Taxation, Health Insurance, and Market Failure in the Medical Economy

Mark V. Pauly

Journal of Economic Literature, Vol. 24, No. 2 (Jun., 1986), 629-675.

Stable URL:
http://links.jstor.org/sici?sici=0022-0515%28198606%2924%3A2%3C629%3ATHIAMF%3E2.0.CO%3B2-4

Journal of Economic Literature is currently published by American Economic Association.

Journal of Economic Literature
Vol. XXIV (June 1986), pp. 629–675

Taxation, Health Insurance, and Market Failure in the Medical Economy

By MARK V. PAULY
University of Pennsylvania

I. *Introduction*

THE MEDICAL CARE INDUSTRY is often said to be in "crisis" (especially by politicians who would rather direct Medicare and Medicaid tax dollars elsewhere). To a considerable extent, this crisis reflects a feeling that, compared to other sectors of the economy, consumer-taxpayers are not receiving full value for the 10.8 percent of GNP that is poured into the medical care sector. One of the major themes of recent research in health economics is that an ultimate cause of this sector's performance is to be found far removed from the world of hospital beds and physicians' offices. The chain of causation instead begins with the tax subsidization of health insurance and medical care payments. It is this practice, rather than the peculiarities of medical services per se, or the attempts by physicians and other professionals to enlist government help in maintaining monopoly power, that has been the central focus of much recent research.

The major purpose of this article is to summarize the literature that develops (or opposes) the theme that tax subsidies to health insurance are a major cause of the behavior in the medical care industry, behavior that many feel to be inefficient.

First I relate the normative foundations of the tax subsidy question to the broader issues of efficiency and market failure in medical markets. Then I examine the empirical foundations for the positive and normative analysis of the tax subsidy. The link between tax treatment and the level (and possible efficiency) of medical care expenditures consists of three propositions:

1. The tax treatment of health insurance has warped the choice process for such insurance, resulting in the purchase of insurance that, at the market level, is both excessive in quantity and distorted in form.
2. Increases in the level and form of insurance coverage have the effect of distorting the demand for medical care.
3. This distorted demand has interacted with a set of supply-side responses, some conventional and some affected by the peculiarities of the industry, to lead to a level of expenditures and costs that is excessive in terms of level and rate of growth.

Behind these tenuous normative propositions lie a series of positive propositions. Some of the positive propositions too are a mixture of scraps of evidence and professions of faith, but some of them have in-

creasingly firm empirical foundations. These empirical propositions—the impact of taxation on insurance demand, the impact of demand determinants, especially insurance coverage and the associated money user price, on the quantity of medical care demanded, and the movements of market equilibrium, will be surveyed. Because a sizeable portion of all medical expenditures is covered by public (tax-financed and government-administered) insurance, it will be useful to develop some normative and positive aspects of the impacts of such public insurance.

In order to understand the relevance of what may appear to be a debate on an esoteric subject, it will be useful to review some information on the importance of uncertainty, and therefore of insurance, in the medical care sector. Medical care differs from conventional commodities in at least two ways. First, the demand for medical care depends in part (though not entirely) on a person's state of health, and that state is to some extent stochastic. Hence, insurance against the cost of this commodity might be expected to emerge, and has indeed done so, to such an extent that it is the predominant form of payment for most types of medical services. But second, insurance markets in medical care are themselves different from conventional insurance markets. Many of these differences trace back to other kinds of uncertainty. Not only is the level of the loss-probability for a particular individual sometimes unknown to the insurer, the actual loss in real income or well-being he experiences is also unknown. Ignorance about probabilities means that the insurer may have difficulties in determining which purchasers of insurance are at which levels of risk, so adverse selection may result. Ignorance about the actual loss in well-being suffered and/or the actual illness of the individual may force insurers to base benefits on the level of expenditures (which are proxies for the effects of illness but which are partly under the control of the insured), so moral hazard may result.

Both adverse selection and moral hazard caused by insurance may also have market-wide effects. Indeed, it is not too much of an exaggeration to read many diagnoses of the reason for "excessively" rising medical care costs as going back to a single cause—the nature, amount, and form of health insurance. Tax policy, as a primary influence on that insurance, and as a primary policy instrument for changing the quantity and form, therefore assumes enormous importance.

In addition to tracing out the tax-insurance links, I will address the question of whether there are some important correctable market imperfections *other than* tax-side distortions. I will review answers to the question of whether there are ways of improving welfare in addition to changing health insurance tax policy. In particular, two major candidates will be considered: dealing with adverse selection in health insurance markets, and correcting imperfections deriving from incomplete consumer information. Finally, I will briefly discuss the recent important changes in the way the Medicare program pays hospitals, and the probable effects of those changes on efficiency.

What is the meaning of market failure in the medical care context, and why has so much of the discussion focused on the role of insurance and hence of taxation as a determinant of insurance? One reason has doubtless to do with the external political or policy concern for the problems of the medical care sector. For many policymakers, the problem at least has a simple definition: The sector is a problem because the last two decades have witnessed rising relative prices for medical care services and rising shares of GNP devoted to medical care. At the policy level, no explicit criterion of optimality is invoked to show why such a shift in relative prices

should be a problem. But if the normative criterion is vague, it has nevertheless been sufficient to motivate studies of the determinants of this inflation, the implicit presumption being that once those determinants are found, it may be possible to correct their levels.

A second reason for focusing on insurance and taxation is more directly concerned with welfare economics concepts. If Pareto optimality is used as the criterion, a tax subsidy to a product for which there is no obvious reason to suspect underconsumption will be inefficient. Not only will the distorted incentives be inefficient, but, as compared with other possible imperfections in the market, they should be easy for public policy to correct. Although there has been no attempt to demonstrate rigorously that, in a second best world, removal of the health insurance tax subsidy would improve welfare, the simple intuitive welfare economics associated with inducing "greater cost consciousness" from participants in the medical care market, and the fact that the subsidy was perceived to redistribute income from the poor to the rich, combined to make it an attractive suggestion for economists to give when asked by policymakers to suggest how to "do something about runaway medical care cost." Arguments that the tax system contained a correctable defect were generally informal, though Pareto optimality could be read between the lines (Alain Enthoven, 1980, furnishes an example).

Much of the basis for the analysis of insurance, taxes, and medical markets can be found in Kenneth Arrow's classic 1963 article (although the tax treatment of health insurance is not mentioned in the article). The concept of moral hazard, further elaborated by Mark Pauly (1968), has been used to explain both the patterns of insurance coverage and the link between a taxation-induced increment in coverage and increases in the levels of medical expenses. The link is somewhat less direct, but strong antecedents of arguments concerning the cost-reducing advantages of prepaid group practices (later rechristened Health Maintenance Organizations), and some discussion of the demand for such types of insurance relative to the conventional sort, reappeared in economists' "procompetition" proposals (Enthoven 1980). The conditions of *supply* of health insurance, which involve some not-for-profit firms with special tax advantages of their own, the possibility of provider control, and some hints of market power, are also relevant to the efficiency debate. They were mentioned briefly by Arrow and will be considered in this paper.

It is, however, remarkable that the major thrust of Arrow's article did not stimulate much subsequent work. His major point was that many of the peculiar aspects of the medical care sector—not only the existence of insurance but its form, not only the licensure of providers but their professed deviation from profit-maximizing behavior—can be explained as public and private institutional substitutes for the absence of a competitive market insuring against all uncertainties. In other words, the market failed to offer individuals insurance against the occurrence of many illnesses. There has been little attempt to elaborate on this proposition, and little empirical testing. The implicit question of whether *new* or *additional* steps could be taken to repair the failure of markets to provide for risk has also been largely dormant although, as will be discussed below, some recent work on adverse selection and on the behavior of physicians as agents may be interpreted in that way.

One more introductory word about the relationship of the efficiency concept to the discussion of health insurance. The great bulk of medical care in the United States is purchased by people who are not poor. A majority of medical care costs are

paid for in the private sector by upper- and middle-income people, either via insurance or out-of-pocket payments. Virtually all of the tax subsidy goes to the nonpoor. The debate over reduction in the current tax subsidy has therefore not involved considerations of the welfare of low-income people. Instead, the notion that some explicit subsidy for the poor is desirable, because of philanthropic externalities or some other externalities (Pauly 1970; Anthony Culyer 1971; Cotton Lindsay 1969), has been accepted on all sides.

Such a subsidy for the poor can be designed in ways that do not concern the tax treatment of employer payments for health insurance or individual income tax deductions. However, the recent debate over the removal of the subsidy has caused some related equity and externality considerations to resurface, and these will be briefly discussed. In addition, the recent changes in how the Medicare program pays for hospital care for the elderly will also be reviewed because they affect the private market. But this paper will focus primarily on private sector markets for insurance and for medical care.

II. *Recent Trends in Health Insurance, Health Care Costs, and Taxation*

The majority of health care dollars comes from insurances, public or private. As shown in Table 1, private direct consumer payments for health care in 1983 were less than one-third of total spending. This low share occurs in part because of the public insurances—Medicare (primarily for the elderly) and Medicaid (primarily for the nonworking poor)—pay about two-fifths of all health care costs. It also occurs because inpatient and outpatient hospital expenses are heavily covered by insurances (both public and private), and because hospital expenses at $147.2 billion in 1983 amounted to 47 percent of all personal health care expenses.

This pattern of finance has been remarkably stable over the past decade, as shown in Table 2. Private payments as a percent of the total personal health care spending were 60.5 percent in 1975; they are virtually at the same level now. There has been a modest fall (from 32.5% in 1975 to 27.2% in 1983) in the fraction paid out of pocket. In the last two years, the out-of-pocket fraction has been stable, and may actually have increased, but the definitive data are not yet available. Of course, this stability follows a period of rapid change between the mid-1960s and the mid-1970s, as the out-of-pocket fraction fell when public share jumped dramatically with the passage of Medicare and Medicaid in 1965–66 and continued to trend upward for about five years thereafter. This period was also one of growth in the share of insurance in private sector expenditures.

But the most dramatic growth over the past two decades, as shown in Table 3, was in real health care spending per capita (adjusted for general inflation) and the share of GNP devoted to health care. This sharp upward trend was actually under way even before Medicare and Medicaid, accelerated somewhat in the late 1960s after those programs began, and was slowed only temporarily by the Nixon administration Economic Stabilization Program (ESP) in 1973–1974, and by the American Hospital Association Voluntary Effort (VE) program to limit hospital costs increases at the time when the Carter administration was proposing a federal hospital cost containment law.

There is one other dramatic change, but it is too recent to show up in national health expenditures data. There was a dramatic falloff in the real rate of growth in hospital expenditures (and in the services those expenditures represent) in 1984. Table 4 shows the recent path of real hospital expenditures, expenditures per admission, admissions, and average length of stay

TABLE 1: NATIONAL HEALTH EXPENDITURES, BY TYPE OF EXPENDITURE AND SOURCE OF FUNDS: CALENDAR YEAR 1983

Type of Expenditure	Total	Private					Public			Private Direct Consumer Payments	Public Payments
		Total	Consumer				Total	Federal	State and Local		
			Total	Direct	Insurance	Other					
	Amount in Billions of Dollars									Percent of Total	
1983 Total	**355.4**	**206.6**	**195.7**	**85.2**	**110.5**	**10.9**	**148.8**	**102.7**	**46.1**	**24.0**	**41.9**
Health services and supplies	340.1	199.8	195.7	85.2	110.5	4.1	140.3	96.8	43.5	25.1	41.3
Personal health care	313.3	188.8	185.2	85.2	100.0	3.7	124.5	93.0	31.5	27.2	39.7
Hospital care	147.2	68.8	67.3	11.1	56.2	1.5	78.4	60.6	17.8	7.5	53.3
Physicians' services	69.0	49.7	49.7	19.6	30.1	—	19.3	15.6	3.7	28.4	28.0
Dentists' services	21.8	21.2	21.2	13.9	7.4	—	.6	.3	.3	63.8	2.8
Other professional services	8.0	5.6	5.5	3.3	2.1	.1	2.5	1.9	.5	41.3	31.3
Drugs and medical sundries	23.7	21.6	21.6	18.4	3.2	—	2.1	1.1	1.1	77.6	8.9
Eyeglasses and appliances	6.2	5.2	5.2	4.5	.7	—	1.0	.9	.1	72.6	16.1
Nursing home care	28.8	14.9	14.7	14.4	.3	.2	14.0	8.1	5.9	50.0	48.6
Other health services	8.5	1.8	—	—	—	1.8	6.6	4.5	2.1	—	77.6
Program administration and net cost of insurance	15.6	10.9	10.5	—	10.5	.5	4.6	2.6	2.0	—	29.5
Government public health activities	11.2	—	—	—	—	—	11.2	1.2	10.0	—	100.0
Research and construction of medical facilities	15.5	6.8	—	—	—	6.8	8.4	5.9	2.6	—	54.9
Research	6.2	.4	—	—	—	.4	5.8	5.2	.6	—	93.5
Construction	9.1	6.5	—	—	—	6.5	2.6	.7	2.0	—	28.6

Source: Gibson et al. (1984).

TABLE 2

AGGREGATE AMOUNT AND PERCENT DISTRIBUTION OF PERSONAL HEALTH CARE EXPENDITURES, BY SOURCE OF FUNDS: SELECTED YEARS 1960–83

Year	Total Payments	Total Private Payments	Patient Direct Payments	All Third Parties: Private: Health Insurance	All Third Parties: Private: Other	All Third Parties: Public: Total	All Third Parties: Public: Federal	All Third Parties: Public: State and Local
	Amount in $ Billions		Percent Distribution					
1960	23.7	18.5	54.9	21.1	2.3	21.8	9.3	12.5
1965	35.9	28.0	51.6	24.2	2.2	22.0	10.1	11.9
1970	65.4	42.9	40.5	23.4	1.7	34.3	22.2	12.1
1971	72.2	46.6	38.9	23.8	1.8	35.5	23.2	12.3
1972	80.5	51.6	38.0	23.6	2.5	35.8	23.5	12.3
1973	89.0	56.9	37.4	24.0	2.5	36.1	23.7	12.4
1974	101.5	62.4	35.3	24.8	1.5	38.4	25.9	12.6
1975	117.1	70.8	32.5	26.7	1.3	39.5	26.8	12.7
1976	132.8	81.4	31.6	28.3	1.4	38.7	27.2	11.5
1977	149.1	91.4	31.1	28.8	1.3	38.8	27.5	11.3
1978	167.4	101.8	30.3	29.3	1.2	39.2	27.7	11.5
1979	189.6	115.0	29.4	30.0	1.2	39.3	28.1	11.2
1980	219.1	132.4	28.5	30.7	1.2	39.6	28.5	11.0
1981	253.4	152.6	27.9	31.1	1.2	39.8	29.3	10.4
1982	284.7	171.4	27.1	31.9	1.2	39.8	29.5	10.3
1983	313.3	188.9	27.2	31.9	1.2	39.7	29.7	10.1

Source: Gibson et al. (1984).

(days divided by admissions) from 1981 through the first quarter of 1985. Downward trends in the latter two indicators were apparent even before the new Medicare hospital Diagnosis Related Groups (DRGs) payment system was begun in October 1983, but there was a stronger decline thereafter. The figures for early 1985 contain a suggestion that this falloff has moderated. Finally, the level of expenditures per patient day has not shown a downward trend.

Table 4 also permits some comparison of recent trends in the Medicare program for people over 65 compared to trends for other hospital users. The same trends as in the aggregate data are apparent, including a 1984 post-DRG falloff for *both* sets of people, those who were subject to the Medicare program and those who were not.

The health insurance coverage pattern in Table 1 makes broad economic sense. Private direct consumer payments represent only 7.5 percent of hospital care expenditures. Only about one person in ten under age 65 enters the hospital as an inpatient each year, but the mean inpatient bill in 1983 was about $3,000 (American Hospital Association 1984). In contrast, ambulatory (mostly physician) services are consumed in some positive amount by more than 80 percent of the population each year, the mean annual expenditure of those who have expenditures is only about $445, and the fraction uninsured is much higher than for hospital care. A similar pattern of high-probability, small-loss

TABLE 3
PER CAPITA PERSONAL HEALTH CARE EXPENDITURES AND NATIONAL HEALTH CARE EXPENDITURES: SELECTED YEARS 1960–83

Year	Per Capita Expenditures	Real Per Capita Personal Health Care Expenditures*	National Health Care Expenditures
	Nominal Dollars	1983 Dollars	Percent of GNP
1960	129	434	5.3
1965	177	559	6.1
1970	305	783	7.6
1971	333	819	7.7
1972	368	876	7.9
1973	403	903	7.8
1974	456	921	8.1
1975	521	964	8.6
1976	586	1,026	8.8
1977	651	1,070	8.9
1978	724	1,106	8.8
1979	811	1,113	8.9
1980	927	1,121	9.4
1981	1,062	1,172	9.7
1982	1,181	1,219	10.5
1983	1,286	1,286	10.8

Source: Calculated from Gibson et al. (1984).
* Deflated by Consumer Price Index (excluding medical care).

experience characterizes dental care, prescription drugs, and most other medical expenses (except for nursing home care), which are also less heavily insured. In broad outline, therefore, insurance coverage is highest for the low-probability, high-loss event of hospitalization, and lowest for ambulatory care costs.

Let us consider the private insurance market for people under 65. (There is also an active market in private coverage to supplement Medicare, so 60 percent of the elderly had some private coverage in 1979 [calculated from Chollet 1984, Table 1.11.]) The bulk of private insurance for people under 65 is provided as employment-related group insurance. In 1980, individual (nongroup) insurance premiums were about 10 percent of total private insurance permiums for both commercial and Blue Cross insurers (Blue Cross/Blue Shield Association 1981; Ross Arnett III and Gordon Trapnell 1984); the rest was almost all group insurance provided at the workplace.

