Nordic Journal of Political Economy

Volume 25 1999 Pages 95-114

Imperatives in Health Care: Implications for Social Welfare and Medical Technology

Carl Hampus Lyttkens

Carl Hampus Lyttkens *

Imperatives in Health Care: Implications for Social Welfare and Medical Technology

The relationship between a doctor and his patient is in many ways a peculiar one. Being an economist, you like to see patients as consumers who purchase health care in order to improve their health. However, as a health economist, you have to realise that patients are very poorly informed consumers. "The noteworthy point is not simply that it is difficult for the consumer to judge quality before the purchase [...] but that it is difficult even after the purchase" (Weisbrod, 1978:52). In other words, we often do not know if we have a medical problem, we do not know what good health care can do for us, and, if we receive care, we do not know *ex post* how much it has contributed to our change in health status.

There is thus a fundamental informational asymmetry in the doctor-patient relationship, because the doctor is usually much better informed about these matters. Or, as a surgeon once said to me, "I can of course always persuade the patient to choose the treatment which I think is best for him." As a patient, faced with a choice of treatment, you frequently find yourself asking your physician, "Doctor, what would you have chosen in my place?" It appears that the informational asymmetry is especially difficult to overcome and more keenly felt in matters of health and health care, even though the individual faces similar problems when dealing with, for example, real estate agents or lawyers.

The situation is often viewed as the relationship between a principal (the patient) and his agent (the doctor) (Mooney & Ryan, 1993). Information on health care apart, this raises the question whether the patient can communicate his preferences to the physician, and whether the physician has an incentive to act as a perfect agent. Since economists assume that actors are governed by an instrumental rationality in their actions (strive towards certain goals), the issue is

* Department of Community Medicine and Department of Economics, Lund University, Sweden. This research has received financial support from the Swedish Medical Research Council, the Swedish Council for Social Research, the Council for Health Care Research (HSF), and the Siamon foundation. Helpful comments from an anonymous referee, the participants in the conference "Innovation, Policy and Society" of the Nordic Journal of Political Economy, as well as from many other colleagues, are gratefully acknowledged.

whether the doctor's goals coincide with those of his patient.

It is in the meeting between doctors and patients that many of the important decisions in health care are taken, both with respect to the health and welfare of patients and with respect to resource use and resource allocation in the health care sector. The health care sector constitutes some 6–14% of GDP in industrialised countries, and the budget of a single production unit such as the Radiology Department at the Lund University Hospital is of the respectable magnitude of SEK 150 million (≈ $20 million).

Consequently, economists have been very interested in the characteristics of the doctor-patient relationship. Already in 1968, the American health economist Victor Fuchs coined the term "technological imperative," when he argued that resource allocation within or to the health care sector reflects the principle that one should "always give the best care that is technically possible." Most people in the business would probably agree that this is still an apt description of medical decision-making in practice in industrialised countries.

There are several aspects of the beliefs, preferences, and incentive structure surrounding medical decision-making, which may help explain the existence of such a principle. It is important to note, however, that there are several possible interpretations of the technological imperative, each with its own rationale. Hence this essay begins with an exploration into the possible interpretations of the technological imperative. After presenting a simple theoretical framework, we then focus on the effects of the different versions of the imperative on the use of medical technology, on social welfare, and medical research. The paper ends with some concluding remarks.

The relationship between the institutional characteristics and incentives of the health care sector and the use of technology is of considerable interest. In the short and medium run, the incentives to adopt new technologies are important both for health care costs and for the ensuing effects on health, and the factors which determine the diffusion of technologies are increasingly being investigated (Cutler & McClellan, 1996; Escarce, 1996). In the longer run, the incentives to develop new technologies are among the most fundamental factors in the future development of society, together with the factors, which shape institutional change (North, 1990). There is also a mutual dependence, in that we can only use such technologies as come into existence and the propensity to use new technologies today influence the activities of those who develop the technologies of tomorrow.

Three interpretations of the technological imperative: the professional imperative, the capital-biased imperative and the health imperative

There are at least three possible and not mutually exclusive interpretations of the technological imperative. Firstly, it is often said that everybody who is sick should be given all possible medical care, all care that may improve his health. We can label this principle *the professional imperative*. Implicit in the statement is also a disregard for the costs involved.

One obvious reason for this way of thinking to permeate the health care sector is that neither the patient nor the physician usually has an incentive to bother about the costs of a treatment. From the patient's point of view, there is invariably some sort of third-party financing, some insurance system, which ensures that you only pay very little at

the point of consumption. Similarly, the reimbursement of the producer (physicians, hospitals, etc.) has often in practice taken the form of retrospective reimbursement of any costs that have been incurred. This is true both for the public-budget systems, like the Nordic ones, and private systems, like in the US. Hence several moral hazard effects have been operational. To a certain extent this has been changing, with Health Maintenance Organization-type arrangements in the US, an increasing use of DRG payment schemes, and the introduction of capitation payment, financial incentives and quasi-markets in Europe.[1] However, in many countries much of the "soft budget-constraint" remains.

Another obvious aspect of the professional imperative is that it corresponds to the ethics traditionally imbued by medical training (in fact, this is the reason for its label). It satisfies the ethical principle of positive beneficence – to do good. Furthermore, one reason why economists and the medical profession have problems in communicating could be that whereas economics assumes instrumental rationality, a good part of medical ethics looks more like duty ethics (deontological ethics), i.e., ethics not concerned with outcomes or consequences but with the goodness of acts in themselves. The professional imperative could be seen as a reflection of duty ethics, the duty to always help the person in need of medical attention.[2] The Good Samaritan also comes to mind.

There are also other relevant aspects of the physician's decision environment. Physicians are sometimes accused of engaging in "defensive medicine." For example, it may be better to order an additional laboratory test, rather than running the risk of later being blamed for having failed to order it. This tendency may exist in a soft form, but also in the nature of physician liability for malpractice or direct regulations of medical practice.

Regarding the different origins of the professional imperative, it is the latter that seems most likely to generate a conflict of interest between the physician and his patient. Given the physician's influential position with respect to medical decision-making, it has often been noted that he can create additional demand for his own services, by deciding how much treatment to give to patients. The possibility of supplier-induced demand (SID) has been extensively discussed. It appears well-established that physicians sometimes react to financial incentives, so that treatment decisions in the aggregate are not only governed by the health of patients (Cromwell & Mitchell, 1986; Grytten et al., 1995; Rochaix, 1993). The best way to define SID is probably that it occurs when a patient receives more care than he would have wished to consume if he had had the same information as the physician.[3]

1. A Health Maintenance Organization provides health care on the basis that payment is received one year in advance with the subsequent obligation to deliver whatever health care is needed during the year. The term capitation system – while denoting the same principle of prepayment – is often used for payment to providers with a more limited responsibility to provide care, such as GPs. DRG (Diagnose Related Groups) payment represents a prospective reimbursement scheme in the sense that the amount a provider will be paid for a patient with a particular diagnosis is determined in advance.
2. The Swedish Government Committee on Priorities in Health Care (*Vårdens svåra val*) seems to adopt a similar position, when it is argued that all health care for which there is need should be collectively financed – where "need" apparently signifies everything that improves your health or your quality of life – and that the cost-effectiveness of care should *only* be an issue when choosing between treatment alternatives for the same disease.
3. Given this definition, there can obviously be inefficiently high levels of health care consumption even in the absence of SID.

This could become the result of defensive medicine, if that incentive is added to the (lack of) financial incentives for doctors and patients.[4]

Before leaving the professional imperative, we should note that this term evokes another perspective on resource allocation in the health care sector. As already suggested above, the professional imperative has a natural affinity to the notion of treatment being allocated according to need, which appears as a prominent objective of health care policy in many countries. While the technological imperative could be said to apply equally to all patients, treatment according to need suggests how resources should be allocated between patients. There are many possible interpretations of the principle of "treatment according to need" (Culyer & Wagstaff, 1993), but one that seems to be of particular relevance in health policy is that those in bad health should be given priority.[5] The implications of such an interpretation of the professional imperative will also be briefly discussed below.

