

Incentives to be healthy: An
economic model of health-related
behaviour

Stephen Birch D. Phil.

Greg Stoddart Ph.D.

Centre for Health Economics and Policy Analysis
Department of Clinical Epidemiology and Biostatistics
McMaster University, Health Sciences Centre
1200 Main Street West
Hamilton, Ontario
Canada
L8N 3Z5

Based on a paper presented at the First European Conference
on Health Economics, Barcelona, Spain, 19-21 September 1989.

Abstract

Increasing attention is being paid to the non-health-care influences on population health status. Policy recommendations emerging from this interest have tended to focus on providing equal opportunity to attain good health. In this sense policies have tended to be 'enabling' with little attention paid to the incentives facing individuals to adopt the healthy behavioural patterns.

In this paper we develop a model of individual behaviour based on standard economic theory. First order conditions are derived for the optimal level of participation in health-related activities. Particular attention is focused on the impacts of behavioural patterns on future health and the consequences for individual well-being. The literature of other behavioural sciences is drawn upon to enhance the ability of the model to explain and predict individual behaviour. The predictions are found to be consistent with observations of individual behaviour reported in the literature.

The policy implications of the model are discussed. Attention is drawn to the need to understand the health-related behaviour of the victims of poor health as opposed to blaming them for not responding to opportunities to behave in ways more conducive to producing health. Otherwise well-intentioned policies aimed at reducing inequalities in health may be counterproductive. If policy makers are interested in changing individual behaviour then it is essential that policies are designed which make such behavioural change in the best interests of the individuals concerned.

1. Introduction

As policy makers around the world struggle to find ways of controlling the increasing demands on health-care systems, attention has been extended to non health-care influences on population health status. This increasing policy interest in environmental and behavioural influences on health status is probably the result of inter alia

- a) observed continuing differentials in health among population subgroups (in particular social classes) despite alleged improvements in access to health care (Royal Commission on the National Health Service 1979, UK Department of Health and Social Security 1980, Whitehead 1988);
- b) an increasing acceptance that the effect of health care on health status at the margin is low in developed economies, both in absolute terms and relative to the marginal impact of other social programmes (i.e. flat of the curve medicine, McKeown 1976, Auster et al. 1969, Enthoven 1980);
- c) observed differentials in health status among 'the affluent classes' (Marmot 1986, Marmot and Morris 1984) suggesting the poverty per se is not the only causal factor of health status differentials and hence that considerable reduction or elimination of health differentials will require more than just anti-poverty programmes.

Nevertheless this shift of attention away from how health care can be used, or its production reorganised to improve health-service performance, and towards how individual and societal behaviour can be changed to improve population health status, appears to have had little impact on health economics research. Research interests appear to remain concentrated on the analysis

of health care (i.e. 'downstream') problems, whether at the level of the individual procedure or the health-care system. What has become known as the 'topic' of health economics (Culyer 1981) might therefore be more accurately described in most cases as health-care economics (Evans 1984).

The primary objective of this paper is to consider what the discipline of economics has to offer in understanding the causes of health-status differentials. We develop a model of individual behaviour, based on an extension of existing economic theory, and consider its potential for use in predicting differentials in lifestyles and identifying efficient policies for reducing health-status differentials with particular reference to health-related behaviour.

The context of the paper is the emerging policy debate in Canada concerning inequalities in health across income groups and the role of incentives in the determination of lifestyles. Recognition of non health-care influences on health status in Canadian policy debate goes back at least to the Lalonde Report (Lalonde 1974). Yet it is only in the last few years that the reduction of inequalities in health status has been identified as a major challenge facing Canadian policy makers (Health and Welfare Canada 1986). Our concern is that this policy debate could learn from the UK experience by avoiding simple 'single cause - single effect' type explanations that implicitly underlay policies of health-care provision free at the point of delivery (e.g. UK National Health Service, Canadian Medicare) and instead focus on multifactorial models of health-related behaviour.

The Black Report (UK Department of Health and Social Security 1980) highlighted the weaknesses (and consequences) of following this 'single cause - single effect' approach by emphasising the unequal distribution of various other factors which may be

determinants of health. Notwithstanding this, we argue that the policy recommendations contained in the report were based on an implicit model which took a similarly restrictive approach to behaviour determination. In particular the observed poorer health status of less affluent groups is assumed to be caused by their lower incomes and knowledge, and the provision of more income and/or education is consequently assumed to generate behavioural changes which enhance health status.

In this paper we focus attention on the determinants of health-related behaviour (i.e. 'upstream' problems), defined as behaviour which although undertaken primarily for reasons other than the impact on health status, nevertheless has an impact, either positive or negative, on the health status of the individual undertaking the behaviour. We examine whether the analytical tools of economic theory can be applied to the determination of individual behaviour in order to predict the differential incidence of health-related behaviours across social groups and the consequential differentials in health status ceteris paribus.

We do not suggest that the economic model to be developed can provide explanations of all observed behavioural patterns. Indeed we shall begin to draw upon the subject matter of other behavioural sciences as part of the analysis. Further we argue that the economic methodology, when applied broadly to encompass features of the sociological and psychological literature, can be used to explain some of the observed differentials in ways which have not to date been recognised or identified, in either the health economics literature or the policy debate.

