

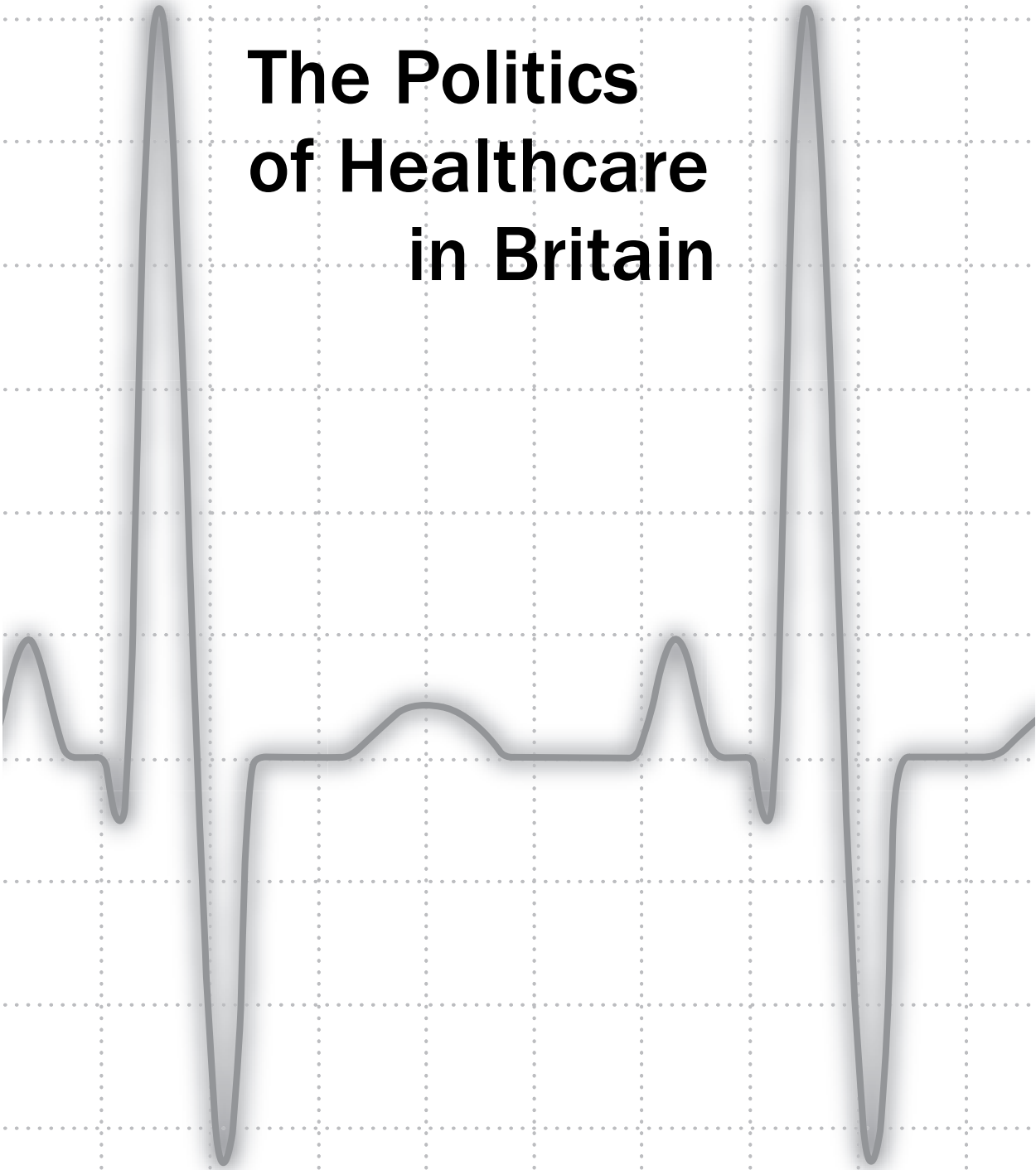


The Politics of Healthcare in Britain

Stephen Harrison
and Ruth McDonald



The Politics of Healthcare in Britain





The Politics of Healthcare in Britain

**Stephen Harrison
and Ruth McDonald**



SAGE Publications

Los Angeles • London • New Delhi • Singapore

© Stephen Harrison and Ruth McDonald 2008

First published 2008

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as permitted under the Copyright, Designs and Patents Act, 1988, this publication may be reproduced, stored or transmitted in any form, or by any means, only with the prior permission in writing of the publishers, or in the case of reprographic reproduction, in accordance with the terms of licences issued by the Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to the publishers.