The professional imperative seems to capture what Fuchs primarily had in mind. However, a second possibility is that there should be an emphasis on the word "technological", and that we can interpret the technological imperative as a *capital-biased imperative*. This implies a tendency to use too much capital in the production of health care.

The most obvious reason for this to occur is perhaps that patients erroneously equate good care with high-tech care. Such a perception would not be difficult to explain. Remember first that patients often do not have exact information on the quality of different health care options. Hence they have to use proxies. In the new institutional economics, it is emphasised that in order to properly analyse individual behaviour, we have to take the individuals' belief systems into account (North, 1990). In the present case, the pertinent fact is that it seems reasonable to assume that – in people's minds – the high standard of living in the industrialised world is intimately associated with the development of productivity enhancing capital equipment. Hence it is not surprising if patients tend to take the use of capital as a proxy for quality. The same argument may apply to physicians.

Furthermore, it has been suggested that physicians have a penchant for capital equipment: "...there is no denying that physicians have been trained to favor sophisticated gadgetry" (Harris, 1977:480). Possession of the latest technological equipment can in our society easily turn into a status matter. Acquisition of new technology can be a way for hospitals to attract patients, referrals or physicians to their staff.[6] The existence of a capital-biased imperative is supported by some empirical evidence which suggests that hospitals tend to be over-capitalised (Jensen & Morrisey, 1986).[7] There is also informal evidence which suggests, for example, that there are reputedly

4. A related aspect is the unpleasantness of unpleasant choices. It is not nice to have to explicitly give priority to one group of patients over another. A natural reaction of the decision-maker is to try to get more resources, and this may happen from the political level down to individual physicians (Broomé et al., 1994; Hernes, 1975). At the political level, for example, explicit prioritising will create dissatisfaction among pressure groups.
5. Cf., e.g., the report of the Swedish Government Committee on Priorities in Health Care (*Vårdens svåra val*).
6. Dozet et al. (forthcoming). A high degree of competition among hospitals sometimes generate higher average costs, unless the competition is carefully managed (Robinson & Luft, 1988; Pope, 1989; Melnick et al., 1992).
7. If producers compete for patients in order to make better use of idle capital equipment, the capital-biased imperative may lead to supplier-induced demand and thereby an increase in health care expenditure.

far too many MR-scanners in the Stockholm area in Sweden, so that public hospitals find themselves looking for patients.

The third and final interpretation of the technological imperative is that is there is a *health imperative.* A new technology will *always* be chosen if its implementation in practice leads to a higher level of health, and will *never* be chosen if it leads to a lower level of health than the old technology.[8] This would suggest that when choosing between two technologies, there is a tendency to only take their effects on health into account. This seems to be a natural extension of the technological-imperative concept, and implies a sort of bounded rationality (Simon, 1979) which does not seem implausible. For example, it may suggest that the relationship between health and welfare is ignored (since costs are ignored), and perhaps that health is more easily observed than welfare and therefore also taken as a proxy for welfare. The expression "in practice" above is used to suggest that perhaps only the solutions/applications of a technology that would actually occur are contemplated – not all the possible scenarios. This is a natural way to simplify a decision problem (bounded rationality once again). Finally, we may note that the health imperative accords well with medical ethics.

Taken together, these different aspects of the technological imperative have profound implications for the diffusion and development of medical technology. For example, a new technology will usually be first applied to those patients who are expected to benefit the most. There will however be a strong pressure to extend the application of the new technology to all other groups that might benefit from it, for example, to the elderly. The technological imperative may thus be one factor behind the quite noticeable diffusion of technology across age groups (Broomé et al., 1994; Dozet, Nystedt & Lyttkens, 1997). A related and important factor is the principle of equality of access to health care, which is found in many policy documents (Mooney et al, 1991). Let us now take a closer look at the implication of these different imperatives for social welfare and the use of medical technology.

A simple theoretical framework[9]

Think of time as consisting of two periods. Social welfare is a function of consumption and health, and health is produced by allocating resources to health care. Health is assumed to be independent of the level of consumption. Both health care and consumption are produced only by current expenditure in each period (no physical capital survives from period 1).

In addition, in the first period, it is possible to allocate resources to medical research. It is reasonable to assume that we engage in research because we expect that this will have a positive impact on the relationship between health care and health in the second period.[10] For the choice in period 2, it is

8. In other words, an increase in health is a sufficient condition and a non-decrease is a necessary condition.
9. In the main text, the analysis will be presented in an intuitive form. An algebraic version of the model is presented in the Appendix, together with an analysis of the effects of the professional imperative in period 1.
10. It is also assumed that research has a positive impact on the marginal utility of health care. This seems reasonable, since research is assumed to have a positive impact on the relationship between health care and health. There is, however, a countervailing force, in that the positive impact of medical research implies that it increases the level of health and this in itself tends to reduce the marginal utility of health (and health care). The latter force seems unlikely to dominate, given that new medical technologies only affect incremental resources in health care.

assumed that the result of the research (in period 1) is that precisely one new technology has become available. Consequently there is a choice to be made between the old and the new health care technology.[11] The nature of the new technology is not known, however, until the second period. Note that since the health imperative concerns a choice between technologies it will only be relevant in period 2.

The social objective is to maximise welfare over the two periods, by allocating resources to consumption, health care, and (in period 1) medical research. The total amount of resources available for spending in the two periods is fixed. While both consumption and health increase welfare, they are assumed to do so at a decreasing rate. Furthermore, health care improves health at a decreasing rate. Resources are freely transferable between periods, and there is no rate of interest.[12]

The choice in period 1

The professional imperative

With the simplifying assumptions outlined in the previous section, the choice in period 1 can now be illustrated by Figure 1. For ease of presentation, the Figure is drawn for a given level of spending on medical research. The horizontal axis in the Figure represents the remainder of the available resources, which is to be spent on health care and consumption. Health care is measured from the left and consumption from the right, so every point on the axis represents a distribution of resources between health care and consumption.

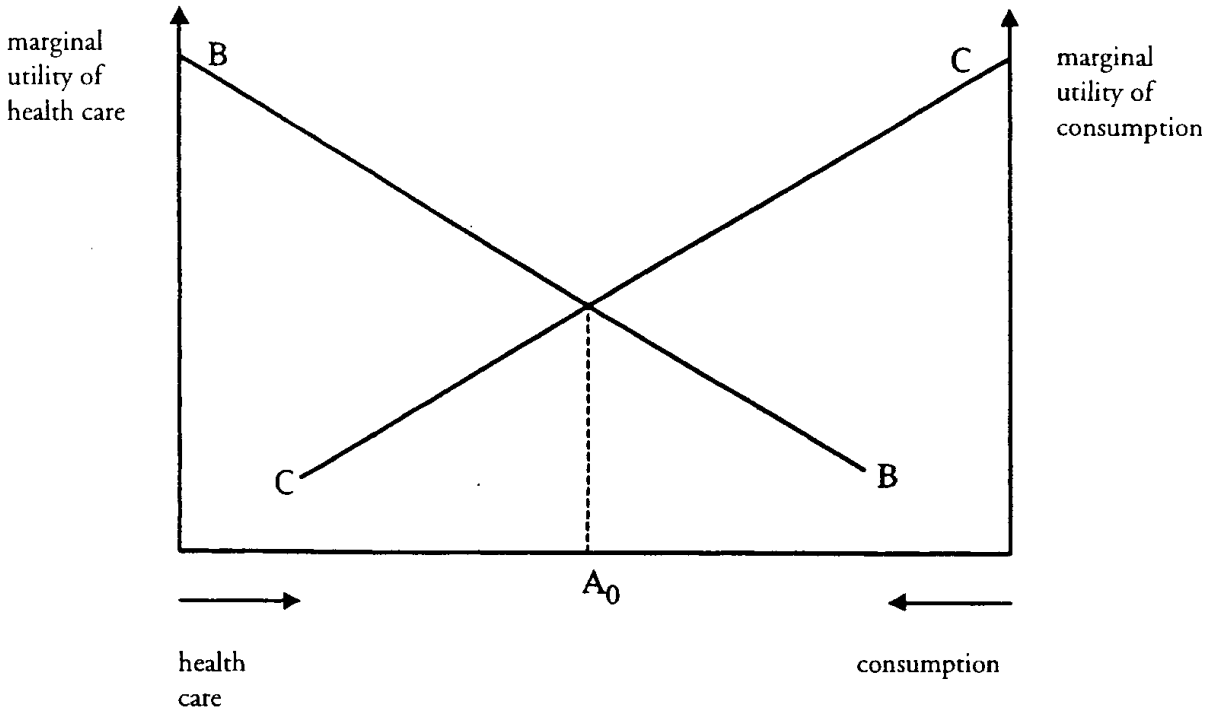
When making choices in period 1, account is taken of what will happen in both periods. Hence the resources are allocated between periods and between the different uses, where we should think of resource allocation for period 2 as a plan made up in period 1. (When we come to period 2, the amount of resources available is determined by how much was spent in period 1, but a new choice will be made with respect to the distribution of these resources between health care and consumption.)

