Back to Basics in the HEALTH CARE DEBATE

Richard A Epstein

NEW ZEALAND BUSINESS ROUNDTABLE

Back to Basics in the HEALTH CARE DEBATE

Richard A Epstein

NEW ZEALAND BUSINESS ROUNDTABLE
SEPTEMBER 1996

The paper 'The Perils of Health Care Provision' is based on a seminar given at the Institute of Policy Studies, Wellington, on 18 December 1995

Also included in this publication is an article 'Why is Health Care Special?' which was published in *The University of Kansas Law Review*, Volume 40, no. 2, Winter 1992. The agreement of the publisher of the *Law Review* to the reproduction of this article is gratefully acknowledged.

First published in 1996 by
New Zealand Business Roundtable,
PO Box 10–147, The Terrace,
Wellington, New Zealand

ISBN 1–877148–18–0

Design and production by *Ruth Munro, Daphne Brasell Associates Ltd, Wellington*
Printed by *Astra DPS, Wellington*

Contents

Richard A Epstein

RICHARD A EPSTEIN is the James Parker Hall Distinguished Service Professor of Law at the University of Chicago, where he has taught since 1972. Previously, he taught law at the University of Southern California from 1968 to 1972.

He has been a member of the American Academy of Arts and Sciences since 1985 and a Senior Fellow of the Center for Clinical Medical Ethics at the University of Chicago Medical School. He served as editor of the *Journal of Legal Studies* from 1981 to 1991, and since 1991 has been an editor of the *Journal of Law and Economics.*

His books include *Bargaining With the State* (Princeton, 1993); *Forbidden Grounds: The Case Against Employment Discrimination Laws* (Harvard, 1992); *Cases and Materials on Torts* (Little, Brown, 5th ed, 1990); *Takings: Private Property and the Power of Eminent Domain* (Harvard, 1985); and *Modern Products Liability Law* (Greenwood Press, 1980).

Professor Epstein has written numerous articles on a wide range of legal and interdisciplinary subjects and taught courses in contracts, criminal law, health law and policy, legal history, property, real estate development and finance, jurisprudence and taxation, torts, and workers' compensation.

His latest book, *Simple Rules for a Complex World* (Harvard, 1995), grew out of a series of lectures and seminars given in New Zealand and Australia in 1990. This lecture touches on some of the themes from his forthcoming book, *Mortal Peril: The State Control Over Health Care* (Addison-Wesley, 1997).

THE PERILS OF HEALTH CARE PROVISION

HEALTH CARE HAS BECOME an enormously contentious political and intellectual issue over recent years, both in the United States and elsewhere in the world. Previously the debate had taken place mainly among technical people – academics and health specialists. Now everyone is discussing health care. Indeed, it has moved so far up the political agenda that a dispute over the size of the Medicare premium recently shut down the United States government. I cannot do justice to all of this debate in a short address. I will give a brief overview of the big picture as I see it, and then focus on one part of that picture in greater detail.

The health debate can be divided into two broad sets of questions. One set consists of the bioethical questions. Here issues of law and economics and scarcity are always relevant, but in common discourse they tend to relate to isolated circumstances or discrete transactions. Bioethical questions cover issues such as organ transplantation – how organs should be supplied, whether sale should be permitted and gift encouraged, and so on. They also cover difficult issues concerning death and dying, from voluntary euthanasia to questions about liability – for example, the way in which health providers should be held responsible for negligence or misadventure. I will make only one comment today on the bioethical issues, which is that the reason they cause such enormous difficulty is that we have tended to treat them as special. Standard common law rules on the rights of disposition and self-control by contract have been applied to these areas erratically at best. Instead, we wrongly prohibit the sale of organs, just as we wrongly prohibit voluntary euthanasia.

While these interferences with voluntary choice and exchange constitute vital substantive issues, it is the second set of questions – those concerning access to health care – that I plan to discuss today. Who gets to be treated? Under what circumstances is the treatment to be supplied? Can anybody be excluded from the coverage of health systems? Are there special obligations to treat persons who are in emergency conditions? Faulty answers to these questions can bring a health system down, while sound responses may help to save systems that are otherwise destined to implode from over-use.

Private provision: fee-for-service and insurance

I propose to start with private provision, and to examine how the provision of health care has been transformed in the United States – and doubtless elsewhere – through private market transactions. Understanding this pattern will provide us with lessons which can be useful when assessing a variety of public health care systems, from the national health services of countries such as New Zealand or Britain to America's version of Medicare – an institution probably unrivalled in the world for its baroque administrative complexities.

In the United States there has been a major revolution in private health care – a revolution driven largely, if not exclusively, by market forces. We have seen fee-for-service medicine lose its role as the dominant mode of medical provision. Once it became clear that there were no legal impediments to the corporate practice of medicine – that firms could supply medical services as well as independent physicians – the pressures placed upon fee-for-service medicine in the United States became enormous. At the primary care level, fee-for-service had held the field for a long time because it had a number of obvious advantages. It was extremely cheap to set up. There was no need for large investments in a system which involved a physician seeing a patient in an examination room and charging a fee for the services rendered. The economic unit providing the service was relatively small, and the patient enjoyed the exclusive loyalty of the physician. While the fee-for-service system raised some conflicts of interest between patient and physician, at least it introduced no third party to whom both owed obeisance. A strong element of trust could develop, and the relationship could last for many years.

Yet in the end this relationship came under pressures that it could not withstand. First, with time and the advance of technology, the inherent conflicts of interest grew and proved progressively more difficult to overcome. A fee-for-service physician who recommends more services will generally do better financially than one who recommends fewer. There is thus a temptation for the physician to over-prescribe services, if he or she can get away with it. And a patient may not have – or be in a position to acquire – sufficient counterbalancing information to monitor a physician who is motivated by financial considerations rather than professional or medical considerations. Formerly, detailed professional codes were designed to curb such tendencies on the part of physicians. Technological change, however, has generated more, and more expensive, procedures and tests, and thereby has increased the scope and intensity of these conflicts of interest. The informal mechanisms developed in the early fee-for-service relationships have tended to lose their effectiveness as the technological and financial stakes have become larger.

The development of conflicts of interest is one reason for the shift away from fee-for-service medicine. A more important factor, however, relates to the very nature of medical intervention and sets health provision apart from services such as education. Most individuals are highly uncertain as to the health costs they are likely to face over the course of their lives. All of us generally have a low probability at any given time of experiencing a catastrophic health problem and incurring very high costs of treatment. If faced with such a contingency, however, fewer than 1 percent of New Zealanders or Americans could cope with it within their household budgets. Yet in quieter times they do have sufficient resources to buy insurance against that risk, owing to the low probability of a catastrophic event occurring. The arguments for pooling risk are thus compelling, and they create a powerful impetus for developing insurance markets for health care.

However, it turns out that any insurance backing to fee-for-service medicine creates a losing combination. The physician still has his or her discretion and the incentive to over-prescribe, which is magnified since the costs are borne by an external payer. The patient and the physician, acting primarily in their own interests, can pile up costs almost without limit until they meet with sustained external resistance. Thus, in the United States, the trend to add insurance on to fee-for-service medicine, in order to handle the problem of the low probability-high loss event, in time caused huge losses for insurers. The need to manage this second generation of risks became one of the major challenges in the marketplace and led to a major restructuring of the practice of medicine. The insurers essentially refused to run a system in which they were meeting the bills unless they also had some control over the outputs and inputs. So by degrees the pattern of handling residual risk changed to one in which an individual, or employer, or some other group, paid a health care provider a fixed – or largely fixed – fee to take on the risk of variations in health status. Yet the assumption of risk by a health care provider went hand in hand with new forms of supervision and review. This shift of residual risk from the patient – or from the employer on behalf of the insured employee – to the health care provider has completely transformed the practice of medicine in the United States. In markets generally, control follows risk – and health care has proved no exception.

