Four Essays on Technological and Organizational Change in Health Care

Inaugural-Dissertation
zur Erlangung des Grades

Doctor oeconomiae publicae (Dr. oec. publ.)

im Jahr 2004
an der Ludwig-Maximilians-Universität München
vorgelegt von

Astrid Selder

Referent: Prof. Ray Rees
Korreferent: Prof. Dr. Andreas Haufler

Promotionsabschlussberatung: 21. Juli 2004

Acknowledgements

This doctoral thesis was written under the supervision of Professor Ray Rees at the University of Munich from 2000 to 2004. I would like to thank Professor Rees for the encouragement and scientific advice which he provided during my studies in Munich. Working at his chair was an extremely pleasant experience due to an excellent work atmosphere and a low non-scientific work load.

I would also like to thank Professor Achim Wambach who was the first to encourage me to pursue this doctoral thesis. By offering me to co-author the essay on willpower and information acquisition (chapter 5) he allowed me to learn from his experience and excellence in research.

I am indebted to my colleagues Tobias Böhm, Florian Englmaier, Ekkehard Kessner, Andreas Knaus and Ingrid Königbauer, who supported my work with helpful comments and suggestions, and with daily and sometimes hourly doses of tea and chocolate.

I would also like to thank Irmgard von der Herberg for handling most of the administrative duties which allowed me to concentrate on my scientific work. Our numerous student helpers provided excellent support in scientific and administrative matters. I would like to mention in particular Christa Dallat and Brigitte Stieghorst who ran our library in an independent and efficient way, thereby relieving me of most supervision duties.

The first study (chapter 2) of this thesis was in part conducted during a research visit to the Centre for Health Economics, University of York, and with financial support from Deutscher Akademischer Austauschdienst (DAAD) grant no. D/02/08137. I am very grateful to both institutions. During my stay in York, Professor Hugh Gravelle and Michael Kuhn accompanied my research project with very helpful advice.

Furthermore, I want to thank Professor Dr. Andreas Haufler and Professor Dr. Reiner Leidl, who kindly agreed on joining the board of referees for this thesis.

Finally, I would like to thank my family and friends who supported me during the last years in my non-academic life and who thereby made a great contribution to this work.

March 2004 Astrid Selder

Contents

1 Introduction

This thesis contains four essays on technological and organizational change in health care. Although each chapter can be read independently, there is a central thread resumed in each chapter. The interaction of patient, physician and insurer in the health care market is analysed in different contexts. Chapter 2 studies the physician's decision to adopt new technologies in a situation where patients have control over the amount of health care consumed. This setup of ex-post moral hazard is complemented in chapter 3 by an analysis of supplier-induced demand. Providers' technology choice and adoption of innovations here depend on the profitability of treatment and patients' willingness to consent. A conclusion which may be drawn from these sections is a call for a more integrated provision of health care in the form of managed care. Chapter 4 contains an analysis which explains why physicians, particularly in Germany, are often very much opposed to such new organizational forms of provision of care. In chapter 5, it is shown that patients may not always want to make use of the opportunities offered by technological progress. Information on one's health status which may be acquired, for instance, by genetic tests, may be declined by individuals who fear a breakdown of will. This finding has important implications for the disclosure of information by physicians and for the information acquisition policy of insurers.

In the following, an extended summary of each essay is provided. A motivation of the research question under consideration and a discussion of the related literature follow in each chapter individually. Chapter 6 contains a brief conclusion of this thesis.

In chapter 2, the link between Physician Reimbursement and Technology Adoption is analysed. The importance of technological change in the health sector has been a widely discussed topic in the economic literature, but an important player in the health care market has so far been neglected: the provider of health care services. A better understanding of the incentives for technology adoption implied by different reimbursement schemes is crucial for the state to induce welfare improving behaviour by providers. This study combines two branches of the literature: Baumgardner (1991) describes a technology by three parameters: monetary costs of treatment, non-monetary costs of treatment to the patient, and the level of healing which is possible using this specific technology. He asks which types of technological advances are welfare improving in a world without and with ex-post moral hazard. The second line of the literature deals with the interaction of physician, patient, and insurer (Ellis and McGuire (1990, 1993), Ma and McGuire (1997)). This literature combines the phenomenon of ex-post moral hazard with strategic actions of the physician. In the current setup, the provider decides on the technology which he adopts and offers to the patient, and afterwards the patient decides on the amount of health care consumed.

Assume that at first technology can be contracted upon, but the treatment intensity is chosen by the patient. Under second best insurance coverage, decreases in monetary and non-monetary costs are welfare improving, whereas an increase in the technologically feasible boundary of healing is socially desirable only if this boundary does not limit overconsumption. Building on these results, the provider's incentives for adopting new technologies under cost sharing, cost reimbursement, and fee-for-service systems are analysed. The adoption of welfare enhancing technologies can only be achieved in very special cases by the simple remuneration schemes under consideration. This is the case if the physician receives a fee for service and the patient has to bear a positive or only weakly negative copayment. In this setting, the physician has an incentive to reduce monetary and non-monetary costs to the patient. An increase in the feasible boundary of healing has to be desirable, which is most likely for extremely severe illnesses. In general, though, social welfare will be lower than in the second best case.

Next, the welfare effects of technological change in a world with a given standard coinsurance contract which is not second best efficient are studied. Effects of changes in the technological boundary stay the same as in the second best case. But increases in costs may now be welfare improving if the patient's demand for health care responds to such increases in a very elastic way. Then, the subsequent reduction in premiums offsets the negative effects of increased treatment costs. One can conclude that changing from cost reimbursement to cost sharing drives marginal monetary costs of treatment down only if demand responds inelastically with respect to costs. At the same time, non-monetary costs are driven up and the technological boundary of treatment is likely to be reduced. Only incentives with respect to monetary costs are a clear welfare improvement.

Finally, the optimal choice of reimbursement schemes can be analysed. Technologies should be classified in subgroups which are reimbursed differently. If the technical boundary is a binding constraint and welfare increases if it is shifted outwards then a fee-for-service system is optimal. If the boundary is binding but increases will reduce social welfare then cost sharing should be implemented. If the technical boundary is not binding then the optimal reimbursement mechanism depends on the demand elasticities with

respect to monetary and non-monetary costs.
Earlier versions of this study were presented in seminars at the University of Munich and at the University of York, at the Spring Meeting of Young Economists 2002 and at the Fourth European Conference on Health Economics 2002. It benefited greatly from discussions with Florian Englmaier, Hugh Gravelle, Sebastian Kessing, Ingrid Königbauer, Michael Kuhn, Ray Rees, and Achim Wambach.

Demand Inducement for Health Care Technologies is studied in chapter 3. The analysis complements the preceding chapter by investigating into providers' technology choice and adoption of innovations in a situation where patients can observe the investment decision and treatment recommendation but not their own health status. Once more, Baumgardner's (1991) parameterization of technologies is taken up, but used to model supplier-induced demand for treatment in the line of Dranove (1988). The physician gains an informational advantage by diagnosing the patient. He can to a certain extent induce demand for treatment which the patient accepts but would not want if he was fully informed. The patient is restricted to a "consent" or "no consent" decision to a treatment recommendation. Taking the physician's behaviour into account the patient will only consent to treatment if his symptoms are severe enough. He may be willing to consent to certain treatment technologies but not to others which are more costly or less effective. Therefore, the provider will offer different technologies to different ranges of patients. His offer will be led by profitability considerations and patients' willingness to consent.
If technology choice can be contracted upon then the first best technology choice and treatment pattern can be achieved. The physician receives a fee-for-visit and zero-profit reimbursement for treatment. If technology choice cannot be contracted upon the physician only has an incentive to offer different types of treatments if treatment is profitable for him. The patients with most severe symptoms will then be offered the most profitable technology irrespective of their true health status. The patients whose symptoms are a bit less severe and who would not accept the most profitable technology are offered a less profitable one which they are willing to accept, and so on. If there exist zero-profit technologies, an intermediate range of patients receives the correct amount of treatment: These patients are only offered treatment if they really need it. The fraction of patients with the least severe symptoms may decide not to visit the physician if diagnosis is costly for them or if they fear overtreatment. Optimal reimbursement of technologies induces the

first best selection of technologies by the adequate ranking of profitable reimbursement. Yet the first best treatment pattern is not achievable because patients are overtreated according to their symptoms instead of receiving the correct treatment according to their true health status.

When an innovation comes on the market the physician adopts it if it is more profitable and/or if it is accepted by more patients than the incumbent technology. If only one technology parameter changes (which may be the case for many process innovations) then the physician adopts the innovation given equal profitability because it is accepted by a wider range of patients. If more than one technology parameter changes this result may no longer hold. Optimal reimbursement then requires adjusting the adequate ranking of profitable reimbursement.

The adoption of an innovation implies ambiguous effects on social welfare. A share of patients gains from better treatment, but another fraction of patients loses because they are now willing to accept being treated with this new technology which is more costly or less effective than the one they would have got previously. The state can counteract these welfare losses by introducing technology specific copayments, which limit patients' willingness to consent. This prevents the physician from offering these treatments inadequately.

I would like to thank Tobias Böhm, Florian Englmaier, Ingrid Königbauer, Michael Kuhn, Tom van Ourti, John Posnett, Ray Rees, Achim Wambach, and the participants in several seminars at the University of Munich, at the HESG Winter Meeting 2003, at the Spring Meeting of Young Economists 2003, at the 4th iHEA World Congress 2003, and at the Jahrestagung des Vereins für Socialpolitik 2003 for helpful comments on this chapter.

A more integrated provision of health care in the form of managed care may solve some of the problems discussed in chapters 2 and 3. Yet providers of health care services are often opposed to such organizational changes. Chapter 4 analyses MEDICAL ASSOCIATIONS, MEDICAL EDUCATION, AND TRAINING ON THE JOB and thereby provides a rationale for this observation.

It deals with the question whether the state (resp. the health authority) should set standards for medical education and medical training or delegate the implementation of these standards to a medical association. The study combines Shapiro's (1986) approach of input regulation with the credence goods characteristics of health care services. Physicians in the first period of their working life receive a license if they meet a specified standard of medical education. A higher standard increases the probability of adequate treatment. In later periods, training on the job is necessary to keep up this probability of correct treatment. Patients are randomly matched with physicians and experience a disutility from inadequate treatment.

If the health authority sets standards for medical education and training on the job, these standards are chosen such that the expected utility of a representative patient is maximized under the constraint that a physician is reimbursed for his expenses in each period. This situation is compared to a setup in which physicians are organized in a medical association. This association is assumed to control market entry and to set itself standards for education and training. Since older generations of physicians constitute a majority within the medical association, this institution maximizes the lifetime income of older generations of physicians. Consequently, older generations of physicians extract rents from the younger generation. If the medical association is allowed to pay out different prices to different generations of physicians, then a system of seniority wages is set up. Standards for education and training do not differ from the situation in which these standards are set by the health authority because the efficient choice allows the biggest surplus to be distributed. If on the other hand, the medical association is bound to a fee-for-service system it pays out the same price for treatment irrespective of the age of the physician. Then rents can only be transferred from the younger to the older generation of physicians by imposing higher standards of education on the young generation, thereby reducing the need for training on the job. The German outpatient sector provides evidence for these results.

For both setups one can conclude that the rents accrue only to the older generation of physicians who set up the system of medical associations. Still, switching back to individual contracting is rejected by the majority of physicians in later periods because they want to recover the investment in education which they have made in earlier periods of their working life.

I would like to thank Tobias Böhm, Florian Englmaier, Ingrid Königbauer, Sonja Ossig, and Ray Rees for helpful comments on this chapter.

Sustaining Will by Staying Ignorant is the title of the last chapter which is joint work with Achim Wambach. In this paper we discuss the decision whether or not to obtain information on one's own health status, e.g. in the form of genetic testing, when individuals discount the future hyperbolically. The basic intuition for our results is the following: People who have time-inconsistent preferences and who are aware of this fact, might devise strategies to overcome the resulting inefficiencies. The psychologists Ainslie and Haslam (1992) mention personal behavioural rules as one means of making one's behaviour more consistent. Such a behavioural rule can be modelled as a trigger strategy in a dynamic game: Stick to the long term equilibrium path, otherwise you will be punished in future periods by always making the short run optimal choice. Based on this definition of will and given the analogy with dynamic games, it becomes clear that will can only be sustained if the future is valued sufficiently highly. This fact may have an impact on the acquisition of information about the value of future periods.

Consider tests on your health status, for example genetic tests. A person who is offered the possibility to undertake such a test has the opportunity to get better information on her illness risk, so she can plan her future life accordingly. On the other hand, a positive test result may indicate a very large illness probability. Thus, the strategies devised to overcome the problem of time-inconsistency might no longer be stable. If this latter effect dominates the first, then it can indeed be rational not to undertake the test. The latter effect only sets in if the consequences of the negative information are sufficiently severe, otherwise willpower can be sustained. Interestingly, if the consequences are extremely bad, the former effect might dominate the latter. If someone is told that she will live for only one more period, breakdown of will does not matter any more, as there are no future periods. Thus there is a non-monotonicity in the value of information: If the illness is light or extremely severe, information is positively valued. For illnesses which are severe, but not extremely so, information has a negative value.

Apart from this non-monotonicity in the value of information, we obtain a third claim from our model. If before the acquisition of information the health status is already quite bad, people are more willing to obtain additional information. This may happen because even before taking a test a breakdown of will is imminent. Then more information is likely to have a positive value. We provide evidence for the implications of the theory based on studies reported in the literature and on our own investigations.

An earlier version of this chapter was published as "Breakdown of will and the value of information", CEPR Discussion Paper 3111. We thank Ray Rees, Klaus M. Schmidt, David Laibson, Friedrich Breyer, and the seminar participants at the Jahrestagung des Vereins für Socialpolitik 2000, Berlin, at the Tinbergen Institute, Amsterdam, and the participants in SET 2001, Venice International University, for helpful comments.

2 Physician Reimbursement and Technology Adoption

2.1 Introduction

The importance of technological change in the health sector has been a widely discussed topic in the economic literature[1]. Especially the discussion of the reasons for expenditure increases and cost explosion has recognized the key role of technical advances. The existing theoretical literature in the field has so far focused on welfare implications of new technologies under various insurance schemes. Yet one important player in the health care market has so far been neglected in the discussion: the provider of health care services. After all, it is the physician (or the hospital) who decides whether a technical innovation is adopted and offered to the patient or not. The technology choice and treatment recommendation depend on the financial incentives associated. A better understanding of the incentives implied by different reimbursement schemes is crucial for the state to induce welfare improving behaviour by providers. This study builds on the existing literature and analyses the incentives of health care providers to adopt new technologies.

A technology is described by three parameters: monetary costs of treatment, non-monetary costs of treatment to the patient, and the level of healing which is possible using this specific technology. First, the results of the existing literature are replicated when asking which types of technological advances are welfare improving in a world without and with ex-post moral hazard. In the first best case both the provider's technology choice and the treatment intensity chosen by the patient are contractible. In this situation, decreases in monetary and non-monetary costs are welfare improving, whereas an increase in the technologically feasible boundary of healing is socially desirable. In the second best case technology still can be contracted upon, but the treatment intensity is chosen by the patient and cannot be determined by the state any longer. Welfare effects in this setting of ex-post moral hazard remain the same except for the technological boundary. If this boundary helps to limit overconsumption then an increase would induce even more overconsumption and in this case reduce social welfare. These results have already been derived by Baumgardner (1991). Building on them, the providers' incentives for adopting new technologies are investigated for various reimbursement schemes. Cost sharing,

[1] See for example Fuchs (1996), Hillman (1986) and Weisbrod (1991).

cost reimbursement, and fee-for-service systems are analysed. The welfare implications of these schemes in a world which is second best efficient with respect to insurance coverage are summarized in Proposition 1. It turns out that only in very special cases can the adoption of second-best efficient technologies be achieved by a simple remuneration scheme. This is the case if the physician receives a fee for service and the patient has to bear a positive or only weakly negative copayment. In this setting, the physician has an incentive to reduce monetary and non-monetary costs to the patient. Finally, an increase in the feasible boundary of healing has to be desirable, which is most likely for extremely severe illnesses. In general, though, social welfare will be lower than in the second best case.

An important contribution of the analysis concerns the ongoing shift towards capitation systems for reasons of cost savings. In order to study the welfare effects of these changes, one first has to analyse the effects of technology parameter changes in a world with a given standard coinsurance contract which is not second best efficient. Effects of changes in the technological boundary stay the same as in the second best case. But increases in costs may now be welfare improving if the patient's demand for health care responds to such increases in a very elastic way. Then, the subsequent reduction in premiums offsets the negative effects of increased treatment costs. Proposition 2 summarizes the corresponding welfare implications of the above mentioned reimbursement schemes in this situation. One can conclude that changing from cost reimbursement to capitation drives monetary costs down only if demand responds inelastically with respect to costs. At the same time, non-monetary costs are driven up and the technological boundary of treatment is likely to be reduced. Only incentives with respect to monetary costs are a clear welfare improvement.

Finally, the optimal choice of reimbursement schemes can be analysed. Technologies should be classified in subgroups which are reimbursed differently. If the technical boundary is a binding constraint and welfare increases if it is shifted outwards then a fee-for-service system is optimal. If the boundary is binding but increases will reduce social welfare then cost sharing should be implemented. If the technical boundary is not binding then the optimal reimbursement mechanism depends on the demand elasticities with respect to monetary and non-monetary costs.

In the empirical economic literature, there are a few studies which present direct evidence for the impact of technological change on cost growth. Some studies show at an accounting level that static supply and demand factors can only explain less than half of the growth of medical spending. They attribute the residual to technological change. Newhouse (1992) analyses health care expenditure growth per capita in the US from 1929 to 1990. He finds that aging, increased insurance, increased income, supplier-induced demand, and factor productivity changes account for less than 50% of the increase in expenditure. Growth rates of health care expenditure between 4 and 12.2% exceeded by far the growth rates of real GNP per capita in that period. Newhouse believes that the bulk of the residual increase which cannot be explained by the above factors is attributable to technological change. This kind of analysis cannot provide an explanation, though, for what actually drives technological change. It merely states its important role as one of the determinants of cost increases in the health sector.

One of the few studies which addresses this issue more closely is the study by Cutler and McClellan (1996). The authors examine the sources of expenditure growth in heart attack treatment. They first show that essentially all of the cost growth is a result of the diffusion of particular intensive technologies. Then they distinguish six factors that may influence technology diffusion: organizational factors within hospitals, the insurance environment in which technology is reimbursed, public policy regulating new technology, malpractice concerns, competitive or cooperative interactions among providers, and demographic composition. The authors conclude that insurance variables, technology regulation, and provider interactions have the largest quantitative effect on technological diffusion[2]. These factors affect both technology acquisition and the frequency of technology use. The study thus shows that in a micro-level analysis the impact of technological change on cost increases is probably even larger than suggested by the studies based on residual analysis.

The existing theoretical literature in the field does not explain the diffusion of new technologies nor look at the impact of regulation and the role of providers of health care. There are only a few studies which try to model technical advances in the health sector. They focus on the role of insurance coverage for welfare effects of technological change. Goddeeris (1984a, 1984b) derives conditions under which a costly technological change that increases capabilities increases or decreases welfare. Each technology is defined by a "healing function" which depends on health care expenditures. A technical advance

[2] This is in line with an early study by Coleman et al. (1957) who looked at the role of physician networks for technology diffusion.

changes this functional form such that a higher level of healing is possible using the same financial resources. This special structure does not allow differentiation between various types of technological advances. One could for instance imagine a new technology which uses the same financial resources for producing the same healing level as the old standard technology but which causes less suffering to the patient during the treatment. Certainly, we would want to capture this as a real "advance" from a welfare point of view. Baumgardner (1991) addresses this issue by describing a technology by three parameters: monetary costs of treatment, non-monetary costs, and the technical boundary up to which healing is possible using this technology. He analyses welfare effects of technological change under different insurance systems, but the type of technical advance can now be specified more clearly. Each parameter change can be analysed separately, which allows a better qualitative and quantitative measuring of technological change.

Differentiating between these separate effects is an important contribution of Baumgardner's study. We learn which types of technological advances are welfare improving under full insurance, under co-insurance, and in a special HMO system. Nevertheless, neither of these studies has yet explained which types of new technologies will actually be used in the health care market. We still cannot tell whether welfare improving or welfare decreasing technologies will be adopted. One should not neglect the role of an important player in the health care market: the role of the provider of health services, i. e. the physician or the hospital. It is the providers who actually are in the position to adopt a new technology and to offer the corresponding treatment to the patient. In fact, even if the state was able to classify new technologies according to Baumgardner's suggestions it would still have to consider the physicians' incentives to adopt these technologies. Therefore, if we want to answer the question of which innovations will experience a wider diffusion in the health care market, and how we can implement the diffusion of welfare improving technologies, we have to consider the incentives of physicians and hospitals to adopt these technologies. Analysing providers' incentives is the first respect in which I have extended Baumgardner's analysis. Secondly, I not only look at a world with second best efficient insurance contracts but also analyse the welfare implications of technology parameter changes in a situation which seems to fit the real world more closely. In the second part of the paper I study welfare effects assuming a standard coinsurance contract which is in general not second best efficient but accounts for truthful reporting requirements[3]. Although the re-

[3]See Ma and McGuire (1997).

sults become more ambiguous, the more descriptive approach appears be a necessary step for deriving policy implications.

I follow in my analysis the lines of the literature on the interaction of physician, patient, and insurer (see, for instance, Ellis and McGuire (1990, 1993) or Ma and McGuire (1997)). This literature combines the phenomenon of ex-post moral hazard with strategic actions of the physician. Although information is complete, the provider's action is not contractible and induces a demand response by the patient[4]. In my setting, the physician chooses a technology and offers this technology to the patient. The patient then chooses the treatment intensity which maximizes his utility given the technology offered[5]. Taking these actions into account, the insurer (or social planner) designs a remuneration scheme for the physician and an insurance contract for the patient. He cannot contract upon technology choice and treatment intensity. Although it seems reasonable that the state can indeed prohibit technologies which are obviously welfare decreasing, it will still leave some choice to the physician because the physician may be better informed about innovations entering the market or about the patient's needs. The assumption that the patient can choose the amount of care he consumes may seem stronger. It is in line, though, with the literature on ex-post moral hazard[6] and it seems to be a reasonable assumption for the analysis of certain institutions and certain types of treatment. As patients are increasingly better informed about health care products and treatments this demand response may even be strengthened in the future. The current setup fits any situation where the patient has some influence on the amount of treatment he receives. Still, as also pointed out in section 2.5, there is scope for future research relaxing this assumption.

The rest of the chapter is structured as follows: In section 2.2, the basic model in the first best situation is set up and the optimal contracts for the patient and for the physician are derived. Then, the welfare effects of changes in technology parameters are analysed. In section 2.3, ex-post moral hazard is modelled by assuming that the treatment intensity

[4]See McGuire (2000) for a survey on this literature. I focus on the case of a single provider's decision without taking into account strategic interaction of providers. Kesteloot and Voet (1998), for example, study cooperative vs. non-cooperative outcomes when hospitals compete in quality and price. The dimension of competition of providers with respect to technology adoption clearly leaves room for future research but goes beyond the scope of this paper.

[5]I will use the terminology of "treatment intensity" and "amount of health care" in an equivalent way. The technical boundary of healing can be thought of as an intensity boundary.

[6]See Ma and McGuire (1997) for a discussion of "demand response".

is chosen by the patient. Second best efficient contracts are derived and welfare effects of parameter changes through new technologies are established. Then, the incentives of providers for adopting new technologies are studied and analysed with respect to their welfare implications. Section 2.4 repeats the analysis for a more realistic standard coinsurance contract with a fixed copayment scheme. Section 2.5 concludes and considers topics for further research.

2.2 Model setup in a First Best world

I analyse a five stage game. In the first stage, the insurer offers contracts to the physician and to the patient. The insurance contract specifies a premium and a copayment rate, the remuneration contract for the physician specifies a fixed payment and a fee per unit of treatment. In the second stage the patient and the physician decide whether to accept the contracts. Then nature decides whether the patient falls ill or not. In case of illness, the physician chooses the treatment technology in the fourth stage. A technology is specified by three parameters, namely monetary costs per unit of treatment, non-monetary costs per unit of treatment, and the technical boundary of treatment. This boundary is defined as the maximal treatment intensity which can be chosen by the patient and it captures the point up to which healing is possible using this special technology. Finally in the fifth stage the patient chooses the treatment intensity which is of course less than or equal to the technical boundary.

In this section, the first best insurance contract for the patient and the first best remuneration scheme for the physician are derived, taking the technology as given. Then it is shown how variations of the technology parameters affect social welfare. From this we can conclude which types of new technologies the social planner would like to implement.

The social planner maximises the patient's expected utility subject to the zero-profit conditions for the insurer and for the physician[7]. The optimal amount of treatment may be constrained by the technologically feasible boundary B. The planner's problem

[7]This approach implies the assumptions of either competitive markets for both physicians and insurers or a state insurer who does not want to make profits and who extracts the physicians' rents.

therefore is given by

$$\begin{aligned}
&\max_{m,P,a,R,p} (1-\pi)U(Y-P) + \pi U(Y-P-a\cdot c\cdot m-\epsilon+G(m)-n\cdot m) \\
&s.t. \\
&P+\pi\cdot a\cdot c\cdot m-\pi R-\pi\cdot p\cdot m\geq 0 \qquad \text{(zero-profit condition insurer)} \\
&R+(p-c)\cdot m\geq 0 \qquad \text{(zero-profit condition physician)} \\
&m\leq B \qquad \text{(technology constraint)}
\end{aligned} \tag{2.1}$$

with $G'(\cdot) > 0, G''(\cdot) < 0, \lim_{m\to 0} G'(\cdot) \to +\infty, U'(\cdot) > 0, U''(\cdot) < 0$

The patient has a strictly concave utility function. With probability $(1-\pi)$, he is in the good state of the world where no illness shock occurs. In that case, only the premium P is deducted from his initial endowment Y. The illness state occurs with probability π. In this case, the patient incurs an illness shock with a monetary equivalent of ϵ. In order to cure the illness shock the patient receives a certain amount of treatment m. Via the strictly concave production function G the chosen level of m is transformed into the monetary equivalent of restored health. Despite the fixed premium P the patient also has to bear part of the monetary costs $c\cdot m$ of treatment, with c representing constant marginal monetary costs of treatment. The share a he has to incur is determined by his insurance contract which specifies the fixed premium P and the copayment rate[8] a. Finally, the patient has to suffer from non-monetary costs of treatment $n\cdot m$ where n are the constant marginal non-monetary costs.

If the patient falls ill, he receives treatment from the physician. This holds true if the physician's profit from signing the contract and treating the patient is at least as big as his outside option, here normalized to zero. The physician's contract specifies a fixed payment R and a fee per unit of treatment p. $p > c$ depicts a fee-for-service system. $p = c$ is cost reimbursement, whereas cost sharing occurs for $p \in [0, c[$. $p = 0$ is the extreme case of capitation without any reimbursement of treatment costs.

The insurer receives the premium P in both states of the world. In the illness case, he additionally receives the patient's copayment but has to pay the physician. The technology in this setting is determined[9] by marginal monetary costs of treatment c, by marginal non-monetary costs n to the patient, and by the technological boundary B. From the First

[8] Note that a can also be negative, which implies a subsidy to the patient for every unit of treatment.

