Journal of Health Economics 12 (1993) 125-135. North-Holland

Agency in health care: getting beyond first principles

Gavin Mooney and Mandy Ryan*

Health Economics Research Unit, University of Aberdeen, Aberdeen, UK

This paper is concerned with the application of the theory of agency to health care. It is argued that the basic theory of agency raises more questions than it provides answers when it is applied to the doctor–patient relationship. More research is needed into the nature of both the patient's and the doctor's utility functions. Only then can we begin to devise optimal incentive structures to encourage doctors to take adequate account of patient preferences.

1. Introduction

Many economists have argued that the assumption of consumer sovereignty does not apply in the market for health care because of the asymmetry of information between the patient and the doctor. The doctor–patient relationship has thus frequently been modelled within the theory of agency, with consideration of the associated incentive problems.

While the concept of the agency relationship has become widely accepted in health economics, there appears to be some lack of agreement about its precise role, how it operates and how it might be made to operate better. This may reflect the fact that, whilst economists have analysed the standard theory of agency in great detail (as is evidenced by the considerable literature in this area), agency in health care has to date been used primarily as no more than a description of behaviour. That is, there is comparatively little analytical work on it in health care.

Whilst the standard theory of agency, as it is applied to any market characterised by an asymmetry of information [MacDonald (1984)], assumes that the principal's and agent's utility functions are independent, in health

Correspondence to: M. Ryan, Health Economics Research Unit, University of Aberdeen, Department of Public Health, University Medical Buildings, Foresterhill, Aberdeen AB9 2ZD, Scotland, UK.

*We are grateful to an anonymous reviewer for comments on an earlier version of this paper. Early work on this was funded as part of a COMAC (HSR) project. The Unit is funded by the Chief Scientist Office of the Scottish Home and Health Department; however, the opinions expressed in this communication are those of the authors, not SOHHD.

care it has been recognised that the doctor's and patient's utility functions are, to a certain extent, interdependent [Evans (1984)].

Outside health care the standard theory of agency has concentrated on devising optimal fee schedules, with complicated fee structures that are related to the outcome being predicted. However, such complicated fee schedules have not emerged in health care, and current methods of paying doctors appear highly simplistic when compared to the predictions of agency theory. Research in the area remains limited (see [Donaldson and Gerard (1989), Krasnik et al. (1990)] and it is not surprising that health economists do not agree on optimal methods of remuneration for doctors.

The purpose of this paper is to examine the principal–agent relationship in mainstream micro-economics with a view to identifying more precisely its relevance and transferability to health care. It is argued that, given the interdependence of the principal and agent's utility functions in health care, transferability is dependent on identifying the nature of these utility functions. Leading on from this, the paper highlights the need to see remuneration systems as a mechanism for influencing doctors to identify more accurately what is in the patient's utility function and to act on that information.

2. The theory of principal and agent

2.1. The standard theory of agency

The agency relationship is characterised by a principal (ill-informed individual) and agent (informed individual), both of whom are attempting to maximise their independent utility functions. Because of this diversity of interests, and the fact that there is an asymmetry of information, the principal has to attempt to devise a contract to ensure that the agent does not cheat. In the absence of such a contract:

> 'the agent's activities are known only to himself and there is no immediate incentive for him to reveal them to the principal' [MacDonald (1984)].

The principal has to motivate the agent to choose his/her activities in a way that benefits the principal. This is done through the devising of *compensation rules* or *incentive compatibility constraints* [MacDonald (1984), Arrow (1986)]. Given that the agent has the choice of whether or not to accept the contract, the principal has to ensure that the contract is attractive to the agent. This is known as the viability or participation constraint [Arrow (1986)].

The principal–agent literature suggests that, on the assumption that the outcome is the only source of information to the principal, the agent's payment will be a function of this outcome. The optimal fee schedule will, however, depend in part on whether the agent is risk neutral or risk averse.

If the agent were risk neutral, the resulting fee schedule would be simply the outcome minus the principal's share (which would be some constant):

> 'if the agent were risk-neutral, the principal–agent problem would have a trivial solution: the agent would bear all the risks, and then the differential information would not matter. That is, the principal would retain a fixed amount and pay all the remainder to the agent, who therefore would have no dilution of incentives' [Arrow (1986)].