There have been two instructive consequences of this change. The first development is that the new third party relationships mean that physicians lose a great deal of the financial and professional independence they had formerly enjoyed. They become employees or – possibly even worse from their perspective – owners of various businesses, where they hire other individuals to be employees and monitor their outputs and inputs in a detailed fashion. Monitoring does not consist of telling a doctor that the prescription was right or wrong in any given case. The standard review is a utilisation review: it looks, after the fact, at the book of patient business, and examines, say, the number of tests that were ordered by a physician for thousands of individuals relative to the numbers ordered by other physicians. If these numbers are out of line, the physician will face sanctions. There is no reason to think that utilisation reviews are misguided. So far as we can tell, not even a small improvement in health outcomes results from interventions by physicians who provide vast amounts of additional services. The health sector in the United States has seen an enormous expansion – from 10 million subscribers in the non-Medicare population in 1976 to over 50 million people in the course of 18 years, with most of that growth in the last five years. Yet the data suggest that a huge proportion of medical procedures are producing relatively small increments in health.

The second element that has changed in the system is the nature of the services physicians supply. A presentation was recently given at the University of Chicago by a University of California professor of psychiatry whose specialty was the study of empathy in physician-patient relationships. She had discovered that empathy was not in plentiful supply in the Californian health plans she had reviewed. It is not difficult to see why empathy is a big loser. Physicians have two kinds of inputs they can supply: empathy and medical expertise. Empathy is a quality that many health care professionals have, or should have. It can be acquired at a relatively low price. But professional training and the ability to do complicated diagnostics is expensive. It is also difficult for a physician to become empathetic in a visit lasting between four and six minutes – the typical consultation. Having somebody with a high level of ability sitting in a capital-intensive clinic and dispensing empathy is not cost effective. So with the shift in control, the shift in behavioural patterns follows.

My point is not that empathy is unimportant in a patient relationship – clearly it is. But people will need to find it somewhere else when physicians do not provide it. One cannot, in effect, keep high-class medical capital idling on the empathetic track, given that it is tying up everything else. Physicians try to resist this trend because they sense, rightly, that it can dehumanise their work. But the cost pressures are so great that they will be able to make headway in their concern only if they can show emphatically that treating the whole patient translates into superior diagnosis and treatment that can help to lower system-wide costs. That result might well be true in dealing with cases that have persistent psychiatric components, but in the present environment that conclusion cannot be presumed as a matter of course. It has to be established, not by anecdotal evidence but by the now-standard comprehensive review of data bases, how the additional effort yields superior diagnosis, shorter treatment periods or improved outcomes. The system-wide approach to medical care evaluation taken by large health care providers shows no exceptions for psychological conditions. They are subject to the same kinds of review as surgeons or internists. Yet, here, some additional uneasiness is the order of the day because the difficulties of documentation are hardest with mental conditions, however real, that have no obvious physical correlates. The problem is too large just to disappear. Indeed it promises to be one of the major bones of contention in the years to come.

The insistence on continuous system-wide evaluation is the new constant of the current system. Indeed, in studying the trends in American health care today, one major task is to understand, within the capitation system, how medical providers should now optimise their behaviour. What types of services should they cover? What level of care should they provide? What kind of referrals should they authorise? The brunt of these concerns is directed to such questions as experimental treatments and alcoholism, and psychiatric care again comes into real scrutiny. But it would be a mistake to assume that the shift to a general capitation system dictates some uniform answer. One can have a low fee system or a very high fee system, and within any structure all manner of elements can vary – the scope of covered services, the length of queues patients will face, the kinds of facilities in which they will be treated, the level of physician competence, the amount of work going to specialists or to nurse pracititioners, and so forth. Another issue is whether a given plan should be 'open' or 'closed' – whether patients can only see physicians participating in the plan, or whether the affiliations are looser. These are all factors currently being determined within a market framework. In such choices there is never a single, exclusive, optimum; a great deal depends upon the population base and individual preferences.

Government failure in health care

These trends in private provision give us valuable insights when we come to examine the changing nature of government intervention in health. It is a serious mistake to assume that, simply because governments are involved, their interventions and the outcomes will be similar to those of the private sector. Exactly the reverse is true: the outcomes of public health plans depend as much on their specific structures and incentives and internal design as they do in the private sector. One example comes from the now familiar area of psychiatric care. Several major American corporations have found it possible to supply some psychiatric care to their employees and their families at reasonable cost. It is sometimes assumed that if private firms can take on this burden any comprehensive public plan should be able to do so as well. The failed Clinton health plan constantly touted its broad coverage for mental illness.

But the underlying conclusion is false. The translation from private example to public programme is hampered by hidden pitfalls. The fact that one, two or 10 corporations adopt a given programme is hardly dispositive. We also know that thousands of corporations, hardly distinguishable in outward appearances, refused to adopt these programmes. Why assume that the practices chosen by a minority of firms are suitable for the population as a whole? It is far wiser to first ask whether some special condition in the workforce – the average age, or educational level, or occupational direction – may have a good deal to do with the nature of the firm-specific risk and with the frequency of claims, both sound and bogus. Underwriters always look to the composition of an insured pool before making their decisions on whether to offer cover and, if so, how much to charge. We should not automatically assume that a low level of market penetration creates large opportunities for government action. It may well signal a genuine insurance disaster if the characteristics of the uninsured populations differ radically from those of the insured populations. We need to learn from private inaction as well as private innovations.

A second lesson is closely allied to the first. Never assume that we can translate the experience from one public programme to another. Once we get past the labels of national health care, the details start to matter, often in unanticipated ways. Britain provides an example. Unlike the American system of public support for health care, capitation was build into the National Health Service virtually from the day it started. Far from leading to an explosion in medical costs, the incentives placed upon physician-providers in Britain produced exactly the opposite result. If a doctor only receives a limited payment per patient, he or she will to some extent attempt to maximise personal welfare by economising on costs. For patients, that means, among other things, waiting in queues. One can easily develop a system of socialised medicine in which too little is spent on health care. In such a regime, too much of the cost of a poor policy will be borne not in the form of financial payments within the system but effectively through private 'payments' by individuals in the form of wasted time and excessive inconvenience and unnecessary pain. These are real costs even if they do not appear on the balance sheets of any official budgets.

The big difference between the rapidly-growing managed health plans in the United States and a government capitation system is that the latter is a one-size-fits-all product. In the United States, within the voluntarily-agreed framework which places the residual risk of catastrophic loss on the health care provider, there is an infinite variety of possible plans to handle different populations, and different fee structures to match. Plans can be sensitive to age, income, the types of diseases to which the population is subject, the accidents they are likely to suffer from, and all manner of other factors which are reflected in the basic logic of insurance contracts so as to maximise the joint gain to the parties. One can minimise all sorts of risks in the provision of medical services by the appropriate contract design. In government capitation systems, however, that degree of experimentation and flexibility is lost. The premiums are set by an external process and the feedback loops are weak since the adverse financial consequences are diffused in a sea of political unaccountability.

