

## **CUTTING HEALTH REFORM DOWN TO SIZE**

Mark V. Pauly, Ph.D.

Health Care Systems Department

The Wharton School, University of Pennsylvania

April 30, 2007

The current state of the medical care system in the United States, and in all other countries, provides more evidence yet again that we do not live in the best of all possible worlds. Even for economists who use a very constricted definition of goodness, there are many deviations from the theoretical ideal. However, as all economists know from Coase onwards, showing that something falls short of the theoretical ideal is at best a useless and potentially a harmful guide to public policy (however much it may get the intellectual juices flowing), because inefficiency only truly exists if *we can* change things in ways that allow for mutual benefit. That is, change must be feasible. I have yet to see any proposed change in health care financing, planning, or organization in the United States that is very likely (much less guaranteed) to meet this goal. More concretely, in analyzing the market like arrangements that cover about half of all U.S. medical costs, one must compare present or possible market performance with alternatives that real world governments can and will accomplish; likewise, in critiquing current government programs of Medicare and Medicaid, one must use as a standard what actual markets, as competitive as we can get them to be, with irreducible ignorance in many (though not all)

parts, and representing a very uneven distribution of income, can accomplish, and not rely on the textbook slogans about the benefits of competition.

In a world where one must therefore choose among second bests, tastes and perhaps even emotions become a factor. Choice among policies in part depends on facts which we know, and in part depends on economic theory, but it also depends on the analysts' preferences, values, and fears. So I will begin with a catalog of "what bothers you the most," as represented in the two longer papers prepared for this meeting and the additional comments. Then I will compare the monsters under these beds with my own fears—not to persuade but to amplify. Having done that, I will briefly outline what I have conceived (since 1990) to be a pretty good though definitely not perfect alternative to the current U.S. system, show how it is possible to get there, and outline the unknowns that remain that may be understood along the way.

Some of what bothers most people about the U.S. medical care and health insurance system:

- 1) The level of expenditures is high relative to GDP and spending in other countries using the typical way of converting currencies.
- 2) While 85% of the population has formal insurance, 15% does not.
- 3) Risk varies, and we cannot conclude that this variation is handled well in either public or private sectors. Economists worry about adverse selection (though they can only find it for sure in government programs) and non-economists worry about risk rated premiums (which they are certain exists in private programs).

- 4) The quality of care could and should be higher.

Some of what bothers me about the U.S. medical care and health insurance system:

- 1) Real costs rising relative to GDP at a rate that is especially difficult for the public sector to sustain, and irritating for the private sector.
- 2) Quality levels for the highest quartile or quintile of the population by income that are inefficiently high in many (though by no means all) dimensions, because of tax subsidies that favor this subpopulation.
- 3) The absence of adequate protection against variation in risk over time in the predominant employment based insurance system.
- 4) While 85% of the population has formal insurance, 15% does not, and yet many of the uninsured could afford coverage and many of the insured are making unfair sacrifices to obtain and retain coverage.
- 5) In the face of strong evidence for variation in both the health care and the type of health insurance that people demand in the U.S., the ideological drive for equity in health (as opposed to the more logical equity in welfare) has populist appeal.

There is no mystery about the reasons for the differences in U.S. medical care spending compared to other countries, and no obvious need to do anything about it. Virtually all of the differences in spending per capita or relative to GDP arise from the higher payments to medical care labor and capital in the U.S. compared to other countries. It is, as Gerry Anderson put it, “prices, stupid,” although he could have said “wages, stupid.” After adjusting for the differences in the prices of inputs, the level of real medical spending in

the U.S. is at about the median of developed countries. The higher wages and prices relative to other countries are in part the result of less aggressive government efforts to redistribute income to consumers of medical care and away from health care specific labor and capital and, to a lesser extent, at a decentralized system for pharmaceuticals which (except for pediatric vaccines) permits higher profit rates; the differences in wages are probably also due in part to higher real output per (educated) worker in the U.S. and the open labor market which requires medical firms to pay wages high enough to attract and retain workers. In either case, there do not seem to be compelling equity or efficiency reasons to change these relative prices. My own judgment is that there should be a ban on using the GDP share to compare medical care systems across countries.