Along the vertical axes, we measure the marginal utility of health care and consumption respectively, i.e., the additional utility we get by spending somewhat more on the item in question. The marginal utility of health care is represented by the BB-curve. It is falling as we increase the level of health care spending (move from left to right), because health affects utility positively but at a decreasing rate (and also because health care affects health at a decreasing rate). The marginal utility of consumption is given by the CC-curve. Correspondingly, it is falling as we increase the level of consumption (move from right to left), because the consumption affects utility positively but at a

11. In the present framework this does not appear restrictive. If more than one new technology becomes available, the old technology will be compared to the best of the new ones.

12. A positive rate of interest seems unlikely to change the *qualitative* nature of the results with respect to the effects on resource allocation of introducing a technological imperative. For example, it would not alter the fact that making the professional imperative more stringent would reduce research and consumption (in both periods) while increasing health care in period 1. The main effect of introducing a positive rate of interest would be to make research less attractive generally, because the cost of research accrues in period 1 while the benefits accrue in period 2; for health care and consumption, costs and benefits accrue in the same period. Some other effects with respect to the intertemporal distribution of health care and consumption would tend to cancel out, because although a positive rate of interest would increase the value of saving resources to the second period, benefits accruing in that period would be discounted at the same rate. The situation becomes more complex if health and financial effects are in practice subject to different rates of discount, as some results suggest (Cairns, 1992).

Figure 1: Resource allocation without imperatives

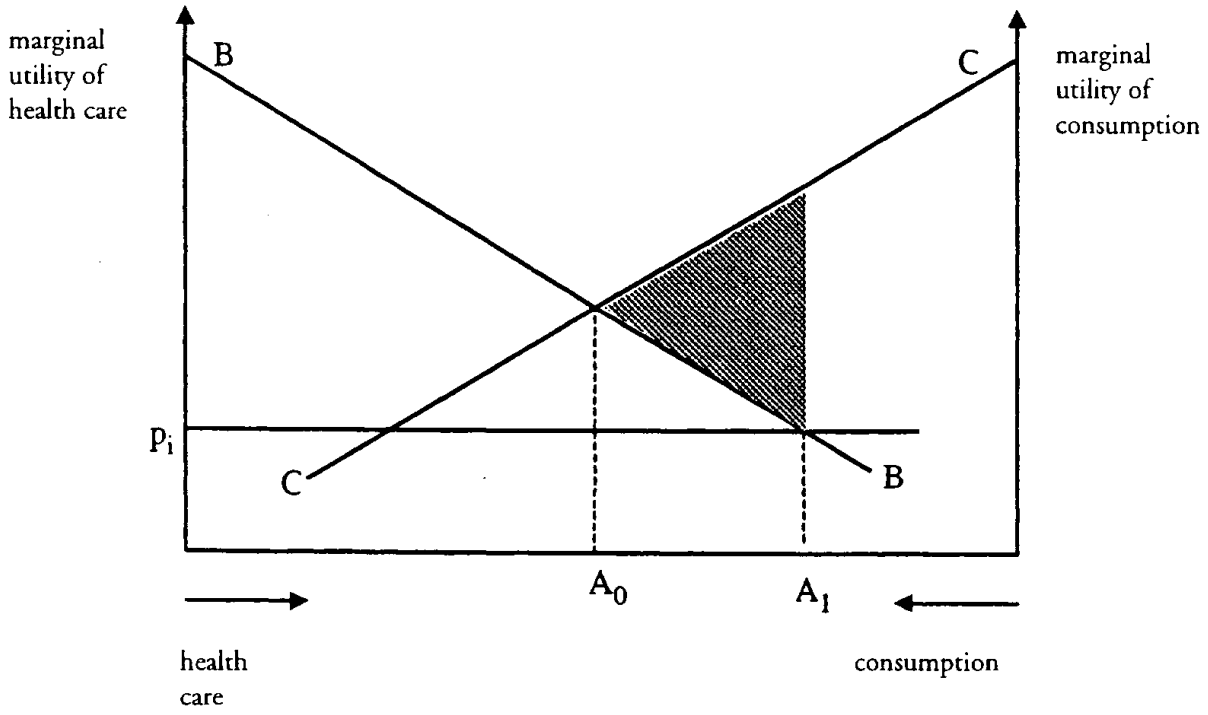


decreasing rate. If this was all, the optimal allocation of resources would be found at point A_0, where the marginal utility of health care and consumption is equal.[13] At all other points, total utility can obviously be increased by moving towards A_0.

Now let us introduce the professional imperative. Analytically, we may think of this as a requirement that the marginal utility of health care resources (the welfare we derive from the last dollar spent on health care) should be "small"; in particular, that it should be smaller than it would have been without the imperative guiding our actions; smaller than a straightforward trade-off between spending on health care and consumption would imply. This seems to be a reasonable way to operationalise the argument that "everything possible" should be done for patients.[14] One should note carefully that this is a statement about the value of the incremental resources spent on health care. This value may be small even though many health care interventions are extremely important for the health and welfare of the population.[15]

13. We will only be dealing with interior solutions throughout the analysis.
14. The interpretation of the professional imperative as a requirement to help those in poor health first will be discussed below in conjunction with the analysis of the capital-biased imperative.
15. Many years ago, Maynard (1983: 31-32), argued that "...large increases in health care expenditure [...] continue to be advocated today as a means of improving health status when the evidence indicates that marginal resources should be allocated on services other than health care if health status improvements are to be maximized [...] it seems more likely that additional health service inputs do little, if anything, to reduce mortality or morbidity indicators of health outcome." With aggregate data, it has been notoriously difficult to show that marginal resources devoted to health care have a positive impact on population health (McGuire et al., 1993), but a recent study shows that such evidence might be forthcoming (Søgaard, 1997).

Figure 2: Resource allocation with a professional imperative (p_i)



In Figure 2, the professional imperative is represented by the level p_i on the axis for marginal utility of health care. To ensure that the marginal utility of health care falls to this level, health care expenditure must be expanded from A_0 to A_1 (consumption being correspondingly reduced).

The straightforward interpretation of the professional imperative is that it is imposed on society by the health care sector. This leads to over-consumption of health care (the difference between A_1 and A_0) and a welfare loss represented by the shaded area in Figure 2. This is a measure, for example, of what we as citizens have to pay for our failure to find a well-functioning incentive structure for ourselves as patients.

