvency the mutual can react in different ways. It can rely on creditors to overcome the liquidity problems. Faced with the risk of bankruptcy creditors demand a relatively high rate of interest for their services. Likewise, the mutual can reduce or suspend payments, which shifts the risk from the mutual to its policyholders. A third option is build up a buffer fund by inflating the premiums asked. All ways imply incomplete risk transformation. Here, we consider the most common way, that is to charge a safety loading.

where Z_η is the standard normal deviate factor which on a 95% confidence interval ($\eta = 0.05$) takes the value $Z_{0.05} = 1.96$. Thus, equation (6) tells us that the average health care cost ($\tilde{\theta}$) will stay in the calculated interval $|\tilde{\theta} - \mu|$ with the probability $(1-\eta)$.[7]

To make sure that all claims will be met, the pool has to establish a buffer fund by charging a loading per insurance policy that amounts to $\lambda = Z_\eta \sigma_\theta$. Hence, the confidence interval prescribes how much loading it takes for the pool to stay solvent.

Under the assumption that health risks are independently distributed, the standard deviation for the per-capita loss is $\sigma_\theta = \sigma/\sqrt{N}$. In this case it is straightforward to calculate an insurance premium in the same manner as in equation (4).

$$P(I) = \frac{\left(1 + Z_\eta\, \sigma/\sqrt{N}\right)\pi I}{1 + r_f} \tag{7}$$

In equation (7), $Z_\eta\, \sigma/\sqrt{N}$ equals the loading factor λ in equation (4).

It is evident in this case, that the loading factor per contract tends to zero, as the number of policyholders tends to infinity.[8] This is also expected since health risks are assumed to be of a solely private character. In this case an insurance pool specializing in health risks will lead to a complete diversification and premiums will be actuarially fair.[9]

We saw earlier that when health risks are not independent the variance of an average policy equals the average covariance between health risks ($\sigma_\theta^2 = \sigma_{L^i L^j}$). To calculate the buffer fund per policy in the presence of social risk, we must replace the standard deviation of the loading term in (7) with the square root of the covariance.

$$P(I) = \frac{\left(1 + Z_\eta\, \sigma/\sqrt{\sigma_{L^i L^j}}\right)\pi I}{1 + r_f} \tag{8}$$

The premium in equation (8) refers to an insurance policy with social risk. The risk of insolvency due to social risk is hedged by the loading

[7] Cf. Goldberger (1991, pp. 123-124).

[8] The variance of per-capita medical cost decreases with the number of policyholders, i.e. $\lim_{N\to\infty} \sigma/\sqrt{N} = 0$.

[9] However, the total capital reserve of the insurance company does not decrease to zero with the number of policyholders. Cummins (1991) explains this by referring to the absolute risk of the insurer not disappearing in large pools.

factor $Z_\eta \sigma / \sqrt{\sigma_{L^i L^j}}$. Regardless of how well diversified the insurance pool is, to stay solvent the premium must always be loaded. Hence, in the presence of social risk, the capital reserve must grow with the number of insurance contracts.[10]

6.3 Health insurance issued by an insurance entrepreneur

6.3.1 Risk management

In organizing health insurance, regulations have forced insurers to specialize on a single health insurance line. In Germany, for example, the mandatory health insurance system, Gesetzliche Krankenversicherung, is managed by a particular state insurer only insuring health risks. In Sweden, newly organized Family Doctors[11] are paid on a capitation basis and have to deliver services to citizens in very small areas, often blocks. A health risk is in these cases diversified within the area of other health risks. In regard to the problem of dependency between health risks, this does not seem to be an optimal solution.

In a single insurance line, the dependency between risks is reduced by enlarging the insurance pool. Evidently, economies of scale in health insurance production exist. However, the economies of scale will possibly eliminate market competition. Diversification gains force health insurers to converge into larger units and we may end up in a monopoly situation. To this problem, a risk management approach gives new conditions for market competition.

Risk management will here be exemplified by discussing how an insurance entrepreneur can use the market institution to diversify health risks. So far we have only discussed a pool consisting of health risks. Another way to reduce the correlation between health risks is to diversify among other kinds of risks, to use the economies of scope of in-

[10] Whether insurers will allow the capital reserve to grow with every new contract is a different matter. Borch (1985) suggests that insurers sense decreasing utility from accumulated capital reserves. If so, an insurer will set an upper limit for the capital reserve and distribute the excess capital to its members rather than let its own capital grow unlimitedly. However, there is a trade off between refunding and avoiding the probability of ruin. The more capital reserves are refunded, the larger the probability of insolvency. On the other hand, a completely risk free insurance contract is costly. Protection for every possible state (if feasible) would amount to a very huge insurance premium.