But while we might accept the *level* of U.S. medical care spending with aplomb, since we largely pay it to ourselves, the long term *growth* trend of rising real costs is of more concern (even though the U.S. is not too far out of line in terms of real growth rates). One reason is the iron law of arithmetic: no category of spending can grow more rapidly than real GDP forever (that is, no good can be a permanent luxury good). This means that we know at some point there will be a change, but we do not have any idea what the form of that change might be and, more importantly, we do not know how to change the system to permit the change to be as graceful as possible.

Even before reaching the inevitable but unknown period of transition from a luxury good to a plain old normal good, medical care real spending growth causes problems for the subpopulations. Most obviously, the Medicare population has so far experienced a

growth rate close to (though perhaps a little less than) overall growth, and a continuation of this practice implies tax rates to finance Medicare benefits two or three times as large as currently; combined with Social Security taxes the net burden on younger workers to pay for programs for seniors is very likely to give us a real world definition of “unsustainable.” With sufficient political will any transfer can be engineered, at least up to the point at which tax rates become so high as to reduce revenues, but the political will to bear the taxes implied by extrapolating Medicare’s growth into the 21<sup>st</sup> century may be absent.

The other fear is that the rising spending accompanied by improvements in quality that satisfies the demands of better-off Americans may be making any health insurance unattractive to lower middle income people, because the market is incapable of offering a variety of levels of quality and amenity in health care, including a lower cost if lower quality product. In part, the difficulty may be the obvious one of the political and legal incorrectness of offering lower quality health care to people who receive the lower incomes society is willing to tolerate. In part there may be some significant impediments to the kind of product differentiation and market segmentation that usually and appropriately occurs for most products and services.

A potential solution strategy for both problems would indeed be the emergence of health plans distinguished by their explicit choice of a variety of alternative policies toward higher cost and higher quality new technologies. Plans might select various thresholds for the cost effectiveness ratio they require new technologies to meet, make systematic

decisions based on those ratios, and clearly inform consumers about the rationing rules that will be applied to care for people who voluntarily choose those plans. Of course, one would expect (and I would hope) that there would be subsidies to the choices of lower income people, based largely on altruistic externalities. But for the middle class and above, the most puzzling aspect of medical insurance markets in the U.S. has been the failure of differentiated plans to emerge, especially those differentiated by their policies toward the adoption of new technology.

Risk variation is worth some consideration, but in my judgment gets much more discussion and more attention than it deserves. One reason why there is so much talk but little effective reform is that economists as economists have different concerns about risk variation than do ordinary people. In economic theory (without supposed equity considerations), the ideal insurance market would involve risk rated premiums. Where insurers cannot observe risk well enough to risk rate premiums (fairly rare) and where they are prevented by law or custom from doing so (fairly common), the equilibrium if one exists will involve harm to *good* risks because they will choose limited insurance coverage to avoid making transfers to bad risks. Non-economists (and economists giving policy advice), in contrast, often think that anything that inhibits such transfers is harmful, based on poorly defined concepts such as “fair” or “affordable” to higher risks, no matter what their wealth.

In the most abstract theory, the ideal economic way to deal with risk variation would be to allow and encourage markets in which insurers risk rate premiums as perfectly as