Our goal is to model the behaviour of the 'victims' of poor health as opposed to 'blaming' them for the consequences of their behaviour. If the reduction of inequalities in health in society is to be adopted as an important social goal, then more attention

should be paid to ensuring that lifestyle incentives to individuals are consistent with this goal. Otherwise unhealthy lifestyles among subgroups of the population may simply reflect behavioural patterns which are in the best interests of these groups, even though they are inconsistent with social goals, and may appear irrational to observers isolated from the influences which generate the behaviour.

2. Understanding inequalities in health: a post war-perspective

In general, concern with inequalities in health has tended to generate policies aimed at the provision of health-care services. In the UK, post war concern with equality of opportunity was reflected, inter alia, in the introduction of the National Health Service (NHS) in 1948.

The objective of the NHS was to ensure that

"every man, woman and child can rely on getting the best medical and other facilities available; that their getting them shall not depend on whether they can pay for them, or any other factor irrelevant to real need."

(Ministry of Health 1944)

The method by which this objective was to be achieved was essentially the removal of user prices at the point of service delivery. Such an approach, it was believed, would lead to previously unmet needs, (the 'iceberg' of sickness) now presenting for health care and hence a gradual reduction in total health-care needs.

Although the basic approach of removing or reducing price barriers to health-care consumption was adopted by many other western industrialised countries, the policy was not based on an explicit behavioural model but on implicit assumptions that

- a) the price of health care was the only (or at least primary) barrier to access to health care, and
- b) access to health care was the only (or at least primary) barrier to access to health.
- c) improvements in health status would generate considerable improvements in well being.

By removing the price barrier, less affluent members of society would have equal opportunities to choose consumption bundles involving similar levels of health-care consumption as more affluent groups and would in fact choose these bundles. Consequently they would achieve similar levels of health status as members of these more affluent groups.

These implicit assumptions have been challenged and shown to be invalid in numerous studies. In the case of the first assumption Le Grand (1978) found that although poorer groups in society had greater rates of utilisation of NHS Services in the UK than more affluent groups, the differential rate of utilisation did not correspond to the differential rate of self-reported morbidity. Similarly Bucquet and Curtis (1986) found that the greater incidence of self-reported morbidity for particular conditions in manual workers in three London boroughs was not reflected in higher rates of consultation for these conditions. In other words, poorer groups used health-care services less per unit illness (or need) even though the services were free at the point of delivery. One explanation of these observations is that the

cost of using the service to the individual (i.e. the opportunity cost or shadow price) exceeds the price at point of delivery. Removing the price barrier of health care therefore deals with only one element of the total barrier to health-care utilisation.

The second assumption was challenged by McKeown (1976) who showed that the major improvements in health occurring over the last two hundred years had less to do with improvements in access to health care and were more closely associated with improvements in dietary practices and sanitary reforms. In earlier work Auster *et al* (1969) used multiple regression analysis on cross-sectional data to show that education and smoking behaviour were more powerful in explaining inter-state variations in US mortality rates than the availability of health-care services. The correlates of health, which were interpreted by the authors as determinants of health, appeared to be multiple, health care being just one such determinant, and by no means the most important one.

The validity of the third assumption is challenged by the observed differential response to information concerning the adverse health effects of smoking. Although the results of studies suggest that individuals in all groups now accept that reducing smoking reduces the risk of illness and death, less affluent groups show much less response to the information. (Townsend and Davidson, 1982, La Vecchia *et al*. 1986, Pierce 1989).

Recognising the limitations of the 'single cause - single effect' framework underlying the NHS approach to reducing health inequalities the authors of the Black Report recommended a comprehensive list of policies aimed at improving the incomes and knowledge of poorer groups in society in order that they might

- a) have a greater understanding of the broader determinants of health and therefore make more informed (and by implication, efficient) decisions concerning consumption bundles,
- b) have sufficient resources to adopt the consumption patterns required to attain the health levels of more affluent groups.

Although these recommendations constitute a much broader approach to understanding factors which influence population health status and the distribution of health throughout society, the report fell into the same 'policy trap' as the founding fathers of the NHS over 30 years previously. Policy recommendations were based on assumptions concerning the response of individuals to the proposed policies, without reference to any underlying behavioural model. In other words, it continued to use a 'single cause - single effect' approach, albeit in a wider framework than previously. Hence it assumed that giving the poor more money and exposing them to information on health risks would lead them to behave (i.e. choose consumption bundles) in similar ways to more affluent groups. If these poorer groups then choose to spend the extra income in different ways (e.g. more cigarettes and hamburgers instead of jogging shoes and fresh vegetables) the 'single cause - single effect' explanation falls down. The model is then unable to challenge the view that, in the presence of similar knowledge about health risks and sufficient resources to achieve a healthy lifestyle (i.e. similar opportunity sets), poorer groups simply choose to follow less healthy lifestyles, and by implication lower levels of health and earlier deaths. If this is the case, inequalities in health become much less important to policy makers concerned with equating opportunities, as opposed to outcomes (i.e. actual health improvements), and focusing attention on health as opposed to well being, i.e. the impact of health on individual utility.

Although the lack of any behavioural model represents a fundamental limitation of the Black Report, it is worth noting that the proposals emerging from the report are remarkably consistent with the comparative static predictions of the existing approaches to the demand for health in the economics literature.