However, another interpretation is that we actually have somewhat peculiar preferences with respect to health and health care. Hence the normal rules for making trade-offs between different welfare-enhancing attributes do not apply. In this case, the shaded area describes in a sense what we have to pay for preferring, for example, that duty ethics govern the use and volume of health care resources. Against the latter interpretation, one must note that we often seem to make trade-offs between our health and other things which we value in life.

The professional imperative will increase health care and reduce consumption in period 1. Interestingly, it will also reduce medical research in period 1.[16] More health care in period 1 means that there will be less resources to spend on research. Furthermore, the professional imperative applies also to period 2 and this serves to make research less

16. This result is not dependent on the assumption that there are only two periods in the model with research occurring only in the first one. An extension to three periods would not change the result.

attractive from the perspective of period 1. The imperative implies that the planned marginal utility of health care spending in period 2 should be reduced, and this can be accomplished either by increasing this spending *or* precisely by reducing spending on medical research. (To increase research in this situation with the professional imperative would in a sense serve to increase the total amount of overspending on health care.) If we think of Figure 2 as representing the situation in period 2, the purpose of medical research (carried out in period 1) is to shift the BB-schedule upwards. If medical research is reduced, the BB-schedule shifts downwards instead, reducing the marginal utility of health care for all levels of spending and contributing to a fulfilment of the professional imperative. In fact, we cannot be certain that the planned health care spending in period 2 will increase as a result of the imperative, though we know that consumption is reduced in both periods and total spending on health care increases.

How costly the professional imperative is in welfare terms depends on the difference between the marginal utilities of health care and consumption and on the size of the change in resource allocation that is induced (cf. Figure 2).

There are two more aspects to be noted about the professional imperative. First, the imperative may apply also to medical research. It is not inconceivable, for example, that some sort of duty ethics also apply to the production of knowledge.[17] In such a case, health care in both periods will increase as does medical research. Hence consumption will be reduced more than in the previous case. The amount of research must increase to fulfil the requirement that its marginal utility be "small". This increase in research serves to raise the marginal utility of health care in period 2 and therefore planned health care spending in period 2 must be increased in order to bring its marginal utility down to p_i.

Secondly, I have implicitly treated the professional imperative as exogenous in the discussion above. This is not necessarily realistic – what goes on in the health care sector may be influenced by the welfare effects in other parts of the economy. A plausible alternative would be to assume that the professional imperative is endogenous, in the sense that the level of the imperative (p_i) depends positively on the marginal utility of consumption, i.e., on the opportunity cost of health care in terms of utility foregone.[18] As we increase health care spending (move to the right in Figure 2), the marginal utility of consumption increases, and this would tend to "drag" p_i along. In other words, the over-consumption of health care is counteracted since the professional imperative becomes less and less rigorous (p_i increases) as the marginal cost of this over-consumption increases. Such an effect would however not change the results qualitatively. The existence of the professional imperative will still increase health care spending in period 1 and reduce consumption in both periods as well as medical research. Not surprisingly, the marginal social loss associated with the imperative will however be smaller in this case, precisely because it becomes less stringent, and causes less of a reallocation of resources, the greater the marginal cost it imposes on society.

17. This somehow resembles the discussion on whether we should be less bothered by the loss of "statistical lives" compared to those of identifiable persons (Broome, 1982; Lyttkens, 1985, app. 3; Möller, 1986).
18. It is assumed here that the professional imperative is period specific, i.e., the level of the professional imperative in each period depends on the marginal utility of consumption in that period.

The capital-biased imperative

Consider next the capital-biased imperative. Since I have interpreted this as an excessive use of capital, it is now taken to imply inefficiency in the production of health care.[19] Since the least-cost combination of inputs is not used, less health care will be produced for any amount of resources devoted to this area.

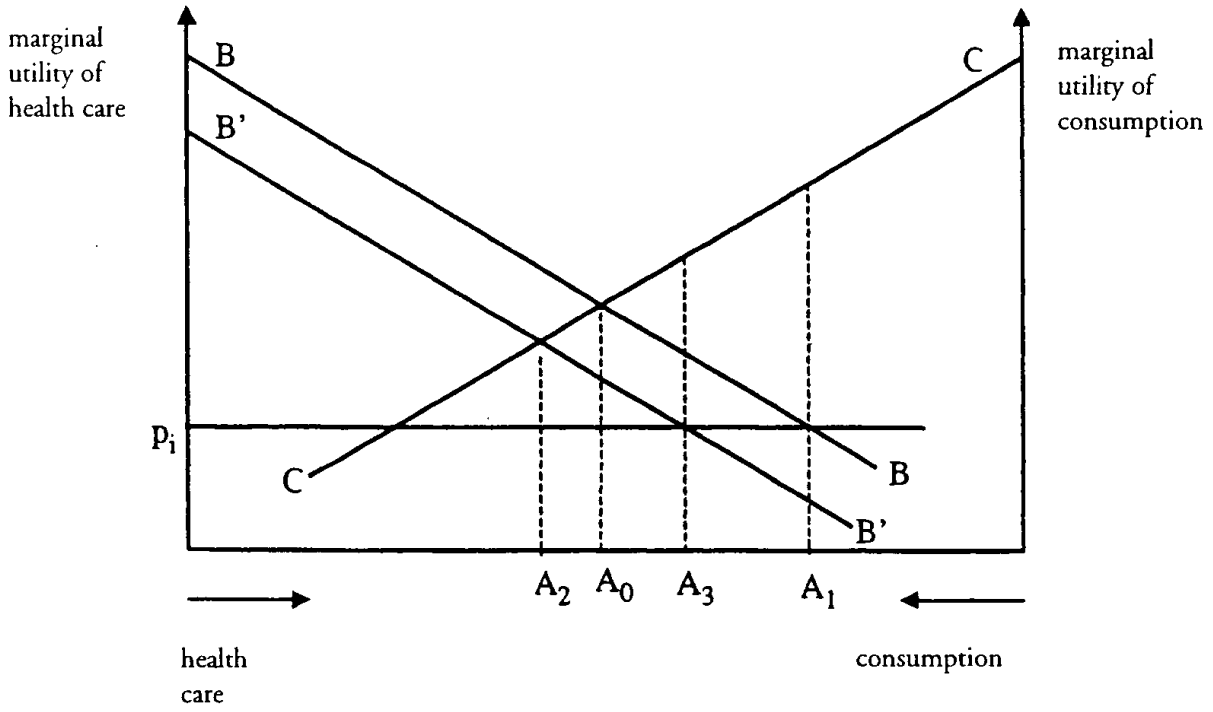
The fact that we get less health care for each dollar spent implies a reduction in the marginal utility of spending on health care. At the same time, however, this inefficiency also implies that – at each level of health care spending – health will be lower than without the imperative. This suggests an increase in the marginal utility of health and therefore of health care spending. Consequently, the introduction of the capital-biased imperative may, at a given level of health care expenditure, cause the marginal utility from such spending either to fall or to rise. The marginal utility of health care will be reduced if the marginal utility of health is relatively constant or if the inefficiency implied by excessive use of capital mainly affects the use of incremental resources in the health care area. If, on the other hand, the obtained level of health is significantly reduced by widespread inefficiency in health care spending the marginal utility of health care is likely to be increased; this also follows if the marginal utility of health changes rapidly as the level of health changes.