One can distinguish four broad sets of suppliers of such insurance: commercial insurance firms (both stockholder and mutual); the not-for-profit Blue Cross and Blue Shield plans, which are usually chartered under special state laws and have special tax and regulatory treatment; the Health Maintenance Organization (HMO) and other independent plans; and the self-insured plans. The pattern since 1960 is one of declines in the shares of the Blues and conventional commercial insurance, and growth in the self-insured and HMO-type plans. However, even though HMOs grew very rapidly in recent years (at a 20 percent annual rate in 1984), the fraction of the population currently enrolled in such plans is only about 7 percent of the total.

Employer contributions furnish most of the premium payments for group insurance. In 1980, 97 percent of employee participants covered by an employment-related insurance plan received employer contributions of more than half of the plan cost, and the employer paid the full cost for 84 percent of participants; for dependents, employers paid the full premium for 60 percent of participants (Chollet 1984, pp. 50–53). The employer's share has been rising over time as well. However, there is some question about whether the substantial jump in employer contributions over time can be attributed largely to tax policy (Chollet 1984).

While the great bulk of people under 65 have health insurance, and while the great bulk of that insurance comes through employment, there are some persons with no insurance coverage or with partial or part-year coverage. The most

TABLE 4
ANNUAL PERCENTAGE RATES OF CHANGE IN COMMUNITY HOSPITAL USE AND EXPENSES, 1981–85

	1981	1982	1983	1984	1985:1
Total expenses					
Current dollars	18.7	15.8	10.2	4.6	6.0
Constant dollars	8.4	9.9	7.3	0.5	1.8
Expenses per admission[a]					
Current dollars	17.4	15.5	10.2	7.5	14.0
Constant dollars	7.1	9.6	7.3	3.4	9.8
Total admissions					
Under age 65	0.0	−1.6	−2.8	−4.5	−10.4
Over age 65	3.0	4.1	4.7	−2.9	−1.2
Average length of stay					
Under age 65	−1.4	0.0	−1.8	−3.6	4.0
Over age 65	−3.7	−2.9	−4.5	−7.6	1.6

Source: American Hospital Association (1985, and earlier years).
[a] Adjusted for the volume of outpatient visits.

reliable survey of the uninsured was the National Medical Care Expenditure Survey, taken in 1977.[1] That survey found that, at any time during 1977, between 22 and 25 million persons were uninsured, of whom 18 million were uninsured for an entire year (Gail Wilensky and Daniel Walden 1981). It is believed that the percentage of the population without insurance is moderately higher now than the 1977 figure of 11 percent (Katherine Swartz 1984). Table 5 indicates the characteristics of the uninsured: They tend to be between 19 and 24 years old, nonwhite, low-income, and employed only part of the year—the people who fall through the cracks between Medicaid and full-time employment which usually provides insurance coverage. Interestingly, the always uninsured show both fewer disability days and fewer ambulatory visits than the always insured (holding age constant), while the part-year insureds show about the same visit rate but more disability days (Wilensky and Walden 1981). (However, the part-year insureds use ambulatory care at twice as high a rate during the time when they are insured as during the time when they are uninsured.)

Taken together, these figures indicate both a rapid growth in real health care costs (somewhat dampened in recent years), accompanied by a growth of employer insurance premium payments that are shielded from taxes. Moreover, for reasons to be discussed in detail below, research suggests a link between the level of insurance coverage and the growth in health care costs.

III. *The Tax Treatment of Health Insurance*

Premium payments for health insurance are given special tax treatment in two ways. First, employer payments for health insurance provided as a fringe benefit are a tax-deductible business expense for the firm, and are not treated as em-

[1] "Uninsured" here means persons not covered by private insurance or Medicare and not currently with incomes sufficiently low or medical expenditures sufficiently high to make them eligible for Medicaid.

TABLE 5

HEALTH INSURANCE STATUS BY SELECTED POPULATION CHARACTERISTICS (NMCES HOUSEHOLD DATA: UNITED STATES, 1977)

		Always Insured	Always Uninsured	Insured Some of the Time
All Persons		**177,830,635**	**18,226,732**	**15,976,550**
Characteristics			Percent	
	Total	83.9	8.6	7.5
Age	Under 6	80.4	8.3	11.3
	6–18	83.8	8.6	7.5
	19–24	70.0	16.0	14.3
	25–54	83.9	8.7	7.4
	55–64	87.4	8.2	4.4
	65+	97.8	1.2	1.0
Sex	Male	83.4	9.4	7.2
	Female	84.4	7.8	7.8
Ethnic/racial background	White	86.0	7.0	7.0
	Black	76.8	9.7	13.5
	Hispanic	75.8	12.8	11.5
	Other[a]	71.8	16.4	11.9
Family income[b]	Poor	72.9	14.1	13.0
	Other low income	78.4	12.0	9.6
	Middle income	86.0	7.6	6.4
	High income	90.3	4.9	4.8
Occupation	Farm	77.7	15.9	6.4
	Blue collar	80.2	11.3	8.5
	Service	79.2	11.9	8.9
	White collar	87.4	5.6	7.0
Perceived health status	Excellent	85.2	7.9	6.9
	Good	83.3	8.7	8.0
	Fair	82.1	9.2	8.6
	Poor	84.4	8.9	6.7
Employment status	Employed all year	84.2	8.7	7.0
	Employed part of the year	78.7	11.7	9.6
	Never employed	86.4	7.6	6.0
Region	Northeast	89.4	5.4	5.3
	North Central	87.5	5.7	6.8
	South	79.6	11.6	8.9
	West	79.3	11.7	9.1
Residence	SMSA	85.2	7.5	7.4
	Not SMSA	80.8	11.2	8.0

Source Walden et al. (n.d.).

[a] Includes all other ethnic/racial groups not shown separately.

[b] Adjusted for family size.

ployee income for tax purposes. Second, consumer payments for health insurance are tax deductible for those taxpayers who itemize. Formerly up to $150 per taxpayer in health insurance premiums per year was deductible without regard to total medical expenses, and the remainder of premiums was deductible if total medi-

cal expenses exceeded a certain proportion of income. The $150 direct deduction has recently been eliminated.[2]

The tax expenditures associated with these two types of exclusions, as estimated by the Congressional Budget Office (1982), amounted to about $28 billion in 1982, or about 10 percent of private health insurance expenditures. Of the total tax expenditure, about a fifth came from the individual deductibility, and the remainder came from employer tax deductibility (Martin Feldstein and Elizabeth Allison 1974; Bridger Mitchell and Ronald Vogel 1975). The tax subsidy increases strongly with income; for a given insurance premium, the subsidy as a percentage of premiums was 8.4 percent for a person with $3,000 gross income in 1970, but was 25.6 percent for persons with income above $50,000 (Mitchell and Vogel 1975). Because higher income workers receive higher levels of employer-paid premiums, the actual tax subsidy is even more strongly related to income. For example, Mitchell and Charles Phelps (1976) furnished 1975 estimates that indicated a tax subsidy per household of $16 for the $3,000 to $5,000 income class, while for the $20,000 to $50,000 class the subsidy was $139.[3] On average for U.S. workers, the marginal subsidy now exceeds 35 percent of premiums (Phelps 1983b). This subsidy includes both the effect of federal and state income taxes, and the social security tax for those employees whose incomes do not exceed the maximum. Because of tax bracket "creep," the strength of the subsidy has been increasing over time until recently.

In general, the premium for a market insurance policy exceeds the benefits expected to be paid or received because of selling costs, administrative cost, and insurer profit. For typical large group insurance, this "loading"—the difference between premiums and expected benefits—is 10 percent or less. Consequently, the effect of a tax subsidy in excess of 10 percent of premiums is to make the after-tax price of insurance protection negative for many workers; it becomes cheaper (after taxes) to pay one's medical care bills via insurance than to pay them directly. ("Insurance protection" here is the transfer of wealth from states of the world with low losses to states with high losses.)

In what follows, the discussion will deal primarily with employer tax deductibility. The initial question of interest is how the level and pattern of insurance coverage is determined, and how it would be affected by tax-induced changes in net loading. To answer this question, one first needs to look at the demand for insurance.

IV. *The Purchase of Insurance*

One of the most conspicuous and unusual features of the medical economy is the prevalence of insurance. In this industry, for several good reasons, insurance has done considerably more than just transfer purchasing power from the lucky to the unlucky; it has fundamentally shaped the structure of the market itself. It should not be surprising, therefore, that tax-induced distortions in the insurance market should be expected to have strong echoes in the market for medical care itself.

[2] Uninsured health care expenses above some limit are also tax deductible, but relatively few households now benefit from this implicit insurance.

[3] Although the dollar amount of the tax subsidy increases with income, the increase is less than proportional to income over the range of middle income, largely because the average level of employer contributions is fairly constant. This means that although a disproportionate *share* of the tax subsidy goes to higher-income households (Jack Meyer and William Johnson 1983), and although the subsidy is progressive up to approximately the median income (Chollet 1984), it is regressive at incomes above that level.

A. *Theory of Insurance*

Most explanations of the medical insurance market begin by postulating an orthodox basis for the widespread purchase of medical insurance—the attempt by risk-averse individuals to maximize expected utility. Such a theory implies that insurance will tend to be purchased if loading or administrative cost is not excessive, and that the amount of insurance purchased will be inversely related to the marginal loading (Isaac Ehrlich and Gary Becker 1972). Under plausible assumptions about the relationship between wealth and risk aversion, and about the pattern of loading, theory generally implies that coverage is more likely to be purchased, and to be purchased in larger quantities, against events where the variance or some other measure of risk is greatest. This implies more extensive and intensive coverage for the larger loss (Phelps 1983a, 1983b). Other things being equal (including moral hazard, to be discussed below), theory also implies that types of medical care associated with low-probability, high-loss events, such as hospitalization, are more likely to be insured than types of medical care associated with high-probability, low-loss events, such as routine physician services.

Expected utility maximization by risk-averse individuals has usually been postulated rather than tested. The most frequently cited piece of contrary empirical evidence to the expected utility hypothesis is the purchasing of first dollar supplemental coverage by Medicare beneficiaries. Even at loading of up to 50 percent of premiums, and with minimal tax subsidies, more than half of all Medicare beneficiaries do buy private coverage against Medicare's initial deductibles and copayments. However, this apparently irrational purchase of low-value insurance at a high price is not inconsistent with expected utility maximization. Individuals who purchase such coverage also obtain more benefits from their publicly provided Medicare coverage (because of moral hazard), but they pay no additional Medicare premium. When these additional benefits are added to those explicitly financed by the private coverage, the ratio of benefits to premiums becomes much more favorable (Paul Ginsburg 1981).

While not all of the implications of these alternative theories of the purchase of insurance have been specified, let alone tested, such tests that have been constructed, and the overall pattern of health insurance purchases, do seem to be generally consistent with orthodox expected utility theory. For example, John Hershey et al. (1982) found that, in contrast to other uncertain events, people were generally willing to choose coverage against low-probability, high-loss medical expenditures over coverage of equal actuarial value for high-probability, low-loss events. Charles Phelps (1973) also provides evidence that indicates an increase in the fraction paid by insurance as a person's total medical bills rise, from 7 percent if the bill was less than $150 in 1970 to 76 percent if it exceeded $1,500. However, Phelps and Marquis (1983) analyzed survey responses to questions about hypothetical insurance purchases, and found results inconsistent with the expected utility model.

There have been some alternatives to a conventional expected utility theory. For example, Victor Fuchs (1979), and Bruce Vladeck (1981) argue that insurance has in part the purpose of freeing the consumer from the necessity of considering prices and alternative uses of wealth in the period of emotional strain that often surrounds the incurring of medical expenses.

Moreover, illness does not cause a direct

reduction in tangible wealth. Yet the orthodox theory of demand for insurance is usually characterized in the context of events that might result in a reduction in such wealth. Therefore, it is not obvious that the conventional theory should apply to health insurance. There are two modifications that have been considered in the literature. The first, discussed briefly, is that insurance necessarily pays in the form of money, not "health," and the marginal utility of money may be affected by the state of health.[4]

The other peculiarity of medical insurance, one that has received a great deal of theoretical and empirical attention, is the existence of moral hazard. Moral hazard arises whenever an individual's behavior that affects the expected loss is altered by the quantity of insurance he obtains (Pauly 1968). Moral hazard can occur because insurance affects either the probability of an event associated with a loss or the size of the loss conditional on the occurrence of the event. Translated into the context of medical care, the first sort of moral hazard arises when the purchase of health insurance encourages individuals to spend less on preventive medical care. The second sort of moral hazard occurs when the purchase of insurance induces an individual who has experienced an illness to spend more resources on its treatment.

Why does moral hazard characterize medical insurance? Put another way, why cannot medical insurance payments be made conditional on an event over which the insured has no control? In part, the answer is the insurer's difficulty of observing the insured's preventive activities that reduce the likelihood of illness, an aspect much emphasized in the principal-agent literature. But a much more important part seems to be related to the difficulty of defining, in an "objective" or contractually specifiable way, the severity of an individual's illness once an illness has occurred (Pauly 1971; Joseph Newhouse 1978a). As a consequence, severity has usually been measured by the amount of medical expenditure, a quantity generally correlated with the severity of illness but also a quantity under the partial control of the insured.

When moral hazard is present, insurance that reduces risk will also cause larger expected losses. In medical care, these losses represent the consumption of units of medical care whose value to the consumer is less than their cost, because the insurance coverage reduces the user price below cost. The welfare loss associated with this distortion can be limited if some uncertain medical events are not fully insured (Pauly 1968).

In addition to moral hazard, the other determinant of the form and structure of medical insurance is loading or administrative cost. Real resources are consumed in selling and administering the insurance contract, and therefore the total premium the insurer receives has to exceed the actuarially fair premium, even if the insurer is assumed to behave in a risk-neutral way. How the loading varies with the amount of type of insurance, and with the setting in which the insurance is purchased, will be an important determinant of market equilibrium. Other things equal, when individuals purchase insurance in a manner consistent with expected utility-maximizing behavior, the greater the strength of moral hazard, the lower the optimal level of coverage. If there were no loading and no moral hazard, full coverage would be

[4] As Steven Shavell (1978) has noted, if the marginal utility of money depends on the state of health, it is not obvious whether the utility-maximizing quantity of insurance should be greater or less than the level of spending on medical care. Somewhat surprisingly, there have been no investigations of the relationship between the demand for medical insurance and the demand for insurance that pays cash in the event of illness or death, such as salary-continuation insurance, disability insurance, life insurance, or the health policies that pay cash per hospital day regardless of other coverage.

purchased against any uncertain event (Kenneth Arrow 1963), but loading (Dennis Lees and Richard Rice 1965) can also cause less than full coverage to be optimal.

The theory concerning the general consequences of a tax subsidy is simple. The effect of a tax subsidy is to offset the effects of loading and moral hazard, and to lead to an individual purchasing more insurance. With higher levels of coverage, not only will actual loading costs be higher but, because of moral hazard, medical care costs will be higher as well. If the true state of health cannot be determined by the insurer, moral hazard cannot be prevented, but will be limited by the way in which insurance premiums will rise disproportionately with coverage (Richard Zeckhauser 1970; Pauly 1974). The effect of a tax subsidy then is to push the market away from even a second best optimum, and make the extent of moral hazard (and welfare loss) greater than would occur if coverage were not subsidized.

The problem is complex, however, because there are many different forms that insurance coverage can take: There are deductibles, coinsurance provisions, maximum limits on benefits, and a host of more specific explicit and implicit benefit limitations that are written into health insurance contracts. It will be useful initially to look at research on each one of these limitations separately.

B. *Deductibles, Coinsurance, and Other Explicit Financial Benefit Limits*

As Arrow (1963) suggested, deductibles in general can be understood as an attempt to reduce the total insurance administrative cost ("loading") actually paid. If the loading is proportional to the total premium, the optimal pattern of coverage will be full coverage above a deductible. A small amount of risk becomes tolerable to the consumer in order to save on the loading. In the real world of health insurance, things are more complicated, partly because loading (or at least administrative cost) is probably not proportional to total premiums or benefits, but rather depends as well on the number of claims. Therefore, the prevalence of deductibles is also alleged to be due to a desire to avoid the cost of processing many small claims.[5]

The impact of deductibles, rather than simple proportional coinsurance, on the use of medical care is complicated, because deductibles make the user price nonlinear: it will be equal to the market price for a while, and then it falls (often to zero). So, as Emmett Keeler et al. (1977) have noted, the impact of a deductible on use depends on whether the spending unit expects to exceed the deductible.

The deductible tends to discourage use if expenses would have been well below the deductible, and to leave large expenses unaffected. In practice, deductibles do seem to have some small depressing impact on health care costs, although there has been virtually no empirical study of the actual administrative cost savings associated with a deductible in medical insurance.

Coinsurance, in which insurer and insured pay fixed fractions of total loss, is easier to understand in theory. Not only does coinsurance reduce the amount devoted to a proportional loading, it also serves to restrain use by raising the user price per unit of medical care. In principle, knowledge of the demand curve and, especially, the user price elasticity, can predict the effect of a linear coinsurance arrangement. A number of such estimates will be discussed below, most of which treat coverage as a simple proportional coinsurance.

[5] One problem with this explanation is that, in theory, such an argument should imply not a deductible, but rather an upper bound—insurance that pays nothing until a claim exceeds a certain amount, but then pays in full (Shavell 1978). Another problem is that, in practice, deductibles are fairly common for inpatient hospital care, even though virtually all inpatient bills exceed the usual deductible.

Unconditional upper limits—usually limiting benefits to a maximum total dollar amount per time period—do occur in both private and public insurances. The rationale for these limits is not so obvious. Given the large number of individual spending units covered by a typical insurer, insurer risk aversion can hardly be a reason (though it could be a reason for limits for small self-insured groups or plans). One could perhaps explain limits as a very blunt instrument to control moral hazard, or as a way of catering to misperceptions of probability—"I could never be that sick."