The contrast between the British and American public health plans is most instructive. When Medicare was passed in 1965, the problem was not that of a socialised medical system resulting in too little medical care rather than too much. It is interesting to compare the rhetoric during the passage of Medicare with the rhetoric 10 years later when the accident compensation scheme was introduced in New Zealand. The Medicare system was designed at a time when the dominant mode of provision was fee-for-service care coupled with voluntary insurance by outside parties, and with relatively little professional management or even utilisation review. The early cost projections for Medicare were made on the basis of some quite extraordinary assumptions. It was assumed that the demand for medical services would not change, despite the enormous change in prices faced by users of the system (who were now paying little or nothing for treatment at the margin). It was also assumed that the supply of services would not change either, despite the fact that payments would now be guaranteed independently of the ability of any individual to pay. There was a projection made in the 1960s according to which the cost of Medicare would level out around the year 1973. That projection turned out to equal only 10 – 15 percent of the eventual cost of Medicare less than 20 years later, in constant dollar terms. The failure to take account of the incentive effects led not just to a small 5 or 10 percent error but to a massive five- or six-fold error.

There are two main structural features of Medicare which help to explain its complete failure to contain costs. First, not a single fee paid by private recipients under Medicare relates in any way to their marginal receipt of services. No payment in Medicare is designed to influence behaviour at the margin. Indeed every measure is taken to separate that connection. Everything is structured so that price increases or decreases merely have distributional consequences, where the only question is how much one group will pay relative to another. All patients, regardless of age, medical history and need, pay a flat monthly premium for the services they receive, and also pay 20 percent of the bill for individual consultations.

The first element of Medicare structure, the front-end premiums, can rise and fall. But no matter how high they are made, no one would drop out of a system that is funded three-quarters by a third party and less than one-quarter by the user. Even if the total percentage payment from the individual patient went up from 25 percent to, say, 35 percent, the dropout rate from Medicare would be extremely small. And since the fee is fixed at the front end, raising it by $5 or $10 per month would not lead to any material change in utilisation. Currently there is a titanic struggle in the United States between the Democrats and Republicans over the level of the fee. Everybody understands that this is simply a funding question – it is solely about how much should be paid by the recipients of health care and how much should be paid by others. The fee does not constitute a price mechanism which will alter the consumption levels of individual patients.

The second part of the Medicare structure is the element which *is* designed to limit utilisation – the requirement for the recipient to pay 20 percent of the cost of service. But we also have a peculiar system of insurance called Medigap, which takes the burden of those costs away from individual patients. The insurer picks up the full 20 percent of the Medicare gap, in exchange for another fixed premium. How does Medigap coverage survive in the context of the American health system? The answer lies in an insurance externality common in any dual or co-payment system. Insurers can work out under Medigap the likely levels of utilisation, and then calculate their premium as a sum which is slightly larger than the probability of the utilisation of service times the average amount per insured occurrence. Since it is only the gap in Medicare that is being covered, fairly high utilisation levels can be tolerated because the Medigap insurers are only making part-payments. With that insulation, the willingness of Medicare recipients to consume heads through the roof again. And when consumption rises further, it triggers not only the Medigap payment but the entire payment from the system at large – the other 80 percent. So the total increment in utilisation attributable to Medigap insurance is huge. Private insurance coming in to fill a gap in Medicare has ended up creating enormous externalities in the form of over-use of the system. Looked at as a whole, the package results in the insured person having no price incentive whatsoever, at any stage of the process, to constrain consumption.

Thus politicians keen to rationalise the whole system are confronted by a sea of costs. How can they go about fixing these problems? In dealing with the costs, they can target essentially three parties – the taxpayer at large, the recipient of medical services, or the health care providers. Those are the only options. And nothing they have tried so far appears to have worked.

The strategy with respect to the practitioners has been perhaps the most interesting. In an attempt to rationalise costs we devised a horrible term known as 'resource-based relative value scales' – RBRVS to weary insiders. This system involves calculating the prices that a competitive market would charge for the entire range of medical services. A medical professor from Harvard University came up with this concept; no one from Chicago would be so naive, or so I would hope. The process involves working out the cost of all the various inputs – facilities, doctors, liabilities and so on – and then adding these constituents together to obtain some standardised price. And to show that the compilers of the index really knew what they were doing, it was calculated to four decimal places. They added in adjustments for regional differences, multiplied it all together, and purported to derive the cost of, say, an angiogram for a given person in a given region.

Calculating prices in this way is guaranteed to lead to extraordinary rigidities in the system. All the numbers need to be cranked in every time you make any changes in procedures. These numbers can never be recalculated quickly enough for the system to respond adequately to market signals. For instance, Medicare was a low payer relative to the private system in the days when we had fee-for-service medicine and Blue

Shield insurance. Then, when we changed over to health maintenance organisations, the price ratios completely switched. The index put together by the Harvard people has not changed much while the market for complicated procedures has undergone a revolution in the last five to seven years. In other words, the standard litany that price controls are a set of rigid constraints, unable to respond to marketplace realities, turns out, unfortunately, to be at least as true in medicine as in any other part of the economy.

The second point is that the American Medicare system does nothing to target the demand side and control utilisation. So long as the private cost of a given service is lower than the private demand, people will over-consume it. Moreover, since the providers face a fixed-price scale, they try hard to introduce a large volume of simple procedures that can be performed at a cost lower than the standard payment for the category, while avoiding all of the difficult medical cases. In other words, a provider makes money under Medicare by ensuring that everybody in an old-age home has, say, an annual eye examination, whether or not it is needed. The strategy is to uprate, or upcode, the simple procedures to a higher category, perform as many as possible, and spread fixed costs over a large volume of interventions. It is another form of maximising income subject to an external constraint.

The biggest expansion on the demand side in the United States happened only after the cost control systems were introduced. This may not have reflected cause and effect, but clearly nothing in the system prevented the increase from taking place. All the efforts to improve the way the government provides payments for individual services have failed, which is why we had the great political impasse in the fall of 1995.

Currently the Republicans are attempting to convert Medicare from a fee-for-service system into a managed care system by paying a fixed dollar amount to health care providers who would then bear the residual risk. Much good might be expected from such a system, but it is both economically and politically very difficult to implement. When Medicare was brought in, the main focus of public concern was the size of subsidy that would be paid by the population under the age of 65 to those over 65. The original social contract called for a 50/50 split between taxpayers and recipients. That did not prove to be. After a technical series of cost-of-living adjustments in the 1970s (which were designed to make sure that recipient medical costs did not increase more rapidly than overall inflation when medical costs were soaring), the recipients are now paying only around 27 percent of the total cost. The taxpayers pay the remainder – a very substantial deviation from the original design, with real adverse consequences on job formation and standard of living for ordinary working families.