For the sake of simplicity, we disregard research for the moment, so that resource allocation is again only a matter of health care and consumption. Figure 3 illustrates the situation where the capital-biased imperative leads to a reduction in the marginal utility of health care for each level of health care spending, corresponding to a shift from BB to B'B'. In itself, the inefficiency in the use of resources obviously represents a welfare loss. It is important to note, however, that the type of situation into which we introduce the capital-biased imperative, namely whether or not it co-exists with the professional imperative, is of considerable importance. If only the professional imperative was operational, we would be at point A_1 (as in Figure 2) with an implied over-consumption of health care. However, the capital-biased imperative reduces the marginal utility of health care resources and this means that the professional imperative p_i will be fulfilled at a lower level of health care expenditure A_3, so that over-consumption of health care is reduced (with A_1–A_3). This represents a social gain which offsets a part of the loss implied by the inefficiency in the use of capital. It is even conceivable that we would be better off with the combination of imperatives than with only the professional imperative.[20]

On the other hand, if we start with only the capital-biased imperative, and then add the professional imperative, health care spending will increase from A_2 to A_3, which would be an unequivocal welfare loss in that situation. Finally, it could also happen, as noted above, that at the current level of health care expenditure, the introduction of the capital-biased imperative instead leads to an increase in the marginal utility of health care. This would generally seem to exacerbate the welfare loss resulting from the professional imperative, since the capital-biased imperative would then add to the over-consumption of health care.

19. This scenario need not be inconsistent with equilibrium on the markets for labour and capital but it could imply that the level of national income will be endogenous to the model. This problem is ignored here.
20. The theorem of second best (Lipsey & Lancaster, 1956) comes to mind. This implies that if there are several market imperfections, we do not necessarily improve social welfare by "correcting" one of them.

Figure 3: Resource allocation with a professional imperative and a capital-biased imperative



To the extent that the capital-biased imperative suggests that less is spent on health care, it will lead to more being spent on medical research, and vice versa.

This is a convenient place to discuss the alternative interpretation of the professional imperative, namely that it requires that those in poor health should be treated first. In contrast to the implicit assumption in the analysis above, this means that health care interventions will not be undertaken in an order solely determined by their effect on health. The effect on health will sometimes but not always be greater among those in worse health. Within a group with the same health status, however, one may still assume that health care will be allocated according to its effect on health. The general implication of allocating health care according to need would seem to be that less health will be produced for any given amount of resources devoted to health care and – somewhat loosely speaking – less health will on average be produced by each additional dollar spent on health care. Consequently, if the professional imperative works this way, it will have effects similar to the capital-biased imperative just analysed and the marginal utility of health care resources could either fall or rise.[21]

The choice in period 2

The health imperative

Period 1 has now expired, some resources

21. Another possibility is that social welfare is actually increased the most by treating those in poorest health. This implies a totally different basis for the welfare analysis. Ranking health care interventions according to the health status of those concerned is then equivalent to ranking according to the marginal contribution to social welfare. In such a case, introduction of this version of the professional imperative would not cause the BB schedule to fall; conceivably it could raise it.

have been devoted to medical research, and this has led to the development of a new medical technology. Hence the distinguishing fact about period 2 is that – in addition to resource allocation – there will be a choice between the old and the new medical technology. The discussion in this section will be focused on this choice and how it is affected by the different versions of the technological imperative. Above, we mentioned the possibility of a health imperative, i.e., that a new technology will always be chosen if its implementation in practice leads to a higher level of health, and that it will never be chosen if it leads to a lower level of health than the old technology. Since the health imperative thus concerns the choice between technologies, it was of no interest for the choice in period 1. However here it will provide the starting point for the analysis. The professional imperative and the capital-biased imperative will be considered in the following subsection.

Resource allocation in period 2 can be illustrated by Figure 4. Resources are devoted to health care and to consumption (remember – no research in period 2). The H(X)-curve represents the old medical technology, and it shows our possibilities for producing health (H) by allocating resources to health care (X). As we move from left to right in the Figure, we increase health care and consequently we reduce consumption, so the H(X)-curve implies a trade-off between health and consumption. The best choice with the old technology is assumed to be at point h_0. This is where an indifference curve (labelled I-I) is tangential to H(X).[22]

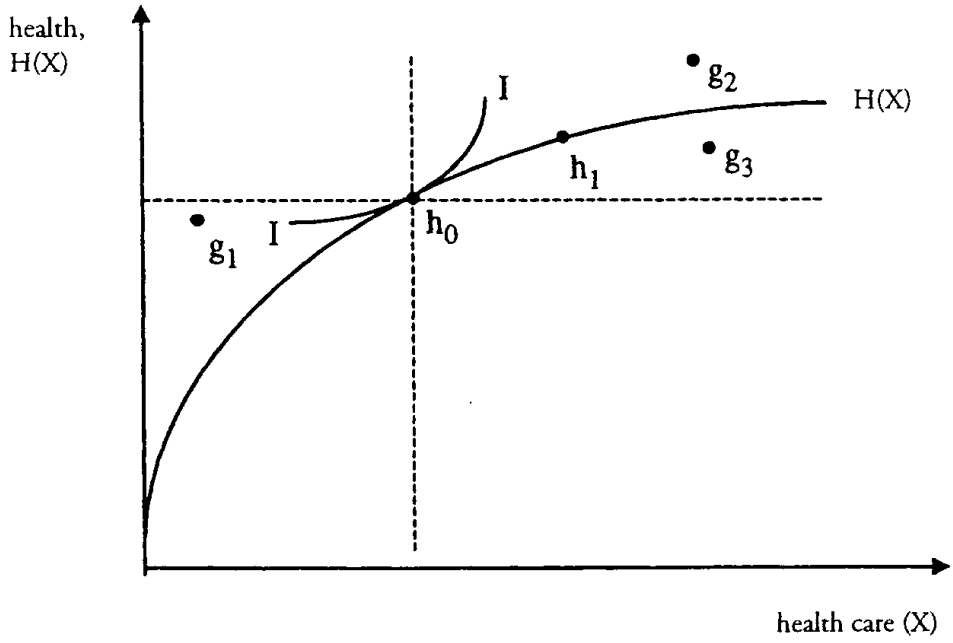
Consider now the new medical technology. Since we have no a priori information about the characteristics of this technology, the best choice with the new technology may end up anywhere in Figure 4. Assume first that the best choice with the new technology is at point g_1, i.e., to choose a level of health that is below the one that would ensue under the best choice with the old technology. Despite the health-loss, such a new technology might be deemed preferable to the old one on conventional criteria for the trade-off between health and consumption, namely if the cost-savings are large enough. This could be the case of a process innovation, i.e., an innovation that mainly serves to reduce the cost of producing existing commodities (in health care, an innovation which mostly affects the cost side usually also produces somewhat divergent effects on health). Hypothetically, we may think, for example, of a new method for women to examine their own breasts, which could be a cheap way of detecting breast cancer, but one which would not detect as many cancers as a continuation of screening with mammography.

However, a choice such as g_1 would be disqualified by the health imperative. The health imperative never allows us to choose a new technology if this implies that we will end up in a situation where we make people less healthy than under the old technology, i.e., all points below the horizontal line through h_0 are prohibited. This means that a whole class of cost-reducing innovations which would be traditionally seen as social improvements are ruled out. They will never be put to any use.

Suppose now instead that the new technology is more in the nature of a product innovation. In the health care area, this would normally mean a new treatment, a new

22. The indifference curve shows combinations of health and consumption that are considered to be equally good. Indifference curves further to the north-west in the Figure would represent higher levels of utility.

Figure 4: Resource allocation and the choice between technologies in period 2



drug, etc., which allows us to make people healthier than with the old methods. Such a new technology could of course entail both higher levels of health and a reduction in costs. However, here we focus on the other, common case where the best choice under the new technology implies both a higher level of health and a higher level of health care spending, such as g_2 in Figure 4. Despite the health gain compared to h_0, it is possible that a conventional trade-off between health and consumption would suggest that g_2 would be too costly and that it would represent a reduction in social welfare (g_2 falling below the indifference curve I-I). However, with a health imperative, we may feel obliged to choose the new technology anyway, since we cannot forego the possibility to make people healthier. This suggests that the new technology will be chosen in a number of cases where the welfare effects are dubious but health care costs will increase.

