

**Charging Peter to Pay Paul:
Accounting for the Financial Effects of User Charges**

**Robert G. Evans
Morris L. Barer
Greg L. Stoddart**

HPRU 93:17D DECEMBER, 1993

HEALTH POLICY RESEARCH UNIT

**CENTRE FOR HEALTH SERVICES AND POLICY RESEARCH
429 - 2194 HEALTH SCIENCES MALL
UNIVERSITY OF BRITISH COLUMBIA
VANCOUVER, B.C. CANADA
V6T 1Z3**

The Centre for Health Services and Policy Research was established by the Board of Governors of the University of British Columbia in December 1990. It was officially opened in July 1991. The Centre's primary objective is to co-ordinate, facilitate, and undertake multidisciplinary research in the areas of health policy, health services research, population health, and health human resources. It brings together researchers in a variety of disciplines who are committed to a multidisciplinary approach to research, and to promoting wide dissemination and discussion of research results, in these areas. The Centre aims to contribute to the improvement of population health by being responsive to the research needs of those responsible for health policy. To this end, it provides a research resource for graduate students; develops and facilitates access to health and health care databases; sponsors seminars, workshops, conferences and policy consultations; and distributes Discussion papers, Research Reports and publication reprints resulting from the research programs of Centre faculty.

The Centre's Health Policy Research Unit Discussion Paper series provides a vehicle for the circulation of preliminary (pre-publication) work of Centre Faculty and associates. It is intended to promote discussion and to elicit comments and suggestions that might be incorporated within the work prior to publication. While the Centre prints and distributes these papers for this purpose, the views in the papers are those of the author(s).

A complete list of available Health Policy Research Unit Discussion Papers and Reprints, along with an address to which requests for copies should be sent, appears at the back of each paper.

**UBC CENTRE FOR HEALTH SERVICES AND POLICY RESEARCH
DISCUSSION PAPER HPRU 93:17D**

**CHARGING PETER TO PAY PAUL:
Accounting for the Financial Effects of User Charges**

**Robert G. Evans
Department of Economics, and
Centre for Health Services and Policy Research
University of British Columbia**

**Morris L. Barer
Centre for Health Services and Policy Research, and
Department of Health Care and Epidemiology
University of British Columbia**

**Greg L. Stoddart
Centre for Health Economics and Policy Analysis, and
Department of Clinical Epidemiology and Biostatistics
McMaster University**

September, 1993

This work was funded by the Ontario Premier's Council on Health, Well-being and Social Justice. Responsibility for the views expressed herein, and any errors or omissions, rests solely with the authors. R.G. Evans is supported by a National Health Scientist award from Health and Welfare Canada, and is a Fellow of the Canadian Institute for Advanced Research. M.L. Barer and G.L. Stoddart are an Associate and a Fellow, respectively, of the Institute. The authors wish to thank the many individuals both inside and outside the health care system who have taken the time to share their views on user charges.

Preface

This is one in a series of articles by the authors about the ongoing debate over user charges in the Canadian health care system.

In this paper we outline a formal and comprehensive analytic framework in which income transfers - the principal effects of user charges - can be traced between groups in the population (e.g. the healthy and the sick, the rich and the poor), between payers and health care providers, and among providers. The accounting relationships in this framework, expressed as equations, are a sub-set of those underlying the economy-wide national income and product accounts. They make it clear that "it is impossible to do only one thing" when charges are imposed -- or removed. The framework is used to analyze the patterns of income transfers associated with different types of user charges.

Other papers in this series focus on other aspects of the user charge debate, including the popular arguments in support of, the common rationales for, the key supporters of, and the principal gainers and losers from such charges. A brief description of each paper follows.

"The Remarkable Tenacity of User Charges" documents the history of the user charge debate in Canada. It reviews the participation, positions and rationales of Canadian interest groups in debates over "patient participation" in health care financing.

"Who Are the Zombie Masters, and What Do They Want?" likens the recurring proposals for user charges to zombies - the so-called 'walking dead' - because although they have been repeatedly rejected by policy-makers and the general public (and the substantive claims of their supporters refuted by analyses of the effects of such charges), these proposals refuse to remain buried. This paper examines why that is the case, and who stands to benefit from the introduction of user charges.

"User Charges, Snares and Delusions: Another Look at the Literature" reviews and extends an earlier in-depth analysis of the effects of user charges which three of the authors published in 1979. The paper assesses whether experience and published literature in the years since then alter any of the (largely negative) conclusions of the earlier study concerning the ability of direct charges to patients to achieve important public policy objectives, including controlling health care costs.

"It's not the Money, It's the Principle" examines why user charges exist for some health care services and not for others. The paper analyzes the characteristics of services which (do or should) underlie decisions to charge in part or in whole for specific types of services. We propose a framework for evaluating the justification for, and feasibility of, user charges for particular types of services in particular situations.

"Why Not User Charges? The Real Issues" examines some of the most frequently heard arguments for user charges and looks at what evidence there is for claims and

counter-claims that are often made. Because statements in the "popular" debate sometimes seem inconsistent with each other, or unrelated to or at odds with the facts, we explore the statements more carefully, asking what they really mean, what values they are based on, and what fundamental issues are at the heart of the user charge controversy. This paper in particular is intended for a wide general audience and assumes that most readers will have heard - or perhaps made - the arguments described, but will have little detailed or technical knowledge of the issues involved.

In addition, a bibliography entitled "User Charges in Health Care" provides an extensive set of references to articles of relevance to the user charge debate in Canada, drawn from diverse sources including academic research and policy analysis literature, the popular press, government documents and reports, and the publications and reports of non-governmental organizations including the professional associations representing a variety of health care providers.

Gainers and Losers: Who Pays, Who Is Paid, and Who Is Cared For?

User charges are one of the ways in which revenues can be raised for the health care sector. Whatever other effects they may have -- encouraging people to be more responsible in their use of care, preventing people from getting access to the care they need, or any other rhetorical claim -- they always transfer funds from those who use care, either to those who provide it, or to those who pay those who provide it. Thus any understanding of their impact must include an awareness of their relative effects on different people in the society. Most of the political debate over user charges includes, implicitly or explicitly, a concern for their distributional effects -- who gains, who loses?

This immediately raises the question, "Gains or losses relative to what?" to which the answer is, "Relative to some other way of financing health care." No one seriously believes that health care, or any other resource-using activity, is "free". Abstracting from minor revenue sources such as charities or lotteries (both of which in fact also include an indirect diversion from government revenue), there are three principal channels through which funds can be raised to pay for health care: taxes, user charges, and insurance premiums.

Modern health care systems are financed primarily from tax revenue. For political palatability certain taxes are sometimes labelled "social insurance premiums", and may be collected either by governments or by agencies which are at half arm's length from governments - in essence quangos.¹ But there is relatively little true insurance in health care finance. The one country which appears to rely primarily on private insurance, the United States, actually provides large tax expenditures (currently estimated at about \$40 billion) from federal government revenues to the private industry.² Table 1 shows the percentage of health expenditures from public sources for the twenty-four countries of the OECD [2]. This percentage is over three quarters for

¹. When payments are compulsory, unrelated to the risk status of the "insured", and not in fact a condition of entitlement to care, then you may call them what you will but they remain taxes. They may, however, be a particularly regressive form of tax.

². This public subsidy is probably critical to the survival of the "private" American health insurance industry at least in its present form. Even with the subsidy, while private insurance covers more people, public programs pay more of the bills. Private health insurance covers only 32.5% of U.S. health expenditures, compared with 43.9% from public sources [1]. If one transfers the \$40 billion in tax expenditures from private insurance to public expenditure, the proportions change to 27.2% private insurance, 49.2% public spending. (The percentages differ from those reported in Table 1 because the OECD uses slightly different spending definitions.)

TABLE 1
International Comparison of Public Share of Health Care Spending

	1980	1990
Australia	63%	68%
Austria	69%	67%
Belgium	83%	83%
Canada	75%	73%
Denmark	85%	83%
Finland	79%	81%
France	79%	74%
Germany	75%	73%
Greece	82%	76%
Iceland	88%	87%
Ireland	82%	75%
Italy	81%	76%
Japan	71%	72%
Luxembourg	93%	91%
Netherlands	75%	71%
New Zealand	84%	82%
Norway	98%	95%
Portugal	72%	62%
Spain	80%	78%
Sweden	93%	90%
Switzerland	68%	68%
Turkey	27%	36%
United Kingdom	90%	84%
United States	42% (?)	42% (48%)*

* adjusted to include tax subsidy

SOURCE: G.J. Schieber, J.-P. Poullier, and L.M. Greenwald "U.S. health expenditure performance: An international comparison and data update" Health Care Financing Review 13:4 (Summer, 1992)

thirteen countries, and between two-thirds and three quarters for another eight. Only the United States and Turkey are under one half.

However it is raised, the total amount of revenue provided for health care in any society is linked by a fundamental accounting identity to two other important aggregates. It must always equal total expenditures on care -- the total cost of the care received and provided -- and it must equal the total of all the incomes earned from the provision of care by those who directly or indirectly participate in -- supply reimbursed resources for -- its provision. Schematically:

TOTAL		TOTAL		TOTAL
REVENUE	equals	EXPENDITURE	equals	INCOMES

Each dollar (pound, franc, mark...) spent on health care must simultaneously have come from someone, and been paid to someone. The relationship need not, of course, hold for any one of the individuals in the society. Most people will contribute either more or less on the revenue side than the cost of the services they use; on the other side of the equation the amount that people earn from providing health care will typically be either much more, or much less, than they contribute to pay for it. But in aggregate, summed over all the individuals making up the society as a whole, the equation must hold as an identity.³

Within this relationship, individual persons may play a number of different roles, using or providing, and paying or being paid for, many different forms of health care in many different ways. In delineating the various redistributive effects of user charges, we identify the reallocation of costs and benefits among the members of society under five heads, as changes in:

- 1) The share of costs borne by different payers,
- 2) The share of services used by different persons or groups,
- 3) The incomes of providers, relative to the rest of society,

³. This presupposes that the society in question is a "closed" economy, at least with respect to health care. The people who live in it neither purchase health care from outsiders, nor sell it to them. This is "almost" true for hospital and medical care in Canada; the cross-border flows in either direction are minuscule (though they receive a disproportionate amount of attention, both in and outside Canada, for political reasons) [3]. For medical equipment and drugs, of course, there is a substantial international trade. This can be represented in the framework below by adding a "Rest-of-the-World" sector as an additional (large) "person" or set of persons in the economy.

- 4) The relative incomes of different providers or provider groups, and
- 5) The numbers and types of 'providers' reimbursable, directly or indirectly, from health care budgets.

In each case, we can show the characteristics of persons which determine whether they will gain or lose from user charges. What is notable is the number of different ways in which gainers and losers can emerge. This in itself is indicative of the breadth and diversity of the potential pro-user charge constituency; there are many potential gainers, in many different roles. On the other hand, the framework which is implicit in the relationship above, and which we will develop more formally below, shows that all these gains are at someone else's expense -- and we can keep track of whose.

People and Services: The Algebra of Utilization

To represent the different individuals in the society of interest, we can assign each of them a unique integer number i between 1 and N (inclusive), where N is the total number of people in the society (for Canada, about 28 million). A randomly selected person from this society can then be referred to as person i .

