THE UNIVERSITY of York

An Economic Framework for Analysing the Social Determinants of Health and Health Inequalities

CHE Research Paper 52

An Economic Framework for Analysing the Social Determinants of Health and Health Inequalities

David Epstein [1,2]
Dolores Jiménez-Rubio [2]
Peter C Smith [1]
Marc Suhrcke[3]

[1]University of York, UK
[2]University of Granada, Spain
[3]University of East Anglia, UK

October 2009

Background

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The new CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

This paper is the outcome of work commissioned by the Department of Health and the Review of Health Inequalities in England. Thanks are due to our colleague Nigel Rice.

Disclaimer

Papers published in the CHE Research Paper (RP) series are intended as a contribution to current research. Work and ideas reported in RPs may not always represent the final position and as such may sometimes need to be treated as work in progress. The material and views expressed in RPs are solely those of the authors and should not be interpreted as representing the collective views of CHE research staff or their research funders.

Further copies

Copies of this paper are freely available to download from the CHE website www.york.ac.uk/inst/che/pubs. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third-parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk, telephone 01904 321458 for further details.

Centre for Health Economics
Alcuin College
University of York
York, UK
www.york.ac.uk/inst/che

Contents

Figures

Tables

Summary

Reducing health inequalities is an important part of health policy in most countries. This paper discusses from an economic perspective how government policy can influence health inequalities, particularly focusing on the outcome of performance targets in England, and the role of sectors of the economy outside the health service – the 'social determinants' of health - in delivering these targets.

Theoretical models

There has been some theoretical work in economics on the interaction between income, personal behaviour, and health. The core of these models is an assumption that individuals pursue a number of objectives, not all related to longevity and health. Within these models, health is valued for its own sake, and also promotes pursuit of other objectives: work, raising family, and participating in the community. Personal choices may therefore be made perfectly rationally to maximize these objectives, but may not necessarily maximize health. Furthermore, these models offer no unambiguous predictions about the relationship between the social gradient and health behaviour or health. However, if income has an increasing influence on health as income increases, for example due to positive lifestyle changes, then under reasonable assumptions it is likely that redistribution of income towards disadvantaged people might reduce Income Related Health Inequality (IRHI), but at the expense of average population health. Overall proportionate income growth would increase average health but increase IRHI.

Empirical estimates of the relation between income, human capital and health

Evidence for a (cross sectional) social gradient in health is strong. However, it is difficult to estimate the causal relationships between income, social factors and health, because of endogeneity, the influence of other factors and long time lags. Health problems and health behaviour tend to be strongly persistent. This makes changing trends in health and health inequalities at the macro level very challenging within the time frame envisaged by national targets.

Micro-level studies strongly suggest that causality of income and health runs both ways. Education and other 'permanent' changes to income have a stronger influence on health than temporary changes. It is difficult to generalise about the relation between health and income at a national level, though perhaps the data suggest a stronger relation from health to GDP growth than the other way around, and a greater size of effect in low income countries. On a macro level, an important question is how the worsening macroeconomic climate will affect health and health inequalities, and the steps that should be taken to mitigate the consequences of the recession on health. However, macro level studies are ambigous on the effect of lower economic activity on health: recessions may improve some indicators of mortality and morbidity (such as road accidents).

Overall income inequality, measured by Gini coefficient, increased substantially in the UK during the 1980s, mainly because of growing differences in earnings, and has not reduced subsequently. However, over the last 10 years inequality has increased most at the upper and lower extremes of the income distribution. For the bulk of the population, there has been income redistribution in favour of the less affluent.

Economists have developed a distinct research literature on income-related health inequality. Both income inequality and health-related inequality have increased since the 1990s. Van Ourti and colleagues have examined the differential impact of increased income on health at different points in the income distribution. They found that this 'income elasticity of health' increases with income in most European countries, offering some explanation for the increased income-related health inequality during the 1990s.

The rationale for government intervention

Micro-economic theory suggests personal decisions about health behaviour might lead to inadequate levels of prevention (from a societal perspective) if (among other reasons): there is inadequate information for citizens; there are externalities (eg passive smoking, alcohol misuse associated with crime, etc.); there are artificially low prices for unhealthy products (eg agricultural subsidies for high fat foods); there is clustering of health problems (eg peer group influence); or individuals are prone to

irrational behaviour or poor self control. From an economic perspective, policies that constrain personal autonomy should weigh these welfare losses against public health and other benefits.

Types of government interventions

Policies to influence individual behaviour include four categories of intervention (from Sassi and Hurst).

- *Increasing healthy options*, where the market fails to provide (eg improving school meals, improving public transport);
- *Influencing preferences*: this might include
    - providing information, such as improved food labelling, personalized health-related advice, and social marketing approaches;
    - incentives: some experiments have been effective, such as the Conditional Cash Transfer experiments in Latin America, that offer small but meaningful cash rewards for compliance with (eg) preventive initiatives or enrolment in school. There are nevertheless numerous design issues to be considered, such as which behaviour to target, which groups (if any) to target, the size of the reward, and how to police the scheme
    - using more recent insights from behavioural economics, there is increased interest in 'liberal paternalism', under which peoples preferences might be influenced by the manner in which options are presented to them.
- *Price controls, subsidies and consumption taxes* have a long history in public policy, for example in the form of 'sin taxes'. Recent studies (eg on mimimum pricing of alcohol) have suggested that these are generally effective in aggregate, though price rises generally have the highest impact on poor people, so the impact on inequalities is less clear cut. There may also be unintended side-effects, such as smuggling and cross-border consumption.
- *Restrictions and bans* and other forms of regulated behaviour can be effective (eg the public smoking ban) but may lead to unintended adverse outcomes (such as illegal avoidance measures).

Government policy on public services

The government can have a profound influence on the shape and performance of local public services through its national policies. These include target regimes, funding mechanisms, performance reporting, and staff contracts.

- *National targets*: since 1998 national government priorities in England have been expressed through the Public Service Agreement (PSA) target regime. This has been successful in some domains (such as waiting times), but less so in others (inequalities, cross-departmental targets). There is a well-developed literature on how to maximize the effectiveness of central targets. English policy is moving towards local priorities, through the Comprehensive Area Assessment initiative being implemented by the Audit Commission.
- *Funding mechanisms*: the English NHS has a well-established resource allocation mechanism that seeks to secure equity in health service access between geographical areas. It has been recently augmented by a major 'health inequalities' adjustment of £7.5 billion intended to direct resources to areas making the biggest contribution to premature mortality and disability. Its effectiveness in reducing health inequalities has yet to be established.
- *Local performance reporting*: there has been increased use of local performance reporting for public services, for example through the Healthcare Commission's annual health check and the Audit Commission's Comprehensive Performance Assessment. These have been effective in focusing managers' attention, although not notably in the inequalities domain. A key issue for the future will be the extent to which Comprehensive Area Assessment succeeds in securing cross-agency collaboration (including public, voluntary and private sector).
- *Staff contracts:* there has been increased interest in the extent to which incentives directed at the practitioner level might secure better outcomes than those directed at organizations. The GP Quality and Outcomes Framework is the most notable English example. It has undoubtedly secured improved focus of GP activity, although the small measured gains that can be attributed to the QOF do not yet seem to justify the large expenditure.

Integrating equity into priority setting

The methodology of priority setting in health care has reached an advanced stage of development, not least through the work of the National Institute for Health and Clinical Excellence (NICE). There are however challenges in integrating public health and social interventions into the traditional cost-effectiveness approach. Drummond and colleagues summarize these as:

- Attributing outcomes to interventions
- Measuring and valuing outcomes
- Incorporating equity considerations
- Identifying intersectoral costs and consequences

An implication of this analysis is that priority setting is drawn towards cost-benefit rather than cost-effectiveness analysis, a much more demanding methodology. Furthermore, analysis of equity requires modelling differential responses by subgroup, again multiplying complexity.

There has been some work by economists on how society values identical health gains for different population groups. There is evidence of strong preference for equity amongst some people, but preferences are highly variable. In principle, this research can be used to adjust cost-effectiveness ratios for equity concerns. However, studies so far have been relatively small scale and tentative in their conclusions.

Given the methodological challenges, policy makers (including the UK government) have developed a more pragmatic approach towards priority setting, in the form of descriptive Health Impact Assessments. These are likely to be especially helpful when examining cross-departmental initiatives.

1. Introduction

Health inequalities are a major concern of government policy in nearly all countries. Health is valued for its own sake. But health also enables participation in other aspects of daily life: work, raising family, and participating in the community. Reducing inequalities in health is considered both a matter of social justice, and a means of opening up other opportunities, particularly for the most disadvantaged.

The World Health Organisation Commission on the Social Determinants of Health has recently completed a two-year investigation into the social causes of health inequalities (CSDH 2008). The report concludes that health inequalities cannot be fully explained by poverty or variation in income alone. Nor can inequality in health be fully explained by the varying capacity of local health services. In addition to these factors, health inequalities are caused by inequitable distribution of more fundemental social, political and economic forces, the 'social determinants of health'.

A central precept of CSDH is that health depends on many factors and policies that are outside of the remit of health ministries. The CSDH make a large number of recommendations for government action at different levels: to improve basic living conditions, health services, education, and working conditions; to reduce inequalities in power and resources; and to create transparency by monitoring and measuring inequalities in health.

The CSDH builds on earlier work by the WHO Commission on Macroeconomics and Health (CMH 2001). However, the CSDH highlights the influence of social conditions on health and inequality, whereas the CMH stressed the value of a healthy population as a means towards national income growth.

This paper discusses from an economic perspective how government policy can influence health inequalities, particularly focusing on the outcome of performance targets in England, and the role of sectors of the economy outside the health service – the 'social determinants' of health - in delivering these targets. The relationship between health and social conditions have been analysed in a number of frameworks, representing different disciplines (Solar and Irwin 2007). We do not argue that an economic framework is more appropriate than other perspectives. Rather, we review how economic analysis has been used to explain the causes and consequences of health inequalities and inform policy making. We focus on:

- Theoretical models of the relationships between human capital, income and health
- Empirical evidence on the relationship between income and health
- The role of individual lifestyle and consumer behaviour for health
- The use of performance indicators and targets relating to health inequalities
- Evaluation and priority setting for policies aimed at reducing health inequalities

While the CSDH report is a tremendous achievement, it downplays some issues that economists would consider important. While the CSDH acknowledge the importance of identifying the direction of causality (Solar and Irwin 2007), many of the data presented in the report are correlations. We review studies that have attempted to identify these causal relationships and the dynamics involved. The CSDH does not analyse the role of lifestyle and choice in much detail, other than implying that behaviour is likely to be determined or restricted by social conditions. Economists have developed theoretical models and empirical work examining the choices people make, and whether these can be considered 'rational'. The Commission's recommendations are not prioritised or costed. This paper will discuss methods for evaluation and priority setting. CSDH stress the need for good governance, transparency and monitoring health inequalities. We review the economic literature on designing performance management targets that may support this recommendation.

2. Economic framework for analysing health inequalities

The causes or determinants of inequalities in health are of course many and varied. The WHO Commission on SDH emphasised that the major intermediate causes of morbidity and mortality in middle and high income countries – undernutrition, obesity, smoking, hazardous alcohol consumption, hypertension, sexual behaviour – are broadly speaking linked to one's socio-economic position (CSDH 2008).

In this section, we describe an economic perspective on the relationship between income inequalities and health inequalities. This conceptual model does not try to explain every aspect of the relationship between income and health – this would be infeasible. Instead it aims to show in general but mathematical terms the effect of income inequality and income growth on income-related health inequality and average health. Many other economic models have been developed to explain other aspects of these relationships, such as national income growth models and household behaviour models, and some of these are described in subsequent chapters.

2.1 A simple economic model of health inequalities

Contoyannis and Forster (1999) develop an economic model based on the relationship between income and health, where the effect on health of a given change in income (or percentage change in income) might not be the same for all social groups. This model shows in general the conditions or the assumptions under which policies aimed at improving health behaviour, proportionate income growth or redistributing income might affect population health and income-related inequalities in health. This model has also provided the conceptual framework underpinning some empirical studies looking at income-related inequalities in health (Van Ourti et al 2009).

2.1.1 The relationship between health and income

For each individual in the population, health Hj is influenced by income Ij.

Hj = h(Ij, Ej) j = members of the population A , B, ..

Figure 1 shows an example of the individual health production functions for persons A and B (H_A and H_B). For simplicity the relationship between individual health and income is linear in Figure 1, but need not be so. The conclusions that Contoyannis and Forster (1999) derive from the model are valid as long as health is increasing with income, even with diminishing individual returns.

The relationship between income and health is not the same for all individuals. Health is also influenced by other factors, Ej. These variables might be intermediate 'determinants' of health, for example, representing lifestyle. Person A (or type 'A') might respond to changes in income by healthy changes to lifestyle, whereas person B might take up some elements of a less healthy lifestyle given the same change in income. Lifestyle is of course not the only intermediate determinant of health. Other factors would include working conditions, housing, social networks/support, access to education and recreation, etc, that to a greater or lesser extent are influenced by income. Figure 1 does not include factors such as genetics that influence health but are unrelated to income. These factors could be included in the model, for example by shifting the 'intercept' for some individuals, but would not change the results.

The 'population' health production function Ph shows the 'average' health of the population at each level of income. Its shape depends on how E varies with income, eg people with higher incomes might be more or less likely to take up healthy lifestyles. Micro-economic theory makes ambiguous predictions about whether preventative health behaviour will increase or decrease with income (Kenkel 2000). A simplified version of this theory is described in Chapter 3. Likewise, the Contoyannis and Forster model does not make any *a-priori* assumption about whether Ej - intermediate determinants of health and/or health behaviour - are increasing or decreasing in income. Given linear individual health production functions, Contoyannis and Forster show the 'population' health production function Ph will be linear (Ph1) if the proportion of people with 'healthy type A' and 'unhealthy type B' lifestyles is the same across the income distribution. Ph will be concave (rate of increase of population health in falling with respect to income) if people are more likely to be 'unhealthy type B' higher up the income scale (Ph2), and convex (Ph3) if people are more likely to be 'healthy type A' higher up the income scale.

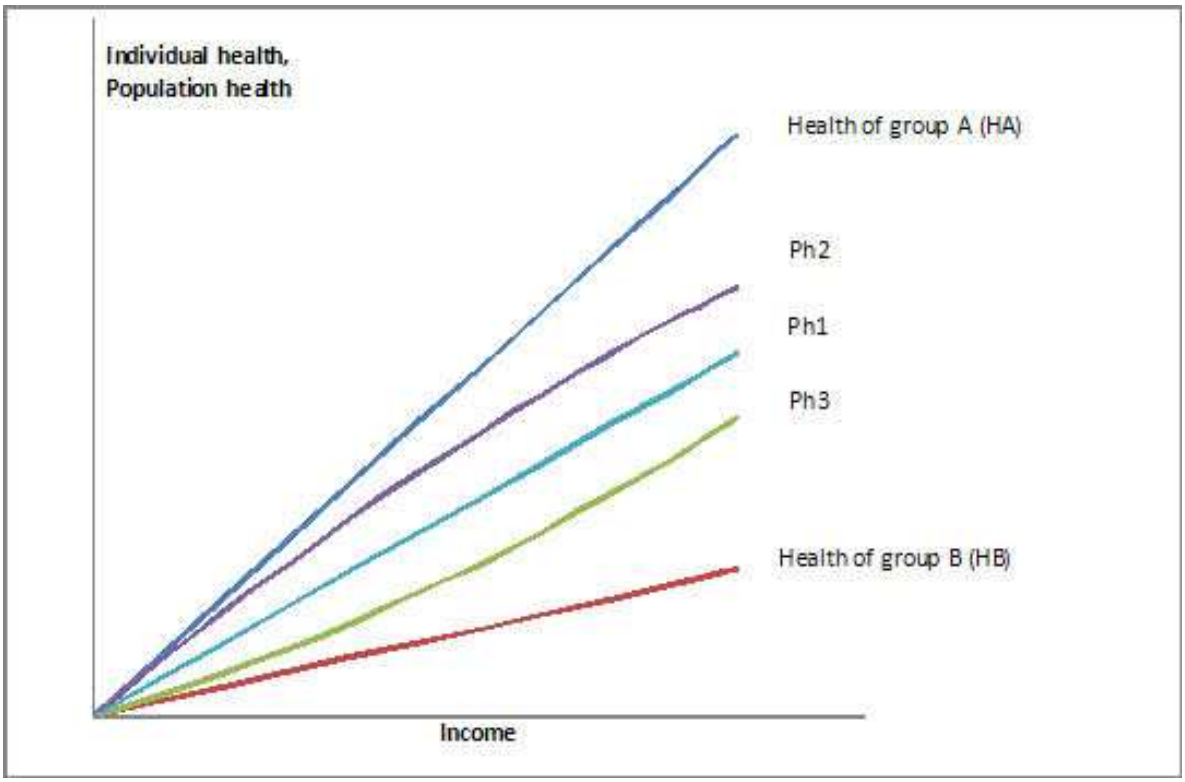

Figure 1. The basic framework (Contoyannis and Forster 1999).

Contoyannis and Forster (1999) show that given their assumption that health is increasing in income, then a redistribution of income

- will reduce income-related inequalities in health (IRHI)
- might reduce mean health if Ph is convex

Growth of income

- will increase mean health
- might increase IRHI if Ph is convex

2.1.2 Strengths and limitations of the simple model

The Contoyannis and Forster model is in many ways closely related to the framework used by the CSDH (Solar and Irwin 2007). It allows examination (in an abstract way) of the influence of income on health, and the interaction between income and other intermediate determinants of health, such as behaviour. It is concerned with explaining income-related inequalities in health as well as average health. It assumes the relationship flows from income to health, rather than the other way round.

Given that this is a general framework, it does not make any *a priori* claim as to whether the population health production function is actually convex or concave. Rather, it shows the conditions under which there may be a trade-off between average population health and IRHI. Public health measures, such as promoting healthy lifestyles, may increase mean health but worsen income-related inequalities in health if, for example, people higher up the income scale are more likely to respond. Similarly, other measures such as redistribution from rich to poor might reduce income-related health inequalities but might reduce overall health if higher-income groups forgo healthy behaviour while recipient groups do not take up healthy behaviour. We look at some of the empirical evidence in the UK concerning lifestyle (smoking, alcohol and obesity), and macro-level evidence on the relationship between income, lifestyle and health in Chapters 3 and 4.

The model has several limitations. First, it assumes that health generally increases with income, though possibly with diminishing returns. There may be limits to this positive relationship (DH 2009).

Work related and road traffic accidents in particular tend to increase with economic activity (Adda et al 2008). Second, the model assumes that the health of individuals is independent of the health and income of others. An important issue is the extent to which health is influenced by one's relative position in the social hierarchy rather than one's absolute resources. Furthermore, the effects of inequality of income might not be confined to the most disadvantaged but could affect the health of the majority of the population. These arguments are comprehensively reviewed by Wilkinson and Pickett (2009). Third, the model is descriptive. It suggests there might be a relationship between income, lifestyle and health but does not try to explain what factors might influence this relationship. Fourth, the model assumes the effects are instantaneous and act in one direction: income to health. Causality may run in the opposite direction, and there are lags between changes in income to behaviour, behaviour to health, and health to mortality. This means there may be a strong (cross sectional) gradient in health across society at any point in time and, at the same time, a weak relationship between changes in income and changes in health during the time frame over which policy makers set targets. Finally, although the model shows how there might be trade offs between improving IRHI and population health, it does not offer a framework to evaluate the overall costs and benefits of policies. Approaches to evidence-based priority setting are discussed in Chapter 6.

3. Empirical estimates of the relationship between health, human capital and income

This chapter examines the evidence on the relationship between health, human capital and income. Clearly, this is a huge and complex topic, and no single study could hope to address all the possible questions that might be asked, let alone provide definitive answers. We only present here a selection from a large literature. For other reviews see for example Suhrcke et al (2005) on the contribution of health to the economy focusing in particular on the European Union, and Suhrcke et al (2007) on health and economic development in Eastern Europe and central Asia.

In this chapter, we look at studies that have assessed the relation from income to health, and health to income. Many of the data presented by the CSDH did not identify the direction of causality and the authors tended to take for granted that the data showed social factors were causes of ill health and health inequality rather than the other way round. Furthermore, if social interventions are found to improve health, policy makers might be more likely to implement them if it can be shown that this would feed through to other national objectives such as improved labour productivity or economic growth.

At the risk of generalising, it is quite plausible that causality runs in both directions but as health tends to be persistent it might take a long time before the full effect is realised (Contoyannis et al 2004). Another question is the interaction between income, health and other human capital (eg education). Changes in income might be only transitory, whereas human capital represents a person's long term productive potential. A third question is whether individual level effects might be dampened or enhanced when scaled up to national level. For example, improvements in health might release household or state resources that could be used in other healthcare services.

To examine these issues, we review a selection of studies in four broad areas. First, we review individual or household level studies looking at the relationships between health and income and vice-versa. Second, we examine the relationship between population health and economic growth. Third, we review the relationship at national level between income inequality and income-related inequalities in health. Fourth, we look at the effect on health of economic downturns.

3.1 Using micro (individual and household) level data

There are several channels by which health might influence income, and vice-versa. The effect might be mediated by other factors or social determinants which might be correlated with income. Some of these channels might include:

- *Health affects productivity*. Healthy adults have a greater capacity for work, and healthy children are more likely to become healthy adults
- *Health affects decisions to enter or leave the labour market*, for example for early retirement.
- *There is a relationship between health and education*. Healthy children and adults have a greater capacity for education and training and a reverse effect as better education can increase earnings and can promote healthy behaviour.
- *Income may influence subsequent health*
- *Other social determinants*, such as the workplace environment influence health for better or worse.

In this section we review empirical evidence from household and individual level data concerning these relationships.

Effect of health on productivity and earnings

The first studies about the relationship between health and productivity focused on developing countries and the association between nutrition and productivity. More recent research has started to analyse the link between health and earnings in high-income countries as provided by the steady emergence of extensive panel data sets, such as household surveys.

There is an extensive literature that supports the existence of a direct relationship between low adult height (partly as a consequence of poor child nutrition) and reduced adult wages in developing countries (Alderman et al 2005). Thomas and Strauss (1997) analyses the impact of adult height on adult wages for urban Brazil and finds that a 1% increase in height is likely to lead to a 2–2.4% increase in earnings. The study accounts for causality between health measures and productivity or earnings by using relative food prices as instrumental variables for health.

Behrman and Hoddinott (2005) use the anthropometric-earnings estimates estimated by Thomas and Strauss (1997) to predict the long term impact of the nutritional component of the Mexican OPORTUNIDADES scheme, a cash transfer programme benefiting poor rural families. They find a 2.9% potential increase in adult earnings due to the increased adult height following the improvement in children's nutrition induced by the programme.

Rivera and Currais (2005) analyse the effect of health on earnings at different levels of the wage distribution using quantile regression for a sample of Brazilian workers. Housing conditions and health infrastructures are used as instrumental variables to identify health in the wage equation, and control for the reverse effect of income on health. The study shows that health (as measured by the Body Mass Index and two other self reported health indicators) has a significant positive impact on earnings, and that the effect is greater at lower levels of the income distribution.

Suhrcke (2005) identifies many studies showing poor health negatively affects wages and earnings in high income countries. Several studies find that physiological proxies for health affect earnings in high income, as well as developing countries. Height appears to increase earnings, while obesity appears to depress wages and earnings. However, Suhrcke concludes that this is more likely to be due to social meaning attributed to height and obesity particularly in adolescence than a direct effect on productivity.

Suhrcke (2005) finds a large number of studies from high income countries showing that health increases the probability of participating in the labour force or retiring. Men tend to reduce their labour supply in response to their wives illness, but in the reverse case women tend to increase their labour supply. These results are very sensitive to the institutional framework in the country, such as pension rules, availability of benefits and access to occupational and health insurance. García Gómez and López Nicolás (2006) find that workers in Spain who suffer a health shock are around 5% less likely to remain employed and 3.5% more likely to become inactive. Social insurance does not fully compensate the fall in labour income.

Influence of health on education

Malnourished children tend to delay school enrolment, have lower school attainment and have poorer performance on cognitive tests (Alderman et al 2005). This might be because their parents may invest less in their education, because schools may accept students on the basis of their physical size, or because poorly nourished children have higher rates of morbidity and are more likely to be absent from school.