[9] Technologies are also likely to differ with respect to non-monetary costs to the physician like costs of informing the patient on the procedure or the time needed to carry out an operation. Since these factors do not influence the patient's demand response they are neglected in the following. The intuition

Order Conditions with respect to the relevant variables m, P, a, R, p we obtain the well-known result that marginal utility is equalized over states, i.e. $U'(S) = U'(H)$ where S is income in the sickness state and H is income in the health state. The premium P thus includes all expected monetary and non-monetary costs incurred by the patient, i.e. $P = \pi(c \cdot m_{FB} + \epsilon - G(m_{FB}) + n \cdot m_{FB})$, whereas the physician is reimbursed for all his expenses: $R = -(p - c) \cdot m_{FB}$. The patient's share in monetary costs a can be positive or negative because the patient is compensated for non-monetary costs of treatment and for the difference in utility between illness shock and healing. If, for instance, the maximal healing level is lower than the illness shock then the patient receives a monetary compensation for the residual of the illness shock which is left after the consumption of health care. For the characterisation of the first best amount of care, we have to distinguish two cases:

If $m_{FB} < B$, then the first best amount of care is characterised by the condition $G'(m_{FB}) = n + c$, i.e. marginal utility equals the sum of marginal non-monetary and monetary costs. m_{FB} decreases if n or c increase, and is unaffected by marginal variations of B.
If $m_{FB} = B$, the condition $\pi U'(S)(G'(m_{FB}) - n - c) - \mu = 0$ holds where μ is the Lagrange multiplier with respect to the technology constraint and S is income in the sickness state. m_{FB} then increases with a marginal increase in B and remains constant when monetary costs c or non-monetary costs n are varied.

Now the effects of changes of technology parameters on social welfare can be analysed.

Lemma 1 *Welfare effects of parameter changes - First Best*
In the first best situation, social welfare decreases in monetary costs c and in non-monetary costs n. Marginal changes in c and n are perfect substitutes in their effect on welfare.
When the technological boundary B increases
i) welfare remains constant if $m_{FB} < B$
ii) welfare increases if $m_{FB} = B$.

---

concerning the adoption decision with respect to non-monetary costs to the physician is straightforward. Irrespective of the reimbursement system has the physician an incentive to adopt technologies which reduce his non-monetary costs. Yet the introduction of variable non-monetary costs into the model would complicate the adoption decision with respect to the other technology parameters, giving in each case additional weight to technologies which induce a reduced demand for health care by patients.

Proof Using the Envelope Theorem, one can show that the effects of changes in technology parameters on social welfare are unambiguous. When denoting $V(c, n, B)$ as the value function the following results are derived:
$\frac{\partial V}{\partial c} = -\pi \cdot m_{FB} \cdot U'(S) < 0$
$\frac{\partial V}{\partial n} = -\pi \cdot m_{FB} \cdot U'(S) < 0$
$\frac{\partial V}{\partial B} = \mu \geq 0$
where S indicates the income in the illness state. This replicates the results derived by Baumgardner (1991). **QED**

2.3 Provider incentives and welfare in a world of ex-post moral hazard

In the following, welfare implications of new technologies in a situation where patients exert ex-post moral hazard and the insurer designs second best efficient contracts are studied. As a starting point, technology is still contractible but the treatment intensity chosen by the patient is not. Throughout the paper it is assumed, though, that total costs of treatment are verifiable ex post. This allows us to abstract from an otherwise necessary reporting subgame[10]. Having replicated the results derived by Baumgardner (1991) for the case of conventional coinsurance, incentives of providers for adopting new technologies and the corresponding welfare implications are analysed.
The social planner has to take into account the patient's response to a given insurance contract and technology. In case of illness the patient maximizes his utility over m. This is equivalent to maximizing his income in the illness state:

$$m_{SB} = \quad \text{argmax} \quad Y - P - a \cdot c \cdot m - \epsilon + G(m) - n \cdot m \tag{2.2}$$

s.t. $m \leq B$. The First Order Condition defines m_{SB} as an implicit function of a and the technology parameters. In the case of an interior solution, the condition $G'(m_{SB}) = n + a \cdot c$ holds. That means that the patient chooses m such that marginal utility equals marginal costs faced by him. m_{SB} decreases in the coinsurance rate a and in the non-monetary costs n, and is unaffected by variations of B. Variations in monetary costs c have an

[10] Ma and McGuire (1997) consider such a subgame and therefore restrict their analysis to a positive copayment rate for the patient and to copayment rate and physician's share in total costs to less than or equal 1, an arrangement which induces truthful reporting to the insurer of the amount of care consumed.

ambiguous effect on m_{SB} because this depends on a: If the patient is subsidized for each unit of treatment, i.e. if $a < 0$, then an increase in monetary costs in fact reduces total marginal costs faced by the patient and induces him to consume more, $\frac{\partial m_{SB}}{\partial c} \geq 0$. If the patient bears a positive copayment rate, $a \geq 0$, then he suffers from increases in monetary costs and consequently reduces the chosen amount, thus $\frac{\partial m_{SB}}{\partial c} \leq 0$. If there is a corner solution, m_{SB} increases in B, but is unaffected by marginal changes of other parameters.

The social planner takes the patient's response into account when designing the optimal insurance contract and remuneration scheme for the physician. He maximizes

$$\begin{aligned}
&\max_{P,a,R,p} (1-\pi)U(Y-P) + \pi U(Y - P - a\cdot c\cdot m - \epsilon + G(m) - n\cdot m) \\
&s.t. \\
&P + \pi\cdot a\cdot c\cdot m - \pi R - \pi\cdot p\cdot m \geq 0 \qquad \text{(zero-profit condition insurer)} \\
&R + (p-c)\cdot m \geq 0 \qquad \text{(zero-profit condition physician)} \\
&m = m_{SB}(a,c,n,B) \qquad \text{(patient's reaction function)}
\end{aligned} \tag{2.3}$$

This yields the First Order Condition for a_{SB}

$$-[(1-\pi)U'(H) + \pi U'(S)]\cdot[(1-a)\tfrac{\partial m}{\partial a} - m] = m\cdot U'(S) \tag{2.4}$$

which replicates Baumgardner's results for the coinsurance case[11]. a_{SB} is set such that the marginal expected gain from a marginal increase in a, which would lead to a reduction in P in both states of the world, is equalized to the marginal expected loss from a higher copayment in the illness state. a_{SB} can be derived from this condition. It still can be positive or negative. This allows us to calculate m_{SB} and P_{SB}. R and p can be chosen as one likes but such that the physician's participation constraint is binding.
Given that the Second Order Condition holds one can show[12] that a_{SB} increases in π and decreases in ϵ. The second best optimal copayment is thus higher if the illness case is more likely, and lower if the illness is more severe. a_{SB} increases (decreases) in income Y for DARA (IARA) utility functions[13].

Note that only if $a_{SB} = 1$ is the first best amount of restored health chosen by the patient. If $a_{SB} < 1$ then $m_{SB} > m_{FB}$ holds[14], the well-known result under ex-post moral hazard.

[11] See Baumgardner (1991), equ. (6). Again, S is income in the sickness state, whereas H indicates income in the health state.

[12] See appendix.

[13] With respect to the technology parameters n and c no clearcut predictions can be made.

[14] Theoretically, m_{SB} may be smaller than m_{FB} if $a_{FB} \geq 1$ and consequently $a_{SB} > 1$. This may

Now the effects of changes in technology parameters can again be analysed.

Lemma 2 *Welfare effects of parameter changes - Second Best*
In the second best situation, social welfare decreases in monetary costs c and in non-monetary costs n. Marginal changes in c and n are perfect substitutes in their effect on welfare.
When the technological boundary B increases
i) welfare remains constant if $m_{SB} < B$
ii) welfare increases if $m_{FB} = m_{SB} = B$ (B does not limit overconsumption)
iii) welfare decreases if $m_{FB} < m_{SB} = B$ (B limits overconsumption).

Proof Using the Envelope Theorem, one can derive the effects of changes in technology parameters on social welfare. When denoting $V(c, n, k, B)$ as the value function the following results are obtained[15]:

$$\frac{\partial V}{\partial c} = -\pi \cdot m_{SB} \cdot [(1-\pi)U'(H) + \pi U'(S)] < 0$$

$$\frac{\partial V}{\partial n} = -\pi \cdot m_{SB} \cdot [(1-\pi)U'(H) + \pi U'(S)] < 0$$

$$\frac{\partial V}{\partial B} = \begin{cases} 0 & \text{if B is not a binding constraint} \\ \pi \cdot U'(S) \cdot (G'(m_{SB}) - n - c) & \text{if B is a binding constraint} \end{cases}$$

The latter expression is positive if $m_{FB} = m_{SB} = B$ and negative if $m_{FB} < m_{SB} = B$. This again replicates the results derived by Baumgardner (1991). **QED**

Let us now go one step further: Within this setting of ex-post moral hazard the physician chooses the technology which he offers to the patient. Assume that there is a standard technology in the market which the physician uses. Now, an innovation takes place and a new technology is available. Which types of innovations the physician will adopt depends on his remuneration. From the second best analysis, it has become clear that the social planner would want him to adopt technologies which reduce costs, but the desired activity with respect to the boundary B is unclear.

In the last stage of the game the patient adjusts his preferred amount of restored health,

happen in a situation where the patient is made better off by a small or even negative premium but has to bear more than the full costs of treatment in case of illness. For such a situation to be optimal the illness shock has to be very small compared to the first best amount of treatment. That is, the patient is "overtreated" in case of illness. Along with Baumgardner (1991) I restrict the analysis to the empirically more relevant case where medical treatment cannot raise health (too far) above its no-illness level.

[15] Again, S is income in the sickness state, whereas H indicates income in the health state.

	Cost reimbursement $p = c$	Fee-for-service $p > c$	Cost sharing $p < c$
$\frac{\partial \Pi}{\partial c} > 0$		$\eta_{m_{SB}/c} > \frac{c}{p-c}$	$\eta_{m_{SB}/c} < \frac{c}{p-c}$
$\frac{\partial \Pi}{\partial c} = 0$		$\eta_{m_{SB}/c} = \frac{c}{p-c}$	$\eta_{m_{SB}/c} = \frac{c}{p-c}$
$\frac{\partial \Pi}{\partial c} < 0$	$\forall \eta_{m_{SB}/c}$	$\eta_{m_{SB}/c} < \frac{c}{p-c}$	$\eta_{m_{SB}/c} > \frac{c}{p-c}$

Table 2.1: Impact of changes in monetary costs on profit (second best)

m_{SB}, to changes in the technology parameters induced by the physician. The physician takes these adjustments into account when taking his decision. The physician maximizes his profit

$$\max_{n,c,B} \Pi = R + (p - c) \cdot m_{SB}(a, c, n, B) \tag{2.5}$$

over n, c, B where $m_{SB}(\cdot)$ is the patient's response function. The first order derivatives are

$$\frac{\partial \Pi}{\partial n} = (p - c) \cdot \frac{\partial m_{SB}}{\partial n} \begin{cases} > 0 & \text{if } p < c \\ = 0 & \text{if } p = c \text{ and/or } \frac{\partial m_{SB}}{\partial n} = 0 \\ < 0 & \text{if } p > c \end{cases}$$

$$\frac{\partial \Pi}{\partial B} = (p - c) \cdot \frac{\partial m_{SB}}{\partial B} \begin{cases} < 0 & \text{if } p < c \\ = 0 & \text{if } p = c \text{ and/or } \frac{\partial m_{SB}}{\partial B} = 0 \\ > 0 & \text{if } p > c \end{cases}$$

$$\frac{\partial \Pi}{\partial c} = (p - c) \cdot \frac{\partial m_{SB}}{\partial c} - m_{SB} = m_{SB} \cdot [\tfrac{p-c}{c} \cdot \eta_{m_{SB}/c} - 1] \begin{cases} > 0 & \text{if } \frac{p-c}{c} \cdot \eta_{m_{SB}/c} > 1 \\ = 0 & \text{if } \frac{p-c}{c} \cdot \eta_{m_{SB}/c} = 1 \\ < 0 & \text{if } \frac{p-c}{c} \cdot \eta_{m_{SB}/c} < 1 \end{cases} \tag{2.6}$$

with $\eta_{m_{SB}/c} = \frac{\partial m_{SB}}{\partial c} \cdot \frac{c}{m_{SB}}$ denoting the demand elasticity with respect to monetary costs. Note that $\eta_{m_{SB}/c} \leq 0$ if $a \geq 0$ and ≥ 0 if $a < 0$. The impact of changes in monetary costs on profits depends - as summarized in table 2.1 - on the demand elasticity and the reimbursement system.

The incentives with respect to the adoption decision can be summarized as follows:

Lemma 3 *Adoption of innovations*
Non-monetary costs n: The physician is indifferent under cost reimbursement, and adopts an innovation which increases n under cost sharing and decreases n under fee-for-service.
Boundary B: The physician is indifferent under cost reimbursement, and adopts an innovation which decreases B under cost sharing and increases B under fee-for-service.
Monetary costs c: The physician adopts an innovation which decreases c under cost reimbursement, decreases (increases) c under cost sharing if $\eta_{m_{SB}/c} > (<) \frac{c}{p-c}$, and decreases (increases) c under fee-for-service if $\eta_{m_{SB}/c} < (>) \frac{c}{p-c}$.

Note that $\frac{c}{p-c} < 0$ under cost sharing and > 0 under fee-for-service. Therefore, $\frac{\partial \Pi}{\partial c} \geq 0$ under fee-for-service only if $a_{SB} < 0$ with the special case of a positive elasticity of demand with respect to monetary costs.

$\eta_{m_{SB}/c}$ decreases (increases)[16] in the illness probability π (severity of illness ϵ) if $\eta_{m_{SB}/c} < 0$. If $\eta_{m_{SB}/c} > 0$ and big enough the effects go in the opposite direction. $\eta_{m_{SB}/c}$ decreases (increases) in income Y for DARA (IARA) utility functions if $\eta_{m_{SB}/c} < 0$. Again, if $\eta_{m_{SB}/c} > 0$ and big enough, these effects change.

Under a fee-for-service system the physician has an incentive to induce a high demand for health care since he receives a positive net payment for every unit consumed. Therefore, he wants to reduce n and increase B. The opposite holds under cost sharing, because the physician is a net payer at the margin for every unit of health care. Consequently, he increases n and decreases B in order to reduce the patient's demand for health care. Only under cost reimbursement is the physician indifferent concerning variations in these parameters. Concerning monetary costs c, two effects have to be distinguished. On the one hand, an increase in c increases the physician's costs at the margin in any of the three reimbursement systems. He therefore has an incentive to reduce c. On the other hand, an increase (decrease) in c induces a reduction (increase) in the demand for health care by the patient. If, under fee-for-service, the elasticity of demand is negative, both effects go in the same direction and the physician reduces c. If the demand elasticity is positive and very high then the demand response may outweigh the direct effect and the physician increases c. Under cost sharing, this may happen if the demand elasticity is (negative

[16] See appendix. It is furthermore assumed that $G'''(\cdot) \cdot m + G''(\cdot) \leq 0$, which holds if $G'''(\cdot) \leq 0$ or (positive but close to) zero.

and) very elastic. Here, both effects go in the same direction if the demand elasticity is positive.

Combining Lemma 2 and 3 yields the welfare implications of the adoption incentives:

Proposition 1 *Welfare implications of adoption decision - Second Best*

- *Cost reimbursement ($p = c$): The physician has welfare increasing incentives with respect to c, and is indifferent with respect to changes in n and B.*
- *Cost sharing ($p < c$): The physician has welfare reducing incentives with respect to n, to B if B does not limit overconsumption, to c if demand is elastic with respect to costs. He has welfare increasing incentives with respect to B if B limits overconsumption, to c if demand is inelastic (positive or weakly negative) with respect to costs.*
- *Fee-for-service ($p > c$): The physician has welfare reducing incentives with respect to B if B limits overconsumption, to c if the demand elasticity with respect to costs is highly positive. He has welfare increasing incentives with respect to n, to B if B does not limit overconsumption, to c if the demand elasticity with respect to costs is negative or weakly positive.*

Ideally, the state would determine a_{SB} for each illness, and then calculate patients' elasticity of demand. Then, the welfare implications under the differing reimbursement schemes would have to be weighted against each other, and the most convenient one be chosen. Yet one can conclude that only in very special cases may there exist a simple linear remuneration scheme, namely a fee-for-service, which induces correct incentives with respect to all possible parameter changes. This result is summarized in the following:

Corollary 1: *Second best efficiency*
Second best efficiency can only be achieved in the following situation: The state implements a fee-for-service system which induces the correct incentives with respect to non-monetary costs n. A higher B has to be socially desirable, this is the case if the illness shock is sufficiently severe or if the current technical boundary is very low. Furthermore, $\eta_{m_{SB}/c} < \frac{c}{p-c}$ has to hold (copayment for the patient positive or weakly negative).

In all other cases we are not able to achieve second best efficiency with a simple remuneration scheme. Expected utility of the patient will be lower than in the second best case.

2.4 Provider incentives and welfare under "Standard Coinsurance"

The analysis in the previous section relied on the assumption that the state implements second best efficient insurance contracts. This is of course not what we observe in the real world. We observe insurance contracts which do neither differentiate between various diseases in terms of copayment rates nor adjust copayment rates as a response to technical changes in treatment. For our approach to provide policy implications it seems reasonable to analyse technology parameter changes for such a given real world coinsurance contract[17].

Let us assume in the following that the state sets a copayment rate[18] $\bar{a}$, with $0 \leq \bar{a} < 1$. The state then determines $\bar{p}$, $\bar{P}$, and $\bar{R}$ such that the participation constraints of the physician and of the insurer are just binding. This insurance contract will in the following be referred to as "standard coinsurance" contract. The patient's utility under this given contract is

$$\begin{aligned}\bar{V} = {} & (1-\pi)U(Y - \pi[(1-\bar{a})c \cdot m_{SC}(\cdot)]) + \pi \cdot U(Y - \pi[(1-\bar{a})c \cdot m_{SC}(\cdot)] \\ & -\bar{a} \cdot c \cdot m_{SC}(\cdot) - \epsilon + G(m_{SC}(\cdot)) - n \cdot m_{SC}(\cdot))\end{aligned} \tag{2.7}$$

where $m_{SC}(\bar{a}, c, n, B)$ is the patient's reaction function to parameter changes and is equivalent to the second best case but takes the constant $\bar{a}$ into account.

Now the effects of changes in technology parameters can be analysed for this standard coinsurance contract[19].

[17] See Zeckhauser (1970) for an illustration of the related welfare losses.

[18] The restriction may be seen as a tribute to the requirement of truthful reporting.

[19] It is assumed that P is always adjusted such that the zero-profit conditions for the insurer and for the physician are fulfilled.

Lemma 4 *Welfare effects of parameter changes - Standard Coinsurance*
For a standard coinsurance contract with parameters $0 \leq \bar{a} < 1$, $\bar{P}$, changes in technology parameters have ambiguous effects on welfare. The more elastic the patient's response is to increases in monetary and non-monetary costs, the more likely is such an increase a welfare improvement because a high reduction in premiums is induced.
An increase in B allows on the one hand more treatment but this may be offset on the other hand by higher premiums to be charged.

Proof Differentiation of $\bar{V}$ with respect to the technology parameters gives[20]

$$\frac{\partial \bar{V}}{\partial c} = -[(1-\pi)U'(H) + \pi U'(S)] \cdot \pi(1-\bar{a}) \cdot m_{SC} \cdot (1 + \eta_{m_{SC}/c}) - \pi \cdot \bar{a} \cdot m_{SC} \cdot U'(S)$$

$$\frac{\partial \bar{V}}{\partial n} = -[(1-\pi)U'(H) + \pi U'(S)] \cdot \pi(1-\bar{a}) \cdot \frac{m_{SC} \cdot c}{n} \cdot \eta_{m_{SC}/n} - \pi \cdot m_{SC} \cdot U'(S)$$

$$\frac{\partial \bar{V}}{\partial B} = \begin{cases} 0 \quad \text{if B is not a binding constraint} \\ -[(1-\pi)U'(H) + \pi U'(S)][\pi(1-\bar{a})c] + \pi \cdot U'(S)[-\bar{a} \cdot c + G'(B) - n] \\ \qquad \text{if B is a binding constraint} \end{cases}$$

The latter expression is ambiguous in its sign. The first term is negative and represents the expected marginal disutility caused by a higher premium payment when B is marginally increased. This effect may offset the second positive term which stands for expected utility from marginally increased treatment.
Concerning changes in c and n, the first term represents the marginal utility which stems from a lower premium P when costs increase and consequently the amount of care goes down. The second term represents the marginal disutility caused by higher treatment costs. If B is a binding constraint, then $\eta_{m_{SC}/c} = \eta_{m_{SC}/n} = 0$ and consequently $\frac{\partial \bar{V}}{\partial c}$ and $\frac{\partial \bar{V}}{\partial n}$ unambiguously negative. **QED**

The physician's maximisation problem is essentially the same as in the second best case. Since by assumption $0 < \bar{a} \leq 1$ and therefore $\eta_{m_{SC}/c} \leq 0$, the impact of changes in monetary costs on profit can now be summarized as in table 2.2.
It is interesting to note that the exogenous parameters π, ϵ and income Y do not have an impact on the demand elasticities $\eta_{m_{SC}/c}$ and $\eta_{m_{SC}/n}$ because $\bar{a}$ is exogenously given. On the other hand, the impact of the exogenous parameters $\bar{a}$, c, and n is unambiguous.

[20] $\eta_{m_{SC}/c} = \frac{\partial m_{SC}}{\partial c} \cdot \frac{c}{m_{SC}}$ denotes the demand elasticity with respect to monetary costs; $\eta_{m_{SC}/n} = \frac{\partial m_{SC}}{\partial n} \cdot \frac{n}{m_{SC}}$ denotes the demand elasticity with respect to non-monetary costs.

	Cost reimbursement $p = c$	Fee-for-service $p > c$	Cost sharing $p < c$
$\frac{\partial \Pi}{\partial c} > 0$			$\eta_{m_{SC}/c} < \frac{c}{p-c}$
$\frac{\partial \Pi}{\partial c} = 0$			$\eta_{m_{SC}/c} = \frac{c}{p-c}$
$\frac{\partial \Pi}{\partial c} < 0$	$\forall \eta_{m_{SC}/c}$	$\forall \eta_{m_{SC}/c}$	$\eta_{m_{SC}/c} > \frac{c}{p-c}$

Table 2.2: Impact of changes in monetary costs on profit (Standard Coinsurance)

Both $\eta_{m_{SC}/c}$ and $\eta_{m_{SC}/n}$ decrease in these parameters[21].

Combining Lemma 3 and 4, the following proposition can be established concerning incentives and welfare effects of technology parameter changes for a standard coinsurance contract:

Proposition 2 *Welfare implications of adoption decision - Standard Coinsurance*

- *Cost reimbursement ($p = c$): The physician has welfare increasing incentives with respect to c if demand is inelastic with respect to monetary costs. He has welfare reducing incentives with respect to c if demand is elastic with respect to monetary costs, and is indifferent with respect to changes in n and B.*

- *Cost sharing ($p < c$): The physician has welfare reducing incentives with respect to n if demand is inelastic with respect to non-monetary costs, to B if gains from additional treatment are high. He has welfare increasing incentives with respect to c, to n if demand is elastic with respect to non-monetary costs, to B if gains from additional treatment are low.*

- *Fee-for-service ($p > c$): The physician has welfare reducing incentives with respect to c if demand is elastic with respect to monetary costs, to n if demand is elastic with respect to non-monetary costs, to B if gains from additional treatment are low. He has welfare increasing incentives with respect to c if demand is inelastic with respect to monetary costs, to n if demand is inelastic with respect to non-monetary costs, to B if gains from additional treatment are high.*

[21] See appendix. Again, it is assumed that $G'''(\cdot) \cdot m + G''(\cdot) \leq 0$.

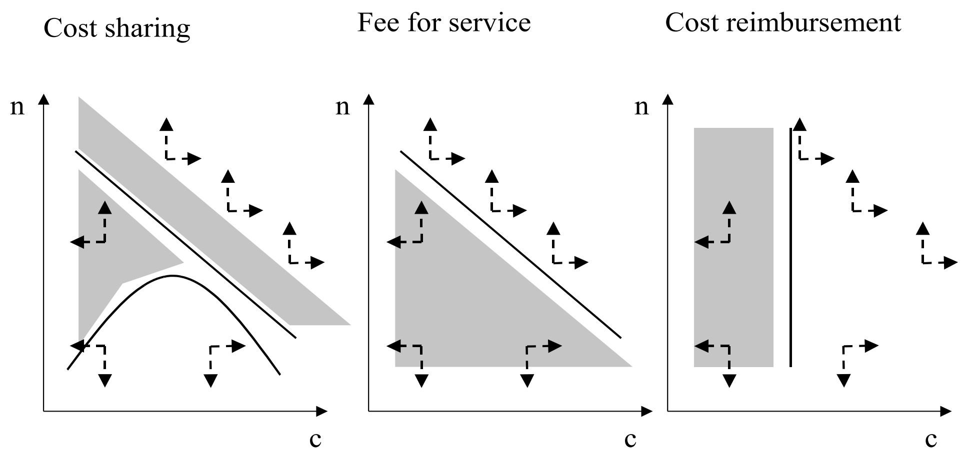

Figure 2.1: Simultaneous changes of non-monetary costs n and monetary costs c

Several insights can be gained from this. Interestingly, lower monetary costs c are adopted for sure under cost reimbursement and fee-for-service, but not necessarily under cost sharing. But only under cost sharing, these incentives are a clear welfare improvement. Secondly, c and n generally are complements in their effect on welfare. Under fee-for-service, a reduction in n makes the demand for health care with respect to monetary costs less elastic which implies that reductions in c are more likely to be a welfare improvement. The same holds for the impact of reductions in c on the demand elasticity with respect to non-monetary costs. Under cost sharing, the analogue can be established for increases in c and n which have a welfare increasing impact on the demand elasticity with respect to the other parameter. Only for the case of cost reductions, the resulting increase in the demand elasticity with respect to non-monetary costs makes increases in n more likely to be welfare reducing. Under cost sharing, c and n are strategic complements in the sense that an increase in n makes the adoption of technologies which increase c more likely. Finally, under fee-for-service it is more likely that the technical boundary B is reached since the reductions in monetary and non-monetary costs increase patients' demand for health care.

In reality, most technological advances will imply simultaneous changes of various technol-

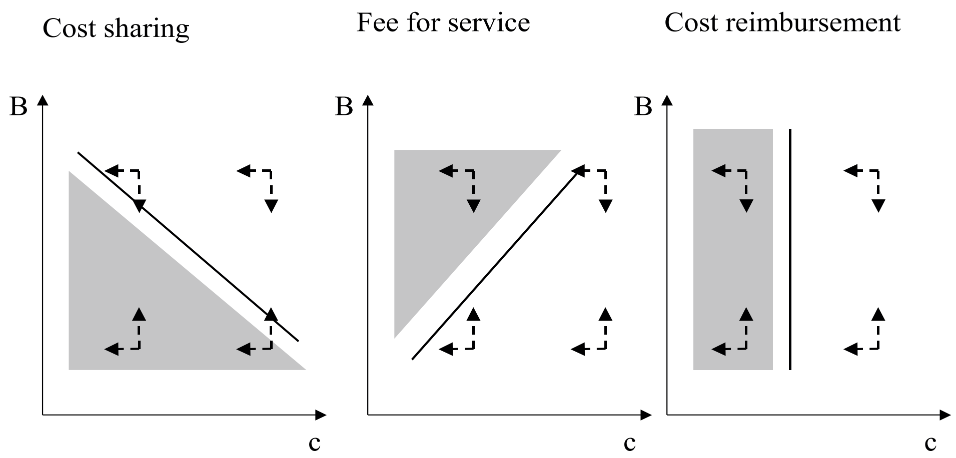

Figure 2.2: Simultaneous changes of non-monetary costs c and boundary B

ogy parameters. The results from Proposition 2 can be used to illustrate such a situation. In Figure 2.1, simultaneous changes of non-monetary costs to the patient n and monetary costs c are considered, assuming that B is not binding. In Figures 2.2 and 2.3, the tradeoff between B and the cost parameters if B is a binding constraint are depicted. The black lines represent isoprofit curves and the shaded areas indicate how the physician can increase profits compared with the usage of a technology on the curve[22]. The arrows indicate utility increases according to Lemma 4. Note that in Figure 2.1, the direction of utility increases depends on the demand elasticities with respect to monetary and non-monetary costs. Both are inelastic for low values of n and c, but decrease (become more elastic) in both parameters. For high values of c and n, the demand for health care is very elastic with respect to non-monetary and monetary costs, consequently an increase in n and / or c increases patients' utility. The reduction in premiums which stems from a lower demand for care outweighs higher treatment costs. It may seem implausible to follow an adoption path which reduces the amount of health care consumed further and further. But there are natural limits to this path, in the sense that only technologies will be admitted for which the condition $\lim_{m \to 0} G'(m) > n + c$ holds. That means that at

[22]For calculations of the slopes of the isoprofit functions see Lemma 3 and appendix.