SAGE Publications Ltd
1 Oliver's Yard
55 City Road
London EC1Y 1SP

SAGE Publications Inc.
2455 Teller Road
Thousand Oaks, California 91320

SAGE Publications India Pvt Ltd
B 1/I 1 Mohan Cooperative Industrial Area
Mathura Road
New Delhi 110 044

SAGE Publications Asia-Pacific Pte Ltd
33 Pekin Street #02-01
Far East Square
Singapore 048763

Library of Congress Control Number: 2007939449

British Library Cataloguing in Publication data

A catalogue record for this book is available from the British Library

ISBN 978-0-7619-4159-0

ISBN 978-0-7619-4160-6 (pbk)

Typeset by C&M Digital Pvt Ltd, Chennai, India

Printed in Great Britain by The Cromwell Press Ltd, Trowbridge, Wiltshire

Printed on paper from sustainable resources

Contents

<i>Preface</i>	vii
1 The Politics of Healthcare Resources and Rationing	1
• Key concepts in third party payment	2
• An historical sketch of supply and demand in the NHS	9
• Contemporary approaches to managing supply and demand	18
• Interpreting third party payment	20
2 The Politics of Health Professionalism	27
• Key concepts in professionalism	28
• An historical sketch of health professionalism in the UK	35
• Contemporary challenges to professionalism	42
• Interpreting the challenges to professionalism	47
3 The Politics of Clinical Knowledge	55
• Key concepts in knowledge and evidence	56
• An historical sketch of evidence-based practice	60
• Contemporary evidence-based practice	65
• Interpreting scientific–bureaucratic medicine	73
4 The Politics of Organisation and Management	79
• Key concepts in social co-ordination	80
• An historical sketch of NHS formal organisation	88
• Contemporary NHS organisation	93
• Interpreting the ‘third way’	98
5 The Politics of Democracy and Participation	103
• Key concepts in participation	104
• An historical sketch of public and patient participation	110
• Contemporary developments in participation	118
• Interpreting participation as policy	121
6 The Politics of Healthcare Policy Making	127
• Key concepts in the policy process	128
• An historical sketch of NHS Policy Making	135

• Contemporary healthcare policy making	139
• Interpreting the contemporary policy process	142
7 The Politics of Contradiction: Three Case Studies	149
• Consumerism and evidence-based medicine	150
• Diversity and devolution	156
• Choice, responsibilisation and health inequalities	161
<i>Glossary of abbreviations</i>	167
<i>References</i>	169
<i>Index</i>	201

Preface

Our aims in this volume are to introduce readers to a fairly wide range of descriptive and analytical material about the past and present politics of healthcare in Britain. Our working definition of ‘healthcare’ is conventional; the majority of our material relates to the funding, organisation and delivery of diagnostic and therapeutic services to individuals, rather than to broader conceptions of public health. Our working definition of ‘politics’ is conventional too, though it perhaps merits some explanation. The funding and provision of healthcare in Britain, and indeed all western countries, is a central concern of public policy irrespective of widely differing degrees of public funding and public ownership of healthcare institutions. This concern extends beyond these matters of finance and provision to the extent of shaping other aspects of politics and public policy, an observation summed up in Moran’s notion of the ‘healthcare state’:

There is more to health care politics than health care policy; the scale of health care institutions means that they have ramifications for the modern state well beyond conventional health care arenas. Like any state, the health care state is about governing; and in the act of governing states shape health care institutions, and are in turn shaped by those institutions ... Health care systems pose problems for statecraft; but they also offer ways of solving problems, often problems whose origins lie beyond health care systems themselves. (Moran, 1999: 4–5)

This description is certainly apt in a Britain where not only does public expenditure on healthcare and direct provision of such care massively outweigh what is privately funded or privately provided, but also where this arrangement is routinely used by governments to enhance their political legitimacy. This connects with contemporary ideas about ‘governance’ which emphasise that, despite the formal provisions of national constitutions, states and governments do not simply govern in a top-down fashion. Rather, they seek to steer society through a variety of channels, some of which are indirect (Pierre and Peters, 2000: 4–5). In order to govern in this sense, governments must by various means enlist the efforts of other social actors. In Britain, the latter of course include ‘official’ public bodies, such as the institutions of the National Health Service (NHS), numerous professional, academic and other interest groups, and less easily definable ‘social movements’ (such as the ‘patient movement’ or the ‘evidence-based medicine movement’) based as much on shared identities as shared interests (Byrne, 1997). Mapping and analysing the interactions between governments and such actors are a central focus of this volume.

The third element of our book title locates our work in the context of Britain, that is England, Scotland and Wales. There have long been organisational differences (and differences

of official terminology) between the three countries (Williamson and Room, 1983). As we show in Chapter 7, these have widened in recent years as a result of devolution to the Scottish Parliament and the Welsh Assembly (Greer, 2004), which is also beginning to result in differences in patient entitlement. Elsewhere in the book, however, we have confined our official terminology to that of the English NHS, whilst trying to ensure that the overall thrust of our analysis is applicable to the whole of Britain. The continuing rapid rate of NHS organisational change precludes any attempt to include definitive organisation charts and even the Department of Health no longer seems to attempt this. In order to avoid becoming mired in the minute history of changes in the titles of statutory bodies, we have generally referred simply to 'health authorities' where the context does not require precision.