The relationship between children health and education is a complex one as both child nutrition and schooling reflect unobserved household attitudes regarding investments in human capital. Many studies that have found associations between nutrition and schooling (see e.g. Behrman 1996). However, only a few established causality. For example, using a longitudinal dataset for rural Pakistan between 1986 and 1991, Alderman et al (2000) explore the influence of nutrition in childhood on later school enrollment decisions by controlling for the behavioural determinants of pre-school malnutrition. They find considerable positive effects of improved nutrition on school initiation and school attainment, particularly for girls.

In addition to malnutrition, contagious health diseases, such as worm infections, are another potential source of low school performance outcomes by children in developing countries. Miguel and Kremer (2004) studied the impact of a school-based mass treatment on student schooling participation based on a randomised evaluation of Kenyan schools. Their results indicate that the deworming treatment led to improved health and school attendance for treated students, and interestingly, also for untreated students. This is thought to be because school attendance is influenced by social norms.

Overall, they found a 25% reduction in absenteeism in treated schools and an increased schooling by 0.15 years per pupil treated.

Suhrcke (2005) finds that, while there is considerable evidence linking childhood health and education in developing countries, there is little empirical work on this link in high-income countries.

A number of studies have looked at health as a potential mechanism through which economic status flows intergenerationally. For instance, the study by Case et al. (2005) using a longitudinal data set of British individuals followed from birth to age 42 shows that children born into poorer families were more likely to have lower childhood health, worse schooling outcomes, and lower health in early adulthood. All those factors were found to be associated with lower earnings in adulthood.

Influence of education on health

Mackenbach (2006) found that Europeans with lower levels of education tend to die younger and to report lower levels of self assessed health, and the authors attribute this partly to exposure to risk factors such as excess alcohol consumption and inappropriate diet.

Cutler and Lleras Muney (2007) also argue that the protective effect of education on health after controlling for income, occupation, or ethnicity, might be mediated through behaviour. The better educated are more likely to consume healthy goods (e.g. preventative health care, use safety devices such as seat belts), and are less likely to consume unhealthy goods such as cigarettes or alcohol.

Influence of income on health

Parental income and wealth has a very strong association with subsequent adult health, both for families in developing countries (van de Poel 2008) and developed countries (Case et al 2005). However, showing a strong relationship between parental income and subsequent adult health does not entirely eliminate reverse correlation. For example, there may be circumstances where the poor health of a child reduces family income, from out-of-pocket healthcare payments or reduced labour hours of parents, and poor child health might be a cause of poor adult health.

Contoyanis et al (2004) used successive waves of the British Household Panel Survey (BHPS) to identify causality. The results suggest that permanent income (average household income across the 8 waves of the BHPS) is more important in influencing health than current income. More recently Jones and Wildman (2008) found a significant effect of current income on health after controlling for education though the magnitude of the effect was small. Case (2001) showed that the health of older people in South Africa improved after they started receiving pensions at the age of 65.

However, other studies have found a weaker causal relationship in developed countries between income and health after controlling for other social determinants. Using panel data for the US, Smith (2007) does not find a significant link between financial resources (whether income or wealth) and the onset of new health conditions in adults, after controlling for education; rather, education appears to be the primary socio-economic influence on health.

Influence of other social determinants on health

Case and Deaton (2003) found that US manual workers have lower self reported health than non manual workers and that their health declines more rapidly. However, unemployment emerges as the main cause of differences in health and the rate of health deterioration, particularly mental health (eg. García Gómez and López i Casanovas, 2005). García Gómez and López Nicolás (2006) found that workers who became unemployed were 2.9% less likely to report good health than a matched worker who remained employed.

Bambra et al (2009) conducted a ‘review of reviews’ of social interventions and found that the effects of employment change (such as privatisation) are experienced differently by employees in different occupational categories, and that the workplace may be an important setting in which health inequalities may be addressed. They found some evidence that housing improvements may positively affect physical health, but the effects may be quite small.

3.2 The relationship between population health and economic growth

This section examines the evidence that improving average population health might increase economic growth, in developing and developed countries. The Commission for Macroeconomics and Health argued that policy makers would be more likely to implement improvements to health and health services if it could be shown that this would feed through to improved labour productivity or economic growth.

Micro studies offer strong evidence that health influences individual productivity and earnings. However, better health may not be translated into long run growth at an aggregate level. This dampening might occur if there are diminishing returns to labour, for example, if land and/or physical capital are limited. Another mechanism might be that increasing health generates population growth or increases the dependency ratio, with more infants surviving to childhood and more elderly people surviving to retirement age, and depresses GDP per capita until the population returns to equilibrium (Ashraf et al 2008). Early childhood development may be the most important determinant of adult health and productivity. In this case, the full effect of improved population health on earnings and GDP per capita might take a generation to be realised.

We compare in detail the empirical evidence from four studies that have estimated growth rates using cross-country or panel data: Bhargava et al (2001), Bloom et al (2001), Acemoglu et al (2007) and Doppelhofer et al (2004).

Bhargava et al. (2001) estimated the determinants of growth at 5-year intervals from 1965-1990 using panel data on 92 countries. They found that in the poorest countries a 1% change in adult survival rate (ASR) was associated with a 0.05% increase in growth rate. The authors consider this positive correlation is because of the productivity gained by labour in prime years. However, beyond a threshold, increases in ASR are difficult to achieve and will increase the proportion of elderly people in the economy. They calculate that, in a model including lagged GDP as a fully endogenous variable, the impact of ASR on growth approached zero when the GDP per capita was 1,714 US$ at 1985 international dollars, and was negative for higher levels of GDP per capita. However, the results are sensitive to the choice of functional form and the data used, including the purchasing power parity weights. The authors stress that ASR is only a proxy for health and has a different significance in richer and poorer countries. Variation in ASR between richer countries is likely to reflect genetic factors and access and cost of preventative and curative health-care, while in poorer countries, it is likely to be influenced by a wider set of factors including level of nutrition, smoking prevalence, infectious diseases, health infrastructure and accident rates.

The model used by Bhargava estimates the overall correlation between ASR and GDP growth, but does not by itself indicate causality. Countries suffering from short life expectancy and ill health are also disadvantaged in other ways, and so many such macro studies may be capturing the effect of other omitted variables.

To investigate the direction of causality, Bhargava et al estimated the effect of lagged GDP growth rates on ASR. They found that lagged GDP growth rates do not influence the current ASR, at least in the short time frame of 5 years. They interpret this to mean that the positive association between ASR and GDP growth rates for low income countries are more likely to reflect causality running from ASR to growth rates.

Bloom (2001) assumes a production function for country i in year t of the form

$$\log Y_t = a_t + \beta \log K_t + \alpha \log W_t + \phi_1 x_t + \phi_2 s_t + \phi_3 h_t + \phi_4 h_t^2$$

where Y_t is the aggregate GDP of country i at time t, x_t is life expectancy at birth in country i at time t, W_t is the size of the workforce at time t, st is the average number of years of schooling and h_t is a measure of human capital as the average number of years of working experience. The (log of) total factor productivity (TFP), representing the level of technological progress, of country i at time t a_t is unobserved. Bloom assumes that

$$a_t = a_t^* + v_t$$

where

$$v_t = \rho v_{t-1} + \varepsilon_t$$

Each country has a long-run, steady state value of TFP of a*. Actual TFP deviates from this value by v_t, which has a systematic and an idiosyncratic component. It is assumed that a country will return to its long run steady state over time following a random shock.

Bloom et al fit two models which differ in the assumptions made about a*. In the first model it is assumed that all countries are converging towards a common level of TFP, albeit at different speeds. In the second model, it is assumed that some countries enjoy long run advantages, so that a* differs between countries, and that this variation can in part be explained by the quality of governance and the proportion of land area in the tropics.

Bloom estimates the change in log output Δy_t by taking first differences of the production function and substituting for the TFP term a_t. They use an instrumental variable to try to capture the causal effect of health on growth in GDP. They assume that lagged levels and growth rates of inputs serve as valid IVs. They estimate the equation using a panel of countries every 10 years from 1960 to 1990. The report does not state which countries were included.

Table 1. The factors associated with change in log output. Coefficients estimated by the analysis of Bloom 2001

	Common long run level of total factor productivity	Country specific long run TFP
Capital	0.342*	0.190
Labour	0.708*	0.824*
Schooling	0.082	-0.025
Experience	0.266	-0.059
Experience^2	-0.005	n/a
Life expectancy	0.013	0.040*
Number of countries	175	147

In the first model, the coefficient on life expectancy is 0.013, suggesting that on average raising life expectancy by 1 year increases growth by 1.3%, though this effect is not well determined and is not statistically significant. In the second model, the coefficient is 0.040, which is significant (Table 1).

The authors conclude that health has a positive and significant effect on economic growth. However, these results are sensitive to the functional form of the model chosen. Other functional forms are also possible, and health might affect other variables including life cycle savings and returns to investment in education. As there are relatively few countries in the world but many potential explanatory variables for growth, macroeconomic data tend to suffer from few degrees of freedom. Furthermore the variables in the models tend to move together over time and show a high degree of multicolinearity. This weakens their power to detect effects.

Acemoglu and Johnson (2007) set up the Solow neo-classical growth model to include health as a determinant of population growth, human capital and aggregate productivity (Barro and Sala-i-Martin). Economy i has the constant returns to scale aggregate production function

$$Y_t = (A_t.h_t.P_t)^{\alpha}.K_t^{\beta}.L^{1-\alpha-\beta}$$

Where

Y is output, A is the total factor productivity (TFP), L is land area, K the capital stock, P the size of the population and h is the average efficiency per worker (that is, human capital per person), α the partial output elasticity of labour and β the partial output elasticity of capital. It is assumed that health (proxied by life expectancy) may increase output per capita through a variety of channels, including more rapid human capital accumulation through greater incentives to invest in human capital (h) or direct positive effects on total factor productivity (A). These effects can be captured in reduced form relationships:

$$A_t = \bar{A}.X_t^{\gamma}$$
$$h_t = \bar{h}.X_t^{\eta}$$

where Xt is average life expectancy in country i at time t, and $\bar{A}$ and $\bar{h}$ indicate baseline values of productivity and human capital (in 1940) for country i.

Greater life expectancy leads to greater population P, both directly and also indirectly by increasing births as maternal health improves and more women live to childbearing age, so

$$P_t = \bar{P}.X_t^{\lambda}$$

Dividing output Y by population P to obtain output per person, taking logs, and substituting for health effects on growth of h, A and P gives

$$y_t = \beta \log \bar{K} + \alpha \log \bar{A} + \alpha \log \bar{h} - (1-\alpha) \log \bar{P} + [\alpha(\gamma+\eta) - (1-\alpha)\lambda]x_t$$

where x_t = log X_t or log (average life-expectancy) and y_t = log (Y_t/P_t) or log(output per person)

This shows that an increase in (log) life expectancy will raise income per capita if the positive effect of health on TFP and human capital measured by $\alpha(\gamma+\eta)$ exceed the potential negative effect arising from the increase in population because of fixed land and capital supply $(1-\alpha)\lambda$.

Acemoglu and Johnson extend the model so that the supply of capital adjusts as life expectancy, population and productivity of factors of production change.

$$K_{t+1} = K_t - \delta K_t + \sigma Y_t$$

where δ is the depreciation rate of capital and σ is the rate of savings (and capital accumulation) . After population and the capital stock have adjusted, the steady-state capital stock with no population growth is

$$K = \sigma Y \ / \delta$$

The long run relationship between log life expectancy x and log income per capita y is

$$y_t = \alpha/(1-\beta)\log \bar{A} + \alpha/(1-\beta)\log \bar{h} + \beta/(1-\beta)\log \sigma - \beta/(1-\beta)\log \delta$$
$$-(1-\alpha-\beta)/(1-\beta)\log \bar{P} + [\alpha(\gamma+\eta) - (1-\alpha-\beta)\lambda]/(1-\beta)x_t$$

Capital now adjusts to the increase in population and productivity resulting from the improvements in life expectancy. If land plays a small role in production (e.g in developed countries) then (assuming constant returns to scale $1-\alpha-\beta \approx 0$) the potential negative effect of population disappears as the effect of life expectancy on growth in GDP is given by $\alpha(\gamma+\eta)/(1-\beta)$. This quantity is expected to be positive. For countries with a substantial agricultural sector $1-\alpha-\beta > 0$, the effect of growth in life expectancy on growth in GDP per capita is given by $[\alpha(\gamma+\eta)-(1-\alpha-\beta)\lambda]/(1-\beta)$ which depends on the positive externalities of greater health γ and η versus the negative effects of the population response λ.

Acemoglu and Johnson estimate the effect of increasing life expectancy on population growth and economic performance (GDP and GDP per capita), using an instrumental variable to capture the causal effect. They compare data from 47 countries from 1940 to 2000. They find that a 1% increase in life expectancy leads to a 1.7-2% increase in population, but a much smaller and insignificant effect on total GDP, and a negative but insignificant effect on GDP per capita.

The instrument is 'predicted mortality'. Disease-specific mortality M_{dit} in each country i at time t was obtained for a set of 15 diseases: tuberculosis, malaria, pneumonia, influenza, cholera, typhoid, smallpox, whooping cough, measles, diphtheria, scarlet fever, yellow fever, plague, typhus and dysentery. It was assumed that a 'global intervention' for each disease d became widely available at time t. For example, streptomycin became available globally in the 1940s to treat tuberculosis. Predicted mortality is constructed as the interaction between baseline mortality for disease d in country i in 1940 (M_{di40}) and the global intervention date for that disease.

$$M_{it}^{I} = \sum_{d \in D} [(1 - I_{dt}) M_{di40} + I_{dt} M_{dFt}$$

where I_{dt} is a dummy for intervention for disease d at time t. M_{dFt} is the mortality rate from disease d at the health frontier of the world at time t. Predicted mortality is thus the country's mortality rate in 1940 from the 15 diseases until there is a global intervention; after the global intervention the mortality rate from that disease declines to the frontier mortality rate. The authors suggest this makes it a good instrument for health as variations in predicted mortality are unrelated to any actions, population shocks or economic events in the country which might be influenced by the dependent variables (GDP or GDP per head).

The study finds that for all countries in the base sample a 1% increase in life expectancy between 1940 and 1980 leads to a 1.67% (se 0.5) increase in population. For low and middle income countries in 1940, a 1% increase in life expectancy leads to a 2.04% (se 1.01) increase in population. For GDP, however, a 1% increase in life expectancy between 1940 and 1980 leads to a non-significant 0.32% (se 0.84) increase in GDP. For low and middle income countries in 1940, the elasticity is -0.39% (se 1.44). For GDP per head, for all countries a 1% increase in life expectancy between 1940 and 1980 leads to a fall in GDP per head of -1.32% (se 0.56), and for low and middle income countries, the elasticity is -2.35 (see 1.13).

Acemoglu and Johnson attempt to reconcile their estimates with the neoclassical growth model to obtain estimates of the effect of health on total factor productivity and human capital accumulation ($\gamma+\eta$). Using the estimates of the elasticity of the response of the population to life expectancy (λ) of 1.67, and the elasticity of response of GDP per head to life expectancy of $[\alpha(\gamma+\eta)-(1-\alpha-\beta)\lambda]/(1-\beta) = -2.35$, and assuming the share of land, labour and capital is each one-third of production for low and middle income countries ($\alpha=\beta=1-\alpha-\beta=1/3$), then $\gamma+\eta$= (-2.35 x 2/3)/(1/3) + 1.67 = -3, that is, the Solow model can only be reconciled to the empirical estimates if health has a negative effect on the rate of total factor productivity growth and/or human capital growth. The authors conclude that the data shows there are other important factors for understanding the effect of health on growth that are not captured by the neoclassical growth model.

The problem faced by all the papers reviewed above is that economic theory is not explicit enough about the set of explanatory variables to include in cross-country economic growth regressions. The studies by Acemoglu and Johnson (2007), Bhargava et al. (2001) and Bloom et al (2001) include similar sets of countries over similar time periods, but differ in the explanatory variables and functional form employed. The approach suggested by classical statistics is that all potential regressors should be included allowing the data to reject the insignificant ones. This is usually not a feasible procedure because the number of potential regressors exceeds the number of countries in the world.

Doppelhofer et al (2004) attempt a novel solution, known as Bayesian Averaging of Classical Estimates. In this, they average across potential models by attaching probabilities that each is the 'true' model. The dependent variable is the annualised growth rate 1960 to 1996 and the independent variables are calculated as close as possible to 1960. The model is fitted using OLS. They find that the strongest predictor of growth is primary school enrolment in 1960, the relative price of investment goods and the initial level of GDP (in 1960). The latter variable supports the theory of conditional convergence, with growth more rapid in low income countries. Human capital-related variables (life expectancy in 1960, the proportion of the country in the tropics, and the prevalence of malaria) are also important. The authors suggest that the prevalence of malaria could be acting as a catch-all and picking up the influence of other variables.

The study by Doppelhofer et al differs from the others reviewed in this section in that they use OLS to estimate growth rates as a function of variables in a baseline year. The advantage of this approach is that they do not need to be concerned about lagged effects or the direction of causality. The disadvantage is that they are not making use of the full panel of data available over time, and are ignoring advances made in health and other factors after 1960.

3.3 The effect of income inequalities on health inequalities in the UK

This section reviews trends in income inequalities and income-related inequalities in health in the United Kingdom during the last decade. There has been considerable attention in the UK on reducing health inequalities following the establishment of cross-departmental public sector performance targets in England (DH 2009). These established that the gap between life expectancy at birth between the bottom quartile of health authorities (the 'spearhead' group) and the national average in England should reduce by 10% over 10 years. The system of performance targets is discussed in more detail in Chapter 5.

3.3.1 Trends in income inequalities

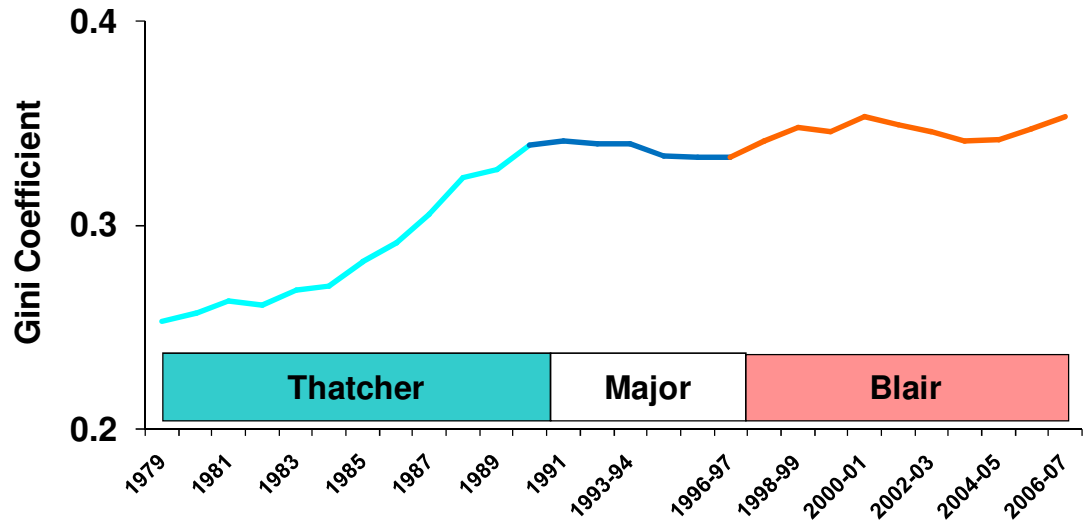

Figure 2. Gini coefficient in Great Britain from 1979 to 2006/07. (Brewer et al 2008: reproduced with permission)

The Gini coefficient has been largely unchanged or become slightly more unequal during the Labour government in the UK from 1996, despite a progressive tax and benefit regimen and a decline in relative poverty (Figure 2). Figure 3 shows why this might have occurred. Inequalities in income have narrowed in the bulk of the distribution, but the tails have diverged considerably. Incomes for high earners have increased much more rapidly than any other group in percentage terms, while incomes for the very bottom few percent of earners have fallen in absolute terms compared with 1996/97. These trends have cancelled out the reduction in inequality in the bulk of the population, leaving the Gini coefficient largely unchanged.

The factors influencing income inequality can be broadly classified as wage inequality and technological change, fiscal policy and demographic trends. There is no overarching research study that has isolated the relative importance of each of these factors. The rapid divergence in earnings between more and less educated workers is likely to have been the driving force behind the rapid rise in income inequality during the 1980s. Different factors may explain the trend from 1997 to 2007. First, the supply of educated workers caught up with rising demand, reducing the upward pressure on wages of educated workers. Second, technological change (such as computerisation of non-manual tasks) may have depressed wages in the middle of the wage distribution more than the top and bottom. Third, in general the tax-benefit system since 1997 is more progressive than in the 1980s. Fourth, globalisation has magnified the rewards available to those at the very top of the income

distribution. Fifth, for those at the very bottom, the trend is complex, but suggests that poverty has increased among particular groups, such as unemployed working age adults without children (Sefton et al 2009).

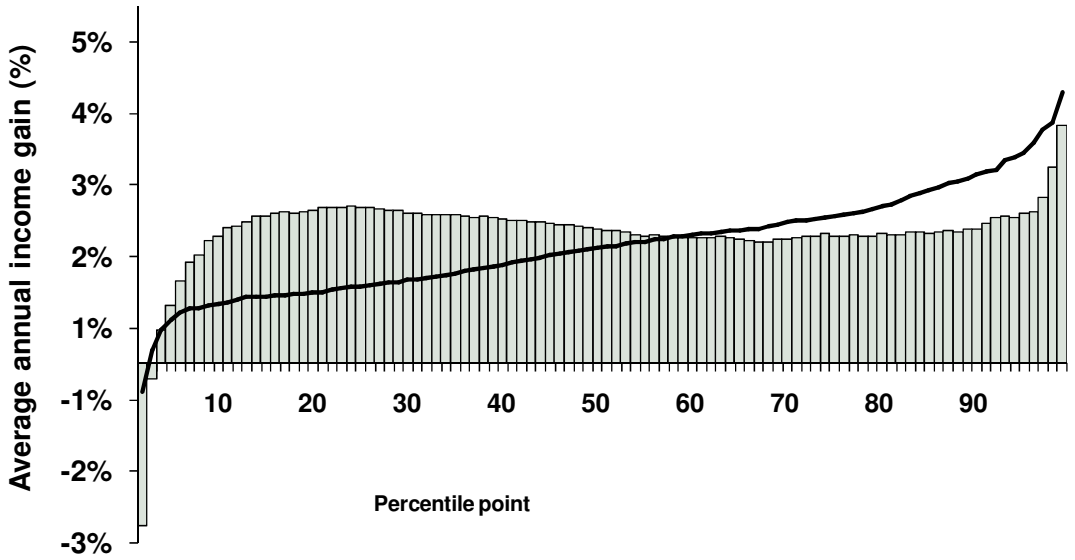

Figure 3. Change in income 1996/97 to 2006/07 by income percentile in Great Britain. Solid black line shows change in income for 1979 to 1996/97. (Brewer et al 2008)

3.3.2 Trends in health inequalities

Data for 2004/2006 show the relative gap in life expectancy between England and the Spearhead group is wider than at the baseline (1995–97) for both males and females, with year-on-year fluctuation (Table 2). Despite the considerable resources and policy attention that inequalities in health have received, it appears to be extremely difficult to make even relatively a modest impact on macro-level trends in life expectancy.

Table 2. Life expectancy at birth and for the 'Spearhead' group.
Source: DH 2007a

	1993-95	***2004-06***	***Target 2010***
Men			
England average	74.2	77.3	
Spearhead average	72.3	75.3	
Difference	1.9	2.0	
% gap	2.51%	2.63%	2.32%
Women			
England average	79.4	81.6	
Spearhead average	78.0	80.0	
Difference	1.4	1.6	
% gap	1.74%	1.96%	1.59%

The model outlined in Chapter 2 showed that changes in IRHI depend on the evolution of the income distribution, the 'proportionate' effect of income on health at different points in the income distribution, and the evolution of other social determinants.