Out of all the commodities, goods and services, produced and used in this society, a certain subset are designated as "health care". How these commodities come to be identified as "health care", while others are excluded from the designation, is an important question. It lies behind sometimes very heated arguments over the "correct" definition of health, and over which 'providers' should therefore be entitled to reimbursement from public programs. "Entitlement" means different things to people on opposite sides of the health care transaction. "Health care", and its providers, receive very special regulatory and financial treatment, so the designation carries important and coveted advantages.

But for the purposes of this discussion we can bypass this issue. (We address it elsewhere, see [4].) Whatever the current definitions happen to be, we let the number of different commodities which they include -- such things as doctor visits, surgical procedures, days in hospital, prescriptions filled, and so on -- be represented by M . We can make the descriptions of these commodities as specific or as general as we wish, by making M large or small.

For a realistic description of a modern health care system M would have to be quite large to distinguish different types of professional contacts, procedures, etc. A medical fee schedule, for example, or a list of diagnosis-related groups of hospital admissions, or a catalogue of prescription drugs, each represents a sub-set of M ; we would have to add the numbers of different items in each of these lists together to begin to approach a value for M . On the other hand, for some rather crude representations we may use a very short list, in which a contact with a physician is treated as a single kind of commodity, a filled prescription is another, and a hospital bed-day a third. Small values of M correspond to highly aggregated and heterogeneous "items" of service.

As we did for the persons, we label each of these commodities with a unique integer j between 1 and M , and we designate a randomly chosen commodity from this (long or short) list as health care of type j .

We can then designate by q_{ij} , the total quantity of health care of type j received by person i (during a particular time period, say the year 1993 -- one could add another subscript t to designate this period). The total quantity of health care of type j produced in this society in this time period will then be the sum of the amounts provided to each of its N members. If we call this total Q_j , then:

$$Q_j = q_{1j} + q_{2j} + q_{3j} \dots + q_{ij} + \dots + q_{Nj}$$

or in a more compact notation:

$$Q_j = \sum_i q_{ij}$$

In any particular year, many of the q_{ij} will be zero -- if person i did not use any of service j . If she were really healthy q_{ij} might be zero for all values of j -- person i received/used no care of any kind. But the total volume of services of type j provided and received in this society will be represented by this summation.

Mathematically we can carry out a corresponding summation to represent the total volume of services received by person i , as by:

$$Q_i = \sum_j q_{ij}$$

But this procedure is invalid because we are now adding up apples and oranges -- or much worse. As the index j takes on different values, it counts numbers of office visits to doctors, laboratory tests, days in hospital, open heart procedures... The meaning of a simple total of "services" is not obvious. Two doctor visits, three lab tests, and two prescriptions in the course of a year equals seven "services", as does a triple by-pass operation and six days in intensive care. But we would not want to treat them as equivalent "packages" of services, in any sense. In this case aggregation leads to statistical nonsense; it can also be a potent source of conceptual nonsense.

Nevertheless it is quite common for payment agencies to report the "average number of (fee-for-service) procedures received per beneficiary" in a particular year, and for newspapers to report this as if it were a meaningful number. And it may be, provided the mix of services has not changed -- but one cannot tell from the service count itself.

The legitimate calculation is to assign each procedure a 'weight' to make them comparable with each other. For surgical procedures, for example, a "hernia-equivalent" scale has been developed such that each operation can be assigned a value reflecting its time and complexity relative to the repair of a hernia. Hernia repairs themselves are given a value of 1.0, More complex operations are multiplied by weights greater than 1.0, and less complex ones by lower weights, before adding up a total of "hernia-equivalent" procedures. This set of weights is an example of a "Relative Value Scale" or RVS, which converts items measured in different units into a common and therefore summable unit.

A more common source of weights is prices, which have the dimension "dollars per" whatever the unit may be in which a good or service is measured. If each of the M different health care services in this society can be assigned a price, say P_j , then the

product $P_j \cdot q_{ij}$ is the value, in whatever currency the price is defined in, of the services of type j received by person i , in this society (during the time period under consideration). This leaves open, of course, the question of the underlying basis for the fee schedule or price list itself; do the fees reflect any meaningful characteristics of the different commodities or services themselves? The current interest in "Resource-Based Relative Value Scales" (RBRVS), which attempt to develop weighting systems for physicians' services correlating better with the actual resources -- time, effort, skills, and materials/equipment -- used up in their provision, arises out of this concern.⁴

The summation: $\sum_j [P_j \cdot q_{ij}]$

now has a perfectly valid meaning; it is the total value in dollars of all the health services of whatever type received by person i , the total cost of her care. And we can sum these expressions across all individuals, thus:

$$\sum_{ij} [P_j \cdot q_{ij}] = \sum_j [P_j \cdot \sum_i q_{ij}] = \sum_j P_j \cdot Q_j \quad \dots(1)$$

The total costs of health care in this society are equal to the costs of care of each type j -- doctors' services, hospital care, prescription drugs... -- added across all M types. And the cost of each type of care j is defined by the total volume of that form of care, Q_j , measured in physical units, multiplied by its price (cost, value) per unit, P_j .⁵

⁴. Of course this process itself requires the use of an implicit or explicit weighting of the relative values of the different resources used in providing each type of care or 'fee item'. And these, too, may be based on prices which do not necessarily reflect the relative values that might emerge in "free and perfect" markets. Nor would we necessarily want them to.

⁵. This of course assumes that the P_j used in the summation actually measure the unit costs of the corresponding services. For some services, those of physicians in a fee-for-service reimbursement system, for example, the selection of the appropriate P_j is straight-forward -- they are the fees paid (by whomever) to the service provider. But for organizations supported by budgetary allocations, such as Canadian hospitals or medical clinics in the British or Swedish systems, the unit prices/costs of the many different types of service provided will not be defined. Nevertheless, in principle these implicit "prices" exist, and can be extracted with a sufficiently sophisticated accounting system. (But perhaps not too sophisticated. We are aware of the conceptual problem of allocating inherently joint costs, but in the real world those co-exist with explicit pricing systems and few analysts balk at using the associated prices.)

People and Payments: The Algebra of Revenues

As pointed out at the outset, however, this total will also and by definition be identically equal to the product of two other ways of aggregating up to the total costs of health care. It will be equal on the one hand to the total revenues which are raised to pay for health care, and on the other hand to the total incomes earned by those who supply the services used in health care production.

Taking the revenues first, we pointed out above that the most common way of paying for health care is through taxation. Funding may be drawn from general revenue -- e.g. Canada, the U.K., or Sweden -- or there may be a separate "earmarked" tax system with its own base and rates, often referred to as "social insurance" -- e.g. Germany, France, the United States -- which is then supplemented out of general revenue.

(Quite) roughly speaking, the total burden of taxation (not just the personal income tax) can be thought of as falling upon individuals more or less in proportion to their incomes.⁶ If we consider again person i in this society, and refer to her income as Y_i , and to the components of the general average tax rate which go to pay for health care as t , then person i pays tY_i for health care, through the tax system. Summing across all persons will then yield the total of public spending on health care.⁷

The proportion of health care revenue which is raised through direct charges to users -- in Canada, this includes much of dental care and of prescription drugs in ambulatory care, and most non-prescription drugs -- can then be represented by attaching a parameter, C_j , to person i 's use of health care of type j , equal to the out-of-pocket user charge per unit for goods or services of type j . For services which are fully covered by government or private insurance, this C_j will be zero; for those which are wholly unsubsidized or uninsured it will be equal to P_j . Partial "coinsurance" of, say, 20% would be represented by a value of C_j equal to $0.2P_j$.⁸

⁶. Much more could be, has been, and will be said on this topic -- but not here, and not by us.

⁷. But what about the deficit? Suppose governments borrow to meet part of their expenses? Well, obviously health care is an intra-marginal public expenditure.... Or perhaps t is the discounted future value -- with perfect foresight -- of the tax rate required to repay future borrowing, with interest.... These considerations can be consistently represented by complications to the algebra which do not add enlightenment.

⁸. Things can of course get quite complicated with deductibles. Moreover, the value of C_j can be different across persons, if some (those below a particular income level, for

Finally there may in fact be some component of true insurance, usually private, which charges a fixed premium based upon some estimate of the risk level of the group to which person i belongs. This will relate neither to her income, nor to her actual use of services -- at least not in the current period. Insurance premiums are based on the expected use and cost of services, and this in turn may be affected by past use, or income level, or any other correlate of cost. But for each period taken in isolation, premiums are a constant. They vary from person to person, of course, both because different persons have different levels of coverage, and because the same coverage will be priced differently for different persons. We label this premium R_i .

Person i 's total contribution to the financing of health care will thus be the sum of these three components -- taxes, user charges, and true insurance premiums. If we add across all persons in the society, we get total revenues raised to pay for health care:

$$\sum_i \{tY_i + \sum_j (C_j * q_{ij}) + R_i\} \quad \dots(2)$$

This sum is identically equal to the total cost of health care, expression (1) above.

example, or above or below a certain age) face different charges. This would be reflected in a larger set of parameters, c_{ij} , but again, complexity without enlightenment.

People and Earnings: The Algebra of Incomes

So where does the money go? To pay for health care, of course, but that is a process, not a destination. Revenues are drawn from someone's pocket, or bank account, and must end up in someone else's. Whose? Since there are N people in this society the payments must go, in aggregate, to those same people -- but not in the same proportions.

People are paid for providing health care, in amounts which bear some relation to the time, effort and skill which they put into providing it, or the other resources -- materials and supplies, services of physical capital such as buildings and equipment, or "intellectual property" which they provide. Those inputs of time, effort and skill, and other resources, are not in general the goods and services themselves, though we often speak loosely (and collect data) as if they were. Any particular health care service will usually involve, in its production, the services of several different types of people, as well as various forms of supplies, and capital services. In the economist's jargon, these are the "factors of production", and the price/cost of health care services is the source of the payments to the owners of those factors.

As with the different types of commodities included as health care services, we can identify a large number of different factors which are required for and used in their production. Some are quite specific to health care -- the services of doctors, dentists, and nurses, for example. Others are quite general -- the construction company which builds a hospital building or the pension fund which invests in the shares of a drug company or a commercial laboratory. Still others are in between -- the retail pharmacist who spends part of her time dispensing prescription drugs, and part running a general retail store which is in no sense part of the health care sector (and may even sell tobacco). Such a pharmacist is providing services specialized to health care, but even if working full-time, is only employed part-time in health care.

If we let L represent the total number of different types of inputs or factors of production which are used in producing health care services, then we can designate the total quantity of a randomly selected type of input used by Z_k , where k can take on any integer value between 1 and L . Since these inputs (factors of production, resources) are each owned by someone, who is paid for its use, we can designate by z_{ik} the quantity of input of type k supplied by person i . There will be $N \cdot L$ such values, but most of them will be zero, because most of the people in this society do not work in the health care sector or supply services to it, and of those who do, most provide a very specific range of services.

The z_{ik} refer to some physical measure of input -- person-hours, or furnished square feet per year, or kwh, for example -- and will be reimbursed at some rate of payment per the relevant unit -- which may be explicit (e.g. salaried persons) or implicit (e.g. fee-for-service practitioners). This rate of payment can be labelled as W_k , the going

average rate of payment per unit for input services of type k. As in the case of the P_j above this rate will not always be identified specifically by the prevailing institutional arrangements, but it will be there.⁹

Accordingly we can set up the sum:

$$\sum_k \{W_k * z_{ik}\}$$

which is the total income received by person i from the supply of all types of inputs used in the production of health care services. Person i might be a hospital-based nurse, earning all her income from providing a single form of input -- for her, only one value of z_{ik} will be non-zero. Or she might be a physician, owning a share in the building in which she practices. Her income would include a share of the return to the use of this form of capital. But she might also be a pensioner, whose pension fund owns shares in a commercial laboratory.