Most of the economic literature on the demand for health is based on the model and related work of Grossman (1972) and uses the household production theory (Becker 1965). Under this approach individuals use marketed and non-marketed resources to produce commodities in the household in combinations which maximise utility. Health capital is one such commodity which is produced from inputs of own time and health-care inputs and which generates (utility-bearing) flows of health.

The first order conditions for utility maximisation derived from the model imply that individuals continue to invest in health capital until the benefit of the flow from additional health capital equals the opportunity cost of its production (i.e. the familiar condition of marginal benefit equals marginal cost). Grossman's main interests did not lie in modelling health-related behaviour as defined in this paper (activities having impacts on health but undertaken primarily for non-health reasons) but in understanding and estimating the relationships between health status, health-care utilisation (which Grossman interpreted as health-care demand), wage rates, education and age.

Issues of joint production were introduced briefly in the final section of the work. In particular some inputs in household production functions were modelled as entering the health capital production function, where they have a negative marginal productivity, as well as entering the production function for which they are primarily used (and hence have a positive marginal productivity). For example cigarettes enter the production

function for relaxation and the health capital production function. The effect on the first order conditions is to increase the marginal opportunity cost of relaxation by an amount equal to the impact of the marginal production of relaxation on health capital.

The purpose of this section of Grossman's analysis was not to identify and explain correlations between socio-economic variables and health-related behaviours, but to show more generally how a negative income elasticity of health does not imply that health is an inferior good - merely that the income elasticity of commodities with health-detrimental inputs (such as relaxation by smoking) exceeds the income elasticity of health capital production. The corollary of this, however, is that if individuals or groups differ in their patterns of health-related behaviour, it is the result of differences in income elasticities. No consideration is given to why income elasticities differ but the implication, as with the implicit framework of the Black Report, is that groups choose different consumption bundles per se and hence the different income elasticities are a reflection of different utility functions.

Although the original Grossman work has generated considerable further research (see Birch 1987), most of this has focused on refining or improving the specification of the model or the estimation of the derived equations. Little attention has been paid to the determinants of health-related behaviour or alternative approaches to modelling such behaviour. Muurinen (1982) introduced 'environmental' variables as influences on the rate of depreciation of health capital stock. In other words, such influences 'used up' scarce health capital (use-related depreciation).

This had the advantage of allowing for adverse impacts on health occurring even in periods when health capital production was zero (i.e. health-care consumption zero in terms of the Grossman model). Muurinen later expanded the environmental variables category to include 'hazardous consumption and production activities'. While this 'use-related' depreciation approach to health-related behaviour provides a more general method of including inter-relationships, it was not a primary focus of the analysis and was not endogenously determined.

A feature of the demand for health research which deals more directly with the issue of health-related behaviour is the specification of the education-health relationship. The relationship between health and education has been the subject of a large number of empirical studies (e.g. Grossman 1972, Grossman 1975, Edwards and Grossman 1979, Shahotko, Edwards and Grossman 1981). However our interest lies in the alternative conceptual foundations of the relationships. While Grossman's initial work modelled education as enhancing the technical efficiency of household production (i.e. producing greater outputs from a set of given resource inputs) (Grossman 1972), an alternative approach used in later studies is that education enhances the allocative efficiency in terms of the choice between production processes.¹ In other words, more educated individuals are more aware of the adverse health effects of certain activities and hence are in a better position to make informed decisions between health-detrimental and health-neutral methods of production of a given commodity. In this case more educated individuals have healthier lifestyles ceteris paribus because of their greater knowledge.

However the important point here is that the prediction is based on ceteris paribus conditions. To move from here directly to a list of policy proposals of the type suggested in the Black Report assumes that providing those individuals who

following unhealthy lifestyles with health education will lead to their choosing similar consumption bundles (i.e. lifestyles) as the more educated - and hence back to the 'single cause - single effect' approach. Again no consideration is given to why the less educated are observed to choose less healthy lifestyles and their behaviour is not modelled - it is merely assumed that they are subject to the same incentives and constraints as their more educated counterparts.

An alternative approach to explaining the health - education relationship, and by implication, to explaining differentials in health, is that education and health both represent investment activities involving current opportunity costs in return for future returns. Farrell and Fuchs (1982) observed that differentials in the prevalence of smoking between young adults of differing education levels existed before the additional education was received. In other words the additional education could not explain the differential smoking behaviour but was more likely an additional reflection of an underlying third variable which affected the willingness and/or ability to invest in health and education in the same way. They suggested that time preference was one possible underlying variable, the value of which might differ between individuals and thus cause the observed differential behavioural patterns.

In separate work Fuchs attempted to measure time preference using individuals', declared preferences between money alternatives occurring now and at points in the future (Fuchs 1982). Although the determinants of time preference were not explored, alternative explanations of genetic transmission of preferences and conditioning early in life were suggested. In the work with Farrell (Farrell and Fuchs 1982) it was suggested that differences in rates of time preference inferred from observed behaviour might be the result of imperfections in the capital

market giving rise to unequal access to intertemporal transfers (i.e. differences in time preference at the margin.) Although this argument implies that observed lifestyle differentials might be the result of differential opportunity costs, the explanation was not based on the prediction of any behavioural model but was generated by post hoc rationalisation of observed behaviour.