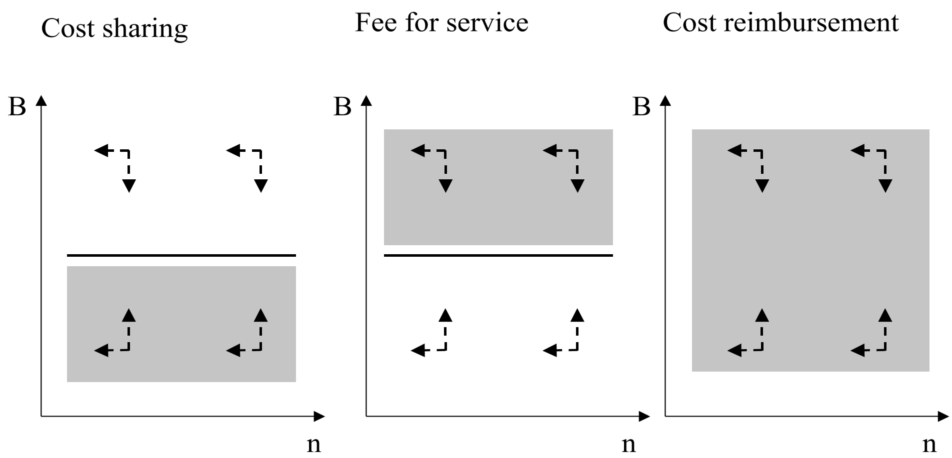

Figure 2.3: Simultaneous changes of non-monetary costs n and boundary B

least for the first unit of treatment the benefits must outweigh the costs. This obvious requirement for global efficiency prevents a killing-off of the insurance market.

The optimal choice of reimbursement sytems can now be concluded from the analysis. Technologies should be categorized according to several criteria. The following subgroups can be set up:

- If B is a binding constraint and increases in B are welfare enhancing then a fee-for-service system should be used. Increases in B and reductions in c and n are induced all of which are welfare improvements.
- If B is a binding constraint and increases in B are welfare reducing then a system of cost sharing should be implemented which induces reductions in B and c and gives no clearcut incentives with respect to n. Incentives with respect to B and c are thus optimal, and incentives with respect to n cannot be improved upon.
- If B is not binding the optimal reimbursement system depends on the demand elasticities with respect to costs. The optimal choice is summarized in table 2.3.

In many countries, there is a tendency to shift from cost reimbursement to capitation

	$\eta_{m_{SC}/c}$ inelastic	$\eta_{m_{SC}/c}$ elastic
$\eta_{m_{SC}/n}$ inelastic	Fee-for-service	FFS / CS dep. on relative impact of n and c
$\eta_{m_{SC}/n}$ elastic	Cost sharing	Cost sharing

Table 2.3: Optimal reimbursement if B is not binding

systems. In **Germany**, for example, patients are in general publicly insured with a low copayment rate for treatment. Physicians in the out-patient sector receive cost reimbursement ($s = 0$). This implies that the only incentive they face with respect to new technologies is to choose technologies with lower monetary costs. Consequently, technical advances should not account for a large share in cost increases in the German out-patient sector.

The more interesting case is the German in-patient sector: Since 1996, elements of prospective payment systems have been introduced for certain treatments, but today they still account for less than 25% of all cases[23]. Currently, the system of Diagnosis Related Groups (DRGs) is introduced for a wider range of treatments. DRGs classify patients according to the diagnosis they have received into a certain category which implies a fixed payment to the hospital. That means that hospital reimbursement changes for an increasing number of treatment categories from cost reimbursement to cost sharing ($p < c$). Assuming that physicians are at least imperfect agents for the hospital, their incentives change. According to Proposition 2, incentives with respect to c are indeed welfare improving in this setting. But the ongoing reforms also include higher copayments for patients which implies a more elastic demand for health care with respect to monetary costs. Therefore, monetary costs may well be increased under cost sharing. The desired reduction in health care cost increases may only indirectly be achieved via less treatment and lower premiums. Furthermore, the providers will now have an incentive to adopt technologies with higher non-monetary costs to the patient and with a lower technical boundary. These effects may or may not be socially desirable. According to this analysis, the DRG system should only be put into place for treatments where an increase in the technical boundary is a clear welfare reduction. For the case of severe illness shocks a tendency to lower B is in most cases not welfare improving. It is also very likely that for severe illnesses patients have a relatively inelastic demand for treatment

[23]See European Observatory (2000).

with respect to non-monetary costs. These treatments should therefore be reimbursed under a fee-for-service system.

In the last decades, managed care systems have found widespread utilization in the **United States** health care system[24]. These plans are so diverse that it is impossible to categorize them into one single scheme. Most plans require small copayments from their patients (and sometimes larger ones for treatment outside the network). But they use a wide range of methods to pay physicians and hospitals, including salaries, fee-for-service, and capitation. The variety of reimbursement schemes could be used for an empirical analysis of technology adoption. Depending on the payment scheme, certain technologies should experience a faster and wider diffusion under some managed care plans as compared to others. Case studies which analyse certain diseases and try to focus more explicitly on the type of technological progress which is introduced by a new technology seem to be a field for fruitful further research.

Overall, one should also observe more innovations with respect to those technologies which are adopted by physicians. This may provide another useful approach for empirical investigations.

2.5 Conclusion

Provider incentives and welfare implications for adopting new technologies under various reimbursement schemes have been analysed. It has been shown that in a second best world of ex-post moral hazard, a simple remuneration scheme which implements the adoption of second best efficient technologies exists only in very special cases. Furthermore, the analysis of the current tendency to shift from cost reimbursement to capitation in countries like Germany has proved that the direct incentives for cost savings are stronger under fee-for-service and cost reimbursement than in a capitation system. In addition, cost sharing may imply negative welfare effects with respect to non-monetary costs of treatment and to the technical boundary of treatment. The state might classify diseases or treatments into subgroups which are reimbursed differently in order to achieve the desired welfare effects in each subgroup. Especially for the case of extremely severe illness shocks the introduction of a fee-for-service system may be socially desirable. It is these very severe diseases where

[24]See Glied (2000) for an overview.

the capitation system not only seems to induce negative welfare effects with respect to non-monetary costs but also with respect to the technologically feasible boundary of treatment. Cost sharing/capitation systems are most suitable for treatments where a reduction of the amount of health care consumed is desirable. If the technical boundary of treatment is not a binding constraint then the demand elasticities with respect to monetary and non-monetary costs have to be analysed and the respective welfare effects of changes in n and c have to be weighted against each other when deciding on a reimbursement scheme.

Future research on provider incentives can be directed in many ways. The US system of managed care plans seems to provide due to its diversity of reimbursement schemes an ideal base for empirical research in this field. But also further theoretical investigations are due. Ongoing research is dedicated to the analysis of incentives in a setup of supplier-induced demand. This would relax our assumption that the patient chooses the amount of care he consumes and might fit certain institutional settings and treatment categories in a more adequate way. It would in particular take informational asymmetries between physician and patient more explicitly into account.

What has also been neglected in the analysis so far is the role of fixed costs. Especially when reimbursement systems change or new technologies come to the market physicians might experience a lock-in effect if they have adopted an expensive technology. The incentives for adopting technologies in the first place will certainly change in this setting. But physicians could also use the investment decision as a signalling device for patients in order to reduce informational asymmetries.

Finally, a closer look at strategic interactions of providers might be worthwhile. As pointed out in the empirical literature, competition as well as networks between providers seem to have a strong influence on the adoption of new technologies.

3 Demand Inducement for Health Care Technologies

3.1 Introduction

Economic evaluation is becoming an important tool in political decision making. Health technology assessment and other types of economic evaluation are meant to ensure that health care expenditure is used in an efficient way. Nevertheless, a better theoretical understanding and empirical investigation of the factors which influence the diffusion of new technologies is of crucial importance for explaining which innovations will actually be used in the health care market, and how the usage of cost effective technologies can be implemented. An economic evaluation is necessary before a technology is allowed to be used by providers of health care, but it is not sufficient for ensuring that the technology is actually adopted and offered to patients, nor that it is offered to those patients who should receive this treatment. A prominent example of technological progress is the usage of bypass surgery and angioplasty in heart attack treatment. Cutler and McClellan (1996) provide data for the US which show that the introduction of angioplasty in the 1980s did by no means reduce the usage of the more invasive and more costly yet also more effective treatment of bypass surgery. Both procedures were increasingly acquired and used throughout the period under consideration[25]. One of the explanations can be found in the following analysis of demand inducement.

This study focuses on the crucial role of providers in this process. It intends to contribute to the discussion by analysing a monopoly provider's technology choice and incentives for adopting a new technology in a setting where patients can only observe the investment decision and the provider's treatment recommendation but not their health status. It turns out that the monopoly provider invests into a range of technologies according to their profitability and acceptance by patients. If technology choice cannot be contracted upon the first best treatment pattern is not achievable but the state can induce the first

[25] The authors examine the sources of expenditure growth in heart attack treatment. They first show that essentially all of the cost growth is a result of the diffusion of particular intensive technologies. Then they conclude that insurance variables, technology regulation, and provider interactions have the largest quantitative effect on technology diffusion. These factors affect both technology acquisition and the frequency of technology use. See also Phelps (1992) and Phelps and Mooney (1993) who find that the use of interventions that could serve as medical substitutes are instead positively correlated.

best selection of technologies by the adequate scheme of profitable reimbursement. Yet patients are treated according to their symptoms and irrespective of their true health status.

A cost effective innovation is adopted if it is more profitable than the reference technology and / or accepted by a wider range of patients. Optimal reimbursement of innovations takes these effects into account. An innovation generally makes some patients better off due to better treatment and makes some worse off due to more costly or less effective overtreatment. The state may reduce inefficiencies associated with the introduction of a new technology by introducing technology specific patient copayments.

The first part of the study analyses the provider's technology choice and treatment pattern. The physician is modelled as an expert who gains an informational advantage by diagnosing the patient. He knows the patient's health status more exactly than the patient and therefore can to a certain extent induce demand for treatment which the patient accepts but would not want if he was fully informed[26]. This informational asymmetry cannot be resolved even after treatment because from the ex post health status the correct treatment recommendation cannot be concluded with certainty. The patient is restricted to a "consent" or "no consent" decision to a treatment recommendation. Taking the physician's behaviour into account the patient will only consent to treatment if his symptoms are severe enough. He may be willing to consent to certain treatment technologies but not to others which are more costly or less effective. Therefore, the provider will offer different technologies to different ranges of patients. His offer will be led by profitability considerations and patients' willingness to consent.

If technology choice can be contracted upon then the first best technology choice and treatment pattern can be achieved. The physician receives a fee-for-visit and zero-profit reimbursement for treatment. If technology choice cannot be contracted upon the physician only has an incentive to offer different types of treatments if treatment is profitable for him. The patients with most severe symptoms will then be offered the most profitable technology in a pooling equilibrium, which means that they receive treatment irrespective of their true health status. The patients whose symptoms are a bit less severe and who

[26]Already in 1963, Kenneth J. Arrow wrote that "...the special economic problems of medical care can be explained as adaptations to the existence of uncertainty in the incidence of disease and in the efficacy of treatment." Among other issues he mentioned moral hazard problems on the side of the patient and on the side of the physician.

would not accept the most profitable technology are offered a less profitable one which they are willing to accept, and so on. If there exist zero-profit technologies, an intermediate range of patients receives the correct amount of treatment in a separating equilibrium, that is they are only offered treatment if they really need it. The range of patients with the least severe symptoms may decide not to visit the physician if diagnosis is costly for them or if they fear overtreatment. Optimal reimbursement of technologies induces the first best range of technologies by the adequate ranking of profitable reimbursement. Yet the first best treatment pattern is not achievable because patients are overtreated according to their symptoms instead of receiving the correct treatment according to their true health status.

The second part of the analysis is dedicated to the reimbursement of innovations. When a new technology comes on the market the physician adopts it if it is more profitable and / or if it is accepted by a larger range of patients than the currently used technology. Each technology is described by four parameters: monetary costs of treatment, non-monetary costs to the patient and to the physician, and effectiveness of treatment. If only one technology parameter changes (which may be the case for many process innovations) then the physician adopts the innovation given equal profitability because it is accepted by a wider range of patients. If more than one technology parameter changes certain cost effective new technologies may not be adopted given equal profitability because a smaller share of patients is willing to accept them. In this case, the technology has to be more profitable than the reference technology. Optimal reimbursement then adjusts the adequate ranking of profitable reimbursement.

These incentives do not translate into clear welfare improvements. Whereas there is always a range of patients who gain from better treatment another fraction of patients lose because they are now willing to accept being treated with this new technology which is more costly or less effective than the one they would have got previously. Consequently, the adoption of an innovation implies ambiguous effects on social welfare. The state can counteract these welfare losses by introducing technology specific copayments. This limits patients' willingness to consent and therefore prevents the physician from offering these treatments inadequately.

Several extensions are discussed in the last section of the paper. The above results have been derived by taking the premium which the insured patients pay as given. Obviously, cost saving or cost increasing new technologies will induce a premium adjustment. It can be shown that the above results still hold for CARA and for DARA utility functions when taking premium adjustments into account. If overall profitability is not easily observed by the state then cost sharing provides stronger incentives for process innovations than cost reimbursement. Finally, competition of providers does not change the results in general but may lead to specialization equilibria in which some technologies are not invested into by all physicians.

The existing theoretical literature in the field does not explain the diffusion of new technologies nor look at the impact of regulation and the role of providers of health care. It has not attempted to explain the actual adoption of technological innovations and their use by providers. There are only a few studies which try to model technical advances in the health sector. They focus on the role of insurance coverage for welfare effects of technological change[27] This study takes up Baumgardner's (1991) approach who describes a technology in terms of three parameters: monetary costs of treatment, non-monetary costs, and the technical boundary up to which healing is possible using this technology. He analyses welfare effects of technological change under different insurance systems. Each parameter change can be analysed separately, which allows a better qualitative and quantitative measurement of technological change as compared with the previous literature.

This analysis follows the models of demand inducement in a profit maximization context as opposed to those which limit inducement by a disutility of acting against the best interest of the patient[28]. The setup is related to Dranove's (1988) approach where the physician recommends treatment and is in his inducement limited by the patient's willingness to consent. However the focus of this study is on the regulation of reimbursement schemes, insurance contracts, and technological change. I study reimbursement schemes and an insurance contract which are set by the state whereas Dranove considers a situation in which the physician sets a price. Also, the inclusion of patient copayment adds the problem of ex-post moral hazard to the setup. Policy recommendations with respect to reimbursement and insurance contracts can thus be given. The similarities and differences to Dranove's results will be mentioned in due course. Some of the results are also related to the literature on credence goods[29]. Most studies on markets for credence goods focus

[27]See, for example, Goddeeris (1984a, 1984b).

[28]See McGuire (2000) for the distinction between these approaches.

[29]See De Jaegher and Jegers (2001) for a distinction between the game-theoretic approach of cheap-talk versus the credence goods approach in the economics of information, which both have an interpretation in the context of supplier-induced demand.

on price competition by providers (see Emons (1997), Wolinsky (1993, 1995)). These models stress the role of customer search for disciplining physicians' fraudulent behaviour and analyse the extent of fraud in equilibrium. Emons (2001) considers a credence goods monopolist who chooses capacity and prices for diagnosis and repair. As long as either capacity or services are observable the fraudulent expert problem is solved. More closely related are the studies of markets where prices are regulated. Pitchik and Schotter (1987, 1993) look at regulated markets where prices are set and physicians recommend treatment. Again, the threat of rejection of recommendation in the monopoly case and of search for second opinion in the competitive setting prevent the physician from cheating always. Finally, Sülzle and Wambach's (2002) analysis of insurance in a market for credence goods is related to my study as they analyse the effects of insurance on fraudulent behaviour. I also incorporate an insurance parameter which can explicitly be used by the regulator to enhance welfare. Therefore, the analysis captures, though in a rudimentary way, both informational problems of supplier-induced demand and ex post moral hazard.

The rest of the chapter is structured as follows: In section 3.2, the model is set up and technology choice and treatment pattern in equilibrium are analysed. Optimal reimbursement with respect to technology choice is studied. Section 3.3 deals with the physician's incentives to adopt an innovation, and with the optimal reimbursement of innovations. Section 3.4 discusses several extensions of the model, and section 3.5 concludes.

3.2 Technology choice and treatment pattern

This section sets up the model and looks at technology choice and treatment pattern in equilibrium. In the first place, the state insurer offers an insurance contract to the patients and a remuneration contract for diagnosis and all available types of treatment to the physician. It is assumed that the physician is a monopolist. His contract specifies for each treatment technology i a fixed payment R_i and a net share s_i in total costs of diagnosis/treatment that he has to bear. There is only one type of diagnosis which comes at a cost D and is reimbursed according to a scheme (R_D, s_D). Diagnosis is assumed to be independent of the treatment technology. These parameters allow to model the different contracts which are used in the health care sector. R_D is a fee per visit which the physician may receive. If $0 < s_i \leq 1$ the physician has to undergo cost sharing for treatment, with $s_i = 1$ implying the case of a fee for service. If $s_i = 0$ his contract is characterised by cost reimbursement. The fixed payment R_i allows variety in overall profitability of technologies even under cost reimbursement.

The insurance contract for the patients contains a premium P and a copayment rate[30] a with $0 \leq a < 1$. The patients accept or reject the insurance contract in the next stage, whereas the physician decides on participation. Then, the physician chooses the range of technologies which he wants to invest into and consequently can offer to the patients in case of illness. Each technology requires a fixed investment cost F_i which can be thought of as the purchase of new machinery or the acquisition of knowledge on the correct application[31]. If the technology is applied it is characterised by four parameters: monetary costs of treatment C_i, non-monetary costs of treatment to the patient N_i, non-monetary costs of applying the technology to the physician K_i, and a probability parameter $0 \leq p_i \leq 1$ which indicates the probability that health is completely restored in the case of applying the technology to an ill patient. With probability $1 - p_i$ the illness level is unchanged after treatment.

It is assumed that each patient can observe the range of technologies which the physician has invested into and consequently could offer him in case of illness. Furthermore, patients are perfectly informed on reimbursement schemes, on the costs and on the effectiveness of all available technologies. When falling ill, patients are uniformly distributed between 0 and 1. Their type r is drawn by nature from this distribution and indicates the (perceived) severity of their illness. r is the probability of being ill and can be interpreted as the symptoms that the patient himself observes before deciding whether to consult a doctor or not. It could for example be the severity of chest pain. The physician then has to determine whether this is caused by a severe heart problem or by less severe tensions. At the same time, r is assumed to be a common prior because the patient tells the physician all the symptoms that he experiences when he visits him for the first time[32]. When a patient visits a doctor, the physician not only learns the type r but also diagnoses the

[30] The restrictions imposed ($a \geq 0$, $s_i \leq 1$) are necessary due to truthtelling requirements which have been discussed elsewhere (see Ma and McGuire (1997)).

[31] These fixed costs do not enter the calculations in this study because they would add complexity without being at the centre of the analysis. They merely represent a signal to the patient for the availability of the technology. In Germany, e.g., investment costs in hospitals are born by the regional government as opposed to costs of usage which the hospital bears. The fixed costs could also be seen as incorporated in the overall reimbursement (neglecting scale effects).

[32] Of course, one could also interpret the symptoms of a patient as his private signal on the illness state (see Dranove (1988)). But it is reasonable to assume that the doctor's signal after diagnosis is better than

patient and finds out whether he is truly ill or not[33]. A separation between diagnosis and treatment is not possible. The patient cannot observe the result of the diagnosis but only receives a signal when the physician recommends "treatment with technology T_i" or "no treatment". Based on this recommendation the patient decides whether to consent to treatment or whether to refuse treatment. If "no treatment" is recommended the patient can only accept this decision as here the monopoly case is assumed. There is thus no scope for a second opinion in this model.

Let us briefly recapitulate the stages of the game: In the first stage, contracts are offered to patients and physician and the parties decide whether to accept these contracts. Thereafter, the physician invests into a range of technologies. In the next stage of the game, nature decides whether patients fall ill or not, and patients learn their type r by experiencing certain symptoms. Subsequently, the patients decide whether they visit the physician or not. Finally, if they visit the physician a signalling subgame is played in which the physician diagnoses the patients and then recommends treatment with a certain technology or no treatment. Each patient consents to treatment or not. The game is now solved by backward induction.

The signalling subgame

As a first step, the signalling subgame once the patient has fallen ill and decided to visit the physician is analysed. I will concentrate on the equilibria in pure strategies, i. e. on two pooling equilibria where the physician independently of the true health status of the patient recommends "treatment" $(T_i(r), T_i(r))$ or "no treatment" (NT, NT) respectively, and on two separating equilibria $(NT, T_i(r))$ and $(T_i(r), NT)$. The first recommendation in brackets refers to the illness case, the second one to the no-illness case. Under reasonable assumptions, the two equilibria (NT, NT) and $(NT, T_i(r))$ are not stable. In line with the literature (see e. g. Pitchik and Schotter (1987)) it is assumed that a liability rule is in place which prevents the physician from recommending "no treatment" in case the patient is diagnosed as ill. The liability rule works through a punishment which makes the physician deviate in either case. One could also interpret this punishment as the outcome

the patient's signal on the illness state. Therefore a limit case is analysed which captures the underlying effects very clearly and allows us to neglect an otherwise necessary screening of the patient's type.

[33] A more realistic model would assume that the patient experiences symptoms which are in turn the expected value of the distribution of the patient's true type over a continuum of conditions. This true type is then diagnosed by the physician. This assumption would not change the following analysis.

of medical ethics. The physician suffers from not treating an ill person whereas he does not suffer from treating a healthy person. Similarly to Ma and McGuire (1997) medical ethics is represented by a lower bound on the health benefits that a physician is willing to provide to a patient. In order to destabilize these equilibria the punishment does not have to be very harsh - it just needs to be less profitable / more costly than treating the patient.

Let us consider the pooling equilibrium $(T_i(r), T_i(r))$. By getting a treatment recommendation the patient does not learn anything about his type, he still assumes that with probability r he is ill. He therefore consents to the treatment recommendation if

$$\begin{aligned} & r \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)] + \\ & (1 - r) \cdot U(Y - P - a(C_i + D) - N_i) \geq \\ & r \cdot U(Y - P - a \cdot D - \epsilon)] + (1 - r) \cdot U(Y - P - a \cdot D) \end{aligned} \tag{3.1}$$

The left-hand side of this equation is the expected utility from receiving treatment. With probability r, the patient is ill and correctly receives treatment. The treatment is successful with probability p_i, and he then has a utility U which depends on his initial income Y minus the premium P, the share in costs of treatment and diagnosis he has to bear $a \cdot (C_i + D)$, and the non-monetary costs of treatment N_i. With probability $1 - p_i$ the treatment is not successful and he still suffers from the illness shock ϵ. ϵ measures state-dependent disutility from illness in monetary terms. In case he receives treatment although he is actually healthy (probability $(1 - r)$) he still has to incur the costs of treatment. The right-hand side of the equation is the expected utility from rejecting treatment. Again, with probability r the patient is ill and suffers from the illness shock, with probability $1 - r$ he is healthy and only incurs the costs from having been diagnosed.

If this condition holds then the physician does not have an incentive not to recommend treatment as long as treating the patient does not come at a cost[34], i. e. $\Pi_i = R_i - s_i \cdot C_i - K_i \geq 0$. If the parameter constellations are such that the patient does not consent (i. e. condition (3.1) does not hold) then $(T_i(r), T_i(r))$ is not an equilibrium. It is assumed that the physician slightly prefers to recommend "no treatment" (in the case of a healthy patient, where there is no punishment for recommending "no treatment") to recommending treatment and then being rejected. One could think of the explanation of

[34]Throughout the analysis it is assumed that reimbursement schemes are not changed after the physician has invested into the technologies.

treatment coming itself at a little cost.

In the separating equilibrium $(T_i(r), NT)$, the patient when being recommended treatment knows that he is ill. He therefore consents to treatment if

$$p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon] \geq U(Y - P - a \cdot D - \epsilon) \tag{3.2}$$

The left-hand side here is the expected utility from being treated in case of illness, whereas the right-hand side is the utility from not being treated when ill. But this can only be an equilibrium if treating the patient does not make profits, i. e. $\Pi_i = R_i - s_i \cdot C_i - K_i \leq 0$, because otherwise the physician has an incentive to deviate and to always announce "treatment". If the parameters are such that the patient chooses not to consent then the equilibrium is stable if not recommending treatment in case of illness is punished harshly as described above. Of course it has to be borne in mind that this will affect the patient's decision to visit the physician in the first place as will be discussed next. But before that the results which have been obtained so far should be summarized.

Lemma 5 *Equilibria "Signalling subgame"*
Under reasonable assumptions, there exist only two equilibria of the signalling subgame in pure strategies:
Under condition (3.1) and $\Pi_i \geq 0$, *a pooling equilibrium exists in which the physician recommends treatment independently of the true health status of the patient. The patient consents to treatment.*
If $\Pi_i \leq 0$, *a separating equilibrium exists in which the physician recommends treatment if the patient has been diagnosed as ill and does not recommend treatment if the patient has been diagnosed as healthy. Under condition (3.2), the patient consents to treatment. If the condition does not hold, the patient does not consent to treatment.*

The decision to visit the physician

Let us now proceed in the analysis of the game by backward induction. Before the signalling subgame is played the patient decides whether to visit the physician or not. He compares the payoffs of the signalling subgame with the option of not visiting the physician at all. Therefore, the above conditions must be modified. The patient prefers to be treated in a pooling equilibrium to not visiting the physician if

$$\begin{aligned}&r \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)] +\\&(1 - r) \cdot U(Y - P - a(C_i + D) - N_i) \geq\\&r \cdot U(Y - P - \epsilon) + (1 - r) \cdot U(Y - P)\end{aligned} \tag{3.3}$$

He prefers to be treated in a separating equilibrium to not visiting the physician if

$$\begin{aligned}&r \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)] +\\&(1 - r) \cdot U(Y - P - a \cdot D) \geq\\&r \cdot U(Y - P - \epsilon) + (1 - r) \cdot U(Y - P)\end{aligned} \tag{3.4}$$

The right-hand side of both equations is the expected utility of not visiting the physician. Clearly, the costs from diagnosis do not have to be born in this case. The left-hand side is the ex ante expected utility from being treated in a pooling and separating equilibrium, respectively. Obviously, condition (3.3) is more restrictive than condition (3.1) in the pooling equilibrium of the signalling subgame, as long as $a > 0$ and $D > 0$. If $a = 0$ and/or $D = 0$ both conditions coincide.

Rearranging condition (3.4) yields

$$\begin{aligned}&p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon] \geq\\&U(Y - P - \epsilon) + \tfrac{1-r}{r}[U(Y - P) - U(Y - P - a \cdot D)]\end{aligned} \tag{3.4'}$$

From this it can be easily seen that condition (3.4) is more restrictive than condition (3.2) in the separating equilibrium of the signalling subgame, as long as $a > 0$ and $D > 0$. If $a = 0$ and/or $D = 0$ both conditions coincide. Conditions (3.3) and (3.4) thus hold for a smaller or equal proportion of patients than conditions (3.1) and (3.2), respectively. The intuition behind these results is that bearing costs from being diagnosed will prevent the healthier patients from visiting the physician in the first place. The analysis of this stage of the game can be summarized as follows:

Lemma 6 *Equilibria "Decision to visit the physician"*
Under reasonable assumptions, there exist three equilibria in pure strategies of the subgame at the stage of the patient's decision whether to visit the physician:
Under condition (3.3) and $\Pi_i \geq 0$*, a pooling equilibrium exists in which the physician recommends treatment independently of the true health status of the patient. The patient visits the physician and consents to treatment.*
Under condition (3.4) and $\Pi_i \leq 0$*, a separating equilibrium exists in which the physician recommends treatment if the patient has been diagnosed as ill and does not recommend treatment if the patient has been diagnosed as healthy. The patient visits the physician and consents to treatment.*
In any other case the patient does not visit the physician and the game ends.