[11] Husläkarsystemet.

suring different risks. For an insurance entrepreneur it is important to analyze how insured risks behave in comparison to other risks in the portfolio. Given that a risk reduces the variance of the portfolio there are reasons to include this risk regardless of its origin. The claim for independence is less important, since negatively correlated risks have a positive value for a risk-averse portfolio manager.

This change from a mutual pool with a fixed insurance line to an entrepreneurial company with different diversification opportunities will illustrate how an insurance company can organize health insurance according to the risk management concept.

6.3.2 *Insurance CAPM*

From the time premiums are paid until an average reimbursement occurs, there are opportunities for capital investments. If the insurance entrepreneur invests parts of his capital in the capital market the Law of Large Numbers will be applied to a broader extent. Health insurance contracts will no longer be judged with respect to how they are related among members in an insurance pool, but how they are related to the market as a whole.

To describe the equilibrium condition between risk and return in a competitive market, we use the Capital Asset Pricing Model (Sharpe 1964, Lintner 1965). The CAPM says that the return on an asset is a function of the market risk premium and the covariability of the asset with the market portfolio. If we use the same technique of modeling and apply it to insurance, we can gain some insights as to how health risk are valued by a competitive market.

Several authors have applied the original CAPM for insurance purposes (Biger and Kahane 1978, Fairley 1979, Hill and Modigliani 1987). In these models, the value of an insurance contract from the underwriters' point of view is a function of two sources, the riskless rate of interest and the systematic risk of a policy. Following the Fairly model, equation (9) gives the key-features of the insurance CAPM.

$$r_U = -kr_f + \beta_U (r_m - r_f), \tag{9}$$

where

r_U = expected rate of underwriting return on an insurance contract,

k = funds generating coefficient,[12]
r_f = risk free return,
r_m = expected return on the market portfolio,
β_U = underwriting beta of the insurance contract.[13]

Premiums are paid in advance and the funds generating coefficient k indicates "the average amount of investable funds created by the cash flow per dollar of annual premium" (Fairley 1979, p. 196).[14] Underwriting return (r_U) is a linear function of two parts. The first part of equation (9) represents interest payments to the policyholders for giving the insurer a "loan". The factor ($-kr_f$) reduces required premiums and policyholders receive an implicit interest payment in parity with the riskless interest rate. The second part of the underwriting return is the insurer's compensation for bearing risk. This part is equal to the underwriting beta (β_U) multiplied with the market risk premium ($r_m - r_f$). In the same manner as investors in the original CAPM only get compensation for bearing systematic risk, insurers are also only rewarded for bearing systematic risk, which is expressed by the beta-value.

From the insurance CAPM in (9), some interesting insights in the functioning of competitive insurance markets can be attained. If underwriting profits are positively correlated with the market return ($\beta > 0$), then the insurance company will get compensation for bearing systematic risk. In the opposite case when underwriting returns are negatively correlated with the market return ($\beta < 0$), the model states that the insurer should pay a risk premium to the policyholders!

Underwriting profits can take negative values. This is due to the delay of reimbursements. The insurer can invest the premium funds at an interest rate at least as good as the risk free rate of return, and policyholders are implicitly refunded in parity with the risk free interest rate for lending to the insurer.[15]

---

[12] k is also known as the 'liabilities-to-premiums ratio',

$$k = \sum_{i=1}^{N} \alpha L^i \Big/ \sum_{i=1}^{N} P^i; \; 0 \leq \alpha \leq 1,$$

[13] $\beta_U = Cov(r_U, r_m) / Var(r_m)$

[14] Cummins (1991) interprets k as the average time between policy issue and claims payment or the average holding period.

[15] In this context it should be noted that the underwriting return does not depend on the composition of the investment portfolio the insurer actually takes, nor on the outcomes of the investments. As equation (9) clearly points out, the underwriting return on a health insurance contract solely depends on the risk free return, the length of the cash flow and the social risk. The insurer can earn a higher (expected)

6.3.3 *Social risk and premium setting by an insurance entrepreneur*

By using the insurance CAPM, we can explain how premiums are set in a competitive market. If perfect competition prevails premiums equal discounted expected medical costs (cf. equation (4)),

$$P = \frac{\pi \alpha L}{1+r_f} = \frac{\pi I}{1+r_f}, \qquad (i=1,...,N) \qquad (10)$$

where α is the rate of reimbursement, $0 \le \alpha \le 1$. The yield of an insurance contract can be calculated according to ordinary rate of return formulas:

$$r_U = \frac{P - \pi I}{P^i} = \frac{\frac{\pi I}{1+r_f} - \pi I}{\frac{\pi I}{1+r_f}} = \frac{\pi I - \pi I(1+r_f)}{\pi I} = -r_f. \qquad (11)$$

From equation (11) it is possible to draw at least three interesting conclusions. First, the profit from underwriting is highly correlated with the interest rate; second, an insurance entrepreneur can increase his profit until the underwriting profit on an insurance contract equals $-r_f$; and third, in equilibrium all insurance contracts must generate the same rate of return.