3. A model of health-related behaviour

In this section a model of health-related behaviour is developed. In contrast to the existing literature on the demand for health, we focus our attention on deriving conditions for the optimal production of commodities which, although having impacts on health status, are primarily consumed for other (utility-bearing) reasons. We conceptualise health status as a quality-adjusted (or equivalent-efficient) time variable which is utility bearing in addition to being a scarce resource for use in other commodity production functions. In other words, illness does not represent time lost (i.e. of no use in other activities) but a reduction in the marginal productivity of time as an input to work, rest and play.

Consider the individual's intertemporal utility function

$$u = u(x, y, h) \quad (1)$$

where $x = \{x_j\}$

$y = \{y_j\}$

$h = \{h_j\}$

and $j = t, \dots, t + n.$

Vectors x and y represent the individual's planned household production of commodities x and y in future periods and h is the vector of health in future periods. Commodity x differs from y in so far as x has no direct impact on quality-adjusted time

production, h , but y impacts directly, either positively or negatively on h . This impact may be immediate (e.g. driving while under the influence of alcohol) and/or delayed (e.g. excessive alcohol consumption). The individual is assumed to have some expectation of the time of death based on life expectancy estimates. Uncertainty does not feature in this simple model at present, although the assumption could be relaxed at a later stage (Gafni and Peled 1984). The production of each of the commodities is represented by the following production functions,

$$x_j = x(X_j, T_j^x) \quad (2)$$

$$y_j = y(Y_j, T_j^y) \quad (3)$$

$$h_j = h(M_{j-1}, T_{j-1}^h, h_{j-1}, \dots, h_0, Y_{j-1}, \dots, Y_0, N) \quad (4)$$

Where X , Y and M are market inputs, T^x , T^y and T^h are inputs of efficient time and N is an exogenous endowment variable. Because x and y are commodities in the form of smoking, jogging or watching T.V. they are 'consumed' as they are produced within the household.

Following Grossman (1972) the individual has a single period full resource constraint, R_j , given by²

$$R_j = h_j w_j + v_j = (h_j - T_j^v) w_j + p_j^x X_j + p_j^y Y_j + p_j^m M_j \quad (5)$$

where w is the marginal opportunity cost of (quality-adjusted) time spent in household production, v is a wealth stock variable, T^v is the quality adjusted time spent in non-household production and p^x , p^y and p^m are unit prices of market inputs x , y and m respectively. Generalising the resource constraint over time gives

$$R^* = \sum_{j=t}^{t+n} \frac{(h_j - T_j^v) w_j + p_j^x X_j + p_j^y Y_j + p_j^m M_j}{(1+r)^{j-t}} \quad (6)$$

Maximising the intertemporal utility function subject to the

Maximising the intertemporal utility function subject to the generalised full resource constraint generates first order conditions as follows,

$$U_t^x = \varphi \left(P_j^x \frac{dX_j}{dx_t} + w_j \frac{dT_t^x}{dx_t} \right), \quad (7)$$

$$U_t^y + \sum_{j=t}^{t+n} \frac{U_j^h \frac{dh_j}{dy_t}}{(1+r)^{j-t}} = \varphi \left[\left(p_t^y \frac{dY_t}{dy_t} + w_t \frac{dT_t^y}{dy_t} \right) - \sum_{j=t}^{t+n} \frac{w_j \frac{dh_j}{dy_t}}{(1+r)^{j-t}} \right] \quad (8)$$

$$\sum_{j=t}^{t+n} \frac{U_j^h \frac{dh_j}{dh_t}}{(1+r)^{j-t}} = \varphi \left[\left(p_t^m \frac{dM_t}{dh_t} + w_t \frac{dT_t^h}{dh_t} \right) - \sum_{j=t}^{t+n} \frac{w_j \frac{dh_j}{dh_t}}{(1+r)^{j-t}} \right] \quad (9)$$

Equation (7) is the familiar first order condition that consumption of x continues until marginal benefit equals marginal cost. Although the first order conditions in equations (8) and (9) are less straightforward in algebraic terms, the interpretation of the conditions is exactly the same. Equation (9) is the first order condition for the production of health previously derived by

order condition for the production of health previously derived by Grossman (1972) and reproduced by Muurinen (1982). The interpretation offered by these authors was that health production continues until the sum of the marginal direct (i.e. utility-generating) and indirect (i.e. resource-generating) benefits equals the marginal opportunity cost of health production.

Although the interpretation of the future resource consequences of improved health status as a benefit, as used by both Grossman (1972) and Muurinen (1982), is inconsistent with welfare economics theory (Birch 1987, Birch and Donaldson 1987), this is not crucial to the present analysis. Of more interest to us here is the first order condition for the production of the health-related commodity, y , given in equation (8) which has not previously been derived in the economics literature as far as we are aware. Although the algebra again appears complex, the interpretation is more simple: under utility maximising assumptions, household production of y continues until marginal benefit (on the left hand side of the expression) equals marginal cost (on the right hand side).

The benefit of y production is given by the net effect of the marginal utility of y consumption and the utility consequences of this consumption of y on health status h . It is worth emphasising that both elements of the benefits expression are subjectively determined. The marginal cost of y production, on the other hand, is objectively determined (although the value of the marginal cost depends on the subjectively determined marginal valuation of wealth, φ) and given by the net effect of the opportunity cost of factor inputs and the impact of this consumption of y on future resource constraints (e.g. income-earning capacity).