Choice of profitable technologies

Building on these results it can now be analysed how the physician decides on his technology range before the patient's decision to visit him. If reimbursement is such that all available technologies make a profit when they are applied then Lemma 6 states that a separating equilibrium does not exist. That means that whenever they visit the physician and are recommended treatment patients are "pooled" and do not acquire further information on their health status. Accordingly, they visit the physician and consent to the treatment recommendation with technology i if condition (3.3) holds. For each technology, a cutoff r_{ci}^p can be determined when (3.3) holds with equality:

$$\begin{aligned} &r_{ci}^p \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)] + \\ &(1 - r_{ci}^p) \cdot U(Y - P - a(C_i + D) - N_i) = \\ &r_{ci}^p \cdot U(Y - P - \epsilon)] + (1 - r_{ci}^p) \cdot U(Y - P) \end{aligned} \tag{3.5}$$

All patients with symptoms indicating an illness probability of r_{ci}^p or higher will visit the physician and consent to being treated with technology i. All patients with less severe symptoms will not visit the physician. Now the physician can rank all available technologies according to their reimbursement $\Pi_i = R_i - s_i \cdot C_i - K_i$, starting with the most profitable one. This technology implies a cutoff r_{c1}^p and will be offered to all patients of type $r_{c1}^p \leq r \leq 1$. Then the physician works through his ranking, looking at the second most profitable one. If it induces a cutoff r_{c2}^p which is below r_{c1}^p, then the physician offers it to all patients of type $r_{c2}^p \leq r < r_{c1}^p$. If the cutoff is above r_{c1}^p the technology is not applied - and consequently not invested into, either - and the physician analyses the next technology on his list in the same manner. This procedure continues until only the healthiest fraction of patients does not receive treatment because they reject overtreatment for any available profit making technology[35]. These patients prefer no treatment to being treated in a pooling equilibrium. As long as diagnosing a patient does not make a loss (or more exactly, as long as diagnosing and treating the range of patients below the cutoff does not make a loss in expectation) the physician would want to treat also the healthier range of patients but as he cannot credibly commit not to "pool" them these patients prefer not to visit him. Consequently, the range of patients below the cutoff receives too little treatment whereas the less healthy patients - who are willing to be treated irrespective of their true health status - receive too much treatment when all available treatments are profitable for the physician. In Figure 3.1, patients' expected utility without and with treatment with two technologies is depicted[36]. Technology 2 induces a lower cutoff and yields a higher expected utility than technology 1 for all patients who are willing to be treated with this technology. If only these two technologies are on the market then it is socially desirable that the physician only adopts technology 2 and treats all patients with a severity of illness $r \in [r^p_{c2}; 1]$. Yet if technology 1 is more profitable than technology 2 then the physician is going to adopt both technologies and patients with an illness severity of $r \in [r^p_{c1}; 1]$ will be treated with this less efficient technology instead.

Choice of zero-profit technologies

Let us now consider a situation where not only profitable technologies exist but also some which break even, i. e. where $\Pi_i = R_i - s_i \cdot C_i - K_i = 0$. Lemma 6 shows that for technologies which break even a pooling as well as a separating equilibrium may exist in which patients receive treatment only if they really need it. Looking at conditions (3.3) and (3.4) it becomes clear that in this situation, whenever a pooling equilibrium exists then also a separating equilibrium exists but not the other way around. When condition (3.4) holds with equality a second cutoff r^s_{ci} is induced which indicates the range of patients

[35]From condition (3.3) it is clear that a cutoff $r^p_{ci} = 0$ cannot be achieved as long as there does not exist a technology which comes at neither monetary nor non-monetary costs of treatment and diagnosis. Only in this case the full range of patients could be treated in a pooling equilibrium.

[36]It is shown in appendix B1 that the slope of the expected utility without visiting the physician (outside option) is always steeper than the slope of the patients' expected utility under treatment. Otherwise conditions (3.3) and (3.4) are not satisfied.

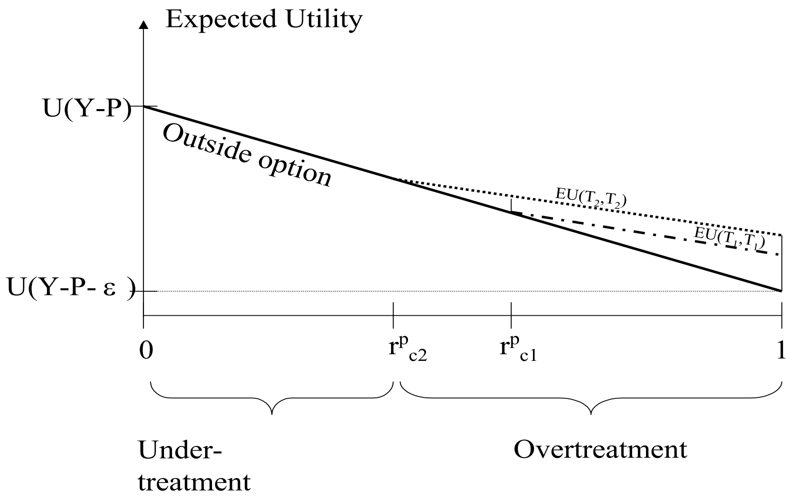

Figure 3.1: Choice of profitable technologies

willing to visit the physician in a separating equilibrium:

$$r^s_{ci} \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)]+$$
$$(1 - r^s_{ci}) \cdot U(Y - P - a \cdot D) =$$
$$r^s_{ci} \cdot U(Y - P - \epsilon)] + (1 - r^s_{ci}) \cdot U(Y - P) \tag{3.6}$$

According to conditions (3.3) and (3.4) it must be that $0 \leq r^s_{ci} \leq r^p_{ci}$. For the same technology, the least healthy patients accept overtreatment ($r^p_{ci} \leq r \leq 1$), the intermediate (and least healthy) patients accept the correct amount of treatment ($r^s_{ci} \leq r \leq 1$), and the healthiest patients do not accept either equilibrium and therefore do not visit the physician at all ($0 \leq r < r^s_{ci}$). Only if diagnosing does not imply a cost to the patient then the whole range of patients accepts a separating equilibrium and $r^s_{ci} = 0$ can be achieved. As long as $a > 0$ and $D > 0$ the healthiest range of patients abstains from visiting the physician[37], $r^s_{ci} > 0$. Figure 3.2 depicts a situation where all technologies yield zero profit. The physician chooses the technology which induces the lowest cutoff,

[37] A comment should be made on the equilibrium analysis in this section. Comparing conditions (3.1) and (3.3) and conditions (3.2) and (3.4), respectively, makes clear that there exists a fraction of patients for each technology who would not accept treatment with this technology ex ante, but would do so once they have been diagnosed. This implies that the physician would want to deviate and offer them the more profitable technology once they have shown up. Anticipating this deviating behaviour, the patients prefer not to visit the physician. Consequently, certain fractions of patients do not receive treatment.

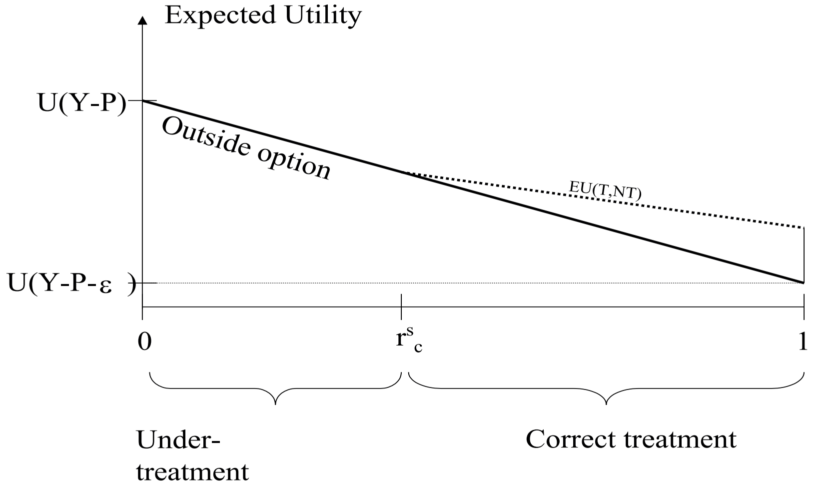

Figure 3.2: Choice of a zero-profit technology

and all patients with symptoms $r \in [r_c^s; 1]$ receive treatment with this technology.

Choice of unprofitable technologies

As long as profitable and budget breaking technologies are on the market a loss-making technology will not be chosen by the physician[38]. Even if a loss-making technology induced a lower cutoff than budget breaking technologies and if at the same time diagnosing patients was so profitable that in expectation the physician would make a profit could he not credibly commit not to induce a rejection of the treatment recommendation once the patients have turned up by offering them a profitable technology.

If all technologies are loss-making and if losses are cross-subsidized via the revenues from diagnosis (fee for visit) then the physician chooses one technology which maximizes his income by trading off the differential of fee for visit and treatment costs and the share of patients who turn up[39]. Investing into more than one technology again does not make sense since the physician cannot commit not to induce a rejection of the recommendation.

This case will be neglected here because it would add complexity to the results without providing further insights into the physician's technology choice .

[38]Kane and Manoukian (1989) analyse the case of cochlear implants. They show that many hospitals ration the availability of the device to Medicare patients because payment for the device remains well below its average cost.

[39]He maximizes $(R_D + \Pi_i) \cdot (1 - r_i)$.

Lemma 7 *Technology choice*
The technology choice by the physician depends on the profitability of treatment and acceptance by patients:
If $\Pi_i > 0 \forall i$*, the physician chooses technologies starting with the most profitable one and adding decreasingly profitable ones which induce lower cutoffs.*
If $\Pi_i = 0$ *for at least one* i*, the physician chooses one technology which induces the lowest cutoff* r^s_{ci}*.*
If $\Pi_i < 0 \forall i$*, the physician chooses one technology which induces the most profitable trade-off of fee for visit, treatment costs and share of patients who turn up. If at the same time* $\Pi_j \geq 0$ *for at least one* j*, no unprofitable technology is chosen.*

Optimal reimbursement
How can the state (the insurer, respectively) implement the optimal choice of technologies and treatment pattern? Figure 3.3 shows an example where five technologies are on the market which induce decreasingly lower cutoffs. The state would want only technologies 3, 4, and 5 to be used and offered to patients. Each patient should receive treatment with the technology which provides the highest expected utility for his type. Technologies 1 and 2 are clearly dominated by the other treatments. Technology 3 should be given to the patients with most severe symptoms but only down to the point where technology 4 provides higher expected utility (i.e. not down to cutoff r_{c3}) and so on. If the state can contract on the technology selection which the physician has to invest into then the first best is achievable. By offering zero-profit reimbursement for all treatment technologies the state achieves that the physician is indifferent between treatments and suggests the correct treatment for everyone. Participation incentives can be given via a fee for visit (R_D).
If the state cannot contract on the technology selection then the first best is not achievable by the simple contracts under consideration. As long as treatment is unprofitable or makes zero profits the physician chooses only one technology. A variety of different technologies is only chosen if treatment is profitable. The state can influence the right choice of treatments by attributing the right ranking of profitable reimbursement. In the example, technology 3 should be more profitable than technology 4, followed by technology 5. Least profitable should be technologies 1 and 2. It is important to note, though, that the physician will then correctly select technologies 3, 4, and 5 but will treat patients according to their symptoms and not to the diagnosis he makes. Furthermore, a larger

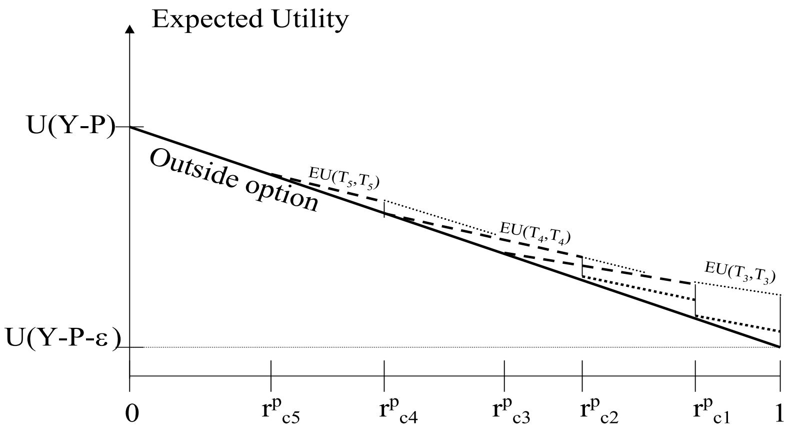

Figure 3.3: Optimal technology choice

share of patients receives treatment with the more profitable technologies because they are given to all patients down to the respective cutoff and not down to the point where the next technology induces a higher expected utility.

These results are summarized in the following

Proposition 3 *Optimal reimbursement with respect to technology choice*
If the technology choice is contractible the physician receives a fee for visit and zero-profit reimbursement for all technologies. The first best is achieved.
If the technology choice is not contractible the first best cannot be achieved. The correct choice of technologies can be implemented by the adequate scheme of profitable reimbursement yet patients are treated according to their symptoms and irrespective of their true health state.

It is interesting to note that an ex-post change of reimbursement schemes may have an important impact on both the selection and usage of technologies. There is no change if relative profitability (i.e. the ranking of technologies according to their profitability) remains unchanged but as soon as the ranking of profitable technologies changes technologies may be offered to different groups of patients or may not be used at all any more. Another important result from the above analysis should be established:

Corollary 2
Cost-effective diagnosis should be free.

Only in this case will a cutoff $r^s_{ci} = 0$ be achieved[40]. Undertreatment is thus eliminated.

3.3 Adoption of innovations

Having studied the equilibrium strategies and technology choice of the game, this section focuses on the physician's decision whether to adopt a new technology or not when it enters the market. A reimbursement scheme is announced before the technology is made available on the market. It is assumed throughout this section that premia are not adjusted as a response to the physician's decision.

If the technology choice is contractible then the state can simply add a new cost effective technology to the set of treatments which the physician has to invest into if he wants to participate.

If the technology choice is not contractible then an innovation may be adopted by the physician only under certain conditions. If all technologies are unprofitable the innovation may replace the one which has been previously chosen if it improves the trade-off between fee for visit, treatment costs and share of patients who turn up. If all technologies make zero profit it may replace the one which has been previously chosen if it lowers the cutoff down to which it can be used.

If technologies are profitable then the adoption decision is more complex: The discussion has made it clear that the new technology in order to be adopted by the physician has to fulfill at least one of the following two criteria: It has to be more profitable than a currently used technology which induces the same cutoff and / or it has to induce a lower cutoff than a currently used technology which is equally profitable.

The direct effect on profits is straightforward. It certainly depends on the announced reimbursement scheme $(R_i,\ s_i)$. But profits $\Pi_i = R_i - s_i \cdot C_i - K_i$ are also directly influenced by the monetary costs of the new technology C_i and the non-monetary costs

[40]One should bear in mind though that this result holds in the partial analysis when effects on the premium P are neglected, and under the assumption that diagnosis is desirable even for the healthiest patients.

to the physician K_i. Let us assume for the moment that profitability Π_i can be directly controlled by the state.

Profits may also change indirectly if changes of the technology parameters influence the cutoff down to which this technology can be used. If the cutoff decreases (increases) a larger (smaller) fraction of patients can now be treated with the new technology. Thus, optimal reimbursement takes direct and indirect effects into account.

Change in one technology parameter

The state (insurer) first has to take the decision which reference technology should be replaced by the innovation. If only one technology parameter changes this is obvious: The more cost effective new technology may reduce (monetary or non-monetary) costs or increase effectiveness of treatment as compared to the currently used reference technology. This type of innovation is likely to take place in the form of process innovations, for example. These changes have an impact on the patient's expected utility and therefore on the cutoff down to which this technology may be offered. The physician wants the cutoff to be as low as possible as this increases the range of patients for whom this technology can be applied. This will increase his profits as a larger fraction of patients is treated with a more profitable technology and a lower fraction with a less profitable one (or left without treatment, respectively).

Since we assume that $a \geq 0$ the effect of a change in monetary costs on the cutoff is $\frac{\partial r_{ci}}{\partial C_i} \geq 0$. That means that lower monetary costs reduce the cutoff for patients' consent to pooling. Intuitively, a reduction in monetary costs increases the expected utility from treatment for the patient which induces him to consent to pooling even with a lower illness probability than before the change[41].

Non-monetary costs of treatment N_i and the probability p_i that health is completely restored also influence the pooling respectively separating cutoff, that is the share of patients who are offered this technology. Whereas a decrease in N_i reduces the cutoff (increases the share), an increase in p_i has the same effect[42]. Intuitively, expected utility from treatment increases when non-monetary costs go down or when the likelihood of

[41] This is in line with Dranove's (1988) results: If the price of treatment decreases, patients are more willing to consent.

[42] Proofs in appendix B2. Superscripts p and s are suppressed in this section because the analysis applies to both the pooling and the separating cutoff. It can also be shown that with only one changed technology parameter, expected utility increases for all patients treated with this technology.

success increases. The physician will therefore adopt technologies which reduce N_i and increase p_i.

Lemma 8 *Adoption of innovations: One changed technology parameter*
If only one technology parameter is improved ($C_i \downarrow$ or $N_i \downarrow$ or $p_i \uparrow$) the cutoff decreases as compared with the reference technology. Given equal profitability, the physician adopts the innovation.

Simultaneous parameter changes

For some process innovations and most product innovations several technology parameters will change simultaneously as compared with the reference technology. Expected utility for the marginal patient at the cutoff r_{ci} increases and the cutoff goes down if[43]

$$(a \cdot dC + dN) < \frac{r_{ci} \cdot [U(Y-P-a \cdot C_i - a \cdot D - N_i) - U(Y-P-a \cdot C_i - a \cdot D - N_i - \epsilon)]}{r_{ci} \cdot [p_i \cdot U'(Y-P-a \cdot C_i - a \cdot D - N_i) + (1-p_i) \cdot U'(Y-P-a \cdot C_i - a \cdot D - N_i - \epsilon)] + (1-r_{ci}) U'(Y-P-a \cdot C_i - a \cdot D - N_i)} \cdot dp \tag{3.7}$$

Some interesting points can be established here. First, the state has to carefully consider which technology is the reference technology to be replaced. For example, a technology which increases effectiveness of treatment (possibly at higher costs) is valued more highly by relatively ill patients than by healthy patients. Thus a change from T_i to T_j with $p_j > p_i$ may increase expected utility of treatment only for rather ill patients[44]. Second and as a consequence, a cost-effective innovation may increase or reduce the cutoff as compared with the reference technology. If effectiveness of treatment increases ill patients may be made better off but at the same time the cutoff may increase[45]. Such a situation is depicted in figure 3.4 where technology 1 is the imcumbent technology and technology 1' is the innovation which makes the least healthy patients better off.

[43] Marginal effects on the cutoff of changes in one technology parameter are simply added up here.

[44] Formal proof in appendix B3.

[45] It is neglected throughout the analysis how cost effectiveness is defined in this setup. A reasonable definition would be that technology T_j is said to be more cost effective than T_i if there exist patients of type r who have a higher expected utility from treatment j than from i when fully internalizing the costs of treatment. Here it becomes clear why it is necessary to make a cost effectiveness analysis before a technology may enter the market: Patients might also be willing to accept cost ineffective technologies because they only internalize a share a of possible cost increases. Furthermore, an increase in cost-effectiveness may reduce patients' expected utility for all r. This can happen if the change in cost effectiveness implies a reduction in monetary costs C. Since patients benefit only partially from this cost

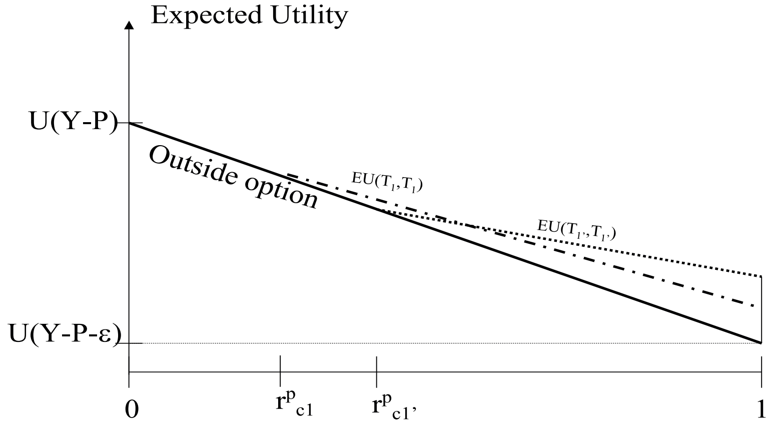

Figure 3.4: Cutoff increasing innovation

Lemma 9 *Adoption of innovations: Several changed technology parameters*
If various technology parameters change simultaneously a new cost effective technology may increase or reduce the cutoff. Given equal profitability, the physician adopts the innovation only in the latter case.

Consequently, if the cutoff decreases the technology can be fitted into the existing scheme of profitable reimbursement by making it equally profitable as the reference technology. Yet if the cutoff increases it has to be more profitable than the reference technology, thereby also changing the reimbursement scheme for all technologies which are more profitable. Let us summarize the results in the following

Proposition 4 *Optimal reimbursement of innovations*
If the technology choice is contractible the innovation is added to the set of necessary technologies and makes zero profit.
If the technology choice is not contractible optimal reimbursement trades off direct and indirect effects on profits. If the cutoff increases the innovation has to be more profitable than the reference technology and the reimbursement scheme for more profitable technolo-

reduction possible increases in non-monetary costs or reductions in effectiveness may outweigh the gains for the individual patient.

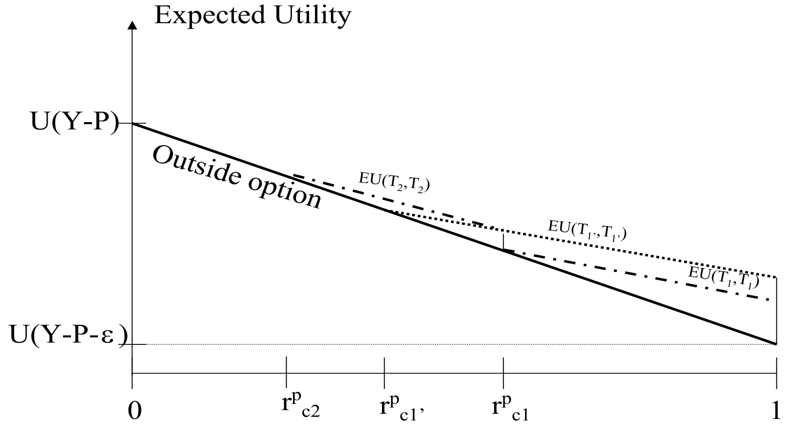

Figure 3.5: Welfare effect of changed cutoff

gies has to be adjusted. If the cutoff decreases the innovation can be made equally profitable as the reference technology.

Some more insights can be gained from the analysis.

Corollary 3
If the physician adopts a new technology which changes the cutoff down to which this technology is used, then in general some patients win and some lose from this change.

The proof can be made graphically. Consider figure 3.5 where technology 1' replaces technology 1. All patients who previously received treatment with technology 1 are now better off. But there is a fraction of patients with $r \in [r^p_{c1'}, r^p_{c1}]$ who are now worse off because the physician offers them the more profitable technology. A similar argument holds when the cutoff increases.

Corollary 4
Welfare losses induced through a change in cutoffs due to the adoption of innovations may be counteracted by technology specific patient copayments.

Throughout the analysis it has been assumed that only a small part of the insurance market is studied. Therefore, copayment a and premium P have been taken as given. A more complete setup would investigate into the optimality of the insurance contracts for the patients. Yet even in this partial analysis one can argue that technology specific copayments for innovations increase welfare by reducing the welfare losses induced through changes in the cutoff. If the cutoff is reduced through an innovation which increases effectiveness of treatment and reduces costs then an increase in patient copayment for this technology will counteract this reduction and thereby reduce the share of patients who are worse off through the treatment with this innovation. For innovations which reduce both costs and effectiveness or increase both costs and effectiveness a clearcut prediction for the optimal choice of copayment cannot be made.

3.4 Discussion and extensions

Taking premium adjustment into account

In general, the adoption of a new technology will have an effect on overall health care costs and therefore on the premium P which the insured individual has to pay to the insurer. Not only does the adoption of cost reducing technologies reduce premiums but also a more extended application of more or less costly technologies due to changes in other technology parameters may increase or reduce premiums. The analysis of optimal insurance contracts goes beyond the scope of this paper but one can show that incorporating changes in the premium P into the partial analysis conducted here does not change the qualitative results.

A change in the premium P is equivalent to a change in income as it reduces or increases an individual's resources in all states of the world. Consequently, the impact on the cutoff which resembles the willingness to accept treatment with a certain technology depends on the patient's attitude to risk, measured by his coefficient of absolute risk aversion. It can be shown that the cutoff down to which a technology can be used is unchanged by premium adjustments if the patient exhibits constant absolute risk aversion, but decreases with higher premium payments if the patient has decreasing absolute risk aversion. From this it can be concluded that for CARA utility functions, the physician's adoption incentives are unchanged. In the case of DARA utility functions, a new counteracting incentive makes the physician less prone to adopt less costly technologies since a decrease in insurance premiums will in addition increase all other cutoffs, therefore all technologies can only be given to a smaller range of patients.

For the welfare analysis, it has to be considered that both outside option and expected utility from treatment are changed when the premium changes. It still holds, though, that some patients gain and others lose through a changed cutoff and an adjustment in copayment for the new technology may reduce this externality.

Unobservable changes in profitability

In the analysis so far it has been assumed that the profitability of a technology can be directly observed and controlled by the state. The special structure of reimbursement therefore had no impact on the behaviour of the physician. Yet it is interesting to consider the situation where the profitability of a technology is not or only after costly monitoring observable for the state. It is a realistic scenario that non-monetary costs for the physician and monetary costs of treatment cannot be easily observed. They are particularly relevant parameters for process innovations at the level of the individual physician. The reimbursement of angioplasty is an example where the authorities had to adjust their perception of cost parameters: After a few years of reimbursement in a higher ranked DRG the procedure was downgraded to a less profitable DRG. On the other hand, effectiveness of treatment and non-monetary costs to the patient are likely to be well known through clinical trials before the admission of a new technology.