A textbook such as the present volume does not have the same sort of aims as the research papers and conceptual reviews on which, as full-time researchers, we spend most of our time. The latter are tightly written in order to make at the most a few points; they therefore tend to employ a narrow range of concepts and literature that relate closely to the argument and/or evidence that is being deployed. In contrast, a textbook has the wider aim of informing readers about the general state of its subject matter. It must introduce a selection of relevant theories, concepts and evidence but it will necessarily leave loose ends and confine itself to indicating general lines of analysis and argument rather than pursuing them rigorously to a single conclusion. In order to meet this wider textbook aim, we have adopted a particular and distinctive structure for each of our substantive themes, that is Chapters 1 to 6. Each has four main sections; the first introduces a range of concepts that we take to be central to the particular theme, the second section provides a summary history of the theme, and the third summarises recent and contemporary developments. The final section of each chapter consists of discussion of how a small sample of theories might be used to address questions relevant to the chapter's theme. It is important to stress that these discussions do not constitute serious 'tests' of the theories; our purpose is rather to suggest to the reader how such abstract material can be related to substantive accounts. Our book would have been unacceptably long had we not been selective in our choice of themes. Important casualties of this selection process have included the politics of public/environmental health, the politics of pharmaceutical manufacture and regulation, and the politics of social care. On community care, see Means et al. (2003) and for public health, see Baggott (2000) and Lewis (1986). On pharmaceuticals, see Abraham (1997), Abraham and Lewis (2001), Davis (1997) and the edited European collection by Mossialos et al. (2004).

Some excellent health policy texts are very limited in theoretical coverage, either employing it only implicitly or treating it as primarily critical. We have taken the opposite view here, employing a wide range of conceptual and theoretical material, drawn mainly from political science and sociology but with important contributions from economics. Some readers may feel that we have been too eclectic, and that we should have undertaken a consistent political analysis, or that important intellectual traditions have been neglected. Others may feel that there is altogether too much conceptual material. We hope, however, that most will find our approach stimulating in terms of generating questions and analyses of their own. Our policy of providing extensive citations and a reading guide is designed to support further study.

Chapter 7 is designed to work in a different way from the other chapters. It follows from our textbook philosophy, summarised above, that we cannot provide a final chapter to summarise the book and neatly tie up the loose ends. Instead, Chapter 7 addresses the risk that our thematic approach diverts attention from potential interaction and tensions between themes. We therefore consider three such tensions that may set the scene for future political and policy conflicts.

Readers will note that our book is extensively referenced in relation to specific points made in the text. The text also sometimes indicates sources for more detailed coverage or overviews of such points. In addition, at the end of each chapter we provide a brief guide to further reading which relates to issues covered in that chapter.

With regard to reading materials on health policy and politics more generally, of the numerous UK health policy texts available, the current editions of Ham (2004) and Baggott (2004) offer wide-ranging and complementary coverage. The two-volume official history of the NHS (Webster, 1988, 1996) is comprehensive from before 1948 to 1979; a shorter but still detailed account of the first 50 years is Rivett (1998). Specifically political histories of the NHS are the current editions of Klein (2006) and Webster (2002). The creation of the NHS in 1948 is the subject of studies by Willcocks (1967) and Pater (1981). The detailed history of NHS organisation through to 1998 can be pieced together from the various editions of Levitt (1976; 1979), Levitt and Wall (1984, 1992) and Levitt et al. (1995, 1999). The history of the Department of Health is selectively covered in Rayner (1994), Day and Klein (1997) and Greer and Jarman (2007). The successive editions of the *Compendium of Health Statistics* (most recently Office of Health Economics, 2007) provide useful tabular and graphical data about the NHS and British healthcare, with some international comparisons.

Academic authors incur numerous intellectual debts, most of which they quickly forget as others' ideas become incorporated in their own thought. The following list of acknowledgements is no doubt grievously incomplete, and we apologise for omissions: Andrew Gray (formerly University of Durham), Huw Davies (St Andrews), Tim Milewa (Brunel University), Christopher Pollitt (Catholic University of Leuven) and our University of Manchester colleagues Kath Checkland, George Dowswell, Mick Moran, John Pickstone and Martin Roland.

Whilst writing this book we have both experienced rather closer contact with the NHS than we would have wished (as patient and carer respectively) so we are grateful to Karen Phillips and Anna Luker of Messrs Sage for their continued patience in the face of our numerous postponements of the manuscript delivery date. Finally, thanks for both support and appropriate diversion are due to our partners Annie Dearman and Tim Payne respectively, and in SH's case to Vic Gammon and the gang at Ryburn 3-Step (www.ryburn3step.org).

Steve Harrison
Ruth McDonald

Chapter 1

The Politics of Healthcare Resources and Rationing

Summary of chapter contents

- Third party payment in healthcare: what is it and why is it used?
- The consequences of third party payment
- Managing healthcare supply and demand
- Ways of interpreting third party payment: neo-Marxist and public choice theories



The phrase ‘out of pocket payment’ denotes the manner in which we purchase most goods and services in developed economies. We may literally take cash from our pockets, write cheques on our bank accounts, or employ debit or credit cards. Even if we are using credit cards, the end result is the same; we pay personally from our own resources, now or in the future. There is no necessary reason why health and medical care cannot be purchased in the same manner, and indeed most of us purchase a range of non-prescription ‘over the counter’ remedies from pharmacists, spectacles from opticians and so on, purchases that account for a little over 7 percent of all UK healthcare expenditure. A further proportion of such expenditure (some 3 percent) is accounted for by private (non-NHS) healthcare purchased out of pocket from doctors, other professional clinicians and hospitals. This figure does not include expenditure on private healthcare insurance (almost 4 percent of the total), which is one example of an alternative approach to healthcare purchase termed ‘third party payment’, to signify the involvement of a third party (in this example, the insurance company) in the transaction in addition to the patient and the clinician or hospital. The NHS is a larger example (some 86 percent of the UK total healthcare expenditure) of third party payment in