There are a number of features of the model derived above that are useful in understanding and predicting health-related behaviour. In particular health-related behaviour is endogenously determined and hence is consistent with the rationality assumptions underlying the model. Utility, and in a broader context social welfare, represent more than just health status and consequently there is an optimal amount of health-related commodity consumption (and more particularly health-detrimental activities) which a utility maximising individual engages in. That is to say, unhealthy lifestyles may be in the best interest of the individual given the distribution of resources in society. But that is not the same as some population groups choosing less healthy lifestyles in the face of equal opportunities for health - a point that we shall return to below.

A second feature of the model is that the determinants of health-related behaviour extend beyond the direct utility consequences and opportunity costs of that behaviour. Identified health-enhancing behaviours may involve competing health risks which are not equally distributed throughout society. Furthermore these competing risks may generate additional consequences for an individual's resource constraints. For example accepting advice to stop smoking may lead to increased food or alcohol consumption, both of which have consequences for health risks and resources. Hunt and MacLeod (1987) found that the chances of contracting lung disease at some hypothetical future date were far less problematic to current women smokers wanting to stop smoking than the possibility of gaining weight immediately. Hence changing health-related behaviour often involves trade offs between health risks both in absolute terms and across time periods (because the timing of the alternative health risks may differ). Moreover these trade-offs may differ individually or in combination across population sub-groups implying that differentials in observed behaviour do not necessarily represent differential preferences in the presence of

equal opportunities. Existing models of the demand for health have the capacity to explore both of these features. Indeed the approach we have taken is to adopt the basic model of these previous analyses and extend it to the analysis of health-related behaviour. In doing so we argue that it helps to explain and predict observed behavioural patterns.

A third feature of the model is the dynamic nature of the determinants of health-related behaviour. In particular the indirect effects of health-related behaviour, represented by the impact of the behaviour on health status and the associated consequences for the opportunity cost of time, occur in future periods. Although the existing models of demand for health incorporate the dynamic nature of the health production process, the data sets used in testing the model's predictions have tended to be cross-sectional and have not considered the impact of expected future values of the consequences of health production, other than by using present values as proxy variables. As a result, implications arising from the dynamic aspects of the model have not been fully explored. Our analysis of health-related behaviour using the current model considers explicitly the individual's utility of expected future health consequences and the expected impact of such consequences on future resources resulting from the marginal production of health-related commodities.

4. Implications of the model

The model developed in the previous section emphasizes the multifactorial nature of the determination of health-related behaviour. In particular the determinants of health-related behaviour can be grouped under four main categories; the marginal valuation of the health-related commodity; the marginal valuation of health consequences arising from the commodity's consumption; the expected marginal opportunity cost of future (quality-adjusted)

time, and the real resources of the individual. We shall take as given and equal across all individuals the marginal valuation of the health-related commodity at each level of consumption (i.e. equal preferences for the commodity per se). Under the current model, observed differentials in behaviour are consistent with this assumption. The observed differentials may be explained by differences in the values of some of the other determinants of health-related behaviour as identified in the following comparative static predictions.

Increases in the real resources of individuals will generate an increase in the consumption of all non-inferior commodities ceteris paribus as the marginal utility of wealth, φ , falls. The income support proposals of the Black Report would, under this line of argument, enable the less affluent members of society to pursue similar patterns of consumption to the more affluent groups. Whether or not this actually occurs depends on the relative levels of the other determinants of health related-behaviour, as is highlighted by the model. In other words, the ceteris might not be paribus and while income support may enable the poorer groups to pursue similar consumption patterns (i.e. lifestyles) it may not be in the individual's best interests (i.e. individual utility may not be maximised) to do so. Assuming for the moment that underlying preferences are the same then only if the lifestyle incentives facing the individuals are also the same across population groups should we expect to observe similar behaviour given similar resources.

The same limitation applies to policies aimed at changing relative and/or absolute prices of health-related commodity inputs. Such policies essentially take a societal viewpoint and apply it to individual behaviour without considering lifestyle decisions from the viewpoint of the individual, i.e. in the context of the incentives and constraints facing the individual. We do not

suggest that such policies have no merit, but that on the basis of an explicit behavioural model, they are not the panacea for observed differentials in health-related behaviour. We should not be surprised, therefore, if poorer groups use such income supplements to increase their consumption of health-detrimental commodities.

Another policy instrument which features prominently in both the policy proposals of the Black Report and associated reports, and the demand for health literature, is the provision of additional education. It is implied that if only the less healthy groups of society had better knowledge of inter-relationships between behaviour and health then, provided their resources were sufficient, they would make more efficient (i.e. healthier) lifestyle choices. In terms of the present model an increase in education would improve the accuracy of individual's perceptions of the impact of current behaviour on future health and consequently lead to changes in lifestyle corresponding to behavioural patterns followed by the more educated groups ceteris paribus.

Improved knowledge of lifestyle-health relationships increases dh/dy and hence increases the marginal benefit and reduces the net marginal opportunity cost of health-enhancing behavioural patterns (and reduces marginal benefits and increases marginal opportunity costs of health-detrimental patterns). As with the income support proposals, the success of this policy depends crucially on the ceteris paribus assumptions holding. Again the proposed policy is essentially enabling less educated groups to behave in different ways without considering whether it is in the individual's best interests to respond to the improved information. For example information concerning the adverse health effects of smoking has been widely disseminated. Yet, as was mentioned above, the published data on smoking prevalence shows consistently that

reduction in smoking prevalence associated with this information provision has been highly concentrated in the more affluent and more educated groups in society (Townsend and Davidson 1982, Pierce 1989).