It is important to ask whether the physician has an incentive to report changes in technology parameters truthfully and whether the structure of reimbursement has an impact on this decision. With respect to the technology parameters p_i (effectiveness of treatment) and N_i (non-monetary costs to the patient) one can conclude that they do not directly influence the physician's profits $\Pi_i = R_i - s \cdot C_i - K_i$. But patient acceptance will increase if these parameters are improved, therefore the physician has a clear incentive to announce relevant changes. Reductions in non-monetary costs to the physician K_i will never be announced because improvements will increase the physician's profits if reimbursement remains unchanged. With respect to monetary costs C_i the reimbursement system matters. If reimbursement is independent of costs (e.g. in the case of a fixed employment) then the physician announces improvements in C_i since this increases patient acceptance. If the physician bears part of the monetary costs of treatment then he trades off increased acceptance by patients and higher profits if reimbursement is unchanged despite the cost savings. Consequently, under cost sharing has the physician a stronger incentive to invest into process innovations but more monitoring is needed to ensure that the savings spill

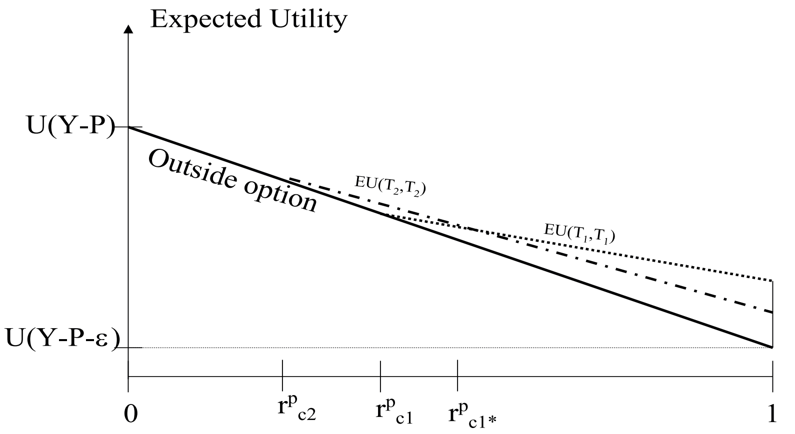

Figure 3.6: Competition of providers - two available technologies

over to patients. Under cost reimbursement, there will be less innovation at the level of the individual physician, but at the same time less monitoring with respect to the profitability of treatments is needed.

Competition of providers

The analysis has been conducted for the case of a monopoly provider. It is interesting to ask in how far competition between providers would change the results. The following discussion will consider a second physician but could be extended to $n < \infty$ physicians. It is a reasonable assumption to consider a limited number of physicians among whom a patient makes a choice.

It is assumed that the physicians simultanously choose the technologies which they can offer to patients. If both physicians offer a certain technology then patients who want to be treated with this technology distribute equally i.e. each physician gets half the share of patients for this treatment. In equilibrium, patients know which treatment they receive before visiting one of the physicians, therefore the changed outside option (second opinion) can be neglected and it suffices to analyse the providers' technology choice. Consider an example where two technologies are on the market, as in figure 3.6. Obviously, as long as treatment is not unprofitable, both physicians always have an incentive to adopt

	1/2	2
1/2	$(1-r^p_{c1})\cdot\frac{1}{2}\cdot\Pi_1+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$ $(1-r^p_{c1})\cdot\frac{1}{2}\cdot\Pi_1+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$	$(1-r^p_{c1*})\cdot\Pi_1+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$ $(r^p_{c1*}-r^p_{c1})\cdot\Pi_2+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$
2	$(r^p_{c1*}-r^p_{c1})\cdot\Pi_2+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$ $(1-r^p_{c1*})\cdot\Pi_1+(r^p_{c1}-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$	$(1-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$ $(1-r^p_{c2})\cdot\frac{1}{2}\cdot\Pi_2$

Table 3.1: Payoffs two physicians - two strategies

technology 2 because they can then offer treatment to an additional share of patients without losing others. On the other hand, it may well be profitable to only offer technology 2 and not to adopt technology 1. If, for instance, the first physician invests into both technologies, it may pay the second physician to adopt only technology 2 because he is then able to attract all patients with severities $r \in [r^p_{c1}; r^p_{c1*}]$ who would be treated with technology 1 if they visited the first physician but have a higher expected utility from treatment with technology 2. Whether such a specialization is an equilibrium depends on the profitability of both technologies and on people's preferences for technologies. The latter condition can be seen from the relative position of r^p_{c1*}. If r^p_{c1*} is close to r^p_{c1} then only a few people among those who accept technology 1 would prefer technology 2. The technologies are not close substitutes. If on the other hand, r^p_{c1*} is close to 1 then many people among those who accept technology 1 actually prefer technology 2. The technologies then are close substitutes. The formal analysis makes these effects clearer.

Table 3.1 shows the payoffs for both physicians depending on the simultaneous choice of strategies 1/2 (choose technologies 1 and 2) and 2 (choose technology 2 only). If physician 1 chooses 1/2 then physician 2 chooses 1/2 if

$$\Pi_1 > \frac{2\cdot\Pi_2\cdot(r^p_{c1*}-r^p_{c1})}{(1-r^p_{c1})} \tag{3.8}$$

If physician 1 chooses 2 then physician 2 chooses 1/2 if

$$\Pi_1 > \frac{\Pi_2\cdot(1-r^p_{c1})}{2\cdot(1-r^p_{c1*})} \tag{3.9}$$

Four cases can now be distinguished:

- Both conditions hold:
  The only equilibrium is (1/2,1/2). This is the case if $r^p_{c1*} \to r^p_{c1}$ ("no close substitutes") and if technology 1 is much more profitable than technology 2.

- Condition 3.8 holds, 3.9 does not hold:
  There are two equilibria in pure strategies (1/2,1/2) and (2,2) as well as one equilibrium in mixed strategies. This is the case if $r^p_{c1*} \to 1$ ("close substitutes") and if technology 1 is much more profitable than technology 2.

- Condition 3.8 does not hold, 3.9 holds:
  There are two equilibria in pure strategies (1/2,2) and (2,1/2) as well as one equilibrium in mixed strategies. This is the case if $r^p_{c1*} \to r^p_{c1}$ ("no close substitutes") and if the profitability of technology 1 is very close to technology 2.

- Neither condition holds:
  The only equilibrium is (2,2). This is the case if $r^p_{c1*} \to 1$ ("close substitutes") and if the profitability of technology 1 is very close to technology 2.

It can thus be established that the profitability of both technologies should not diverge too much. If the technologies are close substitutes only technology 2 is adopted by both physicians which avoids overtreatment of a large fraction of patients with a less desired technology. If the technologies are no close substitutes a specialization equilibrium is achieved. This reduces overtreatment and may in addition save ressources for unnecessary double provision of technologies. Specialization is thus destabilized if the profitability of technologies diverges too much and if too many people prefer the lower ranked technology.

If all technologies make zero-profit then physicians try to maximize the share of patients who turn up. Therefore, both will adopt all available technologies in order to avoid loss of patients. A specialization equilibrium cannot be achieved.

If all technologies are unprofitable it still holds that no physician will invest into more than one technology because he cannot commit not to induce rejection once the patients have turned up. In a similar analysis as for the case of profitable reimbursement conditions can be derived under which specialization may occur yet the maximal number of available technologies is the number of physicians in the market.

3.5 Conclusion

Having studied the provider incentives for technology choice and adoption of innovations, it is interesting to come back to the example of heart attack treatment mentioned in the introduction. Part of the puzzle of increased treatment for heart attack patients may now be resolved. Angioplasty was introduced in the 1980s when bypass surgery was already a well established procedure. Bypass surgery is more invasive and thus more costly both in monetary and non-monetary terms. At the same time it is more effective than the application of angioplasty. One of the reasons why there was hardly any substitution of treatment of potential bypass patients through angioplasty can be concluded from the analysis: Angioplasty is acceptable for a wider range of patients. At the same time, bypass surgery is more profitable for physicians; therefore a similar share of patients as before received bypass surgery, and an additional fraction of patients who would not have accepted the more invasive procedure were treated with angioplasty. The steady increase in both procedures throughout the 1980s can be attributed to process innovations which reduced the cutoff and increased the share of patients willing to accept each treatment. Cutler and Huckman (2002) provide evidence for the 1990s that angioplasty during that period became a stronger substitute for bypass surgery due to learning and technological improvements over time. The authors find that both the expanded use of angioplasty and the substitution for bypass surgery improved quality for patients receiving revascularization. Whereas the quality effect is clearly predicted by this model, the substitution effect should according to this analysis only occur if financial incentives for providers changed during that time span. Focusing on the impact of the profitability of the procedures would be an interesting extension of the empirical investigations.

Future research should be dedicated to a thorough testing of the predictions of the model. Micro-level studies investigating into the treatment of certain conditions seem to be the most promising approach. Further theoretical investigations may complete and extend the analysis by studying optimal insurance contracts and further investigating into competition of providers and provider networks. Some studies suggest that a more integrated provision of care limits wasteful duplication of technologies by hospitals[46]. Also for the current analysis it seems plausible that managed care might solve some of the problems discussed. Allowing insurers to contract with individual physicians and hospitals implies a better control of technology choice and acquisition of innovations. But there is also evidence which indicates that increased competition within and among managed

[46]Baker and Phibbs (2002) show, for instance, that increases in HMO activity slowed the diffusion of neonatal intensive care units in the 1980s and 1990s in the US. Baker (2000) provides a similar study for the case of magnetic resonance imaging.

care organizations may counteract the desired reductions in health care costs[47]. And, as the following chapter shows, there may be strong opposition by physicians against more integrated systems of provision of care.

[47] See Bokhari (2001).

4 Medical Associations, Medical Education, and Training on the Job

4.1 Introduction

Relatively little attention has been given in the economic literature to professional associations in the health care market. This fact hardly reflects the important role which medical associations in particular play in many health systems worldwide. In Germany, for instance, medical associations are not only representatives of their members, they also control market entry for physicians, and serve as the main source of information on health and health care issues for politicians. For the outpatient sector, they distribute the revenue which they receive from public health insurers to the individual physician. The following analysis is mostly concerned with the role of medical associations in the outpatient sector in Germany. But also in other industrialized countries as for instance in Switzerland or France, medical associations have a strong position in the health care system[48]. The results of this study therefore can be seen in a more general context.

Medical associations were founded in Germany in 1931. This act followed a period of struggle between physicians and health insurers. The physicians had been demanding the abolition of individual contracting in favour of collective bargaining, free choice of providers by patients, and a fee-for-service system[49]. The state combined the allowance of forming a supply-side monopoly with several duties like providing adequate supply of health care services for the population. On the one hand, transaction costs of contracting could be saved by collective bargaining, on the other hand a monopoly on the side of the physicians was created with a strong influence on the health care market.

One of the issues raised during the ongoing discussion on health care reforms in Germany is whether public insurance companies should be allowed to return to individual contracting instead of transferring payments to medical associations which in turn distribute physician income. This would be a necessary first step for allowing more far reaching organizational reforms in the direction of managed care. The medical associations are very much opposed to this suggestion. Their members argue that this institution is necessary for maintaining

[48]See Zweifel and Eichenberger (1992) for a comparison of the degree of corporatism in various industrialized countries.

[49]http://www.kvberlin.de/Homepage/publikation/archiv/pk2002/pk041202.html

a high standard of quality of health care. It is not clear at first why quality control should be exerted by the associations. It seems more plausible that this argument is meant to disguise the fear of an expected loss of rents if the supply side monopoly is smashed. Yet this analysis shows that the medical associations use these tasks to redistribute rents within the association. Under a fee-for-service system, young physicians are exploited by putting the burden of high education costs and foregone earnings on them, whereas older generations of physicians reduce their costs by reducing training on the job. If the associations are allowed to pay out differentiated prices they will opt for a system of seniority wages and put other means of redistribution like long working hours for young physicians into place.

This study focuses on the impact of medical associations on the payment structure and quality of treatment in the health care market. It deals with the question whether the state (resp. the health authority) should set standards for medical education and medical training or delegate the implementation of these standards to a medical association. The approach takes up Shapiro's (1986) line of reasoning. He models occupational licensing as an input regulation, where a higher standard of training reduces the costs of producing a high quality product. In this model, physicians in the first period of their working life receive a license if they meet a specified standard of medical education. A higher standard increases the probability of adequate treatment. In later periods, training on the job is necessary to keep up this probability of correct treatment. Patients are randomly matched with physicians and experience a disutility from inadequate treatment. If the health authority sets standards for medical education and training on the job, these standards are chosen such that the expected utility of a representative patient is maximized under the constraint that a physician is reimbursed for his expenses in each period. This situation is compared to a setup in which physicians are organized in a medical association. This association is assumed to control market entry and to set itself standards for education and training. Since older generations of physicians constitute a majority within the medical association, this institution maximizes the lifetime income of older generations of physicians. Consequently, older generations of physicians dominate the younger generation and extract rents from them. If the medical association is allowed to pay out different prices to different generations of physicians, then a system of seniority wages, which can for instance be observed in the German inpatient sector, is set up. Standards for education and training do not differ from the situation in which these standards are set by the health authority because the efficient choice allows the biggest surplus to be distributed. If on the other hand, the medical association pays out the same price for treatment irrespective of the age of the physician[50], then rents can only be transferred from the younger to the older generation of physicians by imposing higher standards of education on the young generation, thereby reducing the need for training on the job. Consequently, training on the job will be lower and medical education higher under a system of medical associations if uniform prices are set, as in the German outpatient sector. For both setups one can conclude that the rents accrue only to the older generation of physicians who set up the system of medical associations. Still, switching back to individual contracting is rejected by the majority of physicians in later periods because they want to recover the investment in education which they have made in earlier periods of their working life. This explains the strong opposition which physicians show towards the introduction of new institutional arrangements.

Only very few studies have so far investigated into the role of professional associations in the health care market. Breyer et al. (2003)[51] pin down the role of medical associations to three broad areas according to the parties they interact with. Concerning their relation with patients, the associations may have a role in maintaining a high standard of quality in the health care market. Since health care is a credence good it may be in the consumers' interest to delegate quality assessment to a member of the profession. With respect to politicians, medical associations may serve several purposes. Votes may be gained by assuring high quality of care and by maintaining the income redistribution through health care which might be undermined by market forces. Furthermore, limiting access to health care markets has a rationing effect on the public health budget. Finally, medical associations may guarantee high incomes to their members.

The current approach focuses primarily on the first and last claim. Since health care is a credence good, the patient is even after consuming the good not able to verify whether

[50]In this analysis, a fee-for-service system is studied, although the payment system in the German outpatient sector is more complex due to a global budget and a system of point values which have certain strategic implications for physicians (see Benstetter and Wambach, 2001). These considerations go beyond the scope of this paper.

[51]See also Zweifel and Eichenberger (1992). Demange and Geoffard (2002) analyse the reform of incentive schemes under the political constraint that a majority of physicians has to be in favour of an organizational change. The focus of their analysis is on the feasibility and welfare implications of contractual arrangements and not on the role of professional associations.

he has been given a high or a low quality product. In particular, the physician's effort choice for producing high quality is in general not observable. The contracts usually under consideration in the health care market range from fee-for-service to capitation and prospective payment systems. The latter systems imply incentives for cost savings whereas fee-for-service is supposed to put a stronger focus on high quality of services. Yet a fee-for-service contract per se does not induce the choice of high effort for high quality by physicians. Various approaches therefore consider additional assumptions which make the physician work hard. In a broader sense, demand may depend on quality, and low quality may induce a negative patient response. A direct link is modelled, for example, in Ma (1994, 1998). But also more indirect links like patient search or reputation building (see, for example, Wolinski (1993) and Emons (1997)) may act as disciplining devices for physicians. On the other hand, there may be a lower bound on effort or quality choice due to intrinsic motivation or for liability reasons. The latter literature usually assumes that a liability standard can be costlessly implemented[52].

A similar assumption yet in a different context is made in the literature on occupational licensing. Leland (1979) assumes that physicians have differing innate abilities. In a lemons model he analyzes how the appropriate setting of a quality standard influences welfare. If the standard is set by a professional group it may be chosen too high due to the rents which then accrue to the group. Licensing as an input regulation is modelled by Shapiro (1986). It may impose a standard on the level of medical training chosen by physicians. Higher levels of training reduce the cost to produce a high versus low quality product and also the cost differential between both products. The product is sold to consumers who differ in their taste for quality. Licensing tends to benefit consumers who value quality highly at the expense of those consumers who do not. If training levels are observable then licensing is Pareto-worsening because it raises prices for low and sometimes also for high quality products by raising the costs of separation.

The current approach combines Shapiro's model of licensing as an input regulation with the credence goods characteristics of health care. Patients receive treatment and experience a disutility if the physician has not chosen the most adequate treatment, but this disutility cannot be verified ex post and therefore cannot be contracted upon. Yet the levels of medical education and training on the job chosen by physicians can be controlled either by the state/health authority or by the professional association of physicians, the

[52] See Danzon (2000) for an overview on medical malpractice and liability issues.

medical association. Higher standards increase the probability of adequate treatment. Physicians always act to the best of their knowledge but make fewer mistakes if they are better trained. Thus, quality depends here on the information acquisition on new treatment technologies or medication by physicians.

The subsequent section introduces the model. The reference scenario is analysed in which the health authority contracts individually with each physician. In sections 4.3 and 4.4, the situation is described in which a medical association exists as a supply side monopoly in charge of setting standards for medical education, training on the job and a payment system for physicians. Differentiated prices and uniform pricing are discussed. Finally, some policy implications and concluding thoughts are given in section 4.5.

4.2 Individual contracting

The following sections will compare several scenarios which differ in the institutional arrangement of the health care market. The next section sets up the reference case in which the health authority (HA) contracts with each physician individually. It sets prices for treatment and requires certain standards of medical education and training on the job. The subsequent sections deal with the role of a medical association (MA) which is established as a supply side monopoly. The consequences of standard and price control by the association are compared with the reference case.

The reference scenario is a situation in which the health authority sets a price for medical treatment and requires standards for medical education and training on the job from participating physicians. Homogeneous patients and heterogeneous physicians, both of mass 1, are randomly matched on the market. Physicians only differ in the age group and consequently in the training level they exhibit. A share α, $\alpha \in \{\frac{1}{3}, \frac{1}{4}, ..., \frac{1}{n}\}$, of physicians has just left university and completed medical education k which costs $c^K(k)$. The remaining share $(1-\alpha)$ constitutes the older generation of physicians in each period. The level of education k represents the probability of prescribing the adequate treatment in the first period. A physician lives for $\frac{1}{\alpha}$ periods, of which he belongs for $(\frac{1}{\alpha}-1)$ periods to the older generation of physicians[53]. During this time span of their working life physicians

[53]This setup has been chosen to make sure that in each period, a fraction α of physicians enters the market, and a fraction α leaves the market. The Federal Medical Association in Germany provides data which show that in 2002, only 0.9% of physicians in the outpatient sector were younger that 35, and only

may keep up the probability of correct treatment by investing in each period in training on the job $t \in [t_{min}; 1]$ at a cost $c^T(t)$, with $c^T(t_{min}) = 0$. Both $c^K(k)$ and $c^T(t)$ are convex cost functions. The resulting probability of correct treatment in the later periods[54] is $t \cdot k$. Patients pay a contribution p in each period to the health care system. This contribution consists of the average payment to physicians in this period plus transaction costs. It is straightforward to show that physicians will exactly be reimbursed for their costs in each period. The HA maximizes the expected utility of a representative patient. It pays out p_i to a physician of generation i, $i \in \{young, old\}$. From the point of view of the individual physician, the interaction with the health authority lasts for $\frac{1}{\alpha}$ periods. In the last period, a standard t is required which implies costs $c^T(t)$. If the payment to the physician in this period p_{old} is lower than $c^T(t)$ then the physician prefers not to participate. If it exceeds $c^T(t)$ the health authority has an incentive to reduce it without harming the physician's participation constraint. The same argument can be made for the second last period and so on, and consequently also for the first period in which the payment covers costs of education $p_{young} = c^K(k)$. It is therefore not possible to establish long-term contracts in which the payment structure differs in any period from the one just described. Since the health authority faces a share α of young and $(1-\alpha)$ of old physicians in each period, the patients' contribution is $p(k,t) = \alpha c^K(k) + (1-\alpha)c^T(t) + R$, where R is the transaction cost which the health authority bears for setting up a contract with each individual physician. The health authority could do better by paying out the lifetime income[55] $LE = c^K(k) + \hat{\delta} c^T(t)$ in the first period and not making any payments in the subsequent periods, but as has just been argued this is not a feasible solution because the physician would then prefer not to participate or to choose t_{min} in the subsequent

9.4% were younger than 40 (http:/www.kbv.de/publikationen/2803.htm). Clearly, older generations of physicians constitute a majority within the association.

[54]This setup implies several simplifications. It focuses on the information acquisition on new standards of treatment by older generations of physicians. Thereby, no differentiation is made among older generations, but only between the youngest physicians and the older ones. Furthermore, it is neglected that physicians may also acquire knowledge over time through practice. This would lessen the need for training on the job but does not change the fundamental trade-off between generations which occurs within a medical association and which is central to this analysis. Also, patients cannot differentiate between physicians because otherwise they would always want to be treated by a young physician who exhibits in this model the highest training level.

[55]$\delta < 1$ is the physician's discount factor. $\hat{\delta} = \frac{\delta - \delta^{-\alpha}}{1-\delta}$ then is the discount factor applying to the total payments made in periods 2 to $\frac{1}{\alpha}$.

periods.
The health authority thus maximizes

$$\begin{aligned}\max_{k,t} \quad & \alpha[kU(Y-p(k,t))+(1-k)U(Y-p(k,t)-\epsilon)]+ \\ & (1-\alpha)[t\cdot kU(Y-p(k,t))+(1-t\cdot k)U(Y-p(k,t)-\epsilon)]\end{aligned} \tag{4.1}$$

with U as the patient's strictly concave utility function, Y initial income, and ϵ the disutility of inadequate treatment in monetary terms.
The First Order Conditions are

$$\begin{aligned}\frac{\partial EU}{\partial k}: \quad & (\alpha+(1-\alpha)t)[U(Y-p)-U(Y-p-\epsilon)] = \\ & [(\alpha k+(1-\alpha)t\cdot k)U'(Y-p)+(1-\alpha k-(1-\alpha)t\cdot k)U'(Y-p-\epsilon)]\cdot\frac{\partial p}{\partial k} \\ \frac{\partial EU}{\partial t}: \quad & (1-\alpha)k[U(Y-p)-U(Y-p-\epsilon)] = \\ & [(\alpha k+(1-\alpha)t\cdot k)U'(Y-p)+(1-\alpha k-(1-\alpha)t\cdot k)U'(Y-p-\epsilon)]\cdot\frac{\partial p}{\partial t}\end{aligned} \tag{4.2}$$

From the First Order Conditions with respect to k and t the following condition can be derived:

$$\frac{(1-\alpha)k}{\alpha+(1-\alpha)t} = \frac{(1-\alpha)\frac{\partial c^T}{\partial t}}{\alpha\frac{\partial c^K}{\partial k}} \tag{4.3}$$

The First Order Conditions determine k^* and t^*. They equate marginal benefits and marginal costs for k and t, respectively. Since both standards determine probabilities of adequate treatment, it can be seen from condition (4.3) that the optimal choice can be reduced to equalizing the relation of incremental success probabilities to the relation of incremental costs when increasing k and t marginally. The results can be summarized as follows:

Lemma 10 *Individual contracting*
Under individual contracting, the health authority sets standards for education k^* and training t^* according to conditions (4.2).
A patient's contribution to the system is $p(k^*,t^*) = \alpha c^K(k^*) + (1-\alpha)c^T(t^*) + R$.
The payment to a young physician is $p_{young} = c^K(k^*)$.
The payment to an old physician is $p_{old} = c^T(t^*)$.

The lifetime earnings of a physician are $LE = c^K(k^*) + \hat{\delta}c^T(t^*)$. This situation is now compared to an institutional arrangement in which medical associations exist.

4.3 Medical associations and differentiated pricing

As is often argued in the literature[56] it is relatively easy for physicians to organize themselves as a supply side monopoly. By forming a monopoly, the physicians may be able to reap a rent. Here, the achievable rent R is the transaction cost which the health authority would have to bear under individual contracting. It is plaubile to assume that the MA pushes the representative patient down to his reservation utility

$$U^* = \alpha[k^*U(Y - p(k^*, t^*)) + (1 - k^*)U(Y - p(k^*, t^*) - \epsilon)] +$$
$$(1 - \alpha)[t^* \cdot k^*U(Y - p(k^*, t^*)) + (1 - t^* \cdot k^*)U(Y - p(k^*, t^*) - \epsilon)]$$

Apart from reaping rents for their members, medical associations often exert control over several other variables in the health care market. In Germany, for instance, the right to determine standards of education and training on the job as well as setting prices for treatment is (at least in part) transferred to the MA. It is not immediately clear why this is the case, and the following analysis provides a rationale for this finding.

A MA consists of physicians of all generations. Clearly, the majority of physicians belong to the older generation and dominates the young generation when making decisions. A first intuition seems to suggest that the association will - as compared with the HA - increase the requirements for medical education, thereby placing higher costs on younger generations of physicians at the expense of older ones, for whom the level of training on the job is reduced. This section shows that this is not necessarily the case.

Assume that the MA receives the payment $p = p(k^*, t^*)$ from the health authority and may distribute it to the physicians. Since physicians of the old generation need to take a joint decision in order to constitute the majority and since they essentially face the same decision problem, it is assumed that they receive a uniform price p_2. The MA maximizes the old generation's income[57]:

$$\begin{aligned} \max_{k,t,p_1,p_2} \quad & p_2 - c^T(t) \\ s.t. \quad & p \geq \alpha p_1 + (1 - \alpha)p_2 \\ & p_1 + \hat{\delta} p_2 \geq c^K(k) + \hat{\delta} c^T(t) \\ & EU \geq U^* \end{aligned} \tag{4.4}$$

[56]See Breyer et al. (2003), Zweifel and Eichenberger (1992)

[57]Considering the remaining lifetime income of older physicians would not change the results. The exact modelling of the maximization problem does not matter as long as enough weight is put on older generations as compared to younger ones. An alternative would for example be the maximization of the average lifetime income of all members of the MA.

The first constraint ensures that the MA breaks even in each period. The second constraint implies that a physician entering the market must be willing to participate. Since it is obligatory to become a member of the MA before entering the market[58], the individual physician may not be able to recover the rent R. The last constraint ensures that the average patient is at least as well off as under the system where the HA sets standards and a pricing structure since the HA might simply switch back to this arrangement otherwise. Solving the first constraint for p_1 and substituting in the second constraint allows reducing the problem to the following First Order Conditions:

$$\begin{aligned}
\frac{\partial L}{\partial k} &= -\lambda\frac{\partial c^K}{\partial k} + \mu(\alpha + (1-\alpha)t)[U(Y-p) - U(Y-p-\epsilon)] = 0 \\
\frac{\partial L}{\partial t} &= -[1+\lambda\hat{\delta}]\frac{\partial c^T}{\partial t} + \mu(1-\alpha)k[U(Y-p) - U(Y-p-\epsilon)] = 0 \\
\frac{\partial L}{\partial p_2} &= 1 + \lambda(-\tfrac{1-\alpha}{\alpha} + \hat{\delta}) = 0 \\
\frac{\partial L}{\partial \lambda} &= \alpha^{-1}p - \tfrac{1-\alpha}{\alpha} + \hat{\delta}p_2 - c^K(k) - \hat{\delta}c^T(t) = 0 \\
\frac{\partial L}{\partial \mu} &= \alpha[kU(Y-p) + (1-k)U(Y-p-\epsilon)] + \\
&\quad (1-\alpha)[t\cdot kU(Y-p) + (1-t\cdot k)U(Y-p-\epsilon)] - U^* = 0
\end{aligned} \tag{4.5}$$

with λ and μ as the Lagrange multipliers for the second and third constraint. From the First Order Conditions with respect to k and t the following condition can be derived:

$$\frac{(1-\alpha)k}{\alpha + (1-\alpha)t} = \frac{(1-\alpha)\frac{\partial c^T}{\partial t}}{\alpha\frac{\partial c^K}{\partial k}} \tag{4.6}$$

Since conditions (4.3) and (4.6) coincide, the MA chooses the same standards k^* and t^* as the health authority. But prices differ from the other system: $p_1 = c^K(k^*) - \frac{\hat{\delta}}{1-\alpha-\alpha\hat{\delta}}R$; $p_2 = c^T(t^*) + \frac{1}{1-\alpha-\alpha\hat{\delta}}R$. Lifetime net income of a physician entering the market is reduced to 0. The results are summarized in the following proposition:

Proposition 5 *Medical associations - differentiated pricing*
With differentiated prices, the medical association sets standards k^* and t^*.
The payment to a young physician is $p_1 < p_{young}$.
The payment to an old physician is $p_2 > p_{old}$.