That problem aside, the effort to rationalise Medicare runs into a second difficulty, one which is less widely understood but which turns out to be at least as important, namely the question of internal subsidies within the Medicare population. Not only do members of the Medicare population pay much less for the care they receive than it costs to provide, but the size of the subsidy received also varies enormously between recipients. Medicare covers everybody from 65 until death, which in many cases is 85 or 90. Clearly the medical expenses incurred by a healthy 66-year-old are much lower than those associated with an ill and infirm 84-year-old. Since Medicare is a defined benefit programme, a low subsidy for one group can co-exist with a high subsidy for another group. There is no pressure to reduce this disparity because the subsidies are all hidden behind the rhetoric of benefit entitlements. But as soon as attempts are made to shift to the other system – where it is not the benefits which are defined but the contributions the government will pay – the unequal nature of the subsidy is brought into the open. If people were given a voucher to buy their own health care, the problem would become obvious. With a voucher equal to the average cost of a health care plan, young and healthy people would be snapped up by private insurers and still have money left over. That same voucher of perhaps $3000–4000 per year given to an 84-year-old on a fixed pension might cover on average only a small fraction of his or her total health costs, especially if the recipient is undergoing treatment for some chronic condition. As in education, all the efficiency experts promote vouchers while the politicians refuse to touch the idea. The politicians understand that the efficient economic solution – vouchers – would force them to be explicit about the amount of redistribution they wish to undertake in health care, and they cannot bring themselves to decide on the parameters of that redistribution. Should the voucher be adjusted for age? Or prognosis? Nobody is sure. And, unlike education vouchers, it is much harder to see our way clear to a sensible solution.

This is the same type of problem that happened in reverse with the capitation programmes. A government works very badly when it assumes simultaneously the multiple roles of an employer, a provider, a payer and a referee. It works much better to the extent that it concentrates on enforcing private contracts. Yet Medicare made it impossible to encourage long-term health contracts – lifetime care provision – which people could enter into at age 30, 50 or 70, much as they buy life insurance with guaranteed renewal provisions. These could have handled the subsidy problem, not by transfers from one group to another, nor by people receiving different amounts from the public purse, but by allowing lifetime providers to say in effect to people: "We will give you a flat payments system for your health care premiums. You will pay more in the early years than you need to, in order to build up the cushion needed for later years". The health problem, I believe, can be handled by long-term contracts if they are allowed. But since we did not allow them, 30 years later we are faced with such a mis-match between ability to pay and the demands for consumption that it can only be handled by transfer payments between individuals and groups. These transfers are much easier to sustain in the dark of night – as under a defined benefit plan – than in the light of day under a defined contribution plan.

This has resulted in some exquisite ironies in American health care. We were told that Medicare would ensure there would not be two tiers of services – that old folks would not be left behind. Now fee-for-service medicine is a rock on which the Medicare system was first built and now founders, so that those people who don't pay their full care go luxury class while those who have to pay for their own medical care are herded – to use the critical term – into managed care for which they pay full premiums. And on top of those full premiums they pay high taxes – an explicit Medicare tax as well as general income tax. So health care money is now spent, in many cases, where it provides the least long-term benefit. The very sick and the very elderly, in the last stages of life, are subsidised by other individuals who receive relatively inferior health care because they are forced to pay for it themselves.

The history of Medicare illustrates yet again a very important principle – that, in assessing how an insurance or social security plan will work, we cannot simply take its first year of operation and assume that the favorable relationship between its costs and benefits in that year will be maintained over its expected life. As schemes grow older, the maturity risks grow large. This is the problem that New Zealand has experienced with the ACC scheme and it is the same problem we in America have had with any long-term entitlement system. These systems survive for five or 10 years, but by the time they are 30 years old we fervently wish we had never started them. But by then everyone is afraid to say so publicly because of the strong constituencies that profit from their continued operation. We have to learn from this cautionary tale two sobering truths. First, ordinary market principles apply to health care arrangements as they do to any other system of human activities. Second, the ordinary maxims of public choice theory apply to the government provision of these health care services.

WHY IS HEALTH CARE SPECIAL?*

Ethical Propositions and Political Discontent

There are several propositions, both normative and positive, on which there is widespread agreement today about the delivery of health care. The first of these is an intangible proposition whose clout is not evident at first blush, but which turns out to be of immense theoretical and practical importance in any event. That proposition is that health care is "special." The second proposition, which is said to follow from the first, is that the special nature of health care calls for the intervention of government into the operation of the market – so that the ordinary intersection of supply and demand should not be allowed to determine the level of resources devoted to medical care, and, more importantly, shall not be allowed to determine who receives how much of that care and why. This view is sometimes captured in the proposition that health care is a right and not a privilege. The distinction between rights and privileges has a somewhat specialized meaning, with an important role to play. To say that health care is a privilege means that it is available only to those who are able to purchase it in the market. Legal protection is provided only to the extent that it prevents third persons by force from interfering with any contract between an ordinary individual and a health care provider. To say that health care is a right changes radically the nature

---

* Koch Distinguished Scholar-in-Residence inaugural lecture, given under the auspices of the Koch Distinguished Professorship in Law and Economics and the Fred C. and Mary R. Koch Foundation, on September 10, 1991, at the University of Kansas.

of the correlative duty. Now each person is entitled not only to purchase services in an open market, but to call upon other persons through government to supply that health care free of charge, or at least at a price below what it costs to fund it. The so-called privilege of obtaining health care in the marketplace imposes no correlative duties on the public at large to fund the care provided. The right to health care through government intervention necessarily requires the imposition of public taxation and regulation, probably both.

In some circles the rejection of the market has taken on the life of a categorical imperative – that is, a first premise that cannot be falsified by any empirical evidence or theoretical arguments to the contrary. That premise is that access to health care should be determined wholly without regard to the ability to pay – where the ability in question is that of the individual recipient, and not that of a robust system which has at its disposal the powers of regulation and taxation. This proposition can be stated in two different ways. First, in its more aggressive form, the proposition states that all persons are entitled to an *equal* level of care regardless of the ability to pay. Second, in its more qualified form, the proposition is that all persons are entitled to some *minimum* level of health care regardless of the ability to pay. Although these are both rejections of the market, they point to very different systems of regulation.

The second set of reasons that health care is sometimes said to be special is empirical. The area of health care is governed by extensive regulation, which in its modern form began with the passage of the Medicare statute in 1965.[1] At that time there were was a genuine optimism about how the system would work. It was assumed that demand for Medicare would be uninfluenced by price, and that price would be influenced by demand. It came as a major surprise that utilization of the system rose when federal funds were pumped into the area. Since that time, other systems of payment and support have been introduced, which have also influenced patterns of system use. The total

---

[1] Health insurance for the Aged Act, Pub. L. No. 89-97, 79 Stat. 290 (1965) (codified as amended in scattered sections of 26, 42, & 45 U.S.C.).

failure to foresee how physicians, hospitals and patients would respond to the change in incentives is one of the great failures of the planning process.

The consequences have been severe and long-lasting. The system as it is now constructed, by happenstance and compromise, does not work. The costs of running the system are inordinately high; the percentage of our gross national product spent on medical care has risen steadily from the 1960s to 13 percent today, and is still rising.[2] Yet the level of coverage afforded by the system is getting lower; the number of persons who are outside the system continues to increase; the Medicaid system is in a state of disarray in many states, with low compensation levels and high delays in repayment; and the breakdowns and bottlenecks – of which the recent studies on access to emergency room care are perhaps the most dramatic – indicate that certain core institutions do not operate the way they are supposed to. At a more subtle level, the relationship between the cost of medical education and the anticipated earnings of most family practitioners and similar doctors is such that most physicians cannot afford to pay back the costs, with a market rate of interest, of their medical education. Indeed, the real salaries for physicians, especially in the later stages of the training, have been static or in decline. In short, the level of benefits that people expect to derive from the system seems to be heading down, while the costs of running the system seem to be heading up. The ethical imperative of universal health care without regard to ability to pay collides with the increased inability to fund the operation of the system at every level.