Gravelle and Sutton (2003) examine the evolution of IRHI in Britain 1979-1995 using a version of the concentration index based on 'standardised' self-assessed health from which the influence of factors correlated with income have been removed. They find that rising income inequality was the primary cause of increasing IRHI in the early 1980s. Subsequently, the main driver of the increase in health inequality was an increasing proportionate effect of income on health (or elasticity) as mean incomes increased.

The results of Gravelle and Sutton are confirmed by subsequent work that looked in more detail at the elasticity of health at different points on the income distribution. Van Ourti (2009) examined IRHI in EU countries during the 1990s, based on the European Panel Health Survey. The measure of health was Self-Assessed Health. They found small increases in IRHI (measured by the concentration index) in the majority of EU countries, with income growth (measured by Gini coefficient) that was slightly pro-poor or distributionally neutral. In the UK, the concentration index was 0.0091 in 1994 (a positive value indicates that the distribution of health is pro-rich), and 0.0111 in 2001.

The Gini coefficient for income distribution was 0.3004 in 1994, increasing slightly to 0.3014 in 2001, consistent with national data shown in Figure 2. The study found that income elasticity of health is increasing in income (Figure 4), from 0.014 for the lowest decile to 0.037 in the ninth decile. While these elasticities are relatively small, cumulative income growth has been substantial over the last 10 years or so up to 2007. Average annual growth in real after-tax income has been about 2.5% per year, implying an increase in self-assessed health over 10 years of less than 0.5% for low income groups but more than 1% for high-income groups. Consequently, the roughly proportionate growth in income seen in the UK during the 1990s has improved the health of the richer groups more than the poorer groups. If one accepts that SAH is correlated with clinical morbidity and mortality, then the finding that income elasticity of health is increasing with income goes some way to explaining the lack of progress on the target for reducing inequality in life expectancy at birth over this period.

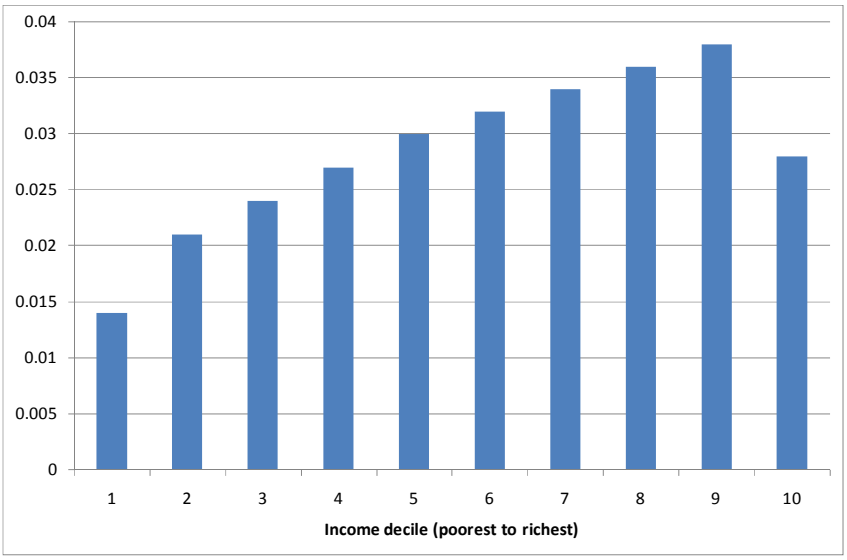

Figure 4. Income elasticity of health in UK by income decile, mean over period 1994-2001.
Source: Van Ourti et al 2009
Note: Income elasticity of health is the percentage increase in health for a 1% change in income. For example, an elasticity of 0.02 means health will change by 0.02% for a 1% increase in income

It should also be noted the headline target chosen by the government is only one of many possible ways to measure and summarise the distribution of health in the population. Sassi (2009) re-analyses ONS data using regression analysis to estimate the 'slope index of inequality', representing the difference in life expectancy between the least and most deprived health authorities. This study found that inequalities increased during the 1990s but have declined somewhat since 2002 in both men and women, which may (at least on this measure) represent a reversal of the trend.

3.4 Health during economic downturns

If, as found by van Ourti et al (2009), average income and average health has improved in the UK during the 1990s, and income inequalities and IRHI have slightly worsened, then one might assume that we can predict the impact of the crisis by running the van Ourti model in reverse. This would conclude that during an economic downturn, average health would be expected to worsen. Given the social protection safety net in the UK, the downturn might be expected to reduce inequalities in income, with greater proportionate decline in income for higher income groups, and possibly improve IRHI.

Ruhm (2006) presents contrary evidence. Health, at least in some dimensions, may improve during an economic crisis in developed countries. Lifestyle changes can be rapid, with less problem drinking, less smoking, and more exercise. There are fewer traffic fatalities and accidents as people drive less. Mental health indicators may worsen, but the evidence that suicides increase during downturns is mixed.

Ruhm (2006) and van Ourti et al (2009) appear to present contradictory results. There are of course many possible reasons, given that these are entirely independent studies. Ruhm evaluates mortality and morbidity indicators while van Ourti evaluates the effect of income on self-assessed health. Behaviour changes may be different when income increases compared to when it falls. Neither van Ourti nor Ruhm distinguishes between 'permanent' and 'transitory' income changes. One might expect transitory income changes to have a limited effect on consumption and health behaviour (Contoyannis et al 2004).

Adda et al (2008) studied the effect of permanent income shocks on health in England, using national cross-sectional survey data for adults aged 30 to 60 years over the period 1978 to 2003. The study aimed to estimate the relation between income and health, controlling for reverse causality (the effect of health on income) and underlying factors influencing both health and income.

The authors used an individual dynamic model of health and income that allowed them to differentiate between permanent and temporary changes to health and income. They aggregate this model to estimate results for age-sex-education cohorts, rather than at individual level. This serves two purposes. First, it allows them to specify an exogenous indicator of permanent income for the cohort, based on the timing of supply side reforms such as decline in unionization and increased competition that is not affected by health. Second, they can combine data for income, mortality, morbidity, consumption and health behaviour at cohort level from several datasets and surveys.

Adda finds that different cohorts were affected by sizable permanent shocks to income over that period, especially those with low education. There are three main findings with respect to income and health. First they find, as in Ruhm (2006), an increase in permanent income decreases mortality in working age adults. A 1% increase in income was associated with 0.7 to 1 additional death per 100,000 prime-aged adults in any year.

Second, they find that these shocks are not transmitted to other health measures, whether subjective ones (self-assessed health, longstanding illness) or objective ones (high blood pressure, cardiovascular diseases, or respiratory diseases).

It should be stressed that Adda is looking at the effect on time-series changes in income on changes in health, rather than the socio-economic 'gradient' of health across a cross-section at a given point in time. The results of Adda should only be compared with studies that have excluded reverse causality (the effect of health on income) and other factors that influence both health and income (such as state dependence).

The following points might be noted:

- Adda's results appear contrary to the results found by Van Ourti (2009), who concluded that change in income does positively affect health. However, Van Ourti does not entirely exclude the effect of health on income when they estimate the effect of income on health, although they state that a dynamic model gives similar conclusions.
- Adda's results are contradicted by Contoyannis et al (2004), who used BHPS data, controlling for reverse causality and state-dependence. Contoyannis found that that permanent income has a much greater (positive) impact on self assessed health than transitory income and also that the impact of permanent income is larger for men than women. Nevertheless, no other studies have looked at changes in income on morbidity and mortality at a level of detail comparable to Adda.
- Adda restricts the study to working age adults. Other studies often quoted in the literature have looked at the relation between health and income for other specific groups, such as pensioners (Case 2001).
- Adda is using a synthetic measure of 'permanent income' (linked to how structural changes in the economy affect potential wages of various cohorts) rather than actual income.

- There are differences between the functional forms used in the econometric models. It may be important whether log(income) is regressed on log(health), or health.
- Adda suggests that their data at cohort level incorporates macroeconomic or general equilibrium effects, such as the effect of income changes on price levels.

Third, Adda found individuals change some of their behaviour such as total expenditure as well as increase expenditures on tobacco and alcohol. Consumption of fresh fruit and vegetables tends to increase with income, but not significantly. These results might be compared with those from the BHPS, which found that those in higher social classes were more likely to be non-smokers (Contoyannis and Jones 2004). The results are not necessarily inconsistent: Contoyannis and Jones are looking at the average gradient of smoking over social class, while Adda is looking at the average effect on smoking of changes in income over time.

Their interpretation of these three findings is that risk behaviours do not seem to transmit directly into mortality or morbidity for the working-aged population. Procyclical mortality, Adda suggests, is probably rather driven by work-related accidents and similar mechanisms. This finding appears contradicted by the BHPS, which found health behaviours (sleeping well, exercising and not smoking) have a dramatic positive effect on SAH in the proximate wave of the survey 7 years later (Contoyannis and Jones 2004). While one would expect SAH to be predictive of mortality (Gravelle and Sutton 2003), they are of course different indicators.

Adda concludes that income redistribution towards particular age-education cohorts of working age adults is unlikely to lead to improvements in health at least in the short to medium run (up to three years). This conclusion is based on their finding that increasing (permanent) income has on average been associated with worse health behaviour around smoking and alcohol, and little effect on health or morbidity. The study has a number of limitations. Income shocks might take longer to feed through into health outcomes than the study allowed for (3 years). The results may be sensitive to the particular functional form chosen. The model only picks up the effect of a very broad measure of income, linked to structural changes that are thought to affect (potential) earnings of working age adults in the economy. The model only shows the average effect on health. Results are not presented disaggregated by socio-economic group, so it is difficult to assess the implications for health inequalities. Finally, the results appear to be inconsistent with data from the BHPS which found on average a positive effect of lifestyle on SAH 7 years later, and a positive overall effect of changes in income on changes in SAH.

4. Initiatives to change lifestyle and consumer behaviour

The model in Chapter 2 suggested income might be one of the variables that influence health behaviour, either positively or negatively. However, it did not explain why the prevalence of healthy or unhealthy behaviour differs between income (or other socio-economic) groups. One of the contributions of economics to the study of public health and prevention has been to try to provide plausible explanations of what factors influence lifestyle choices. Such models may

- provide understanding of which social and economic factors influence health behaviour and health
- indicate the circumstances under which private preventative effort might be considered inadequate from an economic perspective
- help design policies that improve health and health inequalities
- help inform methods to evaluate the effectiveness and cost-effectiveness of policies and guide priority setting.

The model reviewed in this chapter is a standard microeconomic framework of household behaviour based on the work of Becker (1965) and extended by Grossman (1972). In this framework, people choose a level of health promotion activity for themselves and their families, taking account of their economic circumstances, prices, other time committments, their preferences, the available options etc. The model aims to make predictions about the variables that influence health behaviour.

Whether or not this level of 'private' activity is considered 'optimal' (from a societal point of view) of course depends on one's prior normative position. Welfare economics usually begins from a position that market allocations are efficient unless shown otherwise ('market failures'). However, the microeconomic framework can also analyse other normative positions. For example, equity concerns might be included in the social welfare function, discussed further in Chapter 6. Regardless of one's prior normative position, it can nevertheless be helpful to analyse whether there may be a market failure because this strengthens the case *(a-priori)* that a public health intervention could increase overall social welfare.

In this chapter, we show a simple microeconomic model of household behaviour. We briefly review the epidemiological evidence on the relationship between lifestyle/risk factors (alcohol abuse, obesity and smoking) and social class in the UK, and how these indicators have evolved in recent years. We discuss the kinds of market failures might result in inadequate levels of prevention. Finally, we review some of the kinds of public health policies that aim to change lifestyle and health behaviour. We leave a discussion about evaluating the effectiveness and cost-effectiveness of these policies, and deciding which policies should be prioritised, for Chapter 6.

4.1 Models of demand for health

The conventional microeconomic approach is to assume that people make choices in allocating their time and other resources in a rational manner. Health can be considered a component of human capital. People enjoy health for its own sake, and because health enables and enhances participation in work and other activities (Becker 1965). Achieving a desired level of health requires some investment by the individual in terms of her time and consumption of goods and services (health inputs).

We show a simple one-period microeconomic model of individuals' or household decision-making. There is a health production function:

$H = H(N,L; E_0)$

where H is health, N is a (vector of) health inputs and L is hours worked. N might be use of preventative and curative health-care, or consumption of calories, or other goods, services and activities which are determinants of health and are consumed in the current decision making period and are under the control of the individual. L is assumed to expend energy and is negative to health. The components of vector N might be positive or negative to health. E_0 represents exogenous factors which affect the relationship between the inputs and health outcomes. As this is a one period model, these might include environment, education, childhood health or assets.

Productivity w is assumed a function of health, also given previous endowments such as schooling. It is assumed that the wage commanded by the individual is equal to their marginal productivity.

$w = w(H ; E_0)$

Individual utility is a function of health inputs (N), consumption of things which do not affect health (C), health itself and (negatively) labour hours worked

$U = U(N, C, H(N,L), L)$

The budget constraint for the individual is

$p_cC + p_nN = wL + V$

where V is other sources of income (for example from assets or social security) and p_c and p_n are (vectors of) prices of consumption goods and health inputs respectively. The individual's decision is to maximise utility subject to the budget constraint by choosing the optimum consumption of N, L and C.

The first-order condition for any one health input Nj can then be written

$(1/ \lambda) [U'(Nj) + U'(H).H'(Nj)] - (pn – w'(H).H'(Nj).L) = 0$

Where λ is the marginal utility of income, and w'(H) is the partial derivative *δw/δH.* As with all first-order conditions, this is equivalent to saying that individuals will consume health-affecting goods and services up to the point where marginal benefit equals marginal cost. Marginal cost to the individual has two components: the monetary price of the input (pn), less the marginal effect on labour income caused by the use of Nj on health, and consequently productivity. Marginal benefit is the utility arising from additional consumption of the good itself U'(Nj) plus the additional health induced from consuming additional Nj.

In this model of rational behaviour, individuals might undertake some preventative health activity, for example consume vitamins, because they believe they will derive some benefit to health and productivity (H'(Nj)>0) even if they derive very little utility or even negative utility directly ($U(Nj) \leq 0$), and at some financial cost ($pn \geq 0$). On the other hand, people might undertake harmful activity, such as smoking, because they derive direct utility ($U(Nj) > 0$) even though they recognise the cost to health and productivity ($H'(Nj) < 0$).

In the following sections of this paper we show how these kinds of microeconomic models have been used to suggest circumstances where private preventative effort might be inadequate (market failures), to design policies to improve health behaviour and reduce health inequalities, and help priority-setting.

The next section presents statistics based on cross-sectional national survey data showing the social gradient in smoking, obesity and alcohol use. . These are descriptive statistics and therefore do not show causality. Health may affect occupation and social class, or other factors may influence the association. We refer in later sections to examples (eg Contoyannis and Jones 2004) of microeconometric studies (based on the model described above) that have tried to estimate such causal relationships. Nevertheless, these descriptive statistics are useful indicators of the general direction and magnitude of the social gradient in health behaviour and how it might be modified by other variables such as gender.

4.2 Social gradient in health behaviour

4.2.1 Alcohol

Social class is a risk factor for alcohol-related mortality, with men in manual occupations being significantly more likely than professional men to die of alcohol-related causes. It is suggested therefore that problem use is linked to social structural factors such as poverty, disadvantage and social class. However, the picture is not a simple one.

Figures 5 and 6 show how the impact of class on alcohol-related mortality is mediated by age and gender. For example, in men, the difference in alcohol-related mortality between unskilled manual and professional is greatest in the 25-39 age group, declining thereafter. However, in women, for those in paid employment there is no consistent class gradient: in the young, those in manual occupations have raised mortality, but in older women it is the professionals who have the highest risk of dying from alcohol-related causes (Harrison and Gardiner 1999; McNeill 2003).

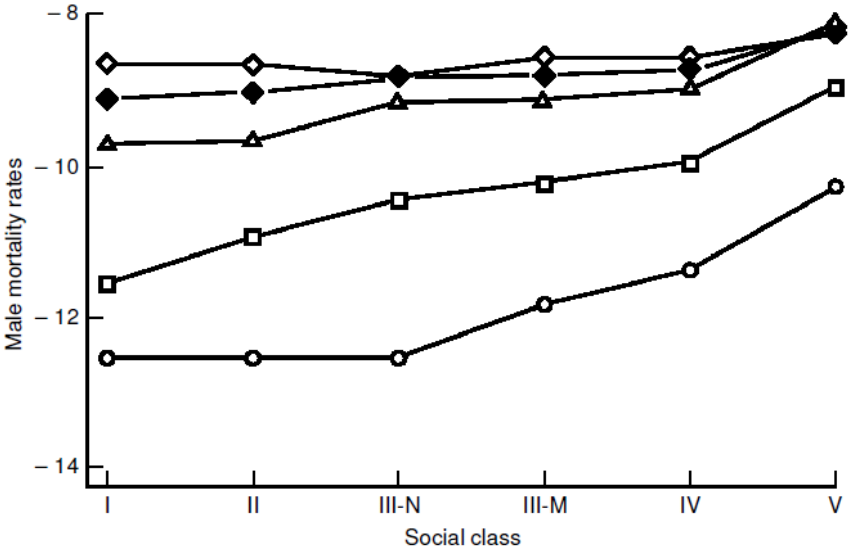

Figure 1. *Logarithms of male mortality rates. Age group: ○, 16–24; □, 25–39, △, 40–54; ◆, 55–64; ◇, 65 + .*

Figure 5. Logarithms of male mortality rates from alcohol (Harrison and Gardiner 1999) by age and social class. (Permission sought)

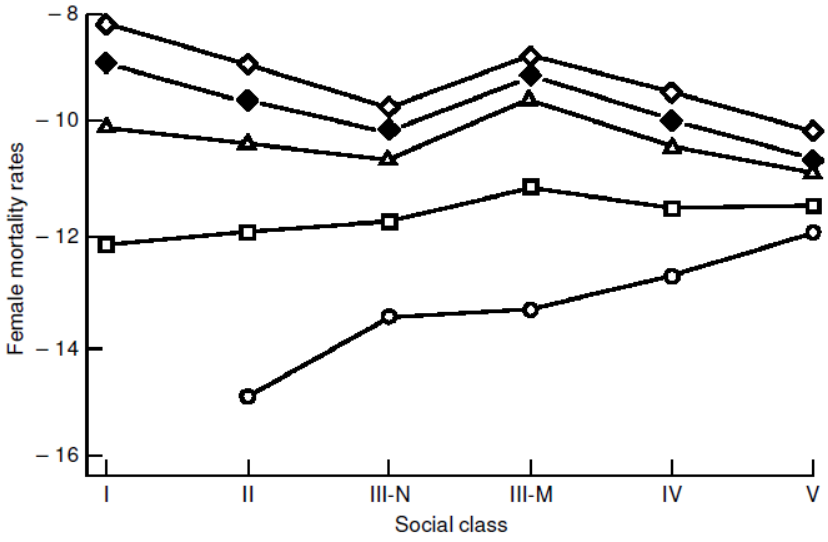

Figure 2. Logarithms of female mortality rates. Age group: ○, 16–24; □, 25–39; △, 40–54; ◆, 55–64; ◇, 65 + .

Figure 6. Logarithms of female mortality rates from alcohol (Harrison and Gardiner 1999) by age and social class. (Permission sought)

There is little evidence of strong class differences in relation to average consumption of alcohol or the prevalence of hazardous drinking, though there is a somewhat raised risk of hazardous consumption in manual occupations compared with non-manual. Being on benefit appears to reduce the risk of hazardous consumption (McNeill 2003).

4.2.2 Smoking

HSE data show that smoking rates have been generally declining in England, particularly among men, and in all socio-economic groups over the period 1991-2004. Sassi (2009) compared trends in three-year moving averages of smoking rates in manual and non-manual groups and found that smoking rates appeared to stagnate or even increase slightly during the period 1994-97 but a sharp declining trend resumed in 1997-2004, with both groups achieving similar absolute reductions in smoking rates. Smoking rates are still higher in men than women in disadvantaged groups but are now lower in men than women in higher social classes.

Bauld (2007) assessed the impact of NHS Stop Smoking Services, a programme specifically designed to reduce social inequalities in smoking rates. Funding in the most deprived tenth of PCTs was almost 70% greater on a per capita basis than the least deprived tenth. The proportion of smokers in Spearhead areas who accessed such services (16.7%) was higher than in other areas (13.4%). Despite success rates in Spearhead areas being slightly lower (52.6%) than elsewhere (57.9%), a higher proportion of smokers reported success in Spearhead areas (8.8% v 7.8%). However, given that such services only reach a minority of smokers the authors conclude that this programme will only have a modest impact on overall geographical health inequalities. The study did not address how the programme might have affected socio-economic health inequalities, the indicator measured by the PSA target.

4.2.3 Obesity

The social gradient for obesity is more pronounced for women than men, though prevalence has increased rapidly in all classes between 1998 and 2006 (Figures 7a and 7b).

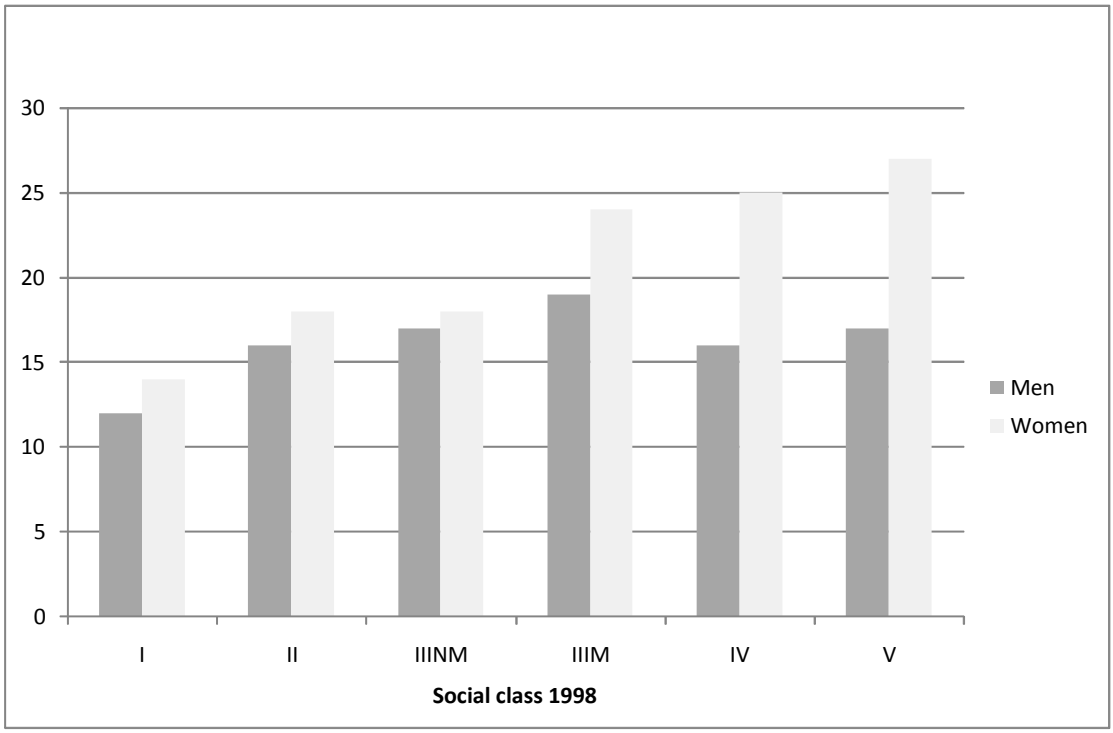

Figure 7a. Prevalence of obesity by social class
Source: Health Survey for England 1998

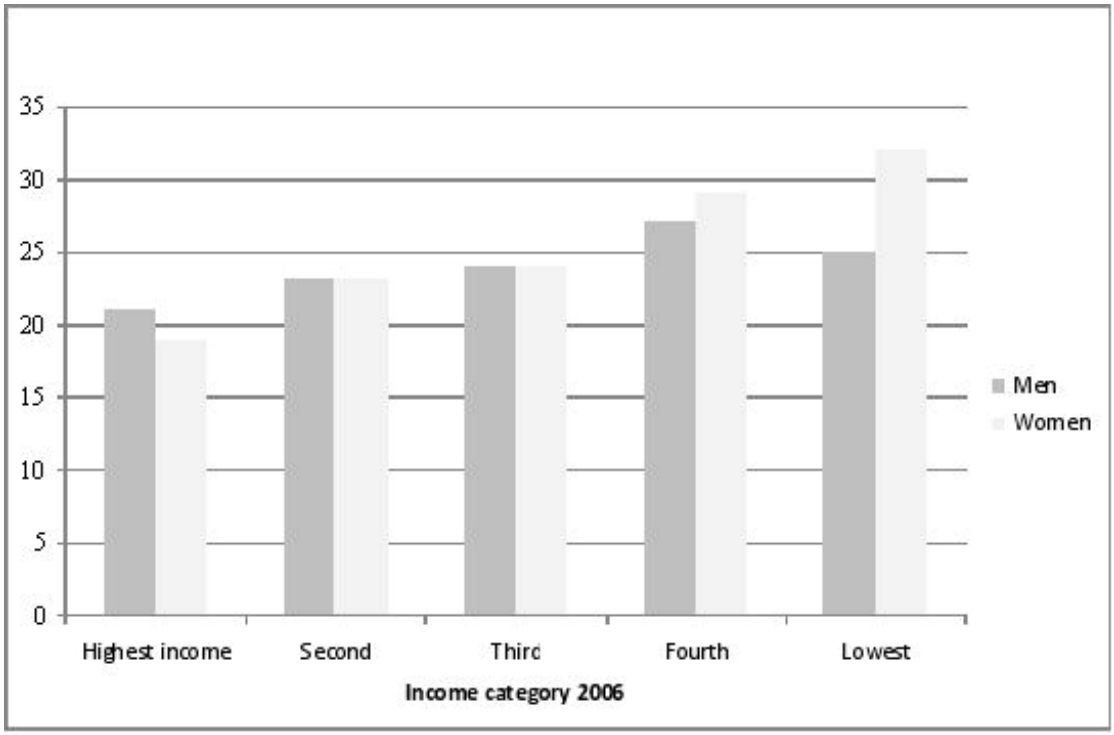

Figure 7b. Prevalence of obesity by income quintile
Source: Health Survey for England 2006

Assessment of the social gradient of health and health behaviour is confounded to a greater or lesser extent by reverse causality. Not only do people, or at least women, who earn less tend to be more likely to be obese, but people with obesity tend to earn less (House of Commons Select Committee 2004). A similar endogeneity problem might affect assessment of alcohol abuse.