Income earned from providing inputs to health care production is thus not exactly the same as the income of those generally regarded as health care workers. Some share of their income is in fact from other sources -- think of the retail pharmacist above, or the physician with investments in non-medical real estate -- and some proportion of the incomes earned (directly or indirectly) from providing health care go to people not commonly identified as directly connected with this sector. When we add up the incomes earned from providing health care, across all the N persons in this society, thus:

$$\sum_{ik} \{W_k * z_{ik}\} \quad \dots(3)$$

we arrive by a third route at an exact measure of the total cost of health care.

We would in fact find that the values of i which corresponded to persons generally recognized as working in health care, do account for most of this total of incomes earned from its production. But the issue of who draws income from this sector is not a trivial one. For example, most of the problem of excessive "overhead" costs in the U.S. health care system -- and they are monumentally excessive, estimated to be as much as \$100 billion greater than in a more conventional system [5] -- corresponds to

⁹. As a generalization, the more clearly the W_k are identified, the more obscure are the P_j , and conversely. Thus in Canada, the prices of physicians' services -- fees -- are generally well defined, but the payment per hour of physicians' time and skill is implicit in their net incomes. On the other hand, while the prices/costs of the different services provided by Canadian hospitals are not explicitly identified, the amounts they pay for each of their inputs are.

the growth in the numbers of non-clinical personnel -- administrators, lawyers, accountants, insurance salesmen, health care researchers, economists -- who have managed to insert their services into this sector, as "necessary" for the production of health care, and thus to draw part or all of their incomes from this source. This significantly increases the range of values of i and k for which the z_{ik} are non-zero.

These three alternative ways of aggregating up to the society-wide total of health care expenditures are different ways of expressing the same thing, not merely different ways of estimating it. They must therefore be equal in principle, even if the inadequacies of real-life data systems may lead to different values from different approaches. And since they are identical, any change in any of the variables in the relationship must be matched or offset by corresponding changes in some of the others.

In particular the conversion of some of the z_{ik} from zero to positive -- new inputs being used and paid for -- must correspond to either a fall in the incomes of some of those currently drawing their incomes from health care, or an increase in the total cost of care (or both). The first might be observed if, for example, more or more highly paid administrators, or better health care research, led either to "greater efficiency" -- a saving in inputs used which would be reflected in a fall in some of the other values of Z_k , and of the incomes of those who supplied those inputs -- or more effective bargaining to lower the values of the W_k . In concrete terms, some other people lose their jobs or have their wages cut, and/or some suppliers lose part of their markets, or receive lower prices for their products (with implications for their workers or shareholders).

If, however, new income claims are added without reducing the established claims, then total costs must rise. This has been the U.S. experience; increases in overhead costs have been translated into increased unit costs of clinical services -- higher P_j . Efficiency gains, regularly predicted on the basis of local experience, have yet to appear in aggregate. Alternatively some concede that costs are increased, but claim that there are significant benefits, very real though difficult to measure (and in any case not currently measured) which correspond to the increase in costs [6,7]. The nature of the services has changed so that the Q_j are different and if properly accounted, would represent "more" output, not just higher prices. This is also a common argument by providers of clinical services. Whatever the pathway, any increase on the income side must correspond to an increase in total expenditures -- a rise in P or Q or both -- and an increase in revenues to support them.

On the outlay side, if the Y_i and q_{ij} are not changed then some or all of the t , C_j , and R_i must increase. Rising costs place upward pressure on tax rates and insurance premiums, and feed calls for increased user charges (supported, for example, by the rhetoric of the need "to encourage consumer responsibility"). Recall that the parameter t in this case is the share of total tax revenue going to support health care. It can increase, while holding total tax rates constant, either by reducing other components of public expenditure or by deferring an increase in overall rates (or a reduction in other

expenditures) through borrowing.

Putting the pieces together:

$$\sum_i \{tY_i + \sum_j (C_j * q_{ij}) + R_i\} = \sum_j [P_j * \sum_i q_{ij}] = \sum_{ik} \{W_k * z_{ik}\} \quad \dots(4)$$

Total revenues raised equal total health care costs equal total incomes earned from the provision of health care. Thus a change in any one of these variables must be accompanied by a change in one or more of the others, in such a way as to maintain this relationship.

Introducing User Fees: What Else Happens?

As hospital and medical care are currently funded in Canada, C_j and R_j are effectively zero. One can pay extra for preferred accommodation in hospitals, and can buy Extended Health Benefit insurance to cover this, but the amounts involved are trivial. We can thus treat such services as entirely tax-financed without any significant distortion. If user charges were then to be introduced for such services, C_j would increase from zero to some positive amount.

It follows then that at least one and perhaps several other variables in equation (4) would have to change. There is room for debate about which variables would in fact change, and also over whether such changes would be desirable in some general sense. But there can be no debate about whether some other variables would change -- as the ecologists point out, it is impossible to do only one thing.

Much of the debate focuses on the relation between user charges and the utilization of health care. Treatments of this topic in the economic literature, working with the elementary "supply-and-demand" framework of the textbooks, are virtually unanimous in assuming that any increase in user charges will result in a decrease in overall utilization, a fall in Q .¹⁰ They are not always specific as to the expected effects on P , but most applications of the "supply-and-demand" framework "predict" a fall (or at least no change) in P . Hence the standard story -- user charges will bring down health care costs.¹¹

Other commentators argue for user charges as ways of moving more money into health care, of raising expenditures in a system which they claim to be "underfunded" [8]. The implicit "theoretical model" underlying this argument is not spelled out in

¹⁰. Many versions of this theoretical framework have been offered which appear to be more sophisticated, but so long as these preserve the central assumption of exogenous "consumer demand", they are simply variations on the same elementary theme.

¹¹. This framework is also used as the basis for an argument that health care costs are escalating because care is "free" to the user, but in fact this extension is illegitimate. The "supply-and-demand" framework is a description, (right or wrong) of a static equilibrium process, and says nothing at all about why the values of variables should be changing over time. There are several examples in the economic literature of writers making up for this silence by assuming that "demand" is simply increasing over time -- "consumers' tastes are changing" -- which would seem to be a classic example of the logical fallacy of petitio principii or begging the question. "Changing tastes" not otherwise qualified, is simply a label for the ignorance of the analyst.

detail, and there are several different lines of argument (all with somewhat more sensitivity to the institutional realities of health care delivery than the naive "supply-and-demand" framework) which support the "prediction" that user charges will raise the overall costs of care.

Each of these different lines of argument can be represented by different patterns of movement of the variables in equation (4). We will refer to some of them below. But in the end, since different theories are available to "predict" that Q will move up, or down (or remain unchanged) in response to a change in C_j , the question is ultimately an empirical one. We review some of this empirical evidence elsewhere [9].

What is not ambiguous, however, is that any changes in the C_j will have redistributive effects. User charges will take money from some, and give it to others. They may also have significant effects upon the distribution of health care, more for some and less for others, whatever their effects upon the overall level of care utilized. Within the framework of equation (4), which must hold for all health care systems, we can represent the distributional effects of alternative ways of financing health care, and show how the burden of payment would be redistributed by various proposed forms of user charges.

User Fees or Taxes: Redistribution among Payers

When there are no user charges or insurance premiums, people contribute to financing health care (roughly) in proportion to their incomes. If user charges do not affect the overall volume of care used,¹² then (holding all prices P_j constant) any introduction of such charges must result in a fall in t . Increased user charges lead to lower tax payments.¹³

Whether person i gains or loses from this change will depend upon the relationship between her income and her use of the various forms of health care, -- Y_i and q_{ij} . It should be obvious that in general those who are healthy and wealthy will gain, and those who are poor and sick will lose, while the effects upon those who are wealthy and sick or healthy and poor are more ambiguous. But from equation (4) we can derive a somewhat more precise description.¹⁴

¹². As discussed elsewhere [9], this is, in fact, the most plausible assumption for hospital care, and probably for physicians' services as well. This is not the same as assuming that people's decisions to seek care are insensitive to price, or in the economists' jargon, a zero elasticity of demand. On the contrary, the empirical evidence is quite clear that prices do affect these decisions, and negatively, just as simple "supply-and-demand" models of behaviour would predict. But the evidence is equally clear that individuals' decisions to seek care do not in themselves determine the overall level of care provided, which depends rather on the capacity and objectives of providers, and the advice they give patients. So user charges, at least within the ranges observed in modern health care systems, may influence which people receive care, but do not appear to influence overall levels of care provided. No doubt some level of charges would be high enough to limit overall use, but if these are politically intolerable, and in any case are beyond any yet seen, then this theoretical possibility is of little practical relevance.

¹³. That is, lower tax payments for health care, and lower than they would otherwise have been. Recall that in equation (4) t is that portion of the overall tax rate which goes to finance health care. It can be raised or lowered, without a change in overall tax rates, if tax revenues are diverted from or to other purposes. In particular a reduction in any public borrowing which would otherwise have occurred, amounts to a reduction in future, rather than present, tax rates.

¹⁴. We are here and throughout referring only to ex post gains and losses in financial position or in some cases in access to services. There will also be ex ante losses of utility resulting from the increased exposure to risk associated with any shift from tax to user charge financing. Insofar as ex ante risk

Holding total expenditures, and therefore total revenues, constant (and assuming no expansion of private insurance), and letting the changes in C and t be designated by ΔC and Δt , we can see that:

$$\sum_i (\Delta t Y_i + \sum_j (\Delta C_j * q_{ij})) = 0 \quad \dots(5)$$

That is, individual i will pay more in user charges and less in taxes, in amounts which are determined by the increases in the various C_j and the decline in t , applied to her income and pattern of use. If we sum over all individuals, the total revenues from user charges will exactly equal the reduction in payments to health care from tax revenue. From this equation we can derive the reduction in tax rates that is made possible by the imposition of user charges. A little bit of algebraic fiddling shows this to be:

$$\Delta t = - \sum_{ij} (\Delta C_j * q_{ij}) / \sum_i Y_i \quad \dots(6)$$

Not very surprisingly it is equal to the total revenue from user charges divided by the total across all individuals of taxable income. But it has a negative sign, because if charges go up, tax rates go down. Using this measure of the change in tax rates, we can then represent the change in person i 's financial position resulting from the introduction of user charges. Her net gain is:¹⁵

$$\sum_{ij} (\Delta C_j * q_{ij}) Y_i / \sum_i Y_i - \sum_j (\Delta C_j * q_{ij}) \quad \dots(7)$$

The first term is the amount by which her taxes used for health care will go down, and the second term is the additional amount she must now pay in user charges. While these amounts cancel out over the population as a whole, it is quite obvious that for most people, expression (7) will be either positive or negative, and can be quite large. Furthermore a simple rearrangement, extracting individual i 's income, Y_i , from both sides of (7), yields:

$$[\sum_{ij} (\Delta C_j * q_{ij}) / \sum_i Y_i - \sum_j (\Delta C_j * q_{ij}) / Y_i] * Y_i \quad \dots(8)$$

in which it may be seen that the first term is the ratio of total user charge payments,

status is correlated with actual use, and uncorrelated or negatively correlated with income, this loss of utility will also bear more heavily on those at lower income levels.