It does not seem plausible to explain this observation in terms of large sections of the population not understanding or not believing the information. A far more plausible explanation would seem to be that while reducing smoking is in the best interests of the more affluent groups in society, the same may not be the case for members of the less affluent groups under current circumstances. Hunt and MacLeod (1987) provide support for this type of explanation in the context of alcohol abuse in Scotland. Many problem drinkers tended to give up drinking of their own accord once the drinking begins to pose a problem for them as opposed to for other people.

Consideration of the other determinants of health-related behaviour as identified in the model helps us to understand why this might be the case. In particular the model emphasises the investment nature of the lifestyle-health relationships. Consequently individual decisions concerning lifestyle choices depend not only on the relative costs of inputs in alternative commodity production functions, but also on the perceived value of the future returns to that behaviour.

Farrell and Fuchs (1982) suggested that differentials in the rates of time preference are connected with observed differentials in the levels of investment behaviour undertaken (see section 2 above). Such differentials in time preference might reflect differentials in preferences for present and future consumption per se between subgroups of the population, in which case they would be of little interest to policy makers concerned with promoting equal opportunities for health. Alternatively they might reflect

differentials in access to intertemporal transfers (e.g. imperfections in the capital market) in which case policy instruments would be required to promote equality of opportunity. In terms of the present model a reduction in r , the marginal rate of time preference, would increase the impact of future health consequences of lifestyle and the associated resource consequences on lifestyle choice. So for example in the first order condition for health-related behaviour (equation 8), with $dh/dy > 0$, a reduction in r increases the marginal benefit and reduces the marginal opportunity cost of the (health-enhancing) behaviour while the opposite effects occur if $dh/dy < 0$ (i.e. health-detrimental behaviour).

Differences in the rate of time preference at the margin generate differences in the present values of future effects of equal future value. Marginal benefits of health-related behaviour would also be affected if the future value of the returns to the behaviour differed across individuals. In other words incentives to adopt healthy lifestyles may differ. In terms of the model derived above, an increase in U^h , ceteris paribus, increases the marginal benefit of health-enhancing behaviour and reduces the marginal benefit of health-detrimental behaviour leading to a substitution of healthier lifestyles. In a simple model of individual utility maximisation this effect reflects differentials in preferences for health between individuals. An alternative explanation suggested here is that the individual utility function is not separable between health and other commodities but the marginal utility of health at any level of health depends on levels of consumption of other commodities which are not equally distributed in society, i.e. $du/dh_j = f(k_j)$, $f_k > 0$, where k is an input to the utility function complementary to h . Increasing the consumption of such commodities (e.g. social support) would therefore increase the marginal benefits of healthy lifestyles and

generate a substitution of healthy for less healthy behavioural patterns ceteris paribus.

Although the notion of utility functions non-separable in health has not featured prominently in the health economics literature to date, there is a considerable and growing body of research in sociology which emphasises the relationships between health and health-related behaviour on the one hand and social relationships, support mechanisms and coping capacities on the other (see for example Antonovsky 1979, Berkman 1984, 1986, Berkman and Breslow 1983, Berkman and Syme 1979, Cassel 1976, Blazer 1982, House et al. 1988, Orth-Gomer and Johnson 1987) which would be consistent with this approach. As Hunt and McLeod (1987) emphasise "Health related behaviours are essentially social acts embedded in the structure which defines and delineates reality. Such behaviours are created and maintained by social relations, including those which grow out of the economic and political ethos, and are made possible by norms, values, rules and common history."

In addition it is consistent with the concepts of surplus powerlessness (Lerner 1986) and learned helplessness (Seligman 1975) which appear in the sociological and psychological literature respectively, in which the loss of control, or ability to cope, in one or more areas of life generates a tendency for individuals to give up control in other areas. Hence if support mechanisms break down, individuals may lose the incentive to follow health-enhancing patterns of behaviour.

This notion of multiple determinants of the marginal utility of health is therefore consistent with both an extended economic model and contemporary models of individual behaviour in other behavioural sciences. Furthermore it has face validity in that we might expect individuals to be influenced by their expected future

social circumstances in determining their consumption portfolios between present commodities and future health.

Finally, consider the impact of an increase in the expected marginal opportunity cost of future (quality-adjusted) time (w in the model). As w increases the net marginal cost of the health-enhancing (health-detrimental) commodity falls (rises) and hence more healthy patterns of behaviour will be adopted ceteris paribus. In other words the current (or potential) lawyer has a greater incentive to avoid future illness from a pure earnings point of view (i.e. the return to healthy behaviour is greater and hence more healthy behaviour is substituted for less healthy behaviour at the margin), as well as from the point of view of any greater expected social support, than the current (or potential) long term unemployed individual. It is important to note that this effect does not relate to current incomes, and hence cannot be addressed by income support policies, but is concerned with the expected value of future (quality-adjusted) time. Providing more income, either now or in the future to those with lower marginal valuations of future time does not affect the incentive to 'produce' (quality-adjusted) time.