possible, and then make ex post transfers to those bad risks whose incomes are so low that their after-insurance-premiums income is thought to be too low. In reality, there are two feasible but somewhat imperfect ways of dealing with risk variation. One strategy, always imperfect but still worth doing, is to risk adjust any public subsidies. The other strategy is to encourage and support insurance policy provisions guaranteeing renewability of private coverage at class-average premiums—in plain English, to have insurers promise that they will not “re-underwrite” and therefore will protect those who did the right thing—bought insurance when they were healthy—from “reclassification risk.” Both of these policies are in theory equally effective, and it is incorrect to anoint one or the other as the only “solution.” The case for either depends in large part on the broader social context in which insurance is financed. I will elaborate more later, but the general idea is that for those populations for which public subsidies to insurance cover almost all of the cost of coverage (which in the U.S. would be those on Medicare or Medicaid), then a good arrangement is one in which the public contribution or de facto voucher made available to finance the individual’s choice would be risk adjusted as would the premiums insurers would collect. In contrast, for those populations for which subsidies are relatively small, ensuring that the policies contain guaranteed renewability provisions will be the best solution.

Because guaranteed renewability is less well understood in theory and practice (although it is a very common provision even in unregulated insurance markets), some additional explanation is in order. In effect, an insurance that includes guaranteed renewability charges a two-part premium to good risks in any time period: one part of the premium

covers the cost of unexpected large expenses in that period, while the other part covers the present discounted value of the cost of protection against “classification risk”—the risk that future premiums will be greater than good-risk premium because of the onset of a chronic condition in the time period in question. Under such an arrangement, the premium charged in every period is a premium good risks are willing to pay. In effect, the transfer to those who are higher risks in a given period does not come from those who are lower risks in that period; instead, it represents funds already collected from good risks in previous periods. Thus there is no way for an outside insurer to profitably draw off the good risks (though it would not harm the bad risks if it did, since the money the bad risks need to stabilize their premiums has already been collected). There is a larger potential problem (as in the case of all long term contracts) of “chiseling” from the current insurer, but reputation effects seem generally sufficient to prevent this problem from being severe.

While neither perfect risk adjustment nor bullet-proof guaranteed renewability are possible, each may be able to work reasonably well. In the U.S. the empirical evidence on risk adjustment of public subsidies mostly comes from the Medicare program (which provides about 90% of the premium cost for seniors). We know that the very crude risk adjustment mechanisms that were used until recently did not work well, but the more recent policies seem to work better. Guaranteed renewability for insureds is only available in the individual insurance market in the U.S. (there is no such protection for group insurance), and it has led to patterns of lifetime premiums consistent with the “front-loading” suggested by the theory as well as a relationship of risk to both premiums

and coverage that cannot be consistent with period-by-period risk rating but is consistent with guaranteed renewability. In round numbers, it appears that this feature results in pooling of about 85% of identifiable chronic-condition risk (given policyholder age), so that premiums and the acquisition of voluntary private coverage are nearly unrelated to risk. In fact, controlling for income, older persons who carry higher-than-average risk due to age are more likely to have coverage than lower risk younger persons.

I am in favor of structuring insurance markets so that consumers have the choice of a wide range of different insurance policies, defined both by the policy provisions and forms and by supply by competing firms. The ideal arrangement in my view is one in which plans produced by investor-owned private firms, by nonprofit private firms, and by some governmental are all offered and are all eligible for spending the subsidy if any (though in equilibrium one or more of these categories may disappear). However, the primary conceptual reason for offering options is *not* to ensure either lower cost and higher productive efficiency, or higher quality, than would prevail under a monopoly public system. While we do have some theoretical reasons to expect that competing for-profit private firms will minimize costs given quality, we do not have an equally well-developed theory of governmental or nonprofit firms that unequivocally predicts higher costs for them. And we do not have any theory to predict that either competitive firms or monopoly government will offer optimal quality. Making the appropriate adjustments for real world deviations from perfect competition and perfect government plunges us into even more ambiguity about productive efficiency or quality.

What I think can be argued is that government is intrinsically inferior to markets in another task: offering products that meet varying consumer demands, as augmented in some cases by public subsidies. Governments can be expected to provide uniform-across-individuals levels of quality and rationing mechanisms. If everyone wants more or less the same thing when it comes to how health insurance provides financial protection and how it rations care, this uniformity may not be a serious defect (and it may even help to lower administrative costs). But when (subsidized, where appropriate) citizen demands vary, markets make much more sense.