¹⁵. In representing person i 's net gain, we change the sign of the Δt from (6). A fall in tax rates, a negative value of Δt , translates into a positive change in person i 's position -- she pays less in taxes -- which must be set against any increase in the user charges she must pay.

society-wide, to total taxable income, and the second is the same ratio for person *i*. Thus the algebraic fiddling leads to a result which may in fact be intuitively obvious -- person *i* gains from the introduction of user charges if the proportion of her income which she spends on such charges is less than the proportion for the community as a whole. Otherwise she loses.

Yet another way of putting the point, by rearranging terms in (7):

$$[Y_i / \sum_i Y_i - \sum_j (\Delta C_j * q_{ij}) / \{\sum_{ij} (\Delta C_j * q_{ij})\}] * \{\sum_{ij} (\Delta C_j * q_{ij})\} \dots (9)$$

The first term is her share of total taxable income, the second is her share of the increase in user charges for health care. Thus the gainers from user charges are those whose share of taxable income exceeds their share of out-of-pocket expenditures for care. Obviously this includes the "healthy and wealthy", with above average shares of income and below average use of health care (weighted by the imposed charges), while low income, high use persons lose financially. But if one is wealthy enough, one can be an above-average user of health care, and still gain financially from a user charge policy; conversely if one is poor enough, one can be a below-average user and still lose (unless of course one uses no care at all, in which case one is always made better off by increasing user charges and lowering taxes.)

Expression (9) also indicates that the size of ones' gain or loss is proportionate to the total amount collected, society-wide, in user charges. If one gains (loses) from such charges, increasing them simply increases the gain (loss).¹⁶

As is well known, health and wealth are correlated. To the extent that utilization of health care also correlates with ill health, Y_i and q_{ij} will be negatively correlated. Going back to expression (7), this implies that over the population as a whole, the first and second terms will be negatively correlated. Persons with high incomes -- large values of Y_i -- will tend to have lower levels of health care use, and hence lower outlays on user charges. And the reverse will be true for people at lower incomes. Thus the closer the correlation between health and wealth, the greater will be the financial gains of the gainers, and the losses of the losers. The redistributive effects of user charges are thus accentuated.¹⁷

¹⁶ It follows immediately that those who advocate the introduction of user fees in the expectation of personal gain [10] could be expected to support increases in their magnitude once they had been introduced.

¹⁷ Illness is correlated with age, and there are, among the elderly, people with a good deal of wealth even though their current incomes may be relatively low. But these "whoopies" -- well-heeled older persons -- do not represent a large share of

the age groups -- the "oldest old" -- that make heaviest use of health care.

Relaxing the Assumptions, Strengthening the Conclusions

These results follow logically from the assumptions of the model, the critical ones being that health care is funded either from taxes or from user charges, and that overall costs of health care do not change. But they also depend on the construction of equation (4), in which taxes have been specified as proportionate to income, and user charges as fixed amounts of $\$C_j$ per unit of health care of type j . Since the relevance of the results from a model depends upon the validity of that model's underlying assumptions, one might reasonably ask what happens if these assumptions and specifications were changed.

If we relax the assumption that overall utilization does not change, then the impact of user charges on the distribution of the financial burden will be different. But if one were to make the assumption which economists commonly make (on a priori rather than empirical grounds), that overall utilization and costs will go down, then the right hand side of equation (5) will be, not zero, but some negative number equal to the hypothesized fall in total health care expenditures. Carrying this into equation (6), then, tax rates will fall even further. If one replaced the zero with $-K$, then in equation (6) tax rates fall by an additional $K/(\sum_i Y_i)$, and in equation (7) this generates an additional positive term proportionate to person i 's income. The financial savings from reduced utilization are translated into larger gains for higher income people.

Thus assuming a negative response of utilization to user charges reinforces the (financially) redistributive impact of user charges, from lower to higher income people. Only if the introduction of such charges were to increase the total costs of health care, and if these increases resulted in tax increases, rather than in further increases in user charges, could one change this conclusion. And even then, if the overall tax system remained roughly proportionate, the result of an increase in both taxes and user charges would be that while people at all incomes would pay more, the greatest relative increases would be borne by those at lower incomes, because being less healthy, they are more likely to use care.¹⁸

Allowing for private insurance has equally straightforward results. If overall use and cost do not change, then one simply adds a positive $\sum_i (\Delta R_i)$ term to the left hand side of equation (5), while the right hand side remains at zero. The fall in tax rates will

¹⁸. If they can afford it. These comments assume similar changes in use by all individuals. But user charges which were high enough to force lower income people to reduce their utilization significantly, relative to those at higher incomes, or to push them out of the system entirely, could reverse the redistributive conclusions here. After all, the burden of paying for Mercedes-Benzes falls disproportionately on wealthy people. Changes in the pattern of the q_{ij} are considered further below.

be matched by some combination of increases in user charges and private premiums, with the mix depending upon the coverage and comprehensiveness of the private plans. To the extent that private insurance substitutes for user charges, of course, the probability of any (negative) utilization response is reduced. And since experience indicates very strongly that private insurers are incapable of imposing effective cost controls on health care systems, the consequences of a shift to such coverage are likely to be an increase in both the quantity of care used, and the prices paid for it.

But private insurers price coverage according to the risk status, the expected cost, of the person or more commonly the group covered. Thus any increase in private insurance premiums and corresponding reduction in tax rates will transfer financial burdens away from those at high income, and toward those at high risk of care. Since illness, and care use, tend to persist over time, past use is one of the better predictors of future use. Thus private insurers tend to link premiums charged to past use -- experience rating -- and either charge the highest premiums to the least healthy, or simply refuse to insure them. Allowing for the reintroduction of private insurance along with user charges thus reinforces the redistributive conclusions above.

Integrating User Charges with Taxes: More Algebra, Same Result

There are many forms of user charges, and many forms of taxes, and the analysis above concentrates on the simplest form of each. One could impose charges which exempted those at lowest incomes, for example, in which case the redistribution from less to more healthy and wealthy would operate only over part of the population. An often suggested alternative would not impose user charges as such, but would add (all or some portion of) the outlays made on behalf of person i to her taxable income. The left hand side of equation (4) would then become:

$$\sum_i \{t(Y_i + r \sum_j [P_j * q_{ij}])\}$$

The proportion of the cost of care of type j which is added to taxable income is represented by r , whose value would lie between zero and unity. Health care would still be fully tax financed, but ones' tax payments would depend upon use of care.

With this expanded tax base, of course, t must fall to hold revenues constant. Once again, some gain and some lose. If the health expenditures made on your behalf represent a proportion of your taxable income which is less than the initial (i.e. pre-scheme) tax rate to finance health care, your taxes fall, regardless of the value chosen for r . The size of your gain will, however, be greater, the larger is r ; moreover your gain will also be larger, the larger is your taxable income, and the smaller is your outlay on health care.

In the simplest terms, let Y be total taxable income in this society, and X be total outlays on health care. Prior to including them in the tax base, these outlays were fully funded at a tax rate of t , so $tY = X$. If some proportion r of these outlays is included in the tax base, then $t'(Y + rX) = X$, where $t' < t$. Health care outlays can now be funded by a lower tax rate on this larger base. Substituting tY for X in the second expression, we get $t'(Y + rtY) = tY$, or $t' = tY/(Y + rtY) = t/(1 + rt)$. The new tax rate to support health care is lower than the previous rate, by an amount which increases with r .

Now consider person i , whose gain or loss from including a share of health care outlays in taxable income will be the change in her taxes. These will be:

$$t'(Y_i + rX_i) - tY_i = [t/(1 + rt)][Y_i + rX_i] - tY_i,$$

or regrouping:

$$\{t[Y_i + rX_i] - (1 + rt)[tY_i]\}/(1 + rt).$$

Cancelling the terms in tY_i and taking out the common factor rt , her gain becomes:

$$[rt/(1 + rt)][X_i - tY_i]$$

Thus her taxes go up if outlays on health care on her behalf exceed her tax payments prior to the inclusion of such outlays as a taxable benefit, and go down otherwise. The amount of her gain or loss will be greater depending on the value of r , the proportion of outlays so included. But for any given values of r and t , the net gain will be larger, the larger is Y_i and the smaller is X_i .

As a numerical illustration, consider the hypothetical data in Table 2. Persons A and B have incomes of \$20,000 and \$80,000 respectively, and each of them has illness which results in health care costs of \$10,000 during the year. Thus taxes of all types must raise \$20,000 in total to cover the costs of care. Assuming proportionate tax rates (all types of taxes taken together), person A pays \$4,000, person B \$16,000. The overall tax rate (20%) is lower than the proportion of person A's income represented by health care, and higher than person B's. If we add health care to each person's tax base as a taxable benefit, then the overall tax rate needed to cover the \$20,000 cost of care is 16.67%. Person A's tax burden rises; person B's falls.¹⁹

Thus, even though such a proposal appears to respond to distributional issues, it still transfers funds to the healthy and wealthy, at the expense of the poor and ill. And again, the fact that in the Canadian population (and most, probably all, others) health and wealth are positively correlated accentuates the degree of redistributive impact.

What is rather ironic, however, is that proposals for integrating user charges into the tax system originated in the United States, and continue to be discussed there, as ways of redistributing income in the other direction -- from higher to lower income users of care [11,12]. And indeed they would have this effect -- in a system which presently has very substantial user charges. But if one starts from a tax-financed system, as in Canada or most of the rest of the developed world, such proposals have distributional effects which are the exact opposite of those intended by their American designers. It is an unfortunate aspect of the Canadian mentality that ideas and policies are so often imported uncritically from the United States, even when the Canadian environment is obviously and radically different.

¹⁹ This result holds even when different individuals are assumed to have different tax rates t , so long as all tax rates are adjusted proportionately when the base is increased.

TABLE 2: Health Care as a Taxable Benefit: The Rich Get Richer

	<u>Person A</u>	<u>Person B</u>	<u>Total</u>
[A] Taxable Income	\$20,000	\$ 80,000	\$100,000
[B] Health Care Costs	10,000	10,000	20,000
[C] B/A	50%	12.5%	20%
[D] Total Taxes Payable	4,000	16,000	
[E] Income + 'Benefits'	30,000	90,000	
[F] New Tax Rate			16.67%
[G] Total Taxes Payable	5,000	15,000	

What About Tax Credits? A User Charge Is a User Charge Is...

One can also show the impact of another proposed "alternative" to user charges, in which it is suggested that each person be assigned some dollar value of health care "credits" for each year, and required to pay some proportion of all expenditures beyond that level. Those who did not exhaust their credit level, would get a rebate of a portion of the unused credit.

Letting X_i represent person i 's credit level (and thus allowing for the scheme to recognize that different people have very different levels of risk), let s represent the proportion that people must pay of the amount by which their actual expenses exceed the credited amount, and also the proportion of the "unused" credit which is rebated. Then the left hand side of equation (4) becomes:

$$\sum_i \{tY_i + s(\sum_j [P_j * q_{ij}] - X_i) - s(X_i - \sum_j [P_j * q_{ij}])\}$$

The first term multiplied by s is the amount by which health care outlays for person i exceed her credit, and is set at zero if they do not. The second term multiplied by s is the amount of her "unused" credit, and again is set at zero if there is none. But then the two terms taken together reduce to:

$$\sum_i \{tY_i + s(\sum_j [P_j * q_{ij}] - X_i)\}$$

where the second term can be either positive or negative, or setting $sP_j = C_j$:

$$\sum_i \{tY_i + \sum_j [C_j * q_{ij}] - sX_i\}$$

So one is back to a plain, common or garden-type user charge with only the addition of a flat-rate credit of sX_i . This is unrelated to either income or actual use of care, but might be based on risk or expected use.