The final insurance payment device is what one might call "quasi-indemnities." For example, insurances may pay "up to" a certain amount for every hospital inpatient day experienced, or up to a certain amount for a normal delivery, etc. Quasi-indemnities can be viewed as approximations to true (nondistortive) indemnities. One puzzle is why indemnities are not more common. Only in dental insurance are benefits typically conditional on the submission and approval of a diagnosis and plan of treatment, which in a sense defines the "illness" state. Perhaps the explanation is that for medical conditions, such pretreatment determination of the illness state may be more costly, or more likely to be in serious error; such procedures are quite uncommon in medical insurance. Other explanations are possible, of course—the medical profession may have been more successful in resisting such limits (Warren Greenberg 1981), or moral hazard in unconstrained dental insurance may be that much worse.[6]

So far we have discussed what might be called "contractually explicit" devices to constrain benefits and use. Although such devices may be administratively costly, they do permit the insured to use any provider in a market area, making supplementary cash payments if necessary; they preserve free choice of provider. An alternative approach to financing health care is to have no explicit limits on benefits and care, but to link payment of benefits to the use of certain providers. The style of practice or care protocols used by those providers then furnishes a kind of implicit contractual limit on use and cost. The ultimate incentive to providers to limit the amount of care in this way is the need to keep the tied premium low to attract customers.

The Health Maintenance Organization (HMO), which provides comprehensive coverage for services of a specific set of providers in return for a capitation payment, furnishes the best-known example of such implicit benefit limits. Such plans, which currently enroll about 7 percent of the population but are growing rapidly, typically have no user charges or explicit benefit limits if physicians and hospitals associated with the plan are used. They do, however, limit enrollees to obtaining care from approved providers. HMO physicians are offered incentives to avoid costly treatment by such devices as financial incentives (for example, payment by salary, or by shares in the profit of the plan), by administrative devices, and by selection of less aggressive types of providers (for example, fewer surgeons, or fewer physicians trained in high cost areas).

In addition to the specific organizational structure of an HMO, arrangements have recently emerged to offer similar incentives with a less formal organizational structure, in the sense that providers are not necessarily employees of or have binding contractual arrangements with the organization that receives the premium.

[6] It is also possible that insurance will affect use even if there is no moral hazard. As Martin Feldstein and Bernard Friedman (1977), John Marshall (1976), and David de Meza (1983) have suggested, even pure indemnity insurance is likely to have an income effect, and the income effects of benefits relative to premiums do not necessarily cancel out.

Such arrangements are usually called "Preferred Provider Organizations" (PPOs), and usually involve low out-of-pocket payments for use of a designated set of doctors and hospitals, price concessions by those hospitals, and the monitoring and exclusion of high-expenditure physicians.

C. *The Choice of Insurance*

Even from the viewpoint of a potential individual demander of insurance, the choice among various types of insurance is not an easy one. (That 80 percent of private insurance is chosen by and for employee groups will be considered below.) The buyer must choose among various types of explicit plans. In addition, each HMO or other health plan represents a different, and difficult to ascertain, implicit benefit limit. For example, for some illnesses the individual considering an HMO may prefer to use amounts of or types of providers different from those furnished by any particular plan. But he must estimate how likely such a situation will be when he chooses which plan to buy, or whether to join an HMO at all (as compared to conventional insurance, which pays for any level of use). Here we will consider the choice as being made in two steps:

1. Among all conventional insurances, the individual chooses the one he likes best, likewise for all potential health plans.
2. He then chooses between the best conventional insurance and the best health plan.

We will consider the work on incentives and performance of HMOs in more detail below. The critical point here, however, is that there has been little satisfactory modeling of the choice between HMO and conventional insurance that takes the implicit benefit limitation feature into account, and no modeling at all of the impact of changes in the tax subsidy on this choice.

D. *The Impact of Price on Demand for Conventional Insurance*

With regard to the influence of price or loading on demand for conventional insurance, empirical and theoretical research has been made difficult because there is no obvious direct measure of quantity of coverage. What is more important, a given expected benefit level (measured in dollars of expense covered) can have quite a different loading cost—the proper general measure of the price of insurance—depending on the circumstances in which the expenses were incurred.

Despite the real world complexity, empirical researchers have tended to model the price of insurance simply as a percentage loading on benefits (or proxies for it), and to measure quantity in very crude ways. Nevertheless, there is fairly consistent evidence that insurance coverage is sensitive to proxies for its price. Surprisingly enough, there have been no estimates of the impact of loading per se. Instead, most estimates have relied on proxies for the loading factor. The most commonly used proxy in cross-section analysis is the size of the employment group. Feldstein (1973) used as a proxy for the availability of group insurance the proportion of workers in a state employed in manufacturing or public service, and found that it was positively related to the level of coverage. More direct evidence has been obtained from firm or individual data to indicate that the level of coverage is positively related to group size (Charles Phelps 1973; Gerald Goldstein and Mark Pauly 1976; Amy Taylor and Gail Wilensky 1983). Paul Ginsburg (1981) has used data on the way loading varies with group size on average to convert his estimate of the group size effect into an estimate of price elasticity.

The results generally support the view that the impact of loading or loading proxies on insurance purchases is significantly negative. The actual numerical estimates of the elasticity of insurance with respect to the loading "price," however, vary considerably, ranging from about −0.2 (Taylor and Wilensky 1983) to numbers greater than unity (Ginsburg 1981).

E. *The Effects of a Tax Subsidy on the Amount of Conventional Insurance Purchased*

The tax subsidy has the effect of reducing the net loading on insurance. Removal or reduction of the tax subsidy is likely to have all of the following effects:

1. Some types of coverage for which net benefits from insurance are small will be dropped entirely (prescription drugs, dental care).
2. Deductibles will be raised.
3. Copayments will be increased.
4. Other adjustments in the benefit package will occur.

Most estimates of the impact of tax side changes on insurance coverage are not based on direct observation of the impact of such changes on coverage because such changes have been rare. Phelps (1973) therefore generated estimates of effect of tax-induced changes in net price by using price elasticity estimates obtained from variation in the group size proxy. These estimates were in the range of −0.4 to −0.7. Feldstein and Friedman (1977) simulated impacts using a variety of assumed parameters for risk aversion and price elasticity of the demand for medical care, and found sizeable quantitative impacts under reasonable assumptions.

Phelps' most recent cross-sectional estimates (1982), using the elasticities from his 1973 study, indicate that making half the employer-paid premium taxable income for the employee would cause desired premiums to fall by about 50 percent. This fall in premiums and coverage would then induce about a 5 percent drop in total spending. Feldstein and Friedman's simulation results, for plausible estimates of risk aversion, forecast an even larger drop.

The recent study by Taylor and Wilensky (1983) gathered (individual) data that provided separate measures of the marginal tax rate and the level of disposal income. Because of nontaxable sources of income and variations in tax deductions and exemptions, the two measures were sufficiently uncorrelated to permit separate statistically significant estimates of the effect of income per se and the tax rate. The estimated price elasticity is −0.21; the price is measured by (1-marginal tax rate). (The income elasticity is 0.02.) Simulations using the price elasticity indicate a fall in premiums of 7.5 percent when all employer premiums are subject to tax. This estimate is considerably lower than those of the earlier studies; Taylor and Wilensky in fact suggest that "because employees currently have so little choice among health insurance plans" there might be a larger response over a longer run when they can better adjust to new conditions.

An alternative interpretation, leading to the same conclusion, is that when insurance is chosen by and for the group, some summary measure of the characteristics of the firm's work force presumably best explains the choice of premium and benefit level. Any individual worker's characteristics would be an imperfect measure of those summary characteristics, so the estimated relationship of coverage to the individual worker's characteristics could be subject to an errors-in-variables bias toward zero. One piece of empirical evidence consistent with this supposition was Phelps' (1973) finding that the estimated insurance price elasticity for families that purchased some individual insurance (and

therefore had more control over the total quantity they received) was larger than for those families who obtained only group insurance.[7]

Another potential source of error in the Taylor-Wilensky study is the fact that the marginal tax rate proxies the effective price of insurance (loading held constant) only in those cases in which the employer pays the entire insurance premium. A follow-up study by Farley and Wilensky (1984) adjusted for the portion of the premium paid by the employer, and found a more substantial price elasticity estimate of −0.41.

Martin Holmer (1984) used information on the choices between coverage options by federal employees to estimate the impact of the tax rate on the level of coverage, obtaining a very low price elasticity of −0.16. However, Holmer had no direct measures of either employee family income or marginal tax rates, and so was forced to estimate both quantities using regression coefficients (from the Current Population Survey data) relating earnings and tax rates to employee sex, age, earnings range, and presence of an employed spouse. This procedure obviously raises again the possibility of errors-in-variables bias. Moreover, one may be concerned about estimates obtained from individual choices of insurance by self-selected workers who choose a particular employment in a larger labor market context. If total labor compensation is related on average to marginal product, then the worker who prefers a low-coverage package should not choose to work for an employer who offers a set of options in which the high-coverage package is "subsidized" by larger employer contributions, as occurs in the federal employees' plan. Hence, those workers who do in fact choose such employment will tend to be workers with strong "tastes" for insurance, tastes that can offset any tax subsidy effects.

There have been some time-series estimates of the impact of taxation on employer contributions to health insurance. Of course, employer contributions can increase either because the amount of insurance increases or because the employer contributes a larger share of the premium for a given insurance policy. The tax price elasticity of the employer contribution tends to be moderately high, although the quantity of insurance does not necessarily vary with the employer contribution. James Long and Frank Scott (1982) obtained time-series elasticity estimates of −1.4, and Steven Woodbury (1983) obtained similar figures.

Finally, Charles Phelps (1985) generated time-series estimates of the impact of changing the tax rates *and* loading on the demand for group insurance based on the theory that group purchases are chosen to satisfy the preferences of the median worker. The price elasticities estimated from either source of variation are fairly similar, ranging from approximately −1.0 to −1.9.

In summary, current estimates of the price or tax subsidy elasticity permit considerable confidence in the conclusion that it is different from zero in the expected direction; the tax subsidy *is* known to matter to the choice of group insurance. How much it matters is less precisely known. In large part, this is because estimates have been based on proxies, and sometimes proxies for proxies, of the actual net (of tax) loading or price facing the individual whose demand is (in some sense) decisive. To improve estimates, one would need both better data and better modeling of the insurance choice process in a world of competing tax-subsidized employment groups. Looking specifically at the impact of the recent

[7] One might also note that the relevant measure of the "quantity of insurance" is based on the total set of policies the family buys—the Taylor-Wilensky study is unclear on how the quantity is actually defined in two-worker families.

tax cuts would seem to be especially fruitful.

F. *The Impact of Taxation on the Mix of Insurances*

Can we gain some insight fom the literature into the form of cutbacks that would be induced by various changes in tax deductibility? Which types of coverage will be cut? This question has been important, both because some types of medical care are thought to be more of a policy problem than others, and because—short of a complete elimination of tax deductibility—the tax subsidy can be reduced in different ways.

The most commonly discussed method of cutting tax deductibility is to include as taxable employee income the amount of employer payment above some fixed dollar amount. For example, the initial version of the 1985 Treasury Tax Reform program would have taxed family contributions above $2,100 per year. Such a procedure makes the net price of insurance nonlinear—there is a jump in the marginal price once total premiums exceed the limited amount.

The most frequently predicted response to such a change is elimination or reduction in coverage that is concentrated on "marginal" coverages—either those coverages that high-benefit firms provide (as compared to low-benefit firms), or those types of coverage that employers have been adding over time as income grew and tax subsidies increased. The usual conclusion then is that a limited cut in tax deductibility will cause cuts primarily in dental insurance and in coverage of vision care, but will not much affect coverage of hospital care or physician services (Pamela Farley and Wilenksy 1983; Phelps 1983a, 1983b). If the cost increases in the latter two services are regarded as more serious, such partial cutbacks may be judged to be ineffective (Phelps 1983b). However, during the recent recession (which was also accompanied by a cut in marginal income tax rates), it appears that many employee groups moved from first-dollar coverage into so-called comprehensive (major medical type) benefits, with positive deductibles and copayments, rather than cut dental insurance (Swartz 1984). For example, the Health Insurance Association of America Survey of new group health insurance policies found that the percentage of employees receiving "first-dollar" (no deductible or copayment) hospitalization insurance as new group insurance fell from 36.4 percent in 1980 to 6.6 percent in 1982 (Chollet 1984).

In theory, it would seem that insurance coverage of those services with greater moral hazard, given some level of variance or risk at current levels of coverage, should display greater price responsiveness (Feldstein and Friedman 1977). Such a conclusion arises from the notion that the level of coverage depends on marginal welfare cost, so the size of the marginal welfare cost at any level of coverage should vary with the (absolute value of) the medical care price elasticity. A second suggestion is that the change should also be greater, given a level of moral hazard (possibly zero), for coverage of those events associated with a smaller variance in expenses (Phelps 1983b). Because any risky event would be fully covered at a fair premium, but coverage would vary directly with risk at any given level of loading, the responsiveness of coverage to loading for the less risky event must, in some sense, be more responsive over at least part of the range of net loadings.

The difficulty with looking at data on the types of coverage that are added in groups that spend more on insurance is that we do not know which of several reasons—price, income, region, or occupation—accounted for the additional spending. For instance, if higher income is the primary reason for choosing to spend on

dental benefits, so that the *income* elasticity of demand for such benefits is high, an increase in loading will not necessarily result in an appreciable cut in dental benefits if the (insurance) *price* elasticity is relatively low. One would want to relate the type and amount of coverage directly to the net loading, but thus far this task has not been accomplished.

Because (net of taxes) insurance premiums are often actuarially favorable, given the tax subsidy, it is plausible that the major limit to some types of insurance coverage is the usual practice of limiting benefits to no more than 100 percent of medical care costs. (Although insurances that pay cash over and above medical expenses are sold, such insurance benefit payments would probably be treated as taxable income if provided by an employer.) Hence, the employee group may well not be satisfying the marginal conditions with regard to the purchase of such insurances, and tax-induced changes in price may provoke no change in quantity until the marginal net-of-taxes premium becomes actuarially unfair.

In general, if the employee group were on the margin, theory would predict some reduction in benefits for *all* types of insurance in response to limits that make marginal dollars of employer-provided insurance fully taxable. Therefore there should be some effect on hospital and physician coverage; how large this effect will be is an important research topic.

G. *How Health Insurance Is Actually Chosen*

We could end the story of the effects of tax subsidies on health insurance with this puzzle over the form and the magnitude of the adjustment in desired coverage if, as with most other goods, the amount of the good each consumer obtained equaled the amount he desired.

However, about 80 percent of private health insurance is purchased through the workplace;[8] and in the great majority of groups there is no choice by the individual employee. A large number of employers pay the full premium, requiring no voluntary health insurance payment by the individual; there is almost always at least partial premium payment by the employer. Evidence is quite convincing that people have different desired levels of health insurance, although the source of this difference is open to question. How are these differences reconciled when the employer or the union must choose a single quantity or a limited variety of quantities for a group of many potentially different employees? Because elimination of tax deductibility at the margin will change the net price of insurance *differently* for employees in a group who pay different marginal tax rates, this question is especially crucial.

The arrangements under which employers pay for health insurance are varied. Many employers offer only a single plan, and often pay the full premium for it. Of those that pay less than the full premium, there appears to be a pattern of payment of a uniform *fraction* of the premium, rather than requiring the employee who chooses a more expensive plan to pay the full incremental premium (Wilensky, Farley, and Taylor 1984).

The first question to be settled is that of the incidence of employer payments. It is almost always supposed (Phelps 1983b) that such payments ultimately come from employee wages, although there has been no direct investigation of whether the incidence does fall entirely on the wage bill. The theoretical and empirical foundation for this assumption is the work on the payroll tax (John Brittain 1972). What makes employer payments

[8] The self-employed do not have access to the employer-contribution tax subsidy, nor is there an analogous tax break; there is no Keogh plan for health insurance.

for health insurance potentially different, however, is the possibility for endogenous variation in employer insurance payments across firms, in contrast to the uniformity of the payroll tax. Whatever this does to the aggregate incidence, it may surely affect the incidence across firms. (One might also note that, despite the unanimity of economists that the incidence is on labor, employers are consistently told that their profits are eroded by high health benefits costs [Chollet 1984].)

There is the additional question of how that pattern of reductions is distributed across employees within firms. The usual arrangement is that any worker classified as full-time is eligible for full benefits. This means that, for such employees, an employer-paid monthly health insurance premium presumably reduces total earnings per month but not the marginal wage per hour. (It may also affect the firm and employee's decision about full-time versus part-time work.)

What is not known is whether differences across employees in expected losses have full incidence. For example, if a particular class of workers in a firm is dominated by older workers with large families and larger expected expenses, would employer-paid insurance coverage result in a reduction in monthly earnings for them that is larger than that for other groups of workers, and should cuts in coverage give them larger increases in monthly income? If part-time workers in high fringe benefit firms are not eligible for health insurance, are their hourly wages higher relative to those for full-time workers, as compared to low fringe benefit firms?

This within-firm incidence is important for understanding both work force composition and the level of insurance benefits. For example, it has been alleged that firms with relatively many part-time workers would be expected to have lower levels of insurance benefits (Suresh Mahotra et al. 1980; Wilensky, Farley, and Taylor 1984). However, if part-time workers can be paid at different rates, this result need not hold.

A consistent finding in studies of group health insurance is that unionization is associated with higher levels of benefit (Victor Fuchs and Marcia Kramer 1972; Feldstein 1974; Goldstein and Pauly 1976). "Ideology" may be the cause, or it may be the case that union choices are affected by the interests of older workers with larger families with higher incomes receiving larger tax subsidies who are subsidized by younger workers, and who therefore desire higher levels of coverage (Goldstein and Pauly 1976).

Even if we ignore for the moment the issue of how the quantity and type of insurance at each firm is determined, the answer to the question of the impact of tax changes obviously depends on the model of a labor market in which different levels of fringe benefits are offered to people of different types—a sort of labor market version of monopolistic competition, in which the number of options varies across markets and occupations. What is even more confusing is that the "type" of insurance bought depends not only on the ostensible provisions of the contract, but also on the type (sometimes not fully observable) of employees covered.

Goldstein and Pauly (1976) argued that a way to represent the "movement equilibrium" of workers among firms is analogous to that of local governments, where voters "vote with their feet" (Charles Tiebout 1956). In the ideal version of such models, given the usual large numbers assumption, one gets the Tiebout conclusions—work forces should become homogeneous within a firm with respect both to the desired form of the insurance contract and to the observable loss-associated characteristics of workers. The "public good" should equal the desired quantity for every group member.

