Commentary

Managed Competition: Little Cost Containment Without Budget Limits
by Henry J. Aaron and William B. Schwartz

Confronted with a false choice between the two techniques most commonly advanced for controlling the growth of health care spending—managed competition and a global budget cap—President Bill Clinton has responded as Winnie the Pooh did when asked whether he wanted honey or sweetened condensed milk on his bread. “Both,” said Winnie and Clinton. This response is fortunate, because, as we argue here, growth of health care costs cannot be slowed significantly without global budget limits. Neither economic theory nor less formal observation supports contentions that competitive mechanisms alone will generate the right level or mix of health care spending. However, budget limits carry the risk of bureaucratic rigidity. The right mix of these two approaches to cost control is still unknown and is likely to vary within the United States depending on local conditions. Therefore, the federal government should set targets for health care spending but leave considerable latitude to the states on how to meet those targets.

A Brief History Of Cost Control

The history of cost control is one of a powerful inflationary force meeting little resistance. The future holds both the promise and the threat that inflationary forces will grow in strength.

The nearly irresistible force. Rising medical costs are a problem because of the pervasiveness of third-party payment. These payments drive a large wedge between social costs and private costs at the time care is rendered. This distortion of incentives affects both parties primarily responsible for purchasing medical care: patients and physicians. Because insured patients bear little cost at the time of care, they have incentives to consume services that produce any amount of benefit, no matter how small. Responsible physicians acting as agents for their patients validate these prefer-

Henry Aaron is director of economic studies at The Brookings Institution in Washington, D.C. William Schwartz is a professor of medicine at the University of Southern California and the Pacific Center for Health Policy and Ethics.
ences. Society must bear the costs of such care other than the small fraction that patients pay.

These incentives can explain excessive, but not rising, medical outlays. Some observers have attributed a sizable part of rising costs to such forces as the extension of insurance coverage, increasing administrative complexity, defensive medicine induced by fear of malpractice litigation, or growth in relative compensation of providers. In our opinion, they are simply wrong. These forces have contributed to a needlessly high level of medical expenditure but are responsible for only a small fraction of the growth of total spending. The explanation for rising outlays and for the increasing amount of low-benefit, high-cost care is to be found in the extraordinarily rapid pace of technological advance in medical care.

The very movable object. In the face of these rising outlays, public policy reacted first with lassitude, then with ineffectuality, and most recently with marginal effectiveness. At first, rising medical costs mostly troubled those who were uninsured. Rapid growth of worker productivity covered employers’ health costs and supported rapidly rising monetary wages as well. Governments easily paid for health benefits and still cut taxes periodically to offset revenue growth from unindexed tax systems.

This sense of limitless possibility gradually faded in the 1970s as economic growth slowed. It expired abruptly in the early 1980s after federal taxes, were indexed. Health care outlays had become major cost items on government and business budgets. In the mid-1980s government began to shift costs to private payers by increasing diagnosis-related group (DRG) payments much less rapidly than actual costs rose (cost shifting under Medicaid had long been the norm). With growth of labor productivity down and the size of employers’ health care budgets up, rising health costs ate up 58 percent of additions to compensation from 1980 through 1991. The private sector relied on managed care. Most elements of managed care merely shifted costs. The entire effort increased administrative costs. The rapid spread of managed care suggests that private benefits outweigh private costs; whether social benefits exceed social costs remains much in doubt.

Policies by business and government may have retarded growth of health care costs. During the mid-1980s the annual increase in per capita hospital spending fell from 6 percent to about 2 percent, cumulating to a nearly 30 percent savings in hospital days. These trends seemed to partially offset the cost-increasing effects of demography, input prices, and new technology. Nevertheless, the overall rate of growth of acute care spending during the 1980s closely resembled that of the preceding decade, perhaps because the “successful” cost-control efforts of the 1980s merely sufficed to hold spending growth to what it had been in earlier periods when such controls were not in effect, or perhaps because cost savings in the hospital sector were
offset by cost increases elsewhere. Either theory suggests rapid future growth of costs under current policies.