This discontent has spilled over into the political arena. There is today an odd coalition that favors some radical reform of the current system. It includes many prominent businesses and unions that are anxious to find some way to place on the public payroll their existing but unliquidated obligations to current and retired workers. These obligations are substantial: it is said, for example, that the cost of health care per automobile

---

[2] Nancy Watzman, *The Democrats' Health-Care Plan—A Nixon Leftover*, THE CHRISTIAN SCIENCE MONITOR, Aug. 9, 1991, at 19.

is greater than the cost of steel. The discontent comes too from health professionals who are unable to meet their current financial requirements from existing sources of funds. It comes from legislators and analysts who are taken with the ostensible success of the centralized Canadian system in coping with the problem of health care. It is too early to know whether these forces will be able to overcome the opposition of traditional health care providers, or indeed to forge an alliance with them wherein comprehensive health care builds on the present system of basic employee coverage. But it is not too late to stress that the issue is now "in play" and promises to dominate, without closure, the debate for the remainder of the decade (I could say millennium with equal accuracy!). At a time when planned economies are in retreat in Eastern Europe they have made a sector-specific renaissance in the United States.

Is Everything Special?

If this set of medical and political outcomes is what one means by special, then there is no reason why anyone should be pleased with the exalted status attached to health care in the United States. The sequence of errors that has led to the current impasse will take a long time for anyone to unpack, but I hope that I can give some little sense of what has gone wrong, and then point, cautiously, to some controversial changes in outlook and policy that might lead to a redefinition of the current system. It is a tall task, and one not easily discharged in a lifetime, let alone a lecture.

The first place to begin is with the title of this talk: Why is health care special? Those of you who have watched the operation of the regulatory system over the generations will know that the term special, for all its emotive power, has a certain tired quality. The term has been used in countless contexts to pave the way for some form of government intervention. The intervention of government regulation into labor markets has often been justified on the ground that labor is special—special because it should not be regarded as a "commodity" or an "article of commerce" to be bought and sold in the marketplace. The point here is not merely one of rhetorical flourish; it is also one of statutory construction, for the Clayton Act[3] makes just such an assertion when it exempts labor from the operation of the antitrust laws. Indeed the entire question of labor regulation has usually rested on this assumption, and competitive markets have been displaced by complex regulatory structures and statutes which have fulfilled the promise of their original supporters. They have taken markets that were perfectly ordinary, and have made them into something special – special because they are costly to navigate, inefficient and unresponsive.

Real estate markets have also been regarded as special, and the outgrowth has been a complex system of rent controls and land use regulation, which again has impeded the efficient flow of capital, and has created bruised and hostile relationships as political figures make constant adjustments between warring factions, all of which have strong incentives to misbehave for their private advantage. There are surely cases in which forms of regulation are warranted – oil and gas pooling, the control of nuisances and the like – but one does not have to rest on the argument that these markets are special. Instead there can be some very precise demonstration of the social losses that follow when rivalrous private behavior is not constrained by an appropriate set of legal rules. Indeed, it is to control the use of force that governments are instituted among people.

We must therefore apply some modest degree of scrutiny to the proposition that health care is special: surely it is important, but so is food, clothing, shelter, education, entertainment, and all the other goods and services that are necessary to sustain life and to make the life sustained worth living. Importance, however, is not an argument for government subsidy or support, for if it were then socialism would apply to things where it matters most, and lead to the most ruinous of consequences. Instead the importance, so to speak, of importance is simple: it is important to get the right set of solutions, be it private or public, to the problem

---

[3] Pub. L. No. 63-212, 38 Stat. 730 (1914) (codified at 15 U.S.C. §§ 12-27, 44 (1988 & Supp. 1989)).

at hand. Importance does not create a presumption in favor of government, or for that matter against it. It only raises the stakes for making a correct decision in the matter at hand.

Two things are missing in the critiques of labor and real estate markets. The first is an accurate account of what constitutes a market failure. Too often that failure is found in the inequality of income in the population at large – an outcome for which equal access to medical care regardless of the ability to pay is regarded as an accurate countermeasure. Yet to the extent that inequalities arise because persons with greater productivity receive higher returns for their labor, then so much the better. The second missing element is awareness that there are failures in collective decision making that are every bit as great as those which exist in private markets. It would be easy to assume that collective responses are preferred when markets are corrupt and governments virtuous. It is far harder to reach that conclusion when self-interest and corruption creates difficulties from both quarters.

Two Market Defects

In order to find, therefore, what is special about the market for health care, it is necessary to determine why the ordinary rules of supply and demand do not yield the optimal result, and then to see whether something can be done to redress that balance *without* undoing all the good things which markets are capable of achieving in disciplining suppliers and consumers, and in encouraging cost savings and the economizing of resources. As I see it, there are two problems here that do require some attention, and both yield somewhat unfortunate results.

1. Imperfect Information

It is difficult in many cases to get accurate information about the cost and effectiveness of medical care. The point is important, for where markets labor under systematic imperfection of information it is likely that people will make the wrong choices. The usual logic of the market is such that people surrender what they have in order to obtain something that they value more highly. If people do not know the proper values to attach to the services that they need, then it is quite likely that they will make the wrong choices, and be left worse off after the provision of medical care than they were before. Bad information distorts the relative evaluation of goods and services that is essential for markets to operate.

In the area of health care, the problem is apt to be of major proportions. If you buy a dozen oranges, and the first one does not taste good, it is possible to return the rest to the grocer or even throw them away. But if the first surgeon messes up an appendectomy, it is highly unlikely that a second surgeon will come along in time to set matters right. Information deficits are always likely to occur in a setting where you cannot test the services you receive before you use them.

To state the problem in that forceful sense, however, is not to indicate a clear solution to it. First, the question of imperfect information not only arises with private providers, but can also arise, and with equal severity, with public providers as well. Anyone who works under a budget constraint is likely to try to cut corners, and a government bureaucrat, who knows that his patients have nowhere else to turn, has no real incentive to impart the information that is necessary to make appropriate choices either. So in both settings there will be a certain number of transactions that will shipwreck, and the proper procedure is to compare imperfect alternative with imperfect alternative. It is not to act instead as though government intervention which is directed to the problem of information will solve that problem. It is always risky to assume that if the ends are laudable, the chosen means are efficient to achieve them.

In addition, the problem of imperfect information should not be understood as one that stymies the operation of markets. The people who work in markets are aware of the problem, and one of the tasks of a sound system of contracting is to develop those institutions that deal with information breakdown. Imperfect information is one reason why brand names are important to establish reputation, or why persons may decide to enroll in a group medical program. It is useful to hire third persons who know more about the business at hand to navigate the perilous waters one faces. The world is filled with middlemen and brokers, with factors and agents, and many persons assume such roles as part of their other responsibilities. One job of the architect is to supervise the work of the general contractor just as the general contractor is responsible for the supervision of the subcontractors.