Failing this possibility, there has at least to be equilibrium at the margin. That is, the marginal worker's tradeoff between two "adjacent" insurance packages governs the amount by which wages offset. Of course, competitive equilibrium requires that total compensation per labor hour be the same everywhere, but should one think that labor markets are perfectly competitive when fringe benefit variation is taken into account?

The empirical evidence on the fraction of the premium for a group insurance policy contributed by the employer bears on this question. If labor market behavior could produce firm labor forces with identical preferences for health insurances, one would never expect to find the employer paying less than 100 percent of the premium for that coverage. Having the employee pay something loses some tax subsidy, and this would only be worth doing if some employees would prefer not to make the employee contribution and instead either buy no insurance or buy some other insurance policy. Chollet (1984) estimates that 15.8 percent of establishments nationwide required some employee contribution for the employee's coverage in 1980, and 39.9 percent required some employee contribution for dependents' coverage.

What would happen in the employee group if some part of the "employer paid" premium for health insurance is made taxable income for the employee? This change would increase the marginal price for all workers, but the amount of the increase would depend on the incidence of the tax payment and each worker's marginal tax rate. If wages and premiums were left unchanged, firms offering high levels of benefit would be less attractive than they were before. Altering their benefit packages to the most preferred level would probably mean, for most firms, a package with a cost just at the level at which premiums become taxable. Equilibrium would be reestablished, but there could be some movement of workers among firms. That is, changes in the tax treatment of fringe benefits will affect labor supply differently at different firms, and a migration of employees among firms and alterations of insurance offerings could result. Note that the cash wage would probably not rise to the full extent of the premium reduction; some of it would be diverted into other fringe benefits.

If unionized firms do tend to be dominated by workers facing larger tax subsidies, one might expect that, other things equal, cutbacks would be larger for unionized firms. More generally, it seems clear that an adequate explanation of the impact of taxation on group insurance choice will require a more sophisticated model of that choice than we now have.

A final interesting question here is whether firms whose health insurance offerings remain below the taxable limit might be affected if the level of tax-free benefits were limited. It seems likely that increased taxability would reduce the relative attractiveness to workers of high-benefit firms, on average, causing an increase in money wages there to rise in order to compensate. The general equilibrium incidence of tax deductibility may differ from the initial pattern.

H. *Adverse Selection and Tax-side Changes*

It is possible that individuals know more about their expected health expenses over a given time period than can be discerned by an insurance firm. Such asymmetry of information can give rise to the problem of adverse selection, with well-known negative consequences for efficiency if equilibrium does exist, an embarrassing richness of potential equilibrium concepts, and the possibility of nonexistence of equilibrium (Pauly 1974;

Michael Rothschild and Joseph Stiglitz 1976; Charles Wilson 1977). Although welfare judgments are often difficult, one can at least conclude that there is no presumption for optimal equilibrium when adverse selection is present. This does not, however, imply that the market outcome with adverse selection (which may or may not represent a equilibrium) is *necessarily* inferior to feasible alternatives, or that retention of a tax subsidy is justifiable on adverse selection grounds.

There is as yet little direct evidence on whether or not adverse selection is or must be an important problem in health insurance, largely because it is difficult to define any kind of strong test. The intuitively convincing observation that people who use more care have more insurance is, for example, also consistent with *homogeneous* risks (i.e., no adverse selection) and moral hazard. James Price et al. (1983) found that for the Blue Cross plans for federal employees, total medical expenses were about 38 percent greater under the high-option plan than under the low-option plan. But does this mean adverse selection or moral hazard? The difference is roughly of the order of magnitude suggested by recent studies of moral hazard (Newhouse et al. 1981).

Despite the outpouring of theoretical models of markets with adverse selection, there has as yet been no conclusive application of those models to the health insurance industry. There are several reasons for this:

1. Insurers seem reluctant to vary premiums with many of the characteristics they *can* observe that predict medical expenditures, such as age and family size. A typical premium will depend only on whether the insurance covers the individual or his family, but little else. This is true of insurance sold directly to individuals, and is even more true for multiple-choice employee groups. Indicators of chronic conditions are even less frequently used. While there can be some reasons for such behavior, it raises some doubts as to the general applicability of models of competitive profit-seeking insurers maximizing against an information constraint.
2. As noted by Pamela Farley and Gail Wilensky (1983), employers tend to pay fixed-premium shares or in other ways to cross-subsidize the high risks. This practice probably has the result of making adverse selection less likely to occur; some reasons for its existence will be discussed later.
3. The public policy question (also discussed by policy-oriented economists like Alain Enthoven [1980]) does not generally concern what economists mean by the *inefficiency* of adverse selection. Rather, it concerns the alleged *inequity* of charging higher premiums to those who can be identified as high risks. It is regarded as unfair to charge sickly individuals actuarially based premiums, whether adverse selection occurs or not.
4. The general model that seems most realistic for the very unconcentrated nongroup (individual) health insurance market, the model of Rothschild and Stiglitz, either predicts nonexistence of equilibrium or an equilibrium in which bad risks buy full coverage (at zero-profit premiums) but good risks buy incomplete coverage. Obviously a model with no equilibrium is not especially useful either for predicting market outcomes, or for making welfare judgments.

There has been no comprehensive empirical study of adverse selection impacts on the stability of health insurance markets. The market for individual coverage appears to be stable; there is little evidence that more

comprehensive plans are driven out of existence by switching of good risks into less comprehensive plans, with this situation then followed by reemergence of the more comprehensive plans. That is, there is no evidence of the absence of equilibrium which is the major feature of the Rothschild-Stiglitz model, and which is alleged by Michael Spence (1978) to be even more likely when the spectrum of risks approaches a continuum. Among the multiple-choice employee groups, the Federal Employees Health Benefit Plan (FEHBP) is the largest group plan that permits employees to choose among a number of coverages, HMO and non-HMO. Enthoven (1980) has asserted that, at the 60 percent employer payment in that plan, "there is no evidence that any of the competitors have suffered seriously from adverse risk selection." James Price, James Mays, and Gordon Trapnell (1983), however, do claim that more recent data show a substantial amount of switching, and that switching tends to be toward those plans that provide larger net benefits, and against full coverage conventional insurance in favor of both HMO coverage and conventional "low-option" coverage. The net benefits appear to vary primarily because there is lower than average use under some of the low-option plans, although it is impossible to tell whether that lower use comes from risk selection or from moral hazard.

5. Many recent models of adverse selection suppose that each insurer will anticipate further reactions to its attempt to pick off the good risks, and so a profit-maximizing "non-Nash" equilibrium will exist (Wilson 1977; Haijime Miyazaki 1977). Given the very small market share of any commercial insurer, this scenario seems implausible. The only insurer with a large enough market share to engage in such behavior is Blue Cross, but Blue Cross is not-for-profit and often associates its interest with those of hospitals (Frech 1974). It does not seem to behave in a profit-maximizing way. Indeed, some of the empirical examples of alleged adverse selection in health insurance appear to be caused primarily by Blue Cross' and employer failure to take profit-maximizing steps to discourage bad risks *who could have been identified* from buying coverage at prices below those appropriate to their expected loss.

   For example, in the Federal Employees Health Benefit Plan, the complete absence of any restrictions on transfers into the comprehensive plan could hardly be consistent with competition. In FEHBP, the most comprehensive plan is said to display adverse selection because it contains more older workers (Marcia Gold 1982). But worker age is surely observable by the insurer, and so premiums ought to vary with age if it is a good predictor of medical care use. Moreover, adverse selection against high-option benefits is a recent (and unexplained) phenomenon; initially federal workers selected overwhelmingly into the more comprehensive plan (Ronald Anderson and J. Joel May 1971). What is needed therefore is an explanation of why things changed to make adverse selection a problem (Pauly 1986). Although the limits on the ability of insurers to discriminate in FEHBP are in large part required by the employer, adverse selection when an observable risk indicator is ignored does not establish the existence of overall market failure.

6. As noted above, the bulk of insurance is provided in employee groups. Although there has been movement to more flexible benefit plans that permit some choice, the great bulk of group insurance still involves a single quantity of insurance offered to all employees. Such an arrangement may have other defects, but it certainly inhibits adverse selection within the firm.

Why do actual insurance markets in which there is choice, both the direct individual market and the multiple choice employer group, seem to be relatively little affected by the instability that would characterize a Rothschild-Stiglitz world with no competitive equilibrium? One reason is offered by Joachim Neipp and Richard Zeckhauser (1986), who suggest that there can be "persistence," an unwillingness to switch insurance plans even when relative premiums change, because of switching costs. Another answer might be found in the recent demonstration by Jonathan Cave (1984, 1986) that, in the case of insurers who offer multiple plans and who only need to break even on the *full set* of plans, a zero profit equilibrium not only always exists, but it is always (second best) Pareto optimal. Such equilibria usually require that the low coverage plan purchasers subsidize the higher coverage plan purchasers just enough to keep the latter from switching into lower-coverage plans. In fact, Cave's "multiple-choice group" equilibrium is exactly the Miyazaki-Spence version of Wilson equilibrium. This important result has obvious direct applications to group multiple-choice insurance plans; it means that by permitting the initial choice of options and premiums to be selected by the group, the inefficiency of adverse selection can probably be effectively eliminated. No recourse to external intervention in the form of tax subsidies or tax credits limited to "qualified" (high coverage) plans, such as these suggested by Enthoven (1980), need be used. What is needed is a corresponding theory of group choice—which sets of breakeven plan-premium combinations will a group (or an employer) optimally choose, given that adverse selection is possible?

How adverse selection might be affected by changes in the tax treatment of health insurance has been a matter of some speculation. The response of a group's coverage decision to changes in the tax treatment is complex, because it involves heterogeneity in risk, heterogeneity in demand for insurance given risk, the market for various types of labor, and the mix of labor types used to produce a particular good.

Obviously, where a firm's labor force is homogeneous with respect to risk, there is no problem of adverse selection. If the firm's labor force is also homogeneous with respect to the demand for insurance, the response to renewal or reduction of tax deductibility would be a cutback in the single uniform level of benefits required for all employees; there would be no advantage to permitting a high-option benefit or opting out of the group, because there would be no benefit from doing so but there would be higher administrative costs. If the firm's labor force is homogeneous with respect to risk but heterogeneous with respect to demand for insurance, then removal or reduction of the tax subsidy removes an artificial stimulus to uniformity.[9] This effect will be stronger the higher the marginal tax rate, the greater the heterogeneity in demand for insurance, and the lower the cost-reduc-

[9] Heterogeneity of preferred quantities can arise from reasons other than tastes for risk. For instance, some employees may also be covered under a spouse's plan. The tax stimulus to homogeneity is weaker when "cafeteria" insurance plans, which permit each employee to choose how much of his income to put into a variety of tax-free benefits, are available.

ing effect of uniform coverage. High-wage workers in large groups might be especially likely candidates for offering multiple-option coverage if the tax subsidy is limited.

What if expected losses do vary among employees, so that adverse selection is possible? Removal of the tax subsidy would in some cases lead to a greater variety of choices and probably more adverse selection (Chollet 1984). Cave's theory appears to suggest, however, that the group may choose the (second best) optimal arrangement in the absence of a subsidy. That is, even if there is more choice and more adverse selection with a cut in the tax subsidy, the outcome may still be preferable to no adverse selection but no choice. Research is needed to clarify this issue.

The efficiency arguments assume, of course, that the menu of choices available to the group will in fact be chosen optimally by the employer or by the group. If a union chooses, such an assumption could be incorrect, depending on the decision process of the union; unions may not make the welfare-maximizing choice for their members. Employer choice as well may not maximize welfare, here as elsewhere, because it ignores the preferences of inframarginal workers.

A final important point to note is that, while adverse selection may be undesirable in a world with no other distortions, it can actually be welfare-improving when there are other incentives for excessive coverage (such as tax subsidies). Adverse selection drives high-option coverage for people who are not high risks out of the market. But if one has the objective of reducing medical care costs, that demise contributes to the objective. For example, the strongest empirical evidence we have for adverse selection is the finding that persons who purchase private insurance supplemental to Medicare have expenditures that are so much higher than those without supplemental coverage that moral hazard is unlikely to explain the full difference (Charles Link et al. 1980). However, discouraging the purchase of supplemental coverage by adverse selection may be a beneficial offset to the implicit subsidy to such coverage mentioned above.

At present, the situation for which there is the strongest empirical evidence for inefficient adverse self-selection is not between conventional insurance policies with different levels of coverage. Rather, as will be discussed below, selection has been most frequently documented in comparisons between HMO coverage and conventional high-option coverage, but with the better risks choosing the more (nominally) comprehensive HMO coverage (Bernard Friedman et al. 1984; Paul Eggers 1980; Paul Eggers and Ronald Prihoda 1982; Marilyn Jackson-Beeck and John Kleinman 1983). Here the Federal HMO Act *requires* groups to provide the HMO option, so there is no presumption that the package is chosen optimally for the group. It appears that the favorable selection in the HMO is due to the necessity to switch doctors to join the HMO, the availability of pediatric benefits that appeal to healthy young families, and the recognition by buyers with a taste against hospitalization that HMOs less frequently recommend aggressive forms of care. By permitting HMOs to pick off good risks, multiple-choice settings could lead to less desirable outcomes for those good risks who really would have preferred conventional coverage at premiums reflective of risk.

There have been some studies that addressed the question of adverse selection by inquiring about *intended* insurance-purchasing behavior. Charles Phelps and M. Susan Marquis (1983) inquired about hypothetical purchase of full-coverage insurance and the subjective expected expenses of some of the health insurance experiment participants, and found a posi-

tive relationship between statements about expected expenses and statements about intentions to buy insurance. In contrast, Bernard Friedman et al. (1984) found little relationship between hypothetical purchases and objective measures of health care use in the recent past. The existence and the magnitude of adverse selection in health insurance still remain open to question.

I. *Expenditure Control by Insurers*

As noted above, moral hazard means that people use medical care whose value to them is less than the cost they have to pay in premiums. Various forms of cost-sharing try to reduce this inefficient use by financial incentives to consumers and doctors, but they all have some cost in terms of greater exposure to risk. An alternative approach is to maintain full coverage but control use by using nonprice methods to ration. The "control" over use would generally take the form of refusal to pay benefits in certain circumstances, so that one can look upon all such cost-control devices as a form of indemnity payment. The control could be even stronger than in a conventional indemnity, by virtue of a refusal to let providers collect anything from patients for unapproved services, or a refusal to pay any benefits for care from unapproved providers or providers who engage in unapproved behavior.

In addition, such controls can be imposed by and at various levels. They could be imposed on an entire market by some public authority; this is what state rate regulation approximates. Or they could be imposed only on those who voluntarily choose to submit to them; this is what an HMO embodies. In this sense, posing the choice as one between "regulation" and "competition" is somewhat misleading. The choice really is between having the form of those regulatory controls determined politically or by the market.

A number of states have imposed some limits on hospital costs or prices, and the evidence suggests that, when these controls have been in place for a while, they do have an impact on the rate of growth in costs (Paul Joskow 1980; Frank Sloan and Bruce Steinwald 1980; Craig Coelen and Daniel Sullivan 1981). The fundamental question that has yet to be answered is—why was this collective action necessary (Clark Havighurst and Glenn Hackbarth 1979)? That is, why did not market forces induce individual insurance firms to supply "cost control" of some type, and so obviate the need for government regulation? The related question is whether tax-side effects might be at work here in discouraging such efforts, and whether the function and the usefulness of public rate regulation might be changed if the tax treatment is changed.

There have been some attempts to pose this question, although many have simply accepted evidence of effectiveness of cost control by government as proof of its desirability. There are three kinds of reasons (and supporting evidence) for the lack of cost control efforts by the private sector.

One answer is a conspiracy or pressure theory. The notion is that providers of hospital and physician services would prefer fewer to more restrictions. Physicians in particular have been alleged to move collectively in overt and covert ways to resist insurer cost controls that exist and to discourage their introduction (Lawrence Goldberg and Warren Greenberg 1977, 1978). It is true that the early history of health maintenance organizations was one of concerted (and eventually illegal) medical society opposition (Reuben Kessel 1958). But such coordinated refusal to deal is not nearly so strong now, and has been further weakened by the recent court decisions (Charles Weller 1983).

A second argument relates to the nature of the industry. As already noted, the bulk of "inflation" in health care has been in

qualitative forms. The ability of consumers to make different choices about quality may be more restricted than their ability to make different choices about quantity. The reason is that the level of quality may be affected when some consumers change their insurance, but the quality level may change for all consumers. More precisely, an insurer that imposes a cost-control program on its insureds, at some explicit or implicit cost to them, may be unable to recoup all of the benefit from lower-service intensity in the form of lower premiums, because service intensity falls for all users of health care, not just for the firm's insureds.

Partly this spillover would be caused by the average cost pricing policies of many hospitals. Attempts to limit the number of nurse hours devoted to the care of a few patients would generally produce a reduction in the room rate charged to all. Partly it is due to the inability of firms, either doctors or hospitals, to offer different quality levels at the same time to those with different types of insurance.

A third explanation finds the cause of undersupply of cost-control activity to be the tax subsidy directly. Suppose consumers can choose whatever level of cost-control activity they want, and suppose that such activities reduce premiums while imposing some nonmoney costs (inconvenience) on them. The conversion from untaxed benefits to taxable financial savings will tend to discourage such efforts.

At the moment, all we can say is that any (or all) of these three explanations could be true (as could a fourth hypothesis, that there are no effective conventional insurance cost-control programs).

J. *Private Sector Indemnities as a Cost-Control Device*

There is a form of insurance that is intermediate between cost-sharing and controls as a way of controlling medical costs. That approach involves the use of *indemnities* of various types. Indemnities can be per unit of care or per illness, can depend on the nature of the illness, and can either be made independent of the amount of actual medical expenses or be limited to no more than those expenses.