In the remainder of this chapter, we reflect on the ways in which private choices about health-related activity might not be 'efficient' from a societal perspective, and the types of government policy that have been proposed to encourage greater preventative effort.

4.3 Types of market failure for prevention

People may have inadequate information about risks to their health of the goods they consume, environmental, workplace and other hazards, the benefits of healthy behaviour, or knowledge of how to achieve a healthy lifestyle. This might indicate some role for government to disseminate information; however, considerable preventative information is already available and the extent to which information by itself leads to sustained behavioural change remains a debated point in health promotion and public health. Information alongside other incentives may have more effect and are discussed later. In health-care, with patent protection, there are clearly incentives for pharmaceutical and medical device companies to undertake research into both curative and preventative medicine. Outside of the medical-pharmaceutical sector, there is not a clear 'industry' producing preventative activity. Consequently the market may under-invest in research into prevention and policies on the social determinants of health, indicating a more active role for government (Kenkel 2000).

Prices may not correctly reflect marginal (opportunity) costs in all relevant markets. Because insured or publicly-financed health-care is free at the point of need, this may diminish the personal (financial) incentives people face to prevent ill-health, leading to inadequate private levels of prevention. This is a classic 'second-best' problem: the zero price of health-care might induce moral hazard, and justify government intervention in health-care or other markets to encourage greater prevention (Lipsey and Lancaster 1957). One response can be to tax unhealthy products. Kenkel (2000) speculates that one of the reasons why tobacco and alcohol taxes are higher in Europe than the US may be to offset the ex ante moral hazard created by their public sector health insurance systems. Another type of intervention to offset moral hazard might be personal copayments for expensive health care where

the recipient is judged to not have taken adequate preventative measures e.g. increased co-payments for COPD drugs, total knee replacements in those overweight etc. It is worth noting that these examples imply that effective preventative measures exist (for example reducing obesity) that are very likely to avoid future expensive treatments. This would make it cost-effective and possible cost-reducing to implement such measures to avoid future disease. Another policy response might be for NICE to set a higher cost-effectiveness threshold for technologies where patients are thought to be 'responsible' for their condition, thereby making it more difficult for such technologies to be funded by the NHS. This measure is discussed in Chapter 6. It is difficult to judge the importance of moral hazard from health-care insurance on prevention in England. Theory suggests that even if insurance does induce ex-ante moral hazard, the impact on health behaviour is likely to be dampened because unlike other types of insurance, health-care offers an uncertain cure, meaning there are still incentives for the insured individual to look after her health. Furthermore, primary care offers low or zero cost screening and other preventative activity. Most empirical evidence seems to be from US cohorts, identifying people who have or do not have health insurance for reasons unrelated to their health. A study of people obtaining Medicare coverage at age 65 found limited evidence that obtaining health insurance reduces prevention and increases unhealthy behaviors among elderly persons. The study found more robust evidence that physician counseling is successful in changing health behaviors. Other evidence from the US (using employment status as an instrument for health insurance status) found working-age adults who obtained health insurance were more likely to be obese and less likely to exercise, but no more likely to smoke (Golcuk 2008). Individuals understanding and estimation of the benefits of prevention on their health may not take account of the benefits (or costs) to those around them. A classic externality is the benefit of 'herd immunity' offered by vaccination against infectious disease. Another example might be the impact on the household if a member falls ill, particularly a chronic illness, such as the need for personal care, stress, and perhaps reduced opportunity for education of other family members. Suhrcke et al (2006) stresses that it should not be assumed that such externalities are large – they might be or not.

There is now widespread agreement that 'passive' smoking imposes considerable external impact on the health and well-being of others. Heavy use of alcohol and binge drinking have been shown to have significant external costs: on crime, the urban environment, family life and children's education (CMO 2008). Obesity is considered to be a growing public health problem and future liability for the health service (Foresight 2007). Costs on the health service might be considered a type of externality, in the sense that taxpayers will be asked to fund these costs. A loss to personal productivity is partly a private cost, but chronic incapacity imposes an external liability on the taxpayer. The House of Commons Select Committee 2004 report estimated the public cost of lost personal productivity due to obesity as £2.4-2.6 billion per year, based on early mortality and incapacity benefit claimants with obesity.

One example of an externality with particular relevance for inequalities is that health behaviours tend to be clustered geographically. To some extent, this may simply reflect that people tend to react to similar social circumstances in similar ways. Microeconomic models of consumer behaviour tend to focus on individual and households. However, clustering may persist even after adjusting for observable individual and household variables. This indicates some factor outside the household is influencing behaviour. One candidate is peer-group influence: cultural or social norms. Cultural (group) behaviour tends to persist even when social circumstances change. Such clustering was observed in patterns of obesity in a detailed study of a community in the United States (Christakis 2007). This is not necessarily a market failure, but a kind of 'multiplier effect' or feedback, possibly indicating that interventions might need to consider targeting peer groups, rather than just individuals, to be effective.

Externalities can be found in government policy as well as private activity. There is considerable evidence that education is a key social determinant of health (DH 2009: Early Child Development Task Group). At an aggregate level, therefore, the health impact of education ought to be included in estimates of the optimal size of education and the type of activity undertaken. One of the objectives of the 'Every Child Matters' (Department for Children 2008) strategy is to ensure education has more influence on child health. The Common Agricultural Policy (CAP) has some negative externalities. The CAP has subsidised production of foods that are high in saturated fat (dairy, meat) and kept prices low. The effect on health and health inequalities has not been fully taken into account, or even properly evaluated (Salay and Lincoln 2008).

These economic models of demand for health and prevention have made some important contributions to the debate on public health. They attempt to explain behaviour and its causes, rather than simply describe it. There is a wide literature of econometric analyses in public health that have based empirical work on conceptual models such as the one outlined above (eg Thomas and Strauss 1997; Contoyannis and Jones 2004), to estimate the relative causal influence of factors that determine lifestyle choices and health.

Economic analysis can indicate situations where private preventative effort might be inadequate from a societal perspective, and can therefore offer *ex-ante* theoretical support for advocates of government action to encourage prevention. This type of analysis also suggests reasons why policy makers might be cautious before implementing preventative policies. The framework is based on an assumption that health may not be the only or even the most important objective to many people when they make everyday decisions about consumption, working and leisure for themselves and their families. They remind policy makers that in many cases people may be making rational choices in this respect, even if these choices are unhealthy.

Theoretical models of the demand for health have several limitations. First, they do not make clear hypotheses about the direction in which age, social class and education might affect preventive effort, depending on the assumptions made and which factors (eg life expectancy) are considered endogenous (Kenkel 2000).

Second, it can be difficult to translate empirical work based on these models into policy. Empirical studies have shown that lifestyle behaviour cannot be entirely explained by observed socio-economic variables, for example, see Contoyannis and Jones (2004). These unobserved factors might represent differences in time preference, differences in childhood circumstance, attitude to risk, differences in health knowledge or opportunity costs in terms of unobserved wage rate and time costs of each lifestyle choice. For example, Fuchs proposed that individuals with a high 'time preference rate' attach relatively greater importance to outcomes occurring now (the pleasure of smoking) as compared to future outcomes (decreased life expectancy). These time preferences might influence other long term decisions, such as education, and may be one of the reasons why people with better education appear to live longer. Some studies have tried to estimate time preference and its role as a determinant of health behaviour, eg for obesity (Komlos et al 2003). This kind of analysis is of interest to researchers who wish to investigate heterogeneity in the population, but would be useful for setting policy only if is possible and acceptable to identify subgroups of the population with high or low discount rate and target policy accordingly.

Third, behavioural economists have attempted to test the fundamental assumption that consumers behave in a rational and consistent manner. Clearly, there are extreme cases such as some kinds of mental illness or learning difficulties where a person cannot make everyday decisions for themselves, and the state must take additional responsibility for them. However, behavioural economists claim myriad ways in which individual decision-making departs from the standard model for a large proportion of the population. Consumers may be influenced by habit; they may have poor self control; they may be addicted; or they may not have the cognitive skill to interpret or act on relevant information. For example, the standard model assumes people discount future events at a constant rate (exponential discounting). However, experimental work has found that people tend to discount events in the near future at higher rates than the far future (hyperbolic discounting), with large variation in the population in the degree to which discount rates diverge over time (Sassi and Hurst 2008). Hyperbolic discounting has been found to relate to real-world examples of lack of self control and addiction. For example, drug dependent individuals discount delayed consequences more than matched nondependent controls. Severe hyperbolic discounting can also lead to inconsistent decisions or procrastination: such as a person would continually put off some long term decisions. For example, a person can know through firsthand experience that drinking is not in her long-range interest and accordingly plan not to drink, but go on a binge when the opportunity arises. This lack of willpower can be seen as a form of irrationality and presents a challenge to conventional microeconomics, which would assume a rational person would learn from their error and adjust their time preferences to be more consistent.

These lines of economic research may nevertheless be useful for policy makers. For example, where research shows a 'lack of willpower' is fairly widespread with respect to a particular health behaviour, this implies that such people accept that a healthier lifestyle is in their own long term interests but are

unable to implement it. This strengthens the case for some kind of government intervention that motivates people without restricting personal autonomy or stigmatizing them. These measures, termed 'liberal paternalism' by Le Grand (2008), might include personal coaching, financial rewards, tax incentives, or measures to make healthy choices the 'easy option' and are discussed later in this chapter.

Addicts are usually thought of as being irrational or out-of-control, and therefore in need of extra cajoling, but there is a school of economic thought that suggests addiction might be rational, though of course this is controversial (Rogeberg 2004). According to this theory, policy should be focused on limiting self-harm and minimising spillover effects on others.

4.4 Policies to promote prevention

Once a *prima facia* case for some kind of intervention has been established, Sassi and Hurst (2008) suggest a taxonomy of polices in order of the degree of restriction of choice:

(a) Increasing healthy options, where the market fails to provide;
(b) Influencing preferences;
(c) Price controls, subsidies and consumption taxes;
(d) Restrictions and bans.

4.4.1 Increasing healthy options

These policies might include government provision of something that the free market would be unlikely to provide, or is not accessible or affordable to low income groups. For example, initiatives have aimed to increase the nutritional quality of school meals and improve children's play provision to address the child obesity target in England. Sassi and Hurst 2008 point out that, in general, because the market does not provide these services, subsidies will probably be required, and so these initiatives might be very financially costly to public services and ultimately taxpayers. On the other hand, such 'entrepreneurial public activity' start-ups might seed future demand, stimulate research and development and allow a private market to develop in the future (eg pay-as-you-go bicycle rental).

4.4.2 Influencing preferences

A wide class of policies aim to modify preferences. Targets of such policies are thought to be people who would in general accept the desirability of personal preventative effort, but lack information, or might need feedback, support, and/or motivation to carry out their intentions. We briefly review some of these measures, and assess their potential to change health behaviour and health inequalities. Examples of such initiatives include provision of information, setting the default option, and financial and non-financial incentives.

Many of the detailed operational initiatives set out in "Choosing Health" (DH 2004) centre around provision of information through food labelling and health-related websites, and 'marketing' health to try to influence preferences. Many policy documents stress the importance of targeting information and advice to meet people's needs. For example, "Choosing Health" stresses the importance of supporting health care staff to communicate complex health information to different groups in the population, and providing additional support for people who lack basic skills to help them use health information. The NICE guideline on obesity in adults (NICE 2006) recommends that advice needs to be tailored for different groups. This is particularly important for people from black and minority ethnic groups, vulnerable groups (such as those on low incomes) and people at life stages with increased risk for weight gain (such as during and after pregnancy, at the menopause or when stopping smoking).

Although tailoring information, advice and support to individual needs appears to be common sense, targeting public health effectively on a large scale may be a complex and costly activity. Boyce et al (2008) review some approaches to identifying the target population and putting in place appropriately tailored interventions. Many specialist prevention programmes receive the majority of referrals from GPs. Some PCTs are using 'geodemographics', mapping software that uses Health Surveys, census data etc and algorithms to identify populations thought to be at high risk or most likely to benefit from

prevention activity. Another initiative is 'social marketing': the use of marketing techniques such as consumer research, segmentation and targeting to persuade people of the personal benefit of greater preventative effort and voluntarily change their behaviour. Kikumbih et al (2005) evaluated a social marketing scheme to increase the use of insecticide-treated mosquito nets in Tanzania. Many of the findings of the study may have relevance for the evaluation of social marketing schemes in the UK, despite the obvious differences in context. The study found that the total cost of each net was greater than those sold through unsubsidised commercial distribution channels, taking account of the user contribution, project administration, advertising and price subsidy. However, the authors considered that the benefit to public health and health inequalities was greater, as social marketing led to greater use in higher coverage of target groups (for example, pregnant women and children under 5 years), in the lowest socioeconomic group, and to reach into more peripheral areas, compared with a pure commercial sector model.

Boyce et al (2008) review other targeted approaches. They find that tailoring written information to the individual, for example, provision of personalised information and practical advice based on answers given in health assessment questionnaire, may help increase motivation, but there was a small evidence base. Piloting is underway of care co-ordinators for patients with chronic illness *Next Stage Review* (DH 2008e). Measures to provide individual support such as health coaches and group sessions might also raise motivation and confidence, support behaviour change as well as co-ordinate access to services but were labour intensive and costly. The authors note there might be scope for online interactive services, or telephone based support, perhaps at a lower cost.

There may be other disadvantages to targeting, as well as the cost. Targeting children in particular who are obese can be stigmatising, worsening the problem. Partly for this reason, the strategy in the UK to reduce child obesity is aimed at all children, rather than the obese or most at risk.

Financial incentives are another way of motivating socially desirable behaviour, for example, stopping smoking. Boyce et al (2008) found that these may work best if used to change simple, one off behaviour, and are most effective if combined with other social support. In most cases, the financial rewards that have been offered to encourage preventative behaviour in the UK are very modest, and the reward may work by clarifying the objective for the individual, recording progress and ensuring that regular feedback is maintained. Nevertheless, Boyce et al points out that financial incentives need to be carefully designed. There is a risk that people return to previous habits when the incentives cease. Some PCTs are piloting schemes where patients hold 'personal budgets' to spend on preventative services.

A type of financial incentive scheme is the Conditional Cash Transfer initiative (CCT), variants of which operate in Mexico, Brazil and many other countries. Families enrolled in CCT programmes receive 'small but meaningful' cash rewards in exchange for complying with certain conditions: preventative health requirements, nutrition supplements, enrolment in school and monitoring. A systematic review found CCT programmes effective in improving school enrolment, nutritional and anthropometric outcomes, and preventative behaviour (Lagarde et al 2007).

Given the success of the Mexican initiative, New York City implemented CCTs as a set of pilot programmes in 2007. The cash payments go to the family, almost always the mother or other female head of the household. There is a detailed schedule of rewards (New York City 2007). There are three pilot schemes: a family focused scheme, with incentives for child education, for child health and prevention, and for an adult in the family working or attending a course; an adult-focused scheme, which only has incentives for the adult working; and a child focused scheme, which only has incentives for the child to take and perform well in school tests. Each pilot will be evaluated using a random assignment design to assess the impact of these incentives on families, adults, and children, as well as its effect on overall poverty reduction.

CCT programes have been criticised on several grounds. The conditions can be stigmatising, or create peverse incentives. Morris et al. (2004) found lower weight gain in participants in a CCT programme in North-east Brazil. This may have been because some parents (wrongly) thought that benefits would be discontinued if children started to grow well. Incentive schemes can also be costly. In many cases, people are being paid to do what they would have done anyway. However, most of the families enrolled in CCTs would receive some form of social security transfer or tax credit in any case. The conditions, as well as giving the recipient additional incentives to invest in human capital,

help make such social security payments more acceptable to the middle classes whose taxes finance them.

Studies of CCT usually compare participation with non-participation. An important question is the relative importance of the cash transfer, disaggregated from the effect of the conditions and other components of the programme, such as nutritional supplements and health interventions. Fernald (2008) assessed outcomes with respect to the length of time families were on the Mexican programme, which is a measure of the cumulative transfer received. They find that the size of the cash transfer is associated with better child health, independently of other components. However, this study did not assess the effect of the other components.

Other evidence suggests the conditions might be effective, alongside the cash transfer. Bourguigon (2002) evaluated alternative programme designs to the Brazilian Bolsa Escola CCT using a microsimulation model. They considered a scenario with a means-tested transfer exactly as Bolsa Escola but with no condition requiring children to attend school. This simulation suggests that the condition to enrol in school in order to achieve the benefit – rather than the pure income effect from the transfer - was the primary cause of the extra demand for schooling. Of course, actual *ex-post* evaluation of the relative effect of the conditions and/or other components *versus* the cash alone in a setting more relevant to the UK would be required before firmer conclusions could be made. Conditional transfer programmes are discussed further in the Social Exclusion and Social Protection Task Groups (DH 2009).

Le Grand (2008) discusses a set of policies which he terms 'liberal paternalism'. Behavioural economics has shown that for some people, the 'default' or 'starting position' influences their behaviour. This is not necessarily irrational. There are costs to making decisions in terms of time and effort to obtain all relevant information, and people use habits and rules of thumb instead. Policy makers may be able to influence some people´s behaviour by setting the default position, while still preserving individual autonomy, or at least, the individual's perception of their autonomy. For example, supermarkets have invested considerable research into understanding how the layout of the shop influences consumer choices. This knowledge could be used to promote healthy options.

Le Grand (2008) suggests an example might be a smoking permit, so that smokers have to 'opt in' and obtain one each year in order to purchase tobacco. Sellers of tobacco would have to see a permit before any sale. The money raised would go to the NHS. Smokers would have to make an explicit decision to continue smoking for the following year, and obtaining a permit would be costly in terms of money and time. This reverses the default position from the current one where they have to make a conscious decision to opt out. One other (perhaps positive) consequence of the scheme is that it creates a *de facto* comprehensive register of smokers. If data protection issues could be resolved, this register could be used to link with primary care services and target Stop Smoking interventions, and for research. Some commentators have suggested extra conditions such as having to sign that they understand the risks, or to agree to make a co-payment for treating future smoking related diseases. One could argue that the scheme could be made more acceptable by offering participants benefits, such as subsidised nicotine patches or rewards for stopping smoking.

There may be negative consequences. It may be perceived as highly intrusive, though proponents would probably argue that it is meant to be at least inconvenient and possibly stigmatising. People may try to avoid the permit and use black markets or purchase supplies abroad. The scheme may have a high administrative and enforcement cost. The legality of the policy might be challenged. Smokers may convince themselves that their habit is socially acceptable, as they have purchased the 'right' to smoke in the same way that car tax buys the right to use the road.

4.4.3 Consumption taxes, subsidies and price controls

The third category in Sassi and Hursts' taxonomy is (consumption) taxes and price controls. These measures have a long history in public policy, for example in the form of 'sin taxes'. However, it is notable that the "Choosing health" White Paper (DH 2004) does not propose any such measures to encourage healthier behaviour. This may in part reflect the Labour government's memory of the disastrous experience of fuel tax in 2000, prompting widespread public protest, and mistrust of price controls, which are associated with perceived failure of Labour governments in the 1970s to understand and regulate markets. However, taxes and price controls have a role in prevention. The DH recently commissioned a study on price controls on alcohol. The Sheffield study (DH 2008a)

simulated the effects of different policies around pricing and promotion of alcohol on health, crime and employment in England. The study looked at the effects on the general population and specific groups: drinkers aged under 18 years, young adult binge drinkers, and harmful (heavy) drinkers. Several types of policy were appraised

- Across-the-board price increases
- A minimum price per unit of alcohol
- Banning promotions and discounting

The study derives own-and cross-price elasticities for 16 beverage categories using data from national surveys. The study finds that, on average, hazardous and harmful drinkers (combined elasticity of -0.21) are less price elastic than moderate drinkers (elasticity of -0.47).

The general pattern found by the study is that the more restrictive the policy, the greater the benefit to population health. Higher minimum prices lead to greater harm reductions, and this goes up steeply – for example, there is relatively little effect for a 20p minimum price (per unit of alcohol), but 30p, 40p, 50p and 60p have increasing effects. The report estimated that benefits for crime and employment would be large and relatively rapid.

Most policy options affect moderate drinkers in a very minor way, simply because they consume only a small amount of alcohol and also because they do not tend to buy as much of the cheap alcohol that is targeted by minimum pricing and promotion bans. Harmful drinkers buy more alcohol and also tend to choose cheap alcohol; therefore these would be most affected.

The authors estimate that both off-trade and on-trade retail sectors would increase revenue from minimum pricing or bans on off-trade discounting, as the loss in sales volume would be more than offset by increases in price (the absolute price elasticity is less than 1). The report does not discuss how a policy of alcohol price control would be implemented or monitored. There might be incentives for some retailers to try to capture greater market share by undercutting the controlled price in some way. There may also be greater incentives for smuggling and unlicensed brewing. Restrictions on alcohol sale in supermarkets in Sweden encouraged people to drive across the border to neighbouring countries or take 'booze cruises'. This indicates the need for a concerted and coordinated approach, addressing illicit supplies.

The Sheffield report does not discuss the consequences for income inequalities or income-related health inequalities. In general, the distributional consequences of price controls (rises) are likely to be worse for the poorest, who usually spend a greater proportion of their income on consumption than the better-off. As shown in Figures 5 and 6, the relationship between alcohol use/misuse, and income/social class is complex.

4.4.4 Restrictions and bans

The public health initiative most restrictive of choice is a ban. A ban on smoking in public was introduced in England on 1 July 2007. Smoking prevalence appears to have decreased by over 4% in the year following the smoking ban (West 2008). This reduction may not be entirely attributable to the ban itself, given the declining trend in smoking in previous years, but many observers suggest that the policy has been effective (CMO 2008). Other examples of public health bans in the UK include restrictions on tobacco advertising or restrictions on advertising and promotion to children of unhealthy foods and drinks. Most of the financial cost to public sector of a ban is usually in enforcement. Producers, and advertisers, will be expected to suffer losses to revenues. The losses to consumers are mainly welfare loss as the broad sweep of the policy affects those who were targets as much as those who consider their choice rational. 'Rational addicts' (Sassi and Hurst 2008) in particular might resent such restrictions. Bans may have unintended consequences. As with taxes, a ban may make unhealthy behaviour harder to observe, for example, criminalising those who were previously law abiding. A concern raised before the smoking ban was implemented was that people might switch to substitutes that were not necessarily healthier, for example, smoking more at home rather than in the pub. In fact, national survey data suggests that on average people smoke less at home in 2007 than 2006, and are less tolerant of others smoking in their home (Lader 2007). There appears to have been a shift in public attitudes against smoking and in support for the ban, though this seems strongest in managerial and intermediate social classes.