In the medical area, one function of the internist is to operate a specialist-referral network for the benefit of his or her patients – to function, as it were, as Virgil in Dante's Inferno. The internist knows more than the patient about the arcane workings within the medical area, and, by virtue of having a large practice, is able to exert some clout over specialists that individual patients cannot. Similarly, agency relationships are at work when patients sign up with a Health Maintenance Organization (HMO) or other type of health provider, only now other intermediaries are at work. The company or union that selects the HMO may be better able to monitor the HMO's activities than individual employees; and the HMO may be able to monitor the activities of individual physicians better than either the company or its employees. The feedback loops are often complex, and the solutions are not perfect. Still the risk that the agent will betray the principal is normally worth running for the gains on the other side, and the risks are in any event similar to those faced when government agents, whom one may not be able to select, are called upon to furnish assistance. The question of imperfect information is common to many markets; and while it is severe, it does little to incline one to a government solution relative to a private one.

2. Patient Solvency

The standard arguments in favor of the market assume that the willingness to pay is the standard to determine who gets what goods. Those persons, therefore, who approach a particular situation with no wealth are entitled to receive nothing in exchange. The reasons why wealth and bids are used to allocate resources are many and powerful. In many cases there is a positive correlation between what one expects to gain from the receipt of a good and how much one is prepared to pay in order to get it. In certain markets, such as food, clothing, and shelter, the prices that are charged for goods normally are are sufficiently low and recurrent that they can be built into ordinary budgets. The systems that might be put in the place of markets – rent control for housing, price supports, or price ceilings for farm goods – lead to contentious wrangling for no discernible social purpose. People who just say how much they need, but who are not prepared to back up their words with deeds or dollars, will always be tempted to place as high a valuation on their needs as they can, at least with a straight face. Utility may be the philosopher's touchstone, but it is not a measurable quantity in plain view; and it is easier to determine whether someone has paid for a particular service than it is to determine whether that person needs services that can be furnished without the necessity of personal payment.

Unfortunately, in the case of medical goods and services, it is easy to envision situations where wealth is a poor proxy for utility or need. Matters are only complicated because medical expenditures, unlike food and clothing, often come in bunches, which are very hard to finance out of current earnings. If insurance markets are not perfect, and they are far from it, then there may be major difficulties in smoothing out the expenditures for medical care. Given these imperfections, we can say with complete confidence that wealth is not what one seeks to measure in the provision of medical care, but it is harder to say that it is not a proxy that in some cases at least helps to ration care in an appropriate fashion.

Charitable Care and Mandatory Access

The twin difficulties associated with information and solvency cannot just be dismissed here any more than they can be dismissed in other contexts. But the question is what type of response ought to be made. At this point, it is useful to distinguish between two different kinds of responses, both of which share the idea that contracts for service do not always offer the right way to provide medical care but differ in virtually every other way. The first of these is to organize some voluntary charitable institutions whereby medical care is given out free of charge. Before the rise of the public assistance programs, hospitals and private physicians routinely provided health care on just this basis, in a private effort to bridge the gap between utility and wealth. I can recall the extensive amount of free care that my father gave in his medical practice during the late 1940s and early 1950s; and hospitals also provided extensive amounts of free care. Where care is given on a free basis one worries less that there is an effort to provide someone with services that he does not need, even though there is still some concern whether the quantity or quality provided is sufficient unto the day.

To argue that voluntary responses are a complete answer to the problems of the health care market is, I think, a piece of idle optimism; but to demand that they function as a complete social response creates dangers of its own: the requirement of perfection places excessive demands on a system that cannot be made in any universe where resources are scarce and individual misbehavior a serious social problem. In effect, the system of private voluntary assistance is an incremental program whose sole claim for our affections is that it makes a bad situation better. It cannot claim to make a bad situation perfect, for it tolerates both differences in the level of care that is provided – those who can pay may well get better medical care – and more critically, it tolerates the possibility that some individuals will fall through the net, that is, the possibility that persons with serious and curable illnesses will die for want of appropriate medical treatment.

These weaknesses have not led to the elimination of the charitable and voluntary side of medical care: today it is still easier to raise money in the private sector for health programs than to do so for legal ones. But the weaknesses surely have led to a very different view of what the role of government is. The gaps that are left in the private system are regarded as gaping holes that have to be plugged, and systems of government provision are now used to take their place. In some sense, these systems are judged by the ethical imperative that I mentioned above, namely, that all persons are entitled to the same level of health care regardless of the ability to pay. The differences in levels of care that are routinely observed in health settings are a tribute to the elusive nature of the goal in question. What I want to state here is something somewhat stronger: the ethical ideal carries with it so much implicit baggage that it can only lead, if aggressively pursued, to many of the problems associated with health care that one sees in the United States today. After I have discussed this issue, I will give my most unethical prescription for what should be done.

Moral Hazard and the Egalitarian Impulse

The first question is why is the ability to receive care regardless of the ability to pay limited to health-related services. In principle, one could argue that all the other important rights that I referred to above – education, food, shelter, clothing, work – should be received on the same ground. Indeed, I think that the moral imperative here could easily be translated into an egalitarian prescription that the distribution of the goods of the world should be equal, or that deviations from equality should be based not on wealth, but on the severity of conditions: those with large appetites shall receive more food; those with more health problems should receive more medical care; and so on down the line. To state this is to see in the grand idea a deposit of the traditional socialist doctrines: to each according to his need, and from each according to his ability. In a world of angels, I could think of no moral postulate that better captures our true ideals. It is all for one and one for all. The levels of production are, in this happy universe, wholly unaffected by the distribution of the goods that are so produced, and we all count ourselves the better when any one of us prospers, so that each views his own satisfaction as though he were stranger to himself. A utile of satisfaction for me has the same influence on my conduct as a utile of satisfaction for anyone else.

To state the proposition in this form is, I think, to refute it, at least in the general case. The difficulty is that we have to organize our social institutions to take human beings as we find them, and not to assume that they will act, or can act, as we would like them to behave. The only regime in which to each, regardless of ability to pay, will work is one devoid of the self-interest that motivates most people, most of the time. Without that sense of disinterested benevolence, the risk is that people will behave worse where the level of medical care is guaranteed regardless of ability to pay than they would if the ethical imperative were ignored. The wholesale decline of production and initiative under socialism should not be regarded as a great surprise. It is the predictable consequence of a set of rules that falsely tells all persons at once that the amount that they get from the whole is independent of the amount of contribution that all people make toward the whole. If my share of the pie is fixed no matter what I do, then I will cut back on labor and save an enormous amount of personal grief, secure in the knowledge that 99 percent of the cost of my self-indulgence will be borne by others. Now the ethical imperative of minimum protection operates in an unintended and unwelcome fashion – as an upper bound on private efforts of advancement and self-improvement.

If I were the only person for whom that strategy of malingering were open, then the system could survive. But when all other persons see that the same opportunity for advancement is available to them, then that same decision is played out over and over again; and when the results are summed, all come out the worse. Each person of a group wins that game precisely one time, and loses it all the other times it is played. The one gain is very large, but it is surely smaller than the product of the small losses multiplied by the enormous number of losses that have to be borne. Thus, suppose that the game is played 100 times in a group of 100 persons. The one time that I win I gain 50 units of health care, and all other persons lose exactly one unit. The social losses appear to be 50 (100-50) but my private gains appear to be 49 (50-1). Yet the next 99 times the game is played, I sustain losses of 1 unit per round, for a total of 99, which means that my overall loss from malingering is 50. Oddly enough there is a perfectly equitable distribution of the poor outcome, for each person suffers the same fate: one glorious round of 49 units of profit and 99 rounds of 1 unit of loss. The systematic failure of socialism stems from the repeated occurrence of just this game.