The rationale for indemnities is obvious: They face the consumer with a *marginal* price equal to full cost, while still providing protection against risk. Indeed, if it were possible to define illness states precisely enough, an insurer could pay an indemnity equal to the expenditure in the absence of moral hazard (adjusted for income effects), and the result would be first best efficient. While the difficulty of determining exactly the illness state would prevent such an arrangement in general, it does seem feasible to have an indemnity payment that depends to some extent on the nature of illness for at least some illnesses (Pauly 1971). One possible strategy is to have an indemnity that is illness-specific and then, in effect, have a supplemental insurance (with deductibles and possible copayments) to cover the possibility that actual expenses exceed or fall short of the indemnity. Frank Gianfrancesco (1983) has discussed such an arrangement, and noted that the recent development of supposedly homogeneous diagnostic categories for Medicare patients—the Diagnosis Related Groups—may provide a vehicle for appropriate indemnities in *private* health insurance as well.

Such private indemnities could permit the individual to pay more than the indemnity if he wishes. However, some private insurers have expressed interest in concluding agreements with some doctors and hospitals to accept the indemnity payment in full. In return, patients would be channeled in various ways to those providers. These are exactly the Preferred Provider Organization discussed earlier.

Let us consider the bargaining possibilities in detail in examining the unit price. A for-profit enterprise currently charges

$ X per unit of some medical service. Could an insurer alter its benefit policy so that less is paid? One possibility is that it might pay an indemnity less than $ X. The seller could "balance bill" for the excess, but would presumably face increased buyer resistance compared to the alternative policy in which any price was covered. An alternative for the insurer would be to set its payment rate at $ X and *forbid* its insured from paying any additional amount; in order to receive insurance payments, the provider would have to agree to accept $ X as payment in full.

At first glance, such an arrangement appears sure to be unattractive to consumers; why should they wish to prevent themselves from engaging in voluntary transactions? But on further reflection, it is easy to see that if $ X is above the provider's marginal cost but below its current price (e.g., if the provider is currently earning monopoly rents), and if the patient's demand from this firm drops to zero with no insurance coverage, then the provider may accept the offer rather than lose the business, particularly if patients covered by this insurer represent a nonnegligible share of his total business. Here we see the paradox of "competition"; this strategy will be most successful when the insurer has a large enough market share to give it market power in the market for medical care—but then it is also likely both to have monopoly power in the market for insurance and may exercise its power in ways that do not advance consumer welfare (Pauly 1985).

K. *The Market for HMO Insurance*

If the benefits from cost-control activities under conventional insurance are dissipated over the entire fee-for-service market, then one would have an additional explanation for the emergence and gradual growth of the Health Maintenance Organization (HMO) and related concepts. This explanation would focus not on financial incentives or the use of preventive and ambulatory care, but rather on the closed-panel nature of the prototype HMO, as prepaid group practice. In such an arrangement, the individual pays a premium for ostensibly complete coverage. However, the care has to be produced in a set of firms controlled, so to speak, by the insurer. Hence the full benefit of any insurer activities to control costs are fully reflected in lower costs for the plan, and presumably ultimately lower premiums.

The form that HMO cost control actually takes is quite well documented. HMOs have lower total cost only because they have lower hospital admission rates than other insurers (Harold Luft 1981). They have not had lower rates of inflation. From the early 1960s up to 1981, HMO costs have increased at approximately the same rate as those in the private sector (Luft 1980; Newhouse et al. 1985). HMOs also do not appear to have more efficient production of a given batch of services, or healthier patients (Luft 1981).

How the lower hospital admission rates are brought about is much less clear. There is some evidence that salaried or profit-sharing closed-panel practices display larger savings than Independent Practice Associations, which typically include virtually all of an area's physicians on a fee-for-service basis. But whether use is a response to direct profit sharing, the mere absence of positive payment for additional use, the staffing pattern of the HMO, or simply the consequence of group practice per se (Anne Scitovsky, Lee Benham, and Nelda McCall 1979) is not known.

There is also evidence, however, that at least some of the reduction in use comes about through self-selection: As discussed above, people who join HMOs typically had lower than average medical expenses in the fee-for-service system (and hospital admission rates) even before they joined.

The admission rate does eventually rise even for them, but they may be self-selected to have a permanent taste against aggressive, hospitalization-oriented care. Willard Manning et al. (1984) find, however, that even when individuals are randomly assigned to free (conventional) care and an HMO, there is still a sizeable reduction in hospital use and cost associated with the HMO, on the order of 40 percent for hospital admissions and 25 percent for total expenditures. Such a result does not hold, however, when the HMO is compared with a less comprehensive conventional insurance. Results from the experiment vary depending on which type of conventional insurance with deductibles and copayments is provided, but overall Manning et al. conclude that the saving from HMO choice compared to conventional insurance is "near the low end of Luft's (1981) range" of 10 to 40 percent.

In addition, although the Manning study found that persons assigned to the HMO had the same expected expense as those who joined voluntarily, their study failed to settle the question of selection. The selection argument as applied to HMOs is that those who choose the HMO would be at lower risk in a fee-for-service situation than those who do not. A direct test of this hypothesis would be to assign some HMO members to conventional insurance, and compare their expenditure levels with those who voluntarily retain conventional coverage. Instead, the experiment compared voluntary and assigned members of the HMO, which is not the relevant comparison (Pauly 1986).

There has been relatively little research that bears directly on the question of how consumer or employee group choices of HMOs versus alternatives would be expected to respond to changes in tax treatment. The most important empirical fact is that the typical HMO premium equals or exceeds a typical high-option conventional insurance premium, even though total expenses under the HMO may be lower (Ginsburg 1981). Consequently, it is not generally the case that the tax treatment distorts the choice in favor of an HMO relative to a conventional high coverage option, because the employee must often pay any extra HMO premium with taxable dollars. Taxing employer premium contributions above a certain level will probably reduce the demand for HMO insurance for those firms that already offered an HMO option, precisely because more of the HMO premium will be subject to taxation. However, there is also reason to believe that more options will be added, which should benefit HMOs. If HMO premiums more frequently fall below those for conventional insurance, as they now seem to be doing, and if the amount of premium subject to tax is not set at a level below the HMO premium, then there will be some additional incentives to select HMO coverage.

In addition to tax-side effects on the quantity of insurance chosen for those firms that offer just one insurance plan, the tax treatment of insurance may also affect the number of options offered. If the employee is required to make a contribution to the plan, that already implies that the option of no coverage (and therefore a saving of the employee's contribution) is available. In addition, other levels of coverage and one or more HMOs may be offered. Indeed, the federal HMO Act required all employers with more than 25 employees to offer an HMO option if a qualified HMO is available, but Ginsburg (1981) speculates that this rule is not typically enforced.

In principle, the level of the employer contribution to multiple-option plans can be varied in several ways. According to Farley and Wilensky (1983), the option of choice is rare—available only to 17 percent of all workers when choices are available. The typical pattern appears to be

for the employer to contribute a fixed share to each of several plans.

There has been little research directed at explaining the decision to offer multiple choices. Suresh Maholtra et al. (1980) found the presence of multiple-option plans to be related to indicators of work force heterogeneity, such as the fraction of part-time workers. If single-option plans are chosen in order to protect tax deductibility, it should also be the case that multiple-choice plans should be more likely to occur in groups where the marginal tax rate is low.

One implication of tax deductibility is that, if the employer does choose to offer multiple options, tax considerations alone may lead him to cross-subsidize between low-coverage and high-coverage plans. (As noted above, inhibiting adverse selection may be another motive; in practice, it may be difficult to attribute causation to one motive or the other.) The reason is that, in the absence of cross-subsidization, the maximum amount that could be shielded from taxation, for any employee, is the premium for the least costly option.[10] However, if the employer makes larger (tax-free) contributions to the more expensive options, employees who prefer those options receive larger tax benefits. Depending on labor supply as a function of fringe benefits and tax subsidies, there may be a large enough fraction of workers who benefit from such a procedure to make it worthwhile.

Removal of tax deductibility should increase the number of multiple-choice offerings. It is also possible that more options will mean the opportunity to buy more coverage, especially the more discretionary coverage, but this seems unlikely. Even now a surprisingly large number of employees choose the least expensive option, even though the premiums they save are much often less than the reduction in average expected benefits (Farley and Wilensky 1983).

The impact of removal or limitation of the tax subsidy in a group insurance market is probably different from the simple individual-choice model, precisely because the tax subsidy brought about an increase in both the quantity and the uniformity of coverage. Of course, if there is perfect sorting, in a full Tiebout world, there will be no difference. But if there is some failure of the tax-subsidized market to satisfy individual worker preferences exactly, removal of the subsidy may cause those preferences to be more closely satisfied, because eligibility for the subsidy often requires some limitation of individual employee choice. The estimates of the impact of the tax subsidy on demand discussed above provide a conjecture for the worker who is in some sense decisive. But if people have different preferences for insurance and if marginal tax rates differ, then different workers and groups of workers may respond differently.

For firms for which the administrative cost-reduction effect of group purchasing is strong, mandatory additional options will be unlikely and purchase of the single insurance policy is likely to remain after removal of the tax subsidy; firms are not, after all, required to offer choices. Small to medium-sized work forces are more likely to retain the single plan. A sufficiently large work force can, however, offer several options and still obtain the administrative cost advantages of group coverage.

The only estimates of the impact of tax changes on the level of insurance purchased by persons who initially received group insurance unfortunately are based on data concerning individual insurance purchases. These analyses in effect assume that removal of the tax subsidy will bring about the loss of any of the advantages

[10] The presence of a cafeteria plan weakens this incentive. In such a plan, an employee may choose to shield income by means of salary reduction in order to obtain a more comprehensive plan.

of group purchase and would require the offering of a full range of individual options to employees. Chollet (1984), for example, assumes that, absent a tax subsidy, not only would workers purchase all insurance individually, but also their purchase patterns would match those of people who currently buy individual coverage. Not only does this ignore the much lower administrative cost of group insurance, a cost and price advantage that would probably be enough to cause some group coverage to be retained even without a tax subsidy, it also ignores the possible sorting of workers based on demand for insurance into jobs offering high or low (or no) group insurance. James Lee (1979) uses the even more atypical sample of workers who lost group coverage through unemployment and who therefore use group continuation coverage, which permits individual choice but some of the advantages of group rates. Estimations based either on the purchases of people with individual coverage or group continuation coverage suggest large declines in coverage; people who do buy individual coverage do buy less coverage than those in groups at all income levels, and the shortfall is particularly pronounced at low income levels. These types of estimates have, therefore, been used to suggest that, if the tax subsidy is ended, low-income workers who currently have generous group coverage may cut back so much that there will be strong political pressure for direct subsidization of their medical care and insurance purchases.

While that message seems a bit overstated, especially if one is considering only a cutback in the tax subsidy above some level, there is an important qualitative result in these analyses that might well be replicated in more realistic estimates. By removing the incentive to uniformity of coverage, coverage may be cut by a larger amount than average in some segments of the work force. The individual coverage data suggest that family income may be inversely related to individually desired coverage. Although removal of tax deductibility will have a smaller than average effect on the implicit price of insurance for low-income people, it may permit them greater opportunity to choose the lower levels of coverage that they prefer—in effect, it removes what is currently a kind of mandatory purchase financed by a head tax. However, the availability (since 1978) of cafeteria plans mitigates this uniformity effect.

Would such reductions in coverage lead to inefficiently low levels of coverage by low-income workers? It is possible that lower-wage workers would choose less extensive insurance coverage when they are free to do so. If there are no social concerns about their level of coverage or use of medical care, however, this choice will still be efficient. It is likely that, at some level, there are altruistic externalities attached to consumption of medical care (Pauly 1970), and so it is possible that levels of coverage may sink so low that these externalities are manifested. A tax-financed subsidy would then be needed, and the cost of this subsidy would offset to some extent the tax proceeds from increasing the level of income subject to tax. Of course, if tax deductibility is simply limited to an amount that is still high enough to buy adequate coverage, and if groups tend to cut back to that level, then there will be no offset. Moreover, the distributional consequences of the current system approximate a set of lump-sum per employee taxes offset by a tax subsidy that varies directly with the marginal tax rates. Therefore low-wage employees currently pay themselves for the "excess" coverage that satisfies social preferences, and they pay according to a very regressive tax structure. In any case, the information on what groups would do, what low wage individuals would do, and what "society" would do in response is presently un-

known; all of these questions would appear to be important.

V. *The Impact of Insurance on the Demand for Medical Care*

A. *Introduction*

The previous section described how the level and form of health insurance coverage is determined and how that level might be affected by changes in the tax treatment of insurance. Altering the level of insurance coverage will affect the demand for and cost of medical care, because of moral hazard. There have been a large number of attempts to estimate the size of this effect.

Rather than go into the detail required to explain the full set of empirical results, a task that has already been performed by others (Feldstein 1977; Newhouse 1978c, 1981), I will limit my summary here to a description of the broad outlines of results and techniques, and a summary of the most recent and most definitive empirical results, especially those from the Rand Health Insurance Experiment.

B. *Estimating the User Price Elasticity from Nonexperimental Data*

Estimation of the user price elasticity of demand from nonexperimental data was the initial procedure. The first estimates came from data aggregated to the market level or even larger areas, such as states; the measure of user price was related to the average percentage of expenses covered by insurance. The results were consistent with moderate to high effects of insurance coverage on use and expenditure (in the range of −0.4 to −1.0), with a larger elasticity for physicians' services as compared with inpatient hospital services. There also appeared to be negative cross-price effects between ambulatory and inpatient care.

Newhouse, Phelps, and Marquis (1980) argue that using an average price measure in such aggregated data, rather than the correct marginal price measure, leads to estimates of price elasticity biased away from zero.[11] Individual nonexperimental data, as analyzed by Phelps and Newhouse (1974), and Newhouse and Phelps (1976), and most recently by Wilensky and coworkers using the detailed information from the recent National Medical Care Expenditure Survey (Taylor, Wilensky and Rossiter 1981), do indicate price elasticities in the range of −0.1 to −0.2 overall; these figures are lower than the earlier estimates from aggregated data such as those of Martin Feldstein (1971) or Richard Rosett and L. F. Huang (1973).

C. *Estimation of Price Elasticity from Experimental Data*

However, in order to assure that the level of insurance coverage is exogenous and at the same time has both a precise and detailed measure of coverage, the federal government has funded the Rand Health Insurance Experiment for the last decade. The $74 million experiment randomly assigned a sample of individuals to 14 different insurance coverage levels, using financial incentives to reward participation. In order to induce families to accept the low-coverage plans, they were given at the outset a cash payment equal to the maximum possible additional out-of-pocket payment. The maximum out-of-pocket payment was limited to $1,000 per family (less for the poor), so that there were unlikely to be large income effects. The experiment provided a variety of different levels of coverage, with different patterns of copayment and deductibles (but all with an upper limit on out-of-

[11] It is the common presence of deductibles that primarily causes this effect. When a deductible is present, a large percentage change in the marginal user price will tend to be associated with a smaller percentage change in the average user price. When average price is used, it will then appear that the smaller change is associated with a given change in use.

pocket expenses) to several thousand families in several sites in the U.S.

The results of this experiment indicate substantial effects of copayments on use. For medical care, expenditures with complete coverage are about 50 percent greater than under the least generous plan (Newhouse et al. 1981). For dental care, a fall in the coinsurance rate from 95 percent to zero is associated with a 56 percent increase in expenses, with an even larger surge in demand during the first year of more generous coverage. The price elasticities implied by these estimates are close to the −0.1 to −0.2 range obtained earlier by Phelps and Newhouse (1974).

Expenditures conditional on initiation of an "episode of care" (by a visit to physician or hospital), and changes in the frequency with which episodes of care were initiated (Emmett Keeler et al. 1982), were also analyzed. For dental care, there was an increase of 54 to 70 percent in the probability of use of any dental services and a small but statistically significant tendency to use more expensive forms of care. For medical care, in contrast, there was no statistically significant effect of coverage on expenditures per episode of care; all of the impact of insurance on expenditures, an impact that appears in both physician and hospital expenditures, appears to come from increases in the frequency with which episodes of care are initiated. Later work by Marquis (1985) also indicated no effect on the price level or expensiveness of the provider selected. One should not, however, conclude therefore that care after the episode is begun is unaffected by the patient's insurance coverage, perhaps because it is wholly under physician control. The average severity of illness for which care is sought may also be affected by insurance coverage, so that higher copayments screen out less severe illnesses. The resulting increase in average severity of treated episodes may offset any insurance effects on the intensity of treatment of an episode of given severity.

Most episodes of care are initiated by visits to a physician's office or hospital outpatient clinic, but can result in hospitalization. Rand results suggest—contrary to frequent suggestions—that a deductible applied only to ambulatory visits, far from encouraging hospitalization as a (more expensive) substitute for ambulatory care, actually discourages hospitalization. Ambulatory care and hospital inpatient care, on balance, appear to be complements.

A final result of considerable importance from the experiment is the failure to detect any strong negative effects on health of higher user charges, as opposed to free care (Robert Brook et al. 1983). Out of more than a dozen health measures, only vision and blood pressure (especially for persons at elevated risk) were affected, and then only by a small amount.

The Health Insurance Experiment's strongest results are for the comparison between fully comprehensive conventional insurance and a varied set of plans with copayments and deductibles. Here there is a (practically and statistically) significant effect of free care on expenditures, but a small effect on indicators of health. However, such a system of free care combined with fee for service is not common even in the most heavily tax-subsidized insurance plans. The experiment is unfortunately less definitive in indicating whether these are significant effects on expenditures as deductibles and copayments are increased, although the estimates are in the expected direction. The point estimates do suggest, however, that the effect of user price on use is nontrivial.

On a priori grounds it seems unlikely that out-of-pocket payments of the conventional deductible-copayment sort can have much of an effect on inpatient hospital expenses per admission, because most practical schemes would still leave the in-

dividual facing a zero user price at the margin (Phelps 1982). And yet virtually all of the hospital cost "inflation" has come about via higher costs per admission, and not from changes in either average stay or the hospital admission rate. In this sense, neither HMOs nor deductible-copayment strategies seem targeted well at what has historically been the fundamental source of "inflation." Of course, cutting admission rates by enrolling more people in HMOs or by discouraging more hospital and ambulatory care use with deductibles and copayments might offset the inflation in cost per stay, and such measures do seem to have been primarily responsible for the slowing of hospital cost increases in 1984 (American Hospital Association 1985). The important question here is whether cuts in volume can continue to offset increases in unit costs and prices.