These results have an interesting interpretation: If medical associations can differentiate

[58] In order to be allowed to treat patients under the public health insurance system, physicians in the German outpatient sector have to get their licence as a practitioner, and then get registered with the regional medical association. Physicians in the inpatient sector obviously have to be employed by a hospital.

prices among their members, they will choose the same standards of medical education and medical training as the health authority would. This is the case because a deviation from these "second best" standards would cause further distortions at the expense of the total surplus which the medical association can distribute. This will become even clearer in the next section. Still, prices for the younger generation are lower and for the older generation are higher than in the HA system. Additional redistribution may occur via the working hours required. In German hospitals, young physicians often have extremely long working hours, which was considered illegal by European courts in 2003. At first, it is obvious that the old generation manages to extract rents from the young generation. But it also becomes clear that this is only possible when the system is set up. The only generation who wins is the old generation who introduces the system of MAs. Generations who start to work after that are born into a system where no more rents can be extracted but which cannot be broken up so easily any more. The old generation who is in the majority maintains the system because the physicians want to recover the investment which they made in the first period of their working life. Only the young generation would be willing to abolish the system but they constitute a minority.

4.4 Medical associations and uniform pricing

If a system of medical associations (MA) is introduced in which the MA may control the standards for education and training on the job but has to pay a uniform price for treatment to each physician irrespective of his age, then it can be shown that standards will differ from k^* and t^*. The payment then of course equals the patients' contribution to the health care system p.
The MA is once more assumed to maximize the old generation's income.

$$\begin{aligned} &\max_{k,t} && p - c^T(t) \\ &s.t. && p(1+\hat{\delta}) \geq c^K(k) + \hat{\delta} c^T(t) \\ & && EU \geq U^* \end{aligned} \tag{4.7}$$

First of all it is straightforward to show that such a system cannot exist if the physicians don't earn a rent. Imagine the situation in which the HA pays out $p' = \alpha c^K(k^*) + (1-\alpha)c^T(t^*)$ to the MA. The MA in turn offers this payment to each physician. Given the same standards k^* and t^* which are necessary to leave the average patient indifferent between the two systems, the participation constraint of a physician of the first generation is not fulfilled because $p' \cdot (1+\hat{\delta}) < c^K(k^*) + \hat{\delta} c^T(t^*)$ as long as $c^K(k^*) > c^T(t^*)$. In order to satisfy this constraint, k and/or t would have to be reduced, thereby harming the patient's willingness to participate in this arrangement.

The following result can therefore be established:

Lemma 11 *Uniform pricing - minimum rent*
A system of medical associations and uniform pricing can only exist if physicians receive a rent $R \geq \bar{R}$.

$\bar{R}$ is defined as

$$\bar{R} = \frac{c^K(k^*) + \hat{\delta} c^T(t^*)}{1+\hat{\delta}} - \alpha c^K(k^*) - (1-\alpha) c^T(t^*) \tag{4.8}$$

If R exceeds $\bar{R}$, then there is once more scope for the older generations to reap rents from the younger generation. Since differentiated prices are not feasible, the extraction of rents now has to be done via adjustments of k and t. The MA will choose $k' > k^*$ and $t' < t^*$ such that the participation constraints for physicians of the first generation and for patients hold with equality. Older generations of physicians urge the young generation into higher levels of medical education and put the burden of higher education costs on them. At the same time, they reduce the requirements for training on the job and incur as a consequence lower training costs in later periods of their working life. This leads to the following proposition:

Proposition 6 *Medical associations - uniform pricing*
With uniform prices, the medical association sets standards $k' > k^*$ and $t' < t^*$.
The payment to a physician equals the patients' contribution $p(k^*, t^*)$.

The German health care system provides evidence for these results: Physicians are particulary old as compared with other industrialized countries when they take up their first fully paid job[59]. This indicates a long time span for education and foregone earnings. At the same time, levels of training on the job are rather low, which can be seen from recent attempts to improve on them: The health care reform put into place in January 2004 tries to implement the requirements for training on the job more strictly.

[59] See Zweifel and Eichenberger (1992).

Once more, physicians will not be able to acquire rents in the long run. The rent $\bar{R}$ is depleted in this system because of the uniform pricing requirement. Any payment in excess of $\bar{R}$ will be depleted by the old generation who is in the market when the system is set up. This generation will be able to push future generations down to their reservation utility. Therefore, the outcome is the same as under differentiated pricing, but the mechanism of depletion of rents is different.

4.5 Policy implications and conclusion

The analysis has made it clear why physicians are opposed to certain reforms of the institutional arrangements in the health care market. During the discussion of the reforms which are being put into place in Germany at the moment, the issue of individual contracting for physicians was a hot topic. The current system with medical associations in charge of payments and control of standards was considered to inhibit more far-reaching organizational reforms. The abolition of these arrangements in favour of individual contracting of the health authority with physicians was considered to open the way for managed care systems like health maintainance organisations or preferred provider organisations which have been developed elsewhere. The majority of physicians was strongly opposed to any reforms of the contractual arrangements, and finally managed to prevent a reduction of powers of the medical associations.

This study has shown that the main reason for the opposition to organizational reforms lies in the fact that older generations of physicians need to recover the investments they made into medical education at the beginning of their career. Additional rents which the physicians may achieve by forming a supply side monopoly in general go to the old generations of physicians who introduce the system. Yet an abolition of the system is difficult because later generations of physicians achieve their reservation utility only over their lifetime: The payment structure is such that they make an investment in the first period which is only paid back in later periods of their life. Abandoning the current system therefore requires us to consider how currently older generations of physicians should be compensated.

The analysis considers different payment structures which the medical association may use. Differentiated prices can be observed, for example, in the German inpatient sector.

Here, young physicians are clearly "exploited" via seniority wages and recover these investments in later periods[60]. In the German outpatient sector, the same prices are paid to physicians irrespective of seniority considerations. This model suggests that in the outpatient sector, older generations reap rents by lower investment in training on the job and possibly higher levels of education. The often-discussed problem of too low levels of training on the job in Germany should therefore be particularly pronounced in the outpatient sector.

The current reform has taken up this problem by requiring higher standards of training on the job. Every five years, members of the medical association have to prove that they have undertaken a specified schedule of training on the job. Although these requirements have existed in the past, MAs have neglected their implementation. This measure will harm older generations now but lead in the long run to a more equitable outcome between different generations of physicians. By increasing these requirements step by step, one can at least for the outpatient sector conclude that the costs of institutional reform are thus distributed over more generations of physicians. Concerning the inpatient sector, the same can only be achieved by gradually adjusting the income of young physicians at the expense of older ones.

Of course, this analysis has only dealt with a specific role of medical associations. It seems plausible to assume that in reality, additional rents can be achieved under this system which have not yet been taken away by the introductory generation but which accrue to all physicians participating in this organizational arrangement. Yet the punchline of the argumentation seems to hold that older generations dominate younger ones, that they force them to take investments in the first periods of their working lifes which are only paid back later, and that therefore it is particularly the older generations of physicians who are strongly opposed to new organizational arrangements.

[60]Although MAs are not in the position to redistribute income of physicians in the inpatient sector, new organizational forms bear the danger of undermining a seniority wage system by implementing, for instance, performance based wages. This is not in the interest of older generations.

5 Sustaining Will by Staying Ignorant

5.1 Introduction

Von Creutzfeldt-Jakob (vCJK) disease, a fatal illness, can potentially be transmitted via blood transfers.[61] In 2000, the British National Blood Authority revealed that at least seven people who had died from vCJK so far donated blood regularly. They also found out the names of several people who obtained this contaminated blood. However, the physicians in charge decided not to inform the patients about this potential threat to their health. The argument was that without knowing that they were at risk of getting vCJK the persons would be more likely to live a happy life. Furthermore, the illness can not be treated.

Even if people can decide for themselves whether to obtain information about their health status or not they sometimes do not want to. This is particularly the case if this information might potentially be really bad news. For example, out of 396 individuals, 169 (46%) declined costless genetic testing for the breast cancer gene (Lerman et al., 1998).

These observations cannot readily be explained by standard expected utility theory. In this theory, more information is individually always better than less, as long as this information is private, i.e. not revealed to other parties. Less information might be better from a social point of view, the famous Hirshleifer paradox. Information might also be undesirable if two parties interact in a strategic relationship (see e.g. Schmidt, 1996). However, these models are not applicable to the situation we want to discuss here: Why do people in some cases reject obtaining better information on their health status?

We show that ignorance may be rational when taking into account the problem of time-inconsistency. There is abundant evidence that many individuals exhibit time-inconsistent preferences.[62] They have a strong preference for the present over future periods which may cause their decisions to depend on the exact date when they are taken. In recent years, many economists (see e.g. Laibson, 1997, 1998; O'Donoghue and Rabin, 1999a; Carillo and Mariotti, 2000) have formalized and applied this concept of boundedly rational

[61] Due to the recent cases of mad cow disease in Great Britain, which potentially leads to vCJK, American, Canadian and German authorities forbid blood transfusion from people who spent more than six months in Britain during the time period 1980-1996.

[62] For further references on empirical evidence, see Loewenstein and Elster (1992).

behaviour. Their models come under the heading of "hyperbolic discounting" or "present biased preferences". The basic model is very simple: An individual discounts the next period very strongly (e.g. by $\lambda\delta$ with $0 < \lambda, \delta < 1$), while all subsequent future periods are marginally discounted by δ only. In such a model, time inconsistent behaviour occurs: E.g. if someone is asked today whether she prefers one apple in five days or two in six days, she might choose the latter (if $\delta > 1/2$). However, when the day arrives, she might prefer obtaining one apple today instead of two tomorrow (if $\lambda\delta < 1/2$). The theory has been applied to numerous phenomena like savings behaviour (Laibson, 1996, 1997, 1998; Laibson, Repetto and Tobacman, 1998; O'Donoghue and Rabin, 1999b), self-confidence (Benabou and Tirole, 2002), haste (Brocas and Carillo, 2000), and also to the value of information (Carillo and Mariotti, 2000), on which we will comment below.

In this paper we discuss the decision whether or not to obtain information on one's own health status, e.g. in the form of genetic testing, when individuals discount the future hyperbolically. The basic intuition for our results is the following: People who have time-inconsistent preferences and who are aware of this fact, might devise strategies to overcome the resulting inefficiencies. The psychologists Ainslie and Haslam (1992) mention personal behavioural rules as one means of making one's behaviour more consistent.[63] As an example, a person might be willing not to drink alcohol today, not because she does not like to drink, but because she is aware that if she once starts to consume alcohol, she will never be able to resist the temptation of drinking alcohol again. Ainslie (1992) calls such a behaviour *willpower*. Such a behavioural rule can be modelled as a trigger strategy in a dynamic game: Stick to the long term equilibrium path, otherwise you will be punished in future periods by always making the short run optimal choice.[64] This is in line with Ainslie's and Haslam's (1992) observation who state that every repetition of the desired action tends to fix conduct in the right direction whereas one single deviation causes disproportional damage by undoing the effects of many correct choices.

Based on this definition of will, and given the analogy with dynamic games, it becomes clear that will can only be sustained if the future is valued sufficiently highly. With the example given above: If a person does not value the future very much, she might as well

[63] Personal rules are also discussed more extensively in Ainslie (2001).

[64] Note that in the present case, the different players are different incarnations of the same person. But due to time-inconsistency, different "selves" have different preferences for future consumption streams and behave according to their own preferences.

drink today. The threat that she will drink tomorrow and in the future, which from today's point of view is not desired, is not enough to forgo the pleasure of consuming drinks today. This fact may have an impact on the acquisition of information about the value of future periods.

Consider tests on your health status, for example genetic tests.[65] In contrast to standard diagnostic tests like e.g. measuring blood pressure, most genetic tests indicate, if the test result is positive, a high probability for getting a severe illness. Three well-known examples are the tests for Chorea Huntington, which is a terminal illness, Alzheimer's disease and breast cancer. A person who is offered the possibility to undertake such a test might reason the following way: If I undertake the test, I have better information on my illness risk, so I can plan my future life accordingly. E.g., if I have a high chance of obtaining Alzheimer's disease I might plan to undertake a trip around the world earlier in life, and not after retirement. On the other hand, a positive test result indicates a very large illness probability. Thus, there is a real danger that my will might break down, that I will lose faith in myself. Or more formally, the strategies devised to overcome the problem of time-inconsistency might no longer be stable. If this latter effect dominates the first, then it can indeed be rational not to undertake the test.

The fundamental trade-off in our theory is that on the one hand more information is better because one can structure future life better.[66] On the other hand, more information is worse because it might lead to a breakdown of will. The latter effect only sets in if the consequences of the negative information are sufficiently severe, otherwise willpower can be sustained. Interestingly, if the consequences are extremely bad, the former effect might dominate the latter. If someone is told that she will live for only one further period, breakdown of will does not matter any more, as there are no future periods. Thus there is a non-monotonicity in the value of information: If the illness is light or extremely severe, information is positively valued. For illnesses which are severe, but not extremely so,

[65] Although our model is suited for different forms of information gathering, genetic testing fits the analysis extremely well for the following reasons. First, the test result is either positive or negative, either the person has the genetic mutation or not. In our model we also discuss only two possible states of the world with regard to the health status. Second, for many genetic tests a positive result indicates a severe health problem, which may plausibly lead to a breakdown of will.

[66] In addition, a better diagnosis might lead to better precautionary behaviour and more specific treatment. We do not model this effect explicitly. However, for our results to go through we only require that someone whose test result is positive is worse off than an average person who is untested.

information has a negative value.

Apart from this non-monotonicity in the value of information, we obtain a third claim from our model. If before the acquisition of information the health status is already quite bad, people are more willing to obtain additional information. This may happen because even before taking a test a breakdown of will is imminent. Then more information is likely to have a positive value.

We provide evidence for the implications of the theory based on studies reported in the literature and on our own investigations. We questioned students on their willingness to gather information on genetic and blood tests. The data were used to test three hypotheses derived from our model. The first claims that people are more willing to get information about mild diseases than about severe ones, as for light illnesses breakdown of will does not occur. The second hypothesis states that the individuals are more likely to want information about extremely severe diseases compared to severe ones, as in the first case a breakdown of will does not matter any more. It turned out that the data were in good agreement with these predictions.
The third hypothesis tested says that individuals are more likely to seek information if their ex-ante probability of illness is already high. We captured this by including two questions for Chorea Huntington in the questionnaire, one describing a situation with and one without a family history (high vs. low ex-ante probability of falling ill) of the relevant genetic mutation. Our data suggest that those individuals whose attitude towards information gathering differs in both cases are indeed more likely to want this information in the case of a high ex-ante probability. Evidence from the literature is presented which confirms our results.

Before going into the details of our model, we briefly discuss the work of Carillo and Mariotti (2000). These authors also show that if people have time-inconsistent preferences, it might be optimal for them not to obtain further information. One instructive example which the authors present is the following: Suppose you are a non-smoker and you believe that the long term consequences of smoking are very bad. Now you are offered the possibility of obtaining further information about the health risks of smoking. Suppose that with some probability the information can be such that smoking is less risky than you thought but still quite risky. In that case you might like to smoke today, but you would prefer to forbid your future selves to smoke. However, due to time-inconsistency, future incarnations will also smoke. Under these circumstances it might be rational not to obtain this information at all. So Carillo and Mariotti argue that more information might be undesirable because it leads to a more severe problem of time-inconsistent behaviour. Our model differs from the Carillo and Mariotti model in several respects. First, information in our case does not concern the severity of the consequences of the action, but the probability of occurance, and therefore the survival probability of an individual. Second, more information in our model does not lead to different degrees of time-inconsistent behaviour. If the illness probability is low, individuals succeed in overcoming the inefficiencies due to time-inconsistency. However they are afraid that in case of a positive test result these strategies will no longer be stable. As a consequence, our model predicts that information has negative value mainly if severe illnesses are tested. In contrast, in subsequent work to Carillo and Mariotti (Brocas and Carillo, 2000) it is shown that the value of information is more likely to be negative in their model if the flow of information is 'small'.

The paper is structured as follows. In the next section we will discuss the basic model of hyperbolic discounting and establish the concept of will. The model is then applied to genetic testing, and comparative statics results are derived. In Section 3, empirical evidence from our own investigations and from the literature is presented. Section 4 concludes and points at policy implications of the model.

5.2 A model of *will*

5.2.1 Hyperbolic discounting

In recent years, the following model of time inconsistent preferences has been widely used (see e.g. Laibson, 1997, 1998; O'Donoghue and Rabin, 1999a; Carillo and Mariotti, 2000). Let $U(c_t, c_{t+1}, ...)$ be the utility in period t, depending on consumption in periods t, t+1, ... Then:

$$U(c_t, c_{t+1}, ...) = u(c_t) + \lambda \sum_{\tau=1}^{\infty} \delta^{\tau} u(c_{t+\tau}) \tag{5.1}$$

where $0 < \lambda < 1$ and $0 < \delta < 1$. The only difference to the standard exponential discounting model lies in the parameter λ. $\lambda < 1$ implies a strong discounting between today and next period, namely $\lambda\delta$, while two subsequent future periods are only discounted by δ.

The model used here follows the line of Phelps and Pollak (1968). Let there be a capital stock K_t in period t. If a person chooses consumption $c_t = (1-\sigma_t)K_t$ in period t, where $0 < \sigma_t < 1$, then $K_{t+1} = \beta\sigma_t K_t$. Here σ_t is the savings ratio in period t and β measures productivity, which is constant over time. This is the interpretation of the variables we will use in the rest of the paper. However, the model goes far beyond such a simple consumption/savings problem. c might be interpreted as the size of any decision which is such that a person enjoys pleasure from it today, but choosing a smaller c allows a larger pleasure in future periods. For example, c could be the amount of cigarettes a person smokes, while K would then be interpreted as a stock of health. Or, c could be leisure, or time and effort not spent on studying. K would then be human capital. Therefore the model is general enough to encompass many situations where a person has discretion over one variable and she faces the trade-off between a larger utility today and a larger utility in the future. To make the model tractable, we further simplify by assuming that utility is logarithmic, i.e. $u(c_t) = ln(c_t)$.

Next we consider the case where no time-inconsistency problem is present. Formally, this is obtained by setting $\lambda = 1$. In decision theoretic terms, this is the utility that the $t = -1$ person, who has already chosen her consumption level, would maximize. We call this person the planner. Let K denote the initial capital stock. If this planner could decide on a savings rate σ for all future periods, her optimization problem would be:

$$\max_{\sigma} U_P(\sigma) = \sum_{t=0}^{\infty} \delta^t \ln[(1-\sigma)\beta^t\sigma^t K] \tag{5.2}$$

Here we use the fact that if the savings rate is σ, then $c_t = (1-\sigma)\beta^t\sigma^t K$. $\beta^t\sigma^t K$ is the capital stock at time t, and $(1-\sigma)$ is the proportion of this stock the person consumes.[67]

Simple calculation gives:

$$U_P(\sigma) = \frac{1}{1-\delta}\ln[(1-\sigma)K] + \frac{\delta}{(1-\delta)^2}\ln[\beta\sigma] \tag{5.3}$$

Taking the first order condition with respect to σ yields:

$$\frac{dU_P(\sigma)}{d\sigma} = -\frac{1}{1-\delta}\frac{1}{1-\sigma} + \frac{\delta}{(1-\delta)^2}\frac{1}{\sigma} = 0 \quad \Rightarrow \quad \sigma = \delta \equiv \sigma^c \tag{5.4}$$

If the planner could decide how much she will save in the future, she would choose the savings rate equal to the discount rate, a well-known result. To make the analogy with a

[67] Although we let the person choose the same savings rate for every period, it is easy to show that this will be optimal from an ex-ante perspective even if different savings rates were allowed.

dynamic game, we call this the cooperative savings rate σ^c. Note that λ does not enter the optimal savings rate, as the planner does not suffer from short-sightedness.

In contrast to this, consider now the case where future selves cannot be forced to save σ^c. We then look for the "non-cooperative"-savings rate, i.e. the amount of savings each incarnation would choose if she took the savings of future selves as given. Let s be the savings rate of the period t individual, where $t \geq 1$.[68] Then in period 0, the person would optimize her consumption according to the following maximization problem:

$$\max_{\sigma} U_0^{\delta}(\sigma) = \ln[(1-\sigma)K] + \lambda \sum_{t=1}^{\infty} \delta^t \ln[(1-s)\beta^t s^{t-1} \sigma K] \tag{5.5}$$

Reformulating this expression gives:

$$U_0^{\delta}(\sigma) = \ln[(1-\sigma)K] + \lambda \frac{\delta}{1-\delta} \ln[(1-s)\beta\sigma K] + \lambda \frac{\delta^2}{(1-\delta)^2} \ln[\beta s] \tag{5.6}$$

The optimal "non-cooperative" savings rate is then given by:

$$\frac{dU_0^{\delta}(\sigma)}{d\sigma} = -\frac{1}{1-\sigma} + \lambda \frac{\delta}{(1-\delta)} \frac{1}{\sigma} = 0 \quad \Rightarrow \quad \sigma = \frac{\lambda\delta}{1-\delta+\lambda\delta} \equiv \sigma^{nc} \tag{5.7}$$

If $\lambda = 1$, i.e. in the limit of exponential discounting, $\sigma^{nc} = \sigma^c$. This is the well known result: If discounting is exponential, decisions will be taken consistently. However, if $\lambda < 1$, σ^{nc} will be smaller than σ^c, i.e. a hyperbolic discounter will save less than she would prefer from an ex-ante perspective.

5.2.2 Will

In the previous subsection we have shown that each incarnation prefers to save less than is optimal from an ex-ante point of view. It is a much-reported observation in the psychology literature that individuals devise behavioural rules to overcome this inconsistency.[69] Ainslie argues that such a behaviour is similar to Kant's categorical imperative: Behave as if this were a universal rule. Here this is not a rule for different persons, but for a single person with different selves at different points in time. Each incarnation behaves as he or she would want all others to behave. One way to get around the problem of undersaving

[68] It is common in the literature to restrict attention to stationary equilibria which are symmetric across incarnations.

[69] See Ainslie and Haslam (1992) for further references.

might be the following strategy for the person in period t:
Save σ^c if in all previous periods σ^c was the savings rate. If in any period before, the savings rate was different from σ^c, then save σ^{nc} instead.

This behavioural rule operates like a trigger strategy. As long as all incarnations cooperate, i.e. choose σ^c, there will be cooperation in the future. However, if someone deviates, then the inefficient non-cooperative equilibrium will be played in all future periods. Although this is an extreme form of punishing oneself, namely by never returning to the cooperative path, it is not uncommon. For example, the Alcoholics Anonymous proclaim a strict "no-drink" policy, as they fear that any drink will lead the person back into being an alcoholic. This is in line with observations by psychologists that a single deviation from a behavioural rule causes disproportionate damage by offsetting many correct choices.

The psychologist Ainslie (1999) calls such a behaviour *willpower*: Although the agent prefers to consume more immediately, she will save instead in order to support her future selves to behave in the same way.

We are therefore lead to the following definition:

Definition 1
An individual is said to possess will, if there exists an equilibrium of the savings game where σ^c will be chosen in every period.

Before formalizing this notion of will, let us first comment on the specific properties of the equilibrium concept. There seems to be widespread agreement in the literature that 1) hyperbolically discounting people suffer from time inconsistency, and 2) that they might devise strategies against this inconsistency. However, how this should be modelled explicitly is far from obvious. In particular, the interaction between different selves is not well understood. The trigger strategy we propose here suffers from the fact that future selves cannot cooperate again. This seems a rather harsh assumption, as future selves are incarnations of the same person, so at least they should be able to communicate with each other. This possibility is taken up by Kocherlakota (1996). He uses the notion of "reconsideration-proofness" where future selves are allowed to reconsider their strategies. However, the concept he proposes is defined only for stationary decision problems. An alternative would be the concept of "revision-proof" decision rules introduced by Asheim (1997). Here an axiomatic setup is proposed which captures the notion that a person will, by devising a decision rule, take into account that she can reselect the rule at any later point in time. However, also in this approach the general case of an infinite horizon capital accumulation problem has not been analysed. While the debate on the "right" solution concept is interesting and deserves more consideration,[70] the focus of the present paper is different: We intend to model the notion of will in a simple, tractable way. This is done by allowing only trigger strategies of the form described above. Formally, we restrict the agent to use only one of two strategies: Either to save the cooperative savings rate σ^c, or to save the non-cooperative savings rate σ^{nc}.

Based on this, we can now show that a person can only possess will if the future is not discounted too much. This is done by comparing the utility of e.g. the period 0 type when she chooses σ^c and σ^{nc}. Noting that dependent on her choice of σ^i, $i = c, nc$, all future selves will choose the same σ^i in equilibrium, her utility is given by:

$$\begin{aligned} U_0^\delta(\sigma^i) &= \ln[(1-\sigma^i)K] + \lambda \textstyle\sum_{t=1}^{\infty} \delta^t \ln[(1-\sigma^i)\beta^t(\sigma^i)^t K] \\ &= (1+\lambda\frac{\delta}{1-\delta})\ln[(1-\sigma^i)K] + \lambda\frac{\delta}{(1-\delta)^2}\ln[\beta\sigma^i] \end{aligned} \tag{5.8}$$

Therefore:

$$\Delta U = U_0^\delta(\sigma^c) - U_0^\delta(\sigma^{nc}) = (1+\lambda\frac{\delta}{1-\delta})\ln[\frac{1-\sigma^c}{1-\sigma^{nc}}] + \lambda\frac{\delta}{(1-\delta)^2}\ln[\frac{\sigma^c}{\sigma^{nc}}] \tag{5.9}$$

We can now formulate the following proposition:

Proposition 7
There exists a δ^1 and δ^2 with $0 < \delta^1 \leq \delta^2 < 1$ such that if $\delta < \delta^1$ ($\delta > \delta^2$) the agent cannot (can) possess will.

The proof is relegated to the appendix. There we show that for small values of δ we get $\Delta U < 0$, while for large values of δ it holds that $\Delta U > 0$. Note that for both $\delta = 0$ and $\delta = 1$ the savings levels σ^c and σ^{nc} are the same, which implies $\Delta U = 0$. In Figure 1, we provide a numerical example for $\Delta U(\delta)$ where breakdown of will occurs for $\delta < 0.69$.

Having outlined the model to discuss will, we now analyse the decision to acquire private information on one's health status.

[70] See also Caillaud, Cohen and Jullien (1999)

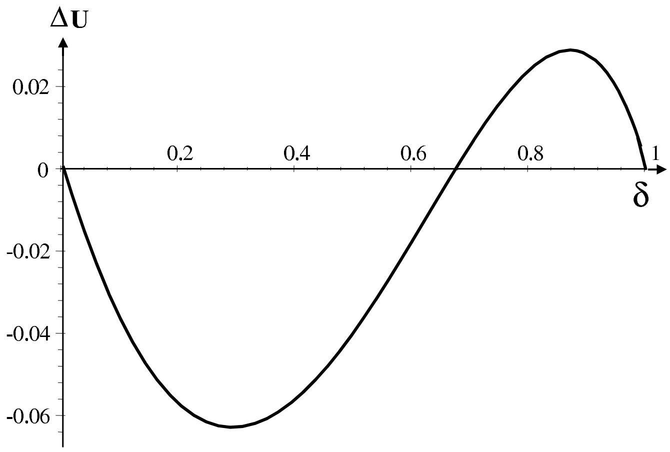

Figure 5.1: Breakdown of will for small δ (numerical example: $K = 2, \beta = 1.5, \lambda = 0.2$)

5.2.3 The value of information

Assume that the planner[71] has the choice to obtain information about her health status. Let us consider as an example that this person might undertake a genetic test. The ex-ante probability of being tested positive is γ. If the test result is positive, the person has a lower survival probability. Denote by δ_l the corresponding discount rate. If however, the test turns out to be negative, the survival (discount) rate will be $\delta_h > \delta_l$. Note that the size of δ_l measures two effects: First, the severity of the illness itself, i.e. the survival probability once the illness breaks out. Second, the probability with which the illness occurs if the test turns out to be positive, i.e. how precisely the test works. The overall impact of a very precise test for a less severe illness is similar to an imprecise test for a very severe illness.[72]

In the previous subsection we modelled a situation with a constant discount rate. The model here is slightly more complicated, as the person, if she remains untested, learns over time about her illness risk. If she survives another period she is more likely to be

[71] It is the t=-1 person who considers the impact of the information acquisition on the behavioural rules adopted.