In fact, with sufficient bounded rationality – only looking at the health outcomes in practice and not considering alternative uses of a technology – it is even conceivable that the health imperative could make us choose a point such as g_3, even though something better could have been attained under the old technology.

The professional imperative and the capital-biased imperative

Suppose now that the health imperative co-exists with the professional imperative. The latter implies that our choice with the old technology is pushed towards a higher level of health care expenditure, and something like h_1 will be the point of comparison for the new technology. At h_1, we already have over-consumption of health care, and this makes the existence of the health imperative more problematical, because, as we have just seen, the health imperative raises the probability that health care costs will increase further rather than being reduced (some cost-reducing solutions are disqualified and some cost-increasing solutions are encouraged). A

specific result of this combination of imperatives is that a choice such as g_2 could represent a welfare loss, even if the new technology is clearly superior to the old one in the sense that the marginal effect of health care on health is higher for all levels of health care spending. With only the health imperative, such a new technology would necessarily be welfare improving.[23]

The capital-biased imperative, finally, in itself causes a welfare loss, but in addition it may either counteract or exacerbate the effects of the professional imperative, as in period 1. It is worth noting at this stage that the capital-biased imperative may also favour cost-increasing solutions. So far this imperative has implicitly been treated as a question of capital intensity (the capital-labour ratio). However, it is also possible that it would rather be a question of the absolute amount of capital involved in the production of health care, for example, because patients are influenced by the capital equipment that they can see and experience. A focus on the absolute amount of capital would favour technologies with a high capital intensity; it would also favour cost-increasing technologies, because the higher the costs involved in a given technical solution, the more likely that it consumes more physical capital than the alternatives, other things being equal.

Implications for medical research and the development of new technologies

There are two overall conclusions from the preceding analysis of the choice in period 2. First, cost-increasing solutions tend to be favoured by the technological imperative. Secondly, some new technologies will be adopted even though this could be said to reduce social welfare. This is a direct result of the medical research which produced the new technology. We would have been better off without this new knowledge, just as individuals who are adverse to genuine uncertainty can be better off without information that confuses rather than enlightens them (Andersson & Lyttkens, 1997).

These effects are important in themselves. In addition, however, the tendency to favour cost-increasing solutions in the diffusion process of new technologies will affect the direction of R&D efforts. Even though the results of research projects are in a sense unpredictable, the production of knowledge is not a random process. Rather it is influenced by the demand for different kinds of knowledge (Stoddart & Feeny, 1986; Weisbrod, 1991). Not only private for-profit firms, but also publicly salaried physicians, can be expected to focus on the development of technologies that have a high probability of being put to practical use. Hence we would expect the development of new medical technologies to be more focused on possibly cost-increasing product innovations than on cost-saving process innovations. It is not just that the prevailing reimbursement systems and the concomitant lack of incentives to question an increase in health care costs have profound implications for technological development in this area (Weisbrod, 1991), but also that certain cost-reducing efforts are disqualified. One may add that the rule of positive

23. If the new technology entails a higher marginal effect of health care on health, then the professional imperative implies that health care spending must increase from h_1. The marginal utility of health care is the product of the marginal utility of health and the marginal effect of health care on health. If the latter is increased, health care spending must increase to ensure that the marginal utility of health care falls to the level required by the professional imperative. In contrast, with such a favourable new technology and only the health imperative, health care expenditure would be reduced from h_0.

beneficence suggests a focus on product innovations, and that physicians are trained to concentrate on ways of helping the individual patient rather than to consider the anonymous welfare effects implied by cost reductions.

The influence that the technological imperative in health care will have on the direction of R&D efforts is arguably its most important effect. In addition, however it will tend to reduce the amount of research which is being undertaken in our simple framework. In period 2, the technological imperative will sometimes lead to the adoption of a new technology even though it is too costly (and some cost-reducing innovations will be disqualified). Seen from the *ex ante* perspective of period 1, awareness of this effect will reduce the expected value of the research being undertaken, with a reduced volume of research as a further consequence.

We noted above that the professional imperative may reflect a set of somewhat peculiar preferences with respect to health and health care, rather than being imposed on society by the health care sector. This may be true also of the health imperative. If this is the correct interpretation, an interesting dilemma is suggested: we may be perfectly happy with a technological imperative that governs the choice between different medical technologies, and at the same time be unhappy with the effect that this has on the focus of medical research (which somewhat resembles the familiar equity-efficiency trade-off). These issues may have a particular policy relevance in a country like Sweden, where a significant part of public resources for research are allocated on the presumption that it is possible and desirable to influence the direction of research.[24]

Concluding remarks

We may conclude that there are several possible interpretations of the technological imperative in health care, and that the professional imperative, the capital-biased imperative and the health imperative can be important both in a static and in a dynamic perspective. These different imperatives may influence the allocation of resources between health care, consumption and research, as well as the choice between medical technologies. This has a resulting effect on the level of social welfare, and it may also affect the magnitude and direction of R&D efforts. The effects of one of the imperatives can be reinforced by the existence of another one, but sometimes they also counteract each other. The presence of such imperatives can help explain, for example, why the technological development is often seen as a major factor behind the upward pressure on health care costs.

It is important to note that we have analysed the effects of the technological imperative in a situation where the aim is to maximise a simple measure of aggregate social welfare. It is what a single decision-maker would do if endowed with the specified social welfare function and ignoring distributional issues. Obviously, however, resource allocation in practice is the result of a complex interplay between decisions taken by many different actors at different levels. Their decisions will be influenced by the incentives they face and these in turn are determined, for example, by the organisation and financing of the health care sector.

In general, one would expect many of the effects of the technological imperative to be present also in a more complex setting. Even

24. Calltorp (1986) found a high degree of agreement within the health care sector (politicians, administrators, scientists) in setting priorities for research. It would come as no surprise, however, if this harmony has no counterpart in the relationship between the health care sector and the rest of society.

when it is not discussed explicitly, there is often an obvious trade-off between consumption and health care, irrespective of whether an individual citizen is allocating his own budget or whether this is partly done on his behalf by politicians, employers, etc. In many settings, the professional imperative is then likely to produce an overspending on health care, at least if it is operative at the level of the physician-patient relationship, which arguably is where the most influential decisions are taken on the use of health care resources. Similarly, the capital-biased imperative will reduce the efficiency in the production of health care, and the two imperatives may interact if these effects are discernible at level where the professional imperative is operative.

The tendency of the health imperative to favour cost-increasing solutions is also likely to feed back into the direction of research efforts. With respect to the total amount of spending on medical research the situation with respect to the imperatives is however clearly more complicated. The decision to spend on research is often taken by a private actor and it is at least one step removed from the trade-off between consumption and health care. Instead it is governed by the private incentive for a company, for an individual physician, etc., to engage in research. This means, for example, that it is no longer necessarily true that a professional imperative (which applies only to health care) reduces medical research; even if the professional imperative reduces the "social" marginal value of spending on research, it can increase the private profitability of doing so by expanding the potential market.[25]

A closely related point is that the strength of the technological imperative and its effects on resource allocation is likely to vary across health care systems, as already implied by our discussion above of the rational for the different imperatives. The technological imperative can be seen as an explicit or implicit decision rule, as an attitude, which partly embodies and is affected by factors such as the financial incentives for consumption and production of health care, the institutional setting, and the belief systems of individual actors. All of these vary between countries, and the relative prominence of the technological imperative is one aspect of the well-known international differences in, for example, use of technology and aggregate health care costs.