Arrow gives the example of a landlord (principal) renting land to a farmer (agent). The landlord would charge a fixed rent independent of the output (this fixed amount would depend on the participation constraint). The agent would then take on board all the risk.

However, risk aversion changes the position:

> 'this solution ceases to be optimal as soon as the agent is risk averse. Since all individuals are averse to sufficiently large risks, the simple solution of preserving incentives by assigning all risks to the agents fails' [Arrow (1986)].

The optimal fee schedule with a risk averse principal and agent will be a function of the outcome, where both actors share the risk. Under such a situation, Shavell (1979) has shown that there is a trade-off between incentives and efficiency with a second best solution being achieved.

2.2. Extensions to the standard theory of agency

Much of the literature on the agency relationship has concentrated on the basic one period non-cooperative situation (as presented above). Extensions to this basic theory have been concerned with devising optimal reward systems.

2.2.1. Monitoring and optimal compensation rules

The basic theory of agency assumes that information about the agent's action is available only to the agent. However, the principal might seek to acquire information about the agent's actions and incorporate this into the contract i.e. the principal might monitor the agent's actions. If monitoring reveals extra information about the action of the agents, then the fee schedule should depend on this information [Holmstrom (1970)].

However, perfect monitoring is usually difficult and/or costly, and 'imperfect estimators of actions' [Holmstrom (1979)] are employed. Examples include rewarding a salesperson by reference to the number of miles they travel as well as their sales, and paying a professional by considering the number of hours they work as well as having a set fee.

2.2.2. Multiple agents and optimal compensation rules

Holmstrom (1982) considered optimal incentives with multiple agents. He was concerned with how relative performance indicators could be used as a

system of rewards to promote efficiency (i.e. salesperson of the month awards where relative performance is measured by rank alone). He shows that such systems will be efficient if, and only if, agents' outcomes are dependent. In the extreme case, if agents' outputs are completely dependent, in that the output of individual *i* reveals all information about the state of nature facing individual *j*, the first best solution will be achieved by using relative performance evaluations. However, if outputs are independent, the optimal fee schedule is one where individual *i*'s payment depends on his/her output alone.

In the above it was assumed that individual output could be observed. Holmstrom shows that if only total output can be observed, and output is shared equally among the agents, the result will be inefficient because free riding will occur. Alchian and Demsetz (1972) suggest that under such a situation efficiency will be improved if a principal is brought in to monitor the actions of agents. However, Holmstrom argues that the role of the principal is not one of monitoring, but rather to impose group incentives i.e. penalty or bonus schemes.

He goes on to argue that in firms where ownership is separate from the means of production, the free-rider problem will be less than in partnerships where output is equally distributed. Fama (1980) also showed how the 'separation of security ownership and control, typical of large corporations, can be a more efficient form of economic organisation.'

2.2.3. Dynamic models and optimal compensation rules

When we consider the possibility of repeated contracts between a principal and an agent, other methods for reducing inefficiency arise. Rubenstein and Yaari (1983) argue that, in these circumstances, the role of rewards and penalties in devising the optimal contract is reduced. They attempt to explain the fact that insurance companies offer discounts to clients who possess favourable past claims. Such insurance contracts offer full payment at a prespecified price, for loss incurred. It is this prespecified price that can be varied in different periods and which becomes a function of the insured person's past record of claims. They show that:

> 'such discounts provide a mechanism which enables both insurer and insured to counteract the inefficiency which arises from moral hazard.'

They go on to argue that dynamic agency relationships lead to more efficient outcomes than are achieved under the single period models. The crucial point in achieving this is that the players interact more than once.

3. The theory of agency in health care

3.1. Independent utility functions?

As we have seen, the standard theory of agency assumes that principals

and agents have independent utility functions, and has concentrated on identifying optimal reward contracts to attain efficient outcomes.

In the market for health care it is generally recognised that the doctor (agent) acts, to some extent, in the patient's interests.

> 'What distinguishes the professional agency relationship is that the professional includes part at least of the patient's/client's interests in her own objectives' [Evans (1984)].