Furthermore this effect is separate from increasing the marginal utility of future health by changing the levels of the determinants of marginal utility, although in some cases the two effects may be co-incident (e.g. individuals with greater probabilities of future support mechanisms may also have greater expected valuations of future time). Indeed the separate nature of each of the determinants of health-related behaviour is an important feature of the model because it provides an explanation of the heterogeneity of risky behaviour and the clustering of health risks both across and within subgroups of the population.

5. Summary

Oversimplification of human behaviour leads us towards descriptions of reality which are considerably at variance with human experience and cast doubt upon the validity of our information and thus upon its applicability

Hunt and MacLeod (1987)

We set out in this paper to consider whether economic analysis could be applied to the determination of individual health-related behaviour and so help avoid oversimplification of explanations of individual behaviour. Our work provides a number of important messages both for economists and researchers of other disciplines interested in understanding differentials in health-related behaviour and the implications for inequalities in health.

The first message is that economic analysis can be applied to the determinants of individual behaviour. In order to provide a comprehensive understanding of observed behaviour, however, we need to use a broad economic framework and go beyond the traditional boundaries of the contents of economic analysis drawing upon concepts, ideas and subject matter of other behavioural sciences.

We do not suggest that the rationality assumption underlying the traditional economic approach is applicable to individual decision making. It is sufficient to assume that individuals behave as if they were rational economic decision makers. Neither do we suggest that economic analysis in general, or the current model in particular can provide all the answers. Other approaches, which do not use any rational economic modelling, might perform as well, or even better than the proposed model in explaining and predicting individual behaviour. Indeed uncertainty, risk taking and addiction could be useful extensions to the model.

Nevertheless it represents an extension to the current application of economics to lifestyle decision making and provides some insights in understanding the determinants of individual behaviour.

The second message is that health-detrimental behaviours are not 'bad' per se, nor is the consumption of such commodities an indication of irrational behaviour. The extended economic model derived here analyses behaviour from the individual's point of view explaining behavioural patterns in terms of what is in the individual's own best interests. It is therefore important that policy makers concerned with reducing inequalities in health take this as a starting point and consider the changes required for the individual to alter behaviour in a way consistent with a social objective of reducing inequalities. Indeed the implication behind the analysis is that policies aimed at changing population health status should first consider what the implications are for individual utility. Determinants of health-related behaviour are multiple. Encouraging or forcing individuals to reduce unhealthy behaviour may reduce the well being of those individuals, even in the long term, unless the incentive structures and resource distributions in society are aligned in ways which make such behavioural changes in the individual's own best interests. This point is further emphasised by Hunt and Martin (1988) who note that health-related behavioural change, while having impacts on health, often stems from non-health considerations. So, for example, changes in smoking behaviour are often the result of individuals changing jobs or getting married as opposed to considerations of the health implications of the change.

A third message is that the determinants of the marginal valuation of health may be multiple and extend beyond just pure preferences and the level of health. As such, the relative importance of health to well being may be smaller than generally perceived but, more importantly, may depend on the availability of

other, complex social forces which are not distributed equally in society.

A final message is that taking a partial view of health-related behaviour, such as that underlying current policy proposals, risks leading to policies which not only fail to achieve their stated objectives but are counterproductive. For example income supplements may lead to the poor consuming more of all goods including cigarettes while generally available health education may generate a greater response from the more affluent and healthy groups. Under these circumstances inequalities increase.

The predictions of the model are, as yet, untested. Nevertheless they are consistent with many observations of differentials in health and health-related behaviours reported in the literature. Furthermore they serve to warn us not to use simplistic answers to complex questions which can lead to premature conclusions being drawn such as 'victim-blaming', or preference-based explanations of observations.

In previous work, one of the authors emphasised that in the context of supplier-induced utilisation, we cannot blame providers for responding to incentives presented to them in ways which contravene neither ethics nor regulations (Birch 1988). Similarly we should avoid blaming individuals for behaving in ways which represent rational responses to current incentives as generated by current social forces. The causes of the socially undesirable outcomes, whether it be overutilisation of services or the excessive consumption of tobacco, are the incentives inherent in the system, not the individuals who respond to these incentives.

Footnotes:

1. The two approaches need not be mutually exclusive as shown by Welch (1970).
2. Derived by substituting the quality adjusted time constraint $h_j = T^x_j + T^y_j + T^n_j + T^w_j$

and the income constraint

$$I_j = T^w_j w_j + v_j = p^x_j X_j + p^y_j Y_j + p^n_j M_j$$

into the expression for full wealth, i.e.

$$R_j = h_j w_j + v_j = (h_j - T^w_j) w_j + p^x_j X_j + p^y_j Y_j + p^n_j M_j$$

Acknowledgements

We wish to thank Bob Evans at the University of British Columbia and our colleagues in the Polinomix group at McMaster University for constructive comments on drafts of this paper, and Lori Houghton for translating our handwritten manuscripts into final form.