By combining the insurance CAPM in equation (9) with (11) and solve for the premium, we get an equation showing how market prices of a health insurance contract vary according to risk,

$$P^i = \frac{\pi^i I^i}{1 + kr_f - \beta_U(r_m - r_f)}. \qquad (12)$$

Consider the case when $k = 1$. Then if underwriting beta equals zero, i.e. there exists no systematic risk, equation (12) is reduced to (10). The premium is in this case equal to its actuarial value and policyholders demand complete coverage ($I = L$). If systematic risk exists, an insurer wants compensation for the non-diversifiable risk he takes. In that case the underwriting beta is larger than zero and the premiums exceed the corresponding expected value. In other words, the premium is loaded. This implies that policyholders demand an incomplete insurance coverage ($I < L$).

---

return by investing in riskier assets, but it is the insurer, not the policyholders, who bears the extra risk.

In conclusion, in a competitive insurance market the insurer is only compensated for non-diversifiable systematic risk. The price of an insurance contract will be adjusted until the rate of return above the riskless interest rate reflects the systematic risk of the contract. By combining health risks with other capital assets, the insurer diversifies health risks in a wider insurance pool. Hence, economies of scope reduce the need for scale production in health insurance.[16]

6.4 Social risk and efficiency

In this paper, we consider the problem of dependencies between health risks. The essential question is in which diversification system social risk is smallest - in a mutual insurance pool with only health risks or in an entrepreneur firm with different kinds of risks.

A mutual insurer uses pooling among similar health risks. When risks are on the average correlated with each other, social risk becomes a problem for premium setting. As is described in 6.2.3, the social risk in a mutual insurance pool equals the covariance between an average pair of medical costs in the distribution ($\sigma_{L^iL^j}$).

An insurance entrepreneur uses the capital market to diversify health risks. The main question is therefore not how the insured risk covaries with other health risks but how it covaries with the market portfolio. In this respect, underwriting beta states how a health risk covaries with the market return. If underwriting beta is multiplied with the variance of the market return, we get a comparable measure of social risk by the insurance entrepreneur, i.e.

$$\beta_U Var(r_m) = \left[\frac{Cov(r_U, r_m)}{Var(r_m)}\right] Var(r_m) = \sigma_{r_U r_m}. \tag{13}$$

[16]We must notice, however, that the premium calculation in (12) is done to reflect risks. Even though premiums exceed expected costs in the presence of systematic risks, there is no guarantee that the insurer stays solvent. The risk of ruin is not priced and policies are treated as free of default risk. If there is a risk of default contract, many of the standard results as to how a policyholder chooses an insurance coverage seems to be invalid (Tapiero et al. 1986, Schlesinger and Schulenburg 1987, Doherty and Schlesinger 1990). However, the problem of insolvency should not be exaggerated. With respect to property-liability insurance Hill (1979) argues that the insurance industry produce "safe" insurance policies, partly because of insurance knowledge and experience and partly due to regulations. Also reinsurance arrangements reduce the probability of ruin, which is worth a price many are willing to pay (Schlesinger and Doherty 1985).

If the covariance between two health risks in the insurance pool is larger than the covariance between a health risk and the market, then the insurance entrepreneur is superior risk-bearer. This is the case when

$$\sigma_{L^i L^j} > \sigma_{r_U r_m}. \tag{14a}$$

On the other hand, if the social risk in a mutual pool is smaller than the social risk in a market portfolio, then the mutual insurer is superior risk-bearer. In this case the covariance between a health risk and the market is larger than the covariance between two health risks in the insurance pool,

$$\sigma_{L^i L^j} < \sigma_{r_U r_m}. \tag{14b}$$

The covariance terms indicate, thus, the diversification efficiency of respective insurer. In a competitive health insurance market, market forces will judge who is the most efficient risk-bearer. For example, if an insurance entrepreneur can exploit the diversification possibilities in the capital market and thereby lower the social risk, he can issue lower premiums and increase his market share. However, the logic of market competition can not be tested with respect to health care financing alone. Few sectors in the economy are so regulated as health care financing. The intention of this paper is to reflect upon the need of regulation in health care financing.[17]

[17] In spite of a likely advantage of an insurance entrepreneur over a mutual insurer regarding social risk, health as well as life insurance is often organized as mutuals. Hansmann (1985) explains the success of the mutual form in the life insurance business from a transaction cost perspective. There are large contracting costs in signing a lifetime contract and the information asymmetry between the contracting partners is huge. These market failures lead, according to Hansmann, either to regulation or emergence of mutual organizations. Smith (1986) also says that the mutual form has comparative advantages in lines which require little managerial discretion, for example in lines with good actuarial data, as in the life and health lines. There are however declining market shares in the US for life insurance companies, which Hansmann refers to as the difficulties for the mutuals to exploit the possibilities in the capital market and their minor success in diversification.