D. *Quantity and Quality*

The framework for virtually all of the empirical work on insurance effects on demand for medical care has been the very simple model of insurance coverage of a homogeneous medical service. In such a model, insurance cuts the user price and thereby increases the physical quantity of services used. However, it is quite obvious that virtually every medical service has a qualitative as well as quantitative dimension; indeed, the issue of how to measure "quantity" often required arbitrary definitions of "quality." In addition, it is undoubtedly true that many aspects of quality are endogenous, so insurance will affect the market equilibrium level and distribution of qualities. This choice can result either from a provider's selection of a particular level of quality, or by entry and exit of firms with different intrinsic levels of quality. The most realistic model is surely that of monopolistic competition.

A possible impact of insurance on the market level of quality raises some question about the generalizability of price elasticity estimates based on the Health Insurance Experiment. However the market pattern of quality is determined, it is unlikely that this pattern (and the associated cost) would be much affected by changing the insurance coverage of a tiny fraction of a market's population, as the experiment did. In contrast, if tax-side changes should reduce the level of coverage for almost everyone, there might well be changes in market-level quality, and therefore cost changes that are larger than those indicated by the experiment.

Existing models of the impact of insurance on market-wide hospital quality are not well developed. Feldstein (1971, 1977) investigated the effect of insurance on the desired quality level for hospital care, but only in a model in which a single hospital chooses quality subject to an arbitrary firm demand curve; there is no consideration of interhospital competitive equilibrium. Jeffrey Harris (1979) and Pauly (1980a) have investigated the impact of insurance coverage on qualitative equilibrium in theory, but there has been no effort to determine it empirically.

There is one case in which the use of individual data to estimate the impact of insurance coverage on use and quality would be appropriate: the case in which a wide variety of quality offerings are available to individuals, who then can choose among them (Sherwin Rosen 1974). Such an "à la carte" world is to be contrasted with one in which quality offerings are limited. While there may be some markets that are examples of à la carte medical quality, in general quality offerings are likely to be limited, especially in the case of hospitals.

The reason why an individual hospital may choose to provide the same level of service intensity or quality to all buyers, even if some of them might prefer different levels, is the presence of fixed costs associated with supplying each level of product variety. In some sense, the hospi-

tal then responds to the average desired level of product quality; if the desired level is determined primarily by insurance coverage, it responds to the average (or perhaps the modal) level of coverage of demanders in its market. In many hospital markets, there are only a handful of hospitals; even in large metropolitan areas limitation to local neighborhood markets and the possibility of a symmetric monopolistically competitive equilibrium may combine to produce a situation in which consumers have a limited range of quality choices. This implies that the level of expense incurred by individuals in different markets may well depend both on their own level of insurance coverage and on some measure of the predominant level of coverage in their market area.

This dependence on the predominant level of coverage is by no means necessarily limited to hospital service intensity. Once the individual initiates care for an illness or a problem, the care is managed by a physician or hospital. The number of visits a doctor recommends, or the number of hospital days a person stays in a given hospital, also represents a kind of dimension of product quality, one for which there may be fixed costs associated with offering different levels by the same physician. So the measure of the quantitative response of use to variation in coverage may be different depending on whether a few individuals' insurance levels are changed or whether the predominant level of coverage is changed.

This kind of spillover process has been discussed in the health economics literature under the name of the norms, demonstration, or bandwagon effect (Newhouse and Marquis 1978; Thomas McGuire 1981; Pauly 1980b). Its theoretical relationship with monopolistically competitive product differentiation, and the use of models based on that theory, has yet to be widely recognized.

While the theory that should be used to explain and evaluate these impacts of insurance coverage in markets with quality that is potentially heterogeneous has barely been developed, there has been some ad hoc empirical testing of the norms hypothesis. The intuitive explanation usually offered is that individual physicians, in deciding what courses of action to recommend for each patient, may not tailor their advice entirely to the insurance coverage of that patient. Instead, they will tend to follow the norms for care in the community. Those norms in turn will be related to the predominant or average demand in the community, and hence will be related to the average or predominant level of insurance coverage. The empirical test then is to see whether measures of the use of medical care by an individual household tend to depend not only on the household's level of insurance coverage but also on the prevailing level of coverage in the community. There is some empirical evidence that such an effect can be observed (McGuire 1981; Pauly 1980b).

A relevant but overlooked consideration is the impact of insurance-induced changes in the quality of medical care on the equilibrium level of insurance purchased. Consider a situation in which there is a uniform impact of user price on consumers' desired level of service intensity, but in which consumers desire different levels. In an à la carte market, in which there are many options, each purchaser of insurance will be able to obtain his desired level, and therefore the marginal insurance premium will reflect this increment in qualitative use and cost. But in a one-hospital model such as Feldstein's (1971) (or in a symmetric monopolistically competitive equilibrium), a change in insurance coverage for the customers of a single insurance firm would have an imperceptible effect on the level of quality, and therefore on the level of hospital cost attributed to them. Insurance premiums

as a function of the level of coverage would then show a much smaller increase in marginal price than in the à la carte model. Even though each individual pays such a nonincreasing marginal price, insurer solvency requires that the average premium paid cover hospital costs. Consequently, each individual will tend to ignore the qualitative moral hazard associated with his purchase of insurance, and so insurance coverage would be excessive, even in a second best sense. Not only can there be a cost-insurance spiral, but the equilibrium level of insurance coverage can be quite high relative to the à la carte case.

In the in-between situation where the number of hospitals is larger than one but fewer than the number needed to provide à la carte choice, the response of insurance premiums to increases in coverage is, as for all models of monopolistic competition, theoretically indeterminant. It is surely possible, however, that the level of insurance coverage is excessive even when there are no tax distortions because of the failure of the experience of small subgroups of individuals to reflect the full incremental cost of services. That is, marginal premiums are ultimately based on cross-sectional information, but that information understates the market-wide effect of insurance on quality and cost. An insurer who bases his premiums on the cost experience of different insureds at a given point in time will fail to take into account the market-wide effects of coverage on quality and therefore on cost.

In summary, if insurance displays qualitative moral hazard but the market for the service to which the moral hazard applies is not the simple perfectly competitive, homogeneous product case, then an ordinary demand elasticity (of quantity) will not be a satisfactory summary of the impact of insurance coverage on total expenditures. Equally important, these characteristics of the market for the service will impact back on the market for insurance—even if the insurance market itself is competitive—and can seriously affect the equilibrium amount of insurance purchased at any net loading. In the case of the uniform-quality equilibrium with qualitative moral hazard and, for example, it may well be that insurance purchasers will buy virtually full coverage even if moral hazard is strong.

E. *Quality Change and the Technological Imperative*

This type of interaction between the market for medical care and the market for insurance has not been well developed in the literature. Feldstein (1973) does discuss a "vicious circle" in which more insurance pushes up the level of medical care costs, and people then respond to those higher levels of losses by purchasing more insurance. He also considers the possibility that, if demand for quantity is increased when quality rises, there may not be an equilibrium (1977). In both cases, however, these issues seem to be more those of the process of moving from one long-run equilibrium to another, and whether that equilibrium is stable, rather than the properties of the equilibrium as such.

A more direct consideration of similar types of questions has been provided by Newhouse (1978a). He focuses on the market for medical care, but does not assume static quality. Instead, he contrasts a conventional competitive (homogeneous product) medical market under insurance with two other models. In one, insurance coverage affects consumers' search for lower prices (but still with homogeneous goods). In the other model, the rate of technical change in quality depends on the level of insurance coverage. That insurance does indeed affect that rate has been shown by Louise Russell (1979).

Newhouse finds it difficult to distinguish

among these three models empirically, in part because of data limitations and the problems of constructing an adequate econometric model, and does not ask the further question how the equilibrium level of insurance coverage is affected by the type of medical market. He does find that, consistent with a model in which higher levels of insurance coverage encourage cost-increasing technical change, there is a suggestion of a relationship between the level of insurance coverage and the rate of change in prices and expenditures. Unfortunately, this relationship also turns out to be consistent in theory with a conventional model, a linear demand curve for medical care, and an environment of increasing levels of insurance coverage.[12] Newhouse also does not consider whether this endogenous technical change can be bought (and covered by insurance) on an à la carte basis, or whether its effects apply to all consumers regardless of their level of insurance coverage—an important consideration in explaining the equilibrium level of insurance coverage.

This question of qualitative supply-side responses to variations in insurance coverage has not been investigated empirically. One implication of such a theory would be that market-wide shifts in the level of insurance coverage might have quite different effects depending on the degree of competition in the particular market. This implication has not been studied directly, but there have been a few preliminary looks at the more general questions of the impact of competition in medical markets. Paul Joskow (1980) found that areas with more hospitals tended to have higher reserve margins of available empty beds. A similar association of greater competition with higher prices was also found for nuclear medicine services by George Wilson and Joseph Jadlow (1982). These results were said to be consistent with the theory of "quality competition" in price-regulated industries, in which firms forbidden to charge lower prices offer higher quality when there is more competition. The theory does not apply directly to the hospital industry, because the prices there are not usually regulated, at least not for all patients. There may be an analogous theory when there is full insurance coverage, a maximum price that is controlled by an insurer, and variation in the degree of competition.

[12] If one could ever observe a situation in which the level of coverage were constant over time, and all lags to past changes had worked themselves out, then the conventional model would suggest no further inflation, while the endogeneous quality change model would be consistent with a rate of growth in costs that is higher for markets with higher levels of coverage.

VI. *Supply Responses to Insurance-Affected Demand*

A. *Introduction*

From the foregoing it is clear that demand can be distorted by tax-side influences, and that the impact of this distortion in part depends on the characteristics of the market for care itself. A further research question, however, is whether the supply behavior of producers of medical care also has direct and potentially welfare-adverse effects on market equilibrium. That is, would there still be correctable market failure emanating from the supply side even if insurance and medical demand could be properly structured?

One reason has already been discussed in detail: the possible responses of a monopolistically competitive market to insurance coverage. There are two broad classes of additional reasons why supply responses in medical markets may differ from those in conventional markets: (1) imperfect consumer information may affect supply responses and (2) public insurance or publicly affected private insurances may pay providers in perverse ways.

B. *Information Imperfections*

Consumers of medical care are not perfectly informed; indeed, often they purchase medical care in order to become better informed. Insurance coverage interacts with consumer information imperfection. The most obvious form of interaction concerns the impact of insurance on search. Even if search costs are modest, the insured consumer may be unwilling to search for lower prices because he obtains at best only a fraction of the benefit from finding a lower-priced seller (Newhouse 1977). When the quality as well as the price of medical outputs varies, and the consumer is imperfectly informed about both, outcomes are more complex. It is also possible that some of the medical care costs covered by insurance themselves represent a kind of search—if insurance covers the cost, it is cheaper to try out a doctor you may like—so there can even be more intensive search, a point made by Georges Dionne (1984).

In addition to interaction between insurance and search, there is the question of the effect of information more generally, and whether public sector activities can improve matters. One place where a positive answer has been given is in the case of advertising restrictions imposed by professional associations. The work of Lee Benham (1972) and Roger Feldman and James Begun (1978) indicates that advertising restrictions favored by physicians and other professional associations tend to be associated with higher prices. But why? Benham argued that advertising permitted the firm to reap economics of scale, something that was true of the optometrists he studied but not necessarily true of health professional services in general. Feldman and Begun indicate that the price declines were associated with reductions in quality, but that quality-adjusted price also fell. A more recent study by John Kwoka, Jr. (1984), using detailed, explicit measures of quality, confirms that quality-adjusted price is reduced by advertising; Kwoka finds no evidence that the time price of care by all firms (a quality measure) is reduced when advertising restrictions are loosened.

A second aspect of consumer information has been raised by Mark Satterthwaite (1979) and by Pauly and Satterthwaite (1981). They argued that prices should be higher in markets in which information about price was difficult to obtain. One does not ordinarily search physician prices by going from office to office. Instead, the evidence in the literature suggests that much of the information about particular physicians—their price and idiosyncratic qualities that may appeal to one person but not to others—comes from friends.

Pauly and Satterthwaite's empirical work indicates that factors associated with the easier access to accurate information from friends are associated with lower prices for (primarily uninsured) primary care physicians services. Some of the proxy measures for access to information are fairly obvious: the number of persons moving to the area, and the fraction of female-headed households (because sociological studies indicate that mothers in such households have less social interaction than average). One proxy for high search cost that is not so obvious is the total number of physician sellers. The notion here is that each friend provides advice based only on his experience with a few physicians, and the smaller the number of physicians in the area, the more accurate and more confident an estimate an individual can make of any one physician's price or quality. Consequently, an increase in the number of sellers can reduce the elasticity of each firm's demand curve, and lead to an increase in price. Given the number of sellers in the market area, Pauly and Satterthwaite found that an increase in the physician stock relative

to demand—as measured by the number of physicians *per capita*—was associated with a lower price, just as one would expect if firm marginal cost curves sloped upward, while an increase in the *total number* of sellers was associated with higher prices.

This explanation of the impact of imperfect information on physician behavior is in contrast to another much debated hypothesis—the demand creation hypothesis. This hypothesis proposes (*a*) that physicians can and do alter the content of the advice they provide to patients in order to increase the quantity demanded (and their net incomes) at any price and (*b*) that physicians alter advice to a greater extent when the number of physicians per capita is high.

To have an interior solution under this hypothesis, one must suppose that physicians suffer some cost (usually subjective) in terms of lost "professional integrity" from altering advice, but that they are willing to substitute income for integrity (Robert Evans 1974; Frank Sloan and Roger Feldman 1978; Pauly 1980b). The primary evidence in favor of this hypothesis is the finding, in some empirical studies, that the use of medical care is related to the stock of physicians, when user price and other demand determinants are held constant. For example, in Victor Fuchs' (1978) study of surgical services, the elasticity of the surgery rate with respect to the number of surgeons per capita (treated as endogeneous) was about 0.27, and similar results were obtained by Janet Mitchell and Jerry Cromwell (1981). In addition, the finding of a positive relationship between physician stock and price is also sometimes interpreted as evidence favorable to demand creation, although there are some theoretical and econometric problems with doing so (Sloan and Feldman 1978; Richard Auster and Ronald Oxaca 1981; Jerry Green 1978).

There is considerable controversy about this phenomenon, however. For one thing, researchers using different data sets or different independent variables do not find evidence of a positive relationship. For example, Pauly (1980b) incorporated measures of patient health status into the demand equation—something most other researchers were unable to do—and found no effect of the surgeon stock on the number of surgically treated hospital episodes. Gail Wilensky and Louis Rossiter (1981) likewise find little evidence of an impact of increased physician stock on physician-recommended care. The most recent attempt to detect inducement, by Thomas McCarthy (1985), found the negative relationship between physician per capita and price that would be expected from conventional economic theory, rather than the positive relationship that is said to indicate demand creation. As Uwe Reinhardt (1985) has noted, however, there is still need for an explanation of the well-documented inter-area evidence on variations in surgery rates (John Wennberg 1984), variation both too large to be explained by price or quality and apparently not well related to proxies for physician incentives.

A second problem of interpretation of positive empirical results, as Newhouse (1978b) has pointed out, is that one can never rule out the possibility of unobserved changes in quality related to the number of physicians per capita. Some such changes are virtually inevitable. With more physicians relative to demand, it should be easier to obtain an appointment, travel distance and delays will be shorter, and visits may be longer but more pleasant. In such a case, it would not be surprising that patients would find more frequent visits attractive.

It may also be that the physician who debases "true" quality by giving inaccurate advice may eventually find his demand shrinking, once this action is understood by consumers. That is, a concern for

reputation may inhibit physician willingness to distort advice for short-term gain (Peter Coyte 1982; David Dranove 1983). The critical question then becomes the consumer's ability to monitor the reputations physicians develop for unbiased advice. When the number of physicians in the market area is small, it may be easy for consumers to detect systematic inaccuracy, so the physician may—even over a range of levels of physicians stocks—be in the "corner solution" of high accuracy (Pauly 1980b). At the other extreme, where the number of physicians is very large, as in a big city, inaccuracy may be very difficult to detect.

In summary, the existence of demand creation remains an open question; recent empirical work fails to find evidence consistent with the joint hypotheses that physicians can create demand and that they will do more of it in response to economic incentives. In contrast, there is strong evidence that consumer information matters for the market. As is usual with imperfect information, however, it is virtually impossible to tell with any confidence how public policy might improve market functioning. Even removing professional society bans on advertising, which does seem to lower price, may make some consumers worse off.

C. *Public Insurance*

The final sort of market imperfection, in addition to tax distortion of the demand for private insurance, would arise directly from public sector activity, specifically from the reimbursement and payment policies followed by public insurance. With the public sector furnishing nearly half of all funds for this industry, its influence is obviously more than marginal.

When Medicare and Medicaid were initiated, they followed the usual practice of Blue Cross, paying hospitals on the basis of costs attributed to beneficiaries. They paid physicians in a new way, however, based on the distribution of actual physician charges rather than according to a fee schedule as Blue Shield and commercial insurers had done. In short order, however, virtually all of the Blue Shield plans and most of the commercial insurers adopted policies qualitatively similar to the government's.

It has been contended that both of these methods of payment have exacerbated upward pressures on costs. The physician payment method in principle causes payments from public insurances to increase whenever charges to private patients are increased. In practice, however, the rate of increase in Medicare physician payments has been held to a level below that of physician prices in general since 1974, while Medicaid programs have drastically reduced what they pay physicians. As a result, for most practical purposes physicians have been paid according to a payment schedule.

In contrast, hospitals had (until October 1983) been paid on the basis of costs incurred and allocated to Medicare patients, with only very generous upper limits being set to those costs. It is suspected that the consequences of this cost-plus or retrospective method of paying for hospital care—which is shared by many (though not all) Blue Cross plans—causes costs to be higher than they otherwise would, and to increase at a more rapid rate.