5. Service delivery of national targets

5.1 National targets

Targets are an important part of government strategy to reduce health inequalities. Public health targets have been introduced to the English NHS previously, but with little impact at a local level (Smith 2007; 2008). Hunter (2002) summarizes the weaknesses of the *Health of the Nation* strategy under six broad headings:

1. There appeared to be a lack of leadership in the national government.
2. The policy failed to address the underlying social and structural determinants of health.
3. The targets were not always credible, and were not formulated at a local level.
4. There was poor communication of the strategy beyond the health system.
5. The strategy was not sustained.
6. Partnership between agencies was not encouraged.

From 1998, the finance ministry (Her Majesty's Treasury) set the health ministry (the Department of Health) challenging strategic targets in the form of 'Public Service Agreements' (PSAs), in common with all other government departments. The 2004 PSAs for the health department are set out in HM Treasury (2004). They are based on four broad objectives, as follows:

- Improve the health of the population. By 2010 increase life expectancy at birth in England to 78.6 years for men and to 82.5 years for women.
- Improve health outcomes for people with long-term conditions.
- Improve access to services, in particular waiting times.
- Improve the patient and user experience.

These objectives are accompanied by detailed targets. The first objective has the most direct relevance for tackling health inequalities at a national level

1. *Substantially reduce mortality rates by 2010:*

   - *from heart disease and stroke and related diseases by at least 40% in people under 75,with at least a 40% reduction in the inequalities gap between the fifth of areas with the worst health and deprivation indicators and the population as a whole;*
   - *from cancer by at least 20% in people under 75,with a reduction in the inequalities gap of at least 6% between the fifth of areas with the worst health and deprivation indicators and the population as a whole; and*
   - *from suicide and undetermined injury by at least 20%.*

2. *Reduce health inequalities by 10% by 2010 as measured by infant mortality and life expectancy at birth.*

3. *Tackle the underlying determinants of ill health and health inequalities by:*

   - *reducing adult smoking rates to 21% or less by 2010, with a reduction in prevalence among routine and manual groups to 26% or less;*
   - *halting the year-on-year rise in obesity among children under 11 by 2010 in the context of a broader strategy to tackle obesity in the population as a whole; and*
   - *reducing the under-18 conception rate by 50% by 2010 as part of a broader strategy to improve sexual health.*

These targets do seem to have been designed to address several of the weaknesses of previous regimes. A distinctive feature of PSAs has been their focus on the outcomes rather than operational activities of public service delivery. They address health, intermediate causes and the structural determinants of health. The targets were accompanied by a milestone public health White Paper *Choosing Health* (DH 2004) that set out a range of initiatives for encouraging healthy lifestyles and reduce risk factors. Responsibilities for targets that require cooperation between different agencies are shared by relevant government departments. There is leadership and strategy from the DH National Support Team for Health Inequalities. The government has identified additional resources,

particularly to the most deprived local areas (the Spearhead health authorities). There is regular monitoring and feedback on progress at national and local level (DH 2008d).

Sassi (2009) identifies some weaknesses of the 2004 PSA targets. First, they are not entirely consistent. Premature mortality focuses on geographical inequalities; the smoking cessation target focuses on inequalities by social class; and the child obesity target refers to overall improvements rather than reductions in inequalities. The focus on geographical inequalities in the headline target for reducing the gap in life expectancy at birth is not necessarily the same as tackling inequalities by social class. While in Spearhead areas a higher proportion of the resident population is disadvantaged, at national level most of those who are disadvantaged live in non-Spearhead areas. The targets probably focused on geographical inequalities at least in part because measurement by social class has to rely on survey data, which takes longer to collate and analyse than ONS data and is not universal. However, many people would expect a strategy on health inequalities would aim to reduce variation by social class.

Second, the public health White Paper *Choosing Health* focused on population health but only addressed in a very cursory way the key question of how overall health improvements might lead to a narrowing of health inequalities. The targets were formulated from an assumption that trends in health would be generally positive in the overall population, and that government action could accelerate the gain in the most disadvantaged without compromising the trend to better health of other groups.

This assumption has not been entirely justified, *ex-post*. In particular, according to HSE data 1991-2005, socio-economic inequalities appear to have worsened among women to a significantly greater extent than among men, across several indicators, including life expectancy, obesity, mental health and cardiovascular disease (Sassi 2009). In some indicators, such as mental health, better-off women have improved health faster than other socio-economic groups, while in other indicators, particularly obesity and cardiovascular disease, there has been a deterioration of the health of the most disadvantaged women in absolute terms.

5.2 Weighted capitation formula

The Department of Health distributes the bulk of its funds to PCTs using what is known as ‘weighted capitation’. This principle has been in force since the celebrated report of the Resource Allocation Working Part (RAWP) in 1976 (Bevan 2008). Traditionally, the equity criterion underlying English resource allocation methods has been to secure ‘equal opportunity of access to health services for those at equal risk’. However, in recent years ministers have sought to introduce a radical new criterion for determining capitation payments, alongside the traditional one: ‘to contribute to the reduction in avoidable health inequalities’.

Equity criteria fall into two broad categories – horizontal equity (equal treatment of equals) and vertical equity (greater priority for those with greater needs). Horizontal equity is concerned mainly with seeking to equalize inputs, whilst vertical equity is concerned mainly with outcomes.

Horizontal equity informs the conventional NHS equity criterion. It reflects a concern that equals should be treated as equals, without regard to whether any differentials between groups that are not equal is appropriate. For example, most health care capitation formulae use ‘age’ as a part of the capitation formula, suggesting that policy makers believe that age is a legitimate reason for variations in expenditure. Most capitation schemes therefore adjust for age according to empirically derived variations in expenditure between age groups. This implies a judgement that those of equal age should receive equal funding (other things equal). It does not however question whether the existing distribution of resources between age groups is in line with policy intentions.

In contrast, the vertical equity criterion is based on the assumption that those in *more* ‘need’ should receive more resources. Its pursuit is therefore more challenging, because it requires a judgement to be made on how much the differential should be. Vertical equity concerns usually signal an interest in variations in outcomes rather than inputs, and the precise measure of outcome deployed is the key policy decision. Once this is specified, it becomes (in principle) possible to determine the amount required to ‘meet’ that need for each individual, and therefore the associated capitation payment.

In practice, almost all systems of capitation funding pursue a horizontal equity criterion only, seeking to allocate funds to health care purchasers so as to enable some 'standard' level of health care to be delivered, given the characteristics of the purchaser's population. This principle is reflected in the original RAWP approach, which sought to allocate the fixed National Health Service budget to geographical regions in accordance with an equity criterion of seeking to secure 'equal opportunity of access for those at equal risk'. The methods adopted by RAWP have since been superseded by more empirically based approaches. However the horizontal equity objective of 'equal funding for equal need' has continued to underpin the empirical methods used in England.

The premise underlying empirically based capitation methods is that current patterns of utilisation should form the basis for making the horizontal equity criterion operational. In common with most empirical capitation methods, the traditional English approach is therefore intrinsically conservative. The resulting formula reflects average health service responses to underlying healthcare needs, as reflected in a rich set of socio-economic and epidemiological indicators considered to be 'legitimate' drivers of expenditure. Using these methods, 'need' is effectively whatever the health service has decided on average to devote its resources to. They therefore cannot capture 'unmet' need.

In contrast to the concern with inputs (in the form of access to care), a vertical equity concern arises if the *outcomes* of current NHS activity are unacceptable from a policy perspective. In particular, the 1999 White Paper *Saving lives: our healthier nation* put in place a public health agenda with the objective of 'improving the health of everyone, especially the worst off' – that is, of improving health and reducing health inequalities. There is of course a question as to whether the NHS is the most appropriate instrument for addressing health inequalities. Indeed the SDH perspective suggests that policy areas such as income redistribution, housing, education, environment, transport, might also have considerable influence in this domain. However, the commitment to reducing health inequalities in turn resulted in a reappraisal of the capitation criterion in use in England. The Advisory Committee on Resource Allocation (ACRA) was instructed by ministers to undertake a fundamental review of its capitation methods, incorporating a new criterion for determining capitation payments: 'to contribute to the reduction in avoidable health inequalities'. This criterion represented a radical departure from that of seeking to offer equal opportunity of access and in effect sought to secure a redistribution of health.

The new criterion implied that prevailing patterns of utilization in the health service were failing to satisfy policy objectives, as reflected in health outcome variations, and reflected a concern with vertical rather than horizontal equity. Of course, the same groups might suffer from both horizontal inequity (poor access to services) and vertical inequity (poor outcomes). However, the two concepts may be quite distinct, in that inequalities in access are only one of a possible number of determinants of inequalities in outcome. It is important to distinguish between the two inequities for the purposes of developing satisfactory capitation methodology (Hauck et al 2002).

In principle, implementing appropriate funding mechanisms to address the health inequalities objective would require the resolution of the following issues:

- Identification of the disadvantaged groups at which the intervention should be directed;
- Identification of effective health care interventions designed to reduce the health inequality (this might of course include improved access for disadvantaged groups, but need in no way be limited to this intervention);
- Identification of the areas where such groups live;
- Designing a formula to direct fair funding to all localities;
- Ensuring that the resources are spent appropriately on the disadvantaged groups and the necessary interventions.

While all of these steps are challenging, identification of effective (and cost-effective) health-care interventions is a particularly time-consuming process. NICE currently undertake detailed appraisals of the effectiveness and cost-effectiveness of a very limited set of interventions each year. Assessing the impact on socio-economic groups would impose additional information requirements.

In the absence of detailed and current information of how specific interventions might reduce health inequalities, a more pragmatic approach has been adopted. The first attempt to introduce a health inequalities element into English capitation methods was in financial year 2001/02, when a 'health inequalities adjustment' comprising £130million was targeted at those health authorities which were

judged to be making the biggest contribution to health inequalities. At first glance, this seems a very small amount when viewed in the light of the £37billion distributed to health authorities on the traditional 'equity of access' criterion. However, it marked a major departure from conventional resource allocation, and the amounts involved increased year-on-year until the adjustment was abandoned when the AREA formula was introduced in 2003/04 (DH 2008b).

The distribution of this early health inequalities adjustment between areas was based mainly on the magnitude of an area's 'avoidable mortality'. This was defined simply as the number of years of life lost under the age of 75 over a three year period, where diagnosis of death was in certain broad categories deemed to be 'avoidable'. Although pragmatic, this adjustment had much to commend it, by distributing a fixed pool between areas in proportion to a justifiable indicator of the size of the health inequality problem in an area.

The adjustment implicitly makes a number of assumptions, amongst which the following are some of the most important:

- The age of 75 is a benchmark. Any variations in life expectancy above that age are ignored. All years of life lost below that age are given equal weight (so one infant's life lost is equivalent to 75 people dying at age 74).
- No consideration is given to variations in the health-related *quality* of life enjoyed. Yet areas with relatively long life expectancy might suffer from poor quality of life that is amenable to health care intervention.
- All mortality attributable to causes deemed 'avoidable' is included, all other mortality is excluded.
- The calculation is based on historical mortality in the current three year period. Yet resources should in principle be directed at people who are expected to suffer premature mortality in the future.
- It is assumed that all avoidable years of life lost are equally amenable to health care intervention. Yet it is quite possible that diminishing returns exist (each extra year added to an individual's life becomes increasingly expensive).
- The amount devoted to the health inequalities adjustment is a political judgement that has little immediate empirical justification.

Hauck et al (2002) give a mathematical formulation of how it might be possible to balance the traditional horizontal equity criterion (equity of access) with the new vertical equity criterion (reducing avoidable health inequalities). It is not possible to develop precise formulae for the required size of the health inequality adjustment. However, under some restrictive but fairly reasonable assumptions, it can be shown that the adjustment should be distributed roughly in proportion to the 'years of life lost' below some benchmark age (such as 75). That is, the original health inequality adjustment does have some face validity.

For 2009/10 and 2010/11 allocations, a new funding formula has been introduced, including the incorporation of a major new health inequalities adjustment (DH 2008). This combines 2005 life expectancy data with 2001 limiting long-term illness data to yield a measure of disability free life expectancy (DFLE), thereby capturing morbidity as well as mortality. It is applied by comparing every PCT's DFLE to a benchmark figure of 70 years. For example, a PCT with a DFLE of 60 years is given an index of 10. This index is applied to each PCT's crude population and the results normalised to the total crude population to give a health inequalities weighted population for each PCT. The major departure for the new adjustment is that Ministers decided to apply it a very large proportion of PCT allocations, amounting to 12.4% of all hospital and community health services expenditure (about £7.5billion in 2009/10), an enormous increase on the first adjustment in 2001/02. It is nevertheless worth noting that the main role of the adjustment appears in practice to have been to compensate PCTs for losses they would otherwise have incurred due to changes in the conventional resource allocation formula.

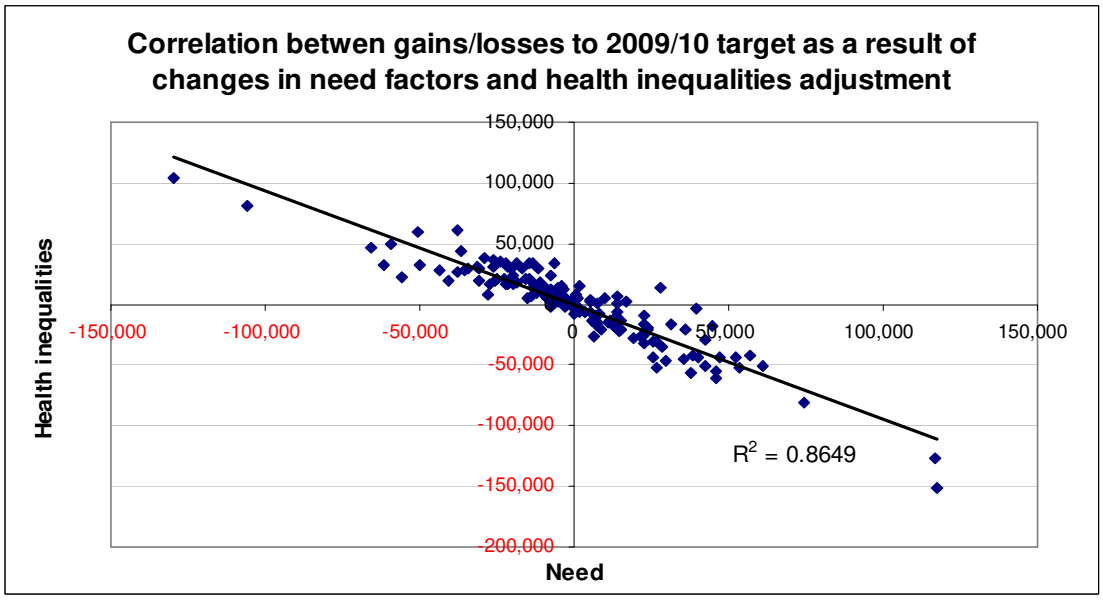

Figure 8. Correlation between gains/losses to 2009/10 target as a results of changes in need factors and health inequalities adjustment

Moreover, there is no guarantee that mere alteration of the funding formula will ensure that additional resources reach deprived populations. By definition, the vertical equity criterion requires that the health sector alters the way in which it delivers health care to those with poor health expectancy. Yet, in general, directing extra 'health inequality' resources at needy areas will not in itself lead to reduction in health inequalities. Rather, it may merely lead to a perpetuation of existing patterns of utilization in an area, albeit at a higher level than before. Additional rules or incentives, such as performance reporting or ring-fenced funding for the disadvantage groups, may be needed in the short term.

It is in any case not at all clear how PCTs should best spend their 'inequalities adjustment' funding. The most immediate instinct might be to seek to improve access to health services for disadvantaged groups. However, this may not be the most efficient, indeed it may not even be feasible if the main causes of health inequalities are mainly due to personal characteristics rather than access problems. If so, the emphasis might shift to health promotion strategies for improving the health of disadvantaged groups. Alternatively (though this may be difficult to implement) policy might consider giving privileged access to health care for disadvantaged groups. Finally, the discussion has emphasised the role of health services in addressing health inequalities, and has made only passing reference to the broader influences of social policy on inequalities. There is no reason in principle why the health inequalities adjustment should be spent on the NHS, if it is felt that other public service interventions might be more cost-effective. But this may require some radical thinking.

5.3 Service delivery at local level

The design of performance management systems is commonly analysed by principal-agent theory, a model of delegated decision-making (Gravelle and Rees 1992). A principal P authorises an agent A to take decisions on her behalf. P might represent the health ministry and A the hospital managers or GPs. If the interests (preferences) of P and A are identical, or if P has complete information about the decisions and consequences and can perfectly observe A's choices, there is no difficulty in ensuring that A acts in P's interest: no performance management system is necessary. Principal-agent theory examines situations in which these conditions do not apply.

In this context, P's objectives might be expressed by the PSA targets, and not to exceed the NHS budget. A might partly share these aims but might also like to be rewarded for greater effort towards these targets. There is likely to be information asymmetry. P may not know whether a target has been satisfied unless A provides the necessary information.

The principal will try to set a system of meaningful local targets and incentives for the agent(s), to deliver national objectives at a reasonable cost, and to obtain (truthful) information. There were very

few targets in the 'traditional' NHS beyond the requirement to break even (73). The number of targets multiplied during the 1990's, but in most cases without much pressure of reward or penalty, either financial or non-financial. Rewards in the NHS were traditionally not directly linked to individual or organisational performance, partly for lack of accurate information, and partly because this was thought to undermine public service values and peer-group monitoring systems. Political impatience with the slow progress of the NHS has led to the creation of sharper incentives, for example, with public reporting of mortality rates in some clinical areas, and the introduction of the new GP contract.

The following sections describe and evaluate NHS systems of performance management aimed at public health and inequalities for hospitals, PCTs, GPs and partnership working with local authorities.

5.3.1 NHS hospital trusts and PCTs

The DH has devised a system of 'performance ratings' for individual NHS organisations (hospitals and purchasers of public health care) that are intended to directly reflect many of the objectives at national level, as embodied in the PSAs (Smith 2008). Managers' success or failure according to these ratings is accompanied by substantial rewards and penalties, such as job loss. However, the outcomes expressed in the PSA public health targets, in terms of smoking, heart disease etc have not been translated into meaningful local targets through the medium of the performance ratings system. This might be considered a weakness of the ratings regime and divert managerial attention away from outcomes and towards service delivery, and has led to growing concern about a growing 'targets culture'.

Some responsibility for target setting has now been devolved in order to give PCTs the flexibility to pursue their local priorities. PCTs work with their strategic health authority (SHA) to agree a set of local targets. These may come from the 'vital signs' list provided by the Department of Health, or they may be developed locally. At present, none of the local indicators detailed in vital signs relates to health promotion activities. Boyce et al (2008) recommends PCTs should develop local 'vital signs' indicators that can be used to assess the impact of their behaviour change interventions.

In order to assess progress at a local level towards national targets, the Healthcare Commission's annual health check assesses whether PCTs have behaviour change initiatives in place, and measures their performance against targets on childhood obesity and smoking cessation. The Healthcare Commission also assesses the extent to which NHS acute trusts promote healthy behaviour in patients and staff through:

- providing services to help people stop smoking and have a smoke-free environment
- providing opportunities for healthy eating
- providing opportunities for physical activity
- encouraging sensible drinking
- improving mental health and well-being
- promoting sexual health.

From April 2009, a new organisation – the Care Quality Commission – will take over the roles of the three organisations that are currently responsible for regulating health care and social care in England.

5.3.2 Staff contracts.

Some PCTs are exploring a 'health gain schedule' as part of all contract specifications to make the prevention agenda part of the business of all frontline staff. This is systematic about training, brief interventions, referral pathways and performance monitoring, and should, at least, cover breast feeding, smoking, alcohol problems and obesity (DH 2008c).

5.3.3 Pharmacy contract

In the current pharmacy contract, public health is an essential service; however, its impact is limited, as it only requires pharmacists to give 'opportunistic' advice. Department of Health research found evidence that pharmacists are apprehensive and cautious regarding proactively raising issues such as smoking or weight loss with their customers (Boyce et al 2008).

5.3.4 GP contract

The new GP contract introduced strong incentives for GP practices to implement national guidelines for primary care. The QOF would be expected to have a positive impact on public health if the indicators are a cost-effective use of health care resources. Cost-effectiveness is the correct criteria (rather than simply effectiveness) because an intervention that has a marginal effect on health, but is very costly, is likely to divert resources away from other interventions that have a greater impact on public health. However, promoting greater overall use of these interventions (as the QOF does) might not necessarily reduce socio-economic inequalities in health, as shown by the Contoyannis and Forster model in Chapter 2. For example, GPs might focus attention on people most likely to respond to the intervention, who might be the highest income groups. We review here the limited evidence on how primary care guidelines might have been implemented differently across socio-economic groups since the introduction of the QOF.

Cost-effectiveness of QOF indicators

Mason, Walker et al. (2008) reviewed costs and benefits of clinical indicators with a direct therapeutic impact, e.g. "The percentage of patients with atrial fibrillation who are currently treated with anti-coagulant drug therapy or an anti-platelet drug therapy". For those indicators where some evidence on costs and benefits was available, all interventions under the 2004/05 QOF were shown to be cost-effective, whereas just one indicator in the 2006/07 analysis was not cost-effective (retinopathy screening for diabetic patients). Whether the QOF payments themselves represent potential value for money depends on baseline utilisation and change in utilisation following the introduction of the QOF payments.

The effect of the QOF on recorded prevalence of CHD and stroke by socio-economic groups in GP practices

As the QOF was implemented in all GP practices simultaneously, there are no studies that have evaluated its impact using an adequate control group. Some studies collected individual patient data from practice registers before and after implementation. These provide some indication of the effect of the QOF, but do not show causality as other coterminous factors and policies will have influenced outcomes.

One requirement of the QOF is for practices to maintain a register of patients with conditions such as diabetes, epilepsy, CHD and stroke. This appears to have been successful in improving practice record-keeping. Simpson (2006) compared two cross-sectional surveys of patients with stroke in sample of practices in Scotland before and after introduction of new contract. McGovern (2008) used a similar design and dataset to compare patients with CHD.

Both studies found large increases in the number of patients with recorded diagnosis of stroke/TIA and CHD respectively in these practices. The proportion of all registered patients with stroke/TIA increased from 1.2% to 1.8%. The recorded prevalence of CHD increased from 3.7% of patients over 16 years registered with the practices to 4.9% after the contract. There were large increases in the number of patients with recorded diagnosis of stroke/TIA and CHD in all socio-demographic groups. Most strikingly, there are a higher proportion of patients registered with stroke/TIA and CHD in the lowest socio-economic group after the QOF than before (Table 3). The authors conclude that the apparent increases in prevalence were probably the result of financial incentives for primary care practices to have accurate disease registers (with opportunities to exclude patients who refused or who were contraindicated treatment or who were too frail or refused to have clinical examinations).

The effect of the QOF on inequalities in utilisation of health care

McGovern (2008) found that the recording of data relating to the care of patients with CHD increased substantially in Scottish practices after the introduction of the GMS contract. The most dramatic increase was observed in the measurement of cholesterol levels in all patients with CHD, although this indicator was the most poorly recorded (non-prescription) item pre-contract. Of those with measured cholesterol, a lower proportion of patients had their cholesterol controlled below 5mmol/l post contract.