**More of the same is coming.** The pace of medical technology advance shows no signs of abating and many signs of accelerating. These advances almost always add to costs, by providing new treatments or new ways to develop medically useful information. Typically, innovations create procedures that are more costly than are the methods they replace. Where less costly per patient, they so vastly increase the number of cases in which the procedure can be used (for example, because they are less risky or painful or invasive) that total expenditures increase. Innovations may eventually curb growth of health care costs, but currently “cost-reducing medical advance” is almost an oxymoron.

We are persuaded that this flow of new technology will more than offset any savings that may be realized from squeezing out entirely wasteful procedures, streamlining administration, eliminating redundant facilities, or curtailing growth of compensation of health care providers. The result, we believe, will be an ever-lengthening menu of ever more costly yet beneficial diagnostic and therapeutic procedures that insured patients will want and their physicians will have powerful incentives to provide.

---

**The Oversimplified Economics Of Cost Control**

Most U.S. residents are covered by high-quality insurance. Since fully half of all health dollars are spent on patients for whom spending averages $30,000 per year, more than half of all care is almost certain to be rendered to patients who have moved past stop-loss limits and are 100 percent insured at the margin. Patients in such situations have no incentive to reject any medical intervention, no matter how small the benefit or how costly the procedure.

This problem can be illustrated graphically with the use of “benefits curves,” which show benefit per dollar spent on a particular medical intervention, arraying dollars spent from highest benefits to lowest benefits. The shapes of benefits curves vary from service to service. Broadly speaking, if all patients were fully insured, spending on health care would be excessive because patients who face no costs have no economic incentive to reject any form of care that generates even minor benefits, no matter how large the total cost. While excessive, health care spending could still be “efficient” in the odd sense used by economists; if all beneficial care were being rendered, no reallocation of spending within the medical sector could improve outcomes.

To control costs, policymakers need to introduce a mechanism that achieves two objectives: saving money and assuring that remaining re-
sources are used in a way that maximizes medical benefits. To achieve savings, spending must be held below the level that fully insured patients (and physicians acting on their behalf) would choose. To achieve savings efficiently, the ratio of expected marginal (social) benefits from consumption of health care to expected marginal (social) costs must be kept equal across all health resources. Reduction in outlays below some baseline is the measure of success of cost control. But egregious inefficiency will discredit the cost-control effort because patients who could have benefited from care are denied services while other patients who benefit little receive care.

What Managed Competition Cannot Do

Plans alleged to embody managed competition are so diverse that, in our view, the term has lost value in organizing public discussion. For example, the plan put forward by Alain Enthoven and Richard Kronick explicitly and emphatically rejects budget limits. In contrast, a bedrock element of the plan offered by California Insurance Commissioner John Garamendi is a cap on tax-favored premiums that individuals or businesses could supplement, from after-tax dollars, up to a maximum limit. (Garamendi’s plan is not presented in this volume, although it is discussed by several of the contributors.) This cap is, in effect, a global budget limit enforced through premium limits. This cap, in our view, is the defining characteristic of the Garamendi plan, a feature that is wholly absent from the Enthoven/Kronick plan and, indeed, is explicitly rejected by them.

Lack of detail on how managed competition would be implemented is alarming. So far as lack of specificity is concerned, the various managed competition plans are not inferior to those incorporating explicit spending limits, which are notably vague on just how the limits would be implemented. Global budgets differ from managed competition, however, in that global budgets and spending targets actually exist in various forms in several other countries where they actually have worked to suppress growth of health care spending. Managed competition exists nowhere. The benefits its advocates assert would follow from its implementation are expressions of faith, not experience. Plans such as the Federal Employees Health Benefits Plan that are sometimes cited as suggestive of life under managed competition give one little cause for optimism, as their costs have risen nearly as fast as those in society at large. Per capita health care costs in California, the state with the largest penetration of health maintenance organizations (HMGs) in the continental United States, remain among the highest in the nation.