There is strong evidence for these arguments in the history of the U.S. Medicare program. That program initially provided uniform nominal coverage. In response, more than 80% of beneficiaries have arranged some kind of alternative more to their tastes—either supplemental coverage or now a private insurance alternative. Some of the variation in demand is due to variation in incomes which society tolerates, but much appears to be due to variation in tastes given income. (Research that documented how taste varies would be highly desirable.)

My final comment concerns the level of quality of care and of insurance. We have enough research to nail down some extremes. We know that one dimension of quality that consumers value is the ability to choose from a range of providers, and we know that this option is most common in the private group (especially large group) insurance obtained by higher income households. Lower income households that do have insurance are less likely to have choice. While the availability of options would doubtless be

correlated with income no matter what, there are strong reasons to believe that the large tax subsidies offered to upper income workers in the U.S. fuels this demand for high quality, and probably imposes some spillover costs on others.

The impact of those subsidies on other dimensions of quality is less easy to document (in large part because of problems of measuring quality). The proportion of national health expenditures paid out of pocket in the U.S. is now about 13% (lower than in many Western European countries) despite the presence of the uninsured, implying that by this measure coverage is very generous to those who have it. Medical care quality is much discussed (and seems to vary about as much as the weather), and is surely higher for the insured than for the uninsured, but beyond that seems to vary for economically unknown reasons and cannot be attributed to exogenous economic factors (and the local availability of specialists or specialized hospital beds is not an exogenous factor). Quality in terms of access to costly but effective new technology also seems roughly related to income but not well understood in general. More to the point, to my knowledge, there is no evidence that any purported inefficiency in terms of quality *can feasibly* be altered by policy interventions now available. I personally would not bet much on the discovery of any effective way to reduce cross sectional variations in measured quality, and would bet nothing at all on the proposition that improving quality would lower the rate of growth of medical spending. But that is no reason not to try.

In summary, the main problems I see in the U.S. health care system are underconsumption by the uninsured minority and overconsumption by the excessively

subsidized majority. For the middle class, the main economic puzzle is the absence of insurance options that offer real choice between cost and quality both in terms of levels and (especially) rates of growth. Even here, I would not guarantee that limiting the tax subsidy or fostering more insurance options would be guaranteed to reduce spending growth or change quality relative to the status quo. If we can restructure incentives better, I am inclined to accept whatever the outcome (no matter how messy) as what we want in an imperfect world. Nevertheless, I do think there are some features of insurance design that might be contemplated as representing improvements, and I now describe some of those.

In my ideal world, Americans would have health insurance that provides sufficient financial protection against unexpected medical care costs and that enables them to consume medical care up to at least the point where the marginal benefit from additional spending, to themselves and others, is equal to the real resource cost of that care. The desire for financial protection may induce additional spending beyond this point; the ideal level of coverage then reflects a tradeoff between increased risk protection and increased moral hazard.

Determining, even approximately, what this optimal point will be is difficult; the main difficulty arises because one person's use of effective medical care, up to a point, may be valued by the person's fellow citizens (either because of contagious disease or altruistic externalities); the "community demand" gets expressed through the political process, and

we know that this process can be imperfect (as well as frustrating and time-consuming) precisely because it requires a collective choice.

If externalities are greater at lower levels of use, and lower levels of use are associated with lower incomes, the ideal subsidy program will offer subsidies that are inversely related to income and other demand shifters, and that diminishes (but does not eliminate) income-related differences in medical care consumption.

How we finance this ideal insurance package depends on economic constraints and equity objectives. Economic constraints arise if mandation of insurance purchase is absent or difficult to enforce. It will be easier to get people to take the optimal insurance if they receive a subsidy that makes the socially optimal level their most desired level as well. Once this is achieved, equity further requires that the net distributional effect of the complete panoply of subsidy programs and taxes be adjusted to produce the social welfare optimum. Beyond what is attributable to externalities, no special consideration needs to be given to equalization of medical care use, health outcomes, or the use of or outcomes from other types of consumption.