Given this assumption, an efficient outcome can only be determined if we know the nature of the doctor's and the patient's utility functions. Only then can incentive systems be devised to try to ensure that the doctor acts in the *patient's* best interests.

There has been a paucity of research by economists in this area. Williams (1988) postulates a 'perfect agency relationship' as one where the doctor gives the patient all the information the patient needs and the patient then makes the decision:

> 'the 'consumers' rely on doctors to act as their agents, in a system which obstensibly works on the principle that the doctor's role is to give the patient all the information the patient needs in order to enable the patient to make a decision, and the doctor should then implement the decision once the patient has made it.'

However, he continues with what appears to be a criticism of the way in which doctors actually do behave:

> 'I am sure that the reader would find the above statement closer to his or her own experience if the postulated roles of the patient and doctor were interchanged, so that the sentence would then read 'the patient's role is to give the doctor all the information the doctor needs in order to enable the doctor to make a decision and the patient should then implement that decision once the doctor has made it'.'

Williams, in his view of the 'perfect agency relationship', appears to assume that the patient's utility can be maximised by the doctor's provision of information to the patient. This view is at some distance from the conventional role of the agent in standard economic theory, where the agent is the decision maker, and does not normally provide information to the principal unless this is specified as part of the contract:

> 'the agent's activities are known only to himself and there is no immediate incentive for him truthfully to reveal them to the principal' (MacDonald, 1984).

Williams' assumption that the patient's utility can be maximised by the provision of information begs the question: information on what? Whilst Williams is not explicit about this, his development of league tables of marginal cost per QALY as a basis for priority setting in health care implies that the only relevant outcome in health care systems is health; that all that is in the patient's utility function is health; and that potentially or eventually

QALYs can be an adequate measure of health [Williams (1985)]. Combined with the above statement on agency, the role of the doctor/agent would then appear to be reduced to giving the patient/principal full information on health outcomes so that the patient can make the rational, informed choices which are to be seen in relation solely to health.

Williams recognises that the agency relationship in practice will not be perfect since:

> 'the doctors ··· may be pursuing interests other than those of the patient in front of them.'

This is an example of 'incentive compatibility constraints', where the fact that the doctor has other competing arguments in her utility function may lead to imperfections in the agency relationship. However, Williams does not deal further with these other interests.

This (Williams') view of the agency relationship in health care differs from the standard theory in that patients take on the decision making role. Culyer (1989) described the agency relationship as involving:

> 'medical doctors, who act as agents on his or her [the patient's] behalf: ideally choosing in the way the individual would, had he or she been possessed of the same informational advantages of the professional.'

This stance is nearer to that presented in standard economic textbooks, since it sees the doctor/agent as the decision maker. The only apparent difference as the basis of choice between the patient, left to his/her own devices, and what the doctor might choose on behalf of the patient, is information. Again, this begs the question, information on what? Other work by Culyer (1988, 1989) suggests that, like Williams, he also sees health as the only relevant argument in the patient's utility function.

Evans (1984), like Culyer, sees the agent as making decisions on behalf of patients. However, he recognises that the 'complete agency relationship' will be that which maximises the utility of patients. He argues that the following factors will lead to, what he calls, 'an incomplete agency relationship':

(i) the doctor may respond incompletely or with misperception of the patient's interests;
(ii) lack of perfect information by the doctor on effectiveness;
(iii) a tendency for the doctor to overstate effectiveness, essentially the ethic of virtue, the desire to do good.

Thus, in Evans' view, the incompleteness of the agency relationship is more concerned with patient utility maximisation than with either maximisation of information or maximisation of health. Whilst Evans does not state explicitly what contributes towards patient utility, he has moved away from the monopoly of health in the patient's utility function. What is particularly relevant here is a quote from Evans and Wolfson (1980) on the question of the relevant arguments in the utility function and the specification of these:

'Preferences are preferences – economic theory is not supposed to pass moral judgements about what should be in a utility function.'