The key to significant cost control from managed competition hinges on the belief that people will not buy health insurance more costly than that
provided through health insurance purchasing cooperatives (HIPCs) if they have to pay after-tax dollars to do so. The key to efficient cost control is the belief that HIPCs will have the incentives and the appropriate levers to curtail spending in a medically efficient manner.

The little evidence that is available on the first issue suggests that people will slightly curtail purchase of insurance if marginal outlays must come from after-tax income. The price elasticity of demand for health insurance seems to be much below one; we use 0.4 as a point estimate. The price elasticity seems to decline with income, so that making insurance taxable will be more effective in discouraging the purchase of supplemental insurance by those with modest incomes than by the well-to-do. The majority of families are subject to marginal tax rates that amount to not more than about 30 percent. Thus, people might curtail insurance purchases on the average by roughly 12 percent, with smaller percentage reductions by the wealthy and larger reductions by the poor.

What form might such curtailments take? Almost certainly people would not forgo coverage for catastrophic illnesses; if they did, such behavior would be regarded as a social problem to be solved. Some might join HMOs. Others might remain under indemnity insurance of one kind or another and accept higher cost sharing. Although higher cost sharing would assuredly reduce the overall cost of insurance, most of any curtailment in insurance outlays would take the form of cost shifting—from third-party payers to patients through out-of-pocket payments. The belief that insurance reform would cause most people to join tightly managed HMOs remains unsupported by empirical evidence. More importantly, evidence to date suggests that costs of HMOs are somewhat lower than, but grow at approximately the same rate as, those of other providers. Savings beyond a one-time reduction in costs will result only if HMOs ration services (HMOs claim not to be rationing now).

We do not believe that advocates of managed competition have come close to shouldering the burden of proof that their proposals would sufficiently curtail growth of medical costs. Our line of argument suggests that any national reform will have to rely, at least at the outset, on other instruments than managed competition to slow the growth of health care spending and that reform efforts should experiment with managed competition as a device to promote efficient resource allocation.

Controlling Growth Of Health Care Spending

Total health care expenditures (E) may be thought of in either of two ways—as a product of prices of services (P) and quantities of services (S), where $E = \Sigma P_i S_i$, or as the product of the cost of services per person (C)
multiplied by the number of people covered ($N_j$), so that $E = \sum C_j N_j$. This symbolic representation of health care spending suggests three ways by which health care spending might be controlled: limits on per capita premium payments; limits on prices paid for services, with quantity feedbacks; or direct ceilings on budgets of key providers. Here we examine each in turn, concluding that all three should be used in different parts of the health care system. We also argue that experiments with managed competition under such a system should be pursued vigorously.

**Premium caps.** Premium limits could be “hard” or “soft.” A hard premium cap would be a capitation payment that could not be supplemented. A soft cap could be supplemented by purchase of extra insurance, presumably on terms that are less favorable than those that apply to basic coverage.

The Garamendi plan is an example of a soft premium cap. Individuals would be required to have coverage under a basic plan, the cost of which would be largely covered by state taxes. Individuals could buy additional coverage out of after-tax income up to some limit. Purchase of insurance beyond that limit would not be allowed. Many arrangements would be possible, regarding the range of supplemental insurance, the types of benefits that it could cover, the penalties attached to purchase of such coverage, the ability of providers to render services under supplemental insurance, the extent to which providers could impose charges directly on patients, and many other aspects of coverage. Such rules would define just how hard a global budget ceiling enforced by premium limits actually is.

The Garamendi plan is based on a natural complementarity between the use of premium caps to impose a global budget and managed competition. While plan details remain obscure, the principle-setting a limit on resources available to the health care system and linking that limit to a system in which individuals or businesses could choose among health care financing vehicles and among providers in order to promote competition among insurance vendors and among health care providers-has considerable intuitive appeal.