which central government acts as the third party, employing tax revenues (rather than insurance premiums) to fund healthcare for UK residents. Third party payment systems of one type or another are the international norm in the healthcare field and although out-of-pocket payment usually exists alongside it, the former usually dominates in terms of public policy and expenditure. As the above figures (estimated from Office of Health Economics, 2003: Tables 2.1, 2.22) show, the latter is certainly true of the UK. The first section of this chapter examines the concept of, and rationale for, third party payment, together with some of the consequences of its adoption as public policy. The second section is an historical sketch of how these consequences have emerged and been addressed in the UK since the foundation of the NHS in 1948, whilst the third section gives a more detailed account of what might be called the 'big ideas' for managing healthcare supply and demand that have emerged in the last few years. The final section of the chapter examines a number of theoretical questions about the basis and consequences of third party payment.

Key concepts in third party payment

The principle of third party payment is that financial contributions are collected from groups, irrespective of the immediate healthcare requirements of the individuals who compose them. Such groups may represent a more-or-less complete national population, or narrower groups such as the members (voluntary or compulsory) of social or private insurance schemes. Contributions are collected by 'third party payers' such as government or quasi-independent agencies or insurance companies, which employ the resources thus obtained to resource or reimburse healthcare providers (individual clinicians and/or healthcare institutions such as hospitals) for the care of individuals considered to be sick. Third party payment thus separates payment for care from its immediate consumption, and to varying degrees separates the financial contribution that the individual makes from the volume of care that they actually consume. In a *tax-financed* system, the government acts as third party payer by employing resources collected through the tax system to pay for citizens' care. Since the taxation system collects revenues to support public expenditure on a wide range of services, health's share may not be hypothecated ('earmarked'), thereby allowing governments to shift their expenditure priorities between different programme areas. In a *social insurance* system, the third party payers are social insurance funds, the number of which varies between countries (Mossialos et al., 2002). They may be non-governmental bodies whose history lies in trade union and voluntary effort or may be managed by the state, but in either case their resources will remain hypothecated ('earmarked') for healthcare and other specified services and not be merged with other revenues. Membership of a fund may be compulsory for some or all citizens. Members make periodic contributions to the fund, typically based on a percentage of earnings. Employers may also contribute, and non-earners may have their contributions to the fund met as a social security entitlement. The fund in turn pays the clinician or hospital for services provided to members, often at rates negotiated annually between organisations representing the various interests in the healthcare industry. A *private insurance*



system treats the cost of healthcare as an insurable risk for an individual. Coverage might be voluntary, or routinely provided as a condition of employment for some workers. In this context, the third party payer is the insurance company, health maintenance organisation, or nonprofit friendly society, to whom the flow of money takes the form of premiums or subscriptions paid by policy holders, which provide the resources to pay for their care. The state may support private health insurance, for instance by allowing employers to provide health insurance as an employment benefit and to treat the premiums as a tax-deductible business expense. Britain and Sweden are typically treated as textbook examples of tax-financed systems (often termed 'national health services'), with Germany as an example of social insurance and the United States as an example of predominant private insurance (Freeman, 2000; Harrison and Moran, 2000; Moran, 1999). A fourth possibility is that governments with substantial direct revenues, such as the income of several Middle Eastern states from oil production, may simply provide healthcare facilities from these revenues.

There is no necessary relationship between a particular funding mechanism and the ownership of hospitals and other healthcare providing organisations. Public funds can be used to purchase care in private hospitals (as in Germany and, as we note in Chapter 4, increasingly in the UK) and private insurance could be used to purchase care in public hospitals (as frequently occurs in Britain). Nor is there any necessary relationship with the manner in which clinicians are remunerated; fee-for-service, capitation or salaries may be used either singly or in combinations. The characterisation of countries in terms of particular types of third party payment can conceal important similarities. For instance over 40 percent of US healthcare expenditure is accounted for by tax-financed public programmes such as Medicare (for older people) and Medicaid (for poor people). It might also be argued that the more compulsory social insurance becomes, and the fewer the different funds involved (as in contemporary Germany), the more it resembles tax financing. Moreover, the categories are not discrete; contemporary France retains many of the institutions of social insurance but increasingly funds them via general taxation (Jacobzone, 2004: 81–2). The most important similarity between these various systems is the underlying principle that third party payment detaches the act of payment for healthcare from that of receiving it when considered necessary.

Uncertainty and equity: rationales for third party payment in healthcare

Third party payment separates the individual's financial contribution from the volume of care that they consume though, especially with private insurance, there may be a relationship between the amount required to be paid and the volume *predicted* to be used by an individual. Third party payment pools (or 'socialises') resources to smooth out the uncertainties of individuals' health states requiring more expenditure than they are able to make, either as income maintenance (Lewis, 2000: 91) or, more usually, in the form of healthcare. Though judgements about their relative merits may be made, all third party payment systems socialise at least some of the financial risks of ill-health across a group that is distinct from current patients. Unlike out-of-pocket payment, third party payment never limits the value of benefits provided to the sum of an individual's contributions. Indeed, all things being equal,

it would be better to save the money in one's own bank, thereby receiving interest and avoiding administration costs, than to participate in a scheme that limited benefits to the value of past contributions.