If the total of the X_i is set equal to overall health expenditures, then user charge revenues and rebates will cancel out over the society as a whole, and tax rates will remain unchanged. The scheme will simply redistribute income among people according to whether their actual use of care exceeds or falls short of the "expected" amounts embodied in X_i . Those who have "unexpected" accidents (the usual kind), or those whose risk status is incompletely reflected in their assigned credit levels -- who have some form of chronic illness, for example, which is not recognized in whatever formula is used to set these individual values -- will transfer funds to those who are healthier than "expected". The more (less) sensitive and sophisticated the process of setting the

individual credit levels, the smaller (larger) will be these gains and losses.²⁰

Like any other user charge scheme, this proposal transfers funds from the more to the less ill. Whether it also transfers from lower to higher income people depends upon (a) whether the rebates do in fact match the charge revenues, and (b) whether the process of setting the individual credit levels provides higher levels to those at lower incomes, sufficient to reflect their higher risk status. If the total of charges exceeds rebates, so that tax rates fall, or if individual credit levels incompletely reflect risk, then there will also be a net transfer from lower to higher income people.²¹

²⁰. Advocates of such a scheme may believe that total utilization and costs of care will fall, which is in fact rather unlikely [9]. Alternatively they may predict, quite plausibly, that even after the credits the scheme will income redistribute from high to low users, and they expect to be among the latter.

²¹. There would also, of course, be administrative problems, and costs, involved in determining individual risk levels. If one relied only on such measures as age and sex, the process would be easy and cheap, and also very incomplete. The result would be large transfers from the chronically ill to the chronically healthy. But if one tried to establish truly individualized risk and credit levels, one would in fact be doing exactly what insurance companies do -- underwriting -- only in reverse. This process is difficult and costly, and provides opportunities and incentives for "false signalling" by both clients and underwriters. In a public system it would also draw considerable litigation and raise some very awkward issues of discrimination. One might predict fairly confidently that the X_i values would not be individualized, and thus that the interpersonal transfers would be considerable.

Recommendations Imply Values -- But Whose?

The principal impact of introducing user charges into a tax-financed system, or for that matter of removing them, is the redistribution of financial resources from some persons to others. One may approve or disapprove of these transfers; that is where the values come in. The expert, economist or otherwise, has no legitimate claim to have his or her values in such matters accorded any higher status than those of any other citizen/voter. But he can, and we do, insist that the fact of this redistribution, and its direction, be recognized.

Anyone who advocates a change in the financing structure of health care, for whatever reason and in whatever direction, is simultaneously advocating an extensive but quite predictable process of income transfer from some members of society to others. It seems only honest that such advocates be prepared to admit, and defend, the necessary consequences of their proposed policies. "A man must be presumed to will the consequences of his own acts" -- this, we believe, extends to "woman", and to recommendations as well as acts.

In particular, those economists who claim that economics as a "science" is value-neutral, so does not make comparisons of utility or well-being among persons, yet who state baldly that user charges are good policy because they "improve allocative efficiency" are in fact advocating such transfers, and therefore implicitly making precisely the interpersonal comparisons which they would have others believe they do not and cannot make. There is nothing wrong with having values -- how else could one recommend anything? But there is something wrong with trying to deceive others (and perhaps oneself), as to what those values are, and trying to cover with the white cloak of "scientific" objectivity, value judgements which one knows or strongly suspects would not be shared by ones' fellow citizens if left naked for their inspection.²²

²². Interestingly, there is empirical evidence indicating that the personal values of economists do differ, on average, from those of the general public [13], making any value-based policy recommendations, or claims of "scientific objectivity", doubly suspect.

Redistribution of Access to Health Care Services

Who pays for health care, and how much, is an important part of health care policy, but obviously not the whole of it. The principal concern of public debate has, in fact, been who gets care, and of what kinds. [One might feel that, since most people want health care, not for its own sake, but because they believe that it will benefit their health, the focus of discussion and of policy should rather be the ultimate objective of health rather than a concentration on one particular pathway to it, through health care, but this idea, though very old, is at best a very new part of health policy [e.g. 14, 15].

A standard argument against user charges is that people will be deterred from seeking care, and that their health will suffer as a result. If we let H_i represent some measure of the health status of person i , then in terms of the variables in equation (4) this argument can be represented as:

$$q_{ij} = F_j(C_j) \quad \dots(10)$$

and:

$$\Delta H_i = G(q_{ij}, H_i) \quad \dots(11)$$

where F and G stand for some functional relationships.

The use of care by person i will be affected by the price she must pay for it, and her health may be improved (depending upon what it was like to start with) by her use of some forms of care. (There are in fact M different relations (10), one for each type of care, and the influence of price on use may be quite different depending upon which is referred to -- strictly speaking there is no such thing as "health care".) Thus user charges, if they deter people from seeking needed care, may be harmful to health, and to their opponents that in itself is a powerful counter-argument. After all, the major purpose of introducing public programs to finance health care in the first place was precisely to enable people to get the care they "needed".

Advocates of user charges, on the other hand, tell a more complex mix of stories. Some, principally those influenced by the mechanical "supply-and-demand" apparatus of elementary economics, accept and indeed celebrate the relationships F_j . (These they identify as "demand curves", which are in reality only a particular special case of a more general class of possible relationships.) But they then ignore completely the relationships (11), simply failing even to meet the arguments of those who are concerned about health consequences. In so doing they are in effect advocating that health care be treated as no different from any other commodity, and implicitly assuming that any influence of health care on health is irrelevant.

As for the impact of user charges on the distribution of health services among the population served, opponents will often assert that charges are more likely to deter people with low incomes, reducing their share of the care provided and accentuating

their health disadvantages. On the other hand those who believe either that health care has nothing to do with health, or that if it does, that should not affect how it is paid for, find such redistributive effects uninteresting. No one worries about access to Mercedes-Benzes, after all. But both lines of argument suggest that by reducing the utilization of particular individuals, user charges will lower overall utilization, and thereby costs, of health care. Each component of equation (4) -- revenues, expenditures, and incomes -- will fall.

Many of the arguments for user charges, however, and particularly those put forward by providers, contemplate an increase in total expenditures [8]. Some are simply trying to add on a user charge to the amounts they are reimbursed by the public plans so as to raise prices of services and their own incomes -- a simultaneous increase in C_j , P_j , and certain of the W_{jk} , with no necessary impact on quantities used. But others tell a more nuanced story, in which the principal impact of user charges is on which services are provided, to whom, rather than on aggregate use or expenditure.²³

The argument is well illustrated by remarks attributed to the premier of Alberta, Ralph Klein, and his Minister of Municipal Affairs, Steve West [16]. The former was said to be proposing changes to the Canada Health Act to permit provinces to impose user charges, without financial penalty, and was quoted: "...steps have to be taken to cut down on abuse and perhaps a small user fee for those who can afford it might be a way to do that." On the other hand Mr. West, in supporting the call for user fees, was more concerned about access to particular expensive services for those willing and able to pay for treatment: "I don't want to retire with half a million dollars in the bank, be 92nd. on the list for heart surgery, and die with all that money in there."

We suspect that these two comments, taken together, represent a much more common view among advocates of user fees than the rather peculiar (to us) position of the neo-classical economist who believes that health care has nothing to do with health, or at least that we should all pretend that it does not. And we single them out, not in any way to target Mr. Klein and Mr. West, but because we believe that they have done us the service of representing this view succinctly and authoritatively. Their rationales do not, however, necessarily add up to a less expensive health care system, only a different one.

²³. The steady accumulation of evidence may also be taking its toll. For those who asserted confidently twenty years ago [8] that user charges were the way, and the only way, to limit cost escalation, the subsequent direct refutation from international experience [9] is at least inconvenient. True believers continue to repeat the litany, but others are modifying the story somewhat to fit the facts.

Those who pursue this line of argument quite clearly believe that health care does matter for health, at least for the health of some individuals in some situations. But they also believe that at present there is a lot of care, and associated expense, that is not effective, and that there is also a lot of care going unprovided, because of capacity constraints in the Canadian system, which would be effective and should be provided. They propose a rearrangement of service use; fewer services for "abusers" (who presumably do not have real needs, or at least not for the care they are seeking) and more for those who really need care (and who might die without it). In terms of the formalism we have been using, they believe that an increase in some of the q_{ij} matched by a decrease in others -- different people, different services -- could lead to a more effective system overall, and that the introduction of user charges would help to bring about this rearrangement.

This leaves open the question of whether overall use and cost would go up or down. Mr. Klein is quoted as saying that the user fee would help offset the growing cost of Medicare, and would only be charged to "those who can afford it", which seems to imply that the primary intention would be to raise revenue rather than to reduce cost. And Mr. West seems to be anticipating an increase in the rate of cardiac surgery, supported by an inflow of private funds. This suggests that the ultimate objective is a health care system with higher overall costs -- which is certainly consistent with the arguments for user charges made by providers -- but more costs borne by users and less out of taxes. A larger number of "needed" complex diagnostic and surgical procedures would be provided for those willing to pay for them, and perhaps less care, hopefully less "unnecessary" care, would be provided to those who respond to charges.

This seems to correspond, point by point, to a move in the direction of the American system. If that is Mr. Klein's objective, then his recommendation seems to make perfect sense -- up to a point. The empirical evidence [9] indicates that user charges do influence peoples' decisions to seek care, in the expected direction, but that they do not appear to influence overall levels of use through this route.²⁴ Rather they

²⁴. It is remarkable how many economists routinely commit the "fallacy of composition" against which we warn students in the first year courses. Relations (10) assert that an individual faced with an increase in the price of a commodity will, everything else being equal, reduce her use of it. And the empirical evidence confirms this. But they do not say that if everyone faces an increase in prices, for all the different commodities, that total use will fall, because in that case everything else does not remain equal for each individual. The aggregate relationship is not necessarily the sum of the individual ones, or, in one of the examples commonly used in first year courses, "If you are in a crowd watching a parade, and you stand on tiptoe, you will see the parade better. Therefore everyone should stand on tiptoe...." The empirical evidence

redirect care from those who are more sensitive to charges, to those who are less.

Their effect on overall use and cost depends upon the base line for comparison. If user charges are a substitute for capacity constraints, price controls, and global expenditure caps (as many in the U.S. have recommended) then they lead to increases in total expenditures, relative to alternative, more effective policies. (As the U.S. has experienced -- but not learned. They continue to rely heavily on user fees.) On the other hand, if user charges are levied along with various forms of direct budgetary control, as in some of the European systems, they have no clear effect one way or another on total costs.²⁵ They do, however, enable the well-to-do to buy their way to the front of any queues that may develop, which seems to meet Mr. West's concern.

Where Mr. Klein's argument leaves the evidence behind and strikes out on its own, however, is in the assumption that the redistribution of services which it implies would represent a more effective health care system. One may be concerned about both the over-provision of ineffective care and the under-provision of effective care, without any presumption that this situation would be improved by user fees. The argument that such charges reduce "abuse" founders on the observation that no one, doctor or patient, seems to be able to define "abuse" [17], although like pornography, they can all think of one or two egregious examples. It should not then be surprising that the available evidence in both Canada and the United States does not support the argument that "free" care leads to more "unnecessary" use, or that user fees lead to less [9].