6.5 Concluding remarks

Arrow (1963) argues that virtually all special features of the medical care sector stem from the prevalence of uncertainty. The uncertainty calls for risk-bearing, but the non-existence of markets for bearing some risks has given rise to non-market social institutions as substitutes. The response of the non-existence of risk-bearing institutions is regulation of private or collective health insurers (as in the United States and Germany) or a completely socialization of the health care sector into a national health service (as in Great Britain and Sweden). It is possible, however, that these social institutions put obstacles in the way of an efficient risk-bearing market where insurance entrepreneurs would be able to exist.

Almost everyone, including Arrow, sees health risks isolated from other insurable risks. In this case, diversification in the health insurance can only be achieved through an increasing scale production. By allowing insurers to assemble a mix of different kinds of insurance lines and investment activities, the economies of scope will offset the need of scale production. Regarding social risk, an insurance entrepreneur can diversify health risks differently than a mutual. It is of great interest for an economy, to allow entrepreneurs to act freely in the market. Insurance entrepreneurs have recognized that the health insurance sector is part of a much broader insurance industry, which in turn is part of a much broader financial industry. Merging health insurance with other financial services, not only know-how about effective insuring is received, but also about efficient capital management.

Our analysis suggests that a competitive market for health insurance would benefit from economies of scope in addition to scale production.

References

Akerlof, G.A. (1970), "The Markets for 'Lemons': Quality Uncertainty and the Market Mechanism", *Quarterly Journal of Economics* 84, 488-500.

Arrow, K.J. (1953), "Le rôle des valeurs boursières pour la répartition la meilleure des risques" in Économétrie, 41-47, Paris, CNRS. Translated in English as, "The Role of Securities in the Optimal Allocation of Risk-Bearing", *Review of Economic Studies* 31 (1964), 91-96.

Arrow, K.J. (1963), "Uncertainty and the Welfare Economics of Medical Care", *American Economic Review* 53, 941-973.

Arrow, K.J. (1965), "Insurance, Risk and Ressource Allocation", reprinted in Arrow, K.J. (1970), *Essays in the Theory of Risk-Bearing*, New York, Elsevier Publishing Company, 134-143.

Arrow, K.J. (1970), *Essays in the Theory of Risk-Bearing*, New York, Elsevier Publishing Company.

Arrow, K.J. and Lind R.C. (1970), "Uncertainty and the Evaluation of Public Investments", *American Economic Review* 60, 364-378.

Beauchamp, T.L. and Childress, J.F. (1989), *Principles of Biomedical Ethics*, 3d ed., New York, Oxford University Press.

Berger, L.A. and Cummins, J.D. (1992), "Adverse Selection and Equilibrium in Liability Insurance Markets", *Journal of Risk and Uncertainty* 5, 273-288.

Bernoulli, D. (1738), "Specimen theoriae novae de mensura sortis", Commentarii academiae scientiarum imperialis Petropolitanae 5, St. Petersburg, 175-192. English translation by Sommer, L., "Exposition of a new Theory on the Measurement of Risk", *Econometrica* 22, (1954), 23-36.

Besley, T. (1989), "Public Provided Disaster Insurance for Health and the Control of Moral Hazard, *Journal of Public Economics* 39, 141-156.

Biger, N. & Y. Kahane (1978), Risk Considerations in Insurance Ratemaking, *Journal of Risk and Insurance* 45, 121-132.

Bjurulf P. et al. (1990), *Sjukskrivningar i förhållande till diagnos och yrke 1985-1987 i Östergötlands län*, Samhällsmedicinska institutionen, Hälsouniversitetet i Linköping.

Boorse, C. (1977), "Health as a Theoretical concept", *Philosophy of Science* 44, 542-573.

Borch, K.H. (1961), "The Utility Concept Applied to the Theory of Insurance", *Astin Bulletin* 1, 245-255.

Borch, K.H. (1962), "Equlibrium in a Reinsurance Makret", *Econometrica* 30, 424-444.