Third party payment is therefore a partial answer to the problem of an individual's uncertainty about their future healthcare needs (Barr, 2001). None of us can be certain of such needs; even those with 'good genes' and healthy lifestyles may walk under the proverbial bus. It would therefore be difficult for anyone to be confident about being able to afford future necessary healthcare from income or personal savings. Third party payment has the rationale of insulating individuals against unknown risks of both illness itself and of unaffordable costs of treatment. Third party payment can also be a partial answer to the problem of equity, that is to the empirical tendency for the poor and the sick to be the same people, as clearly demonstrated by the existence of wide social class differentials in both mortality and morbidity (*Population Trends* 86, 1996: 15–20). Despite having the greatest healthcare needs, under out-of-pocket payment such people would be least able to afford to have them met. However, the extent to which this is ameliorated by third party payment will depend upon how widely the risks of ill-health are spread and it is clear from the descriptions above that different variants of third party payment achieve this to different extents. Tax-financed systems pool the risks across a whole national population and, other things being equal, spread the risk most equitably. Private insurance, and any system of social insurance which employs a multiplicity of third party payers, are likely (other things being equal) to be less equitable since there exists the probability that poorer, sicker people will be found in some risk pools and richer, healthier people in others. The former group may therefore receive a poorer range of benefits than the latter. Moreover, any type of third party payment system (including, as we shall see below, the NHS) might in practice make charges for, or impose access restrictions upon, certain treatments.

Although these matters seem technical, they are also deeply political in that they manifest different normative assumptions about what risks should be pooled. A private insurance system seeks to compensate for the relatively narrow range of uncertainties related to the individual's health over his or her own lifetime and within his or her own social group. In contrast, a tax-financed system, in socialising risk across a single national population, additionally seeks to compensate for a broader range of socioeconomic inequalities. The latter has a *universalist* rationale, implying a notion of citizenship which includes social rights (Marshall, 1950) in which effective participation by individuals in society is to be secured by state action (Flynn, 1997).

The policy consequences of third party payment

Third party payment systems risk the inflation of demand over time. Such demand increases may be conceptualised in terms derived from the economic concept of 'moral hazard'. *Consumer moral hazard* arises only where some or all of the costs of care are met by the third party payers; it encourages a higher rate of use than would occur if full costs had to be met at the point of use (Pauly, 1968), since the demander assumes that the cost of their usage will be spread over a large number of taxpayers, fund members or policy holders. However, if large numbers of people behave in this way, then total demand (for healthcare and hence for



the resources to provide it) *will* rise. Consumer moral hazard in third party payment systems for healthcare is the consequence of divorcing payment for services from their use. Third party payment makes it easier for people to obtain care than would otherwise be the case, but at the same time tempts them to increase their demands. An important qualification to this is that cost and price are not synonymous; the non-money costs of obtaining care can be significant. At the minimum the user must take steps, such as telephoning for an appointment, rearranging a working day, travelling to the surgery or hospital, and perhaps sitting for some time in a crowded waiting area, in order to gain access. Costs can be higher. We may react adversely to the drug which is prescribed or the needle may hurt as it pierces the flesh and the prospect of gastroendoscopy or sigmoidoscopy is hardly pleasant.

Provider moral hazard (or 'supplier-induced demand') arises from information asymmetries; the consumer's lack of knowledge of a highly technical service coincides with a provider's interest in increasing provision and allows the latter to affect demand. Whilst patients do make generalised demands in the sense of arranging a visit to the doctor or being taken to the Accident and Emergency Department, it is typically (though not invariably) the physician or other clinical professional who translates it into a specific demand for antibiotics, pathology tests, a specialist appointment, or a surgical operation. Some accounts cast such professional motivation in material terms. If the clinician is remunerated on a fee-for-service basis, there are clear incentives to maximise supplier-induced demand unless the total fees are 'capped' in some way. The same incentive may exist if the institution which employs the clinician is itself remunerated by the third party payer on any basis which is sensitive to the volume of patients or treatments. From such a perspective, a system in which clinicians were salaried would have the opposite effect of 'underprovision', since there would be no economic incentives to perform beyond the level necessary to retain one's job. However, this seems an unnecessarily narrow perspective on incentives; Donaldson and Gerard (1993: 33) have argued that where providers are salaried and do not themselves have to bear the costs of treatment, simple ignorance of costs may lead to overprovision. There may be professional ethical incentives to provider moral hazard; even if the hospital's budget is not volume-related and clinicians are remunerated by salary or capitation fees, demand might still increase as a result of the clinician's desire to behave ethically, that is to do the best possible (according to their own imperfect opinion) for his or her patient. Supplier-induced demand does occur in out-of-pocket payment systems, but might be limited by the patient's inability or unwillingness to pay; in a third party payment system such limitations are attenuated by the patient's and clinician's mutual knowledge that a third party will meet all or part of the money costs of care (Reinhardt, 1985). Thus, demand in a third party payment system might be expected to increase over time since neither consumers nor providers have the incentive to moderate it. This does not imply the assertion that healthcare demand in such systems is infinite; it is unlikely that demand for healthcare could escalate to the point where all other demands were excluded. It is rather that no real-world third party payment system as yet seems to have experienced an autonomous levelling off of demand.