At the other end of the scale, it should not be surprising that people who have been told that they 'need' heart surgery are willing to pay whatever it costs -- if they have the money. It does not follow that they will necessarily benefit as a result. The U.S. experience with unlimited access -- for those with the money -- is well documented. A lot of people receive services which, in the judgement of external experts, are inappropriate, unnecessary, in some cases harmful.²⁶ But is this "abuse"? Presumably

suggests that overall rates of health care use are primarily determined by the capacity and the objectives of the suppliers of care, not (at least within observed ranges) the levels of charges levied.

²⁵. The only European nation that seems rhetorically committed to the importance of the "ticket modérateur" as a mechanism for cost control, France, is also the country in Europe least successful in overall cost control -- but the direction of causality is ambiguous.

²⁶. Chassin et al. [18] review the relevant literature. Their conclusion is supported by a recent comparison of patterns of management of patients with myocardial infarction (heart attack) in Canada and in the United States. Rouleau et al. [19] found that while such patients in the United States are nearly

all those who underwent such inappropriate procedures believed that they were going to benefit, and some died in the comfort of this belief.

On the other hand, documentation is increasingly emerging to confirm what one might have expected, that those who do not have the money, or the insurance, receive significantly fewer services of specific and identifiable types, and that their health suffers as a result. As supporters of the U.S. health care system often point out, it is not true that people with no insurance receive no care. It is however the case that they receive less care, enough less to make a difference. And so do some who have insurance -- those covered by the Medicaid program for people in poverty.

The relationships between health status and health care are sufficiently diverse and complex to keep busy a large proportion of the international health services research establishment. The point we want to emphasize here, however, is that the impact of user charges on who gets what kind of services is much more significant than their effect on the overall level of use and expenditure. The more a system is funded from user charges, the more the mix of services provided is determined by willingness to pay. But "willingness" is a product of desire and ability.

There is a considerable amount of evidence to support the (rather intuitively obvious) point that user charges reallocate services from those with fewer resources to those with more -- the greater willingness to pay of the latter being a natural consequence of their greater ability to pay. But there is none to indicate that this corresponds to a more effective mix of services, because needs are inversely correlated with ability to pay, and because in any case people do not generally know their own needs. That is after all why they seek professional advice. And while there are good grounds for believing that the quality of the resulting decision-making process could be considerably improved, there is neither a priori logic nor empirical evidence to support the proposition that this would be assisted by user charges.

The rhetoric of "shortages", "cutbacks", and "rationing" which surrounds the process of adjustment of the Canadian health care system to a more slowly growing economy may lead increasing numbers of people to fear that they personally will not be able to get care when they need it. User charges may then come to be seen by more people as a desirable policy precisely because they serve to reallocate care from someone else, to those of us who can afford to pay.

This seems to be Mr. West's point. If there must be a shortage, let it fall on

twice as likely to undergo coronary artery bypass grafting (CABG), there is no difference in mortality rates. Overall CABG rates are about twice as high in the United States, but cardiac mortality is lower in Canada, as it is for most causes of death [20].

someone else. Those willing and able to pay should be served. If many of us do not wish to admit to ourselves what our fears and motivations really are, then the rhetoric of "personal responsibility", "control of abuse", and perhaps a bit of "mutual sacrifice in difficult economic times" will provide a comfortable cover for what might otherwise be a rather embarrassing attempt to get more for "us" at the expense of "them".

In the past, the Canadian community decided, through its political processes, that we wished to redistribute the use of health care from the more to the less healthy and wealthy, and to redistribute costs in the opposite direction. We put in place a financing system to do this. It worked. If now we as a community decide that we want to increase the proportion of health care used by those with relatively more resources, and move the payment burden in the opposite direction, that is obviously also a legitimate political decision. User charges are one way of achieving that result.

There is, of course, nothing sacred or even scientific about the present level or pattern of health care services provided in Canada. Experts of various persuasions have assembled a great deal of evidence to support the general proposition that a reallocation of services -- more of some kinds for some people, less of others for others -- could significantly improve the efficiency and the effectiveness of the Canadian health care system without any increase in overall costs.

Recent Royal Commissions and other provincial public enquiries have reached the same conclusions, recommending in particular a shift in emphasis from acute care hospitals to alternative community-based services. They have also stressed the need for more attention to the outcomes resulting from particular services; much of present activity is inappropriate or unevaluated. Thus trying to rearrange the q_{ij} to improve health outcomes seems widely accepted as a desirable social objective. A number of policies are currently under discussion or in implementation to try to do this.

But there is no basis for any notion that the rearrangement that would follow from user charges would move us in this direction; quite the contrary. Willingness to pay does not correlate with capacity to benefit.

Redistribution between Payers and Providers

Working our way through equation (4), we have considered the redistributive effects of user charges on who pays for health care, and on who receives which services. But any rearrangement of the types of health care provided, or in how they are funded, will also have implications for whose services -- inputs, or factors of production -- will be employed in and paid for providing them. Who gets what jobs, and how much will they be paid?

In Reinhardt's [21] graphic analogy, each person who is paid from total expenditures on health care, represents a "place setting at the health care feast". In our more pedestrian terms the number of places set is the number of z_{ik} with positive values, and the amounts put on each plate are represented by the size of the W_{ik} and the z_{ik} . Any change in health policy will influence the number of settings, the size of the portions, and who is invited to sit.

For example [only], the simple "supply-and-demand" story of the economics textbooks would have us believe that user charges lead to lower use of care overall, which implies fewer resources employed in providing it. The process of adjustment may involve a "short-run" transition stage [whose length is not specified, it could be as long as a working lifetime for some professionals] in which some people find themselves under- or unemployed, some suppliers find their markets shrinking, and competition may lead to a fall in fees, prices, and wages. Average portions, incomes, shrink. But eventually some people leave the health sector entirely -- z_{ik} for them goes to zero -- and/or others choose not to enter. Jobs are lost. The levels of employment and income for the remainder go back to their "equilibrium" level. Some producers of materials and supplies for the health care sector are "shaken out", and move into other lines of business. Portion sizes may be restored, but fewer places are set.

Or so the economics textbooks confidently assert. The reality is usually quite different, and rather more complex. But the general point is valid; any change in the level and mix of output, or in prices, must have balancing effects on the income side. Those who, in defiance of the textbooks, argue for user charges as a "cure" for a perceived problem of underfunding are simultaneously arguing for increased incomes and/or jobs in the health care sector.

The simplest example of a cost-expanding user charge is extra-billing by physicians. When physicians' fees are set by negotiation between physicians and provincial governments, and fully paid by those governments, the introduction of extra-billing would raise both P_j and C_j by the same amount. If, as its physician advocates claim, the process of extra-billing were moderated to ensure that no one was denied care -- only those who "can afford it" being billed -- then the q_{ij} would not change and the whole of the increase in prices would go into increases in W_k -- payments for the

services of physicians.²⁷ This particular form of user charge simply transfers income from patients to physicians.

But this may not be the whole story. Practice overheads may have to increase to cover the extra administrative costs of billing.²⁸ This implies more use of the services of medical office assistants -- an increase in some of the Z_k . The increase in prices and user charges would then be partly absorbed in additional resources used in medical practice, for the same level of service output.

A similar point has been made by hospital administrators in B.C., who have suggested that the small "nuisance" user charges prior to 1986 cost more to collect and account for than they brought in as revenue. Yet larger charges, which might pay for their own collection costs, would in the Canadian context have required the exemption of those who could not "afford" them -- hospital users tend to be elderly, elderly people tend to be poor, poor people tend to be sicker -- and this adds further administrative expense to determine who these are. The ambiguity of hospital administrators on this issue is understandable.

Such increases in input use represent a reduction in the technical efficiency of medical practice. In economic jargon, one refers to a "production function" which links the inputs or resources used in production, with the outputs of goods and services that are of value to users:

$$Q^j \leq E_j(Z_k) \quad \dots(12)$$

Thus the total output of each type of service j is dependent upon the amounts of the different inputs Z_k used in its production. This is in one sense a "technical" relation, like (11) above which was a "production function" for health. Both are based upon the possibilities, in presently known technologies, for combining physical inputs, time, and know-how to produce certain outputs. Some minimal set of inputs is required to produce any specified level of output: one can use more, but not less.

Equation (10), by contrast, purports to describe how people will behave in response to certain circumstances. In fact, however, both (11) and (12) have a

²⁷. Most economists would probably take a less charitable view of extra-billing "only those who can afford it", regarding this as price discrimination according to the elasticity of demand, and recognizing the profitability as well as the political appeal of the tactic.

²⁸. In February 1993, certain physicians in North Vancouver who had opted out of the provincial plan pursuant to a fee dispute, began extra-billing their patients on the grounds that opting out had added to their administrative costs.

significant behavioural component as well. An exact relationship between a set of inputs and a corresponding level of output, (12) written as an equation, is commonly assumed in theoretical economics. But this corresponds to perfect (technical) efficiency, no waste, and state-of-the-art technology. It is an abstraction rarely if ever observed in practice, in any industry. Rather (12) holds as an inequality, a boundary condition, with the extent of departure from potential equality representing the degree of technical inefficiency, wasted resources, in the production process. Behavioural considerations obviously enter into the determination of this degree of inefficiency.

This behavioural component is well illustrated in the extra-billing case. If the inputs used increase -- more office overhead -- and the service output does not, then the right-hand side of relation (12) increases while the left-hand side does not. The production process becomes less efficient. But the costs of this decreased efficiency are borne by users of care, in the form of out-of-pocket costs, while the practitioner gains through fee increases which are large enough to pay for the increased overhead costs, and still yield a higher net income.²⁹

Other adjustments are possible. Physicians sometimes argue that the fees negotiated with provincial payment agencies are too low to permit them to spend adequate time with their patients. To make an adequate income, they must speed patients through too fast to provide adequate care.³⁰ By extra-billing, they are able to

²⁹. In elementary "supply-and-demand" models of economic behaviour, the possibility of the supplier gaining from reduced efficiency is never considered, because it is assumed (usually implicitly and sometimes unconsciously) that suppliers are motivated strictly by profit considerations, and sell their products to informed buyers in competitive markets without regulation or collusion. If all these assumptions held, then indeed such behaviour would simply result in patients moving away from the extra-billing physician. But these textbook assumptions, while analytically convenient, are very far from reality. Moreover physicians are acutely aware of this possibility, and take steps collectively to minimize such "market" responses -- as any rational suppliers would. Nor should one focus only on physicians; when "lean and mean" private sector industries are studied, they too show significant departures from equality in relation (12), because the textbook assumptions are at best an approximation there too.

³⁰. Such an argument is most interesting. If taken literally it implies both that physicians can exercise considerable influence over their own patient flow, thus contradicting the standard economists' assumption of exogenous demand, and that they are prepared to use that influence, even to the detriment of their patients, in order to meet self-determined income targets.

spend more time with each patient. If this argument is valid, then one would expect extra-billing to result in a fall in the total volume of patient contacts -- though not as a result of patient decisions. [As noted above, not every inverse relation between prices and quantities represents a demand curve.] To the extent that increased prices were offset by decreased quantities, the increased costs borne by users of care would be partly offset by reduced costs to taxpayers, while physicians would receive their gains partly in greater incomes, partly in reduced patient loads.