Borch, K.H. (1969), *Wirtschaftliches Verhalten bei Unsicherheit*, München, R. Oldenburg.

Borch, K. (1985), "A Theory of Insurance Premiums", *The Geneva Papers on Risk and Insurance* 10, 192-208.

Borch, K.H. (1990), *Economics of Insurance*, Amsterdam, North-Holland.

Breyer, F. and Zweifel, P. (1992), *Gesundheitsökonomie*, Berlin, Springer-Verlag.

Breyer, F. and Zweifel, P. (1995), *Health Economics*, (in press).

Briys, E. (1988) "On the Theory of Rational Insurance Purchasing in a Continuous Time Model", *Geneva Papers on Risk and Insurance* 13, 165-177.

Briys, E., Schlesinger, H. and Graf v.d. Schulenburg, J.-M. (1991), "Reliability of Risk Management: Market Insurance, Self-Insurance and Self-Protection Reconsidered", *Geneva Papers of Risk and Insurance Theory* 16, 45-58.

Briys, E. and Viala, P. (1995), "Optimal Insurance Design under Background Risk", Paper presented at the 22nd Seminar of the European Group of Risk and Insurance Economists, Geneva, Switzerland, September 18-20, 1995.

Brown, M.J. (1992), "Evidence of Adverse Selection in the Individual Health Insurance Market", *Journal of Risk and Insurance* 59, 13-33.

Brown, M.J. and Doerpinghaus, H.I. (1993), "Information Asymmetrics and Adverse Selection in the Market for Individual Medical Expense Insurance", *Journal of Risk and Insurance* 60, 300-312.

Camerer, C. and Kunreuther, H. (1993), "Making Decisions About Liability and Insurance: Editors' Comments", *Journal of Risk and Uncertainty* 7, 5-15.

Coase, R.H. (1937), "The Nature of the Firm", *Economica* 4, 386-405.

Cook, P.J. and Graham, D.A. (1977), "The Demand for Insurance Protection: The Case of Irreplacable Commodoties", *Quarterly Journal of Economics* 91, 143-156.

Cooper, R. and Hayes, B. (1987), "Multi-period Insurance Contracts", *International Journal of Industrial Organization* 5, 211-231.

Crocker, K.J. and Snow, A. (1986), "The Efficiency Effects of Categorical Discrimination in the Insurance Industry", *Journal of Political Economy* 94, 321-344.

Cummins, D.J. (1991), "Statistical and Financial Models of Insurance Pricing and the Insurance Firm", *The Journal of Risk and Insurance* 58:2, 260-302.

Cummins, J.D., Smith, B.D., Vance, N. and van Derhei J. (1983), *Risk Classification in the Life Insurance*, Boston, Kluwer Nijhoff Publishing.

Dahlby, B.G. (1981), "Adverse Selection and Pareto Improvements through Compulsory Insurance", *Public Choice*, 37, 547-558.

Dardoni, V. and Wagstaff, A. (1990), "Uncertainty and the Demand for Medical Care", *Journal of Health Economics* 6, 283-290.

Dean, K., Holst, E. and Wagner, M. (1983), "Self-Care of Common Illnesses in Denmark", *Medical Care* 21, 1012-1032.

Debreu, G. (1959), *Theory of Value*, New York, Wiley.

Dionne, G. (1983), "Adverse Selection and Repeated Insurance Contracts", *Geneva Papers on Risk and Insurance* 8, 316-333.

Dionne, G. and Harrington, S.E. (1992), *Foundations of Insurance Economics - Readings in Economics and Finance*, Boston, Kluwer Academic Publishers.

Doherty, N. and Garven, J.R. (1986), "Price Regulation in Property/Liability Insurance: A Contingent Claims Approach", *Journal of Finance* 41, 1031-1050.

Doherty, N.A. & H. Schlesinger (1990), Rational insurance purchasing, Consideration of contract nonperformance, *Quarterly journal of economics* 105, 243-253.

Dowie, J. (1975), "The Portfolio Approach to Health Behavior", *Social Science and Medicine* 9, 619-631.

Drèze, J.H. (1987), *Essays on Economics Decisions Under Uncertainty*, Cambridge, Cambridge Universty Press.

Eeckhoudt, L. and Kimball, M. (1992), "Background Risk, Prudence and the Demand for Insurance", in: Dionne, G. (ed), *Contributions to Insurance Economics*, Boston, Kluwer Academic Press.

Eckstein, Z., Eichenbaum, M. and Peled, D. (1985), "Uncertain Lifetimes and the Welfare Enhancing Properties of Annuity Markets and Social Security", *Journal of Public Economics* 26, 303-320.