Whilst the two forms of moral hazard provide the immediate basis for the inflation of demand in third party payment systems, there are a number of secondary factors which may affect the propensity of patients and clinicians to increase their demands. One obvious candidate is the

constant invention and development of new medical technologies, many of which are extremely expensive. The pharmaceutical and medical equipment manufacturing industries are important sectors of the economies of the UK, USA and Germany and significant exporters. The term 'technology' encompasses 'drugs, devices and medical and surgical procedures used in medical care and the organisational and supportive systems within which such care is provided' (Office of Technology Assessment, 1978: 2). New approaches to psychotherapy, new packages of care for the elderly, and new multi-professional approaches to the care of stroke victims are therefore new technologies, and indeed may carry costs just as high as new drugs. The mere existence of technologies does not create a demand, which depends upon patients and/or clinicians perceiving that they might be beneficial. Moreover, the inventors and manufacturers of medical technologies have an economic interest in maximising their sales which, if it can be linked to public and professional perceptions of a condition as a 'disease', leads to the 'medicalisation' of conditions not previously thought of in such terms. In the words of Moynihan and Smith (2002), 'many of life's normal processes – birth, ageing, sexuality, unhappiness and death – can be medicalised', including pharmaceutical treatment of male baldness on the grounds that it might lead to panic, mental health problems or poor job prospects, and the transformation of personal shyness into the condition of 'social phobia', treatable by antidepressants.

Another secondary source of demand for healthcare is demographic shift; many countries have an ageing population in both absolute and relative terms, brought about by increasing life expectancy and a falling birth rate. Indeed one major study has suggested that although policy makers have tended to assume that further increases in life expectancy (currently a little under 76 years for British males and 81 for British females: Wanless, 2002: 42) can only be modest, there is apparently no absolute limit to life expectancy (Oeppen and Vaupel, 2002). Older people currently consume greater amounts of hospital care per capita than do younger people; the 75–84 year age group consumes four times as much NHS resource per capita as does the 16–64 group, a comparison that rises to seven times for the 85 and over age group (Glennerster, 2003: 58). However, there are disputes about whether increases in the older groups as a proportion of population necessarily imply continuing increases in demand. Policy makers in the UK have tended to assume that any such increases will be modest (see, for instance, Wanless, 2002: 42–3) and recent evidence suggests that proximity to death is a much more important determinant of hospital usage than age *per se* (Dixon et al., 2004; Seshamani and Gray, 2004). On the other hand, there is likely to be a very substantial increase in the demand for 'long-term care' for older people in residential and nursing homes and in their own homes over the next 30 years (Comas Herrera et al., 2003). Moreover, as the above discussions of moral hazard and medicalisation suggest, there is not necessarily a linear relationship between individuals' health and their demand for care – it has been reported that UK residents' expectations of 'unhealthy' life before death are increasing faster than their total life expectancy (Hebert, 2004).

The net financial impact of new technologies and demographic shift on a health system will depend partly upon whether the former reduce demand elsewhere in the system, either by substituting for other interventions for the same condition or by helping patients to attain a state of health in which they perceive themselves to need less treatment than would otherwise



be the case. Klein has linked the effects of demographic and technological change in a pessimistic conclusion:

Even if the limitations of medical technology in curing disease and disability are now becoming apparent, there are no such limitations on the scope of health services for providing care for those who cannot be cured. Even if policies of prevention ... were to be successfully introduced, their very success in extending life expectancy would create new demands for alleviating the chronic degenerative diseases of old age. In other words, no policy can ensure that people will drop dead painlessly at the age of eighty, not having troubled the health services previously. (Klein, 1983: 182)

It is also worth noting that developing technology (for instance, in anaesthetics or ‘minimally invasive’ surgery) can make interventions safer and subject to increased demand (Seshamani and Gray, 2004: 67).

A final secondary factor is *public demand*, which operates against a background in which healthcare is a prominent public and political issue in the UK, the USA and many other countries. One element of this is the much greater public availability of information about healthcare, with particular emphasis on high profile media reportage about new technologies. The Internet has begun to contribute substantially to the availability of such information (Coiera, 1996). Alongside this growth of information is an apparent increase in the level of activity by patient pressure groups, as we shall see in Chapter 5, usually organised around a particular disease or health condition (Wood, 2000). Such groups (which may also provide advice and other services to their members) are often supported by health professionals from appropriate clinical specialties and, naturally enough, press the appropriate health service bodies for what they perceive to be better services, including new technologies, for themselves. In the UK, these demand factors operate against a background of considerable public support for the NHS as an institution, even though they may be critical of their own experiences as patients. To cite one set of findings from numerous similar examples, 63 percent of respondents to a 2000 opinion poll rated the NHS as the single ‘most important national institution’, placing it in first place followed by Parliament (12 percent), the police (11 percent), the BBC (4 percent), and (at around 3 percent) the Royal Family, the Bank of England and the Benefits Agency (ICM poll, cited in the *Guardian*, 18 April 2000: 19). In thinking about patient and public demand, it is important to remember that it reflects the *perception* of the demander that healthcare would be good for them. Consequently, consumer moral hazard in healthcare is not necessarily avoided by reference to research into the clinical effectiveness or otherwise of treatments, a matter to which we return in Chapter 3.