The empirical evidence in support of such a comonsense proposition is surprisingly scant. For example, cross-sectional comparisons of hospitals with various proportions of cost-based reimbursement show no significant differences in total cost (Mark Pauly and David Drake 1970; Karen Davis 1972) and the experiments with results from prospective payment likewise were ambiguous (Fred Hellinger 1976); however, the experience with state hospital rate setting does suggest that fixed and controlled prices eventually force costs back into line.

In October 1983 the Medicare program began a four-year transition to payment of hospitals on a prospective per-admission basis. Any Medicare admission must be classified into one of 468 Diagnosis Related Groups. Each DRG has a relative price weight or "points," which are multiplied by a standardized price per point to determine the actual price paid for an admission in each group. There is to be some adjustment of the point price for differences in local wage rates, hospital location, and teaching intensity, but payment is not affected by the level of a particular hospital's cost or resource use.

The stated objective of the DRG program is to provide a financial incentive to hospitals to constrain costs per admission. The unstated objective is to free Medicare from blank check financing, and instead permit it to limit the rate of increase in its hospital payments in a way that was not possible when Medicare was required to pay whatever costs a hospital incurred. With a fixed price, the government can (and does) control payments by controlling the permitted rate of increase in the standardized price per "point."

The introduction of the DRG system, passed by Congress with little debate, has prompted a good deal of speculation about its possible effects. Among these effects, the most important seem to turn on the following questions:

1. What changes in real resource use for Medicare patients will occur?
2. What changes in pricing for non-Medicare patients—so-called cost-shifting—will occur?

Past research provides little guidance. There have been examples of per-admission indemnities—for example, so many dollars for a normal delivery. But these indemnities never covered all patients and never contained a prohibition on additional hospital charges, as the DRG system does.

Under the DRG system, for-profit hospitals would presumably reduce expected resource use for any patient at least to the level of the DRG payment or below, but subject to competitive constraints on quality reduction. For example, because extra days of stay cost something (though probably less than incremental charges), hospitals will gain financially by reducing them, but will be constrained by patient unwillingness to burden finances or family by convalescence elsewhere. If (as appears to be the case) there are constant returns to scale in hospitals, then the for-profit hospital, however much it might wish to shift costs to other payers, would find it impossible to do so (Jack Hadley and Judith Feder 1984). Prices for charge-paying non-Medicare patients would already have been set at the profit-maximizing level, and changes in the price of other unrelated outputs would not change the level of this price. Finally, those patients for whom expected expenses exceed the Medicare payment would be turned away.

With not-for-profit firms, all of these crisp conclusions disappear. The not-for-profit hospital with average cost initially below the DRG payment level (which should be about half of all hospitals) will not necessarily wish to depress costs further, because it may get little benefit from making profits. Indeed, it might find it desirable to push up its costs and quality to enlarge market share or just to increase its utility from quality. The not-for-profit firm may also raise its price to other patients if it was not initially at the profit-maximizing level, a kind of regressive tax to finance Medicare patients (Meyer and Johnson 1983). It may not reject unusually expensive patients if it expects to break even overall, although presumably it will eventually need to do so (Newhouse 1983).

Preliminary empirical results suggest that DRGs made a difference. The length of stay for patients over 65 dropped precipitously after the introduction of this

new payment mechanism (American Hospital Association 1985), although a similar fall was apparent for people under 65. It is too early to tell how much of this change—and other striking declines in the admission rate per capita and in the frequency of low-severity admissions—can be attributed to other causes, such as the upsurge in copayment features in private insurance. It seems likely, however, that this dramatic change in the payment environment will provide a natural experiment for testing the different theories of hospital (and physician) behavior. Analysis of the impact of this step toward more market-like arrangements may help in answering other questions, including that of altering the tax treatment of health insurance.

VII. *Conclusion*

The nature and importance of the distortion from the tax subsidy to health insurance is generally agreed upon, although its magnitude is still open to considerable question. Because the medical market would not approach perfect competition—the best it could do is monopolistic competition—even in the absence of tax-side distortions, second best theory tells us that we cannot be sure that removal of the distortion will improve welfare. This argument has been made by Jeffrey Harris (1979), who finds a potentially offsetting distortion in the public goods nature of biomedical research. But there is little empirical evidence to support this possibility, so the demand-dampening effect of decreased insurance coverage caused by a lower tax subsidy would probably be viewed to be favorable. In addition, the distributional effects of ending or curtailing tax subsidies would generally be regarded as desirable.

There is a possibility that tax subsidies may offset externalities of some type that attach to health insurance. This notion is present in the explanation offered by politicians: "These exclusions are based on a desire of Congress to encourage employers to provide specific benefits, viewed as socially desirable, to a wide range of employees" (U.S. House of Representatives 1978, Statement of Jerome Kurz). While it is not especially useful in offering advice to Congress to assume that the status quo reflects Congress' most preferred point, it probably is true that there are some positive externalities that attach to medical care and to medical insurance that increase the desired quantity of insurance. The externality is in the form of the altruistic externality discussed earlier.

But then the question is whether tax deductibility is an appropriate subsidy device. Tax deductibility pays larger subsidies to higher-income workers, and these workers would probably buy health insurance in socially adequate amounts even in the absence of a tax subsidy. The subsidy fails to affect the coverage of the self-employed and those employed in firms currently offering low fringe benefits. And it gives the lowest subsidies to those low-income workers who would be most likely to be in need of encouragement. A direct tax credit, of the type proposed by Enthoven (1984), would almost surely be a superior replacement for tax deductibility. Only the "local public goods" argument mentioned earlier, in which those low-income workers who work for high-benefit firms are forced by the tax subsidy to buy more insurance than they want, would strongly suggest that the tax subsidy lands near the externality target. Just how low-income workers' coverage would change is an important research topic.

Another potential market imperfection is adverse selection, and it is certainly possible that some type of a tax subsidy might lead to a second best optimum. There is some recent theory that suggests that there may be no room for improvement if choices about insurance offerings are

first made in a group setting such as in employment groups. It is also far from obvious that the current tax subsidy is the optimal tax-side corrective for adverse selection—but that is because we are barely able to define the problem in an empirically relevant context, much less outline the solution. However, there is surely no obvious evidence at present that adverse selection is the problem to which the tax subsidy is the answer.

Although a good bit is known about the effect of tax subsidies and other influences in the health care sector, there are still some anomalies and puzzles in assessing its contribution to the medical care cost crisis. Hospital expenditure, which accounts for more than 40 percent of total health expenditure, has grown very rapidly for decades, and the growth has been primarily because of an increase in the intensity and cost of care per patient day in the hospital, and not because of increases in admissions or patient days, or because of especially large increases in the price of hospital-specific inputs. We also now know two methods of reducing hospital expenditures—the HMO strategy and the strategy of insurance with cost sharing. Both strategies appear to operate primarily by reducing the number of hospital admissions. But although we have a cause and some cures, the two do not quite come together. Specifically the cures operate on something—the admission rate—that was never an important cause of rising expenditures. We know how to get down the mountain (at least part way) but we do not know how to go down the way we came up. It seems likely that the expenditure savings of either strategy will be once and for all, and not affect the long-run equilibrium rate of price increase.

Here is where the explanations based on qualitative moral hazard, technical change, and search may be important. As noted above, neither the Health Insurance Experiment nor existing empirical work deal with these explanations very well, largely because of our ignorance of both the proper characterization of firm behavior and of market equilibrium. What is even less clear is whether there are any other public sector actions that show promise of improving welfare. Improved information flows come closest, but the difficulty here is one of finding feasible and cheap ways to help consumers choose.

Another policy question is how one might alter the tax treatment of health insurance in a way that is politically feasible. By its nature, the benefits from this "loophole" are not distributed uniformly within income levels. If changes in tax treatment are to occur, they are likely to occur as part of an overall tax reform measure. If reform is not undertaken, the industry is likely to remain in a crisis position for the foreseeable future. In any case, even though we are not certain how much good reducing the tax subsidy will do as far as efficiency is concerned, it is quite unlikely that it will do any harm. It will, however, change the distribution of income, and so will be politically feasible only when (and if) opposition from the losers can be overcome.

References

AMERICAN HOSPITAL ASSOCIATION. *Hospital statistics.* 1984 ed. Chicago: Amer. Hospital Assoc., 1984.

———. "1984 Hospital Cost and Utilization Trends," *Econ. Trends,* Spring 1985, *1*(1), pp. 1–12.

ANDERSON, RONALD AND MAY, J. JOEL, JR. *The Federal Employees' Health Benefits Program, 1961–1968: A model for national health insurance.* Chicago: Center for Health Administration Studies, U. of Chicago, 1971.

ARNETT, ROSS H., III AND TRAPNELL, GORDON R. "Private Health Insurance: New Measures of a Complex and Changing Industry," *Health Care Fin. Rev.,* Winter 1984, *6*(2), pp. 31–42.

ARROW, KENNETH J. "Uncertainty and the Welfare Economics of Medical Care," *Amer. Econ. Rev.,* Dec. 1963, *53*(5), pp. 941–73.

———. "Welfare Analysis of Changes in Health Coinsurance Rates," in *The role of health insurance in the health services sector.* Ed.: RICHARD ROSETT. NY: NBER, 1976, pp. 3–23.

AUSTER, RICHARD D. AND OAXACA, RONALD L. "Identification of Supplier Induced Demand in the

Health Care Sector," *J. Human Res.*, Summer 1981, *16*(3), pp. 327–42.

BENHAM, LEE. "The Effects of Advertising on the Price of Eyeglasses," *J. Law. Econ.*, Oct. 1972, *15*(2), pp. 337–52.

Blue Cross/Blue Shield fact book. 1980 ed. Chicago: BCBS Assoc., 1981.

BRITTAIN, JOHN. *The payroll tax for social security.* Washington, DC: Brookings Institution, 1972.

BROOK, ROBERT H. ET AL. "Does Free Care Improve Adults' Health? Results from a Randomized Clinical Trial," *New Eng. J. Med.*, Dec. 8, 1983, *309*(26), pp. 1426–34.

CAVE, JONATHAN A. K. "Equilibrium in Insurance Markets with Asymmetric Information and Adverse Selection." Rand Corporation Report R–3015–HHS. The Rand Corporation, Apr. 1984.

_____. "Subsidiary Equilibrium and Multiple-option Insurance Markets." in *Advances in health economics and health services research.* Vol. 6. Eds.: RICHARD SCHEFFLER AND LOUIS ROSSITER. (Paper presented at Conference on Biased Selection in Health Care Markets, Berkeley, Apr. 1985.) Greenwich CT and London: JAI Press, 1986, pp. 27–35.

CHOLLET, DEBORAH J. AND STAFF. *Employer-provided health benefits.* Washington, DC: Employee Benefits Research Institute, 1984.

COELEN, CRAIG AND SULLIVAN, DANIEL. "An Analysis of the Effects of Prospective Reimbursement Plans," *Health Care Fin. Rev.*, Winter 1981, *2*(1), pp. 1–40.

CONGRESSIONAL BUDGET OFFICE. *Containing medical care costs through market forces.* Washington, DC: U.S. GPO, May 1982.

COYTE, PETER C. "The Economics of Medicare: Equilibrium within the Medical Community." PhD dissertation, U. of Western Ontario, 1982.

CULYER, ANTHONY J. "Medical Care and the Economics of Giving," *Economica, N.S.* Aug. 1971, *38*(151), pp. 295–303.

DANZON, PATRICIA M. "Hospital 'Profits': The Effects of Reimbursement Policies," *J. Health Econ.*, May 1982, *1*(1), pp. 29–52.

DAVIS, KAREN. "Economic Theories of Behavior in Nonprofit Private Hospitals," *Econ. Bus. Bull.*, Winter 1972, *24*(2), pp. 1–13.

DIONNE, GEORGES. "Search and Insurance," *Int. Econ. Rev.*, June 1984, *25*(2), pp. 357–67.

DRANOVE, DAVID A. "An Economic Model of the Physician-Patient Relationship." PhD dissertation, Grad. School of Bus., Stanford U., 1983.

EGGERS, PAUL W. "Risk Differential Between Medicare Beneficiaries Enrolled and Not Enrolled in an HMO," *Health Care Fin. Rev.*, Winter 1980, *1*(1), pp. 91–99.

EGGERS, PAUL W. AND PRIHODA, RONALD H. "Preenrollment Reimbursement Patterns of Medicare Beneficiaries Enrolled in At-risk HMOs," *Health Care Fin. Rev.*, Spring 1982, *4*(1), pp. 55–74.

EHRLICH, ISAAC AND BECKER, GARY. "Market Insurance, Self Insurance, and Self Protection," *J. Polit. Econ.*, July/Aug. 1972, *80*(4), pp. 623–48.

ENTHOVEN, ALAIN C. *Health plan: The only practical solution to the soaring cost of medical care.* Reading, MA: Addison-Wesley, 1980.

_____. "A New Proposal to Reform the Tax Treatment of Health Insurance," *Health Affairs,* 1984, *3*(1), pp. 21–39.

EVANS, ROBERT G. "Supplier-Induced Demand: Some Empirical Evidence and Implications," in *The economics of health and medical care.* Ed.: MARK PERLMAN. London: Macmillan, 1974, pp. 162–73.

FARLEY, PAMELA J. AND WILENSKY, GAIL R. "Options, Incentives and Employment-Related Health Insurance," in *Advances in health economics and health services research,* Vol. 4. Eds.: RICHARD SCHEFFLER AND LOUIS ROSSITER. Greenwich, CT: JAI Press, 1983, pp. 57–82.

_____. "Household Wealth and Health Insurance as Protection Against Medical Risks," in *Horizontal equity, uncertainty, and economic well being.* Eds.: MARTIN DAVID AND TIMOTHY SMEEDING. Chicago: U. of Chicago Press for the NBER, 1984, pp. 323–54.

FELDMAN, ROGER AND BEGUN, JAMES W. "The Effects of Advertising Restrictions: Lessons from Optometry," *J. Human Res.*, Supp. 1978, *13*(5), pp. 247–62.

FELDSTEIN, MARTIN S. "Hospital Cost Inflation: A Study of Nonprofit Price Dynamics," *Amer. Econ. Rev.*, Dec. 1971, *61*(5), pp. 853–72.

_____. "The Welfare Loss of Excess Health Insurance," *J. Polit. Econ.*, Mar./Apr. 1973, *81*(2), pp. 251–80.

_____. "Econometric Studies of Health Economics," in *Frontiers of quantitative econometrics,* Vol. 2. Eds.: MICHAEL D. INTRILIGATOR AND DAVID A. KENDRICK. Amsterdam: North-Holland, 1974, pp. 377–434.

_____. "Quality Change and the Demand for Hospital Care," *Econometrica,* Oct. 1977, *45*(7), pp. 1681–1702.

FELDSTEIN, MARTIN S. AND ALLISON, ELIZABETH. "Tax Subsidies of Private Health Insurance," in *The economics of federal subsidy programs.* Pt. 8. Papers submitted to the Subcommittee on Priorities and Economy in Government of the Joint Economic Committee, 93rd Cong., 2nd sess., July 1974.

FELDSTEIN, MARTIN S. AND FRIEDMAN, BERNARD. "Tax Subsidies, the Rational Demand for Insurance, and the Health Care Crisis," *J. Public Econ.*, Apr. 1977, *7*(2), pp. 155–78.

FRECH, HARRY E. "The Regulation of Health Insurance." PhD dissertation, U. of California, Los Angeles, 1974.

FRIEDMAN, BERNARD S.; LATOUR, STEPHEN A. AND HUGHES, E. F. X. "A Medicare Voucher System: What Can It Offer?" in *Proceedings of the conference on the future of medicare.* Washington, DC: Subcommittee on Health of the Committee on Ways and Means, U.S. GPO, 1984.

FUCHS, VICTOR R. "The Supply of Surgeons and the Demand for Operations," *J. Human Res.*, Supp. 1978, *13*(5), pp. 35–56.

_____. "Economics, Health, and Post-industrial Soci-

ety," *Milbank Mem. Fund Quart.*, Spring 1979, *57*(2), pp. 153–82.

FUCHS, VICTOR R. AND KRAMER, MARCIA. *Determinants of expenditures for physicians' services in the United States 1948–1968.* NY: NBER, 1972.

GIANFRANCESCO, FRANK D. "A Proposal for Improving the Efficiency of Medical Insurance," *J. Health Econ.*, Aug. 1983, *2*(2), pp. 175–84.

GIBSON, R. ET AL. "National Health Expenditures, 1983," *Health Care Fin. Rev.*, Winter 1984, *6*(2), pp. 1–30.

GINSBURG, PAUL G. "Altering the Tax Treatment of Employment-Based Health Plans," *Milbank Mem. Fund Quart.*, Spring 1981, *59*(2), pp. 224–55.

GOLD, MARCIA. "Competition Within the Federal Employees Health Benefits Program." Paper presented at Amer. Pub. Health Assoc. Meetings, Montreal, 1982.

GOLDBERG, LAWRENCE G. AND GREENBERG, WARREN. "The Effect of Physician-Controlled Health Insurance: U.S. v Oregon State Medical Society," *J. Health Politics, Policy and Law*, Spring 1977, *2*(1), pp. 48–78.

_____. "The Emergence of Physician-Sponsored Health Insurance: A Historical Perspective," in *Competition in the health care sector; Past, present, and future.* Ed.: WARREN GREENBERG. Germantown, MD: Aspen Systems Corp., 1978, pp. 231–54.

GOLDSTEIN, GERALD S. AND PAULY, MARK V. "Group Health Insurance as a Local Public Good," in *The role of health insurance in the health services sector.* Ed.: RICHARD ROSETT. NY: NBER, 1976, pp. 73–110.

GREEN, JERRY R. "Physician-Induced Demand for Medical Care," *J. Human Res.*, Supp. 1978, *13*, pp. 21–34.

GREENBERG, WARREN. "Provider-Influenced Insurance Plans and Their Impact on Competition: Lessons from Dentistry," in *A new approach to the economics of health care.* Ed.: MANCUR OLSON. Washington, DC: Amer. Enterprise Institute, 1981, pp. 359–76.