Markets for all their weaknesses do not have the very important one I have just described, for markets create a world in which my prosperity depends on making someone else better off than they otherwise would have been. Each interaction that produces 1 unit of net good for me has to produce at least some net good for everyone else. The more rounds that the game is played, the more likely it is that everyone will prosper. The person who nets $1,000,000 in revenue is someone who, it can be said, has supplied at least $1,000,000 in benefits to other persons through the mechanism of exchange. There is no grand system-wide dilemma that works only when *A* pulls the levers but fails whenever anyone else does the same. All can imitate *A* and the overall net product will surely increase, even if it does so at an uneven rate. Left to their own devices, these markets will not fail if they are not regulated and clubbed into submission. Using the ability to pay to ration medical care therefore has important allocative effects in a world of scarcity – effects that are ignored by an ostensible ethical imperative which can work nowhere else and which will fail in medicine as well.

Disorder and Discontent

How can we see the signs of disorder? Let me give a couple of illustrations. As noted before, the ethical imperative is one that has to be funded by someone. Since, by definition, it is not funded from fees that patients pay for services rendered, it must be funded out of general tax revenues, and these revenues have to come from someone. No matter how the revenues are raised, they will involve taxes on other productive activities, and the taxes will in turn reduce the profitability of those activities, and hence the taxes that can be raised. There are, then, indirect effects of the ethical imperative that do not appear on the books of the medical expense, but which have to be regarded as costs of the program in question.

Next it is necessary to decide how the tax revenues will be spent. In these circumstances, the usual constraints of a system of demand are necessarily rejected from the system as a whole. It is not possible to turn away anyone because of an inability to pay. Given the usual constraint of self-interest, we should expect to see people demand medical care so long as it has positive value to them, that is, value above and beyond their own private costs for that care (e.g., lost time from work, costs of transportation, the risk of adverse medical outcomes, and the like). At the same time that there is an increase in demand, there is likely to be a transformation on the cost side of the business. Marginal cost curves do not always tend downward: in some cases it can become more and more expensive to provide the care in question. Even where wealth does not measure utility, a system that guarantees access regardless of the ability to pay is one that has to deal with the social losses that are normally caused by subsidies: costs that exceed benefits to the recipients, even though the benefits to the recipients still exceed the recipients' own private costs. The usual model for externalization of costs, and social losses through subsidy, does not disappear simply because our system of entitlements tries to handle the solvency question which everyone acknowledges exists. Instead the effort to effectively constrain one type of problem will normally exacerbate other difficulties which did not exist in the original market environment, but which take on strong significance in the altered environment. Subsidies can distort outcomes in one direction even as they may eliminate important imbalances in the other direction.

The question then arises: What systems can be used to handle the problem? One approach is to attack the demand side of the question, and the other is to attack the cost side. The factors that go into dealing with each side are enormously complex, and the standard responses to them are wholly inadequate, and often worse. What follows is only a short sampling of the types of errors that we have to face today.

1. Fee Restrictions

First, on the cost side, the rise of government payment systems has led to an insistence on the control of fees charged by physicians. Note that one hidden consequence of the current situation is that patients are far less willing to monitor charges when they are paid for by third-party providers. This is but another fiendish version of the socialist dilemma that I outlined earlier: what each person gains with respect to his own individual case, he more than loses when others take advantage of the overall situation. In fact, the basic situation here is even worse because there is no fundamental symmetry in the initial positions, so it is quite possible that some people will be able to take advantage of the system far more than others. In any event, the want of close patient supervision of what transpires means that the ultimate payor has to place its own restrictions on the fee structure. One of the current battles, in Medicare and especially Medicaid, has been frightening in its implications. There is no effective way to set the standards in question. If the fees are left too high, physicians flock to provide services that are not needed; if they are set too low, physicians will skimp on care, change specialities, or even abandon the profession. The ethical imperative that medical services be provided without regard to the ability to pay does not repeal the ordinary law of supply on the other side of the market: as payment goes down, the availability of services will shrink.

Or so I would have thought, but the judiciary in some cases has not quite gotten this message. To mention but one case of note and import, *Massachusetts Medical Society v. Dukakis*,[4] the question there was whether physicians could successfully challenge on constitutional grounds a statute that provided, as a condition for obtaining or renewing a license, that any physician who treated Medicare patients must "agree not to charge to or collect from such beneficiary any amount in excess of the reasonable charge for that service as determined by the United States Secretary of Health and Human Services."[5] The statute was upheld against the usual forlorn, rational basis, due process challenge.[6] What concerns me here, however, is not the nature of the legal rulings, but the cool reception given to the empirical evidence that the Society's expert offered to

---

[4] 637 F. Supp. 684 (D. Mass. 1986).

[5] *Id.* at 686. See MASS. GEN. L. Ch. 112, § 2 (1986) for the full text of the statute.

[6] *Dukakis*, 637 F. Supp. at 706-07.

support the proposition that the amount of services provided would be choked off by this statutory requirement.

The point of the expert's affidavit was that lower prices would lead to less access.[7] I should have thought that this was an axiomatic conclusion, but to the contrary Judge Robert Keeton, a former Harvard Law Professor who specialized in tort and insurance, regarded it as so outlandish as to require a public rebuke from the bench.[8] He could not conceive how any reasonable factfinder could find that reduced rates lead to reduced levels of care. Thus, when the Society's experts noted that the inability to resort to so-called balanced billing would crimp incentives to provide for costly high-quality care,[9] the judicial response was that there are other incentives for good care, most notably pride in work, curiosity, nonpecuniary satisfactions in the welfare of patients and the like.[10] I am sure that these inducements do play a role, but no one would say that they are sufficient to keep physicians in business at the limit, that is, without any cash compensation of any kind. So at some point before reaching that limit these financial incentives have to kick in. Do we assume that the response curve is flat for enormous regions of changes in wages, but then, when the straw breaks the camel's back, that all the supply of medical services precipitously disappears? Most people draw supply curves as continuous, and not with sharp precipices and edges, and for an industry I am sure that this view is closer to the truth than any other. The total compensation package to the physician does include all the intangibles, but if these remain constant when fees are reduced, then quantity of services supplied will be reduced. In reality, the situation is doubtless worse than this, given that the increase in external controls, whether by private contractors or government regulation, is likely to reduce the non-pecuniary satisfactions associated with the job.

Judge Keeton was largely oblivious to these difficulties. Instead of thinking that reduced rates would lead to a reduced supply of services,

---

[7] *Id.* at 695.

[8] *Id.* at 695-97.

[9] *Id.* at 697.

[10] *Id.* at 698.

he urged the contrary. There may be some effort in individual cases to increase the amount of work to offset the loss in revenue, but to not expect reductions in supply of services is to ignore the lessons of marginal revenue and marginal cost. The strategy of preserving income by working harder will fail dramatically if marginal costs are in excess of marginal revenues. An additional $10,000 in fees will not enhance physician wealth if it costs an extra $15,000 in time and expenses to produce. Instead of more services, there will be exit from the field, reduced entry into the field, political pressures to change the rate structure, unobservable diminutions in the quality of service, and the like. The supply side of the business is not impervious to price, and the very fact that the providers of health care are protesting cuts in their income should make the salience of financial incentives all too clear. To advocate full access to care without ability to pay, or even the more modest restrictions on the balanced billing system, is an open invitation to disaster.