Managing healthcare supply and demand

Faced with demands for healthcare fuelled by the kind of factors outlined above, third party payers may not simply respond with increased resources, and indeed one recent study has shown that there is no straightforward relationship between public opinion and NHS expenditure (Soroka and Lim, 2003). More generally:

The modern health care problem can therefore be seen as a reflection of the way healthcare financing has been collectivised, through the sort of risk pooling arrangements outlined above and through the way this process of collectivisation, by breaking the direct link between consumption and payment, removes or weakens budget constraints on consumers of health care resources. The problem facing health care systems is therefore how to reimpose, or reinvent, those constraints, in a world where the collectivisation of finance has to be taken as a given. (Harrison and Moran, 2000: 496)

In response to this situation, governments adopt various (not necessarily mutually exclusive) means of matching supply and demand. These can be roughly classified into *supply-side adjustments*, that is those which aim to increase the resources available for healthcare, and *demand-side adjustments* which aim to reduce or stabilise demand for services. Supply-side adjustments may take the form of measures to increase the flow of revenue to third party payers (tax or contribution increases, or co-payments), or to encourage a higher level of out-of-pocket or private insurance expenditure as assumed substitutes for third party payment. In publicly funded healthcare systems, other supply-side measures include toleration of public sector budget deficits and reallocation of public expenditure priorities so as to increase healthcare expenditure at the expense of other public programmes. In any system, policy makers may seek to improve the productive efficiency of the healthcare sector by a range of management and organisational measures aimed at codifying clinical practice and modifying the incentives facing actors in the system, a strategy further examined in Chapter 3. Examples include tighter management in general as well as some more specific techniques such as ‘managed competition’ (Bruce and Jonsson, 1996) and ‘managed care’ (Robinson and Steiner, 1998).

In contrast, demand-side adjustments are aimed at reducing or containing demand for care, a process for which we employ the term ‘rationing’, whilst recognising that this usage is not without its critics, including UK politicians. Some demand-side measures operate *implicitly* so far as the patient is concerned. Examples include cost barriers which partially offset the effect of consumer moral hazard. Such costs may be financial (thus charges for services are a deterrent), but spatial, psychological and procedural barriers may also be effective; remote or highly centralised facilities, user-unfriendliness and strict ‘gatekeeping’ criteria tend to reduce demand. Other demand-side measures are *explicit*, that is consist of more-or-less clear rules about patient entitlement; for instance, such rules may exclude certain procedures or drugs. The desirability of implicitness and explicitness is much debated. Some proponents of implicitness (Hoffenberg, 1992; Mechanic, 1992; for a philosophical discussion, see Calabresi and Bobbit, 1978) have justified their position on the grounds that explicit decisions are too brutal for society to contemplate, whilst others (Hunter, 1993; Klein et al., 1996) have concentrated on their conceptual and practical difficulties. Proponents of various degrees of explicitness (Harrison and Hunter, 1994; New and LeGrand, 1996) often stress transparency as a prerequisite of fairness.

Whether implicit or explicit, rationing mechanisms are necessarily underpinned by one or more of a range of criteria that may themselves be explicit or implicit and may reflect a range of political and other normative positions. Five criteria are perhaps the most widely advocated. First, the *rule of rescue* gives priority to persons in acute or life-threatening conditions,



thereby locating moral content in trying rather than in succeeding. This is likely to generate significant opportunity costs in terms of 'wasted' effort, though it might be argued that this is a legitimate public preference (Goodin and Wilensky, 1984). Second, *deserts* are sometimes used as the basis of an argument for exclusion, often in the context of a health state which is considered to be self-inflicted (for instance by smoking or participation in dangerous sports). Third, prospective *effectiveness* of a healthcare intervention is widely argued to be a common-sensical rationing criterion (Evans, 1990). The existence of uncertainties about effectiveness undermines a good deal of the force of arguments that *whatever* is effective should be provided, though the 'prudent insurance principle' (Dworkin, 1994) provides a thought experiment for dealing with such difficulties. Fourth, *cost-effectiveness* and *cost-utility* are espoused by those who maintain that the cost, as well as the degree of effectiveness, of interventions should be considered. This position has given rise to a number of artificial measures of health outcome such as Quality (or Disability) - Adjusted Life Years, whose theoretical properties are utilitarian in the sense that they aim at the maximisation of health gain in return for any given level of expenditure. Finally, as noted above, third party payment systems are underpinned by a desire to enhance *equity*, that is to ameliorate the position of people who cannot afford the care from which they might benefit. *Equity* and *equality* are therefore concerned with the distribution of services or of health status respectively, a criterion which may trade off against cost-effectiveness.

An historical sketch of supply and demand in the NHS

As noted above, policy measures may focus on either the supply of resources to the NHS or on various means of reducing or containing demand. These are treated separately here, but it should be noted that in practice, governments often develop parallel policies so that both supply and demand are addressed simultaneously.