[72] In the following, we use the notion of "extremely severe", "severe" and "mild/light" diseases as descriptions of the overall impact of a positive test result.

the type who would obtain a negative test result. Or, in other words, γ changes from period to period. In the optimal savings decision without a test this has to be taken into account.

Let us start by calculating the optimal savings rate from the perspective of the planner if no test will be undertaken.

$$U_P^u(\vec{\sigma}) = \sum_{t=0}^{\infty}[\gamma\delta_l^t + (1-\gamma)\delta_h^t]\ln[\beta^t(1-\sigma_t)K\prod_{\tau<t}\sigma_\tau] \qquad (5.10)$$

$\vec{\sigma} = (\sigma_0, \sigma_1, ...)$ is the vector of savings rates. Note that because the discount rate varies between periods, the optimal savings rate will not be constant over time.
Taking the derivative with respect to σ_t gives:

$$-[\gamma\delta_l^t + (1-\gamma)\delta_h^t]\frac{1}{1-\sigma_t^c} + \sum_{\tau=t+1}^{\infty}[\gamma\delta_l^\tau + (1-\gamma)\delta_l^\tau]\frac{1}{\sigma_t^c} = 0 \qquad (5.11)$$

After some calculations:

$$\sigma_t^c = \frac{\gamma\frac{\delta_l^{t+1}}{1-\delta_l} + (1-\gamma)\frac{\delta_h^{t+1}}{1-\delta_h}}{\gamma\frac{\delta_l^t}{1-\delta_l} + (1-\gamma)\frac{\delta_h^t}{1-\delta_h}} \qquad (5.12)$$

The individual uses the information from having survived one more period for updating γ in the appropriate way. Consequently, σ_t^c is increasing over time.[73]

As before, the person still has the temptation to consume more than the cooperative consumption rate. To calculate the non-cooperative savings rates, we maximize the utility of the time t self, if she takes all future savings rates as given. This time not everyone

[73] Take γ as the ex-ante probability of being tested positive in $t = 0$. Then $\hat{\gamma} = \frac{\gamma\delta_l}{\gamma\delta_l + (1-\gamma)\delta_h}$ is the updated ex-ante probability in $t = 1$ if the individual has survived until $t = 1$. If we calculate

$$\sigma_0^c = \frac{\gamma\frac{\delta_l}{1-\delta_l} + (1-\gamma)\frac{\delta_h}{1-\delta_h}}{\gamma\frac{1}{1-\delta_l} + (1-\gamma)\frac{1}{1-\delta_h}}$$

we can use $\hat{\gamma}$ for calculating

$$\sigma_1^c = \frac{\hat{\gamma}\frac{\delta_l}{1-\delta_l} + (1-\hat{\gamma})\frac{\delta_h}{1-\delta_h}}{\hat{\gamma}\frac{1}{1-\delta_l} + (1-\hat{\gamma})\frac{1}{1-\delta_h}}$$

Plugging in $\hat{\gamma}$ yields after some calculations the result of the above formula

$$\sigma_1^c = \frac{\gamma\frac{\delta_l^2}{1-\delta_l} + (1-\gamma)\frac{\delta_h^2}{1-\delta_h}}{\gamma\frac{\delta_l}{1-\delta_l} + (1-\gamma)\frac{\delta_h}{1-\delta_h}}$$

σ_0^c is decreasing in γ, and as $\hat{\gamma} < \gamma$, $\sigma_1^c > \sigma_0^c$. A similar reasoning holds for any other period.

will choose the same savings rate, due to the change in discounting over time.

$$\max_{\sigma_t} U_t^u(\sigma_t, \vec{s}) = \ln[(1-\sigma_t)K_t] + \lambda \sum_{\tau=1}^{\infty} \frac{[\gamma\delta_l^{t+\tau} + (1-\gamma)\delta_h^{t+\tau}]}{[\gamma\delta_l^t + (1-\gamma)\delta_h^t]} \ln[(1-s_{t+\tau})\beta^\tau \prod_{0<\rho<\tau} s_{t+\rho}\sigma_t K_t] \tag{5.13}$$

Taking the first order condition and solving for σ_t gives us the non-cooperative saving rate:

$$\sigma_t^{nc} = \frac{\lambda A_t}{1+\lambda A_t} \tag{5.14}$$

with

$$A_t = \frac{\gamma\frac{\delta_l^{t+1}}{1-\delta_l} + (1-\gamma)\frac{\delta_h^{t+1}}{1-\delta_h}}{\gamma\delta_l^t + (1-\gamma)\delta_h^t} \tag{5.15}$$

The interesting case is where the agent can sustain will if the test result is negative, i.e. $\delta = \delta_h$, but not if the test result is positive $\delta = \delta_l$. We can now formulate a lemma about the possibility to sustain will if the person remains untested.

Lemma 12 *Suppose that δ_h (δ_l) is such that the person can (cannot) sustain her will with such a discount rate. Then there exists a γ^1 and γ^2 with $0 < \gamma^1 \leq \gamma^2 < 1$ such that if $\gamma < \gamma^1$ ($\gamma > \gamma^2$) the agent can (cannot) possess will if she remains untested.*

Proof:

Note that if $\gamma = 0$ we are back to $\delta = \delta_h$. Here the agent at period 0 strictly prefers to save σ^c than σ^{nc}. By continuity, for small values of γ the agent can still sustain her will. The same reasoning applies for γ close to 1. **QED**

This result is straightforward. If γ is small, the agent has a very small probability of obtaining the bad news if tested. So if untested, she can behave as if there is no threat from this illness.

To get some feeling for the relevant parameters take Alzheimer's disease as an example. When you are young, the probability of getting Alzheimer's disease at some point in life is very low. At age 65, the probability is between 12% and 15%. US data show that people of age 85 and more have a 50% chance of suffering from Alzheimer's disease. For Chorea Huntington, e.g., the contrary holds. In the late thirties and early forties, the probability of an outbreak of the disease is highest, getting smaller and smaller as one gets older. That means that young people have a stronger fear of Chorea Huntington inducing a breakdown of will than of Alzheimer's disease.

Now suppose we are in a situation where the agent can sustain will if untested. We can then calculate the gain (or loss) in utility from the planner's perspective from undertaking a genetic test.

$$\Delta = \gamma U_0^{\delta_l}(\sigma_l) + (1-\gamma)U_0^{\delta_h}(\sigma_h) - U_0^u(\vec{\sigma}^c) \tag{5.16}$$

Here, $U_0^{\delta_i}(\sigma_i)$ is the utility of the planner if her discount factor is δ_i and she saves with the savings rate σ_i, where $i \in \{l, h\}$. In (5.16), σ_l is either δ_l if the agent can support her will even if tested positive, otherwise $\sigma_l = \frac{\lambda\delta_l}{1-\delta_l+\lambda\delta_l}$. Because we assumed that the agent can sustain will if untested, it must be the case that she can do so if the test turns out to be negative. Therefore $\sigma_h = \delta_h$.

We can now show that the value of a genetic test might indeed be negative.

Proposition 8
There exist values for δ_l, δ_h and λ such that the value of information is negative.

Proof:
The utility of the person if untested and if she uses the optimal consumption strategy is larger than if she consumes a ratio $(1-\delta_h)$ in every period. Therefore:

$$\Delta < \gamma \sum_{t=0}^{\infty} \delta_l^t \{\ln[\beta^t(1-\sigma_l)\sigma_l^t K] - \ln[\beta^t(1-\delta_h)\delta_h^t K]\} \tag{5.17}$$

Some further calculations give:

$$\Delta < \gamma \frac{1}{1-\delta_l}\{\ln[\frac{1-\sigma_l}{1-\delta_h}] + \frac{\delta_l}{1-\delta_l}\ln[\frac{\sigma_l}{\delta_h}]\} \tag{5.18}$$

With $\sigma_l = \frac{\lambda\delta_l}{1-\delta_l+\lambda\delta_l}$, we obtain $\Delta < 0$ if e.g. $\delta_l = 0.5$, $\delta_h = 0.8$ and $\lambda = 0.2$. Note that for these values expression (5.9) is larger than zero if the test result is negative, but not if it is positive. That is, the person with $\lambda = 0.2$ can sustain will if her discount rate is 0.8, but not if it is 0.5. As the sign of expression (18) is independent of γ, we can according to Lemma 1 find γ small enough such that will is sustained if the person remains untested.
QED

This is our first main result. There exist parameter values such that a person will reject taking a genetic test, even if the test is costless and even if the information is only disclosed to her.

The model allows to say somewhat more about the circumstances under which this effect is most likely (not) to occur. First consider changes in the severity of the illness which is modelled by δ_l.

Proposition 9

(i) If λ and δ_l are such that expression (5.9) is positive, i.e. the person can sustain will independent of the information she obtains, then $\Delta > 0$, i.e. the value of the information is always positive.

(ii) For all values of δ_h, there exists a $\delta^* > 0$ such that if $\delta_l < \delta^*$, then $\Delta > 0$. That is, if the survival probability after a positive test is very low, the agent will prefer to be tested.

Proof:

(i) This can be seen by noting that if the agent can sustain her will always, she might choose the same consumption stream if tested as if untested. Due to the additional choice she now has, namely to condition consumption on the test result, she can only do better.

(ii) We go to the limit of $\delta_l = 0$. As all functions are smooth, one can find a δ^* positive but small enough, such that the following statement holds for all $\delta_l < \delta^*$. In the limit $\delta_l \to 0$, the agent knows that if she survives period zero, she is the type who would obtain a negative test result. Thus she will choose $\sigma_h = \delta_h$ from period one onwards. Her choice in period zero if she remains untested is given by

$$\sigma_0 = \frac{(1-\gamma)\frac{\delta_h}{1-\delta_h}}{\gamma + (1-\gamma)\frac{\delta_h}{1-\delta_h}} = \delta_h \frac{1}{1 + \frac{\gamma}{1-\gamma}(1-\delta_h)}$$

which follows from (5.12) by setting $t = 0$ and $\delta_l = 0$. On the other hand, if the person becomes tested, she will consume K in case the test is positive, while if the test is negative, she will immediately start saving with rate δ_h. In both cases, she is better off if she obtains the test result before deciding how much to consume. Overall, this difference is given by:

$$\Delta = \gamma \ln[\frac{1}{1-\sigma_0}] + (1-\gamma)\frac{1}{1-\delta_h}\{\ln[\frac{1-\delta_h}{1-\sigma_0}] + \frac{\delta_h}{1-\delta_h}\ln[\frac{\delta_h}{\sigma_0}]\} > 0 \qquad (5.19)$$

The first term is the difference in utility if the test is positive. This expression is positive. The second term is also positive, because the savings rate δ_h maximizes the expression.

QED

This result shows that if the severity of a positive test result is either very low, in which case there is no danger of a breakdown of will, or very high, in which case a breakdown of will does not matter very much, the person will prefer to become tested. Figure 2 shows a numerical example for our result. For low levels of δ_l, the value of information is positive because a breakdown of will does not matter very much. For intermediate levels, the negative impact of a breakdown of will is so strong that the value of information becomes negative. Then there is a jump back to a positive value of information when δ_l becomes so large that will can be sustained even with a positive test result.

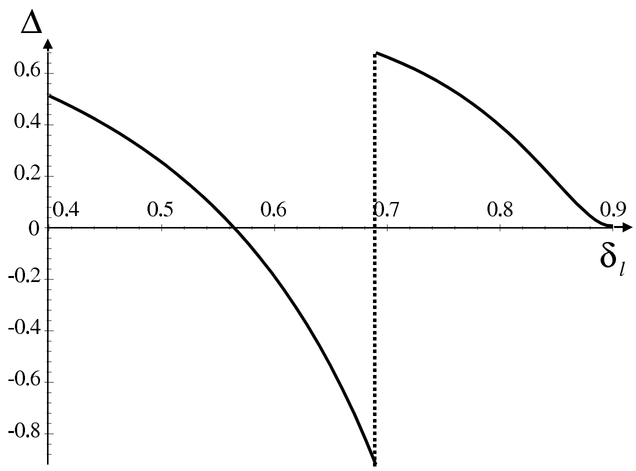

Figure 5.2: Negative value of information for intermediate levels of δ_l (numerical example: $K = 2, \beta = 1.5, \lambda = 0.2, \gamma = 0.5, \delta_h = 0.9$)

Next consider changes in the probability of obtaining the bad news which is given by γ. In particular, if γ is large the person already faces a breakdown of will if she remains untested. It is not possible to make any general claim in this case, as there are several effects which interact. If the person undertakes the test, she might obtain a negative test result which allows her to sustain her will. However, if the test result is positive, the breakdown of will can lead to a much worse situation as savings rates will drop. In this case an effect of the type of Carillo and Mariotti (2000) might set in: It is better not to obtain the information as the problem of time inconsistency without this information is less severe than with the information. However, consider a critical γ^* such that for $\gamma > (<)\gamma^*$ breakdown of will does (not) occur if the person remains untested. Then around γ^* the expected utility of the person if she undertakes the test is continuous in γ.

However, the utility of the planner if she remains untested is lowered by a discrete step when γ reaches γ^*. Thus testing becomes more attractive. We formulate this as a claim:

Claim: *The larger the ex ante probability of obtaining an illness, the more likely it is that the value of information is positive.*

Let us summarize the results:

- Agents might reject undergoing genetic tests if they are afraid that their will will break down if the test is positive.
- Tests for illnesses which are not severe are less likely to lead to a breakdown of will. Therefore these tests are more likely to be taken.
- Tests for illnesses which are extremely severe might also be taken, as here a better consumption planning dominates the negative effects of a breakdown of will.
- If the survival probability is already low before obtaining the information, the test is more likely to be undertaken.

5.3 Empirical evidence

By questioning students on their interest in the results of genetic tests and blood tests we tested the predictions of the model empirically.

Three hypotheses were tested. The first one says that tests for rather mild diseases are more likely to be taken than those for illnesses considered as severe, as they do not induce a breakdown of will.[74] The empirical findings confirm this result. The second hypothesis says that tests for extremely severe illnesses are also more likely to be taken as compared to the middle range of diseases, because a breakdown of will does not matter much. Our empirical results also confirm this theoretical postulate. The third hypothesis tested says that tests are more likely to be taken if the ex ante risk of illness is high, as then the will of an individual is already broken. This may be due to a family history of genetic

[74] Recall that severe illness implies two things: 1) the illness itself is indeed severe, and 2) the probability of obtaining the illness conditional on a positive test result is large.

mutations. Also in this case, the empirical results provide significant evidence. In the literature, similar studies obtained significant results regarding our hypotheses, too.

5.3.1 Data description

We tested the model by asking 222 undergraduate students of economics and business administration whether they wanted to be informed about the results of one blood test and various tests for mutations in genes causing monogenetic diseases. The students were told that this information was purely private and costlessly available to them.

Each disease was described by its major characteristics, i.e. symptoms, course of illness, age, and probability of outbreak. The descriptions allowed a ranking of the diseases by severity.[75] Ranked as rather mild or least severe was "hemochromatosis" (HE, rank 5), an iron metabolic disorder leading among other things to a malfunctioning of organs. Not drinking alcohol can heavily reduce the negative effects of this disease. Rank 4 was occupied by "alpha1-antitrypsin deficiency" (AL), causing a high risk of chronic lung illnesses. To stop smoking can clearly reduce this risk. The middle category contained "Alzheimer's disease" (AZ, rank 3) and "Chorea Huntington" (CH1, rank 2). Carriers of the CH gene mutation experience the outbreak of this illness of the nervous system before the age of 40 with 100 % probability. It leads to death within a few years. CH is the most severe illness among the well-known monogenetic diseases. We feared that the description of CH might not be perceived as severe enough to be able to explain the answering behaviour predicted by our model for extremely severe illnesses. Therefore we included a question about a blood test for the Lassa virus (LA, rank 1) to be taken after a vacation in Africa. Being tested positive means a high probability of dying within one week.

For each disease, the following question had to be answered by "yes" or "no": "Do you want to be informed about the test result concerning the relevant gene mutation (blood test)?". The answering behaviour is described in table 5.1. In the questionnaire, the diseases were listed in the rank order 1, 3, 5, 4, 2. A supplementary question (CH2) was asked with regards to CH1. In addition to the standard question people were asked whether they would want to know the test result if in their family history there had been cases of CH1. This would raise their ex-ante probability to be carriers of the relevant

[75] The ranking we propose has been confirmed by questioning a different sample of students.

		HE	AL	AZ	$CH1$	$CH2$	LA
1=want	frequ.	211	203	145	166	182	212
info	%	95.0	91.4	65.3	74.8	82.0	95.5
0=doesn't want	frequ.	11	19	77	56	40	10
info	%	5.0	8.6	34.7	25.2	18.0	4.5

Table 5.1: Answering behaviour

mutation to 50%. The question (CH2) was meant to provide evidence for hypothesis 3. The questionnaire also asked for personal data which were of interest for the quality of the answers. Some descriptive statistics of the individuals' characteristics are given in table C1 in the appendix .

5.3.2 Model specification

For the analysis, a probit model was used. In the regression framework

$$y_{ij} = \beta' x_j + u_{ij}$$

y_{ij} can take the value of 1 or 0, depending on whether individual i's answer to question j is "yes" or "no". x_j is a vector of dummy variables, one for each of the six questions asked. The dummies take the value of 1 if the observation y_{ij} belongs to question j, otherwise 0. This is the reference case to start with. To check the robustness of our results and to correct for hidden heterogeneity between individuals, a random effects probit model was estimated. In extensions of the standard probit model, further explanatory variables like age, school grade etc. were included or used to control for the results in certain subsamples.

5.3.3 Results

Table 5.2 shows the results of the probit and random effects probit estimations (coefficients and simulations) for hypotheses 1 and 2. The reference case is AZ (rank 3). The data confirm a significant increase in positive answers from AZ to AL and HE, i. e. from the severe to the light diseases. People thus are more likely to want to know about

Dependent Var: Dummy, 1=yes			
Indep Variables	Probit	% Change (Probit)	Random Effects Probit
LA	1.301	17.1	1.837
	(7.638)**		(8.140)**
CH1	.273	5.1	.415
	(2.175)*		(2.699)**
AL	.974	14.2	1.451
	(6.586)**		(7.249)**
HE	1.255	16.8	1.832
	(7.540)**		(8.014)**
const	.394		.601
	(4.552)**		(4.390)**
Pseudo R^2	.1296	.1296	
Log L	-418.1	-418.1	-378.4
# observations	222	222	222
t-values in brackets	**signif. at 1% level	*signif. at 5% level	

Table 5.2: Probit and random effects probit - reference case: AZ

Dependent Var: Dummy, 1=yes			
Indep Variables	Probit	% Change (Probit)	Random Effects Probit
CH2	.247	7.2	.567
	(1.844)+		(2.689)**
const	.667		1.520
	(7.311)**		(5.678)**
Pseudo R^2	.0074	.0074	
Log L	-230.1	-230.1	-198.9
# observations	222	222	222
t-values in brackets	**signif. at 1% level	+signif. at 10% level	

Table 5.3: Family history of Chorea Huntington - reference case: CH1

mild illnesses as compared to severe ones. There is also a significant increase in positive answers from AZ to LA, but also from AZ to CH1 and from CH1 to LA. People are more likely to want information on extremely severe as compared to severe illnesses. The significant difference in answering behaviour between AZ and CH1, which both were ranked by us as "severe" but not extremely so, could indicate that people indeed perceive CH1 as "extremely severe". But, when analysing various subsamples, it turns out that the significant increase from AZ to CH1 is the only result which is not robust in all subsamples, so it is clear that our initial ranking was indeed correct. In fact, the data confirm the predicted U-shape in the probability for positive answers across the five ranks.

Table 5.3 presents the results of the tests of hypothesis 3. This hypothesis says that individuals will be more likely to want to be informed about a disease if their ex-ante probability to develop this disease is already quite high. Our results confirm that individuals with a family history of Chorea Huntington (CH2) are more likely to want to be informed about the test result than without CH1 in the family history.

As already mentioned, including further explanatory variables does not change the overall picture of our results. The significant differences between severe and extremely severe as well as severe and mild diseases can be found in all subsamples. Nevertheless, some significant deviations can be found concerning the answering behaviour in our middle ("severe") category. Age, lower A-level grade, being a smoker, or drinking alcohol increase the propensity to obtain the information about one's health status even in the case of severe illnesses. The effects are between 2% and 5% only, but significant at least at the 5%-level.[76]

Being privately insured has a significant negative influence on the answering behaviour of nearly 6%. People may either fear - although we emphasised the exclusiveness and private availability of the data - that the results might be used by insurers to increase premia, or publicly insured persons might want to use the information in order to improve their insurance coverage.[77]

5.3.4 Further empirical literature

There is a vast literature in the field of genetics about attitudes towards genetic testing. In all studies, there is a significant fraction of individuals who do not want to become tested.

Most studies have only considered high risk groups, i.e. individuals with a family history of genetic mutations. These studies are therefore only partly useful to provide evidence for the hypotheses we derive from our theory. It is convenient, though, to have a brief look at the literature on Chorea Huntington in order to control the magnitude of our results. Kessler et al. (1987) asked 66 individuals at 50% risk of inheriting Chorea Huntington whether they would use genetic testing if available. 79% of them said they would use it. Mastromauro et al. (1987), when asking 131 persons at risk for Chorea Huntington the same question, obtained 66% of positive answers. The last figure is significantly smaller than the 82% which we obtained in our study but the first one is very similar. It is

[76]Also living in a students' accomodation increases the percentage of positive answers by 7%, but as this is only a small subsample, the result is only weakly significant. Being a student of economics or of economics and business education reduces the willingness to obtain information compared to a student of business administration by 5% and 11%, respectively.

[77]In the case of differences between having or not having a family history of CH, we obtain a similar result for the coefficients of age, being a smoker (positive influence) and being privately insured (negative influence). Being married has a strong and significant positive influence on the willingness to obtain information, but again the subsample is very small. A positive influence here means that the gap between positive answers without (lower probability) and with (higher probability) a family history of CH is smaller and therefore less significant.

obvious, though, that a more valid analysis should be made between groups of individuals with and without a family history of Chorea Huntington within a single study.

Only very few studies can be found that compare the interest in becoming tested between high and low risk groups. One such study has been done by Hofferbert et al. (1998). The authors found an increased propensity to genetic testing in the case of a family history of breast cancer. They analysed the behaviour of 52 families, out of which 29 were high risk families with a family history of the relevant genetic mutation. 97% of the high risk families opted for becoming tested but only 39% of the low risk families did so. Lipkus et al. (1999) compared African-American women with and without a family history of breast cancer (sample size of 130 and 136 women, respectively). Women with such a family history were significantly more interested in genetic testing than the other group. Among women with a family history, 11%, 17%, and 72% reported being not at all/slightly interested, somewhat interested, and interested/very interested, respectively. Among women without a family history the corresponding values were 25%, 16%, and 58%.
Both studies thus confirm our third hypothesis. They also confirm that among low risk individuals, a significant proportion refuses to become tested.[78]

5.4 Conclusion

We present a formal model on the basis of time inconsistent preferences which allows us to discuss the existence of will and the possibility of a breakdown of will. An individual has discretion over one variable, and she faces the trade-off between pleasure today and pleasure in the future. Due to her present biased preferences she has a tendency to consume too much too early. She might devise strategies to overcome this inefficiency. Such a behaviour is called will. However, these strategies are only viable if the future is valued sufficiently highly. If this is not the case, will can not be sustained.

The model is applied to the case of genetic testing. Empirically we observe that many people reject undertaking genetic tests, or more generally, that (some) people do not want to obtain certain information about their health status. We provide an explanation for this observation by arguing that people may rationally prefer not to undertake a genetic

[78]See also Neumann et al. (2001) for a study with hypothetical questions on the willingness to become tested for Alzheimer's disease, and for further references confirming our results.

test if they are afraid that in case of a positive test result they will no longer be able to sustain their will. Comparative statics analysis shows that a test might be undertaken if a positive result indicates a mild illness or a very extreme illness, but not if the severity is in between. Tests will also more likely be undertaken if a person's risk of illness even without a test is already high. Experimental results from a questionnaire distributed to students and evidence from the literature confirm the hypotheses derived from the theory.

A better understanding of the reasons why people might not obtain information about their health status is of crucial importance for political decision making. The hospital example on vCJK discussed in the beginning of this paper is increasingly relevant due to better means of diagnosis and better data available to physicians and hospitals on their patients' health status. Should hospitals/doctors/the government disclose such information? Our model suggests that information disclosure should depend on the immediate threat that this information represents for the individual patient's health. Whereas the patient should be informed on less severe diseases and on imminent lethal threaths, the decision to reveal information on severe illnesses should be carefully thought about and probably not be agreed on without further support to the patient.
Another policy issue concerns genetic testing and insurance. There is a long debate in the literature on whether insurance companies should be allowed to ask for previous results from genetic tests, or even demand genetic tests as a prerequisite for an insurance contract[79]. Suppose that it is decided that the insurer should have the same information as the insured when signing a contract[80] . However, if tests can be made secretly at home, the insurer consequently has to be allowed to demand genetic tests to be undertaken before a contract is signed. This scenario forces the insured to bear the risk of being tested positive. According to our analysis, a further welfare loss occurs because people simply might not want to obtain more information on their health status. A welfare improving policy would then require tests only to be undertaken by certified physicians and not to be freely available to everyone. The insurer might then obtain the information whether a test was undertaken or not and, when a new contract is to be signed, he only needs to ask for results of tests already taken. There would be no need to require further tests to be done.

[79]See e.g. Ad Hoc Committee on Genetic Testing (1995), Hall (1996), Strohmenger and Wambach (2000), Tabarrok (1994, 1996).

[80]Such a regulatory scheme applies in the US life insurance market and in the German and British insurance markets (Chuffart, 1996).

6 Concluding Remarks

In 2000, when I started to work on this thesis, an article by Victor R. Fuchs on "The future of health economics" was published in the Journal of Health Economics. Fuchs suggested five areas where in his opinion health economists could make a significant contribution: Endogenous technology and preferences, social norms, principal-agent problems, behavioural economics, and measurement and analysis of quality of life. Having re-read this article before finishing my thesis, I am glad to see that my research has evolved in a way which might - at least concerning the approaches I took - get Professor Fuchs' blessing. The strategic interaction between patient and physician was central to the analysis of technology choice and adoption of innovations in chapters 2 and 3. Chapter 4 was concerned with a special institutional feature of the health care market, namely the role of medical associations. Although not mentioned by Fuchs as one of the major fields of future research, he was convinced that "institutions matter". Finally, chapter 5 borrowed an approach from behavioural economics to explain the reluctance against genetic testing.

Of course, I hope that not only the approaches but also the contents of this thesis will be a contribution to health economics. To me, this field of economic research is extremely fascinating in its diversity and relevance for political decision making. I am deeply convinced of its sparkling and delightful future.