Whether managed competition would promote efficient use of medical resources is open to question. Considerable doubt exists about whether major impediments to the implementation of managed competition could be circumvented. Would it be possible to develop a risk adjustment formula that would prevent insurers from trying to select among patients? Would it be possible to test such a formula to find out if it was “ready for prime time?” If HIPCs enjoy monopoly status, how will one assure that they behave in the public interest? If HIPCs approved only a few plans, would this create barriers to entry for innovative providers? Is there any conceivable circumstance under which a HIPC could deselect a plan that had achieved significant market share? Would HIPCs be subject to “regulatory
capture?” Would it matter if employers, whose liability will be limited to the cost of a basic plan, lose interest in guiding employee choice among plans?

More generally, will the incentives under which limited resources would be administered through conventional insurers selected by a HIPC result in efficient curtailment of medical care? Efficient curtailment requires cutbacks in most services, although by varying amounts. In some cases these cutbacks must be supported by investment and staffing decisions. In the case of medical services not dependent on specialized medical equipment or on providers with specialized skills, the cutbacks require collegial discussion among providers, subject to external review.

In principle, efficient use of medical resources requires that socially weighted benefits and costs be balanced at the margin across all uses of medical resources. In practice, precision in evaluating both benefits and costs is impossible. While HIPCs, insurers, or vendors of services may try to equalize benefits and costs at the margin, we think it more likely that they will engage in the same marketing strategies that unmanaged insurers currently use—trying to select low-cost customers or to entice customers by offering appealing services that may provide few medical benefits.

HIPCs or some other public agency will try to curb such behavior, no doubt with imperfect success. In any event, it is clear that the administrative and transaction costs of operating managed competition will be large—not perhaps so large as those under the current system, but still formidable. These costs would have to be weighed against two kinds of possible efficiency gains. First, managed competition would create an environment favorable to change in a dynamic industry, as it would continually put existing modes of rendering care at competitive risk. Second, managed competition might result in efficient rationing of medical care, although for the reasons presented above, we are doubtful.

In brief, it would be medically foolhardy, in our view, to base an entire national reform on such untried principles. It would be highly desirable, however, for a state or part of a state to put such a plan in operation in combination with premium limits. The Garamendi plan is an intriguing model for experimentation; it is not yet a model for national reform.

**Price controls.** Universal price regulation similar to that embodied in the Medicare system is an alternative approach to cost control. All payers, rather than just the Health Care Financing Administration (HCFA), would base payments to hospitals on a schedule of fees related to diagnoses. The fees produce a case-mix-adjusted hospital budget with a built-in proportional adjustment for admissions. A universal fee schedule for physicians would embody a similar principle.

By themselves, price limits of this kind contain only weak incentives to reduce use. They contain no incentives to promote efficient use of medical
services. The hospital admission or the outpatient treatment of a patient for whom diagnosis or therapy promises large benefits precipitates a payment identical to that triggered by another patient with the same diagnosis who stands to gain little from diagnostic intervention or therapy.

Price limits of this kind can serve as a budget constraint only if the adjustment factor provides for diminishing reimbursement as use increases. Such feedback mechanisms can be applied in various ways. Hospitals or physicians can be given declining payments per case as use rises beyond a certain target. In the extreme case, a price schedule with negligible adjustments for higher-than-planned occupancy rates is tantamount to a fixed budget ceiling. Alternatively, any adjustment could be deferred, as provided under the plans of some Canadian provinces or as has been proposed for Medicare’s resource-based relative value scale (RBRVS).

Price regulation is subject to numerous shortcomings. Indeed, critics of the DRG system expressed fears that hospitals would be tempted to seek out low-cost patients on whom they could make profits that could be used to cross-subsidize more costly patients. While experience under the DRG system so far has not confirmed these fears, this potential problem should not be wholly dismissed when considering whether to universalize price regulation. Despite its shortcomings, price regulation may prove useful for controlling physician expenditure in regions where population is sparse and meaningful managed competition is impracticable.

Imposed budgets. U.S. hospitals could be subjected to fixed budgets, as hospitals are in other countries. A central authority sets a budget for capital and current outlays and may impose some requirements for the distribution of expenditures. A governing body of the hospital determines how this budget is to be spent. The governing body, in principle, could consist of some combination of trustees, physicians, other providers, ethicists, and community representatives. Because hospitals and services related to hospitalization account for most acute medical care outlays in the United States and other countries, controlling these outlays would deal with the lion’s share of rising medical costs.