This view seems more in line with the standard theory of agency where it is the principal who sets the objective. However, only recently have economists begun to ask what, beyond *health*, might be important to patients/principals in the provision of health care. Some are now arguing that patients are concerned not just with the *outcome* of treatment, but also with the *process* [McGuire et al. (1988), Mooney and Lange (1991), Mooney (1991)]. For example, Mooney and Lange (1991), in evaluating pre-natal screening, address the question of what enters the patient's utility function. They argue that:

'many of the existing studies in this field make the implicit and sometimes explicit assumption that the only form of benefit available from such screening arises as a result of the opportunity to abort an affected foetus.'

Thus, the assumption of such studies is that the only relevant argument in the pregnant woman's utility function is outcome in terms of whether the foetus is aborted. However, the observation that some women choose to be screened but do not abort an abnormal foetus [Lubs and Falk (1977), Macnicol et al. (1986)] suggests that there are other relevant arguments in the pregnant woman's utility function. They suggest that one of these is information per se, and that women get utility from 'reduced uncertainty or perhaps just from knowledge'.

3.2. Incentive compatibility constraints in health care

The literature on the economic theory of agency predicts complicated fee structures that are related to outcome [Arrow (1986)]. However, such fee schedules have not emerged in health care, and current methods of paying doctors appear highly simplistic when compared with the predictions of agency theory.

A number of points are worth noting here. In the standard theory of agency the need for incentive compatibility constraints arises out of the separate utility functions of the principal and agent. It may be that doctors, in their professional role, attempt to act in the patient's interests (or that those devising contracts assume this to be the case). Given this, there would be no need to introduce complicated fee structures. It becomes crucial to establish the nature of the patient's and the doctor's utility functions. Having done so, we can then explore optimal reward schedules to establish efficiency in the doctor–patient relationship.

The standard theory of agency assumes that such schedules are set by principals and will be, at least in part, a function of outcome. In health care, methods of remuneration are usually set by some third party (government or insurance companies). The question then becomes: how can we get this third

party to define optimal methods of remuneration? The question is complicated by the fact that, in the market for health care, there is a problem in defining and measuring health (or whatever the relevant outcomes of health care are). And, even if the outcomes were observable, there is the additional problem of establishing the input of the agent in achieving these outcomes.

There is also the issue of *non-monetary* constraints. Economists have been generally agreed that the asymmetry of information in health care creates an incentive problem, resulting in more (or less) treatment being 'demanded' than would have been the case if the patient had full information and knowledge. This results in the possibility for supplier-induced demand [Hay and Leahy (1982), Rice (1984)].

Whilst different studies have come up with different conclusions about the level of demand inducement[1], it is generally recognised that doctors' behaviour is constrained in some way. That is, the doctor does not choose the level of treatment for a patient that would be best for the doctor. One explanation for this is that doctors' behaviour is constrained by ethical considerations [Evans (1974, 1984)]. As Arrow (1986) states with reference to the health care market:

> 'the usual reasons why the market acts as a check to ensure quality operate with very weak force here. It is for this reason that the ethical indoctrination is of such crucial importance. The control that is exercised ordinarily by informed buyers is replaced by internalised values'.

Coyte (1984) argued that the market for health care is like any other market characterised by an asymmetry of information, and that patients can monitor doctors' behaviour in much the same way as they can monitor any agent's behaviour.

Thus Coyte is assuming that patients are *aware* of the need, *willing* and *able* to monitor the behaviour of doctors. But are patients aware of the need to monitor doctors? What do patients perceive to be in their doctor's utility function and are they aware that doctors might act in a way that is not in their interests?

Even if patients recognise the need for monitoring, will they be able to do so? The ability of the principal to monitor the agent's behaviour is limited in the market for health care. The patient might feel that his/her health status has improved after treatment, but often does not know what would have happened with no treatment. Further, given that the principal does not normally set the contract, the assumption that patients can monitor the

[1]Whilst economists have normally defined SID as medical services the doctor orders which the patient would refuse if they were 'fully informed and knowledgeable', there has been a lack of agreement in the literature as to how best to test for SID. Given that the definition above is impossible to test for, numerous indirect methods have been adopted, the evidence from which is inconclusive. In another paper (Ryan, 1992a) alternative definitions of SID are suggested which are capable of being tested directly.

doctor's actions through the devising of contracts is questionable. The patient's ability to monitor the agent's actions are clearly limited, and the costs of monitoring are probably higher in health care than in any other market characterised by an asymmetry of information

Rocnaix (1989) considered the ability of patients to monitor their doctors' behaviour. Her model was characterised by multiple principals and multiple agents. She argued that all that is required for doctors to behave in patients' interests is for a small number of patients to be informed about optimal treatment. The fact that there are multiple agents results in the threat from principals that they will seek a second opinion from these 'other agents' if the treatment decision by the doctor differs markedly from their a priori expectations. This threat forces doctors to act as 'better agents'. The fact that there are multiple principals means that this result holds for all individuals because agents cannot distinguish well informed principals from ill-informed principals.