A Physician Reimbursement and Technology Adoption

A.1 Comparative statics a_{SB}

Given that the Second Order Condition holds, $sgn[\frac{\partial a_{SB}}{\partial j}] = sgn[\frac{\partial FOC}{\partial j}]$ with $j \in \{\pi, \epsilon, Y\}$

$\frac{\partial FOC}{\partial \epsilon} = [\pi U''(S)] \cdot [(1-a) \cdot \frac{\partial m}{\partial a} - m] + m \cdot U''(S) < 0$

$\frac{\partial FOC}{\partial Y} = -[(1-\pi)U''(H) + \pi U''(S)] \cdot [(1-a) \cdot \frac{\partial m}{\partial a} - m] - m \cdot U''(S) \begin{cases} > 0 & \text{if DARA} \\ \leq 0 & \text{if CARA/IARA} \end{cases}$

$\frac{\partial FOC}{\partial \pi} = -[(1-a)c \cdot m] \cdot \{-[(1-a) \cdot \frac{\partial m}{\partial a} - m](1-\pi)U''(H) -$

$-[(1-a) \cdot \frac{\partial m}{\partial a} \cdot \pi + (1-\pi) \cdot m]U''(S)\} - [(1-a) \cdot \frac{\partial m}{\partial a} - m][U'(S) - U'(H)] > 0$

A.2 Comparative statics $\eta_{m_{SB}/c}$

$\eta_{m_{SB}/c} = \frac{a_{SB}}{G''(m_{SB})} \cdot \frac{c}{m_{SB}}$

$\frac{\partial \eta_{m_{SB}/c}}{\partial j} = \frac{G''(\cdot) \cdot m_{SB} \cdot c - a \cdot c \frac{\partial m}{\partial a}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m_{SB}]^2} \frac{\partial a_{SB}}{\partial j}$

$sgn[\frac{\partial \eta_{m_{SB}/c}}{\partial j}] = sgn[G''(\cdot) - \eta_{m_{SB}}(G'''(\cdot)m + G''(\cdot))] \cdot sgn[\frac{\partial a_{SB}}{\partial j}]$ with $j \in \{\pi, \epsilon, Y\}$

Assume that $G'''(\cdot)m + G''(\cdot) \leq 0$.

A.3 Comparative statics $\eta_{m_{SC}/c}$, $\eta_{m_{SC}/n}$

$\eta_{m_{SC}/c} = \frac{\bar{a}}{G''(m_{SC})} \cdot \frac{c}{m_{SC}}$

$\frac{\partial \eta_{m/c}}{\partial \bar{a}} = \frac{G''(\cdot) \cdot m \cdot c - \bar{a} \cdot c \frac{\partial m}{\partial \bar{a}}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

$\frac{\partial \eta_{m/c}}{\partial c} = \frac{G''(\cdot) \cdot m \cdot \bar{a} - \bar{a} \cdot c \frac{\partial m}{\partial c}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

$\frac{\partial \eta_{m/c}}{\partial n} = \frac{-\bar{a} \cdot c \frac{\partial m}{\partial n}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

$\eta_{m_{SC}/n} = \frac{n}{G''(m_{SC}) \cdot m_{SC}}$

$\frac{\partial \eta_{m/n}}{\partial \bar{a}} = \frac{-n \frac{\partial m}{\partial \bar{a}}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

$\frac{\partial \eta_{m/n}}{\partial c} = \frac{-n \frac{\partial m}{\partial c}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

$\frac{\partial \eta_{m/n}}{\partial n} = \frac{G''(\cdot) \cdot m - n \frac{\partial m}{\partial n}[G'''(\cdot)m + G''(\cdot)]}{[G''(\cdot) \cdot m]^2} < 0$

Assume that $G'''(\cdot)m + G''(\cdot) \leq 0$. Subscript $_{SC}$ omitted for reasons of readability.

A.4 Simultaneous parameter changes - Standard Coinsurance

Isoprofit functions: $\Pi = R + (p-c) \cdot m_{SC}(\bar{a}, c, n, B)$

- $\frac{\partial n}{\partial c} = -\frac{\frac{\partial \Pi}{\partial c}}{\frac{\partial \Pi}{\partial n}} = -\frac{m_{SC} \cdot [\frac{p-c}{c} \cdot \eta_{m_{SC}/c} - 1]}{(p-c) \cdot \frac{\partial m_{SC}}{\partial n}} \begin{cases} \geq 0 & \text{if } \eta_{m_{SC}/c} \geq \frac{c}{p-c} \text{ and cost sharing} \\ < 0 & \text{if } \eta_{m_{SC}/c} < \frac{c}{p-c} \text{ and cost sharing} \\ < 0 & \text{if fee-for-service} \\ +\infty & \text{if cost reimbursement or B binding} \end{cases}$

- $\frac{\partial B}{\partial c} = -\frac{\frac{\partial \Pi}{\partial c}}{\frac{\partial \Pi}{\partial B}} = -\frac{m_{SC} \cdot [\frac{p-c}{c} \cdot \eta_{m_{SC}/c} - 1]}{(p-c) \cdot \frac{\partial m_{SC}}{\partial B}} = \begin{cases} < 0 & \text{if cost sharing} \\ > 0 & \text{if fee-for-service} \\ +/-\infty & \text{if cost reimbursement or B not binding} \end{cases}$

- $\frac{\partial B}{\partial n} = -\frac{\frac{\partial m_{SC}}{\partial n}}{\frac{\partial m_{SC}}{\partial B}} = \begin{cases} 0 & \text{if B binding} \\ +\infty & \text{if B not binding} \\ & \text{not def. under cost reimbursement} \end{cases}$

B Demand Inducement for Health Care Technologies

B.1 Comparison of expected utility without treatment and with treatment

It is shown that the difference in expected utiliy between treatment in a pooling equilibrium and no treatment is strictly increasing in r as long as condition (3.3) holds. Rewriting (3.3) gives the differential Δ^p as

$\Delta^p = r \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon)] + (1 - r) \cdot U(Y - P - a(C_i + D) - N_i) - r \cdot U(Y - P - \epsilon) - (1 - r) \cdot U(Y - P) \geq 0$

Taking the derivative with respect to r yields

$\frac{\partial \Delta^p}{\partial r} = p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon) - U(Y - P - a(C_i + D) - N_i) - U(Y - P - \epsilon) + U(Y - P) > 0$

because (3.3) can be transformed to

$p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon) - U(Y - P - a(C_i + D) - N_i) - U(Y - P - \epsilon) + U(Y - P) \geq \frac{U(Y-P) - U(Y-P-a \cdot (C_i+D)-N_i)}{r} > 0$

The same holds for Δ^s although only with weak inequality if diagnosis is costless for the patient. **QED**

B.2 Effects of changes in one technology parameter on the cutoff

The cutoff r^p_{ci} for each technology is implicitly defined by condition (3.5)

$r^p_{ci} \cdot [p_i \cdot U(Y - P - a \cdot (C_i + D) - N_i) + (1 - p_i) \cdot U(Y - P - a \cdot (C_i + D) - N_i - \epsilon] + (1 - r^p_{ci}) \cdot U(Y - P - a(C_i + D) - N_i) =$
$r^p_{ci} \cdot U(Y - P - \epsilon)] + (1 - r^p_{ci}) \cdot U(Y - P)$

Rearranging this and applying the Implicit Functions Theorem yields

$$\frac{\partial r^p_{ci}}{\partial C_i} = \frac{r^p_{ci} \cdot a \cdot [p_i \cdot U'(Y-P-a \cdot (C_i+D)-N_i) + (1-p_i) \cdot U'(Y-P-a \cdot (C_i+D)-N_i-\epsilon)] + (1-r^p_{ci}) \cdot a \cdot U'(Y-P-a(C_i+D)-N_i)}{[p_i \cdot U(Y-P-a \cdot (C_i+D)-N_i) + (1-p_i) \cdot U(Y-P-a \cdot (C_i+D)-N_i-\epsilon) - U(Y-P-\epsilon)] - [U(Y-P-a(C_i+D)-N_i) - U(Y-P)]}$$

The denominator can be simplified using condition (3.5) which yields the following expression:

$$\frac{\partial r^p_{ci}}{\partial C_i} = \frac{r^p_{ci} \cdot a \cdot [p_i \cdot U'(Y-P-a \cdot (C_i+D)-N_i) + (1-p_i) \cdot U'(Y-P-a \cdot (C_i+D)-N_i-\epsilon)] + (1-r^p_{ci}) \cdot a \cdot U'(Y-P-a(C_i+D)-N_i)}{-\frac{1}{r^p_{ci}}[U(Y-P-a(C_i+D)-N_i) - U(Y-P)]}$$

From this it can be seen that $\frac{\partial r^p_{ci}}{\partial C_i} \geq 0$ as long as $a \geq 0$.

Similarly, it can be derived that

$$\frac{\partial r^p_{ci}}{\partial N_i} = \frac{r^p_{ci} \cdot [p_i \cdot U'(Y-P-a\cdot(C_i+D)-N_i)+(1-p_i)\cdot U'(Y-P-a\cdot(C_i+D)-N_i-\epsilon)]+(1-r^p_{ci})\cdot U'(Y-P-a(C_i+D)-N_i)}{-\frac{1}{r^p_{ci}}[U(Y-P-a(C_i+D)-N_i)-U(Y-P)]} \geq 0$$

$$\frac{\partial r^p_{ci}}{\partial p_i} = -\frac{r^p_{ci} \cdot [U(Y-P-a\cdot(C_i+D)-N_i)-U(Y-P-a\cdot(C_i+D)-N_i-\epsilon)]}{-\frac{1}{r^p_{ci}}[U(Y-P-a(C_i+D)-N_i)-U(Y-P)]} \leq 0$$

The equivalent proof holds for r^s_{ci}. **QED**

B.3 Effects of simultaneous parameter changes on the cutoff

Expected utility for the marginal patient at the cutoff r_{ci} increases resp. the cutoff goes down if

$$(a \cdot dC + dN) < \frac{r_{ci} \cdot [U(Y-P-a\cdot C_i-a\cdot D-N_i)-U(Y-P-a\cdot C_i-a\cdot D-N_i-\epsilon)]}{r_{ci} \cdot [p_i \cdot U'(Y-P-a\cdot C_i-a\cdot D-N_i)+(1-p_i)\cdot U'(Y-P-a\cdot C_i-a\cdot D-N_i-\epsilon)]+(1-r_{ci})U'(Y-P-a\cdot C_i-a\cdot D-N_i)} dp$$

$f(r_{ci}) = \frac{r_{ci} \cdot [U(Y-P-a\cdot C_i-a\cdot D-N_i)-U(Y-P-a\cdot C_i-a\cdot D-N_i-\epsilon)]}{r_{ci} \cdot [p_i \cdot U'(Y-P-a\cdot C_i-a\cdot D-N_i)+(1-p_i)\cdot U'(Y-P-a\cdot C_i-a\cdot D-N_i-\epsilon)]+(1-r_{ci})U'(Y-P-a\cdot C_i-a\cdot D-N_i)}$ is strictly increasing in r_{ci}, therefore it suffices to show the effect of certain parameter constellations for $r_{ci} = 0$ and $r_{ci} = 1$. If outcomes differ for these extreme cases then it is clear that they differ for the two groups of healthier and less healthy patients.

$f(0) = 0$ and $f(1) > 0$. Thus, a healthy patients accepts a new technology if $(a \cdot dC + dN) < 0$ whereas an ill patient accepts this technology if $(a \cdot dC + dN) < f(1) \cdot dp$. Assume that at first, everyone is treated with the same technology. Now an innovation enters the market which increases effectiveness, i.e. $dp > 0$. At the same time costs are higher for this technology, but it is still more cost effective than the old one. It is obviously not preferred by the healthy patients because $(a \cdot dC + dN) > 0$ but it may well make the least healthy better off, $(a \cdot dC + dN) < f(1) \cdot dp$. The opposite argument holds for reductions in both effectiveness and costs which will be accepted by the healthy but not by the least healthy patients. **QED**

C Sustaining Will by Staying Ignorant

C.1 Proof of Proposition 7

We first show that for small values of δ we get $\Delta U < 0$, where ΔU is defined in equation (5.9). First note that at $\delta = 0$ it holds that $\Delta U = 0$, as in that case $\sigma^c = \sigma^{nc} = 0$. So it suffices to show that $\frac{d\Delta U}{d\delta}|_{\delta=0} < 0$ to prove the first statement. Using that $\sigma^c = \delta$ and $\sigma^{nc} = \frac{\lambda\delta}{1-\delta-\lambda\delta}$ gives after some reformulations:

$$\frac{d\Delta U}{d\delta}|_{\delta=0} = -\lambda \ln[\lambda] - (1-\lambda) \tag{C.1}$$

which is smaller than zero for all values of $0 < \lambda < 1$.
Next we show that for large values of δ we get $\Delta U > 0$. To do so, multiply ΔU by $(1-\delta)^2$. This expression has the same sign as ΔU. Now this new expression is equal to 0 at $\delta = 1$, so it remains to show that the derivative of this expression with respect to δ at $\delta = 1$ is negative, to prove the statement. We get:

$$\frac{d(1-\delta)^2 \Delta U}{d\delta}|_{\delta=1} = -\lambda \ln[\lambda] - \frac{(1-\lambda)}{\lambda} \tag{C.2}$$

which is also smaller than zero as long as $\lambda < 1$. This proves Proposition 7. **QED**

Variable	*Mean*	*Std.Dev.*	*Description*
persnr			number of person
yesno			1=answer "yes, want info", 0="no"
LA			1=observation refers to question "LA", 0=not
			"LA"=Lassa virus
CH1			1=observation refers to question "CH1", 0=not
			"CH1"=Chorea Huntington
CH2			1=observation refers to question "CH2", 0=not
			"CH2"=Chorea Huntington with family history
AZ			1=observation refers to question "AZ", 0=not
			"AZ"=Alzheimer's disease
AL			1=observation refers to question "AL", 0=not
			"AL"=Alpha1-antitrypsin deficiency
HE			1=observation refers to question "HE", 0=not
			"HE"=Hemochromatosis
age	22.216	1.972	age in years
sex	.627	.484	1=male, 0=female
alevel	.977	.149	1="Abitur"(A-Level), 0=else
alevgrad	2.183	.589	A-Level grade: best grade 1.0, worst 4.0
economic	.359	.480	1=student of economics, 0=else
business	.482	.500	1=student of business administration,0=else
econteac	.141	.348	1=student of economics education, 0=else
diffcare	.018	.134	1=different career from above, 0=else
smoke	2.484	.761	1=smoke regularly, 2=irreg, 3=not
drink	1.927	.490	1=drink regularly, 2=irreg, 3=not
health	7.521	2.134	0=very unhappy with health status, 10=very happy
insur	.688	.463	1=private (or priv. add.) insurance, 0=state insurance only
genmut	.073	.261	1=genetic mutation known in family, 0=not

Table C.1: Description of variables and sample characteristics

References

[1] Ad hoc committee on genetic testing: Insurance issues (1995): Background statement: Genetic testing and insurance, *American Journal of Human Genetics* 56, 327-331

[2] Ainslie, George (1992): Picoeconomics - The strategic interaction of successive motivational states within the person, Cambridge University Press

[3] Ainslie, George (1999): Breakdown of will, Working paper presented at GREMAQ/CEPR conference "Economics and Psychology", Toulouse

[4] Ainslie, George (2001): Breakdown of will, Cambridge University Press

[5] Ainslie, George, Haslam, Nick (1992): Self-Control, in: Loewenstein, George, Elster, Jon, eds., Choice over time, Russell Sage Foundation, 177-209

[6] Arrow, Kenneth J. (1963): Uncertainty and the welfare economics of medical care, *American Economic Review* 53, 941-973

[7] Asheim, Geir B. (1997): Individual and collective time-consistency, *Review of Economic Studies* 64, 427-443

[8] Azfar, Omar (1999): Rationalizing hyperbolic discounting, *Journal of Economic Behaviour and Organization* 38, 245-252

[9] Baker, Laurence C. (2000): Managed care and technology adoption in health care: Evidence from magnetic resonance imaging, *NBER Working Paper No. 8020*

[10] Baker, Laurence C., Phibbs, Ciaran S. (2002): Managed care, technology adoption, and health care: The adoption of neonatal intensive care, *RAND Journal of Economics* 33, 524-548

[11] Baumgardner, James R. (1991): The interaction between forms of insurance contract and types of technical change in medical care, *Rand Journal of Economics* 22, 36-53

[12] Benabou, Roland, Tirole, Jean (2002): Self-confidence and personal motivation, *Quarterly Journal of Economics* 117, 871-915

[13] Benstetter, Franz, Wambach, Achim (2001): Strategic interaction in the market for physician services: The treadmill effect in a fixed budget system, *CESifo Working Paper No. 427*

[14] Bokhari, Farasat A. S. (2001): Managed care and the adoption of hospital technology: The case of cardiac catheterization, *HEW Economics Working Paper Archive* 0110001

[15] Breyer, Friedrich, Zweifel, Peter, Kifmann, Mathias (2003): Gesundheitsökonomie, Berlin, Heidelberg, New York: Springer-Verlag

[16] Brocas, Isabelle, Carillo, Juan D. (2000): The value of information when preferences are dynamically inconsistent, *European Economic Review* 44, 1104-1115

[17] Caillaud, Bernard, Cohen, Daniel, Jullien, Bruno (1999): Towards a theory of self-restraint, Working paper presented at GREMAQ/CEPR conference "Economics and Psychology", Toulouse

[18] Caplin, Andrew, Leahy, John (2001): Psychological expected utility theory and anticipatory feelings, *Quarterly Journal of Economics* 116, 55-79

[19] Carillo, Juan D., Mariotti, Thomas (2000): Strategic ignorance as a self-disciplining device, *Review of Economic Studies* 67, 529-544

[20] Chuffart, Andre (1996): Genetic testing in Europe, *Journal of Insurance Medicine* 28, 125-135

[21] Coleman, John, Katz, Elihu, Menzel, Herbert (1957): The diffusion of an innovation among physicians, *Sociometry* 20, 253-270

[22] Cutler, David M., McClellan, Mark (1996): The determinants of technological change in heart attack treatment, *NBER Working Paper No. 5751*

[23] Cutler, David M., McClellan, Mark (2001): Productivity change in health care, *AEA Papers and Proceedings* 91, 281-286

[24] Cutler, David M., Huckman, Robert S. (2002): Technological development and medical productivity: Diffusion of angioplasty in New York State, *NBER Working Paper No. 9311*

[25] Danzon, Patricia (2000): Liability and liability insurance for medical malpractice, in: Anthony J. Culyer and Joseph P. Newhouse, eds., Handbook of Health Economics, Amsterdam: Elsevier, 1341-1404

[26] De Jaegher, Kris, Jegers, Marc (2001): The physician-patient relationship as a game of strategic information transmission, *Health Economics* 10, 651-668

[27] Demange, Gabrielle, Geoffard, Pierre-Yves (2002): Reforming incentive schemes under political constraints: The physician agency, *CEPR Discussion Paper No. 3589*

[28] Dranove, David (1988): Demand inducement and the physician/patient relationship, *Economic Inquiry* 26, 281-298

[29] Ellis, Randall P., McGuire, Thomas G. (1990): Optimal payment systems for health services, *Journal of Health Economics* 9, 375-396

[30] Ellis, Randall P., McGuire, Thomas G. (1993): Supply-side and demand-side cost sharing in health care, *Journal of Economic Perspectives* 7, 135-151

[31] Emons, Winand (1997): Credence goods and fraudulent experts, *RAND Journal of Economics* 28, 107-119

[32] Emons, Winand (2001): Credence goods monopolists, *International Journal of Industrial Organization* 19, 375-389

[33] European Observatory on Health Care Systems (2000): Health care systems in transition. Germany 2000

[34] Fuchs, Victor R. (1996): Economics, values, and health care reform, *American Economic Review* 86, 1-24

[35] Fuchs, Victor R. (2000): The future of health economics, *Journal of Health Economics* 19, 141-157

[36] Glied, Sherry (2000): Managed care, in: Anthony J. Culyer and Joseph P. Newhouse, eds., Handbook of Health Economics, Amsterdam: Elsevier, 707-753

[37] Goddeeris, John H. (1984a): Insurance and incentives for innovation in medical care, *Southern Economic Journal* 51, 530-539

[38] Goddeeris, John H. (1984b): Medical insurance, technological change, and welfare, *Economic Inquiry* 22, 56-67

[39] Hall, Jane (1996): Consumer utility, social welfare, and genetic testing. A response to 'Genetic testing: An economic and contractarian analysis', *Journal of Health Economics* 15, 377-380

[40] Harris, Christopher, Laibson, David (2001): Dynamic choices of hyperbolic consumers, *Econometrica* 69, 935-57

[41] Harris, Christopher, Laibson, David (2003): Hyperbolic discouting and consumption, in: Dewatripont, Mathias et al., eds., Advances in economics and econometrics (Vol. 1). Theory and applications, Eighth World Congress of the Econometric Society, Cambridge University Press

[42] Hillman, Bruce J. (1986): Government health policy and the diffusion of new medical devices, *Health Services Research* 21, 681-711

[43] Hofferbert, Sigrun et al. (1998): Familiäres Mamma- und Ovarial-Carcinom: Interdisziplinärer Ansatz zur genetischen Testung, *Medizinische Genetik* 10, 253-255

[44] Kane, Nancy M., Manoukian, Paul D. (1989): The effect of the medicare prospective payment system on the adoption of new technology, *New England Journal of Medicine* 321, 1378-1383

[45] Kessler, Seymour et al. (1987): Attitudes of persons at risk for Huntington Disease toward predictive testing, *American Journal of Medical Genetics* 26, 259-270

[46] Kesteloot, Katrien, Voet, Nathalie (1998): Incentives for cooperation in quality improvement among hospitals - The impact of the reimbursement system, *Journal of Health Economics* 17, 701-728

[47] Kocherlakota, Narayana R. (1996): Reconsideration-proofness: A refinement for infinite horizon time inconsistency, *Games and Economic Behaviour* 15, 33-54

[48] Kühn, Astrid, Wambach, Achim (2001): Breakdown of will and the value of information, *CEPR Discussion Paper No.3111*

[49] Laibson, David (1996): Hyperbolic discount functions, undersaving, and savings policy, *NBER Working Paper No. 5635*

[50] Laibson, David (1997): Golden eggs and hyperbolic discounting, *Quarterly Journal of Economics* 112, 443-477

[51] Laibson, David (1998): Life-cycle consumption and hyperbolic discount functions, *European Economic Review* 42, 861-871

[52] Laibson, David, Repetto, Andrea, Tobacman, Jeremy (1998): Self-control and saving for retirement, *Brookings Papers on Economic Activity* 1, 91-196

[53] Leland, Hayne E. (1979): Quacks, lemons, and licensing: A theory of minimum quality standards, *Journal of Political Economy* 87, 1328-1346

[54] Lerman, Caryn et al. (1988): What you don't know can hurt you: Adverse psychologic effects in members of BRCA1-linked and BRCA2-linked families who decline genetic testing, *Journal of Clinical Oncology* 16, 1650-1654

[55] Lipkus, Isaac M. et al. (1999): Relationships among breast cancer concern, risk perceptions, and interest in genetic testing for breast cancer susceptibility among African-American women with and without a family history of breast cancer, *Cancer Epidemiology, Biomarkers & Prevention* 8, 533-539

[56] Loewenstein, George, Elster, Jon (1992): Choice over time, New York: Russell Sage Foundation

[57] Ma, Ching-To Albert (1994): Health care payment systems: Cost and quality incentives, *Journal of Economics and Management Strategy* 3, 93-112

[58] Ma, Ching-To Albert (1998): Health care payment systems: Cost and quality incentives - Reply, *Journal of Economics and Management Strategy* 7, 139-142

[59] Ma, Ching-to A., McGuire, Thomas G. (1997): Optimal health insurance and provider payment, *American Economic Review* 87, 685-704

[60] Mastromauro, Carol, Myers, Richard H., Berkman, Barbara (1987): Attitudes toward presymptomatic testing in Huntington Disease, *American Journal of Medical Genetics* 26, 271-282

[61] McGuire, Thomas G. (2000): Physician agency, in: Anthony J. Culyer and Joseph P. Newhouse, eds., Handbook of Health Economics, Amsterdam: Elsevier, 461-536

[62] Neumann, Peter J. et al. (2001): Public attitudes about genetic testing for Alzheimer's disease, *Health Affairs* 20, 252-264

[63] Newhouse, Joseph P. (1992): Medical care costs: How much welfare loss?, *Journal of Economic Perspectives* 6, 3-21

[64] O'Donoghue, Ted, Rabin, Matthew (1999a): Doing it now or later, *American Economic Review* 89, 103-124

[65] O'Donoghue, Ted, Rabin, Matthew (1999b): Procrastination in preparing for retirement, in: Aaron, Henry J., ed., Behavioral dimensions of retirement economics, Washington D.C.: Brookings Institution Press; New York: Russel Sage Foundation, 125-156

[66] Phelps, Charles E. (1992): Diffusion of information in medical care, *Journal of Economic Perspectives* 6, 23-42

[67] Phelps, Charles E., Mooney, Cathleen (1993): Variations in medical practice use: Causes and consequences, in: Arnould, Richard J., Rich, Robert F., White, William D., eds., Competitive approaches to health care reform, Washington D.C., 139-178

[68] Phelps, Edmund S., Pollak, Robert A. (1968): On second best national saving and game-equilibrium growth, *Review of Economic Studies* 35, 185-199

[69] Pitchik, Carolyn, Schotter, Andrew (1987): Honesty in a model of strategic information transmission, *American Economic Review* 77, 1032-1036

[70] Pitchik, Carolyn, Schotter, Andrew (1993): Information transmission in regulated markets, *Canadian Journal of Economics* 26, 815-829

[71] Schmidt, Klaus (1996): The costs and benefits of privatization - An incomplete contracts approach, *Journal of Law, Economics and Organization* 12, 1-24

[72] Shapiro, Carl (1986): Investment, moral hazard, and occupational licensing, *Review of Economic Studies* 53, 843-862

[73] Sieff, Elaine M., Dawes, Robyn M., Loewenstein, George (1999): Anticipated versus actual responses to HIV test results, *American Journal of Psychology* 112, 297-311

[74] Strohmenger, Rainer, Wambach, Achim (2000): Adverse selection and categorical discrimination in health insurance markets: The effects of genetic tests, *Journal of Health Economics* 19, 197-218

[75] Strotz, Robert H. (1956): Myopia and inconsistency in dynamic utility maximization, *Review of Economic Studies* 23, 165-180

[76] Sülzle, Kai, Wambach, Achim (2002): Insurance in a market for credence goods, *CESifo Working Paper* 677

[77] Tabarrok, Alexander (1994): Genetic testing: An economic and contractarian analysis, *Journal of Health Economics* 13, 75-91

[78] Tabarrok, Alexander (1996): Genetic testing and human welfare: Reply to Hall, *Journal of Health Economics* 15, 381-384

[79] Wolinsky, Asher (1993): Competition in a market for informed experts' services, *RAND Journal of Economics* 24, 380-398

[80] Wolinsky, Asher (1995): Competition in markets for credence goods, *Journal of Institutional and Theoretical Economics* 151, 117-131

[81] Weisbrod, Burton 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

[82] Zeckhauser, Richard (1970): Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives, *Journal of Economic Theory* 2, 10-26

[83] Zweifel, Peter, Eichenberger, Reiner (1992): The political economy of corporatism in medicine: Self-regulation or cartel management?, *Journal of Regulatory Economics* 4, 89-108

http://www.kbv.de/publikationen/2803.htm, 23/02/04

http://www.kvberlin.de/Homepage/publikation/archiv/pk2002/pk041202.html, 05/12/03

Lebenslauf

Persönliche Daten

Name	Astrid Selder
geboren am	22.11.1973
Geburtsort	Augsburg

Schullaufbahn

09/1980 – 07/1984	Grundschule an der Schackstraße, Augsburg
09/1984 – 07/1994	Maria-Ward-Gymnasium Augsburg Juli 1994: Allgemeine Hochschulreife
02/1991 – 01/1992	Colegio Compañia de Maria, Viña del Mar, Chile

Akademischer Lebenslauf

10/1994 – 07/1996	Studium der Volkswirtschaftslehre, Albert-Ludwigs-Universität Freiburg
10/1996 – 06/1997	Third Year Undergraduate Studies of Economics, University of Cambridge, UK
11/1997 – 06/2000	Studium der Volkswirtschaftslehre, Ludwig-Maximilians-Universität München Juni 2000: Diplom-Volkswirtin
09/2000 – 04/2004	Wissenschaftliche Mitarbeiterin, Seminar für Versicherungswissenschaft, Ludwig-Maximilians-Universität München
03/2002 – 08/2002	Forschungsaufenthalt am Centre for Health Economics, University of York, UK