But healthy Canadian taxpayers might be quite short-sighted to conclude that they would gain from a return to extra-billing. Physicians' representatives were quite explicit, in their opposition to the Canada Health Act, in arguing that the right to extra-bill, whether or not exercised, was a "safety valve" to protect them against unfairly restrictive government controls on their fees. This implies that the political pressure created by increased extra-billing could force up negotiated fees paid by the public plans. In this case the prices of services, the P_j , would rise by more than the C_j , implying an income transfer from taxpayers to physicians levered out by the discretionary use of user charges.

This seems to be exactly what happened in Ontario in the late 1970s and early 1980s. Nor did Canadian physicians fail to notice that Quebec, the one province in which extra-billing was effectively banned during the 1970s, also had by far the lowest increase in fees.³¹

Physician extra-billing makes the linkage from user charges to the relative incomes of providers quite simple and direct. But the causality may be more indirect. Consider the process of wage negotiation in a public hospital system. Wages for workers translate directly into implicit prices and then into tax rates -- W , P , and t rise or fall together. But suppose that a hard-nosed or hard-pressed government simply refuses to adjust hospital budgets to cover a negotiated wage settlement, so that total costs do not change. Higher wages now mean fewer positions -- W up and Z down.

³¹. Dentists in Canada illustrate another connection between user charges and overall fee/price levels. Dental insurance is predominantly private, with public subsidies, and includes some degree of coinsurance, allegedly to hold down overall costs. But patients, when they are asked to pay their share of a dental bill, are generally not aware of the relation between the reimbursed amount, the coinsurance rate, and any additional amount which the dentist may have added on. Thus extra-billing is much less noticeable than it is in the physician's office where the patient does not normally expect to pay at all. Requiring the patient to pay some small amount may or may not reduce utilization, but it probably on balance increases prices.

Hospitals may try to do the same work with fewer people; but if they cannot or will not, then P rises and Q falls.

Depending upon the way in which "demands" for service are generated, the fall in Q can lead to perceived "shortages", queuing, and rationing, and pressure for a spill-out of workload into a private, self-paying sector. This may be, in effect, what Mr. West was advocating in his remarks quoted above, and it is a common argument. Let those who are willing to pay for extra services, pay for them, and thus expand the capacity (and cost) of the whole system. In the process, one would see an increase in the incomes of the fee for service practitioners -- surgeons and anaesthetists, for example, or radiologists -- who provide the additional services privately, quite possibly at fees above the provincial schedule, and more employment for hospital support workers in the privately supported facilities.³²

In this case, the development of a private service would in essence be "ratifying" the wage increases in the public sector, converting their effects from a loss of output and jobs, into an increase in total system costs, and a shift in both cost bearing and access. Back at the health care feast, those still at the public table would have succeeded in getting larger portions while those displaced have moved to a newly opened private table.

This form of cost escalation supported by user charges has not been an issue to date in Canada, unless one considers the occasional political theatrics over patients going to the U.S. a nascent "private sector". But it has been a serious problem in the U.K., and a threat in Sweden. The problem is particularly severe if the same persons, usually physicians, are able to work in both the public and the private sectors; they are then able to manipulate access to the former to steer patients to the latter, where prices are higher. The best of all worlds is to be served at both tables.

³². Conversely, fee for service practitioners who are dependent upon hospital services lose income from any restrictions in throughput. This is most obvious during a hospital strike, but any limitation on the levels of diagnostic or therapeutic activity in hospitals will have a negative effect on the billable workloads of some physicians. Fee-for-service physicians are thus economically at risk in negotiations over hospital budgets and capacity levels, so one should not be surprised to find some of them supporting the development of privately funded facilities.

Redistribution among Providers: Gainers and Losers on the Supply Side

Salaried workers in the Canadian hospital sector appear collectively firmly opposed to user charges or the development of a privately funded delivery system. But their economic interests may be somewhat ambiguous. The emergence of such a system might in fact provide powerful support for the contention that the public sector is not adequately funded.³³ On the other hand, even if more private funding did bring in more resources, raising total incomes and employment in the health sector, the distribution of these benefits among different suppliers might be very unequal and few may "trickle down" to the salaried hospital worker. When the mix and volume of care is determined by willingness to pay rather than professionally judged need, the pattern of inputs used, and their rates of pay, will change. Some providers lose, even if total costs/incomes rise.

Again the U.S. is instructive, and in general the changes are rather what one would expect. Willingness to pay is highest among those who have most resources, and feel themselves most at risk and least able to judge the value of the service. American physicians in general have the highest incomes in the world, relative to the general income levels in their society, while those physicians who provide services on referral (sometimes from themselves, self-dealing), and who serve the well-endowed or well-insured -- the diagnostic specialists, cardiologists, thoracic and ophthalmological surgeons -- are particularly generously paid.

Primary care specialists, on the other hand, who depend more on patient-initiated demand and have less scope for recommending mysterious and expensive interventions, have lost a lot of ground economically. Not surprisingly, young physicians are abandoning these areas and specializing where the money is. Those who fear a resulting "shortage" of primary care physicians are applying a professional standard for appropriate levels of care -- patient needs. But the marketplace is instead responding predictably and powerfully to the criterion of willingness to pay. The effects on providers, the pattern of gainers and losers, are the invisible hand's way of encouraging resource owners to redirect their resources to the "most valued" forms of output -- on that criterion.

At the same time, efforts to control costs in the U.S. have led to very substantial

³³. Much of the pressure for major reform in the U.K. National Health Service has arisen from the perceptions, right or wrong, of politicians that physicians and other hospital workers have shared a common interest in low productivity in public hospitals. Waiting lists place political pressure on governments to increase health service budgets, while steering patients to the private side of hospital consultants' practices.

decreases in the use of inpatient care. This has not affected the trend in overall costs, which keep climbing, but jobs and incomes have moved out into the ambulatory care sector. More of the income goes, not only to physicians, but also to technicians, and to the corporate suppliers of equipment, reagents, and drugs -- those who support the "mysterious and expensive" services. The more traditional inpatient roles -- ward nursing, housekeeping, dietary -- are shrinking. There are thus economic as well as ideological grounds for hospital unions in Canada to be very suspicious of user charges; their members may not be the ones who gain. Portions at the private table tend to be much more unequal than at the public, and some of those displaced from the public table find no place there.³⁴

Redistribution to New "Providers"-- Setting New Places at the Health Care Feast

But many others do, who would not be seated at the public table. If extra-billing is "necessary" to cover the extra administrative costs of opting out of provincial medical care programs, then a general re-introduction of physician- or hospital-administered user charges would obviously have a more significant impact on total resources used in this sector. User charges administered by governments could require the setting up of administrative systems to determine who was to be charged, and how much; the enforcement process would also be rather interesting. On the whole, feasibility considerations suggest that any user charge system would be either de facto provider administered, or integrated with the income tax collection process.

But if the prevalence of user charges became so widespread as to lead to the re-introduction of private insurance, then one would also see an increase in L , the numbers of types as well as amounts of inputs Z used in the production of health care. The overhead costs of financing and managing the delivery of health care which are borne by providers, and which are significantly increased when funding comes from multiple private sources, are embodied in the P_j , the unit prices of the services actually provided to patients. But a private insurance system, in which the R_i are increased to take up some of the increases which would otherwise occur in the C_j , could as described above be represented either as adding to the costs of health care services themselves or as adding additional types of services, such as "risk bearing".

These "services" require resources for underwriting, marketing, claims

³⁴. Fuchs and his colleagues have found that the incomes of physicians (relative to the rest of the workforce) are higher in the United States than in Canada, although their workloads are lower [22]. Hospital workers as a group do not have higher (relative) wages in the United States, but this is because higher wages for the higher-skilled in the U.S. are offset by higher wages for the lower-skilled in Canada [23].

administration, etc., none of which are needed in a universal public system. In equation (4) above these activities can be represented either as new Q_j , with their own P and Z counterparts (increasing M as well as L), or as increased prices, the P_j , and resource inputs for the existing pattern of services. In the former representation, the additional costs correspond to additional services, but these do not enter into any of the relations (10). Insurance salesmen do not in themselves make much contribution to anyone's health, separately from the care which their policies may reimburse [nor, for that matter, do economists].³⁵

Alternatively, if the cost of insurance administration is included in the prices of health care services themselves, then it is recorded as a decrease in the efficiency of production of such services. Either way, relative to a universal public system, these activities represent "cost without (health) benefit" [5].³⁶

³⁵. As noted above, advocates of private coverage have on a priori grounds alleged indirect benefits corresponding to the extra cost, in the form of either improved efficiency or effectiveness of health services themselves, or increases in consumer satisfaction. But these "benefits" are seen only with the eye of faith; actual experience is that private insurance coverage is associated in aggregate with high costs, questionable effectiveness, and low expressed satisfaction.

³⁶. So why does anyone ever buy insurance? Recall that at the outset, we pointed out that this paper analyses distribution ex post, at the end of the day or year, when all events are known. And in fact no one does buy or sell insurance for last year. In this sense, we may appear to have rigged the game against private insurance. But the baseline for comparison is critical.

Insurance is not an efficient way of buying health care services, in a world of certainty. What insurance buys, in an uncertain world, is risk reduction for the individual, at the (certain) price of higher overall costs for the group. If our baseline is a world where $t = 0$ and $C_j = P_j$, for all j , then the individual may be subject to a high degree of financial uncertainty, and insurance may be a good buy even if it raises total costs of care.

Indeed, it might seem as if we knew that this was so, because people did voluntarily buy such insurance prior to the public plans, and in the U.S. they still do. Unfortunately we cannot draw this conclusion with complete certainty, because in both cases there were large public subsidies through the income tax system to encourage such purchases. We do not know how much private insurance would have been (would be) purchased, if buyers had to pay its full cost.

But if instead one's baseline is a world of universal comprehensive public coverage, then the (financial) risk which

The massive and rapidly increasing overhead costs of the American health insurance system are a natural consequence of the modes of finance chosen. They do not all follow directly from the important role played by user charges in that system, in the sense that overheads in the American system pay for much more than simply the levying and collection of direct charges to patients. But such charges probably are partly responsible, indirectly, for the extraordinary level of administrative waste in that system.

In the first place a substantial share of private insurance -- the so-called "Medigap" market, is for coverage against large user charges for the care of the elderly. The public system imposes such charges, allegedly to discourage excessive use, and then the tax system subsidizes the purchase of private insurance to cover those charges, at considerable administrative cost. Each component of the policy looks rational: their sum is transparent lunacy. If user charges were eliminated from the public system, the private coverage would disappear, at considerable savings all round. But of course taxes would be increased....

In addition to inducing this inefficient mix of public and private coverage, however, user fees contribute to the relentless escalation of costs in the U.S. health care system, which is unmatched in the rest of the world. This escalation has spawned an extraordinarily wide, and expensive, range of institutions and mechanisms in response. Some are intended to limit overall expenditures; these have failed. Others are intended simply to push the costs onto someone else; some of these have been more successful.

But each such effort either to control overall costs, or to push them onto someone else, against predictable resistance in each case, has added to the overall cost of the system. The struggle draws in more administrators, lawyers, marketers, actuaries, economists -- more and more places, and well-fed ones, being set at the health care feast. These additional incomes added on the right-hand side of equation (4) become reflected, as described above, in increasing values of the P_j and/or the Q_j .

the individual would otherwise bear is already being pooled at the highest collective level. In this context, the introduction of user charges and private insurance coverage corresponds to an increase in individual risk-bearing (since private coverage is rarely as comprehensive as public, and can never be universal). The administrative costs of such a private system then represent not only "cost without benefit", but "cost with loss" for the community as a whole. For those individuals whose taxes fall by more than the increase in their private premiums and user charges, the healthy and wealthy again, the private system may be preferable. As it may be for the wealthy and unhealthy (so long as they are not too unhealthy) who want to be sure that any limitations on access do not bear on them.