Eggertsson, T. (1990), *Economic Behavior and Institutions*, Cambridge, Cambridge University Press.

Ehrlich, I. and Becker, G.S. (1972), "Market Insurance, Self-Insurance, and Self-Protection", *Journal of Political Economy* 80, 623-648.

Eisen, R. (1986), "Wettbewerb und Regulierung in der Versicherung: Die Rolle asymmetrischer Information", *Schweizerische Zeitschrift für Volkswirtschaft und Statistik* 122, 339-358.

Eisen, R., W. Müller & P. Zweifel (1993), "Entrepreneurial Insurance. A New Paradigm for deregulated Markets", *The Geneva Papers on Risk and Insurance* 18, 3-56.

Fairley, W. (1979), "Investment Income and Profit Margins in Property-Liability Insurance, Theory and Empirical Tests", *Bell Journal of Economics* 10, 192-210.

Goldberger, A.S. (1991), *A Course in Econometrics*, Cambridge, Cambridge University Press.

Grossman, M. (1972), *The Demand for Health: A Theoretical and Empirical Investigation*, National Bureau of Economic Research, Occasional Paper 119, New York.

Hansmann, H. (1985), "The Organization of Insurance Companies. Mutual versus Stock", *Journal of Law, Economics and Organization* 1, 125-153.

Hayek, F.A. (1960), *The Constitution of Liberty*, London Routledge and Kegan Paul.

Hedvall, M.-B. (1994), *The Process of Self-Care Decision Making*, Helsingfors, Swedish School of Economics and Business Administration.

Hill, R.D. (1979), Profit regulation in property liability insurance, *Bell Journal of Economics* 10, 172-191.

Hill, R.D. and F. Modigliani (1987), "The Massachusetts model of profit regulation in non-life insurance, An appraisal and extensions", in: Cummins, J.D. and Harrington, S.E., *Fair Rate of Return in Property-Liability Insurance*, Kluwer-Nijhoff Publishing.

Hirshleifer, J. and Riley, J. (1992), *The Analytics of Uncertainty and Information*, Cambridge, Cambridge University Press.

Horgby, P.-J. (1993), "Sjukvård och rättvisa inom hälso- och sjukvården", in: Perselli, J. (ed), "Sju doktorander. Essäer om hälsa, livskvalitet, vård och vårdbehov", Meddelande 1993:1, ISSN 0284-5725, Department of Health and Society, Linköping University.

Horgby, P.-J. (1995), "Sjukvårdsförsäkringar på en privat marknad - en teoretisk analys av informationsasymmetri", *Scandinavian Insurance Quarterly* 1/95, 20-26.

Hu, T.-W., Ren, Q.-F., Keeler, T.E. and Bartlett, J., (1995), "The Demand for Cigarettes in California and Behavioral Risk Factors", *Health Economics* 4, 7-24.

Iannaccone, L. R. (1995), "Risk, Rationality, and Religious Portfolios", *Economic Inquiry* 33, 285-295.

Karni, E. (1985), *Decision Making Under Uncertainty*, Cambridge, Harvard University Press.

Kenkel, D.S. (1995), "Should You Eat Breakfast? Estimates from Health Production Functions", *Health Economics* 4, 25-30.

Knight, F.H. (1921), *Risk, Uncertainty and Profit*, Hougton Mifflin.

Kunreuther, H. et al. (1995), "Ambiguity and Underwriter Decision Processes", *Journal of Economic Behavior and Organization*" 26, 337-352.

Laffont, J.-J. (1989), *The Economics of Uncertainty and Information*, The MIT Press.

Lindbeck, A. (1994), "Uncertainty under the Welfare State - Policy-induced Risk", *Geneva Papers on Risk and Insurance* 19, 379-393.

Lintner, J. (1965), The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets", *Review of economics and statistics* 47, 13-37.

Liss, P.-E. (1993), *Health Care Need*, Aldershot: Avebury.

Machina, M.J. (1987), "Choice Under Uncertainty: Problems Solved and Unsolved", *Journal of Economic Perspectives* 1, 121-154.

Machina, M.J. (1995), "Non-Expected Utility and the Robustness of the Classical Insurance Paradigm", *Geneva Papers on Risk and Insurance Theory* 20, 9-50.

Manning, W.G. et al. (1987), "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment", *American Economic Riview* 77, 251-277.

Markowitz, H.M., (1952), "Portfolio Selection", *Journal of Finance* 7, 77-91.

Markowitz, H.M. (1959), *Portfolio Selection: Efficient Diversification of Investments*, New York, Wiley.

Marquis, M.S. (1992), "Adverse Selection with Multiple Choice Among Health Insurance Plans: A Simulation Analysis, *Journal of Health Economics* 11, 129-151.