HADLEY, JACK AND FEDER, JUDITH. "Hospital Cost Shifting: An Analysis of Hospitals' Markups and Financial Needs." Urban Institute Working Paper 3179–07. Washington, DC: The Urban Institute, June 1984.

HARRIS, JEFFREY E. "The Internal Organization of Hospitals: Some Economic Implications," *Bell J. Econ.*, Autumn 1977, *8*(2), pp. 467–82.

_____. "The Aggregate Coinsurance Rate and the Supply of Innovations in the Hospital Sector." Unpub. M.I.T., Dept. of Econs., July 1979.

HAVIGHURST, CLARK C. AND HACKBARTH, GLENN A. "Private Cost Containment," *New Eng. J. Med.*, June 7, 1979, *300*(25), pp. 1298–1305.

HELLINGER, FRED. "Prospective Reimbursement Through Budget Review—New Jersey, Rhode Island, Western Pennsylvania," *Inquiry*, Sept. 1976, *13*(3), pp. 309–20.

HERSHEY, JOHN C. ET AL. "Health Insurance Under Competition: Will People Choose What is Expected?" Management Discussion Paper Series #D44, Leonard Davis Institute of Health Economics, U. of Penn., Philadelphia, July 1982.

HOLMER, MARTIN. "Tax Policy and the Demand for Health Insurance." U.S. Dept. of Health and Human Services, Washington, DC, Apr. 1984.

INTERSTUDY. *National HMO Census*, June 30, 1984. Excelsior, MN: Interstudy, 1983.

JACKSON-BEECK, MARILYN AND KLEINMAN, JOHN H. "Evidence for Self-Selection Among Health Maintenance Organization Enrollees," *J. Amer. Med. Assoc.*, Nov. 25, 1983, *250*(20), pp. 2826–29.

JOSKOW, PAUL L. "The Effects of Competition and Regulation on Hospital Bed Supply and the Reservation Quality of the Hospital," *Bell J. Econ.*, Autumn 1980, *11*(2), pp. 421–47.

KEELER, EMMETT B.; NEWHOUSE, JOSEPH P. AND PHELPS, CHARLES E. "Deductibles and the Demand for Medical Care Services," *Econometrica*, Apr. 1977, *45*(3), pp. 641–55.

KEELER, EMMETT B. ET AL. "The Demand for Episodes of Medical Treatment." Rand paper R-2829-HHS. Santa Monica, CA, The Rand Corporation, Dec. 1982.

KESSEL, REUBEN. "Price Discrimination in Medicine," *J. Law Econ.*, Oct. 1958, *1*(1), pp. 20–53.

KWOKA, JOHN E., JR. "Advertising and the Price and Quality of Optometric Services," *Amer. Econ. Rev.*, Mar. 1984, *74*(1), pp. 211–16.

LEE, JAMES A. *Employment, unemployment, and health insurance.* Cambridge, MA: Abt Assoc., Inc., 1979.

LEES, DENNIS S. AND RICE, RICHARD G. "Uncertainty and the Welfare Economics of Medical Care: Comment," *Amer. Econ. Rev.*, Mar. 1965, *55*(1), pp. 140–54.

LINDSAY, COTTON M. "Medical Care and the Economics of Sharing," *Economica, N.S.* Nov. 1969, *36*(144), pp. 351–62.

LINK, CHARLES ET AL. "Cost Sharing, Supplemental Insurance, and Health Services Utilization among the Medicare Elderly," *Health Care Fin. Rev.*, Fall 1980, *2*(3), pp. 25–32.

LONG, JAMES AND SCOTT, FRANK. "The Income Tax and Nonwage Compensation," *Rev. Econ. Statist.*, May 1982, *64*(2), pp. 211–19.

LUFT, HAROLD S. "Trends in Medical Care Costs," *Med. Care*, Feb. 1980, *18*(1), pp. 1–16.

_____. *Health maintenance organizations.* NY: Wiley, 1981.

MAHOLTRA, SURESH ET AL. "Employment Related Health Benefits in Private Nonfarm Business Establishments in the United States." Unpublished final report by Batelle Human Affairs Research Center on Contract J9P7-0150. Labor Management Services Administration, U.S. Dept. of Labor, June 1980.

MANNING, WILLARD G. ET AL. "A Controlled Trial of the Effect of a Prepaid Group Practice on Use of Services," *New Eng. J. Med.*, June 7, 1984, *310*(23), pp. 1505–10.

MARQUIS, M. SUSAN. "Cost-sharing and Provider Choice," *J. Health Econ.*, June 1985, *4*(2), pp. 132–57.

MARSHALL, JOHN. "Moral Hazard," *Amer. Econ. Rev.*, Dec. 1976, *66*(5), pp. 880–90.

MCCARTHY, THOMAS R. "The Competitive Nature of the Primary-care Physician Services Market," *J. Health Econ.*, June 1985, *4*(2), pp. 93–118.

MCGUIRE, THOMAS. "Price and Membership in a Prepaid Medical Group Practice," *Med. Care*, Feb. 1981. *19*(2), pp. 172–83.

_____. *Financing psychotherapy: Costs, effects, and public policy.* Cambridge, MA: Ballinger, 1981.

MEYER, JACK A. AND JOHNSON, WILLIAM R. "Cost Shifting in Health Care: An Economic Analysis," *Health Affairs*, Summer 1983, *2*(2), pp. 20–35.

DE MEZA, DAVID. "Health Insurance and the Demand for Medical Care," *J. Health Econ.*, Mar. 1983, *2*(1), pp. 47–54.

MITCHELL, BRIDGER M. AND PHELPS, CHARLES. "National Health Insurance: Some Costs and Effects of Mandated Employee Coverage," *J. Polit. Econ.*, June 1976, *84*(3), pp. 553–71.

MITCHELL, BRIDGER M. AND VOGEL, RONALD. "Health and Taxes: An Assessment of the Medical Deduction," *Southern Econ. J.*, Apr. 1975, *41*(4), pp. 660–72.

MITCHELL, JANET B. AND CROMWELL, JERRY K. "Physician-Induced Demand for Surgical Operations," in *Health care financing administration grants and contracts report.* DHHS Publ. No. (HCFA) 03086–81. Washington, DC: U.S. Dept. of Health and Human Services, 1981.

MIYAZAKI, HAIJIME. "The Rat Race and Internal Labor Markets," *Bell J. Econ.*, Autumn 1977, *8*(3), pp. 394–418.

NEIPP, JOACHIM AND ZECKHAUSER, RICHARD. "Persistence in the Choice of Health Plans," in *Advances in health economics and health services research.* Vol. 6. Eds.: RICHARD S. SCHEFFLER AND LOUIS E. ROSSITER. Greenwich, CT and London: JAI Press, 1986, pp. 47–72.

NEWHOUSE, JOSEPH P. "The Structure of Health Insurance and the Erosion of Competition in the Medical Marketplace," in *Competition in the health care sector: Past, present, future.* Ed.: WARREN GREENBERG. Germantown, MD: Aspen Systems Corp., 1978a, pp. 215–30.

_____. *The economics of medical care: A policy perspective.* Reading, MA: Addison-Wesley, 1978b.

_____. "Insurance Benefits, Out-of-Pocket Payments, and the Demand for Medical Care: A Review of the Literature," *Health and Med. Care Services Rev.*, Mar. 1978c, *1*(4), pp. 1–15.

_____. "The Demand for Medical Care Services: A Retrospect and Prospect," in *Health, economics, and health economics.* Eds.: JACQUES VAN DER GAAG AND MARK PERLMAN. Amsterdam and NY: North-Holland, 1981, pp. 85–102.

_____. "Two Prospective Difficulties with Prospective Payment of Hospitals, or, It's Better to be a Resident Than a Patient with a Complex Problem," *J. Health Econ.*, Dec. 1983, *2*(3), pp. 269–74.

_____ ET AL. "Some Interim Results from a Controlled Trial in Health Insurance," *New Eng. J. Med.*, Dec. 17, 1981, *305*(25), pp. 1501–07.

_____ ET AL. "Are Fee-for-Service Costs Increasing Faster than HMO Costs?" *Med. Care*, Aug. 1985, *23*(8), pp. 960–67.

NEWHOUSE, JOSEPH P. AND MARQUIS, M. SUSAN. "The 'Norms' Hypothesis and the Demand for Medical Care," *J. Human Res.*, Supp. 1978, *13*(5), pp. 159–82.

NEWHOUSE, JOSEPH P. AND PHELPS, CHARLES E. "New Estimates of Price and Income Elasticities for Medical Care Services," in *The role of health insurance in the health services sector.* Ed.: RICHARD ROSETT. NY: NBER, 1976, pp. 313–20.

NEWHOUSE, JOSEPH P.; PHELPS, CHARLES E. AND MARQUIS, M. SUSAN. "On Having Your Cake and Eating It Too," *J. Econometrics*, Aug. 1980, *13*(3), pp. 365–90.

PAULY, MARK V. "The Economics of Moral Hazard: Comment" *Amer. Econ. Rev.*, June 1968, *58*(3), pp. 531–37.

_____. "Indemnity Insurance for Health Care Efficiency," *Econ. Bus. Bull*, Fall 1971b, *4*(4), pp. 53–59.

_____. *Medical care at public expense.* NY: Praeger, 1971a.

_____. "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection," *Quart. J. Econ.*, Feb. 1974, *88*(1), pp. 44–62.

_____. "Overinsurance: The Conceptual Issues," in *National health insurance: What now, what later, what never?* Ed.: MARK PAULY. Washington, DC: Amer. Enterprise Institute, 1980a, pp. 201–19.

_____. *Doctors and their workshops.* Chicago: U. of Chicago Press, 1980b.

_____. "Market Power, Monopsony, and Health Insurance Markets." Paper presented at American Risk and Insurance Assoc. meeting, Aug. 1985.

_____. "What is Adverse About Adverse Selection?" in *Advances in health economics and health services research.* (Paper presented at Conference on Biased Selection in Health Care Markets, Berkeley, Apr. 1985.) Eds.: RICHARD S. SCHEFFLER AND LOUIS F. ROSSITER. Greenwich, CT and London: JAI Press, 1986, pp. 281–86.

PAULY, MARK V. AND DRAKE, DAVID F. "Effect of Third-Party Methods of Reimbursement on Hospital Performance," in *Empirical studies in health economics.* Ed.: HERBERT KLARMAN. Baltimore: Johns Hopkins U. Press, 1970, pp. 315–19.

PAULY, MARK V. AND REDISCH, MICHAEL. "The Not-for-Profit Hospital as a Physicians' Cooperative," *Amer. Econ. Rev.*, Mar. 1973, *63*(3), pp. 87–89.

PAULY, MARK V. AND SATTERTHWAITE, MARK S. "The Pricing of Primary Care Physicians' Services," *Bell J. Econ.*, Autumn 1981, *12*(2), pp. 488–506.

PHELPS, CHARLES E. *The demand for health insurance: A theoretical and empirical investigation.* Santa Monica, CA: Rand Corp., 1973.

_____. "Health Care Costs: The Consequences of Increased Cost Sharing." Rand Corp. Report R-2970-RC, Rand Corp. Santa Monica, CA, Nov. 1982.

_____. "Tax Policy, Health Insurance, and Health Care," in *Market reforms in health care.* Ed.: JACK

MEYER. Washington, DC: Amer. Enterprise Institute, 1983a, pp. 198–224.

_____. "Taxing Health Insurance: How Much is Enough?" Rand Corp. Report P–6915, Rand Corp., Santa Monica, CA, Sept. 1983b.

_____. "Large Scale Tax Reform: The Case of Employer-Paid Health Insurance Premiums." Work. Pap. No. 20, Applied Econs. Workshop, U. of Rochester, Sept. 1985.

PHELPS, CHARLES E. AND MARQUIS, M. SUSAN. "Demand for Supplementary Health Insurance." Rand Corp., Santa Monica, CA, Jan. 1983.

PHELPS, CHARLES E. AND NEWHOUSE, JOSEPH P. "Coinsurance, the Price of Time, and the Demand for Medical Services," *Rev. Econ. Statist.*, Aug. 1974, *56*(2), pp. 334–42.

PRICE, JAMES R.; MAYS, JAMES W. AND TRAPNELL, GORDON K. "Stability in the Federal Employees Health Benefits Program," *J. Health Econ.*, Dec. 1983, *2*(3), pp. 207–24.

REINHARDT, UWE E. "Health Insurance and Cost-Containment Policies: The Experience Abroad," *Amer. Econ. Rev.* May 1980, *70*(2), pp. 149–56.

_____. "The Theory of Physician-induced Demand: Reflections After a Decade," *J. Health Econ.*, June 1985, *4*(2), pp. 187–94.

ROSEN, SHERWIN. "Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition," *J. Polit. Econ.*, Jan./Feb. 1974, *82*(1), pp. 34–55.

ROSETT, RICHARD N. AND HUANG, LIEN-FU. "The Effect of Health Insurance on the Demand for Medical Care," *J. Polit. Econ.*, Mar./Apr. 1973, *81*(2), pp. 281–305.

ROTHSCHILD, MICHAEL AND STIGLITZ, JOSEPH. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quart. J. Econ.*, Nov. 1976, *90*(4), pp. 630–49.

RUSSELL, LOUISE B. *Technology in hospitals.* Washington, DC: Brookings Institution, 1979.

SATTERTHWAITE, MARK A. "Consumer Information, Equilibrium Industry Price, and the Number of Sellers," *Bell J. Econ.*, Autumn 1979, *10*(2), pp. 483–502.

SCITOVSKY, ANNE A.; BENHAM, LEE AND MCCALL, NELDA. "Use of Physicians' Services Under Two Prepaid Plans," *Medical Care*, May 1979, *17*(5), pp. 441–60.

SHAVELL, STEVEN. "Theoretical Issues in Medical Malpractice," in *The economics of medical malpractice.* Ed.: SIMON ROTTENBERG. Washington, DC: Amer. Enterprise Institute, 1978, pp. 35–64.

SLOAN, FRANK A. AND FELDMAN, ROGER. "Competition Among Physicians," in *Competition in the health care sector: Past, present, and future.* Ed.: WARREN GREENBERG. Germantown, MD: Aspen Systems Corp., 1978, pp. 45–102.

SLOAN, FRANK A. AND STEINWALD, BRUCE. "Effects of Regulation on Hospital Costs and Input Use," *J. Law. Econ.*, Apr. 1980, *23*(1), pp. 81–109.

SPENCE, MICHAEL. "Product Differentation and Performance in Insurance Markets," *J. Public Econ.*, Dec. 1978, *10*(3), pp. 427–47.

SWARTZ, KATHERINE. "The Changing Face of the Uninsured." Paper presented at First Annual Meeting of Assoc. for Health Services Research, Chicago, June 11, 1984.

TAYLOR, AMY K. AND WILENSKY, GAIL R. "The Effect of Tax Policies on Expenditures for Private Health Insurance," in *Market reforms in health care.* Ed.: JACK MEYER. Washington, DC: Amer. Enterprise Institute, 1983, 163–84.

TAYLOR, AMY; WILENSKY, GAIL R. AND ROSSITER, LOUIS. "The Role of Money and Time in the Demand for Medical Care." Paper presented at Amer. Econ. Assoc. Meetings, NY, Dec. 1981.

TIEBOUT, CHARLES. "A Pure Theory of Local Expenditures," *J. Polit. Econ.*, Oct. 1956, *64*(5), pp. 416–24.

U.S. HOUSE OF REPRESENTATIVES, COMMITTEE ON WAYS AND MEANS. *Tax treatment of employee fringe benefits.* U.S. GPO, 1978.

VLADECK, BRUCE C. "The Market vs. Regulation: The Case for Regulation," *Milbank Mem. Fund Quart.*, Spring 1981, *59*(2), pp. 209–23.

WALDEN, D. C. ET AL. "Changes in Health Insurance Status." Data Preview 21, National Health Care Expenditures Study, National Center for Health Services Research, n.d.

WELLER, CHARLES. "Antitrust, Joint Ventures and the End of the AMA's Contract Practice Ethics," *North Carolina Cont. Law J.*, Mar. 1983, *14*(1), pp. 3–32.

WENNBERG, JOHN. "Dealing with Medical Practice Variations: A Proposal for Action," *Health Affairs*, Summer 1984, *3*(2), pp. 6–32.

WILENSKY, GAIL R.; FARLEY, PAMELA J. AND TAYLOR, AMY K. "Variations in Health Insurance Coverage: Benefits Versus Premiums," *Milbank Mem. Fund. Quart.*, Winter 1984, *62*(1), pp. 53–82.

WILENSKY, GAIL R. AND ROSSITER, LOUIS F. "The Magnitude and Determinants of Physician-Initiated Visits in the United States," in *Health, economics, and health economics.* Eds.: JACQUES VAN DER GAAG AND MARK PERLMAN. Amsterdam: North-Holland, 1981, pp. 215–43.

WILENSKY, GAIL R. AND WALDEN, DANIEL. "Minorities, Poverty, and the Uninsured." Paper presented at American Public Health Assoc. meeting, Los Angeles, Dec. 1981.

WILSON, CHARLES. "A Model of Insurance Markets with Incomplete Information," *J. Econ. Theory*, Dec. 1977, *16*(2), pp. 167–207.

WILSON, GEORGE AND JADLOW, JOSEPH. "Competition, Profit Incentives, and Technical Efficiency in the Provision of Nuclear Medicine Services," *Bell J. Econ.*, Autumn 1982, *13*(2), pp. 472–81.

WOODBURY, STEVEN. "Substitution Between Wage and Nonwage Benefits," *Amer. Econ. Rev.*, Mar. 1983, *73*(1), pp. 166–82.

WORTHINGTON, NANCY AND PIRO, PAULA. "The Effects of Hospital Rate Setting Programs on Volumes of Hospital Services," *Health Care Fin. Rev.*, Dec. 1982, *4*(2), pp. 47–66.

ZECKHAUSER, RICHARD. "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives," *J. Econ. Theory*, Mar. 1970, *2*(1), pp. 10–26.