User charges play a central role in this cost-shifting process, both because the imposition of such charges is a major component of that process, and because the adoption of more effective and less wasteful institutions for cost control has been held off, politically, by constant repetition of the claim that user charges rather than direct political action will deal with the problem. False hopes lead to false remedies; and the administrative "therapies" have become more and more expensive as the situation deteriorates.

Canada is not the United States; the point is not that "It can happen here!" [although in fact we believe that it could].³⁷ Rather the U.S. environment provides the best illustration of how user charges can contribute to redistributing incomes within the health care sector itself, by providing opportunities for a number of other "factor suppliers" -- providers of administrative and financial services -- to increase the incomes they draw from the health care sector. As they do so, either total costs go up, or the incomes of those who provide clinical services must fall.

Again, it is not accidental that the American College of Physicians, the second largest physicians' organization in the U.S. after the A.M.A., came out several years ago in favour of a national health insurance plan. Its members have noticed their portions being squeezed, despite the continuous increase in overall costs, by the rapid increase in the numbers of non-clinicians being served, as well as by the overflowing plates of the surgeons and radiologists.

Nor are sellers of administrative and financial services the only "health care wannabees" who might quite reasonably hope to gain from some form of private funding for health care. There are also "wannabees" whose ambitions focus more specifically on the Q₂, as they offer services which are on the fuzzy boundary of the health care system.

Chiropractors and naturopaths, for example, have no doubts that they are providers of health care, and these occupations display many of the features of the health professions. But their services are not within the scope of the Canada Health Act; provinces need not cover them to meet the federal standard of comprehensiveness. Many physicians do not recognize them as "members of the health care team"; quite the contrary. And the public are split; while some use such services routinely, others do not contemplate any circumstances under which they would visit a chiropractor or naturopath.

Correspondingly, the services of such practitioners are reimbursable at provincial

³⁷. And we note that there have been, from time to time, statements from people in the Canadian insurance industry in support of a re-introduction of private funding, both user fees and private insurance.

option, on whatever terms provinces choose. Provinces can impose user charges and limits on the number of reimbursable visits and procedures, and do. These practitioners regard such restrictions as unfair, placing them at a competitive disadvantage relative to physicians. They argue that either their services should be fully and comprehensively reimbursed, or -- given the fiscal realities of the times, user fees should be imposed on physicians' services as well "to level the playing field". The explicit intent of the user fee in this case is to increase the flow of income to this group of practitioners -- factor suppliers -- at the expense of physicians and/or their patients.

Such practitioners are at the edge of health care as currently defined by the reimbursement system; many others who claim to offer health-enhancing services are outside that definition. They deal with their clients on a straight-forward market basis. All would benefit from inclusion within the public reimbursement system, and have an economic interest in seeing that system made more expensive, relative to their own services. While very diverse, this group is not large. But it represents still another form of interest in the expansion of user charges for "conventional" hospital and medical care.

Conclusion

The principal effects of user charges are straightforward income transfers -- what Peter loses, Paul gains -- and whether one thinks this is a good or a bad policy depends ultimately on what one believes about the relative deservingness of Peter and Paul. [This in turn will be in part correlated with whether one is Peter or Paul.] In this paper we have developed a comprehensive accounting framework within which one can keep track of all of these transfers, and have used that framework to trace out the effects of several suggested forms of user charges. One can demonstrate quite rigorously that such charges consistently transfer income (net of health costs) not only from the less to the more healthy, but also from those with lower to those with higher incomes. Arguments to the contrary, based on a partial or incomplete specification of the impact of charges, are shown to be erroneous. In effect they fail to address the question "What happens next?" or "How do the accounting identities stay balanced?"

The accounting identities also force one to be explicit about the relationship between user charges, and the incomes of and opportunities for different types of service providers. To the extent that such charges change either the volume or the mix of services used/provided, or the prices received, they must simultaneously change the incomes and/or work opportunities of providers. Who gets paid, and how much, for supplying their time, their skills, or their capital to the health care sector? If user charges are the first step towards an environment with more private forms of financing for health care, this will have significant effects on the mix of inputs used in its provision, and particularly its administration and financing, and on the rates of reimbursement received by those suppliers.

Here the issues are in part distributional -- providers on average will gain relative to the rest of the community, but some will gain and others may lose. But the mode of payment can also have a significant bearing on the total volume and cost of resources used up to provide a given amount of health care -- on the technical efficiency of the system. Some systems cost a lot more to run than others, paying high rewards for dispensing paper as well as pills. And paying people to push either paper or pills, when there is no corresponding improvement in anyone's health, is waste pure and simple. Enhanced market opportunities in the financial sector come at the expense of overall losses, on average, for everyone.

The effects of user charges on access to care, and ultimately on health, cannot be represented within the accounting identities linking financial flows. As noted above, however, evidence seems to support quite strongly the intuitively obvious conclusion, that user charges will increase the access of those with greatest willingness/ability to pay, at the expense of those with less [9]. Some of their advocates are explicit that this is the objective. Such changes in the types of services provided, and to whom, may have effects on the health of recipients -- and non-recipients -- raising questions about the way in which different payment systems influence the translation of health care services into health outcomes. To what extent do the people who "need" care, in the sense of having the capacity to benefit, actually get the care they need, and how might this be affected by user charges?

Again other evidence -- principally but not entirely from the United States -- shows that willingness or ability to pay is not correlated with (expert-judged) need for care [9]. People who are deterred from care-seeking by user charges are as likely to forego "needed" as "unneeded" services. And (again in the United States) the negative health consequences for those "deterred" are increasingly documented. In the Canadian system (unlike the American) roughly ninety percent of all Medicare expenditure is for services requiring the explicit recommendation of a physician/ It is thus difficult to argue that patients making "frivolous" demands, whether knowingly or otherwise, can generate more than one or two percent of overall costs without the support or at least the acquiescence of a physician.

But these considerations, though very important, take us beyond the scope of the financial relationships. What those relationships demonstrate quite clearly, by taking full account of all the necessary responses and adjustments (necessary in a logical, not a normative, sense), is that whatever else they do, user fees redistribute income. And they consistently redistribute, not only from the sicker to the healthier, but from those with lower to those with higher incomes. In this light, the concentration of advocacy among the well-to-do and their representatives makes perfect sense.

REFERENCES

- [1] Letsch, S.W. (1993) "National Health Care Spending in 1991" Health Affairs 12:1 (Spring) pp. 94-110.
- [2] Schieber, G.J., J.-P. Poullier, and L.M. Greenwald (1992) "U.S. health expenditure performance: An international comparison and data update" Health Care Financing Review 13:4 (Summer, 1992) 1-87.
- [3] Bardin, S., I. Morrison, R. Mittman, A. Saveri, G. Schmid, and J. Wayne (1992) Health, Health Care, and Health Policy: An International Perspective (Special Report SR-472) Menlo Park, CA.: Institute for the Future (July 31).
- [4] Evans, R.G., M.L. Barer and G.L. Stoddart (1993) "It's Not the Money, It's the Principle" Toronto: The Premier's Council on Health, Well-being and Social Justice Discussion Paper (forthcoming).
- [5] Woolhandler, S., and D.U. Himmelstein (1991) "The Deteriorating Administrative Efficiency of the U.S. Health Care System" New England Journal of Medicine, Vol. 324, no. 18 (May 2) pp. 1253-8.
- [6] Danzon, P. (1992) "Hidden Overhead Costs: Is Canada's System Really Less Expensive?" Health Affairs 11:1 (Spring) pp. 21-43.
- [7] Thorpe, K.E. (1992) "Inside the Black Box of Administrative Costs" Health Affairs 11:2 (Summer) pp. 41-55.
- [8] Barer, M.L., V. Bhatia, G.L. Stoddart and R.G. Evans (1993) "The Remarkable Tenacity of User Charges" Toronto: The Premier's Council on Health, Well-being and Social Justice Discussion Paper (forthcoming).
- [9] Stoddart, G.L., M.L. Barer and R.G. Evans (1993) "User Charges, Snares and Delusions: Another Look at the Literature" Toronto: The Premier's Council on Health, Well-being and Social Justice Discussion Paper (forthcoming).
- [10] Evans, R.G., M.L. Barer, G.L. Stoddart and V. Bhatia (1993) "Who Are the Zombie Masters, and What Do They Want?" Toronto: The Premier's Council on Health, Well-being and Social Justice Discussion Paper (forthcoming).
- [11] Feldstein, M.S. (1971) "A New Approach to National Health Insurance: The Public Interest (Spring) pp. 93-105.
- [12] Rice, T. and K.E. Thorpe (1993) "Income-Related Cost Sharing in Health Insurance" Health Affairs 12:1 (Spring) pp. 21-39.

- [13] Frank, R.H., T. Gilovich and D.T. Regan (1993) "Does Studying Economics Inhibit Cooperation?" Journal of Economic Perspectives 7:2 (Spring) pp. 159-71.
- [14] Evans, R.G., and G.L. Stoddart (1990) "Producing Health, Consuming Health Care" Social Science and Medicine, 31:12 pp. 1347-63.
- [15] Barer, M.L., R.G. Evans and T.R. Marmor (1993) Why are Some People Healthy and Others Not? Hawthorne, N.Y.: Aldine-de Gruyter, forthcoming.
- [16] The Toronto Globe and Mail (1993) "Alberta Premier to Press for Medicare User Fees" Wednesday, January 13, p. A3.
- [17] Woodward, C.A., J.R. Gilbert, R.S. Roberts et al. (1983) "When is a Patient's Use of Primary Care Services Unwarranted?" Canadian Medical Association Journal 129 pp.822-7.
- [18] Chassin, M.R., R.E. Park, A. Fink et al. (1986) "Indications for Selected Medical and Surgical Procedures -- A Literature Review and Ratings of Appropriateness (R-3204/2-CWF/HF/HCFA/PMT/RWJ)" (May) The Rand Corporation, Santa Monica, CA.
- [19] Rouleau, J.L., L.A. Moyé, M.A. Pfeffer et al. (1993) "A Comparison of Management Patterns after Acute Myocardial Infarction in Canada and the United States" New England Journal of Medicine Vol. 328, no. 11 (March 18) pp. 779-84.
- [20] Nair, C., R. Karim, and C. Nyers (1992) "Health Care and Health Status: A Canada-United States Statistical Comparison" Health Reports, Vol. 4, no. 2 (October) Ottawa: Statistics Canada (cat. no. 82-003), pp. 175-183.
- [21] Reinhardt, U.E. (1982) "Table Manners at the Health Care Feast" in D. Yaggy and W. Anlyan, eds. Financing Health Care: Competition vs. Regulation Cambridge, Mass.: Ballinger, pp. 13-34.
- [22] Fuchs, V.R. and J.S. Hahn (1990), "How Does Canada Do It? A Comparison of Expenditures for Physicians' Services in the United States and Canada", New England Journal of Medicine 323(13):884-890.
- [23] Redelmeier, D.A., and V.R. Fuchs (1993), "Hospital Costs in the United States and Canada" New England Journal of Medicine Vol. 328, no. 11 (March 18) pp. 772-8.