Marquis, M.S. and Phelps, C.E. (1987), "Price Elasticity and Adverse Selection in the Demand for Supplementary Health Insurance", *Economic Inquiry* 25, 299-313.

Mayers, D. and Smith, C.W. (1983), "The Interdependance of Individual Portfolio Decisions and the Demand for Insurance", *Journal of Political Economy* 91, 304-311.

Meyer, J. (1987), "Two-Moment Decision Models and Expected Utility Maximization", *American Economic Review* 77, 421-430.

Mossin, J. (1966), "Equilibrium in a Capital Asset Market", *Econometrica* 35, 768-783.

Mossin, J. (1968), "Aspects of Rational Insurance Purchasing", *Journal of Political Economy* 79, 553-568.

Muurinen, J.-M. (1982), "Demand for Health: A Generalized Grossman Model", *Journal of Health Economics* 1, 5-28.

Nordenfelt, L. (1993), *Quality of Life, Health and Happiness*, Aldershot, Avebury.

Neuhouse, J., Phelps, C.E. and Marquis, M.S. (1980), "On Having Your Cake and Eating It Too: Econometric Problems in Estimating the Demand for Health Services", *Journal of Econometrics* 11, 117-129.

Neuman, J.,v. and Morgenstern, O. (1947), *Theory of Games and Economic Behavior*, Princeton University Press.

Nordquist, G. and Wu, S., (1976), "The Joint Demand for Health Insurance and Preventive Medicine", in: Rosett, R.N., (ed), *The Role of Health Insurance in the Health Service Sector*, Universities-National Bureau Conference Series 27, New York, National Bureau of Economic Research.

Pauly, M.V. (1974), "Overinsurance and Public Provision of Insurance: The Role of Moral Hazard and Adverse Selection", *Quarterly Journal of Economics* 88, 44-62.

Phelps, C.E. and Newhouse, J.P (1974), "Coinsurance, the Price of Time, and the Demand for Medical Services", *Review of Economics and Statistics* 56, 334-342.

Philipson, T.J. and Posner, R.A. (1995), "A Theoretical and Empirical Investigation of the Effects of Public Health Subsidies for STD Testing", *Quarterly Journal of Economics* 445-474.

Persson, Å. (1993), "Information and the Demand for Health Care", Paper presented at the 14th Nordic Health Economists' Study Group Meeting, Tromsø, Norway, August 1993.

Pool, W. (1991), Insurance and the European Community", *The Geneva papers of risk and insurance* 17, 178-199.

Price, J.R. and Mays, J.W. (1985), "Biased Selection in the Federal Employees Health Benefits Program", *Inquiry* 22, 67-77.

Rawls, J. (1971), *A Theory of Justice*, Cambridge.

Riley, G., Lubitz, J. Prihoda, R. and Stevenson, M.-A. (1986), "Changes in Distribution of Medicare Expenditures Among Aged Enrollees, 1969-1982", *Health Care Financing Review*, Spring 1986.

Rothschild, M. and Stiglitz, J.E. (1970), "Increasing Risk: A Definition", *Journal of Economic Theory*, 2, 225-243.

Rothschild, M. and Stiglitz, J.E. (1976), "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information", *Quarterly Journal of Economics* 90, 629-650.

Sandmo A. (1971), "On the Theory of the Competitive Firm under Price Uncertainty", *American Economic Review* 61, 65-73.

Sharpe, W.S. (1964),"Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk", *Journal of Finance* 19, 425-442.

Sharpe, W.S. (1991), "Capital Asset Prices With and Without Negative Holdings", *Journal of Finance* 46, 489-509.

Schlee, E.E. (1995), "The Comparative Statics of Deductible Insurance in Expected- and Non-Expected-Utility Theorems", *Geneva Papers of Risk and Insurance Theory* 20, 57-72.

Schlesinger, H. (1981), "The Optimal Level of Deductibility in Insurance Contracts", *Journal of Risk and Insurance* 48, 465-481.

Schlesinger, H. (1994), "Zur Theorie der Versicherungsnachfrage", *Zeitschrift für die gesamte Versicherungswissenschaft*, 1/2, 113-136.

Schlesinger, H. and Doherty, N.A. (1985), "Incomplete Markets for Insurance: An Overview", *Journal of Risk and Insurance* 52, 402-423.

Schlesinger, H. & J.-M. Graf v.d. Schulenburg, (1987), Risk Aversion and the Purchase of Risky Insurance, *Zeitschrift für Nationalökonomie* 47, 309-314.