Table 3. Recorded prevalence of stroke/TIA and CHD in Scottish GP practices before and after introduction of contract, and comparison of prevalence of recorded disease by deprivation index

	Stroke/TIA (84)		CHD (85)	
	Pre contract	Post contract	Pre contract	Post contract
Disease recorded by the practice, as a % of all adults registered with the practice				
Prevalence	1.2%	1.8%	3.7%	4.9%
% of those recorded with the disease, by deprivation index				
Most affluent quintile	20.8%	19.4%	18.2%	18.2%
Most deprived quintile	11.7%	14.5%	13.1%	15.0%

Both pre and post-contract, the most deprived patients were more likely to receive antiplatelet/anticoagulant or ACE inhibitor therapy than the least deprived. Both pre-and post contract, the most deprived patients were less likely to receive b-blocker therapy. Pre-contract, the most deprived patients had been equally likely as the least deprived to have smoking status recorded or have blood pressure measured. However, post-contract, the most deprived patients were less likely to have smoking status recorded or have blood pressure measured, and tended to be less likely to be given anti-smoking advice.

This data is summarised in Figure 9. While any overall trends are difficult to pick out, given the diversity of these indicators, in general indicators that were pro-rich before the contract appear pro-rich after the contract. For the two indicators that were pro-poor before the contract, use of ACE inhibitors appear to remain pro-poor but provision of smoking advice has changed to be (non-significantly) pro-rich.

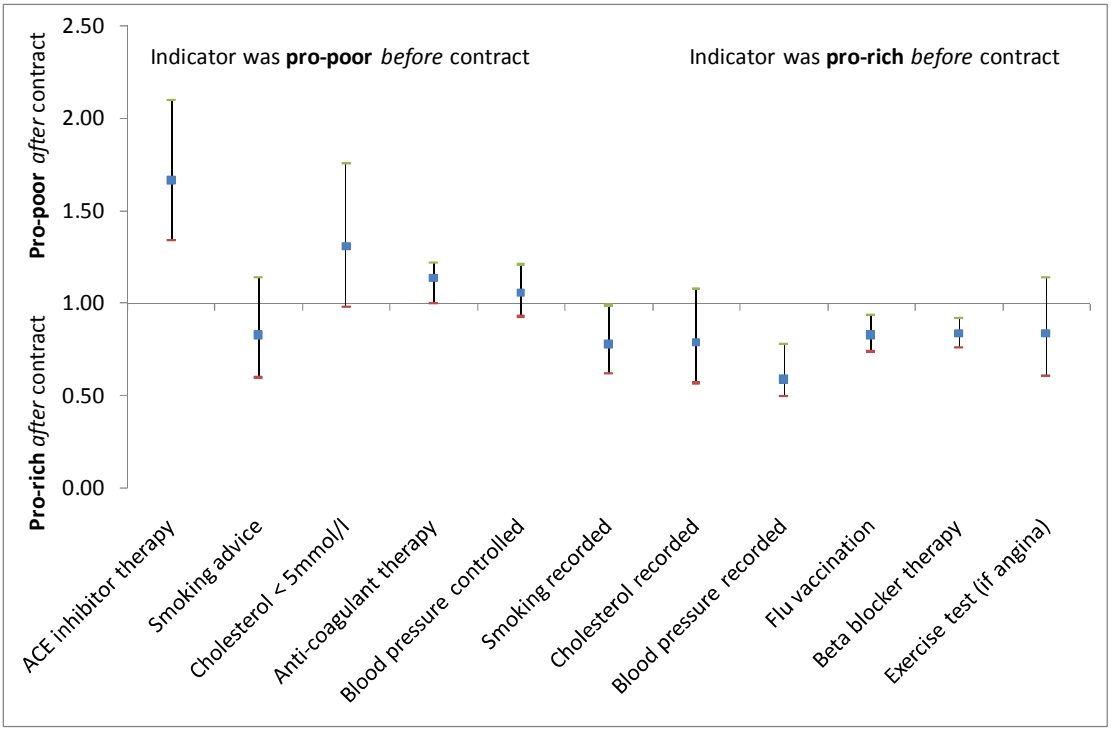

Figure 9. Degree to which delivery of primary care to patients with CHD was pro-poor or pro-rich before and after the introduction of the new GP contract. Source: McGovern 2008

Notes: ORs and 95% confidence intervals comparing frequency of recorded quality indicator data among patients with a recording of CHD living in most deprived neighbourhoods versus those living in the most affluent neighbourhoods. Vertical axis shows OR after the introduction of the new GP contract. Indicators are ordered on the horizontal axis according to whether pro-rich or pro-poor before the introduction of the contract. ORs are adjusted for sex, age, number of co-morbidities, and practice size

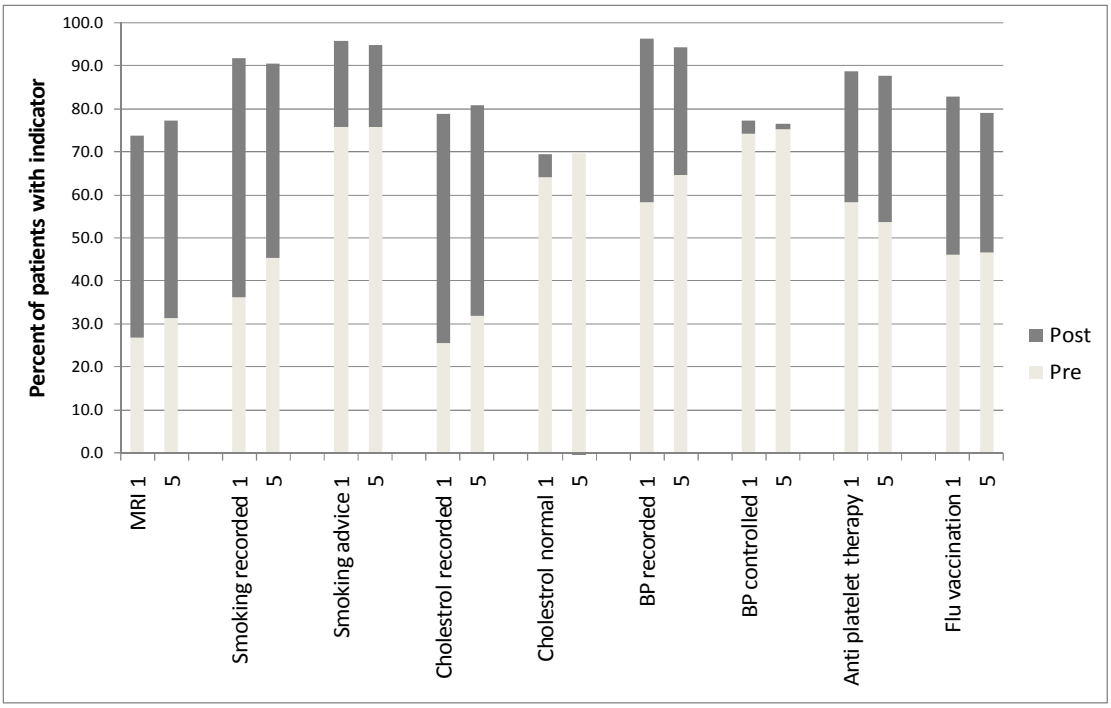

Figure 10. Recorded Quality Indicator Data Among Most Affluent (Group 1) and Most Deprived (Group 5) With a Recording of AnyStroke/TIA, before and after QOF. Source: Simpson 2006

Notes. Vertical axis shows the frequency of the recorded indicator in absolute percentage points before and after the introduction of the new GP contract by social class.

Simpson et al (2006) presented data for patients with stroke in a little more detail, with similar trends shown in patients with CHD. Figure 10 shows the change in selected indicators in percentage point terms. This is relevant because GP income is directly related to changes in the percentage points recorded for each indicator (up to a maximum percentage point achievement of 70% or in some cases 50%). Recording of quality indicators and interventions increased sharply in all socioeconomic groups. However, there was a trend for those groups with the lowest use of primary care services before the QOF to show the greatest change in use. Most indicators that were pro-poor before the contract were still pro-poor after the contract but showed the greatest increase in use in the most affluent group. Of the indicators that were pro-rich before the contract, smoking advice increased in similar percentage-point terms but anti-platelet therapy increased most sharply in the most disadvantaged group.

Millett (2007) evaluated smoking recording, cessation advice and prevalence of smoking in individual patients with diabetes in Wandsworth, London, before and after introduction of new contract (Table 4). Recording of smoking status in 15 months previous to survey: 67.6% before, 86.7% after, no evidence of difference by SES; provision of smoking cessation advice: 48.0% before, 83.5% after, no evidence of difference by SES. Data shows a trend for a negative correlation between SES status and smoking prevalence, though this is not statistically significant. There is a reduction in smoking prevalence between 2003 and 2005, and this reduction seems similar for all SES groups. Inequalities across SES groups in health behaviour do not seem to have been improved or worsened.

Table 4. Prevalence of smoking between 2003 and 2005 in patients with diabetes. Millett et al. 2007 (86)

SES quintile (2004 Index of Multiple Deprivation)	*2003*	*2005*	*Change*	*Adjusted OR*
1 Most affluent postcode	15.8	12.3	-3.5	1
2	21.3	14.8	-6.5	0.84 (0.5-1.42)
3	20.8	17.1	-3.7	1.47 (0.89-2.44)
4	21.3	17.8	-3.5	1.46 (0.87-2.45)
5 Most deprived postcode	22.2	19.5	-2.7	1.39 (0.81-2.38)
Average	20.0	16.2	-3.8	

There are some aspects of the QOF design that address health inequalities. The QOF rewards GP practices for completing a task for a certain percentage of eligible patients on their register. This means practices in higher-prevalence areas have to work harder than in low-prevalence areas for the same percentage point achievement. To reflect this, QOF payments are adjusted for prevalence relative to the national average (DH 2006).

In other ways, the QOF rewards average health improvement but does little to encourage GPs to address health inequalities (Boyce et al 2008). First, practices are responsible for identifying the patients on their list with CHD and other conditions. As the QOF rewards achievement in proportion to interventions given to those on the practice disease register, it does not incentivise practices to spend additional resources actively seeking out the marginally hardest-to-reach patients with the condition and include them on the register. These may be those in lower socio-economic groups. PCTs are responsible for auditing practice registers (DH 2006). The apparent low proportion of patients in lower socio-economic groups on the Scottish practice register with of CHD and stroke might be investigated further.

Second, practices can exclude patients on the disease register from QOF reporting. This may disadvantage those who are not responsive to GP requests to attend the practice, again who might be more likely to be from disadvantaged groups.

Third, the QOF rewards GP practices for completing a task for a certain percentage of eligible patients on their register. This means practices in higher-prevalence areas have to work harder than in low-prevalence areas for the same percentage point reward. Practices in deprived areas would be expected to have higher prevalence of diseases such as CHD.

Finally, practices can receive maximum payments for 70% or in some cases 50% achievement. However, it may be that more patients from lower socio-economic groups are among the most difficult-to-reach 30% of patients.

The evidence reviewed in this paper suggests that there is no clear overall pro-rich or pro-poor bias in primary care, although there might be deficits in specific areas such as GP practice opening times in more deprived areas (Ashworth et al 2007). Nevertheless, some primary care interventions may be effective and cost-effective in reducing health inequalities. Primary care resources could be targeted more directly at the disadvantaged at various levels – at the lowest income patients within each GP practice, or at GP practices in the most deprived areas within each PCT. As this paper has noted, there is currently no mechanism to ensure that resources are targeted within PCTs at reducing health inequalities. It would probably be inappropriate for GPs to target low income patients within a practice for special attention. However, practices could be encouraged to reach the more difficult to serve patients if maximum QOF payments were made for 100% of achievement with tighter exceptions.

5.3.5 Incentives for GPs to undertake health promotion

The QOF and health promotion

There is little incentive within the Quality and Outcomes Framework (QOF) to encourage GPs to incorporate health promotion (Boyce et al 2008). Most indicators included in the framework relate to clinical care, but could be adapted to go further and include health promotion. For example, one indicator asks GPs to measure the cholesterol level of patients with coronary heart disease, but there are no indicators regarding the promotion of healthier diets among these patients. The only QOF indicator relating to health promotion is 'offering smoking cessation advice to patients with coronary heart disease, stroke, diabetes, chronic obstructive pulmonary disease (COPD), and asthma'. Furthermore, this target rewards GPs for offering smoking cessation advice but is not linked to the number of patients who take up that advice or successfully act on it.

Practice-based commissioning

PBC has been slow to take off. Practices have commissioned few services, and few of these have been on preventative care. The financial incentives encourage GPs to undertake services in-house, rather than buy-in. But for NHS as a whole may be more cost-effective to purchase specialist services aimed at prevention and behaviour change. GPs may not have time or skill to either provide services

directly or search for an appropriate external provider. Boyce et al 2008 concludes that 'while PBC provides an opportunity for GPs to commission behaviour change initiatives for their practice populations, it does not provide incentives to do so'.

5.4 Partnership working

Child obesity

NAO evaluated the effectiveness of joint working arrangements between departments and agencies towards the childhood obesity target, shared between the Department of Health (DH), the Department for Education and Skills (DfES) and the Department for Culture, Media and Sport (DCMS). The three Departments are coordinating their action at a national level.

Key programmes:

- School meals
- School sport
- Healthy schools programme
- Improve childrens play provision
- Public health anti-obesity campaign, aimed at children and families

Programmes targeted at children in general, (rather than obese children, or specific SES-groups), though children most at risk may need special help.

Methods for evaluation include:

- Performance targets eg hours of school sport per week
- Qualitative self evaluation by schools
- Standards for school meals (eg 60p /day) (rather than performance management targets)
- Baseline data (height and weight of children) is to be collected by PCTs. The prevalence of child obesity by PCT (and risk factors) are not well understood.

Arrangements for joint working between schools and health care sector at local level include pooled funding, and setting up 'Childrens Trusts' to coordinate action

Partnership working between PCTs and local authorities

Local strategic partnerships (LSPs) extend the responsibility for local targets from the PCT to other partners. The development of local area agreements (LAAs) provides an opportunity for PCTs, local government and other partners to identify local health priorities and build policies to tackle wider determinants of health into the agreement.

Boyce et al (2008) identifies some weaknesses of partnership target-setting arrangements. Some population groups or health issues are not adequately addressed by the LAA indicators. Obesity targets, for example, concern only children, whereas an integrated approach requires that there should also be targets for adults (who influence children's eating and exercise habits). Another significant omission concerns targets related to alcohol. Currently, the indicator measures the extent to which alcohol consumption is a problem. It is a proxy measure of harm, constructed from national attributable fractions. For example, 33% of all epilepsy admissions are assumed to be alcohol related admissions. Boyce et al recommends targets that measure how effective local areas are in promoting responsible drinking.

Private health service providers

DH is piloting new models of integrated care such as polyclinics, salaried GPs, and private-sector primary care co-located with health and well being services. However, commissioners need to ensure that incentives to invest in healthier lifestyles and behaviour change interventions are built into contracts with new providers of primary and community services (Boyce et al 2008). Where short-term reductions in hospital utilisation can be delivered, there is little doubt that organisations with capitated budgets will seek to prevent behaviours that result in such deteriorations in health status. However,

some of the impacts of behaviour change will not be evident (at least in monetary terms) as quickly, and therefore may require a different incentive structure.

5.5 Inequalities in the delivery of primary and hospital health-care

Socio-economic inequality in health may at least in part arise because of unjustifiable inequality (inequity) in the delivery of healthcare. Even if there is no systematic pro-rich bias in healthcare, if the NHS is to play a greater role in overcoming socio-economic inequalities in health, services may need to be prioritised more specifically in favour of disadvantaged groups, rather than offered according to clinical need.

Analysis of equities in the use of healthcare can be difficult because measures of use ought to be adjusted for individual need and possibly unmet need. Morris et al (2005) linked national survey data (HSE) for 1998-2000 to small-area characteristics including indices of health (standardised mortality ratio) and supply (access to GPs, outpatient waiting times and distance to hospitals). The study found people with lower income made significantly greater use of primary care services but less use of hospital inpatient and outpatient services, after adjusting for need and supply variables. There was no clear independent trend for other indicators of socio-economic status.

Ashworth et al (2007) investigated the quality of primary care between areas with different levels of deprivation, based on national data from the UK Quality and Outcomes Framework. This study design does not show how health inequalities have been influenced by the QOF because no pre-QOF or control group was compared, but it does show whether delivery of GP services is currently pro-rich or pro-poor.

Differences in total QOF scores between practices in the most and the least deprived areas were small in the first year of QOF (2004) and smaller still in the second year. Larger shortfalls in the achievement of a few specific QOF indicators in deprived areas were identified, relating to practice opening hours, treatment of epileptics, treatment of people with serious mental illness, referral rates for investigation to secondary care and cervical smear testing. Some of these differences were attributed to the organisation of the practice. The authors suggest that well-organised primary care can largely compensate for substantial social disadvantage.

6. Evaluation and priority setting for health and health inequalities in England

This chapter reviews economic approaches to evaluation and priority setting of health related programmes in England. Drummond et al (2006) summarise the main challenges of evaluating public health and social programmes as:

- Attributing outcomes to interventions
- Measuring and valuing outcomes
- Incorporating equity considerations
- Identifying intersectoral costs and consequences

6.1 A normative framework for priority setting

Approaches to priority setting can be broadly categorised into efficiency-based or equity-based. Efficiency-based criteria are concerned with gains, rather than final outcomes (Tsuchiya and Dolan 2008). In welfare economics, these gains are valued by the individual. In 'nonwelfare' perspectives, the gains are (usually) valued by a social planner. Efficiency is increased in a Pareto sense if some individuals gain and no-one loses as a result of some change. Efficiency is increased in a slightly less restrictive Hicks-Kaldor framework if the beneficiaries could compensate the losers. Cost-effectiveness analysis (CEA) defines a programme to have increased allocative efficiency if the incremental gain in health, as valued by the social planner, exceed the incremental costs:

Net monetary benefit = $\lambda x B - C \geq 0$
where λ is about £20,000 / QALY in England (Rawlins and Culyer 2004)

This framework is illustrated in Figure 11. There are two individuals or homogenous groups A and B. Initial distribution of health is at point N0 (E_A0, E_B0), and total health is N0 = H0 = $E_A0 + E_B0$. The 45° line shows where health is equally distributed between the groups. Initially, health is unequally distributed at N0. CEA considers all outcomes N1, N2 and N3 to be an improvement over N0, and is indifferent between them.

Distribution of gains and losses in different groups is not relevant to health-maximising CEA and not measured in many evaluations. In N1, inequalities in health have widened. In N2, both groups have gained, but inequalities in health remain. In N3, population health has improved compared with N0 and inequalities have narrowed. N3 is not a Pareto improvement, however. Health has improved in the group with lowest initial health but worsened in the group with the greatest initial health.

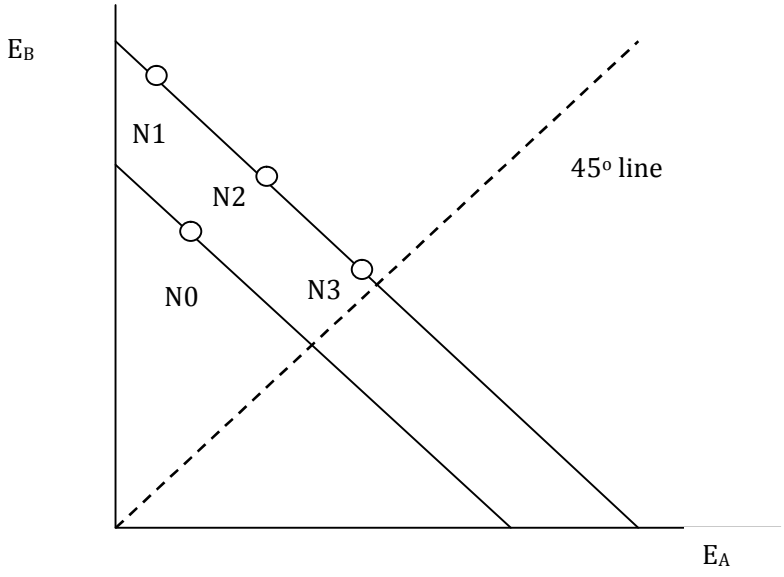

Figure 11. Illustration of gains-based egalitarianism (Tsuchiya and Dolan 2008)

Equity-based criteria involve a concern for the distribution of either gains or final outcomes. These attributes are normative, that is, they may reflect preferences of the general public or policy makers for less inequality and/or priority to be given to certain groups who are felt to be disadvantaged. These attributes might be defined *a-priori* according to an overarching moral/ethical philosophy. For example, Williams (1997) suggests priority should be given in inverse relation to life expectancy (the 'fair innings'). This criterion would give higher priority to the young and lower socio-economic groups (with lower than average life expectancy at birth). Sen (2002) suggests equity is multidimensional and should not be defined too narrowly. Dolan aims to estimate preferences for equality and for certain attributes empirically (of the general public or other sections of society eg clinicians). In Figure 12, people are indifferent to outcome N3 and N0. The distance N3 – N4 represents the total population health people are willing to give up to realise greater equality in outcomes.

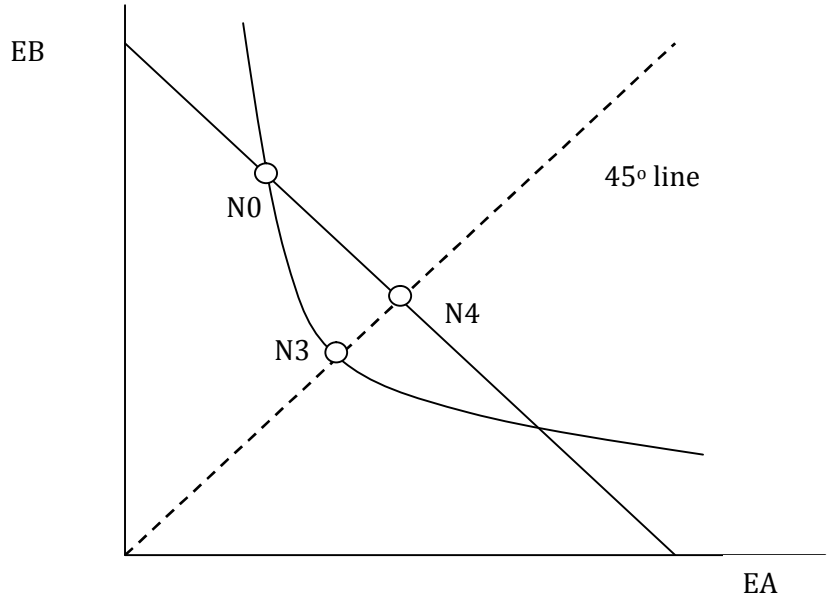

Figure 12. Illustration of social welfare function (Tsuchiya and Dolan 2008)

6.1.1 Estimating equity weights

The Dolan et al (2008) report was commissioned by NICE to investigate whether it was feasible to estimate equity weights to be used in cost-effectiveness analysis. Dolan represents the social welfare function (SWF) parametrically as

$$W = [aE_a^{-r} + (1-a)E_b^{-r}]^{(-1/r)}$$

where r represents the aversion to overall inequality. The slope a/(1-a) represents the weight that should be given to a gain in group A relative to group B. CEA can be represented as values of a= 0.5 and r = 1, that is, a linear SWF with a slope of -1.

Dolan considered the following attributes to be potentially of interest to the general public and policy makers, while recognising that these attributes are not necessarily independent (see box)

Attributes and values considered by Dolan (95) for equity weights	
Attribute	*Value*
Social class	(income, employment status)
Dependents	(number, whether children or elderly)
Age	(children, younger adults, elderly)
Low quality of life	
Length of time with the illness	
Cause of condition	(Personal responsibility, NHS negligence)
Rarity of condition	(vary rare, extremely rare)

6.1.2 Results of the Dolan study

Dolan et al (2008) presents the results in terms of statistically significant differences in the key parameters but as the report acknowledges, at a practical level, what ultimately matters is whether the results would make a difference in terms of the incremental cost-effectiveness of an intervention. This is a complex issue that is beyond the aims of the Dolan study, and thus the report does not (fully) address this.

Dolan estimated parameters for age, quality of life without treatment, responsibility and rarity of condition, and the aversion to inequality. Dependents, length of time with the condition and social class were dropped. *"The main reason for dropping these three attributes is that the complexity associated with them means that they are unlikely to be used in macro level decision-making. For example, it is not the length of time in general that is the issue, but the distinction between length of time since acquiring a condition and since being placed on a waiting list. It is not whether or not one has dependents that matters, but whether or not one is the sole responsible adult for a small child.*

Social class confounds several considerations such as prevalence of ill health, life expectancy, lifestyle choices, social behavioural norms, social status, financial affluence (and access to private health care) etc. Each of these considerations is linked to social class through very complex pathways."