This model makes a number of assumptions that are open to question. Like Coyte, Rochaix assumes that patients recognise the need to monitor their doctors. This is clearly an empirical question. Even if patients do recognise this need, given that the medical profession cannot agree themselves on what is 'best' medical practice (as is evidenced by the extensive variations that exist in medical practice e.g. see Ham, 1988), how can we expect patients to know what 'best' medical practice is?

Even if some patients do recognise the need to monitor the doctor's actions, and know what 'best' medical practice is, the usefulness of this model would be restricted largely to sectors of the health care market characterised by multiple agents. That is, whilst it might be applicable to general practitioners (where patients can, to a certain extent, 'shop-around'), it would not be so readily applicable to specialists.

Rochaix' model assumes that doctors cannot distinguish well informed patients from the less well informed, and that all patients are treated in a homogeneous way. A review of the non-economics literature [Ryan, (1992b)] indicates that patients are not homogeneous, and doctors do not treat them as such. Given this, doctors may respond differently to different patients. That is, patients who they perceive to be relatively well informed may be recommended a different treatment plan than less well informed patients. Or it may be that patients are recommended the same treatment plan, but end up with different treatments. That is, relatively well informed patients may question their own doctor rather than 'shop-around'.

It is at this level – i.e. differences across individual patients and individual doctors – that the issue of risk aversion and its relevance to the optimal fee structure becomes relevant. (See Section 2.1 and Arrow, 1986.)

The models of Coyte (1984) and Rochaix (1989) both assume that monitoring in health care is carried out by principals. Fama (1980) recog-

nised that, within the context of multiple agents, 'vertical monitoring' was present in the firm. Extending this to the market for health care, doctors may be monitored by other doctors. Indeed, monitoring in the health care market by agents is both 'vertical' and 'horizontal'. Members of the medical profession are monitored both by more senior doctors and by doctors at a similar level to them i.e. through medical audit, peer group reviews.

It was shown earlier (Section 2) that dynamic interactions lead to more efficient outcomes. Whilst it is possible that repeated interactions constrain doctors' behaviour, since doctors will not want patients to lose faith in them, such repeated interaction will not take place in all sectors of the health care market. Whilst a dynamic model may be applicable to the GP-patient interaction (since there will be repeated interactions), it is less clear how applicable such a model is to the specialist–patient interaction.

4. Conclusion

Although research in the past twenty years has resulted in an increasingly sophisticated literature on the theory of agency, the application of this literature to the doctor–patient relationship remains somewhat limited.

The literature on standard agency theory has concentrated on the devising of incentive compatibility constraints to reduce inefficiency. Given the interdependence of the patient's and doctor's utility function, a necessary (but not sufficient) condition for an efficient outcome in health care is that we establish the nature of these utility functions. Only then can we attempt to devise optimal reward systems that encourage doctors to consider patient preferences in the decision making process.

A clear conclusion from the principal–agency literature is that more thought needs to go into methods of remuneration of doctors. The theory of agency suggests that complicated fee schedules will prevail in markets characterised by an asymmetry of information. Health economists can be criticised here for a lack of detailed consideration of optimal remuneration systems. Indeed, it is still possible to find health economists arguing against fee-for-service medicine as a system [Culyer (1989)] rather than considering the details of the fee schedule.

The standard theory of agency has concentrated on devising optimal methods of remuneration. When we apply the theory to health care, the emphasis needs to be on researching in more detail into first, the patient's utility function, second, the doctor's utility function and, only then, into optimal methods of remunerating doctors.