Schulenburg, J.-M. Graf v.d. (1987), *Selbstbeteiligung. Theoretische und empirische Konzepte für die Analyse ihrer Allokations- und Verteilungswirkungen*, Tübingen, J.C.B. Mohr.

Schulenburg, J.-M. Graf v.d. (1994), "Grundsicherung durch Gesetzliche Krankenversicherung, Nachfrage nach ergänzenden Krankenversicherungssystemen - einige theoretische Grundlagen", in: Schwebler, R. (ed.), *Dieter Farny und die Versicherungswissenschaft*, Karlsruhe.

Shavell, S. (1979), "On Moral Hazard and Insurance", *Quarterly Journal of Economics* 93, 541-562.

Segal, U. and Spivak, A. (1990), "First Order versus Second Order Risk Aversion", *Journal of Economic Theory* 51, 111-125.

Selden, T.M. (1993), "Should the Government Provide Catastrophic Insurance?", *Journal of Public Economics* 51, 241-247.

Siegel, C. et al. (1992), "A risk-based prospective payment system that integrates patient, hospital and national costs", *Journal of Health Economics* 11, 1-41.

Sinn, H.W. (1980), *Ökonomische Entscheidungen bei Ungewissheit*, Tübingen, J.C.B. Mohr.

Sinn, H.W. (1983), *Economic Decisions under Uncertainty*, Amsterdam, Nort-Holland Publishing Company.

Skogh, G. (1989), "The Transactions Cost Theory of Insurance, Contracting Impediments and Costs", *The Journal of Risk and Insurance* 56, 727-732.

Smith, V.L. (1968), "Optimal Insurance Coverage", *Journal of Political Economy* 79, 68-77.

Smith, C.W., Jr. (1986), "On the Covergence of Insurance and Finance Research", *Journal of Risk and Insurance* 53, 693-717.

Söderström, L. (1992), "A Case for Patient Charges?", SNS Working Paper, Stockholm, SNS Förlag.

Söderström, H. Tson. (ed) (1994): *Välfärdsland i ofärdstid*, Konjunkturrådets rapport, SNS-förlag.

Strassl, W. (1988), *Externe Effekte auf Versicherungsmärkten*, Tübingen, J.C.B. Mohr.

Tapiero, C., Kahane, Y. and Jacque, L. (1986), "Insurance Premiums and Default Risk in Mutual Insurance", *Scandinavian Actuarial Journal*, 82-97.

Tobin, J. (1965), "The Theory of Portfolio Selection", in: Hahn, F.H. and Brechling, F.P.R. (eds), *The Theory of Interest Rates. Proceedings of a Conference Held by the International Economic Association*, New York, 3-51.

Van de Ven, W.P.M.M. and van Vliet, R.C.J.A. (1995), "Consumer Information Surplus and Adverse Selection in Competitive Health Insurance Markets", *Journal of Health Economics* 14, 149-170.

Walras, L. (1874), Translated in English 1954, *Elements of Pure Economics*, Richard D. Irwin.

Wagstaff, A. (1986), The Demand for Health: Some Empirical Evidence, *Journal of Health Economics* 5, 195-233.

Wagstaff, A. (1993), "The Demand for Health: An Empirical Reformulation of the Grossman Model", *Health Economics* 2, 189-198.

Weisbrod, B.A. (1991), "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment", *Journal of Economic Literature* 29, 523-552.

WHO (World Health Organization) (1946), *Constitution. WHO*, New York.

WHO (World Health Organization) (1985), *Targets for health for all.* WHO, Copenhagen.

Wilson, C. (1977), "A Model of Insurance Markets with Incomplete Information", *Journal of Economic Theory*, 12, 167-207.

Zweifel, P. (1992), *Bonus Options in Health Insurance*, Dordrecht, Kluwer Academic Publishers.

Zweifel, P. (1992), "The Production of Health: A Stochastic Formulation", Paper presented at the Colloque Européen, Paris, 16-18 December 1992. A modified version can be found in Breyer, F. and P. Zweifel, 1992, *Gesundheitsökonomie*, Berlin, Springer-Verlag.

Prior reports published by CEFOS:

Behovsbudgetering. Nya budgetprinciper i kommuner. Cecilia Bokenstrand. (Rapport 1 1993)

Självständighet eller statsbundenhet. Den kommunideologiska idédebatten 1962-1974. Urban Strandberg. (Rapport 2 1995)

Finansieringsanalysens dimensioner. Teori och kostnader. Pär Falkman och Stefan Pauli. (Rapport 3 1995)

Utveckling och tillämpning av ekonomistyrning i en decentraliserad kommunal organisation. Problem och möjligheter vid förändringsarbete - en fallstudie om ekonomistyrning. Lennart Jansson. (Rapport 4 1995)