However, Dolan infers social class weights attributable to inequalities in life expectancy at birth. People in social class 1 are expected to live almost 7 years longer than those in social class V. Based on this, Dolan provides indicative estimates how λ would vary from the 'average' (that is, £20,000 per QALY) for the following class-specific thresholds: £17,530 for I; £18,450 for II; £19,130 for IIIN; £20,670 for IIIM, £21,160 for IV); and £23,120 for V. These figures are indicative, and could change for a variety of reasons including where: illness is incorporated, social mobility is incorporated, and the causes of the different life expectancies (lifestyle etc.) are fully accounted for.

The Dolan approach asked the general public for their normative preferences for prioritising certain attributes. Dolan presents these attributes as dimensions of 'equity'. However, in some cases, the distinction between 'equity' and 'efficiency' is not clear. Some of the attributes valued by Dolan could reflect concerns about process, the quality of healthcare, or other objectives that are not completely captured by QALYs. If there are external benefits that are not measured by the QALY, then the threshold 'cost per QALY' may not reflect the correct marginal benefit of treatment to society.

Dolan found the general public gave significantly higher weight correcting NHS error. However, the PSA already includes performance management indicators for clinical error or hospital-acquired infection. It may be more effective to ensure that NHS organisations place a high priority on reducing these risks upstream than allocating more resources to treatments caused by negligence.

Dolan found a (non-significant) lower preference for treating illness caused by lifestyle (other factors equal). It would be difficult for decision makers to make judgements when setting macro-level priorities about the degree of personal responsibility inherent in a condition. It also raises a question about whether a different weight (higher or lower) might be placed on an upstream preventative measure that changed behaviour or reduced the harm caused by such behaviour.

Dolan found a significantly greater weight for extremely rare conditions. This may reflect concern for 'ultra-orphan drugs'; high cost treatments that require specialist facilities, and research that would be unlikely to be funded by normal channels. NICE has indicated that an ICER of £200,000 to £300,000 would typically be needed to fund these treatments (NICE 2008a), considerably greater than the preference-based weight estimated by Dolan of £23,760 per QALY.

6.1.3 Equity in priority setting

NICE sets guidelines for the NHS in England. NICE's objectives are:

- Efficiency : to recommend effective and cost-effective NHS interventions
- Equity: to secure consistent clinical practice for all NHS patients and eliminate differential access to NHS treatments in different parts of the country. This is a concept of 'horizontal equity' (Chapter 5).

Under current NICE guidance, equal weight is given to QALY gains in all members of the population under consideration for a new treatment. The rule is sometimes termed 'a QALY is a QALY' Rawlins and Culyer 2004). This is consistent with an aim of horizontal equity of gains (or 'capacity to benefit'), but not necessarily of outcomes.

NICE tends to only differentiate subgroups within a population according to clinical parameters such as severity of the condition, and where there is a plausible reason to expect this changes the expected effect of the therapy. Current NICE guidance (NICE 2008b) on sub-groups states:

"For many technologies, the capacity to benefit from treatment will differ for patients with differing characteristics. This should be explored as part of the reference-case analysis by the provision of estimates of clinical and cost effectiveness separately for each relevant subgroup of patients. The characteristics of patients in the subgroup should be clearly defined and should preferably be identified on the basis of an a priori expectation of differential clinical or cost effectiveness due to known, biologically plausible mechanisms, social characteristics or other clearly justified factors".

If NICE were to stratify a population by age (for example), some therapies would appear to differ in their cost-effectiveness by age groups, even if there were no reason to expect age to change the effectiveness of the therapy. *A priori*, a particular treatment might be more or less cost-effective in older people, depending on the context. Other things equal, saving the life of an older person will deliver less in terms of life years gained than a younger person. On the other hand, health service costs over the remaining lifetime may be lower in older people.

In most cases, NICE does not consider age and other social characteristics. Indeed, NICE was specifically set up to end differential treatment for patients with similar clinical conditions in different PCTs ('postcode prescribing'). However, public health interventions often target certain social groups with the aim of reducing socio-economic health inequalities (eg see Bambra 2009).

As described in Chapter 5, introducing a concern for outcomes or 'vertical equity' into NHS priority setting implies differential access to treatment, with greater resources or targeted interventions for socially disadvantaged groups. Income-related health inequalities might be addressed by national priority setting at a number of levels:

- *Increasing access of disadvantaged groups to existing NHS treatments or guidelines.* These measures deal with how current guidelines are implemented. Vertical equity implies greater effort should be made towards promoting guidelines in socially disadvantaged groups with a given clinical condition or risk. This may be most acceptable for preventative and health promotion measures such as Stop Smoking.
- *Ensuring that new treatments for conditions that disproportionately benefit socially disadvantaged groups are given higher priority.* This might be made operational by using equity weights such as estimated by Dolan to lower the 'cost-effectiveness threshold' for, say, a new cholesterol lowering drug if gains were distributed in favour of low-income groups.
- *Restricting (or targetting) some effective but relatively costly interventions to socially disadvantaged groups.* This would suggest dividing the population into social groups and evaluating the benefits and costs separately for each group, possibly with a lower cost-effectiveness threshold for disadvantaged groups. This would also probably only be acceptable for preventative interventions.

Shiell (2009) notes that equity-weighting of 'downstream' interventions of the kind usually considered by NICE is unlikely to have much effect on SES-related health inequalities at the macro level. Increasing the threshold for accepting a downstream intervention on equity grounds might pull more

resources towards this sector of health-care and possibly away from upstream programmes. In this respect, the selection of topics referred to (or chosen by) NICE might have a more profound effect on health inequalities than equity-weighting within a topic. NICE can only review a limited number of topics per year, and more appraisals of preventative or social programmes, including perhaps outside of the traditional healthcare remit, might mean fewer appraisals could be conducted of secondary care. NICE may need to be more explicit about how topics are short-listed and selected and whether the choice of topic is aiming to address overall health or health inequalities (or both).

It is not clear if the weights estimated by Dolan should differ for evaluations of preventative interventions compared with palliative or curative treatments. The finding that the public are less willing to fund interventions where people have some personal responsibility for their condition presumably only applies to 'downstream' treatments for that condition. The public might be more willing to fund 'upstream' interventions that helped change lifestyles to prevent such conditions. This might be an area of further research.

6.2 The 'societal' perspective

The traditional cost-effectiveness approach aims to maximise health gains (with or without weighting for equity), given the resources available to the NHS. This restricted perspective has been found to adequately capture the most important costs and benefits in most appraisals of health-care technologies undertaken by NICE. A wider perspective may be required for priority setting for public health and social interventions. Depending on the context, one or more of the following considerations might be relevant:

- Size of the programme
- Wider costs and benefits
- Patient choice and autonomy
- Political considerations

CEA does not usually take account of the size of the programme. A policy is considered cost-effective and should be adopted by the health service if the additional cost per quality-adjusted life year (QALY) gained is less than a given threshold, currently £20,000. This threshold represents value to society of an additional QALY, and is the opportunity cost of (usually unidentified) other programmes which have to be cut in order to implement the new programme. CEA implicitly assumes that the new programme will be of modest size, such that it can be implemented with only marginal changes elsewhere. However, NICE's new responsibility for public health may lead to recommendations that imply a more radical redistribution of funding within the NHS away from secondary care towards more upstream interventions and investigation, if not for individual appraisals, then perhaps cumulatively over time. In this case, NICE would probably need to be more explicit about which programmes or areas should be divested as well as where funding should be directed.

Many of the costs and the consequences of public health programmes fall outside the NHS: on other public services, on individuals and households, and/or on the economy as a whole. Action on social determinants of health, such as the child obesity target, involves several other government ministries. Public health campaigns call on households to make greater preventative effort and expenditure, for example to consume fruit and vegetables. Consumption taxes on alcohol for example may have widespread effects on other markets, retailers and manufacturers, and government revenues.

Claxton et al (2007) considers amending the decision rules where the intervention is delivered in partnership with other agencies, which might be governmental or private, who each have somewhat different objectives and fixed budgets. In this case, the return to health of an additional pound spent needs to consider the opportunity cost foregone by other sectors (eg education, sport).

This approach assumes that each actor has a fixed budget, and therefore the opportunity cost foregone by other sectors can be captured by a simple parameter. In the case of education, this might represent the value to society of an additional GCSE. The framework can be used to assess whether the policy delivers a net benefit to society, and to estimate the size of any inter-sectoral transfers that may be needed to implement the policy.

The assumption that each actor has a fixed (exogenous) budget appears too strong where there are significant costs or benefits (other than health) falling on private individuals, households and firms. This would need some estimate of the 'opportunity cost' of consumption foregone. There is however some theoretical work in this area (Claxton et al (forthcoming)).

CEA is justified by a fundamental assumption: there is no need to consider whether the 'market' could provide a better (ie more efficient) solution. This seems entirely appropriate when evaluating alternative health-care policies in the UK and other countries where there is broad agreement that government should provide comprehensive health-care to all citizens according to their clinical need, regardless of their financial status, and funded from taxation or compulsory health insurance. Generally speaking, though perhaps more weakly, this also applies to education and social services, justifying the inter-sectoral adaptation to CEA proposed by Claxton et al (2007).

In CEA, patient preferences do not have any formal role in the the evaluation of health and cost-effectiveness at a macro level, though NICE 'takes these into account' during the appraisal process. Public preferences are only considered to the extent that they are used by the social planner to value health states. Patient preferences have somewhat more weight in individual doctor-patient relationships. In the NHS, people have the right to refuse treatment, and some choice of GP or hospital, but are usually only given the opportunity to choose from a limited range of alternative health-care options, particularly when the costs are very different.

The CEA perspective seems broadly appropriate for people, that is, 'patients', who have actively sought medical help for their condition. This was the perspective taken by recent NICE guidance on obesity (2006). The interventions compared in this appraisal were generally of a medical nature or administered by physicians.

Where there are thought to be wider costs and consequences outside the public sector, then it ought not to be taken for granted that an intervention will provide the most efficient, or even the most equitable solution, when these broader issues are taken into account. Public health policies such as price controls and consumption taxes are not restricted to ill people, or even those at high risk of illness, but are aimed at (or indirectly affect) a much wider population.

The welfarist normative position is that households and individuals are the best judges of whether their level of economic activity is optimum, unless there is evidence to the contrary. In the welfarist perspective, the burden of proof is on the government to show that there is too little (or too much) of a particular activity in society and some intervention is justified.

Whether or not one accepts this argument, it seems appropriate that an appraisal of a public health or social policy should consider whether there are likely to be important wider costs and consequences. This case might be supported by one or more of the 'market failures' outlined in Chapter 4 (such as externalities). There are many examples in the public health literature that have taken this approach, for example, the Chief Medical Officer's report on alcohol misuse and its influence on crime and other social hazards (CMO 2008). The case might be based on evidence that government intervention would be likely to reverse an important inequity. Even if reducing inequities in health is the primary motivation, it should not be taken for granted that public health or SDH policies will achieve this, or that the policy will not generate inequalities elsewhere. For example, consumption taxes may be regressive.

The effect on personal autonomy may have a greater weight in the public health and social policy context than (macro-level) decisions within health-care. If there is a trade-off between social justice (health or health equity) and individual liberty, then public health decision makers should devise ways to 'push back the point at which the constraint on individual liberty actually bites' (Shiell 2009). The perceived effect on personal autonomy may influence the degree to which policies are effective and/or cost-effective. It may be cost-effective to the health service for people to stop smoking, but to be implemented, any policy must convince individuals who smoke that the costs outweigh the benefits. Bans of course may be more effective than other measures: smoking prevalence appears to have decreased by over 4% in the year following the smoking ban (West 2008). This implies that evaluations of policies should take account of the potential welfare loss from restricting choice, and weigh this against the harm from the behaviour and the expectation of benefit.

Economic evaluation usually aims to identify the 'efficient' allocation of resources, whether efficiency is defined by CEA as maximising health (or equity-weighted health) given a fixed budget, or by Cost-Benefit Analysis, where broader costs are identified and broader benefits are valued using willingness–to-pay techniques. These models rarely consider barriers to implementation or political trade-offs. Tsuchiya and Dolan (2007) found that clinicians showed less willingness to sacrifice average health for equality than the general public. This might reflect clinicians' awareness that devoting resources for the benefit of the worse off might erode long-term support for universal and comprehensive health care. Shiell highlights the political barriers where corporate interests are threatened, which might occur if resources are to be transferred from secondary care to health promotion (Shiell 2006). Economic models can in principle explore public opinion and political will with concepts such as median voter and competing interest groups, though there is as yet little applied work in public health (Goddard et al 2006).

6.3 Incorporating multiple criteria into decision making

An incremental 'cost-per-QALY' within the acceptable threshold indicates that the preventative policy is at least as cost-effective as treatments that are funded by the NHS. This, presumably, is a necessary and minimum criterion for funding any proposed preventative policy from the NHS budget. Implicitly, an economic analysis that calculates a cost-per-QALY assumes that the proposed policy would displace other, less cost-effective interventions from the overall health service budget. This appears appropriate for tackling conditions where current trends are predicting considerable future costs on the health service and other public expenditure, and a high burden of morbidity and mortality. For example, the NICE guidelines on obesity (NICE 2006) started from the point of view that the health service should give greater priority to prevention, given the scenarios on future health service expenditures and public health attributed to obesity outlined in the Wanless Report (Wanless 2002).

However, Wanless also stressed that the impact that could be made by the health service alone was limited (Wanless 2007). CEA is less relevant where policy changes require action from agencies outside the health service and have consequences on the wider economy. Cost-benefit analysis (CBA) is usually recommended to quantify gains and costs to society, where there are effects on the wider economy. This might take account of how the policy might offset some of the 'market failures' discussed in Chapter 4, such as externalities. However, CBA can be complex, especially obtaining estimates of willingness-to-pay. WTP estimates carried out in one context cannot usually be carried over and applied to a different evaluation. The analysis of equity requires modelling differential responses by subgroup, again multiplying complexity. The cost and complexity of CBA implies that it should be carried out, if at all, to evaluate broad, large scale interventions, such as alcohol pricing policy, when there are thought to be substantial effects on the wider economy. Other approaches to priority setting in health care have been proposed where there are multiple objectives, particularly for the appraisal of local initiatives. These include PBMA (Miltton et al 2004) , population cost-impact assessment (Heller et al 2006), and health impact assessment.

6.4 Health Impact Assessment

There has been increasing policy attention in the UK and elsewhere to the broad social determinants of health. The WHO Commission for SDH recommended HIA as a means of assessing the effect on health of government policy.

Most authors refer to the Gothenburg statement for a definition of HIA (WHO 1999):

"Health Impact Assessment is a combination of procedures, methods and tools that systematically judges the potential, and sometimes unintended, effects of a policy, plan, program or project on the health of a population, and the distribution of those effects within the population. HIA identifies appropriate actions to manage those effects."

HIA are intended to inform a decision, rather than evaluate the effectiveness of an existing policy. However, there is no commonly agreed framework, such as CBA or CEA, and most could be described as 'cost-consequences analyses'. Many HIA are conducted in a similar manner to an environmental risk assessment, where health impacts will be identified and described in a qualitative or quantitative way, but no assessment is made as to whether overall benefits exceed costs.

Kemm 2008 suggests that a HIA might include the following components:

- Describing the baseline (health and other characteristics of population before policy and/or without policy)
- Logic diagrams (Causal paths)
- Predicting change in intermediate factors (between policy and final outcomes)
- Exposure and dose response curves (showing relationship between risk factor and health behaviour or health)
- Modelling (of qualitative impact, under expected and extreme scenarios)
- Distribution of impacts

Qualitative or quantitative assessment of impacts might come from one or more of these sources:

- Participative assessment
- Key informants (expert elicitation)
- Literature searches

A review of HIA in England found that studies rarely used a systematic method to identify appropriate policies for conducting a HIA (YHEC 2006). It is now a legal requirement to conduct a screening test for health impacts. Ministers in England must answer three screening questions related to effects on health services, health determinants, and risk factors related to lifestyle to establish whether a full HIA is needed, before undertaking important policies (see box). Salay (2008) recommends the English model of HIA as an example for other jurisdictions, particularly for EU transport and agriculture policy.

The screening tests for England (DH 2007b) are:

Are the potential positive and/or negative health and well-being impacts likely to affect specific sub groups disproportionately compared with the whole population?

Are the potential positive and/or negative health and well-being effects likely to cause changes in contacts with health and/or care services, quality of life, disability or death rates?

Are there likely to be public or community concerns about potential health impacts of this policy change?

Mindell et al (2008) found that the methodologies used for HIA are tending to become more similar over time. A major weakness of most HIA frameworks is they do not quantify health impacts nor discuss their uncertainty. HIA studies varied in their focus on health inequalities, and whether disadvantaged groups should be identified at the start or during the course of the HIA process.

HIA appears to be a pragmatic method of estimating the health impact of policies that cut across several government departments, and where it might not be feasible to conduct a full CBA.

6.5 Attribution of outcomes

RCTs and HTAs do not usually evaluate the effect of treatment according to attributes such as social class. This information would be required to understand the effect on equality. While RCTs are still considered the most internally valid method of estimating the average treatment effect in public health interventions (Drummond 2006), experiments can be difficult to rigorously implement outside a controlled 'clinical' setting, and the results can be difficult to generalise. There are some examples, particularly from developing countries eg (Lagarde et al (2007) and Kremer and Holla (2008)), perhaps an indication that aid donor agencies and private funders are more concerned about evaluation and cost-effectiveness than governments.

One of the challenges for the evaluation of public health interventions is that participants have, to varying degrees, the ability to choose their treatment. In 'randomised' studies this non-compliance does not necessarily bias the trial, but can make results difficult to generalise. In observational studies, some method must be found to control for the possibility that people who choose one treatment are systematically different from those who chose the other.

To provide an example, consider a public health campaign to encourage use of sun-cream, eg by providing information, marketing and/or subsidising the price. An evaluation randomises individuals to an intervention group (offered subsidy) and a control (not offered subsidy). The outcome might be incidence of sunburn. Following the terminology of Angrist, Imbens and Rubin (1996) we can define four types of individual:

- Compliers – would use suncream if given the intervention, but not otherwise
- Always-takers – would use suncream regardless of intervention
- Never –takers – would not use suncream regardless of intervention
- Defiers – would use suncream if not given the intervention, but would not if given the intervention

If we exclude defiers as violating rational assumptions about monotonicity of preferences, then Figure 13 shows the partition of the population into compliers, never-takers and always-takers. These characteristics might be unobservable.

	Randomised to intervention	*Randomised to control*
Complier	1	0
Never taker	0 (not observed)	0 (looks like complier)
Always taker	1 (looks like complier)	1 (not observed)

Figure 13. Treatment received, or behaviour, for different kinds of individuals
Note : 1 shows desired behaviour (eg uses suncream as recommended) and 0 shows no change

A conventional intention to treat ITT analysis would compare the average difference in outcome (or other measure of relative effect) between those randomised to intervention and those randomised to control. Because randomisation has distributed individuals equally between the groups, this will estimate an unbiased effect across the sample.

In many evaluations, especially in public health, participants are randomised but have some degree of choice and may not ultimately take the randomised treatment.

ITT is a useful and often pragmatic estimate of effectiveness, that is, the average effect of a general strategy of offering an intervention to a population, even though only a proportion take it up (and follow desired behaviour), and some of those would have done so anyway.

It may nevertheless be useful to estimate separate measures of efficacy alongside ITT measures of effectiveness. One such measure of efficacy is ‘adjusted treatment received’ which estimates the difference in outcomes only in the complier group (Grant 2008). Modelling might then be used to estimate effectiveness and cost-effectiveness under different assumptions about the proportion of the population who accept treatment. RCTs are infeasible in many settings. Instead, an observational study might obtain data on a group that took the intervention and a control group. If the choice is assumed to be predictable from observed characteristics, then matching methods might be used to compare gain in similar individuals who took and did not take the intervention. If the choice is thought to be based on unobserved characteristics, an instrumental variable might be found.

In some public health settings, such as mass media information campaigns, it will be difficult to identify a control group. Boyce et al (2008) suggests that before and after studies might be sufficient evidence. As a minimum, evaluations should:

- include behavioural outcome measures where possible
- assess impact over the longer term by finding out if the behaviour change was
- be sustained after the intervention finished
- collect information on cost-effectiveness
- include a control group

With randomised experiments or instrumental variables, one can include covariates, and interactions between covariates and treatment assignment, to estimate if the difference in mean outcomes varies across social groups.

Most empirical research on treatment effects focuses on the estimation of difference in mean outcomes. Estimating distributional effects is straightforward with experimental data; we need only to

compare the distributions of the outcomes for the randomised groups. This might show, for example, whether the treatment effect at the median of the population is different from the effect at the 25th or 75th percentile. Methods are also available for estimating distributional effects using instrumental variables (Abadie et al 2002).

The results from RCTs or observational studies might be difficult to generalise to another setting, where for example the population is different, or where some of the parameters of the intervention are changed. Mathematical modelling might be able to predict treatment effects in these cases of a range of policies on different subgroups of the population. The Sheffield alcohol study (DH 2008a) provides one example of a methodology (discussed earlier). Microsimulation is another approach. Bourguignon et al (2002) developed a microsimulation model to simulate outcomes for a conditional cash transfer programme. The model represents households' choice between school attendance, leisure and work for their children based on a microeconomic model of demand for health over time and was parameterized using household survey data from Brazil. The model was used to simulate outcomes in several scenarios. First, it was used to help policy makers estimate the average outcome if the programme were scaled up to the whole eligible Brazilian population. Second, the model estimated the impact of the policy on different socio-economic groups. Third, it assisted in the evaluation of alternative programme designs, such as changing the means test identifying eligible families, the size of the cash transfer or the behaviour conditions. Mathematical modelling is therefore potentially a very powerful tool for policy development, offering a complement to empirical data and allowing policy makers to explore rapidly many different scenarios.

Modelling health and other outcomes for a whole population can be very complex task, particularly when dynamic effects and feedback is considered. The Foresight Project found that constructing a model of how the determinants of obesity might evolve over time at an individual level was infeasible (Foresight 2007). To simplify, they extrapolated current trends for obesity into the future, and modelled the effect on health.

7. Conclusions

This paper has discussed from an economic perspective how government policy may influence social determinants of health and health inequalities. The CSDH showed that the many of the causes of ill health, and therefore the solutions, lie outside the remit of health ministries. Therefore the scope of this paper has been very broad, covering

- Theoretical models of the relationships between human capital, income and health
- Empirical evidence on the relationship between income and health
- The role of individual lifestyle and consumer behaviour for health
- The use of performance indicators and targets relating to health inequalities
- Evaluation and priority setting for policies aimed at reducing health inequalities

Evidence for a (cross sectional) social gradient in health is strong. However, it is difficult to estimate the causal relationships between income, social factors and health, because of endogeneity, the influence of other factors and long time lags. Health problems and health behaviour tend to be strongly persistent. This makes changing trends in health and health inequalities at the macro level very challenging within the time frame envisaged by national targets.

Micro-level studies strongly suggest that causality of income and health runs both ways. Education and other 'permanent' changes to income have a stronger influence on health than temporary changes. It is difficult to generalise about the relation between health and income at a national level, though perhaps the data suggest a stronger relation from health to GDP growth than the other way around, and a greater size of effect in low income countries. On a macro level, an important question is how the worsening macroeconomic climate will affect health and health inequalities, and the steps that should be taken to mitigate the consequences of the recession on health. However, macro level studies are ambigous on the effect of lower economic activity on health: recessions may improve some indicators of mortality and morbidity (such as road accidents).