There is, I think, no way to escape the dilemma with a single payor system, where price flexibility is so hard to introduce and maintain. What competition does is to allow individual groups to decide what mix of restraint on access should be blended with what level of monitoring from without, and what level of fees. I am sure that no one knows the ideal solution to the major dilemma with medical care – that is, how much risk should be left on an individual patient and how much should be covered by insurance – but I am confident that there are no government measures that will ever be able to get within a stone's throw of the correct answer.

2. COBRA Legislation

Let me mention only two other difficulties, and these go to the demand side for services. First, the COBRA[11] legislation requires, to be eligible for Medicare reimbursement, that all hospitals take in, without regard to ability to pay, persons who are either in active labor or who have a

---

[11] Consolidated Omnibus Budget Reconciliation Act of 1985, Pub. L. No. 99-272, 100 Stat. 82 (1986) (codified as amended in scattered sections of 15, 19, 20 & 42 U.S.C.).

serious medical condition.[12] The COBRA statute also imposes severe restrictions on a private hospital that seeks to "dump", that is, transfer its nonpaying patients to public institutions.[13] In effect we have here a situation in which medical care in an important class of cases has to be provided without ability to pay.

There are several major difficulties with the program. The first is that under COBRA the incidence of the public burden on hospitals is determined in large measure by happenstance, reputation, and geographical location. There are many hospitals that are not impacted by this legislation because they are located in affluent communities in which most people have extensive medical insurance. The Mayo Clinic is not likely to have much COBRA business in Rochester, Minnesota; but there are others, such as the University of Chicago and other inner-city hospitals, that receive a large number of cases of this sort, for which they receive little or no reimbursement from anyone. The net effect is that certain institutions run the risk of being saddled with expensive treatment for which they receive insufficient compensation, or indeed no compensation at all. The proper response, which will induce some caution in the amount of care to be provided, is one that requires the government to contract with private and state institutions for the levels of emergency care that it wants provided, and to pay for that care out of general tax revenues. That measure of financial prudence will force an off-budget decision back onto the budget, and will therefore alleviate many of the odd distributional problems. It will also lead the government to think hard about how much care it wants to provide and to whom.

So this brings me to the issue of individual demand for service, which I will introduce with a recent newspaper story about the homeless.[14] The article reports that the head of the homeless operations in New York resigned in frustration after revelations that an increase in the level of housing afforded to the homeless did not clear the market, but only

---

[12] 42 U.S.C. § 1395dd (1988 & Supp. 1989).

[13] *Id.*

[14] *See* Michael H. Cottman, *Advisor on Homeless Quits*, NEWSDAY, Sept. 5, 1991, at 33.

induced persons who lived in cramped quarters to declare themselves as homeless in order to obtain preferred accommodation paid for at taxpayer expense. It was a nice vacation, some reported, especially since the units to which they were assigned had a market rent in the neighborhood of $3,000 per month. In essence, the skeptics were correct. The higher the inducements for being homeless, the greater the number of people who would cross the line. The demand for services was not fixed independent of what was provided.

The same moral hazard problem exists with respect to health care. If one announces that free care is available for all who need it, but only if they are in distress, then there is a private incentive to be in distress. There is an incentive, for example, to avoid good prenatal care because one can go to an emergency room during active labor, without medical records, secure in the knowledge that someone else has to pay. The number of drug-related admissions through emergency rooms is similarly very high, so that a program that is designed to handle poverty ends up subsidizing addiction. The budgetary pressures are enormous, and the displacement of ordinary patients from hospitals can be large. The size of the medical problem thus increases because of the access that is now provided as a matter of right. It has been suggested that we can cure this problem by giving free medical care up and down the line, so that people will not have to wait for distress in order to come to hospitals. But the effort to close down one problem – administrative costs – would only create other problems because the total level of demand from all quarters at zero price is exceedingly high, so the funding problem would still exist. I suggest that at some level we have to "just say no", and be prepared to turn out of emergency rooms persons with, say, drug-related conditions. Efforts of benevolence after the fact have the same effect that they have in the homeless case: they induce more conduct that is in need of medical care. These efforts offer an implicit subsidy to the same activity for which we seek to impose punishment through the criminal system. The complex situation here exists no matter what we think to be the ethical imperative, which now involves the right to receive care not only regardless of the ability to pay, but regardless of the self-inflicted nature of the disease or injury. The moral hazard pressure for overutilization will not be solved by the mere adoption of a system of unlimited access.

3. Terminal Medical Care

My last illustration comes from a different quarter, terminal medical care, and it is here that I think that we can make the greatest progress. The current system has two positions. First, it will not tolerate voluntary euthanasia, a point which I think is a mistake, but will not speak about here. The second is that it will subsidize to the limit any decision to stay alive notwithstanding the quality of life. In consequence, there is an enormous amount of high-technology, high-sophistication medical care that is routinely supplied to patients who want, and sometimes do not want, such care. The net effect is to prolong life for a matter of days or weeks, or even months, but at enormous public expense. I do not know the exact percentage of medical care that is spent on the last ninety days of life, but I suspect that it is enormous. I think that where there is collective payment of individual expense, this is the most promising area to control expenses.

The basic procedure I advocate is this: require all persons to decide in advance the point at which they wish to abandon further medical care. There are today the Acute Physiology and Chronic Health Evaluation (APACHE)[15] tests which give pretty accurate indications of whether recovery is possible. These are cheap to administer and reliable, at least in extreme cases. What has to be done is to say that all public aid presumptively ceases at, say, APACHE 35, and that those who want additional care have to pay for it. This would be an automatic and fixed rule that limits the subsidy and fixes the point beyond which the ability to pay matters. I would give people all the information they want about the odds of success after that point, and I am confident that very few

---

[15] TABER'S CYCLOPEDIC MEDICAL DICTIONARY 123 (Clayton L. Thomas ed., 16th ed. 1989).

would pay the actuarial value, particularly after they have figured out that they are likely to have a longer life expectancy if, rather than buying this coverage, they instead find ways to spend the additional money to improve their nutrition, happiness, and health. The enormous amount of Intensive Care Unit (ICU) heroics will largely become a thing of the past, and the ICU will return to a more sensible function, which is to allow for massive intervention to control acute episodes where the long-term prognosis is good. The system proposed here is an effort to mix subsidy with public restraint while avoiding the problems of abuse that come with below-market pricing. I think that it could work.

Living With Imperfections

So what then is the message? Here I would put it in this fashion. The insistent demand for medical services often leads people to believe that rationing should not be done through the market. At some level, if only for charitable assistance, this is surely correct because of the weak correlation between utility and its proxy, wealth. But the asserted corollary does not follow. The usual rules of supply and demand, the risks of moral hazard, and the dangers of public subsidy cannot be ignored. The *predictive* side of economics, and its underlying cautious view of human nature, cannot be overlooked just because markets do not work without defects. Every form of regulation not only has its intended consequences, which are often unobtainable, but also its unintended consequences, many of which are undesirable. Any effort to plug up one imperfection will bring on others. We are at the dyke with *n* enormous holes and *n* minus *x* plugs, none of which quite fits. We can never plug all the holes. To devote all of our attention to obvious imperfections will lead us astray if we ignore the new problems we create in an effort to deal with the old. Some hard-headed realism must temper the ethical concerns over access to care that are so widely expressed. Our failures in this area are not a result of happenstance. They occur because the laws of economics are not repealed by the laws of the sovereign.