The structure of health care systems is however also subject to change, and during the last decade *organisational* innovations have been an important factor in the health care sector in many countries. As mentioned above, we have witnessed the introduction and expansion of quasi-markets, prospective reimbursement schemes, HMO-type arrangements, and other attempts to change the financial incentives of health care providers and to increase their degree of cost-consciousness.

The motivation behind these institutional changes is a fascinating research subject in itself (Lyttkens & Borgquist, 1995), but here I am concerned with the fact that the effects of these reforms are contingent upon several factors. First, we know that markets and competition do not necessarily reduce costs in the health care area, although they may do so (Robinson & Luft, 1988; Pope, 1989;

25. Then again, such an effect could possibly be countered by an appropriate reduction of publicly funded research. Hence the effect of the technological imperative on research could partly depend on the relative importance of public vs private research.

Melnick et al., 1992). The effect of introducing HMO-type arrangements in less resource-intensive health care systems than in the US largely remains to be seen.[26] Cost-reducing measures create well-defined pressure groups of discontented patients and health care personnel while the winning taxpayers are difficult to organise (Olson, 1965).

Secondly, the effects of the reforms will depend on the strength of the technological imperative. There is a saying that old habits die hard, but informal rules of behaviour may die even harder (North, 1990). Politicians, physicians and patients often seem unanimous in advocating generosity with health care resources, frequently on moral grounds. On the one hand, their arguments carry considerable weight on the basis that these groups are particularly well informed about matters of health care. On the other hand, however, the independent observer may note that the ethical standpoint that a person advocates often seems to coincide with his self-interest (Elster, 1996). This leads to a problem of interpretation which is familiar to historians well-versed in a critical attitude towards their sources.

The tendency for health care costs to take a higher and higher share of GDP appears largely to have been curbed in many countries (with the US as a notable exception). This, however, still leaves room for a considerable expansion in terms of the real cost of health care. Nor should this relative stagnation in health care costs necessarily be attributed to the organisational changes. For the future development of health care costs, it is however important that the efforts made to change the financial incentives for actors in the health care sector largely seem to be an international phenomenon. Information is a public good in the sense of non-rivalry in consumption. Hence the new technologies available in one country will always largely be a function of the research incentives in other countries. With a technological imperative in health care, it is an obvious possibility that cost-increasing new technologies will continue to be adopted as long as they are produced somewhere.

References

Andersson, F. & C.H. Lyttkens, 1997. "Health, Genuine Uncertainty and Information," paper presented at the European Economic Association Congress, Toulouse, August 1997.

Broome, J. 1982. Uncertainty in Welfare Economics and the Value of Life. In Jones-Lee, M. W. (ed.), *The Value of Life and Safety*, North-Holland, Amsterdam, 201-216.

Broomé, P., Lindgren, B., Lyttkens, C. H., & R. Ohlsson, 1994. Health Care and the Elderly. In Bengtsson, T. (ed.), *Population, Economy and Welfare in Sweden*, Springer Verlag, Heidelberg, 155-186.

Cairns, J. A. 1992. "Health, Wealth and Time Preference," *Project Appraisal*, 7:31-40.

Calltorp, J. 1986. Assessment of the Potential Benefits of Research – A Study Among Politicians, Administrators and Researchers in Sweden. In May, A., Smedby, B. & L. Wetterberg (eds.), *Perceptions of the Values and Benefits of Research, Acta Psychiatrica Scandinavia* 74 Supplementum, 137-148.

Cromwell, J. & J. B. Mitchell, 1986. "Physician-induced Demand for Surgery," *Journal of Health Economics*, 5:293-313.

Culyer, A. J. & A. Wagstaff, 1993. "Equity and Equality in Health and Health Care," *Journal of Health Economics*, 12:431-57.

Cutler, D. M. & M. McClellan, 1996. The Determinants of Technological Change in Heart Attack Treatment. *NBER Working Paper* 5751.

Dozet, A., Höjgård, S., Lindgren, A., Öhlin, H. & C. H. Lyttkens (forthcoming). Technological Change in the

26. It is of limited relevance for countries like, for example, Denmark or Sweden to know that HMOs run up lower costs than the US fee-for-service system (Pedersen, 1991).

Treatment of Acute Myocardial Infarction in Sweden. In McClellan, M. (ed.), *Technological Change in Health Care. A Global Analysis of Heart Attacks.*

Dozet, A., Nystedt, P. & C. H. Lyttkens, 1997, "Health Care for the Elderly – Two Cases of Technology Diffusion," paper presented at the 18th Nordic Health Economists' Study Group Meeting, Helsinki, August 1997.

Elster, J. 1996. "Transmutation and Misrepresentation," *Nordic Journal of Political Economy*, 23:3-23.

Escarce, J. J. 1996. "Exernalities in Hospitals and Physician Adoption of New Surgical Technology: An Exploratory Analysis," *Journal of Health Economics*, 15:715-34.

Fuchs, V. R. 1968. "The Growing Demand for Medical Care," *The New England Journal of Medicine*, 279:190-195.

Grytten, J., Carlsen, F. & R. Sørensen, 1995. "Supplier Inducement in a Public Health Care System," *Journal of Health Economics*, 14:207-229.

Harris, J. E. 1977. "The Internal Organisation of Hospitals: Some Economic Implications," *The Bell Journal of Economics*, 8:467-482.

Hernes, G. 1975. *Makt og avmakt*, Universitetsforlaget, Bergen.

Jensen, G. A. & M. A. Morrisey, 1986. "The Role of Physicians in Hospital Production," *The Review of Economics and Statistics*, 68:432-442.

Lipsey, R. G. & K. Lancaster, 1956. "The General Theory of Second Best," *Review of Economic Studies*, 24:11-32.

Lyttkens, C. H. 1985. *Swedish Work Environment Policy. An Economic Analysis.* Lund Economic Studies no. 37, Studentlitteratur, Lund (Ph.D. thesis).

Lyttkens, C. H. & L. Borgquist, 1995. Swedish Health Care in the Late 1980s and Early 1990s – A Story of Institutional Changes. In Alban, A. & T. Christiansen (eds.), *The Nordic Lights. New Initiatives in Health Care Systems*, Odense University Press, 306-327.

Maynard, A. 1983. "The Production of Health and Health Care," *Journal of Economic Studies*, 10:31-45.

McGuire, A., Parkin, D., Hughes, D. & K. Gerard, 1993. "Econometric Analyses of National Health Expenditures: Can positive economics help to answer normative questions?" *Health Economics*, 2:113-126.

Melnick, G. A., Zwanziger, J., Bamezai, A. & R. Pattison, 1992. "The Effects of Market Structure and Bargaining Position on Hospital Prices," *Journal of Health Economics*, 11:217-233.

Mooney, G., Hall, J., Donaldson, C. & K. Gerard, 1991. "Utilisation as a Measure of Equity: Weighing Heat?" *Journal of Health Economics*, 10:476- 480.

Mooney, G. & M. Ryan, 1993. "Agency in Health Care: Getting beyond first principles," *Journal of Health Economics*, 12:125-135.

Möller, G. 1986. *Risker och människolivets värde. En etisk analys*, Uppsala studies in social ethics 8, Almqvist & Wiksell International, Stockholm (Ph.D. thesis).

North, D. C. 1990. *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, New York.

Olson, M. 1965. *The Logic of Collective Action.* Cambridge, Mass.

Pedersen, K. M. 1991. Health Maintenance Organizations. For amerikansk, eller kan danskerne lære noget? In Andersen, P. & T. Christiansen (eds.), *Styring & regulering på udbudssiden i sundhedssektoren*, Odense universitetsforlag, 114-158.

Pope, G. C. 1989. "Hospital Nonprice Competition and Medicare Reimbursement Policy," *Journal of Health Economics*, 8:147-172.

Robinson, J. C. & H. S. Luft, 1988. "Competition, Regulation, and Hospital Costs, 1982 to 1986," *The Journal of the American Medical Association*, 260:2676-2681.