References

- Antonovsky A. (1979). Health, Stress and Coping, Jossey-Bass, London.
- Auster R, Leveson I, Saracheck D. (1969). The production of health: An exploratory study. Journal of Human Resources, 4, 411-36.
- Becker G. (1965). A theory of the allocation of time. Economic Journal, 75, 493-517.
- Berkman L. (1984) Assessing the physical health effects of social networks and social support. Annual Review of Public Health, 413-32.
- Berkman L. (1986) Social networks, support and health: taking the next step forward. American Journal of Epidemiology, 123, 559-62.
- Berkman L, Breslow L. (1983). Health and ways of living. The Alameda County Study. Oxford University Press, Oxford.
- Berkman L, Syme S. (1979). Social networks, host resistance and mortality: a nine year follow study of Alameda county residents. American Journal of Epidemiology. 109. 186-204.
- Birch S. (1987). The demand for health: Extensions to the human capital model. University of York, unpublished, D.Phil. thesis.
- Birch S. (1988). Item of service remuneration in General Practice in the UK: What can we learn from dentists? Family Practice, 5, 265-70.
- Birch S, Donaldson C. (1987). Applications of cost benefit analysis to health care: Departures from welfare economic theory. Journal of Health Economics, 6, 211-25.
- Blazer D. (1982). Social support and mortality in an elderly community population. American Journal of Epidemiology, 115, 686-94.
- Bucquet D, Curtis S. (1986). Socio-demographic variations in perceived illness and the use of primary care: the value of community survey data for primary care service planning. Social Science and Medicine, 23, 737-44.
- Cassel J. (1986). The contribution of the social environment to host resistance. American Journal of Epidemiology. 104, 107-23.

- Culyer A. (1981). Health, economics and health economics, in Van der Praag and M. Perlman (eds.). Health, economics and health economics. North Holland, Amsterdam, 3-11.
- Enthoven A. (1980). Health plan: The only practical solution to the soaring cost of medical care. Addison Wesley, Reading Mass.
- Evans R. (1984). Strained Mercy: The economics of Canadian health care. Butterworths, Toronto.
- Farrell P; Fuchs V. (1982). Schooling and health. The cigarette connection. Journal of Health Economics, 1, 217-30.
- Fuchs V. (1982). Time preference and health: An exploratory study, in V. Fuchs (ed.). Economic aspects of health. University of Chicago Press, Chicago, 93-120.
- Gafni A, Peled D. (1984). The effect of labelling on illness related absenteeism: An economic explanation for the case of hypertension. Journal of Health Economics, 3, 173-8.
- Grossman M. (1972). The demand for health: A theoretical and empirical investigation. National Bureau of Economic Research, New York.
- Grossman M. (1975). The correlation between health and schooling, in N. Terlerkyj (ed.) Household production and consumption. National Bureau of Economics Research, New York, 147-211.
- Health and Welfare Canada (1986). Achieving health for all: A framework for health promotion. Health and Welfare Canada, Ottawa.
- House J, Landis K, Umberson D. (1988). Social relationships and health. Science, 241, 540-5.
- Hunt S, MacLeod M. (1987). Health and behavioural change: some lay perspectives. Community Medicine, 9, 68-76.
- Hunt S, Martin C. (1988). Health-related behavioural change - A test of a new model. Psychology and Health, 2, 209-30.
- Lalonde M. (1974). A new perspective on the health of Canadians, A working document. Government of Canada, Ottawa.
- La Vecchia C, De Carli A, Pagano R. (1986). Educational prevalence of smoking in Italian men and women. International Journal of Epidemiology, 14, 279 (letter).
- Le Grand J. (1978). The distribution of public expenditure: The case of health care. Economica, 45, 125-42.

- Lerner M. (1986). Surplus powerlessness. Institute for Labour and Mental Health, Oakland, Ca.
- McKeown T. (1976). The modern rise of population. Arnold, London.
- Marmot M. (1986). Social inequalities in mortality: the social environment, in R. Wilkinson (eds.) Class and Health. Tavistock Publications, London.
- Marmot M, Morris J. (1984). The social environment, in W. Holland, R. Detels and G. Knox (eds.). Oxford Textbook of Public Health, volume 1. Oxford University Press, Oxford.
- Muurinen JM. (1982). Demand for health: A Grossman model. Journal of Health Economics, 1, 5-28.
- Orth-Gomer K, Johnson J. (1987). Social network interaction and mortality: a six year follow-up study of a random sample of the Swedish population. Journal of Chronic Disease, 40, 949-57.
- Pierce J. (1989). International comparison of trends in cigarette smoking prevalence. American Journal of Public Health, 79, 152-7.
- Royal Commission on the National Health Service (1979) Report (Chairman: Sir Alec Merrison). Her Majesty's Stationary Office, London.
- Seligman M. (1975). Helplessness: On depression, development and death. Freeman, San Fransisco.
- Shakotko R, Grossman M. (1981). An explanation of the dynamic relationship between health and cognitive development in adolescence, in J Van de Gaag and M. Perlman (eds.). Health, Economics and Health Economics. North Holland, Amsterdam.
- Townsend P, Davidson N. (1982), Inequalities in health. The Black Report. Penguin, Hammondsworth
- UK Department of Health and Social Security (1980). Inequalities in health: Report of a Research Working Group. Department of Health and Social Security, London.
- UK Ministry of Health (1944). A National Health Service. Cmnd 6502, Her Majesty's Stationary Office, London.
- Welch F. (1970). Education in production. Journal of Political Economy, 78, 35-59.
- Whitehead M. (1987). The health divide. Pengiun, Hammondsworth.