The chief strength of this approach is that it is the only method of controlling costs that sets up incentives for the efficient use of available resources. Confronted with a fixed budget, physicians who head the various departments within the hospital will be forced by peer pressure to assess systematically whether increased outlays on service $X$ will bring more benefits than the same amount spent on service $Y$. Of course, such judgments are complicated by uncertainties about efficacy and by the difficulty of comparing the value of such disparate outcomes as relief of pain, correction of a disability, and extension of life. The internal politics of the debate about allocation of fixed resources within a hospital inevitably will invoke
personal interests and personalities, particularly when incomes are at stake. Nevertheless, a preliminary assessment, which is far better than no assessment at all, is entirely feasible and would help greatly in assuring that resources are allocated to their highest-valued uses.

A top-down budget limit on hospitals also would have the virtue of imposing relatively modest administrative costs. Some appeals process would be necessary to deal with hospitals that felt shortchanged, but this process would impose only a small burden on the system.

However, budget caps do have important shortcomings. Most importantly, they can encourage a kind of bureaucratic ossification in the health care system. In the absence of competitive pressures, economic incentives to replace old and familiar practices with newer and more productive ones are weak. In addition, methods of controlling physician incomes that maintain incentives to work hard but not to induce demand for care are difficult to design. Rational methods for adjusting budget ceilings and reimbursement formulae for physicians that are not subject to manipulation have not yet been designed.

This list of shortcomings of budget caps, to which other deficiencies could be added, underscores that no easy or perfect method of controlling health care costs exists. We believe that for this reason, and others presented below, the United States should set performance standards for controlling growth of acute care costs but permit states to use different methods of meeting those standards.

A Strategy For Cost Control

Political and technical trade-offs will shape the choice among various strategies for imposing expenditure limits. Price limits are clearly unsatisfactory because they do not control volume except through a complex sliding scale for reimbursement. If one were starting with a blank slate, there would be strong arguments favoring a top-down budget cap. Administrative costs would be lower than with a system of managed competition. The hospital staff and administration, free of the administrative burdens and competitive pressures of a managed competition system, could focus their energies on assuring that resources were allocated to their highest-valued uses. The major drawback would be that bureaucratic inertia would delay the elimination of obsolete or low-valued activities. In the real world, however, it is highly unlikely that a top-down budget system will be implemented at one fell swoop. The current insurance system is not likely to be eliminated abruptly. Should the United States move to top-down budgeting, it is likely to do so slowly and incrementally.

Managed competition has the virtue that it would evolve from the
present health care system. It also preserves competitive forces that would encourage the ruthless elimination of unjustified activities. On the other hand, a large number of difficult problems would have to be addressed before the idea became a reality, including the regulation of HIPCs, the avoidance of risk selection, and the removal of substandard plans. Furthermore, moving to managed competition will be a complicated process because of the diversity within the present insurance system.

Let fifty-one flowers bloom. Unfortunately, the world is a cluttered place. For every California with a highly developed system of HMOs, there are several other states in which HMOs are curiosities. For every state consisting mainly of thickly settled metropolitan areas, there are thinly settled states in which solo practitioners are all that communities can support and often more than they can retain. For every community in which physicians spanning every imaginable specialty vie for patients, there are other communities in which primary care cannot readily be obtained except from emergency room staffs.

Because any practicable reform of health care financing will have to take the world as it is, not as reformers wish it were, it is vital that the reform be designed to accommodate a wide variety of practice arrangements and to permit states to employ a wide variety of cost-control mechanisms. We therefore urge that the federal government allow states wide latitude in choosing methods for meeting federal standards for cost control. California may elect to use a premium control mechanism linked to managed competition similar to that embodied in the Garamendi plan. We suspect that should such a plan become law, those charged with implementing it would promptly discover that although it may work well in San Francisco or Santa Barbara counties, it may work poorly in such thinly settled counties as Mono, Sierra, or Lassen, where it would be necessary to use price controls in dealing with physicians and, perhaps, fixed budgets in dealing with hospitals. Such states as New York or Maryland, rather than embarking on the entirely uncharted seas of managed competition, may choose to build on hospital budget controls they have developed over a decade or more. Other states may use still different methods. Instead of imposing a single approach on an extremely diverse system, Congress should recognize that all reform must begin from where the health care system is and simply insist that each state meet specified standards for limiting expenditure growth and assure financial access to all its citizens.