Economic theory offers ambiguous predictions about the relationship between socio-economic status and health behaviour. The epidemiological evidence shows that the relationship between socio-economic status and lifestyle does not follow a straightforward gradient, and is not the same for men and women. However, if people with higher incomes tend on average to healthier behaviour, then under reasonable assumptions, income growth would increase average health but tend to exacerbate inequalities in health. This might offer one explanation of the lack of clear progress towards reducing health inequalities in the UK. Redistributing income towards the disadvantaged might offset to some extent this trend towards growing inequalities in health, but if people with higher incomes tend to healthier behaviour, at the expense of average population health. Evidence from the Institute for Fiscal Studies shows income inequalities rose sharply during the 1980s, mainly because of sharpening wage differentials between educated and less educated workers. Since 1997 the income distribution has been in favour of the less affluent for the bulk of the population, though inequality has increased at the lower and upper ends of the distribution.

Economic theory assumes that people pursue a number of objectives, not all related to longevity and health. Personal choices therefore may be made perfectly rationally to maximise these aims, but may not necessarily maximise health. Economic theory also suggests reasons why personal choices might not be in the best interests of society or even the individuals themselves. Such reasons include inadequate information, externalities such as passive smoking or alcohol-related crime, artificially low prices for unhealthy products (such as farming subsidies for high-fat foods), clustering of behaviour among peer-groups, irrational behaviour or poor self control. Future liabilities on the welfare state (that is, other people) arising from unhealthy behaviour (for example, of obesity) of an individual might also be considered an externality. Government policies that encourage more healthy behaviour would include increasing healthy options, influencing preferences, price controls and consumption taxes, and bans.

Unequal access or use of health services may be at least in part a cause of health inequalities. Evidence from English national survey data from the late 1990s found people with lower income made significantly greater use of primary care services but less use of hospital inpatient and outpatient services, after adjusting for need and supply variables. There was no clear independent trend for other indicators of socio-economic status. QOF data from 2004 delivery of primary care did not show

that delivery of NHS services is systematically either pro-rich or pro-poor, although disadvantaged areas were poorly served in some specific indicators.

At a minimum, the health service should ensure that disadvantaged groups have equal access to NHS services. The English NHS has a well-established resource allocation mechanism that seeks to secure equity in health service access between geographical areas. It has been recently augmented by a major 'health inequalities' adjustment of £7.5 billion intended to direct resources to areas making the biggest contribution to premature mortality and disability. However, this allocation appears to be a compensation for changes in the needs-based element of the capitation formula, and there are few mechanisms to ensure that these resources are devoted to disadvantaged groups within PCTs.

Addressing socio-economic health inequalities may require more radical action, implying some departure from the NHS principle of equal treatment for equal clinical need, towards differential treatments for different social groups. One way in which health service guidelines such as those produced by NICE could address health inequalities is by lowering the cost-effectiveness threshold for initiatives that impact on disadvantaged groups. There has been some work by economists on how society values identical health gains for different population groups. There is evidence of strong preference for equity amongst some people, but preferences are highly variable. In principle, this research can be used to adjust cost-effectiveness ratios for equity concerns. However, studies so far have been relatively small scale and tentative in their conclusions. Moreover, it is not clear if these weights should differ for evaluations of public health interventions compared with health-care interventions.

A greater impact on health inequalities in England and Wales may be realised if NICE were to evaluate more topics which focus on public health, prevention and particularly conditions and interventions with the greatest impact on social inequalities in health. A greater emphasis on positive action to reduce inequalities implies some departure from the principle of equal treatment for those of equal clinical need. Political economy models such as median voter may offer some insight into how voter preferences might affect decision making.

Where social interventions fall outside the traditional remit of the health care service, priority setting is drawn towards a much wider evaluation of the costs and benefits of a policy than are usually considered by a conventional a cost-effectiveness analysis. This might include taking account of the welfare cost of restricting people's choices, alongside the public health benefits. Public health policies such as price controls and consumption taxes are not restricted to ill people, or even those at high risk of illness, but are aimed at (or indirectly affect) a much wider population.

A full analysis of public health and social interventions such as cross-departmental initiatives is likely to be highly complex where there are a wide range of potential costs and consequences. Evaluating the equity implications requires modelling different responses by subgroup, increasing the complexity further. Cost-benefit analysis can provide one possible methodology for valuing different outcomes using a common (monetary) unit, but estimates of willingness-to-pay can be costly and difficult to obtain. Descriptive evaluations (such as Health Impact Assessments) are likely to be a more pragmatic approach, and should be used both where health-service policies impact on other sectors, and where other programmes such as the Common Agricultural Policy are likely to have a significant health impact.

The government can have a profound influence on the shape and performance of local public services through target regimes, funding mechanisms, performance reporting, and staff contracts.

Since 1998 national government priorities have been expressed through the Public Service Agreement (PSA) target regime. This has been successful in some domains (such as waiting times), but less so in others (inequalities, cross-departmental targets). There is a well-developed literature on how to maximize the effectiveness of central targets. English policy is moving towards local priorities, through the Comprehensive Area Assessment initiative being implemented by the Audit Commission.

There has been increased use of local performance reporting for public services, for example through the Healthcare Commission's annual health check and the Audit Commission's Comprehensive Performance Assessment. These have been effective in focusing managers' attention, although not notably in the inequalities domain. A key issue for the future will be the extent to which

Comprehensive Area Assessment succeeds in securing cross-agency collaboration (including public, voluntary and private sector).

There has been increased interest in the extent to which incentives directed at the practitioner level might secure better outcomes than those directed at organizations such as PCTs and hospitals. The GP Quality and Outcomes Framework is the most notable English example. It has undoubtedly secured improved focus of GP activity, although the small measured gains that can be attributed to the QOF do not yet seem to justify the large expenditure.

While the government does not appear on track to achieve the 'headline' target of reducing the gap in life expectancy between the most deprived areas and the national average by 2010, progress has been made towards other, more specific targets, such as reducing child obesity, which may be a cause for cautious optimism for progress over the longer term. It should also be noted the headline target chosen by the government is only one of many possible ways to measure and summarise the distribution of health in the population, and health inequalities have improved by some other measures.

Other detailed indicators are less promising for future trends. According to HSE data 1991-2005, socio-economic inequalities appear to have worsened among women to a significantly greater extent than among men, across several indicators, including life expectancy, obesity, mental health and cardiovascular disease. In some indicators, such as mental health, better-off women have improved health faster than other socio-economic groups, while in other cases, particularly obesity and cardiovascular disease, there appears to have been a deterioration of the health of the most disadvantaged women in absolute terms.

It is of course difficult to extrapolate future trends, though this does not mean it should not be attempted. Modelling work has been undertaken on a micro level, for example in obesity, heart disease and tobacco control to inform predictions of the burden of disease.

References

Abadie A, Angrist J, Imbens G. Instrumental variable estimation of the effects of subsidizing training on quantiles of trainee earnings. *Econometrica*. 2002;70:91-117.

Acemoglu D, Johnson S. Disease and development: The effect of life expectancy on economic growth. *Journal of Political Economy*. 2007;115(6):925-85.

Adda J, Banks J, von Gaudecker H. *The Impact of Income Shocks on Health: Evidence from Cohort Data* IZA DP No. 3329 Bonn: IZA; 2008.

Alderman H, Behrman JR, Hoddinott J. Nutrition, malnutrition and economic growth In: López i Casasnovas G, Rivera B, Currais L, editors. *Health and economic growth findings and policy implications* Massachusetts: The MIT Press; 2005. p. 169-95.

Alderman H, Behrman JR, Lavy V, Menon R. Child health and school enrollment. A longitudinal analysis. *Journal of Human Resources*. 2000;36(1):185-205.

Angrist, Imbens, Rubin. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*. 1996;91:444-72.

Ashraf Q, Lesley A, Weil N. *When Does Improving Health Raise GDP?* NBER Working Paper No. W14449 National Bureau for Economic Research; 2008.

Ashworth M, Seed P, Armstrong D, Duraba S, Jones R. The relationship between social deprivation and the quality of primary care: a national survey using indicators from the UK Quality and Outcomes Framework. *British Journal of General Practice*. 2007;57(539):441-8.

Bambra C, Gibson M, Petticrew M, Sowden A, Whitehead M, Wright K. *Tackling the wider social determinants of health and health inequalities: Evidence from systematic reviews*. York: Public Health Research Consortium, University of York; 2009.

Barro RJ, Sala-i-Martin X. *Economic growth*. Massachussets: The MIT Press; 2003.

Bauld L, Judge K, Platt S. Assessing the impact of smoking cessation services on reducing health inequalities in England:observational study. *Tobacco Control*. 2007;6:400-4.

Becker G. A theory of the allocation of time. *Economic Journal*. 1965;75:49.

Behrman JR, Hoddinott J. Programme evaluation with unobserved heterogeneity and selective implementation: The Mexican Progresa impact on child nutrition. *Oxford Bulletin of Economics and Statistics* 2005;67(4):547-69.

Behrman JR. Impact of health and nutrition on education. *World Bank Research Observer*. 1996;11(1):23-37.

Bevan G. *Review of the Weighted Capitation Formula RARP33*. London: Department of Health; 2008.

Bhargava A, Jamison DT, Lau LJ, Murray C. Modelling the effects of health on economic growth. *Journal of Health Economics*. 2001;20:423-40.

Bloom D, Canning D, Sevilla J. *The effect of health on economic growth: Theory and evidence* (NBER Working Paper No. 8587) National Bureau of Economic Research; 2001.

Bourguignon F, Ferreira F, Leite P. *Ex-ante evaluation of conditional cash transfer programmes: The case of the Bolsa Escola* (William Davidson Working Paper No. 516) 2002.

Boyce T, Robertson R, Dixon A. *Commissioning and behaviour change: Kicking bad habits* final report. London: Kings Fund; 2008.

Brewer M, Muriel A, Phillips D, Sibieta L. *Poverty & Inequality in the UK: 2008*. 2008. Institute for Fiscal Studies, London.

Case A. *Does money protect health status? Evidence from South African pensions*. Center for Health & Wellbeing Working Paper. Princeton NJ: Woodrow Wilson School, Princeton University; 2001.

Case A, Deaton A. *Broken down by work and sex: How our health declines* (NBER Working Paper No. W9821). Massachussets: National Bureau of Economic Research; 2003.

Case A, Fertig A, Paxton C. The lasting impact of childhood health and circumstance. *Journal of Health Economics*. 2005;24:365-89.

Christakis, Fowler. The spread of obesity over a large social network over 32 years. *New England Journal of Medicine*. 2007;357:370-9.

Claxton K, Sculpher M, Culyer A. *Mark versus Luke? Appropriate methods for the evaluation of public health interventions*. CHE Research Paper 31. York: University of York; 2007.

Claxton K, Walker S, Palmer S, Sculpher M. Appropriate perspectives for health care decisions. CHERP (forthcoming): University of York.

CMH. *Macroeconomics and health: Investing in health for economic development.* Geneva: Commission on Macroeconomics and Health, World Health Organization; 2001.

CMO. *Chief Medical Officers Report. 2008*. Department of Health, London.

Contoyannis P, Forster M. Our healthier nation? *Health Economics*. 1999;8:289-96.

Contoyannis P, Jones A, Rice N. The dynamics of health in the British Household Panel Survey. *Journal of Applied Econometrics*. 2004;19:453-503.

Contoyannis P, Jones A. Socio economic status, health and lifestyle. *Journal of Health Economics*. 2004;23:965-95.

CSDH. *Closing the gap in a generation: health equity through action on the social determinants of health*. Final report of the Commission on Social Determinants of Health. Geneva: World Health Organisation; 2008.

Cutler D, Lleras-Muney, A. Education and health: Evaluating theories and evidence. In: House J, Schoeni G, Kaplan G, Pollack H, editors. *The health effects of social and economic policy*. New York: Russell Sage Foundation; 2007.

Department for Children SaF. Every Child Matters website. http://www.everychildmatters.gov.uk/. 2008.

Department of Health. *Choosing Health*. London: DH; 2004.

Department of Health. *General Medical Services Statement of Financial Entitlements for 2006 onwards*. Leeds: Department of Health; 2006.

Department of Health. *Tackling health inequalities: 2004/06 data and policy update for the National Target*. London: DH; 2007(a).

Department of Health. *Health impact assessment: questions and guidance for impact assessment.* 2007(b). http://www.dh.gov.uk/en/Publicationsandstatistics/Legislation/Healthassessment/Browsable/DH_075622. 2007.

Department of Health. *Independent review of the effects of alcohol pricing and promotion: Part B. Modelling the potential impact of pricing and promotion policies for alcohol in England:* Results from

the Sheffield Alcohol Policy Model Version 1.0. http://www.shef.ac.uk/mediacentre/2008/1128.html. Sheffield: University of Sheffield; 2008(a).

Department of Health. *Resource allocation: weighted capitation formula*. Sixth Edition. Leeds: Department of Health; 2008(b).

Department of Health. *Systematically addressing health inequalities*. Leeds: DH; 2008(c).

Department of Health. *The Department Autumn performance Report.* http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_091854. Leeds: DH; 2008(d).

Department of Health. *Next stage review*. Leeds: Department of Health; 2008(e).

Department of Health. *Post 2010 strategic review on health inequalities: the Marmot Review* http://www.ucl.ac.uk/gheg/marmotreview. Leeds: Department of Health; 2009.

Dolan P, Edlin P, Tsuchiya A. *The relative social value of health gains to different beneficiaries.* London: NICE; 2008.

Doppelhofer G, Miller R, Sala-i-Martin X. Determinants of long-term growth: A bayesian averaging of classical estimates (BACE) approach. *American Economic Review*. 2004;94(4):813-35.

Drummond M, Weatherly H, Mason A. *Assessing the challenges of applying standard methods of economic evaluation to public health interventions*. Retrieved September 2008 http://www.york.ac.uk/phrc/projects_1.htm York: PHRC; 2006.

Fernald L, Gertler P, Neufeld L. Role of cash in conditional cash transfer programmes for child health, growth and development: an analysis of Mexico's Oportunidades. *Lancet*. 2008;371:828-37.

Foresight *Modelling Future Trends in Obesity and the Impact on Health*. www.foresight.gov.uk/OurWork/ActiveProjects/Obesity/Obesity.asp 2007.

García Gómez P, López i Casasnovas G. Socio-economic inequalities in health in Catalonia. *Hacienda Pública Española*. 2005;175:103-21.

García Gómez P, López Nicolás A. Health shocks, employment and income in the Spanish labour market. *Health Economics*. 2006;15:997-1009.

Goddard M, Hauck K, Smith PC. Priority setting in health – a political economy perspective *Health Economics, Policy and Law*. 2006;1(1):79-60.

Golcuk P. *Ex-ante moral hazard: impact of health insurance on health behaviors*. Stockholm: University of Stockholm; 2008.

Grant A, Wileman S, Ramsay C, Mowat A, Krukowski Z, Heading R, et al. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. *British Medical Journal*. 2008;337:a2664.

Gravelle H, Rees R. *Microeconomics*. Harlow: Prentice Hall; 1992.

Gravelle H, Sutton M. Income related inequalities in self assessed health in Britain 1979-1995. *Journal of Epidemiology and Community Health*. 2003;57:125-9.

Harrison, Gardiner. Do the rich really die young? alcohol-related mortality and social class in Great Britain, 1988-94 *Addiction*. 1999;94:1871-80.

Hauck K, Shaw R, Smith P. Reducing avoidable inequalities in health: a new criterion for setting health care capitation payments. *Health Economics*. 2002;11(8):667-77.

Heller R, Gemmell I, Wilson E, Fordham R, Smith R. Using economic analyses for local priority setting: The population cost-impact approach. *Applied Health Economics and Health Policy*. 2006;5(1):45-54.

HM Treasury. *Stability, security and opportunity for all: investing for Britain's long term future.* Public Spending Plans 2005-2008. London: HM Treasury; 2004.

House of Commons Select Committee 2003/2004. http://www.parliament.the-stationery-office.co.uk/pa/cm200304/cmselect/cmhealth/23/2309.htm. 2004.

Hunter. England. In: Marinker M, editor. *Health targets in Europe*. London: BMJ Books; 2002.

Jones A., Wildman J. Health, income and relative deprivation: Evidence from the BHP. *Journal of Health Economics* 2008; *27*(2): 308-324.

Kemm J. *More than a statement of the crushingly obvious: A critical guide to Health Impact Assessment* Birmingham: West Midlands Public Health Observatory; 2008.

Kenkel D. Prevention. In: Culyer, Newhouse, editors. *Handbook of Health Economics* The Netherlands: Elsevier Science BV, Vol 1B, 1676–1720; 2000.

Kikumbih N, Hanson K, Mills A, Mponda H, Schellenberg J. The economics of social marketing: The case of mosquito nets in Tanzania. *Social Science & Medicine* 2005;60:369–81.

Komlos, Smith, Bogin. *Obesity and the rate of time preference: is there a connection?* Available at SSRN: http://ssrn.com/abstract=464105 or DOI: 10.2139/ssrn.464105. 2003.

Kremer M, Holla A. *Pricing and access: Lessons from randomised controlled evaluations in education and health.* Retrieved September 2008, from http://www.economics.harvard.edu/faculty/kremer/files/PricingandAccess080803.pdf

Lader D. *Smoking-related Behaviour and Attitudes*. Omnibus Survey Report No. 36. . Newport: Office for National Statistics; 2007.

Lagarde M, Haines A, Palmer N. Conditional cash transfers for improving uptake of health interventions in low and middle income countries. *Lancet*. 2007;298:1900-10.

Le Grand J. Giants of excess. *Journal of the Royal Statistical Society*, Series A. 2008;171(4):843-56.

Lipsey R, Lancaster K. The general theory of second best. *The Review of Economic Studies*. 1957;24(1):11-32.

Mackenbach J. *Health inequalities: Europe in profile*. London: Department of Health; 2006.

Mason, Walker. *The GMS quality and outcomes framework. Are the quality and outcomes framework (QOF) indicators a cost-effective use of NHS resources?* York: University of York; 2008.

McGovern M, Boroujerdi M, Taylor M, Williams D, Hannaford PC, Lefevre K. The effect of the UK incentive-based contract on the management of patients with coronary heart disease in primary care. *Family Practice* 2008;25:33-9.

McNeill A. *Response to DH consultation*. http://www.ias.org.uk/resources/papers/healthcare.pdf. Institute for Alcohol Studies 2003.

Miguel E, Kremer M. Worms: Identifying impacts on education and health in the presence of treatment externalities. *Econometrica*. 2004;72(1):159-217.

Millett C, Gray J, Saxena S, Netuveli G, Khunti K, Majeed A. Ethnicity and quality of diabetes care in a health system with universal coverage:population based cross-sectional survey in primary care. *CMAJ*. 2007;176:1705-10.

Mindell JS, Boltong A, Forde I. A review of health impact assessment frameworks. *Public Health* 2008 Nov;122(11):1177-87.

Mitton C, Patten S, Donaldson C. Listening to the Decision Makers: Sustainability of PBMA in Alberta. *Applied Health Economics and Health Policy*. 2004;3(3):143-51.

Morris S, Olinto P, Flores R, Nilson E A F. Conditional cash transfers are associated with a small reduction in the rate of weight gain of preschool children in Northeast Brazil. *Journal of Nutrition*. 2004;134:2336-41.

Morris S, Sutton M, Gravelle H. Inequity and inequality in the use of healthcare in England: an empirical investigation. *Social Science & Medicine*. 2005;60(6):1251-66.

New York City.http://www.nyc.gov/html/om/pdf/2007/pr201-7_cct_incentives_schedule_final.pdf. 2007.

NICE. *Appraising orphan drugs*. http://www.nice.org.uk/niceMedia/pdf/smt/120705item4.pdf. London: NICE; 2008a.

NICE. *Methods for the appraisal of health technology*. London: NICE; 2008b.

NICE. *Obesity full guidance final version*. CG43. London: NICE; 2006.

Rawlins M, Culyer A. National Institute for Clinical Excellence and its value judgments. *British Medical Journal*. 2004;329:224-7.

Rivera B, Currais L. Individual returns to health in Brazil: A quantile regression analysis. In: López i Casasnovas G, Rivera B, Currais L, editors. *Health and economic growth findings and policy implications* Massachusetts: The MIT Press; 2005. p. 169-95.

Rogeberg O. *Taking absurd theories seriously: economics and the case of rational addiction theories.* Philosophy of Science. 2004;71:263-85.

Ruhm C. Macroeconomic conditions, health and mortality. In: Jones A, editor. *The Elgar Companion to Health Economics*; 2006.

Salay R, Lincoln P. Health impact assessments in the European Union. *Lancet* 2008 Sep 6;372(9641):860-1.

Sassi F, Hurst J. *The prevention of lifestyle related disease: an economic framework*. OECD working paper 32. Paris: OECD 2008.

Sassi F. Health inequalities: a persistent problem. In: Hills J, Sefton T, Stewart K, editors. *Towards a more equal society?* Bristol: The Policy Press; 2009.

Sefton T, Hills J, Sutherland H. Poverty, inequality and redistribution. In: Hills J, Sefton T, Stewart K, editors. *Towards a more equal society?* Bristol: The Policy Press; 2009.

Sen A. Why health equity? *Health Economics*. 2002;11:659-66.

Shiell A. *Economics of health promotion*. Harvard Health Policy Review. 2006;7(2):21-31.

Shiell A. Still waiting for the great leap forward. *Health Economics, Policy and Law*. 2009;4:255-60.

Simpson CR, Hannaford PC, Lefevre K, Williams D. Effect of the UK incentive based contract on the management of patients with stroke in primary care. *Stroke*. 2006;37:2354-60

Smith JP. The impact of social economic status on health over the life course. *Journal of Human Resources*. 2007;42:739-64.

Smith PC. England: intended and unintended effects. In: Wismar M, McKee M, Ernst K, Srivastava D, Busse R, editors. *Health targets in Europe: learning from experience*. Copenhagan: World Health Organization; 2008. p. 62-82.

Smith PC. *Performance budgeting in England: Public service agreements*. Washington DC: International Monetary Fund; 2007.

Solar O, Irwin A. *Towards a conceptual framework for analysis and action on the social determinants of health*. Discussion Paper for the Commission on the Social Determinants of Health. Geneva: Commission on the Social Determinants of Health. 2007.

Suhrcke M, Nugent R, Stuckler D, Rocco L. *Chronic disease: an economic perspective*. Retrieved January 2009 from http://www.oxha.org/knowledge/publications/oxha-chronic-disease-an-economic-perspective.pdf. London: Oxford Health Alliance 2006.

Suhrcke M. *Health: a vital investment for economic development in Eastern Europe and central Asia*. Copenhagen: WHO; 2007.

Suhrcke M. *The contribution of health to the economy in the European Union*, (http://ec.europa.eu/health/ph_overview/Documents/health_economy_en.pdf) Health and Consumer Protection Directorate General, European Commission 2005.

Thomas D., Strauss J. Health and wages: evidence on men and women in urban Brazil. *Journal of Econometrics*. 1997;77(1):159-86.

Tsuchiya A, Dolan P. Do NHS clinicians and members of the public share the same views about reducing inequalities in health? *Social Science & Medicine*. 2007;64:2499-503.

Tsuchiya A, Dolan P. Equity of what in health? *Health Economics*. 2008;DOI: 10.1002/hec.1355.

Van de Poel E, Hosseinpoor A, N S. Socioeconomic inequality in malnutrition in developing countries. *Bulletin of the World Health Organization* 2008;86(4):241-30.

Van Ourti T, van Doorslaer E, Koolman X. The effect of income growth and inequality on health inequality: theory and empirical evidence from the European Panel. *Journal of Health Economics*. 2009;doi:10.1016/j.jhealeco.2008.12.005.

Wanless D, Appleby J, Harrison A, Patel D. *Our future health secured?* London: Kings Fund; 2007.

Wanless D. *Securing our future health: taking a long term view.* London: Her Majesty´s Treasury; 2002.

West R. *Key performance indicators on smoking cessation in England: findings from the Smoking Toolkit Study*, accessed at: www.aspsilverbackwebsites.co.uk/smokinginengland/ref/smokinginengland.ppt. 2008

WHO. *Health Impact Assessment: Gothenburg consensus paper*. http://www.euro.who.int/document/pae/gothenburgpaper.pdf. Brussels: WHO; 1999.

Wilkinson R, Pickett K. *The spirit level: why more equal societies almost always do better*. 2009.

Williams A. The rationing debate: Rationing health care by age: The case for *British Medical Journal.* 1997;314:820.

York Health Economics Consortium. *Cost benefit analysis of health impact assessment*. 2006. Department of Health: London.