References

Alchian, A. and H. Demsetz, 1972, Production, information costs, and economic organisations, American Economic Review 62, 777–795.

Arrow, K.J., 1986, Agency and the market, in: K.J Arrow and M.D. Intriligator, eds., Handbook of mathematical economics, vol. III (Elsevier, Amsterdam)
Culyer, A.J., 1988, Inequality of health services is, in general, desirable, in: D.G. Green, ed., Acceptable inequalities (IEA, London).
Culyer, A.J., 1989, The normative economics of health care finance and provision, Oxford Review of Economic Policy 5, 34–58.
Coyte, P.C., 1984, The market for medical services and physicians: an application of hedonic price theory, Research Paper (Department of Economics, University of Alberta, Edmonton).
Donaldson, C. and K. Gerard, 1989, Paying general practitioners: shedding light on the review of health services. Journal of the Royal College of General Practitioners, 39, 114–117.
Evans, R.G., 1974, Supplier-induced demand: some empirical evidence and implications, in: M. Perlman, ed., The economics of health and medical care (MacMillan, London)
Evans, R.G., 1984, Strained mercy: the economics of Canadian medical care (Butterworths, Toronto).
Evans, R.G. and A.D. Wolfson, 1980, Faith, hope and charity: health care in the utility function, Discussion Paper (Department of Economics, University of British Columbia, Vancouver).
Fama, E.F., 1980, Agency problems and the theory of the firm, Journal of Political Economy 88, 288–307.
Hay, J. and M.J. Leahy, 1982, Physician-induced demand: an empirical analysis of the consumer information gap, Journal of Health Politics, Policy and Law, 9, 595–609.
Holmstrom, B., 1979, Moral hazard and observability, Bell Journal of Economics 10, 74–91.
Holmstrom, B., 1982, Moral hazard in teams, Bell Journal of Economics 13, 324–340.
Krasnik, A., P.P. Groenewegen and P.A. Pederson et al., Changing remuneration systems: effects on activity in general practice, British Medical Journal, 300, 1698–1701.
Lubs, M.L. and R.F. Falk, 1977, Response of relatives to prospective genetic counselling in severe X-linked disorders, in: H.A. Lubs and F. de La Crux, eds., Genetic counselling (Raven Press, New York).
MacDonald, G.M., 1984, New directions in the economic theory of agency, Canadian Journal of Economics 17, 415–40.
Macnicol, A.M., M.L. Watson and A.F. Wright, 1986, Implications of genetic screening programmes for polysystic kidney disease, Aspects of Renal Care 1, 219–222.
McGuire, A.J. Henderson and G. Mooney, 1988, The economics of health care. An introductory text (Routledge and Kegan Paul, London).
Mooney, G. and M. Lange, 1991, Economic appraisal in pre-natal screening: reassessing benefits. Health Economics Research Unit Discussion Paper No. 08/91 (University of Aberdeen, Aberdeen).
Mooney, G., 1991, Agency in health care: what's the Maximand? Health Economics Research Unit Discussion Paper No. 10/91 (University of Aberdeen, Aberdeen).
Rice, T., 1984, Physician initiated demand for medical services, Journal of Health Politics, Policy and Law, 14, 587–600.
Rochaix, L., 1989, Information asymmetry and search in the market for physicians' services, Journal of Health Economics 8, 53–84.
Rubenstein, A. and M. Yaari, 1983, Repeated insurance contracts and moral hazard, Journal of Economic Theory 30, 74–97.
Ryan, M., 1992a, Supplier induced demand: where are we and where should we go? Health Economics Research Unit Discussion Paper No. 05/92 (University of Aberdeen, Aberdeen).
Ryan, M., 1992b, The economic theory of agency in health care: Lessons from non-economists for economists. Health Economics Research Unit Discussion Paper No. 03/92 (University of Aberdeen, Aberdeen).
Shavell, S., 1979, Risk sharing and incentives in the principal and agent relationship, Bell Journal of Economics 10, 55–73.
Williams, A., 1985, Economics of coronary artery bypass grafting, British Medical Journal 291, 326–329.
Williams, A., 1988, Priority setting in public and private health care, a guide through the methodological jungle, Journal of Health Economics 7, 173–183.