Rochaix, L. 1993. "Financial Incentives for Physicians: The Quebec experience," *Health Economics*, 2:163-176.

Simon, H. A. 1979. "Rational Decision Making in Business Organizations," *The American Economic Review*, 69:493-513.

Stoddart, G. & D. Feeny 1986. Policy Options for Health Care Technology. In Feeny, D., Guyatt, G. & P. Tugwell (eds.), *Health Care Technology: Effectiveness, Efficiency, and Public Policy*, The Institute for Research on Public Policy, Montreal.

Søgaard, J. 1997. "Production of Life-Time. Short Term Models with Lags for Ten OECD Countries 1970 to 1989," paper presented at the Nordic Health Economists' Study Group Meeting, Helsinki, August 1997.

Vårdens svåra val, Slutbetänkande av Prioriteringsutredningen, SOU 1995:5, Socialdepartementet, Stockholm.

Weisbrod, B. A. 1978. Comment on M. V. Pauly. In Greenberg, W. (ed.), *Competition in the Health Care Sector*, Proceedings of a conference sponsored by Bureau of Economics, Federal Trade Commission, Aspen Systems, Germanstown.

Weisbrod, B. A. 1991. "The Health Care Quadrilemma: An essay on technological change, insurance, quality of care, and cost containment," *Journal of Economic Literature*, 29:523-552.

Appendix

This appendix presents the formal model underlying the discussion in the main text, and an outline of the analysis of the effects of the professional imperative in period 1.

Time is discrete and divided into two periods. Subscripts are used to denote time periods. Social welfare in period i (W_i) is assumed to be a function of consumption (C_i) and health (H_i) in that period:

$$W_i = W_i(C_i, H_i), \qquad i = 1,2.$$

Social welfare over the two periods is assumed to be the sum of the period-specific utilities (the level of utility in one period does not affect utility in the other period). All functions are assumed to be continuous and twice differentiable. Welfare is assumed to be concave in both consumption and health, and the marginal utility of consumption is assumed to be independent of the health state (zero cross-partial derivative).[27] Health can be increased by allocating resources to health care (X_i). In addition, in period 1 it is possible to allocate resources to medical research (R), and research is assumed to change the relationship between health care and health in period 2 in a positive direction. Hence we have

$$H_1 = H_1(X_1), \text{ and } H_2 = H_2(X_2, R),$$

where it is assumed that health is concave in both health care and research, and that there are decreasing returns to scale in the production of health.[28] Furthermore, it is assumed that research has a positive effect on the marginal effect of health care on health, and that it has a positive impact on the marginal utility of health care,[29] i.e., that

$$\partial^2 H_2/\partial X_2 \partial R \geq 0,$$

$$\partial/\partial R[(\partial W_2/\partial H_2)(\partial H_2/\partial X_2)] \geq 0.$$

National income (Y) is exogenously given. With resources freely transferable between periods and no rate of interest, there is a joint resource constraint for the two periods:

$$Y = C_1+X_1+R+C_2+X_2.$$

Assuming that we have an interior solution, the first-order conditions (FOC) of the maximisation problem presented so far show that the marginal utility will be the same for all five different uses of resources.

The professional imperative is then introduced as constraints on the marginal utility of health care:

$$\partial W_i/\partial X_i \leq \mu, \qquad i = 1,2.^{30}$$

Here, μ is assumed to be a "small" number, and the same μ is assumed to apply to both periods. To be meaningful, the professional imperative must imply that the marginal utility of health care resources becomes smaller than it would otherwise have been. Hence the two constraints are both assumed to be binding. Assuming again that we have an interior solution, the Kuhn-Tucker conditions are necessary for a maximum.[31] By comparative-static analysis on the corresponding equation system, it can be shown that a marginal reduction in μ have the effects reported in the text, i.e., that $dC_1/d\mu$,

27. While this assumption is not necessarily realistic, it can be motivated by the fact that it seems equally possible to argue that the marginal utility of consumption is higher for a healthy individual as it is to argue that the opposite holds.
28. By means of Euler's theorem, the latter assumption implies that $(\partial^2 H/\partial X^2)(\partial^2 H/\partial R^2) - (\partial^2 H/\partial X \partial R)^2 > 0$.
29. Cf. the discussion in note 10.
30. Note: $\partial W_i/\partial X_i$ is shorthand notation for $[(\partial W_i/\partial H_i)(\partial H_i/\partial X_i)]$. Similar notation is used below.
31. Inspections of the bordered Hessian determinants show that the second-order sufficient condition is fulfilled.

$dC_2/d\mu$, and $dR/d\mu$ are positive, $dX_1/d\mu$ is negative, but $dX_2/d\mu$ is indeterminate in sign.[32] Similarly, inspection of the FOC shows that we will now have

$$\partial W_2/\partial R > \partial W_1/\partial C_1 =$$
$$\partial W_2/\partial C_2 > \partial W_1/\partial X_1 = \partial W_2/\partial X_2$$

which yields qualitatively the same conclusions for the introduction of the imperative as those just reported for marginal changes in μ. The value of the relevant Lagrange multiplier is a measure of how the objective function reacts to a slight relaxation of the constraint, i.e., how costly the constraint is at the margin. At the optimum, the values of the multipliers associated with the professional imperative-constraints (y_i) are given by

$$y_i = [\partial W_i/\partial X_i - \partial W_i/\partial C_i]$$
$$[1/(\partial^2 W_i/(\partial X_i^2)] > 0, \qquad i = 1,2,$$

and comparative-static analysis shows that $dy_i/d\mu < 0$.

If the professional imperative applies not only to health care but also to medical research, we get an additional constraint in the form

$$\partial W_2/\partial R \leq \mu.$$

The FOC now show that we have

$$\partial W_1/\partial C_1 = \partial W_2/\partial C_2 > \partial W_1/\partial X_1 =$$
$$\partial W_2/\partial X_2 = \partial W_2/\partial R,$$

which suggests that consumption has been reduced in both periods whereas spending on all other items have increased (cf. the main text). The marginal social loss associated with the professional imperative is now larger for the constraint with respect to the marginal utility of X_2. We have

$$y_2 = [\partial W_2/\partial X_2 - \partial W_2/\partial C_2 -$$
$$y_R(\partial^2 W_2/\partial R\partial X_2)]\ [1/(\partial^2 W_2/\partial X_2^2)] > 0,$$

where y_R is the multiplier with respect to the constraint on the marginal utility of R. Compared to the previous case, the value of the Lagrange multiplier is increased by the presence of an interaction term in the numerator. As explained in the main text, X_2 has to increase when μ is reduced, and this will increase the marginal utility of research, thus necessitating additional spending also on R.

Finally, consider briefly the case where the professional imperative is endogenous, and assume for simplicity that it applies only to health care.[33] As argued in the main text, a plausible alternative is to let the professional imperative (μ) depend positively on the marginal utility of consumption and therefore on the level of consumption, so that

$$\partial \mu_i/\partial C_i < 0, \qquad i = 1,2.$$

In this context, it makes more sense to make μ period specific. Inspection of the FOC suggests that we derive qualitatively the same effects on resource allocation as with an exogenous imperative. The relationship between C and μ serves to reduce the marginal effect of the constraint on the objective function, because a marginal change in μ is now associated with a smaller reallocation of resources. We have

$$y_i = [\partial W_i/\partial X_i - \partial W_i/\partial C_i][1/(\partial^2 W_i/\partial X_i^2) +$$
$$(\partial \mu_i/\partial C_i)] > 0, \qquad i = 1,2,$$

where $\partial \mu_i/\partial C_i$ serves to increase the absolute value of the denominator.

32. All third derivatives are ignored.
33. It is trivial to extend the analysis to the case where an endogenous professional imperative applies to both health care and medical research.