Timing is everything. All of this indicates that even if managed competition is the road Congress chooses to follow, many years will pass before a national system is in place. Enabling legislation will almost certainly not be passed until late 1993 at the earliest and, more likely, mid-1994. Implementation is likely to take an additional eighteen months to two years, and
considerably longer in regions now without HMOs or other tightly managed systems. Thus, even if managed competition is as effective as its most enthusiastic advocates anticipate, it cannot be expected to suppress health care spending before the incumbent president must stand for reelection. By that time health care costs are likely to have risen from the current level of about 14 percent of gross domestic product to, perhaps, 16 percent.

The budgetary imperatives confronting the Clinton administration may not permit such delay. Some reduction in projected federal spending on health care is likely to be essential to give the new administration a chance of significantly cutting the federal budget deficit; such savings are certainly indispensable if the deficit is to be halved. Yet, if efforts are directed solely to suppressing government spending, they are likely to exacerbate the already serious problem of cost shifting to private payers. For that reason, significant reductions in the growth of government health care spending (other than by curtailing benefits) can be achieved only if linked to similar controls on private spending.

We see but one way out of this dilemma between the need to achieve government (and, perforce, private) savings soon and the inescapable delays in implementing any comprehensive reform. Control of costs in the hospital sector alone, the largest single component of total acute care spending, might well be achieved quickly. While the administrative obstacles to such controls are formidable, the necessary framework for achieving significant reductions in costs is in place in a few states and could be extended nationally in less than one year. Congress could authorize states to use different techniques for meeting national targets for hospital spending. It would be essential to prohibit hospitals from sloughing off various services in order to comply with spending limits or to penalize these providers if they do.

A full-scale national reform would probably supersede these short-term measures, although this short-run approach to cost control might play a role in longer-run reform. Whether or not short-term measures survive as elements of permanent reform, they promise to slow growth of health costs more rapidly than comprehensive reform can do. These savings may help to pay for the extension of insurance within a deficit-reducing program.

An initial program focused on hospital cost control could promise to slow the growth of health care spending at least modestly. More stringent measures to deal with physician payment could be delayed. In a long-term reform, managed competition with its HIPCs is likely to have important uses in promoting efficient use of resources. It is still too early, however, to know how much this untried concept can slow the growth of spending. Let the experiments begin.
NOTES


2. This technological growth has been going on for four decades. In three of these decades annual growth in real health care spending exceeded 5 percent; in the other decade it exceeded 7 percent annually. See also W.B. Schwartz, “The Inevitable Failure of Current Cost Containment Strategies: Why They Can Provide Only Temporary Relief,” *Journal of the American Medical Association* 257 (1987): 220-224.


5. For illustrations of benefits curves, see Aaron and Schwartz, *The Painful Prescription*.


7. Even this cost shifting is speculative, as individual would be subject to blandishments from vendors of targeted Medigap insurance. In a paradoxical reversal of managed care, individuals who were concerned about denial of beneficial services might be tempted to buy second-opinion insurance, high-technology riders, and other coverage to fill in actual or perceived gaps. Such possibilities are admittedly speculative, but no more speculative than the claimed benefits for a system that remains a mental construct, and an incomplete one at that.

8. Most recent growth of HMO membership has occurred within individual practice associations (IPAs), a form of HMO that typically uses fee-for-service payment and that has shown no cost advantage over non-HMO providers.


10. These questions draw on a thoughtful review of how HIPCs would operate under the Garamendi plan. E. Wicks et al., “An Analysis of the Garamendi Plan for Health Reform in California,” (Institute for Health Policy Solutions, September 1932).