This correct but arid theory says nothing about the organizational mechanisms for supplying medical insurance and medical care. If we should have the perfect government that can determine the optimum, this perfect government could then optimally determine the way in which insurance and care would be produced. Theoretically perfect government will tie or beat any market arrangement. Yet common sense tells us that

realistic governments may not get things right. So if market or market-like managements can do reasonably well in terms of efficiency, such arrangements may be permitted and even fostered.

To say more about what kinds of market arrangements might be a good idea, we need to add some content to the optimal coverage-optimal subsidy idea. A simple construct divides people or households based on two parameters—their (lifetime) income broadly defined, and the initial medical risk level. I am going to assume that the deviation between what individuals would choose in terms of insurance and care between the individual optimum and the social optimum falls as income rises. Roughly speaking, the kind of insurance upper-middle income people would demand on their own will lead to use levels for their own care reasonably close to the social optimum. As a person at good risk with a moderately high income, I do not worry about Bill Gates' use of medical care, and he does not (as far as I know) worry about mine. In contrast, insurance coverage and medical care use will be of greater social concern even among currently good risks who are low income. When risk varies, the extent of social concern presumably extends further up to income distribution for bad risks compared to good risks.

In this framework, there are then two separable groups: a set for whom optimal subsidies for insurance are likely to be large, and a set for whom they are small or minimal, perhaps limited to subsidies or programs for rare financially catastrophic illnesses.

The U.S. Medicare and Medicaid programs (one based on risk and one based on income) represent the “nearly full subsidy” programs. The tax subsidy to private employment-based insurance in the U.S. represents an example of the partial subsidy approach. I hasten to add that in neither case is there at present appropriate design, but only offer these as real world examples. Most other countries extend the full subsidy model to a larger fraction of the population than is done in the United States, but there are some southern hemisphere countries (Chile, Australia, South Africa, and Singapore) which also use the “income-conditional” small subsidy approach.

There is no bright line that distinguishes the two groups. In fact, the largest single group among the uninsured in the United States—lower middle income, usually young, people in good health—fall right on the likely dividing line, and thus are a flash point in the ideological struggle that unavoidably comes along with collective choice.

The market-like arrangement I would envision for the large subsidy population is a voucher-like arrangement. In some fashion (discussed in more detail below), there is a collective choice on the configuration of insurance coverage and the amount of the subsidy. Even within this group the size of the subsidy and the terms of coverage may vary with income, as it does in U.S. Medicare with higher Part B beneficiary premiums for the well-off and greater Part D generosity of coverage for lower income people.

I agree with [van de Ven](#) that the value of this voucher should vary as well with risk (expected expense under given nominal coverage). Risk adjustment will never be

perfect, but it should be as good as is cost effective. There will, as Newhouse has told us, inevitably be some tradeoff at the margin between efficient incentives for plans and ex post uniformity for people with different ex ante risk levels. The information and regulatory burden in the public sector can potentially be lowered by several design features:

- 1) Have a publicly provided (in Musgrave's sense) or produced option as one of the choice options.
- 2) Set a default option for every population that involves at least catastrophic coverage.
- 3) Set the terms for required or subsidized coverage very loosely, in actuarial rather than service benefit terms, so that people will be most strongly motivated to choose insurance they like.
- 4) Have both the subsidy and the actuarial value of coverage decline as income/wealth rise.
- 5) Make coverage comprehensive and free for poor people and very high risks.
- 6) Finance any subsidies with the least distortive tax possible, taking into account the excess burden of conventional forms of taxation.
- 7) Remove or cap the tax subsidy to employment based health insurance

These papers and comments contain much more that is noteworthy. The two main populations in the US which need work are, in my opinion, those with incomes below the median and those with incomes in the upper quintile of the distribution. The former need help and the latter get too much help. How to engineer the political process to bring both

equity and efficiency into being is the most puzzling question, but hope here always can be obtained cheaply.