The supply of NHS resources

Many of the ideas that were to constitute the social policy of the postwar Labour government had previously been brought together in the wartime Beveridge Report (Beveridge, 1942). Although the need for a health service occupied relatively little space in the Report (which concentrated largely on social security), it contained the key assumptions that doctors would limit unnecessary demands for healthcare (Webster, 1988: 36) and that the impact of making medical services freely available would be to produce a population that would be healthier than before in absolute terms. As a result, it seemed, the workforce would be more productive and in due course, the costs of the proposed health service would fall (Beveridge, 1942). Early experience of the NHS, established in 1948, suggested a different picture. For each of the first few years, it proved necessary for the government to obtain supplementary estimates from Parliament in order to meet the unanticipated level of demand (Klein, 2001: 26; Rivett, 1998: 110; for a detailed account, see Webster, 1988: 133–43). Contemporary interpretations saw the situation in terms of the bursting of a 'dam' of accumulated demand

from people previously unable to afford care, and especially dentures and spectacles (Rivett, 1998: 80; Webster, 1988: 134). This experience led to the Labour government's decision to legislate in 1950 for prescription, dental and optical charges. Labour introduced the latter before losing office in 1951, though prescription charges were not introduced until 1952, after the election of a Conservative government (Rivett, 1998: 46, 55, 110) and were briefly abolished between 1966 and 1968 (Office of Health Economics, 2003: 39).

As early as 1949, it was argued that NHS cost estimates had ignored the effects of an ageing population and of perpetual technological advance. Moreover, contrary to Beveridge's assumptions, there was not a finite quantity of ill-health into which the NHS would make inroads. Rather, cures for particular diseases simply meant that patients survived to suffer from other, more expensive ones (Roberts, 1949, 1952). However, the Guillebaud Committee on the cost of the NHS, established by the government in 1952, reached the conclusion that the unexpected increases in NHS expenditure were largely the consequence of failure to account for inflation and that after 1951 NHS expenditure had actually fallen as a proportion of gross national product (Committee of Enquiry, 1956). The report did not recommend any further charges, and was unenthusiastic about existing dental and optical charges, recognising that they might act as a deterrent to patients who needed services. Guillebaud's reassuring conclusions can be seen as having set a positive tone for health politics for the next 30 years and perhaps even as having virtually removed the NHS from party politics (Klein, 2001: 25; Rivett, 1998: 114). By its tenth anniversary in 1958 the Labour and Conservative Parties were vying to gain political credit from, respectively, the NHS's creation and expansion (Klein, 2001: 48).

Yet even the founding Minister of Health of the NHS had entertained doubts about the value of 'the ceaseless cascade of medicine ... pouring down British throats' (Bevan, quoted in Webster, 1988: 145), whilst a later Minister of Health subsequently reflected both that rationing was inevitable and that explicit criteria should be employed, but that:

the task is not made any easier by the political convention that the existence of any rationing at all must be strenuously denied. The public are encouraged to believe that rationing in medical care was banished by the [NHS], and that the very idea of rationing being applied to medical care is immoral and repugnant. Consequently when they, and the medical profession too, come face to face in practice with the various forms of rationing to which the [NHS] must resort [such as hospital waiting lists], the usual result is bewilderment, frustration and irritation. (Powell, 1966: 17)

It is therefore hardly surprising that no explicit attempts to manage demand were made, and prescription, dental and optical charges were seen as a temporary necessity rather than as an ongoing means of reducing demand. However, the 1950s and early 1960s were a period of economic austerity for the UK generally and of stagnation in the level of NHS resources more specifically. Table 1.1 shows changes in the latter, expressed both as a percentage of gross domestic product (GDP; roughly speaking, the national income) and in real terms, that is, adjusted for inflation. (It should be noted that these adjustments are made in terms of inflation across the whole UK economy rather than in terms of the usually rather higher

Table 1.1 NHS expenditure (UK, 1949 to 2003)

Calendar year & political party in office	% Gross domestic product	Real terms (adjusted by GDP deflator at market prices) Index (1949 = 100)
1949 Labour	3.46	100
1950	3.62	109
1951 Conservative	3.39	105
1952	3.20	99
1953	3.07	98
1954	3.00	100
1955	3.02	104
1956	3.07	107
1957	3.15	110
1958	3.21	114
1959	3.29	122
1960	3.37	131
1961	3.44	138
1962	3.43	138
1963	3.43	145
1964 Labour	3.43	155
1965	3.56	163
1966	3.68	173
1967	3.79	182
1968	3.81	189
1969	3.70	188
1970 Conservative	3.84	201
1971	3.91	208
1972	4.03	223
1973	4.00	235
1974 Labour	4.58	268
1975	4.85	281
1976	4.84	288
1977	4.63	283
1978	4.53	286
1979 Conservative	4.49	290
1980	4.88	310
1981	5.11	320
1982	5.09	324
1983	5.00	329
1984	4.96	335
1985	4.83	339
1986	4.88	355
1987	4.86	370
1988	5.05	404
1989	5.00	408
1990	5.10	420
1991	5.47	445
1992	5.80	472
1993	5.80	483
1994	5.84	507
1995	5.83	521
1996	5.71	525
1997 New Labour	5.64	535
1998	5.61	548